CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

1.1 GENERAL INTRODUCTION

The nature of viruses only began to be understood just over one hundred years ago, despite the fact that their concept as a natural phenomenon separate from all other organisms had been realised much earlier before the turn of the 19th century. There is evidence in ancient Chinese literature and Egyptian art for what may have been smallpox and other viral infections such as poliomyelitis and measles (Levy *et al.*, 1994).

The early successes in the study of microorganisms were not based on any real understanding of the nature of the infectious agents or pathogens but on the diseases that they caused. Antony Van Leeuwenhoek identified bacteria with the aid of a microscope and referred to them as small animals or "animalscules" (Myrvik and Weiser, 1988; Prescot *et al.*, 1996). The significance of the "animalscules" identified by Antony van Leeuwenhoek only became apparent when Robert Koch and Louis Pasteur pronounced their "germ theory". At the time, Pasteur worked on rabies, but was unable to discriminate between bacteria and other infectious agents of disease, such as bacterial toxins.

At the turn of the 19th century, viruses were recognised as agents responsible for diseases of plants (tobacco mosaic) and animals (foot and mouth) through initial experimentation by Dimitri Iwanowski and Martinus Beijerinick (see historical reviews by Bos, 1999 and Lecoq 2001). Through these studies, viruses were shown to be submicroscopic and obligate intracellular parasites. In fact, it is possible that viruses could have been in co-existence with living organisms since the origin of life (Levy *et al.*,

1994). While many viruses are harmful to their host, others are symbiotic and may give an advantage to the infected host. For example, viruses could provide a gene that mediates a needed metabolic capability or drug resistance. Mammalian evolution would in fact not have been possible without viruses. Placental morphogenesis results from the syncytium-forming properties of syncytin, a gene that was borrowed from an endogenous retrovirus (Mi *et al.*, 2000). Several viruses have however raised great concern in the past century because of serious threats in humans; influenza (1918, 1957, 1968, 1977) (Alexander and Brown, 2000), poliomyelitis (1930s), herpes (1960s), acquired immunodeficiency syndrome (HIV) (1980s) and Ebola and Marburg in the 1990s (Le Guenno, 1997; Brown, 1997; Levy *et al.*, 1994). The diseases caused by these viruses have brought increased attention to the field of virology.

1.2 Rabies: A short history

Globally, rabies is an important and one of the oldest recognised viral zoonoses. Although rabies has never been a disease of pandemic proportions such as smallpox, plague, malaria and diarrhoea, it has always inflicted terror because of its horrific clinical manifestations and lethality (Wiktor, 1985; Charlton, 1988). Rabies is a fatal viral disease of the central nervous system (CNS) that is usually transmitted through bites of animals with infectious saliva. Rabies affects humans and other warm-blooded mammalian species (reviewed in King and Turner, 1993; WHO, 1991, 1992; Chommel, 1993) and several bat species (King *et al.*, 1990; Nadin-Davis *et al.*, 2001; Loza-Rubio *et al.*, 1999; King *et al.*, 1994; Smith, 1988; Hanna *et al.*, 2000).

Rabies is an old disease as illustrated by the relevant descriptive terms. For example, the Latin word "rabies" is believed to derive from the Sanskrit “rabhas” meaning “to do violence”. The French word for rage is derived from the noun “robere”, to be mad (Steele and Fernandez, 1991; Krebs *et al.*, 1995). *Lyssavirus*, the genus to which rabies and rabies-related viruses belong, owes its name to the Greek term "lyssa" or "lytta" meaning "madness". In the 18th century, the cause of rabies was not well understood, and it was thought to be caused by a poison (Latin word "virus") contained in the saliva of rabid dogs. Rabies is widely cited in Greek and Roman mythology (Steele and Fernandez, 1991) and the first recorded descriptions of canine rabies were made by Democritus at *ca* 500 BC (Krebs *et al.*, 1995).

Rabies is a disease of both public health and veterinary importance (Smith and Seidel, 1993) with mortality ranking eleventh of all infectious diseases (Krebs *et al.*, 1995). It is estimated that over 40 000 people die of rabies each year and that dogs are responsible for 94% of these deaths in developing countries (WHO, 1999; Meslin *et al.*, 1994). Furthermore, post-exposure treatment (PET) is administered to thousands of people bitten by animals (rabid or otherwise) annually, making rabies management costly (Meltzer and Rupprecht, 1998a, 1998b). The WHO (1994) estimates that about 4 million persons per year are given at least partial PET. Surveillance of rabies facilitates an identification of viruses originating from various reservoir species (De Mattos *et al.*, 1996; Curtis, 1999), which could be targeted for effective rabies control (Bingham *et al.*, 1993; Matter *et al.*, 1998).

1.3 The *Rhabdoviridae* family

The family *Rhabdoviridae* belongs to the Mononegavirales order, which includes the *Paramyxoviridae* and *Filoviridae* families. The viruses share a non-segmented RNA genome of negative polarity embedded within the ribonucleoprotein (RNP). Two genera that have been defined to classify members of the Rhabdoviridae infecting mammals are the *Lyssavirus* and *Vesiculovirus* genera, respectively (Figure 1). The rabies virus belongs to the family Rhabdoviridae (Greek rhabdos for rod), genus *Lyssavirus*. The rhabdoviridae are a family of enveloped viruses with a wide host range including fish, arthropods (mostly insects), crabs, amoebae and plants (Cann, 1993; Murphy *et al.*, 1995; Benmansour *et al.*, 1994). The virion assumes a typical elongated, bullet-shaped or cone-shaped morphology (Figure 2) (Murphy *et al.*, 1995; Levy *et al.*, 1994) with average dimensions of 75 x 180 nm.

1.4 Vesiculovirus genus

Vesiculostomatitis virus (VSV) is the prototype and the most studied of the Rhabdoviruses, and is thus a model of non-segmented negative-strand RNA viruses. The viral genome is 11.2 kb in size. The two classical serotypes of VSV are the New Jersey (NJ) and Indiana (IN) serotypes (Cartwright and Brown, 1972; Banerjee *et al.*, 1984) which infect cattle, horses and swine. VSV is enzootic in certain tropical and subtropical areas of the Americas, and is probably spread by a variety of insect vectors.

1.5 *Lyssavirus* genus

Classification of the Lyssaviruses is based on serotypic and genotypic characteristics (Schneider *et al.*, 1973; Bourhy *et al.*, 1993a). The *Lyssavirus* genus currently consists of seven genotypes (Figure 1). Rabies, the prototype member of genotype 1 (GT1) of the *Lyssavirus* genus, has worldwide distribution. Genotypes 2-6 correspond to rabies-related viruses including the Lagos bat (GT2), the Mokola virus (GT3), the Duvenhage virus (GT4), European bat lyssaviruses 1 and 2 (EBLs) making GT5 and GT6 respectively (Bourhy *et al.*, 1992; Amengual *et al.*, 1997; Montano-Hirose, 1990). The European bat lyssaviruses have been reported from Europe (Russia to Spain) whilst the remaining Lyssavirus genotypes; GT2-GT4 have been reported exclusively from Africa (Schneider and Cox, 1994; King and Crick, 1988; Meredith *et al.*, 1971). A recently isolated lyssavirus from a bat (*Pteropus alecto*) in Australia was proposed to form a new genotype (GT7) (Gould *et al.*, 1998).

1.6. Rabies virus genome organisation and replication

The molecular biology of the rabies virus has been reviewed extensively (Wunner *et al.*, 1988; Tordo and Kouknetzoff, 1993; Tordo, 1996; Tordo, Charlton and Wandeler, 1998). The rabies virus contains a non-segmented negative-sense, single-stranded RNA genome, approximately 12 kb in size (Tordo *et al.*, 1992). The virus genome encodes five proteins: the nucleoprotein (N), phosphoprotein (P) (formerly called NS), matrix protein (M), the glycoprotein (G) and the polymerase (L) protein (Tordo *et al.*, 1986b). Three of these proteins (L, N and P) are associated with the genomic RNA to form a ribonucleoprotein (Yang *et al.*, 1998; Kawai and Morimoto, 1994). The other two

proteins, the matrix and the glycoprotein, form the inner side and the outer layer of the bilayered lipid envelope respectively. In addition, the glycoprotein forms spike-like projections (Figure 2).

1.6.1 Transcription, translation and replication mechanisms of the rabies virus

The mechanisms of transcription, replication and expression of non-segmented negative strand RNA genomes were originally established with studies on VSV (Barnerjee and Barik, 1992). At the 3' end of the RNA genome is a 58nt long leader RNA which is followed by 5 open reading frames (ORFs) in the order N, P, M, G, L characteristic of Rhabdoviridae (Pringle, 1995). A single promoter is recognised by the transcriptase complex at the 3' end of the genome. Both transcription and replication start at the 3' end of the genome with the leader sequence. Virion RNA (-ve sense) is copied to a complementary mRNA (+ sense) before virus replication occurs (Tordo *et al.*, 1998). The RNP complex attaches near the 3' end, and then moves towards the 5' end producing monocistronic transcripts. The transcriptional complex recognises start, stop and polyadenylating signals flanking the cistrons (Tordo *et al.*, 1988; Sacramento *et al.*, 1992). Translation of each of the mRNA's of the rabies virus occurs immediately after transcription. Structural genes and their mRNAs are copied into five monocistronic mRNA transcripts (Tordo and Kouknetzoff, 1993; Tordo *et al.*, 1992; Tordo *et al.*, 1998), which then encode the respective rabies proteins (described in section 1.6.2.). The lengths and positions of individual genes of the Pasteur Virus strain are shown in Table 1.

Table 1: Individual genes of PV laboratory strain

mRNA	nucleotide position	length of mRNAs
N	59-1482	1424
P	1485-2475	991
M	2481-3285	805
G	3291-4964	1674
L	5388-11862	6475

After Tordo *et al.*, 1986a

1.6.2 Rabies Viral Proteins

1.6.2.1 The Glycoprotein (G)

The glycoprotein (505 amino acids, MW 57 KDa) is composed of a cytoplasmic domain, a transmembrane domain and an ectodomain (Gaudin *et al.*, 1992; Anilionis *et al.*, 1981). The glycoprotein is exposed in the form of 10-nm-long peplomers on the external surface of the virus membrane (Figure 2). Some notable features of the glycoprotein are its glycosylated mannose residues (Tordo *et al.*, 1998) and the hydrophobic amino acid residues 439-462 (Whitt *et al.*, 1991). The ectodomain of the glycoprotein has been shown to be highly conserved whilst the transmembrane and cytoplasmic domains are significantly divergent (Fodor *et al.*, 1994; Baer, 1991; Tordo *et al.*, 1992).

The G protein is responsible for virus attachment to cell surface receptors and is therefore important in determining virulence (Benmansour *et al.*, 1992; Dietzschold *et al.*, 1983; Coll, 1995). In addition, the glycoprotein is responsible for the induction and binding of virus-neutralising antibodies (VNA). Several antigenic domains located on the external domain of the G protein have been described. In genotype 1 rabies viruses,

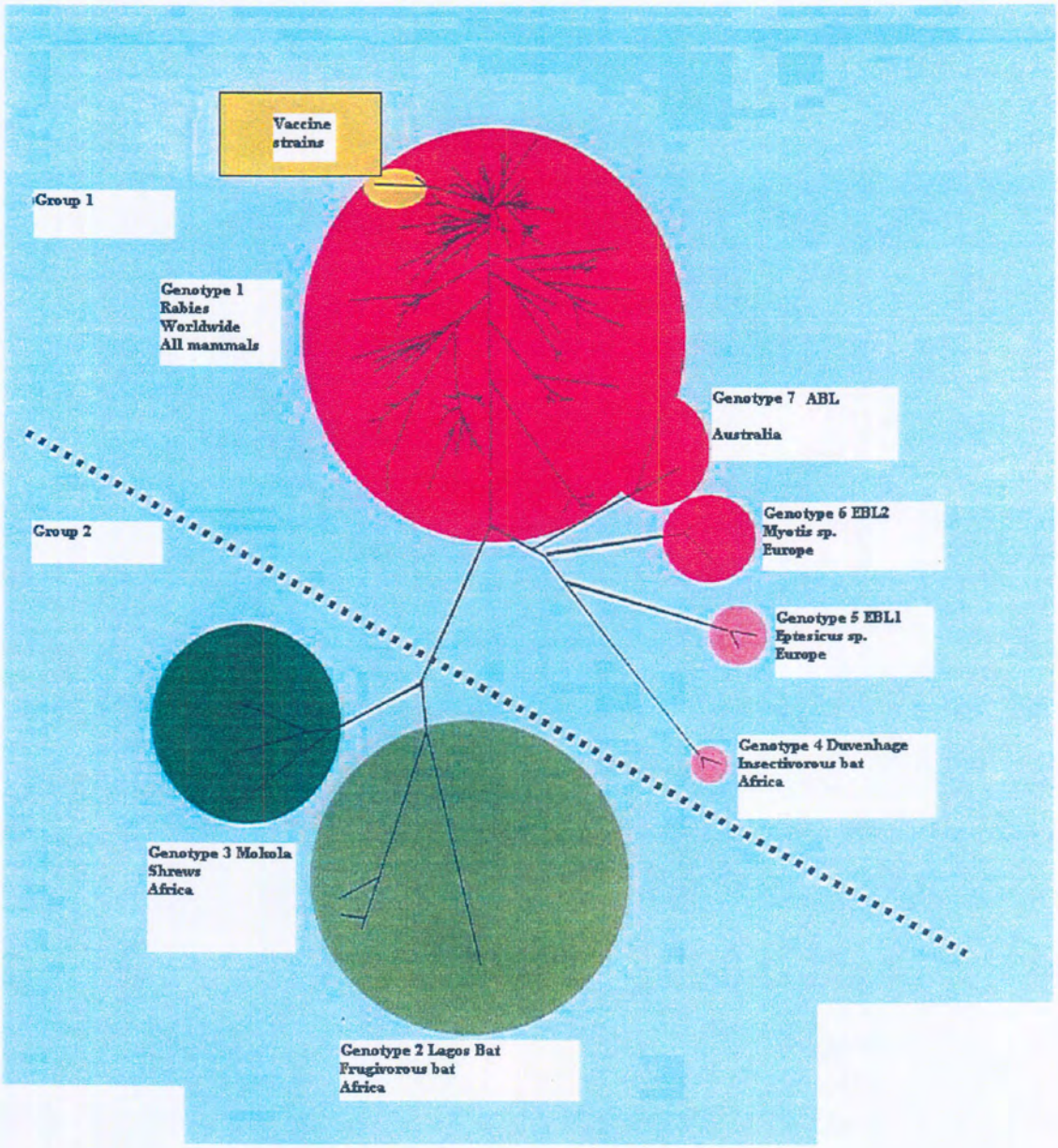

Figure 1: Phylogenetic tree of lyssaviruses and vesiculoviruses based on the alignment of partial nucleotide sequence of the ectodomain of the glycoprotein. Reprinted with permission from McColl *et al.*, 2000.

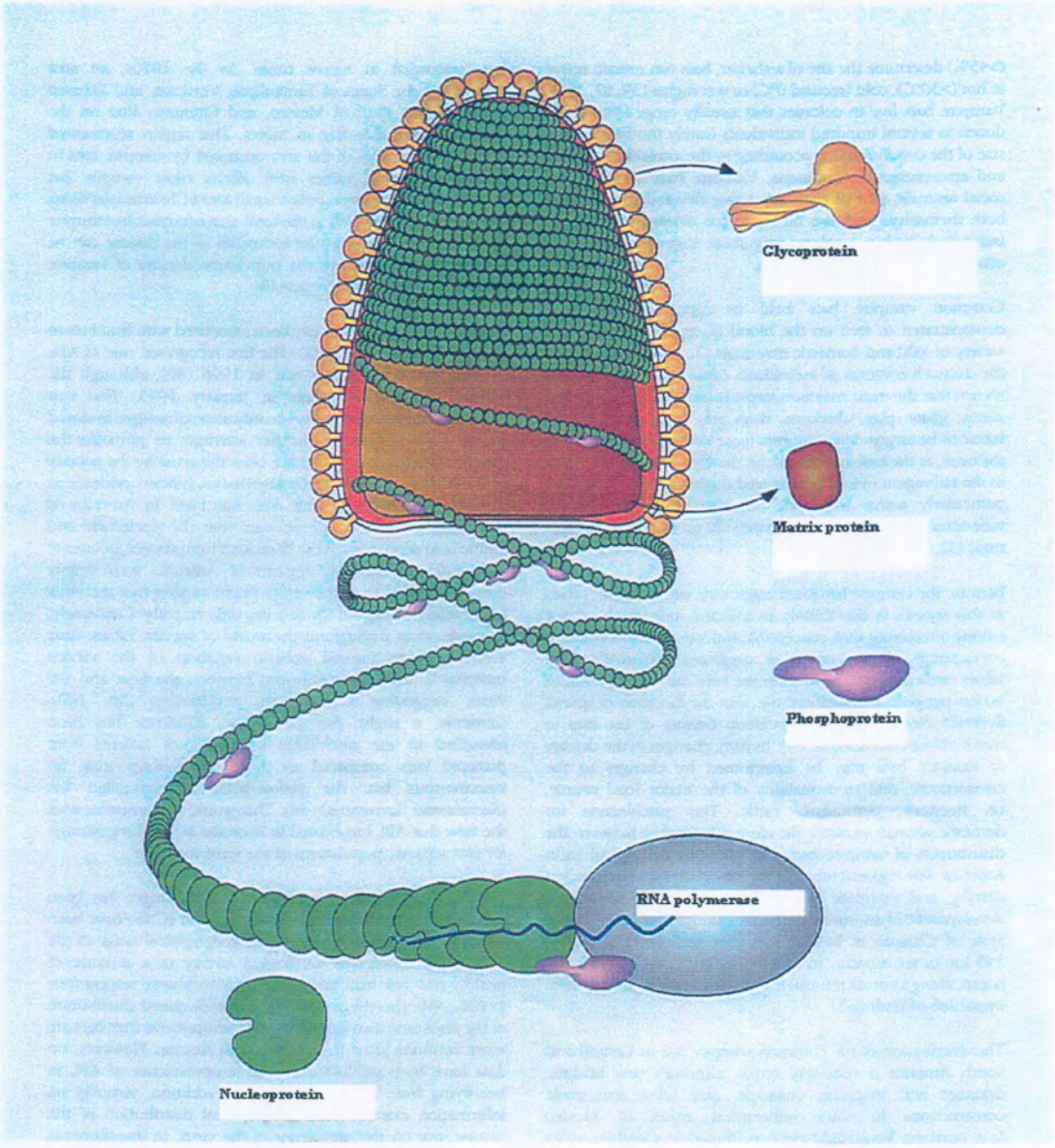

Figure 2: Structure of a rabies virion particle showing surface glycoprotein projecting from the lipid-containing envelope that surrounds the internal nucleocapsid complex. The matrix protein is shown lining the viral envelope with the cytoplasmic domain of the surface glycoprotein. The helical nucleocapsid core is comprised of the RNP. Reprinted with permission from McColl *et al.*, 2000.

neurovirulence seems to be directly related to the maintenance of an arginine in position 333 referred to as site III. Mutagenesis or changes of amino acid sequence of site III results in altered pathogenicity (Dietzschold *et al.*, 1983; Seif *et al.*, 1985; Tuffereau *et al.*, 1989; Benmansour *et al.*, 1991; Ito *et al.*, 2000b), antigenicity and immunogenicity (Wunner, 1991) of the rabies virus. The G protein is involved in rabies pathogenesis and is also thought to be responsible for part of the neurotropism (Morimoto *et al.*, 1998). From studies involving rabies viruses lacking the pseudogene sequence, it was demonstrated that the pseudogene is not involved in viral spread and mutagenesis (Ceccaldi *et al.*, 1998).

1.6.2.2 The Nucleoprotein (N)

The 1350 nucleotide long mRNA transcript of the nucleoprotein gene encodes a protein of 450 amino acids (MW 50 KDa) (Tordo *et al.*, 1986a, 1986b; Tordo, 1992; Kissi *et al.*, 1995), and is the second most extensively studied protein of the rabies virus. The N protein is produced in abundance during virus infection both *in vivo* and *in vitro* (Iseni *et al.*, 1998; Wunner *et al.*, 1988). The N protein of rabies virus is phosphorylated unlike that of the VSV (Tordo, 1986a). Once phosphorylated, the N protein of rabies virus then packages the RNA genome (Kouznetzoff *et al.*, 1998; Yang *et al.*, 1998). Mab studies have shown the nucleoprotein gene to be highly conserved in comparison to the non-coding G-L intergenic sequences of the glycoprotein gene. Classification of Lyssaviruses has been based on reactivity patterns with antinucleocapsid Mabs (Wiktor *et al.*, 1980; Bourhy *et al.*, 1993b).

The N protein may also be involved in immunity since the rabies RNP has been shown to protect animals against a peripheral challenge with infectious saliva (Dietzschold *et al.*, 1987) particularly through the induction of T-helper cells (Ertl *et al.*, 1989) and cytotoxic T cells (Lafon, 1994). An earlier study, however, found that the N protein did not protect against an intracerebral challenge (Wiktor *et al.*, 1973).

1.6.2.3 The Matrix protein (M)

The matrix protein (202 amino acids, MW 23 KDa) is a major structural protein of rabies virus but little is known about its biological and immunological significance. The matrix protein is the smallest rabies virus structural protein (Tordo *et al.*, 1992). This protein is located in the inner surface of the viral envelope (Figure 2) thus facilitating interaction of the membrane-anchored viral glycoprotein and the membrane-associated viral proteins. Evidence from recent studies has shown that M covers the RNP, thus keeping it in a condensed form (Mebatsion *et al.*, 1999). This property of condensing the RNP makes the M protein crucial in viral morphogenesis. The M protein of rabies virus has no sequence homology to that of VSV, despite the fact that both proteins are involved in similar functions (Wunner, 1988).

1.6.2.4 The Polymerase (L)

The multifunctional RNA-dependent RNA polymerase is the largest polypeptide (2142 amino acids; approximately MW 200 KDa) and occupies 54% of the rabies genome (Tordo *et al.*, 1988). The L protein is responsible for activities involved in RNA synthesis such as polymerase activity, capping, methylation and polyadenylation. In

comparison with other rabies structural proteins, it is the L protein, which has the highest proportion of hydrophobic amino acids, but an extremely small number of protein molecules per virion (Tordo *et al.*, 1992). The rabies polymerase exhibits a high degree of homology with that of VSV and to a lesser degree with those of *Paramyxoviridae* family such as Sendai and Newcastle disease virus (Tordo *et al.*, 1988; Levy *et al.*, 1994).

1.6.2.5 The Phosphoprotein (P)

The deduced amino acid sequences of phosphoprotein of the rabies virus reveals that it contains 297 amino acid residues (MW 33 KDa). The phosphoprotein of the rabies virus is part of the RNP complex and is therefore required for transcription and replication of the single-stranded RNA genome (Larson and Wunner, 1990; Tordo *et al.*, 1992). The first 19 amino acid residues of the phosphoprotein have been shown to interact with the nucleoprotein (Chenik *et al.*, 1998; Schoen *et al.*, 2001; Jacob *et al.*, 2001; Fu *et al.*, 1994) and this demonstrates the importance and possible involvement of this protein in virus assembly. Recently, it has been demonstrated that LC8, a dynein, which is important in both actin and microtubule-based neuronal transport, interacts with the N terminal of the phosphoprotein, (Jacob *et al.*, 2000). The phosphoprotein of the CVS strain of the rabies virus is phosphorylated by a unique cellular kinase and this may illustrate the important role the phosphoprotein plays in the life cycle of the rabies virus (Gupta *et al.*, 2000). A recent study on the genetic variability of the lyssavirus P locus showed the presence of both conserved and highly divergent domains that could potentially be useful in diagnosis and serological typing of lyssaviruses (Nadin-Davis *et al.*, 2000).

1.6.2.6 Intergenic sequences

The intergenic regions of the rabies genome, N-P, P-M, M-G and between G-L are of variable length, notably 2, 5, 5 and 423 nucleotides, respectively. The genomic nucleotide sequences of the three short intergenic regions (N-P, P-M, and M-G) are identical in ERA, CVS-II and PV strains of the rabies virus (Wunner, 1988; Baer, 1991). However, in the PV strain, the intergenic region between the G and L genes extends to 423 nucleotides long and has been regarded as a vestigial gene or a pseudogene (Ψ) outside any selective immunological pressure (Ravkov *et al.*, 1995; Tordo, 1986b; 1988; Tordo and Poch, 1988). The pseudogene is a non-protein coding non-functional region that undergoes random mutation and this sequence is therefore considered to be a good neutral indicator of rabies evolution (Tordo *et al.*, 1986b; Wunner, 1988). Comparison of the intergenic gene sequences of rhabdoviruses seems to suggest that non-segmented negative stranded RNA viruses could have evolved by expansion (or contraction) of intergenic regions (Pringle, 1995).

1.7 RNA quasi-species

RNA viruses evolve at much higher rates than DNA viruses or eukaryotic genes. Many of the RNA viruses have the capacity to exchange genetic material with one another and to acquire other genes from their hosts. The lack of proofreading activity and the capacity of exchanging genes demonstrates the enormous capability that RNA viruses possess in adapting to changing host environments (Novella *et al.*, 1995). The RNA genomes normally exist as populations of different genomic variants known as quasi-species (Holland *et al.*, 1992; 1982; Drake, 1993; Domingo and Holland, 1994, 1997;

Morimoto *et al.*, 1998) and notably for rabies virus (Benmansour *et al.*, 1992; Kissi *et al.*, 1999). Evidence for the presence of "quasi-species" is shown by the presence of mutants in viral preparations, revertants in mutant stocks, frequent occurrence of antigenic variants (micro-variants) and genetic variation as observed among natural isolates of a virus.

2 Rabies pathogenesis

The rabies virus is normally introduced into warm-blooded mammals through animal bites with infectious saliva. Unusual forms of transmission of rabies have occurred and these include infection following a laboratory accident (Winkler *et al.*, 1973), human-to-human transmission (Fekadu *et al.*, 1996a) and post-vaccinal and corneal transplant infections (Gode and Bhide, 1988; Hoett *et al.*, 1979; WHO, 1994). Recently, a woman in the U.S. with a chronic skin condition was infected with recombinant vaccinia-rabies glycoprotein virus (Rupprecht *et al.*, 2001).

Entry of the rabies virus into mammalian cells occurs via acetylcholine receptors (Lentz *et al.*, 1982; Wunner and Reagan, 1984), the neural cell adhesion molecule (NACM) (Tholouze *et al.*, 1998) or the nerve-growth factor receptor (Tuffereu *et al.*, 1998). In addition, gangliosides, phospholipids and carbohydrates may be involved in virus entry (Tsiang, 1993). The viral spike glycoprotein (Figure 2) facilitates viral entry at acid pH (Gaudin *et al.*, 1993, 1991, 1999; Gaudin, 2000a, 2000b; Cleverley and Lenard, 1998). Rabies viruses remain at the site of bite for most of the incubation period (which is normally 10-90 days), but longer incubation periods were first described by Smith *et al.* (1991).

Upon entry into mammalian cells, the viruses then replicate initially in muscles and subcutaneous tissues before gaining access to the CNS (Murphy, 1985; Tordo *et al.*, 1998; Jackson, 2000). Microglial cells and astrocytes have been shown to support viral spread or persistence at the site of infection (Ray *et al.*, 1997a). The predominant target of rabies virus infection is the neuron, but virions can infect monocytes prior to reaching neurons in the spinal cord or the dorsal root ganglia (Morimoto *et al.*, 1998).

Rabies infection leads to a clinical syndrome classified as either paralytic (dumb) or the classical (furious/encephalitic) (Baer, 1991; Tordo *et al.*, 1998; Mrak and Young, 1994; Jackson, 1997; 2000). Furious rabies occurs in 80% of the cases and is associated with histopathological changes that include inflammation of the mid-brain, medulla, meninges and possible damage to the salivary and lacrimal glands. Clinical symptomatology of rabies and rabies-related virus infections are similar and include amongst others anxiety, restlessness, depression, feeling of tension, sense of foreboding, nightmares or inability to sleep and lack of concentration (Tordo *et al.*, 1998).

3. Global rabies situation

Rabies viruses are widespread throughout the world, but a small number of countries, including New Zealand, Papua New Guinea, Japan and other geographical entities that include islands, have been reported to be rabies-free (Blancou, 1988). However, bat lyssaviruses have recently been isolated from Australia and the U.K., countries that were previously thought to be rabies-free (Gould *et al.*, 1998; McColl *et al.*, 2000; Hanna *et al.*, 2000; Whitby *et al.*, 2000; Speare *et al.*, 1997).

3.1 Europe and Asia

In most developed countries of Europe, dog rabies has been successfully eliminated although it is now maintained in wildlife reservoir species. Across much of Europe, the red fox (*Vulpes vulpes*) is the main reservoir, while the raccoon dog (*Nictereutes procyonoides*) and the gray wolf (*Canis lupus*) are reservoirs in the more northern regions of Eurasia (Bourhy *et al.*, 1999, Rohde *et al.*, 1997; Crawford-Mikza, *et al.*, 1999; Krebs *et al.*, 2000b).

3.2 The Americas

In North America, raccoons (*Procyon lotor*), striped skunk (*Mephitis mephitis*), the red fox (*Vulpes vulpes*), the arctic fox (*Alopex lagopus*) and several species of bats are the major reservoirs of rabies (Nadin-Davis *et al.*, 2001; Smith and Baer, 1988; Smith, 1988; Ballerd and Krausman; 1997; Greenwood *et al.*, 1997). Rabies virus variants associated with silver-haired bats (*Lasionycteris noctivagans*) have been shown to be responsible for a high proportion of the recent human rabies cases in the United States (Dietzschold *et al.*, 2000; Morimoto *et al.*, 1996). Recent epidemiological data reviewed by Noah *et al.* (1998) and described by other authors (Krebs *et al.*, 2000a; Krebs *et al.*, 2000b) have shown human rabies statistics in the United States to be much lower than in the first half of the century, possibly indicating improved control and management of the disease. In Latin America, one of the haematophagous bats, the vampire bat (primarily *Desmondus rotundus*) is the major wildlife vector for rabies. Vampire bat rabies primarily affects cattle and occasionally humans (De Mattos *et al.*, 1996; Schneider *et al.*,

1996; Loza-Rubio *et al.*, 1999; Delpietro *et al.*, 2001) but it is dog rabies that contributes significantly to human rabies cases in these Latin American nations.

3.3 Rabies in Africa

The domestic dog is the primary reservoir for rabies and source for more than 95% of human exposures in many developing countries of Africa and Asia (WHO, 1992; Blancou, 1988). Wildlife carnivore species contribute significantly to rabies statistics (Swanepoel *et al.*, 1993; Bingham, 1999d). In comparison to Europe, there is less documented literature on rabies in Africa. Rabies is documented from a number of African countries that include; Kenya (Alexander *et al.*, 1993; Kat *et al.*, 1995; Kitala *et al.*, 2000), Nigeria (Aghomo and Rupprecht, 1990), Ethiopia (Mebatsion *et al.*, 1993), Tanzania (Maas, 1993; Cleaveland and Dye, 1995; Gascoyne *et al.*, 1993), Namibia (Laurenson *et al.*, 1997; Anon, 1973), Botswana (Tremlett *et al.*, 1994), Zimbabwe (Foggin 1988; Bingham, 1999a, 1999b; Central Veterinary Records (CVL) records, 2001; this study; Bingham *et al.*, 2001) and South Africa (Swanepoel *et al.*, 1993; von Teichman *et al.*, 1995; Nel *et al.*, 1998; Nel *et al.*, 1993).

3.3.1 Rabies in southern Africa

Foggin (1988) and Swanepoel *et al.* (1993) have reviewed the history of rabies in southern Africa. Southern Africa comprises the following countries; Zimbabwe, Botswana, Namibia, Mozambique, South Africa, Lesotho and Swaziland. The epidemiology of rabies in the southern African subcontinent is characterised by a large number of vector species and the presence of at least two distinct biotypes of the virus

(King *et al.*, 1994; von Teichman *et al.*, 1995; Nel *et al.*, 1997; Nel *et al.*, 1998; Swanepoel *et al.*, 1993). The two rabies biotypes that have been recognised in southern Africa are the canid and viverrid biotypes, respectively (King *et al.*, 1994; Tremlett *et al.*, 1994; Foggin, 1988). The canid biotype cycles in carnivores of the family *Canidae* (dogs, jackals and bat-eared foxes) and the mongoose biotype cycles in carnivores of the family *Herpestidae* (principally the yellow mongoose *Cynictis penicillata* and slender mongoose *Galerella sanguinea*) (King *et al.*, 1994; Swanepoel *et al.*, 1993). The latter is thought to be well adapted to small carnivores of southern Africa and appears unique to the African subcontinent. This virus shows considerable antigenic and genetic diversity in contrast to the canid biotype, which belongs to a lineage disseminated around the world with European colonisation (Von Teichman *et al.*, 1995).

It is postulated that before the 1950s, rabies was limited to the drier western parts of South Africa and Namibia, with species of the *Herpestidae* family being the main vector species (Swanepoel *et al.*, 1993). Dog rabies entered the subcontinent from Angola (in the west) to the Indian Ocean (in the east) in the 1950s (see Figure 3). This epidemic, believed to have been initiated by black-backed jackals, extended to the northwest and the Limpopo province of South Africa in both domestic dogs and cattle.

The first reports of rabies in Zimbabwe date back to 1902 (Edmonds, 1922), when dog rabies first appeared in the Bulawayo area, possibly introduced from Zambia. At the time, the rabies infection from Zambia did not become established in wildlife and was subsequently brought under control by the destruction of all feral and stray dogs. However, it is possible that rabies could have been present in Zimbabwe earlier than 1902, as depicted in traditional folklore (Edmonds, 1922).

Zimbabwe was rabies-free for about three decades, until another rabies infection in dogs at the Victoria Falls (in the 1950s). By 1954, the disease was present in most parts of Zimbabwe despite efforts to control its spread (Adamson, 1954). In the 1960s and 70s, most of the rabies cases were reported in the eastern and southern parts of the country. A notable increase in the incidence of rabies was observed in the 1970s and this was attributable to the disruption of immunisation programmes as a result of the liberation war being waged against the European colonialists (Kennedy, 1988; Lawrence, 1980). In the early 1980s, Foggin isolated rabies-related viruses of genotypes 3 and 4 from domestic mammals in Bulawayo and a common slit-faced bat (*Nycteris thebaica*) from Dorowa (Foggin, 1982; 1983).

Rabies epidemiological trends in Zimbabwe have shown the domestic dog (*Canis familiaris*) and jackal sp. to be the principal vectors of rabies (Foggin, 1988; Bingham, 1999a, 1999b; Central Veterinary Records (CVL) records, 2001; Cumming, 1982). Of the two vectors, domestic dogs are the most frequently diagnosed with rabies, and accounted for approximately 45% of all confirmed rabies cases between 1950 and 2000 (Foggin, 1988; Bingham *et al.*, 1999a; Bingham, 1999b). Dog rabies appears to be maintained in rural communal areas where 71% of the dog population is to be found (Brooks, 1990; Butler, 1998; Butler and Bingham, 2000). Urban rabies however, does not appear to be important except in the region of Mutare (eastern Zimbabwe), which is adjacent to large communal areas with susceptible dog populations. During the same period (1950-2000), approximately 25% of all confirmed rabies cases in Zimbabwe were diagnosed in jackals.

Two species of jackals, namely the side-striped (*Canis adustus*) and the black-backed (also called saddle-backed) (*Canis mesomelas*) jackals occur in Zimbabwe and are also frequently diagnosed with rabies (Bingham and Foggin, 1993; Bingham, 1999d). *C. adustus* occurs over most of the country except the dry western and southern regions whereas *C. mesomelas* occurs in the southern, western and central regions of lower rainfall (Skinner and Smithers, 1990). Jackal rabies has been confined mainly to commercial farming areas and contributed to rabies in other wildlife species (Foggin, 1988) and domestic cattle (Kennedy, 1988).

In South Africa, rabies epidemics have been maintained within various specific host species with defined geographical associations. These include the domestic dog in KwaZulu/Natal, black-backed jackals (*Canis mesomelas*) and bat-eared foxes (*Otocyon megalotis*) in the northern border areas and bat-eared foxes in the drier western regions and the southwest Cape (Swanepoel *et al.*, 1993). Yellow mongooses are the major maintenance species of rabies virus on the highveld plateau and Karoo areas of South Africa (King *et al.*, 1994; Taylor, 1993). The different geographical regions associated with the respective vector species are shown on the map of Zimbabwe and South Africa (Figure 4). Like in Zimbabwe and most likely elsewhere in Africa and the developing world, rabies is primarily diagnosed in dogs (typically more than 50% of the cases per year) (Onderstepoort Veterinary Institute (OVI), 2000).

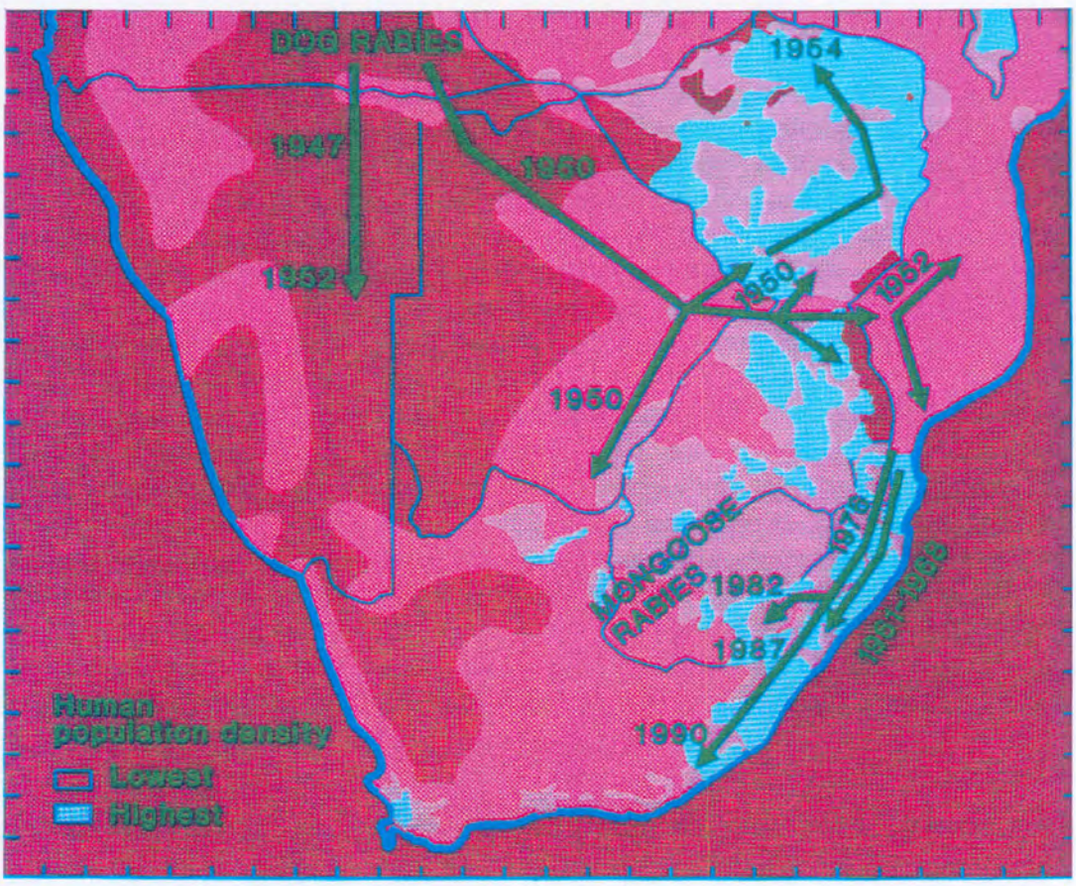

Figure 3: Map of southern Africa showing the routes for rabies movement (Swanepoel, 1994). The map shows the routes by which rabies was introduced into the southern African subcontinent from Angola in the early 1950s. Reprinted with permission from Oxford Publishers.

4 Rabies control

Today, vaccines against viral diseases are amongst the most successful prophylactic agents and their application has led to the control of many diseases including rabies, particularly in those countries that have been able to afford and organise effective immunisation programmes. Vaccination of domestic livestock and other animals, oral vaccination programs targeting wildlife, ongoing education programs and other activities conducted by public health authorities have contributed to reduced transmission of rabies variants to humans (Schneider, 1995; Krebs *et al.*, 1995; Meslin *et al.*, 1994; Haupt, 1999).

Just over one hundred years ago, Louis Pasteur developed a crude dessicated nervous tissue vaccine that resulted in the successful post-exposure protection of Joseph Meister (in 1885) and this initiated the era of rabies prevention (Hoenig, 1986). The most currently used method to reduce incidence of human rabies is vaccination of domestic animals and wildlife (Meslin *et al.*, 1994). Vaccination is particularly recommended for individuals at occupational risk such as veterinarians and staff of laboratories that handle the virus.

An ideal rabies vaccine should be immunogenic, genetically stable, inexpensive, thermostable, effective orally for diverse mammalian species and be capable of providing long lasting immunity (Rupprecht and Kieny, 1988; Wandeler, 1991; Kilm *et al.*, 1992; Artois *et al.*, 1993; 1997). Several types of vaccines have been developed and considered as potential candidates for rabies control. These are inactivated virus preparations currently in use for humans and domestic animals, inactivated nervous tissue origin (NTO), inactivated cell culture vaccines (Bunn, 1988), recombinant vaccines, sub-unit

vaccines (reviewed by Wunner *et al.*, 1983) and DNA-based vaccines (Drings *et al.*, 1999; Bahloul *et al.*, 1998; Perrin *et al.*, 2000; Lodmell and Ewalt, 2001). The major problem with inactivated vaccines is the cost involved in their production. In many developing nations, individuals are vaccinated with rabies vaccines produced in sheep, goat and suckling mouse brain, and UV or phenol inactivated (Lodmell and Ewalt, 2001).

Live attenuated viral vaccines are amongst the most successful and can provide life-long protection. In some cases, severe side effects are associated with live attenuated vaccines. One advantage of live attenuated vaccines is that they often elicit both humoral and cell-mediated immunity (CMI) thus providing better protection. The major disadvantage of live attenuated vaccines is the possible reversion of the virus strains to more reactive and virulent forms (Brown *et al.*, 1993) and as such these vaccines are not used in humans.

Sub-unit vaccines based on specific viral proteins or parts of a protein have been considered as potential candidates for vaccines (Wunner *et al.*, 1983). Sub-unit vaccines may not contain a large number of T-cell epitopes sufficient to elicit immune responses in all the members of the host population and have been used where attenuation has not been successful (Brown *et al.*, 1993).

The success of rabies vaccines has changed the global distribution of the disease. Europe and the Americas are areas of the world where rabies has been successfully controlled through immunisation programmes but equal measures of success have not been achieved in most developing nations of Africa and Asia. Factors that have contributed to the failure of rabies control in the developing nations are primarily the non-implementation of recommended procedures for rabies control (WHO report, 1989).

The failure of rabies control in Africa and Asia is also attributable to i) lack of proper infrastructure ii) limited access to preventive medicine iii) endemicity of rabies and iv) the fact that rabies does not feature as high priority in these nations (WHO report, 1984; Joshi and Bogel, 1988).

Successful field oral vaccination and protection against viral diseases and rabies in particular, have so far been achieved with either live-attenuated or live-recombinant virus vaccines e.g. Flury and Street Alabama Dufferin (SAD) strains. Strains of SAD_{Berne} have successfully contributed to fox rabies control in Europe (Schumacher *et al.*, 1993; Vuillame *et al.*, 1998), but were pathogenic in baboons (Bingham *et al.*, 1992). Clone SADB19 was selected for vaccine production because of its thermostability, high virus titre and low residual pathogenicity in rodents (Schneider, 1995) and found suitable for oral vaccination campaigns for carnivores against rabies (Vos *et* al., 1999). The SAG (SAD Avirulent Gif) strains were developed from SAD_{Berne} (Le Blouis *et al.*, 1990; Lafay *et al.*, 1994) and has been used extensively in France and Switzerland for fox rabies control (Aubert *et al*, 1994; Kilm *et al.*, 1992), but considered unsafe in dogs. A more stable variant SAG-2 was selected and used as an effective immunogen in dogs (Fekadu *et al.*, 1996b; Seghaier *et al.*, 1999; Masson *et al.*, 1996; Orciari *et al.*, 2001) and wild carnivore species (Bruyere *et al.*, 2000; Bingham, 1999c; Bingham *et al.*, 1993; Wandeler and Perry, 1993).

Rabies control in wildlife in Europe, North America and Canada has been facilitated by oral vaccination of free-ranging wildlife with genetically engineered vaccinia recombinants expressing glycoprotein (V-RG) (Wandeler, 1993b; Linhart *et al.*, 1997; Wilhem and Schneider, 1990; Hanlon *et al.*, 1998). Recombinant vaccines such as

VR-G were shown to be very stable, highly immunogenic and safe for target species (Rupprecht and Kieny, 1988). Such vaccines have also been shown to protect against rabies in domestic dogs (Matter *et al.*, 1999; Childs *et al.*, 1998). Field trials of recombinant rabies virus vaccines have led to their widespread use in protecting foxes and raccoons against the rabies virus even as baited food (Fekadu *et al.*, 1991; Masson *et al.*, 1999; Setien *et al.*, 1998; Hanlon *et al.*, 1998; Rupprecht *et al.*, 1993; Wiktor *et al.*, 1984b; Philippe *et al.*, 1998). Oral vaccination campaigns have recently been extended to bats (Setien *et al.*, 2000).

Co-expression of glycoprotein (Fu *et al.*, 1993) and nucleoprotein (Fu *et al.*, 1991) derived from recombinant baculovirus-infected cells were shown to be efficacious as oral vaccines (Fu *et al.*, 1993). However, vaccinia-recombinant nucleoprotein (V-RN) failed to confer similar protection (Fu *et al.*, 1994). Vaccinia rabies glycoprotein virus has recently been shown to cause adverse effects in hosts with altered immunocompetence and in persons for whom smallpox is contraindicated (Rupprecht *et al.*, 2001).

Since the mid-1990s, many researchers have turned their attention to the production of cheaper rabies vaccines. One such approach has been the use of DNA vaccines for pre and post-exposure treatment (Lodmell and Ewalt, 2001; Jallet *et al.*, 1999; Perrin *et al.*, 2000a, 2000b; Bahloul *et al.*, 1998; Ray *et al.*, 1997b). DNA immunisation elicits protective immunity and may have a future in human rabies immunisation (Lodmell *et al.*, 1998).

METHODS USED IN THIS STUDY.

5.1 Polymerase chain reaction (PCR)

Since its inception by Karry Mullis in the 1980s (Mullis, 1990), PCR has found many different applications. Current uses of PCR technology cover a wide spectrum, from the identification of novel genes from pathogens and other hosts (plants and animals), to its use in the fingerprinting of genomes and molecular genetic studies (Kapsa *et al.*, 1997). The major focus of PCR technology to date has been in the field of medicine and more specifically human pathology and microbiology (Vangrysperre and Clercq, 1996; Park *et al.*, 1997; Landergreen *et al.*, 1998). PCR is now routinely used in the detection of infectious diseases and genetic disorders (McDade and Anderson, 1996). In forensic science, PCR has facilitated obtaining highly specific data from very small samples such as hair, blood or skin fragments (Saiki *et al*, 1988). This technique has also found specific applications in rabies molecular epidemiological studies (Ermine *et al.*, 1990; Smith *et al.*, 1991; Smith and Seidel, 1993; Sacramento *et al*, 1991; Sacramento *et al.*, 1992; McColl *et al.*, 1993; Heaton *et al.*, 1997; Whitby *et al.*, 1997a, 1997b; Nadin-Davis *et al*, 1997; Nadin-Davis *et al.*, 1998, 1993; Nel *et al*, 1998; Warner *et al.*, 1996).

PCR is an *in vitro* method for the enzymatic synthesis of specific DNA sequences, using two oligonucleotide primers that hybridise to opposite strands and flank the region of interest in the target DNA. The activity of the polymerase enzyme results in an exponential accumulation of a specific product defined by the primers. The number of target DNA copies doubles at every cycle, 2^n, where n is the number of cycles. PCR was initially carried out using the Klenow fragment of *E. coli* DNA polymerase to extend the annealed primers but because this enzyme was thermolabile, it meant that fresh enzyme

had to be added at the end of each cycle. The automation of PCR was made possible by the introduction of a thermostable DNA polymerase (Landergren *et al.*, 1988; Mullis, 1990). Amplification is performed in a DNA thermal cycler and was as in our case used to amplify rabies and Mokola viral sequences from infected cells and tissues.

5.2 Nucleic acid sequencing

There are several approaches for DNA sequencing of PCR products and genomes of pathogens. Two approaches for nucleic acid sequencing namely the "dideoxy" or Sanger approach (Sanger, 1977) and the "chemical" or Maxam and Gilbert methods (1977) were published simultaneously. The dideoxy sequencing method is currently the method of choice for nucleic acid sequencing in most laboratories. Sequencing using the dideoxy method is carried out using a DNA polymerase enzyme that extends a primer along a single-stranded template in the presence of four dNTPs as substrates. The Klenow fragment of DNA polymerase 1 was originally used but other polymerases such as T7 DNA polymerase and thermostable DNA polymerases such as AmpliTaq DNA polymerase (Perkin Elmer Cetus) are now in use.

Chain elongation occurs in a random fashion and is terminated when a ddNTP is incorporated into the elongating chain, producing chains of varying lengths. The sequencing products will then be separated by polyacrylamide gel electrophoresis. In order to detect these chains, the most common methods employ the incorporation of a radiolabel or a fluorescent label (Newton and Graham, 1994). In our current investigation, we made use of DyeDeoxy terminators (Applied Biosystems) in order to understand molecular evolution of rabies and Mokola viruses from southern Africa.

5.3 Phylogenetic methods

5.3.1 What is molecular phylogeny?

Molecular phylogeny is the study of evolutionary relationships amongst organisms using molecular data such as nucleotide sequences, restriction fragments and amino acid sequences by either distance or character-based methods (Felsenstein, 1988; Miyamoto and Cracraft, 1991). A clear understanding of the genetic relationships of organisms can provide some information about their origin and geographical distribution of particular strains and also establish paths of their transmission (Brown, 1994). Furthemore, molecular phylogeny can assist in the reconstruction of the evolutionary history of specific genes (gene tree) and species (species tree), making it one of the most important aspects in the study of molecular evolution (Fitch, 1995; Nei, 1994; Li and Grauer, 1991). The objective of phylogenetic studies are to i. reconstruct the correct genealogical ties between organisms and ii. to estimate the time of divergence between organisms since they last shared a common ancestor (Smith, 1998).

Definitions of some of the key terms used in molecular phylogeny (adapted from Li and Grauer 1991).

- A phylogenetic tree is a graph composed of nodes and branches, in which one branch connects any two adjacent nodes.
- species tree; a phylogenetic tree that represents the evolutionary pathways of a group of species.
- gene tree; a phylogenetic tree that is constructed from one gene of each species.
- Operational taxonomic units (OTUs); any of the taxonomic units under study.

- Nodes represent the taxonomic units.
- Branches define the relationships between the units.
- The branching pattern of a tree is called the topology.
- When a node has two intermediate descendants lineages it is called bifurcating.
- Rooted tree; there exists a particular node from which a unique path leads to another node.
- Branch length represents the number of evolutionary changes that have occurred in that branch.
- An informative site is a nucleotide site that is phylogenetically informative only if it favours some trees over others.

5.3.2 Multiple sequence alignment

Alignment of nucleotide or amino acid sequences is currently one of the most utilised computational tools in molecular biology (Higgins and Sharp, 1988; Thompson *et al*; 1994). Alignments of molecular sequences are based on some functions that reward matches and penalise mismatches, gaps, insertions and deletions (indels). Sequence alignments are used to demonstrate homology or functional relationships between sequences (protein or DNA), may assist in predicting primers for PCR and restriction sites and are used as the primary data for construction of phylogenetic trees (Thompson *et al*., 1994). In this computational analysis, e.g. the ClustalW programme for DNA/sequence analysis, all pairs of sequences are aligned separately (pairwise alignment) in order to calculate a distance matrix, which is a measure of similarity of the sequences. It makes use of pairwise similarity scoring to find clusters of similar

sequences and pre-aligns the sequences into a guide tree. The final multiple alignment is carried out using the guide tree (dendrogram) and the output saved in different alignment formats (CLUSTAL, GCG, NBRF/PIR, PHYLIP AND GDE), which are compatible with a number of currently available software packages (Phylip, PAUP, GCG). Models to study the dynamics of nucleotide substitution include the Jukes and Cantor one-parameter model and Kimura 2-parameter model (Li and Grauer, 1991; Page and Holmes, 1996).

5.3.3 Distance methods

In distance methods, pairwise genetic distances are computed for all the species or OTUs under investigation in a similarity matrix. A tree is then constructed using an algorithm based on some functional relationships among the distance values. Clustering of OTUs starts with two OTUs with the smallest distance, and more distantly related OTUs are gradually added to the cluster.

UPGMA

One of the distance methods, the UPGMA (Unweighted Pair Group Method with Arithmetic Means) was originally proposed for taxonomic purposes, but it is now possible to use it for tree building (Sneath and Sokal, 1973). The method builds trees on the assumption that the rate of nucleotide or amino acid substitution is the same for all evolutionary lineages. When the rates are unequal in different lineages, the UPGMA method is not satisfactory and is replaced by methods such as the Fitch and Margoliash (Felsenstein, 1993) that allow rates of change to differ within the tree with minimal loss of phylogenetic information.

NJ method

The neighbour-joining (NJ) method is an extremely fast and highly efficient at finding the correct tree (Saitou and Nei, 1987). The NJ method sequentially finds the nearest pairs of neighbouring sequences that result in the shortest overall length of an unrooted tree. This method unlike the UPGMA, does not assume that all lineages have diverged by equal amounts (Swofford and Olsen, 1990). The tree inferred by the NJ method is an unrooted tree and can be rooted by i. assuming a degree of constancy in the molecular clock by placing the root along the longest branch and ii. including an outgroup, which is a sequence that definitely branches outside the tree (Felsenstein, 1981). ClustalW uses Neighbour-joining method which is more robust against the effects of unequal evolutionary rates in different lineages and which gives better estimates of trees (Thompson *et al.*, 1994). In practice, the NJ tree is often the same or very similar to the ME tree.

5.3.4 Character-based methods

In discrete-character methods, a tree is obtained by considering the evolutionary relationships of OTUs or DNA sequences at each character position. Of the discrete-character methods, the most popular method in phylogenetic reconstruction is that based on the principle of maximum parsimony (MP).

Maximum parsimony

In this method, the DNA (or amino acid sequences) of ancestral species are inferred from those of extant species, and the minimum number of evolutionary changes that are required to explain all the observed differences among the sequences is computed. The "best" tree is that which gives the shortest length, with the branch lengths being proportional to the number of nucleotide substitutions. Only those nucleotide sites at which there are at least two different kinds of nucleotides (parsimony-informative), each represented at least twice, are used. The other sites are eliminated from the analysis. This may slightly reduce the informativeness of the data. In practice, there may be two or more topologies that show the smallest tree and these are referred to as equally parsimonious trees (Li and Grauer, 1991; Nei, 1994).

The MP method is considered to be a good estimator of the tree if i. the extent of sequence divergence is small ($d < 0.1$), (ii) the rate of nucleotide substitution is more or less constant, (iii) there are no strong transition/transversion and G+C content bias and (iv) and if a small number of sequences (a few thousand nucleotides) are examined. Maximum parsimony may lead one to choose the wrong tree either because not all changes are equally likely (e.g. synonymous to non-synonymous changes) or because some branches are much longer than others.

Maximum likelihood (ML)

In this method, the nucleotides of all DNA sequences at each nucleotide site are considered separately. Likelihood methods differ from parsimony methods by employing standard statistical methods for a probabilistic model, which assumes the independence of

evolution at different sites. Potentially, the ML method is the most powerful and statistically most reliable method for phylogenetic inference but is of limited use because of the complexity of likelihood calculations and the large computing time taken to analyse data sets. In addition, the currently available ML methods include the assumption that every site in the sequence undergoes evolution at the same rate or that the rates are different among sites have to be specified by the user (Felsenstein, 1989; 1988; Weir, 1990).

5.3.5 Limitations of the tree building methods

Each of the methods for reconstructing phylogenetic trees has its own advantages and disadvantages. David Penny (Page and Holmes, 1996) suggested the following criteria to be desirable properties for a tree building method; i. efficiency (how fast is the method?), ii. power (how much data does the method need to produce a reasonable result?), iii. consistency (will it converge on the right answer given enough data?), iv. robustness (will minor variations of the method's assumptions result in poor estimates of the phylogeny?) and v. falsifiability (will the method tell us when its assumptions are violated?).

The current tree building methods emphasise on one or more of the above criteria at the expense of others. For example, maximum parsimony, though an efficient method, is time consuming and assumes a constant rate of evolution (which may not necessarily be the case). The efficiency of MP in obtaining the correct tree seems to be generally lower than that of the NJ and the ML methods (Weir, 1990). Maximum likelihood incorporates assumptions about evolution appropriate to the data, is time consuming and

limited to small data sets (Bourhy *et al.*, 1995) but is generally much better than the MP method and is nearly the same as or slightly better than the ME and NJ methods (Nei, 1994).

For a group of closely-related organisms such as the two rabies virus biotypes, it is appropriate to target a rapidly evolving part of the genome such as the glycoprotein and the G-L intergenic region, whereas for a more distantly related group, a slowly evolving gene such as the nucleoprotein gene would be desirable. In this study, we have utilised the NJ and Kitsch methods for deriving phylogenetic relationships of rabies and Mokola viruses included in this molecular epidemiological investigation.

5.4 Resampling methods: Bootstrapping

The topology of phylogenetic trees is only an estimate of the relationships amongst a group of strains or sequences. Most methods for estimating tree topologies do not give any information on the statistical significance of their structure. Two methods commonly used for estimating the variability of a true estimate are jacknifing and bootstrapping (Felsenstein, 1985; Dopazo, 1994).

The use of non-parametric statistical procedure referred to as bootstrapping is used to evaluate the significance of a particular topology. Bootstrapping involves resampling the data on which the tree was based to generate a distribution of data sets, from which a new tree is determined (Felsenstein, 1985). The frequency at which particular branches are observed in the resampled data sets then allows probability statements to be attached to them in the original tree. The bootstrap assumes that sites in the alignment have evolved independently on the same phylogeny (which may not be

realistic) and are identically distributed. In the other procedure, jacknifing, subsets are selected without replacement and consequently are much shorter than the original sequence (Weir, 1990)

6 Aims of this study

Rabies is a fatal zoonotic disease that occurs throughout most of the world in a variety of species (Blancou, 1988; WHO, 1992). The disease has been effectively controlled or eliminated in wildlife species in North America, Canada and Europe (Wandeler and Perry, 1993; Linhart *et al.*, 1997) through utilisation of information obtained using molecular biological approaches (Tordo *et al.*, 1986a; Sacramento *et al.*, 1992). To the contrary, rabies is prevalent in most developing nations of Africa, southeast Asia and the Indian subcontinent, and control has been curtailed by lack of infrastructure and proper resources required for combating the disease. In addition, limited access to preventive treatment, poor surveillance and low dog vaccination coverage, have further contributed to problems associated with rabies control in Africa (Smith and Seidel, 1993; Foggin, 1988; Coleman and Dye, 1996; Bingham, 1999d; Butler, 1998).

Recent advances in molecular biology have elucidated the rabies virus genome (Tordo *et al.*, 1986b) and this has facilitated discriminatory diagnosis of rabies variants (Nadin-Davis *et al.*, 1993, 1996; Nadin-Davis, 1998; Heaton *et al.*, 1997; Kamolvarin *et al.*, 1993; Nel *et al.*, 1998; Black *et al.*, 2000). Furthermore, molecular typing methods have led to a better understanding of the molecular epidemiology of the disease in some African countries-notably Ethiopia (Mebatsion *et al.*, 1993), South Africa (Nel *et al*,

1993; von Teichman *et al*, 1995) and more recently, Zimbabwe (this study; Bingham *et al*., 2001).

Comparative sequence analysis of rabies viruses have significantly advanced our knowledge of rabies epidemiology in the world. A very large amount of rabies virus nucleic acid sequence data has become available in the public domain (Genbank). This investigation is the first genetic analysis of Lyssaviruses of genotype 1 (rabies) and genotype 3 (Mokola viruses) from Zimbabwe, in order to gain a better understanding of Lyssavirus epidemiology in southern Africa. This information is crucial to design better public health and veterinary medicine practices.

Primary objectives

The primary focus of this study was to investigate the molecular epidemiology of lyssaviruses from Zimbabwe.

More specifically, we wanted to:

a) Genetically characterise rabies viruses from different mammalian hosts from canine and viverrid species from Zimbabwe and then make a comparison with those from South Africa,

b) Establish the genetic relationships of Mokola viruses from Zimbabwe and South Africa, and

c) Establish the role of the rabies variants in rabies epizootics occurring in the southern African subcontinent.

CHAPTER 2

MOLECULAR EPIDEMIOLOGY OF CANINE RABIES IN ZIMBABWE AND SOUTH AFRICA.

2.1 Summary

The epidemiology of rabies in southern Africa is complex, due to a large number of vector species and the presence of at least two distinct biotypes of the virus. Our objective was to contribute to the understanding of the epidemiology of rabies in the African subcontinent by studying the genetic relatedness of 89 rabies virus isolates from this region. In this study, we have focused on an analysis of viruses specific to canine host species (canid biotype) throughout South Africa and Zimbabwe. All the southern African canid viruses were found to be closely related and no apparent general distinction could be made between viruses from any of the canid species. Despite the general overall similarity between virus isolates, certain phylogenetic groupings were apparent and by association with host species, geography and year of isolation, some of the groups could be identified as particular epidemiological cycles. This pattern would indicate that the canid rabies virus lineage is opportunistic within whatever canid host population is available and ecologically capable of sustaining prolonged cycles. This molecular epidemiological study presents the first comprehensive comparison of rabies viruses from South Africa and Zimbabwe and has demonstrated the need for multinational approaches towards the control of this important zoonotic disease in Africa.

Data from this chapter has been submitted to Virus Research.

2.2 Introduction

The molecular organisation and taxonomy of the rabies virus and disease epidemiology have been covered in detail in Chapter One. Rabies is one of the most important and widespread viral zoonoses and is with the exception of a few countries, a truly global dilemma (Blancou, 1988; WHO, 1989 and 1991). It is an important disease in southern Africa and elsewhere in Africa because of the high costs of prevention and post-exposure treatment. In southern Africa, rabies epizootics involve a large number of mammalian species (reviewed by Swanepoel *et al.*, 1993). In this study, our focus has been on canine rabies in the southern African countries of South Africa and Zimbabwe specifically and the different geographical regions associated with the various reservoir species in these countries are shown in Figure 4.

Of the two rabies biotypes circulating in mammalian species in southern Africa (King *et al.*, 1994) and most probably elsewhere in Africa, it is the viverrid biotype that is well adapted to mongooses of southern Africa and therefore appears unique to the African subcontinent. The viverrid rabies virus shows considerable antigenic and genetic diversity to the canid biotype, which seems to have been introduced recently during European colonisation of southern Africa (Smith *et al.*, 1992).

In order to successfully combat rabies in Africa (and elsewhere in the world), a clear understanding of the disease epidemiology is required. The way in which viruses change and adapt to different host species and the role that time and geographical factors play in virus proliferation, are issues underlying our understanding of the disease (Nadin-

Davis *et al.*, 1993; De Mattos *et al.*, 1996; Bourhy *et al.*, 1999; Badrane *et al.*, 2001).

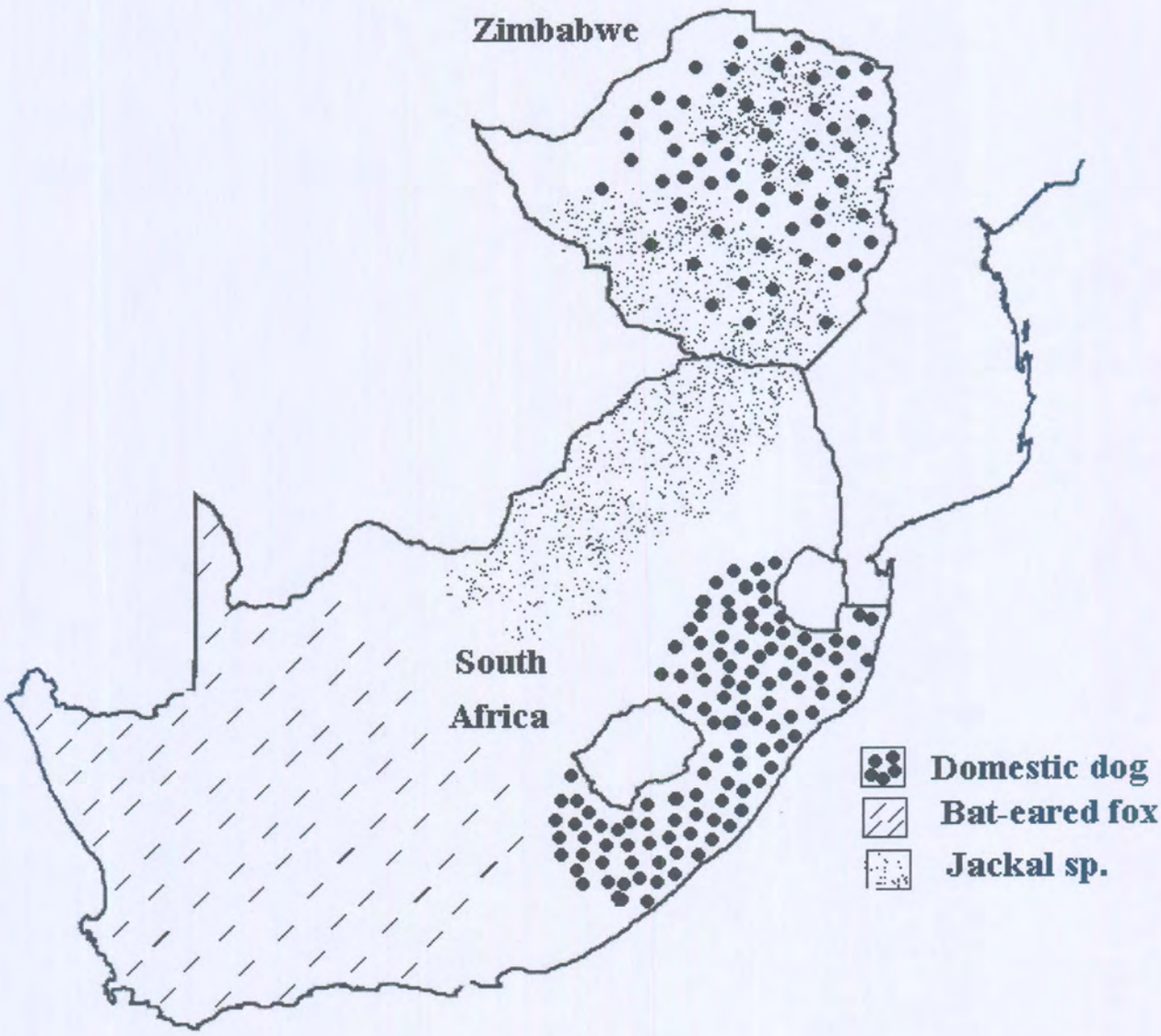

Figure 4: Map of Zimbabwe and South Africa showing the approximate enzootic regions for canine rabies.

In recent years, a most useful approach towards elucidating viral epidemiology has been through molecular sequence analysis of specific parts of the genomes of virus isolates. For rabies, sequences of genes of the nucleoprotein (Smith *et al.*, 1992; Tordo and Kouknetzoff, 1993; Ito *et al.*, 2001a), the phosphoprotein (Nadin-Davis *et al.*, 2000) and the G-L intergenic region (Sacramento *et al.*, 1991; Sacramento *et al.*, 1992) have for instance been useful in determining molecular epidemiology of rabies in Europe (Bourhy *et al.*, 1999), the Americas (Smith *et al.*, 1995; 1992, 1993; 1995; Smith, 1996; Nadin-Davis *et al.*, 1999), the Middle East (David *et al.*, 2000) and Africa (von Teichman *et al.*, 1995; Warner *et al.*, 1996; this study, Nel *et al.*, 1993).

In comparison to Zimbabwe and other southern African countries, the molecular epidemiology of rabies in South Africa is more comprehensively documented (von Teichman *et al.*, 1995; Nel *et al.*, 1997; Nel *et al.*, 1998). However, these studies included a limited number of virus isolates and were focused on the genetic distinction between the canid and viverrid virus biotypes (Von Teichman *et al.*, 1995; Nel *et al.*, 1997; Nel *et al.*, 1998). Although many different canid species in different geographical regions are pivotal in the maintenance and spread of rabies, the genetic relationships between the viruses involved in these instances has never been fully investigated.

The objective of this investigation was to genetically characterise rabies virus isolates obtained from the principal canine hosts throughout Zimbabwe and South Africa. The regional distribution of variants identified is detailed and their phylogenetic relationships determined using nucleotide sequence data of the cytoplasmic domain of the glycoprotein and the G-L intergenic region. This study therefore provides the first description of molecular epidemiology of rabies in Zimbabwe and a first regional

comparison of virus relationships. It was demonstrated from this investigation that viruses from domestic dogs and jackals are closely related and that either or both these species could be the maintenance species of the virus. Specific reservoirs could be identified in certain cases and epidemiological links (i.e. close relatedness of virus isolates involved in rabies epidemics) of rabies cycles in the subcontinent were indicated.

2.3 Materials and methods

2.3.1 Viruses and virus isolation

Eight-nine rabies virus isolates obtained from domestic and wild carnivore species from different regions of Zimbabwe and South Africa were analysed. The isolates were collected by the Central Veterinary Laboratory in Harare, the Onderstepoort Veterinary Institute, Pretoria and the KwaZulu-Natal Provincial laboratory at Allerton, Pietermaritzburg. The year of isolation, host species and geographical origins of the viruses are shown in Table 2 and Figure 4. The virus isolates were shown to be rabies by a fluorescent antibody test (FAT) (Meslin and Kaplan, 1996). The FAT is the most widely used method for diagnosing rabies infections in humans and animals. In brief, the suspect tissue smear were fixed with acetone, dried at 37°C and then incubated with a conjugate that is immunologically bound to rabies antinucleocapsid. The smears were washed in phosphate buffered saline (pH 7.2), mounted and examined with a fluorescent microscope. The viruses were further shown to belong to the canid biotype by Mab typing (Bingham *et al.*, 2001). Rabies virus isolates from South Africa were passaged in suckling mice and stored in the form of lyophilised 20% mouse brain material in the collection of the Rabies Unit, Onderstepoort Veterinary Institute, Pretoria. Isolates from

Zimbabwe were passaged in suckling mice and the infected-whole brains stored in liquid N_2.

Table 2: Rabies virus isolates included in this study.

Virus	Sample origin							
	Lab Ref. #	Species	Country	Yr of	Locality isolation	Coordinates (long-lat)	Genbank	Group
1	19319	Jackal (C. adustus)	Zimbabwe	1991	Selous	30°28'-18°09'	AF177082	A1
2	20548	Canis familiaris	Zimbabwe	1992	Seke CL	31°12'-18°09'	AF177061	
3	19301	Jackal (C. adustus)	Zimbabwe	1991	Chegutu	30°14'-18°09'	AF177081	
4	20896	Canis familiaris	Zimbabwe	1992	Chatsworth	30°48'-19°27'	AF177062	
5	16838	Canis familiaris	Zimbabwe	1987	Shangani	29°17''-19°27'	AF177058	
6	14937	Jackal (C. adustus)	Zimbabwe	1984	Eiffel Flats	30°17'-18°24'	AF177077	
7	17711	Jackal (C. mesomelas)	Zimbabwe	1988	Turk Mine	28°49'-19°27'	AF177078	
8	19344	Jackal (C. mesomelas)	Zimbabwe	1991	Fort Rixon	29°16'-20°13'	AF303077	
9	14702	Canis familiaris	Zimbabwe	1984	Beatrice	30°59'-18°28'	AF177053	
10	20813	Canis familiaris	Zimbabwe	1992	Mount Darwin	31°28'-16°46'	AF177063	
11	14917	Jackal (C. adustus)	Zimbabwe	1984	Mukwichi CL	29°50'-16°31'	AF177076	
12	16234	Canis familiaris	Zimbabwe	1986	Lukanka, Nyamandhlovu	28°09'-19°40'	AF177054	
13	21533	Jackal (C. adustus)	Zimbabwe	1993	Mhangura	30°11'-17°02'	AF177085	
14	24137	Jackal (C. adustus)	Zimbabwe	1996	Mtepatepa	31°22'-16°56'	AF177095	
15	988/94	Canis familiaris	South Africa	1994	Piketberg	18°37'-32°42'	AF303068	
16	306/98	Canis mesomelas	South Africa	1998	Warmbaths	28°07'-24°51'	AF177105	A2
17	479/96	Canis familiaris	South Africa	1996	Thabazimbi	27°24'-24°43'	AF303070	
18	675/99	Canis familiaris	South Africa	1999	Mankwe, Rustenberg	27°23'-25°06'	AF303071	
19	602/90	Canis mesomelas	South Africa	1990	Thabazimbi	26°48'-24°45'	AF177110	
20	598/90	Canis mesomelas	South Africa	1990	Soutpansberg	29°53'-22°50'	AF177117	
21	A90/77	Canis familiaris	South Africa	1990	Lower Umfolozi	32°00'-28°35'	AF177102	
22	504/96	Canis mesomelas	South Africa	1996	Leeupoort, Warmbaths	27°44'-24°56'	AF177108	
23	19286	Jackal (C. mesomelas)	Zimbabwe	1991	Mwenezi	30°58'-21°45'	AF177080	
24	20034	Canis familiaris	Zimbabwe	1991	Bvumba, Mutare	32°45'-19°02'	AF177059	
25	17722	Jackal (C. mesomelas)	Zimbabwe	1988	Gwanda	28°58'-20°41'	AF177079	
26	1403/80	Canis mesomelas	South Africa	1980	Soutpansberg	29°24'-22°32'	AF177116	

27	1627/80	Canis mesomelas	South Africa	1980	Soutpansberg	29°13'-22°38'	AF177104	
28	A90/352	Canis familiaris	South Africa	1990	Durban	30°50'-29°50'	AF177100	A3
29	A95/755	Canis familiaris	South Africa	1995	Amanzimtoti	30°54'-30°04'	AF303081	
30	101/90	Canis mesomelas	South Africa	1990	Soutpansberg	28°55'-22°48'	AF079902	
31	A90/57	Canis familiaris	South Africa	1990	Durban	30°53'-30°03'	AF177101	
32	596/99	Canis mesomelas	South Africa	1999	Piet Retief	30°12'-29°33'	AF303063	
33	224/98	Canis familiaris	South Africa	1998	Ermelo	29°59'-26°31'	AF177098	
34	373/97	Canis familiaris	South Africa	1997	Barbeton	31°48'-25°42'	AF303069	
35	378/90	Canis mesomelas	South Africa	1990	Potgietersrus	28°31'-23°10'	AF177106	
36	A80/56	Canis familiaris	South Africa	1980	Pietermaritzburg	30°25'-29°37'	AFO79904	
37	24465	Canis familiaris	Zimbabwe	1996	Wedza CL	31°43'-18°52'	AF177074	B1
38	23667	Jackal (C. adustus)	Zimbabwe	1995	Marondera	31°41'-18°11'	AF177093	
39	23357	Jackal (C. adustus)	Zimbabwe	1995	Wedza	31°21'-18°35'	AF177090	
40	22642	Jackal (C. adustus)	Zimbabwe	1994	Glendale	31°06'-17°19'	AF177088	
41	21532	Jackal (C. adustus)	Zimbabwe	1993	Arcturus	31°20'-17°41'	AF177067	
42	21147	Jackal (C. adustus)	Zimbabwe	1992	Arcturus	31°25'-17°44'	AF177084	
43	21819	Jackal (C. adustus)	Zimbabwe	1993	Mazowe	31°00'-17°29'	AF177087	
44	21111	Jackal (C. adustus)	Zimbabwe	1992	Bromley	31°16'-17°58'	AF177083	
45	23578	Jackal (C. adustus)	Zimbabwe	1995	Karoi	29°42'-16°40'	AF177092	
46	24307	Jackal (C. mesomelas)	Zimbabwe	1996	Featherstone	31°03'-18°33'	AF177096	B2
47	21467	Canis familiaris	Zimbabwe	1993	Goromonzi	31°22'-17°53'	AF177066	
48	23275	Jackal (C. adustus)	Zimbabwe	1995	Bindura	31°21'-16°56'	AF177089	
49	23652	Canis familiaris	Zimbabwe	1995	Mutare	32°40'-18°58'	AF177070	
50	21428	Canis familiaris	Zimbabwe	1993	Chipinge	32°43'-20°22'	AF177065	
51	24299	Canis familiaris	Zimbabwe	1996	Musikavanhu CL	32°28'-20°21'	AF177073	
52	23895	Jackal (C. adustus)	Zimbabwe	1995	Headlands	32°06'-18°17'	AF177094	
53	24132	Canis familiaris	Zimbabwe	1996	Mutare	32°40'-19°00'	AF177072	
54	19273	Jackal (C. mesomelas)	Zimbabwe	1991	Chiredzi	31°58'-20°54'	AF303076	B3
55	19901	Jackal (C. mesomelas)	Zimbabwe	1991	Triangle	31°29'-21°00'	AF303075	
56	19347	Canis familiaris	Zimbabwe	1991	Zvishavane	30°05'-20°20'	AF303079	
57	19385	Canis familiaris	Zimbabwe	1991	Zaka	31°34'-20°12'	AF303080	
58	24505	Canis familiaris	Zimbabwe	1996	Gutu	31°10'-19°37'	AF177075	
59	19366	Canis familiaris	Zimbabwe	1991	Chiredzi	31°49'-20°46'	AF303078	
60	16387	Canis familiaris	Zimbabwe	1986	Zhombe CL	29°22'-18°41'	AF177057	
61	16254	Canis familiaris	Zimbabwe	1986	Odzi	32°25'-18°47'	AF177055	
62	733/99	Canis familiaris	South Africa	1999	Brits	27°33'-25°14'	AF303067	

63	22547	Canis familiaris	Zimbabwe	1994	Kumutsenzere, Masoso CL	31°47'-16°22'	AF177070	B4
64	21057	Canis familiaris	Zimbabwe	1992	Muzarabani CL	31°12'-16°19'	AF177064	
65	20519	Canis familiaris	Zimbabwe	1992	Lower Gweru	29°22'-19°17'	AF177060	
66	21869	Canis familiaris	Zimbabwe	1993	Nyakasoro, Pfungwe CL	32°15'-16°49'	AF177069	
67	16347	Canis familiaris	Zimbabwe	1986	Troutbeck, Nyanga	32°48'-18°11'	AF177056	
68	21570	Canis familiaris	Zimbabwe	1993	Binga	27°20'-17°37'	AF177068	
69	377/90	Otocyon megalotis	South Africa	1990	Gordonia	21°10'-28°03'	AF177119	
70	773/95	Canis mesomelas	South Africa	1995	Tosca, Vryburg	23°35'-25°47'	AF177111	
71	669/99	Canis mesomelas	South Africa	1999	Pietersburg	29°27'-23°47'	AF303062	C
72	673/99	Canis mesomelas	South Africa	1999	Potgietersrus	28°36'-22°43'	AF303061	
73	717/99	Canis mesomelas	South Africa	1999	Pietersburg	29°29'-23°42'	AF303064	
74	631/99	Canis mesomelas	South Africa	1999	Soutpansberg	30°05'-22°03'	AF303060	
75	8/96	Canis familiaris	South Africa	1996	Roetran, Potgietersrus	29°06'-24°33'	AF177103	
76	23374	Jackal (C. mesomelas)	Zimbabwe	1995	Bulawayo	28°47'-20°15'	AF177091	
77	820/94	Canis familiaris	South Africa	1994	Stella, Vryburg	24°45'-26°08'	AF177118	
78	21579	Jackal (C. mesomelas)	Zimbabwe	1993	Tsholotsho CL	27°26'-19°31'	AF177086	
79	446/95	Canis mesomelas	South Africa	1995	Stella, Vryburg	24°36'-26°20	AF177107	
80	214/98	Otocyon megalotis	South Africa	1998	Uitenhage	25°00'-33°23'	AF177097	D
81	774/95	Otocyon megalotis	South Africa	1995	Carnavon	21°56'-31°13'	AF177115	
82	414/96	Otocyon megalotis	South Africa	1996	Beaufort West	22°47'-32°22'	AF177112	
83	462/90	Otocyon megalotis	South Africa	1999	De Aar	24°08'-30°48'	AF303074	
84	491/98	Otocyon megalotis	South Africa	1998	Petrusburg	25°29-29°23'	AF303059	
85	469/99	Otocyon megalotis	South Africa	1999	Kimberley	24°27'-28°47'	AF303073	
86	578/95	Otocyon megalotis	South Africa	1995	Strydenburg, Hopetown	23°46'-29°55'	AF177113	
87	615/91	Otocyon megalotis	South Africa	1991	Britstown	23°27'-30°50'	AF079906	
88	544/90	Canis mesomelas	South Africa	1990	Piketberg	19°01'-33°03'	AF177109	
89	958/95	Canis familiaris	South Africa	1995	Hofmeyr	25°37'-31°33'	AF303072	

(1) Lab reference numbers: Isolates from Zimbabwe use the Harare Central Veterinary Laboratory rabies reference; isolates from South Africa use the OVI Rabies Unit number or are prefixed by "A" to indicate the Allerton Provincial Veterinary Laboratory reference numbers.

2.3.2 Primers for RT-PCR

All oligonucleotides for RT-PCR and sequencing were produced by Life Technologies (U.K.) and used without further purification. PCR amplification and sequencing were performed with primer sets G(+) (4665-4685) and L(-) (5520-5543) (Table 3). Two additional internal primers P1(+) (4997-5018) and P5(-) (5131-5154) were also used in PCR and sequencing (von Teichman *et al.*, 1995). Annealing positions are numbered according to the published PV sequence used as the reference genome (Tordo *et al.*, 1986a, 1988). The symbols (+) and (-) refer to genomic and messenger sense respectively.

Table 3: Designation, sequence and location of oligonucleotide primers used for cDNA synthesis, PCR and sequencing of the G-L intergenic region of rabies virus isolates.

Name	Sequence (5'-3')	use(s)
G(+)	$_{4665}$GACTTGGGTCTCCCAACTGGGG$_{4687}$	RT- PCR, sequencing
L (-)	$_{5543}$CAAAGGAGAGTTGAGATTGTAGTC$_{5566}$	PCR, sequencing
P1(+)	$_{4997}$CAACTGGGTAGATTGAAGAGTC$_{5018}$	PCR, sequencing
P4(+)	$_{5026}$TTTTCATTAATCCTCTCAGTTGATC$_{5050}$	PCR, sequencing
P5(-)	$_{5131}$TTTGTCTACAACTGTTGGTGTCAG$_{5154}$	PCR, sequencing

(After Von Teichman *et al.*, 1995)

2.3.3 **Total RNA extraction**

Total RNA was extracted according to Chomczsynki (1993). Briefly, 500 µg of lyophilised mouse brain material and 1 ml TRIReagent (Molecular Research Center, U.S.A) were mixed and kept at RT for 5 minutes. Subsequently, 200 µl of chloroform were added to the homogenate, shaken vigorously for a brief moment and then kept at RT for 2-3 minutes. The mixture was centrifuged at 13 000 rpm at 2-8°C for 15 min and 600 µl of the aqueous phase transferred to a sterile 1.5 ml eppendorf tube. Total RNA was precipitated with addition of 500 µl of isopropanol to the supernatant at RT and kept for 10 min. The RNA was pelleted at 13 000 rpm for 15 min, washed with 70% ethanol, air-dried and then dissolved in 50 µl RNase-free water. Five microlitres (0.4-0.8 µg) of purified total RNA was used in the reverse transcription (RT) reaction.

2.3.4 cDNA synthesis

Complementary DNA (cDNA) synthesis was carried out at 37°C for 90 minutes as described elsewhere (Sacramento *et al*, 1991; Von Teichman *et al.*, 1995). In brief, total brain or cellular RNA (approximately 0.4 µg) and 100 ng G (+) primer (approximately 15 pmol) were denatured at 65°C for 5 min and immediately cooled on ice. A previously assembled reaction mixture comprising 1X First strand buffer (Life Technologies, U.K.), 200U M-MuLV (Life Technologies, U.K.), 1 mM dNTPs (Promega, U.S.A) and 10U RNasin inhibitor (Promega, U.S.A.) was mixed with the denatured RNA and reverse transcription carried out according to Sacramento *et al* (1991). The reaction mixture

(cDNA/RNA hybrid) was diluted 10-fold with RNase-free water and 5 µl used for amplification.

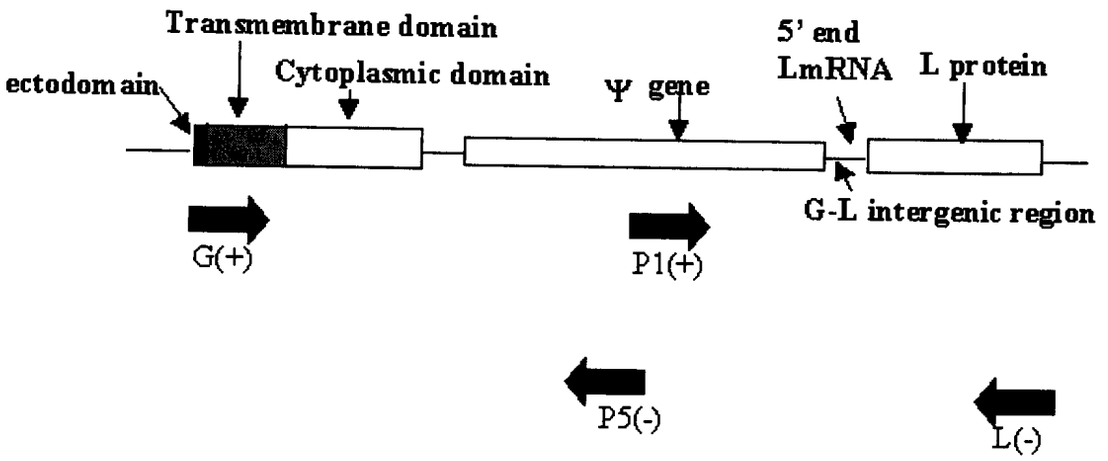

Figure 5: The rabies virus G-L intergenic region and the flanking domains. Ψ is the pseudogene. The location and orientation of the primers used in amplification and sequencing are shown (after Von Teichman *et al.*, 1995).

2.3.5 PCR

Amplification was performed in a DNA Thermal Cycler (Perkin-Elmer GeneAmp PCR 2400). Five µl of the cDNA was added to 19 µl of PCR mixture comprising 10X

PCR buffer (5 µl) (Promega), 200 µM each of four dNTPS (Promega), 40 pmol each of the G (+) and L (-) primers (Gibco BRL) and 1 U Taq DNA polymerase and then made up to 50 µl with sterile distilled water. The amplification involved an initial denaturation at 94°C for 2 min (hot start), then 30 cycles of [denaturation at 94°C for 50 s; primer annealing at 42°C for 90 s and primer extension at 72°C for 2 min], and a final extension of 7 min. On completion of the amplification programme, amplicons were analysed by ethidium bromide-stained agarose gel electrophoresis (Figure 6).

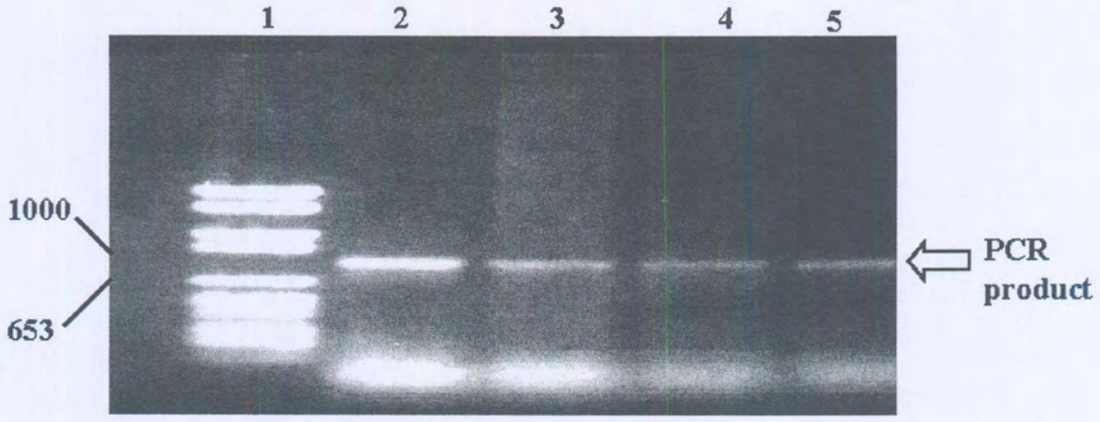

Figure 6: RT-PCR amplification of rabies brain-infected material (described in Methods above). In lane 1 is MW marker vi (Roche Diagnostics, Germany). An amplicon of expected size (approximately 850 bp) was generated (lanes 2-isolate 23652; lane 3-isolate 22547; lane 4-isolate 16347 and lane 5-isolate 16234).

2.3.6 Nucleotide sequencing

PCR amplicons were gel-purified using GeneClean (Bio101, U.S.A) according to the manufacturer's instructions. The excised agarose fragment was mixed with three volumes of 3M sodium iodide solution and melted at 55°C after which 7 µl of the silica matrix, glassmilk, was added to the solution. Following incubation on ice for five minutes, the silica-bound DNA was pelleted, washed three times with ice cold wash buffer and the DNA eluted from the silica matrix at 55°C in a final volume of 10 µl sterile distilled water. One µl of elute was quantified on 1% agarose gels.

Cycle-sequencing reactions were done with the BigDye Terminator system (PE Applied Biosystems) using the BigDye Terminator system and analysed on an ABI 377 DNA sequencer (PE Applied Biosystems) with the same forward and reverse primers as in the preceding RT-PCR step according to the manufacturer's instructions. A reaction mixture comprising terminator ready reaction mixture (8 µl), 50 ng of the purified PCR product, 3.2 picomoles of the primer (forward and reverse) was made up to 20 µl with sterile distilled water and cycle-sequenced in a Perkin Elmer GeneAmp 9600, using the following cycling parameters [96°C for 10 s, 50°C for 5 s, 60°C for 4 min] for 25 cycles and then analysed on an ABI377 sequencer. Forward and reverse sequences were analysed and compared with the use of the Sequence Navigator Software (Version 1, PE Applied Systems).

2.3.7 Phylogenetic analysis

Phylogenetic analysis of the rabies virus isolates based on the nucleic acid sequences of the cytoplasmic domain of the glycoprotein and the G-L intergenic region was undertaken using the PHYLIP software package (Felsenstein, 1993). The viral nucleic acid sequences were aligned in ClustalW (Thompson *et al.*, 1994). Firstly, a distance matrix was constructed by pairwise alignment of the sequences. Similarity scores were calculated as the number of identical residues in the best alignment of 2 sequences minus a fixed gap penalty of 10 (Thompson *et al.*, 1994). A guide tree was then constructed from the distance matrix using the neighbour-joining method of Saitou and Nei (1987). The branch lengths were proportional to the estimated divergence along each of the branches and a guide tree was then used in the final multiple alignments. These alignments were then stored in a NBRF/PIR file format, which is recognised by ClustalW for calculating phylogenetic trees. The bootstrap method (Felsenstein, 1985) was used in combination with the neighbour-joining method to estimate the confidence levels of the phylogenies. Bootstrap supported trees were constructed using a random seed generator of 111 and 1000 replicate data sets to assess the confidence limits of the branching pattern using SEQBOOT (Felsenstein, 1993). Bootstrap values of >70% are generally considered as providing evidence for phylogenetic groups (Hills and Bull, 1993). Similarly, genetic distances were derived with the DNAdist program applying Kimura's (1980) two-parameter distance model and a phylogenetic tree generated as an output option of the Kitsch method. The program TREEVIEW was used to view the

graphic outputs of the NJ (Figure 7) and Kitsch trees (Figure 8) (Page, 1996). The topologies of the NJ and Kitsch trees were similar and the differences are highlighted.

2.3.8 Nucleotide sequence accession numbers

All the sequences generated during the course of this study were submitted to Genbank and assigned accession numbers (Table 2).

2.4 RESULTS

2.4.1 Virus isolates, virus propagation, RNA extraction, cDNA synthesis and amplification.

In order to determine the genetic relatedness among rabies viruses derived from different mammalian species and geographical regions of southern Africa, virus isolates were collected and PCR-generated amplicons of approximately 850 bp from viral cDNA templates were sequenced. The G-L primer pair could amplify all the isolates in this investigation. Analyses of these nucleotide sequences, encoding the cytoplasmic domain of the glycoprotein plus the G-L intergenic region of the glycoprotein, were performed as described in Methods.

2.4.2 Phylogenetic analysis

The sequence variation within the G-L intergenic region of the African rabies viruses showed that the canid virus isolates from southern Africa are closely related with an average sequence homology of >92%, despite the varied geographical origins of the

virus isolates (Appendix 2). For this region of the genome, southern African viruses displayed an average sequence divergence of 18.3% from the PV (Tordo *et al.*, 1988). In addition, all the southern African isolates were found to lack one TTP (Appendix 1, nucleotides 219-235), as previously shown for street viruses (Morimoto *et al.*, 1989; Sacramento *et al.*, 1992). The G stop signal has been shown to be more of a transcriptional, rather than evolutionary significance (Sacramento *et al.*, 1992).

The most important finding of this study is the fact that all the southern African canid viruses were closely related and no apparent general distinction could be made between viruses from any of the canid species. It was clear that the assemblage of southern African viruses is distinct from the Pasteur virus (PV), which was included as an outgroup. Despite the general overall similarity between virus isolates certain groups were apparent. These groups have been labelled as indicated in Figure 7. However, it is emphasised that these groups are not all supported by high bootstrap values; rather than defined phylogenetic classifications they are indicated to assist with epidemiological description in the following text. By association with host species, geography and year of isolation, some of the groups could be identified as particular epidemiological cycles. The term group refers to the loose clusters observed in the phylogenetic tree (Figure 7).

Group A was made up of isolates from domestic dogs (17) and jackals (19) over a geographic region presented as a North-South column through Zimbabwe and South Africa south to the Indian ocean (Figure 9a). The group (A) could be subdivided into 3 groups (A1-A3) (Figure 7). Group A1 was composed of isolates from dogs (9) and jackals (8) from eastern through central and northern Zimbabwe, with one isolate from a

dog in South Africa. Group A2 was made up of viruses isolated from a dog (1) and jackals (2) in eastern and southeastern Zimbabwe and from dogs (3) and jackals (6) from South Africa in areas bordering southern Zimbabwe through to northwestern South Africa. Group A3 was made up of viruses from dogs (6) and jackals (3) from eastern South Africa.

The proposed group B was, with the exception of one South African dog isolate (#62, subgroup B3), exclusively composed of isolates from Zimbabwean dogs (18) and jackals (13) (Figure 7, Figure 9b). A statistically significant bootstrap value of 93% separated the subgroups B1 and B2 (both composed of Zimbabwean isolates), which may otherwise have been regarded as one subgroup. Group B1 was composed of jackal isolates (8) and one dog isolate whereas B2 was represented by more dog isolates (5) together with three jackal isolates. Apart from 1 dog isolate from northern South Africa, B3 included 8 Zimbabwean isolates, 6 from dogs and 2 from jackals. Genetic group B4 was comprised of viruses made from domestic dogs in northeastern Zimbabwe (see geographical map, Figure 9b).

Group C (Figure 9c) was found to contain viruses from southern Zimbabwe (2 jackal isolates) and bordering northern South Africa (2 dog and 5 jackal isolates). Group C was the second phylogenetic group containing viruses from both southern Zimbabwe and the northern bordering areas of South Africa (compare with A). Bat-eared fox isolates appeared to be part of a distinct group (group D), originating from south-central South Africa as indicated in Figure 9d. Thus only two groups of viruses were each

associated with a single host species, *viz*: B4 (dogs, Zimbabwe) and D (bat-eared foxes, South Africa).

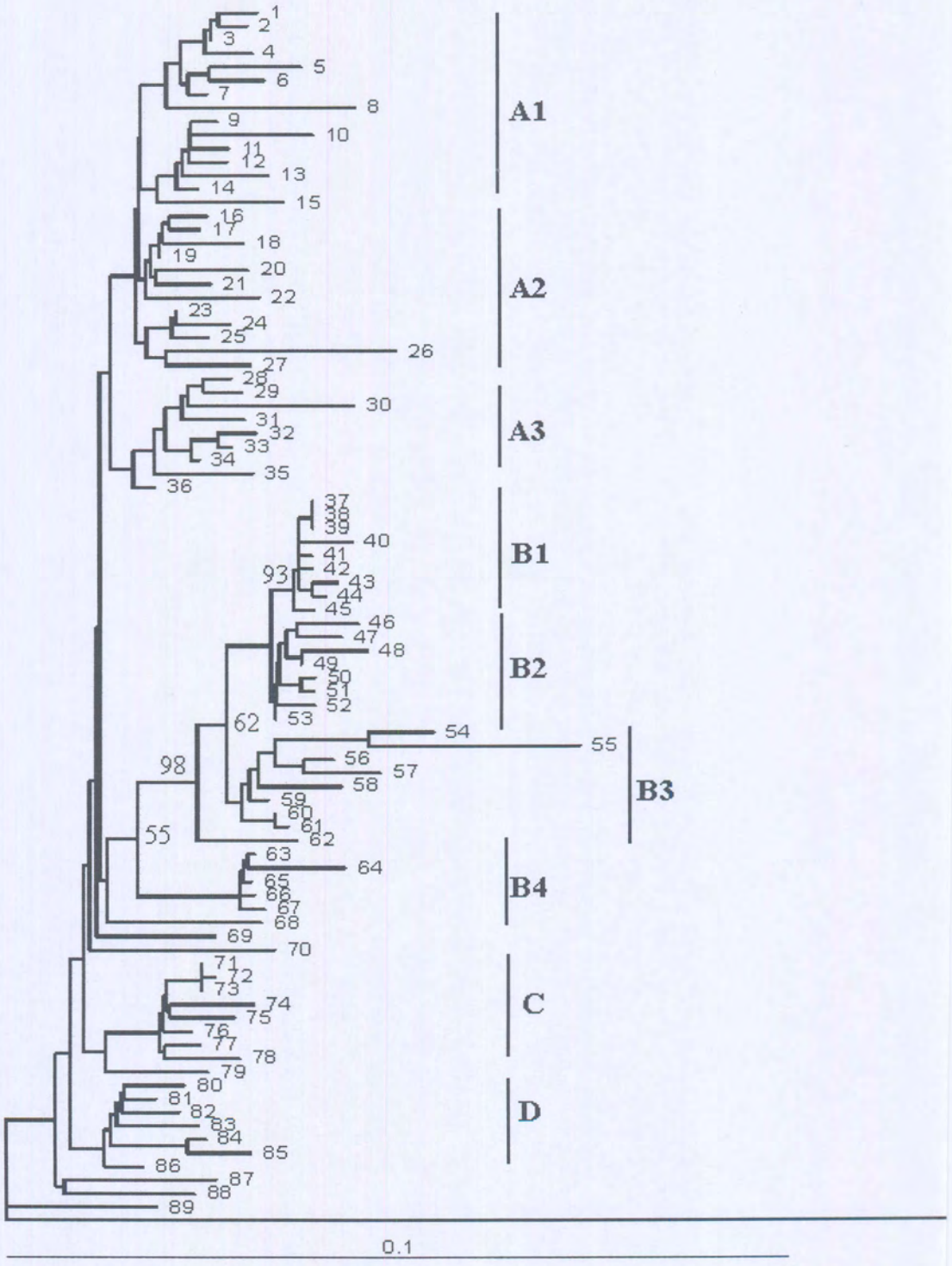

Figure 7: A neighbour-joining tree generated using the cytoplasmic domain of the glycoprotein and the G-L intergenic region sequences of 89 canid rabies virus isolates from Zimbabwe and South Africa based on Kimura 2-way corrected values. The sequence of the Pasteur Virus (PV) was included as the outgroup. Principal bootstrap values are shown at the nodes. The horizontal branch lengths are proportional to the similarity of sequences within and between clusters. Vertical lines are for clarity of presentation only. The scale bar represents 10 nucleotide substitutions per 100 positions.

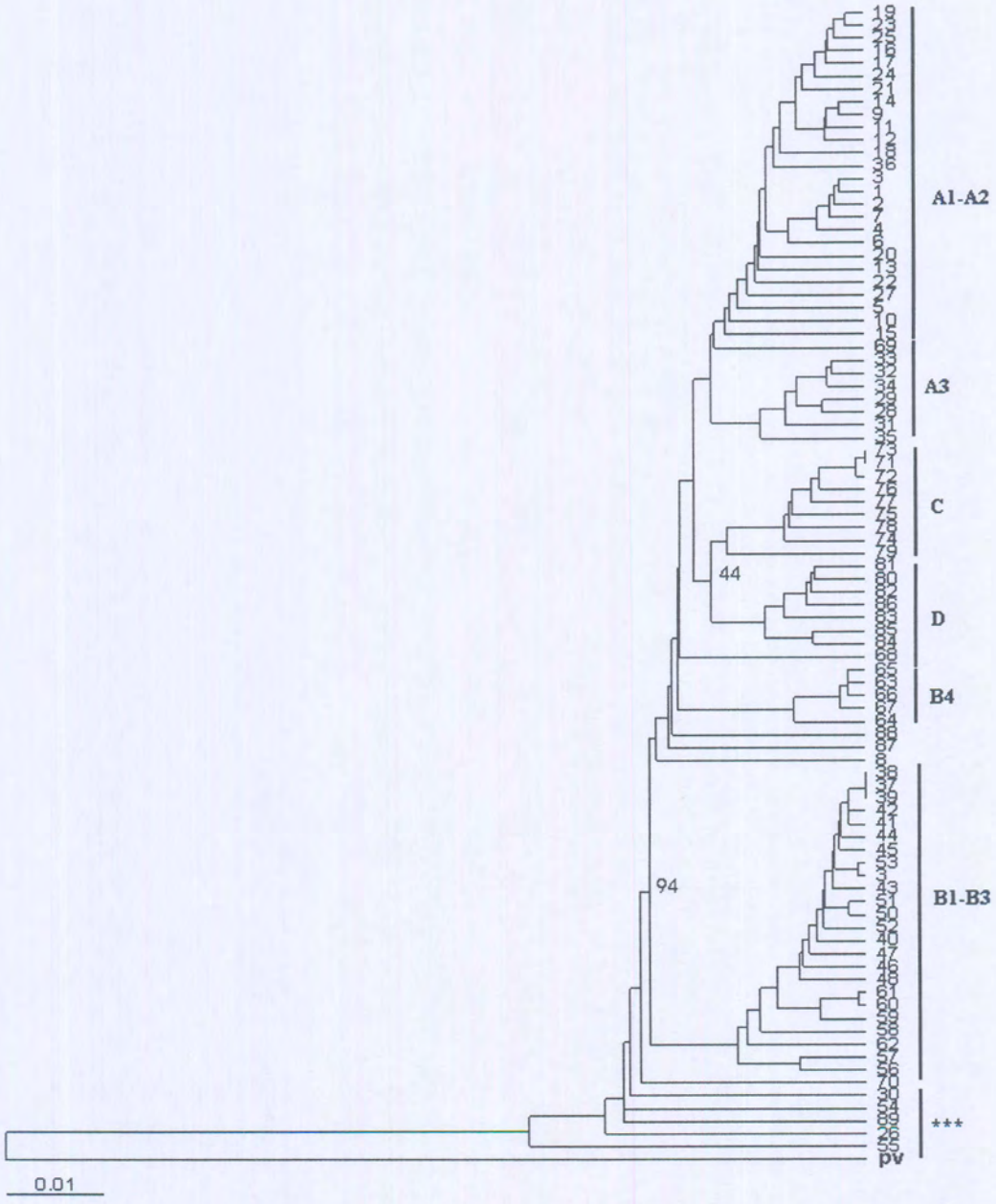

Figure 8: Phylogram generated using the cytoplasmic domain of the glycoprotein and the G-L intergenic intergenic sequences for rabies viruses from Zimbabwe and South Africa using the Kitsch output algorithm (Phylip version 3.5, Felsenstein 1993). Pasteur virus

was used the outgroup. The nodes were supported by weak bootstrap values and with the exception of virus isolates (***), a similar topology was obtained as in Figure 7.

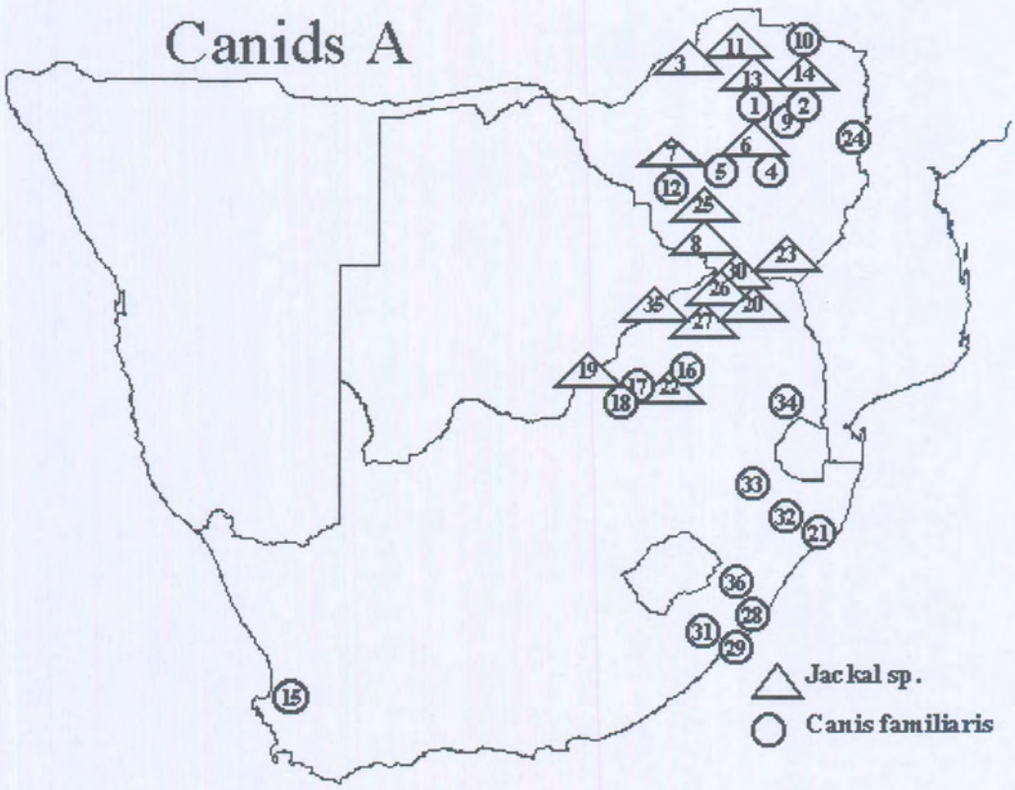

Figure 9a: Map of Zimbabwe and South Africa showing the geographical location of virus isolates contained in group A.

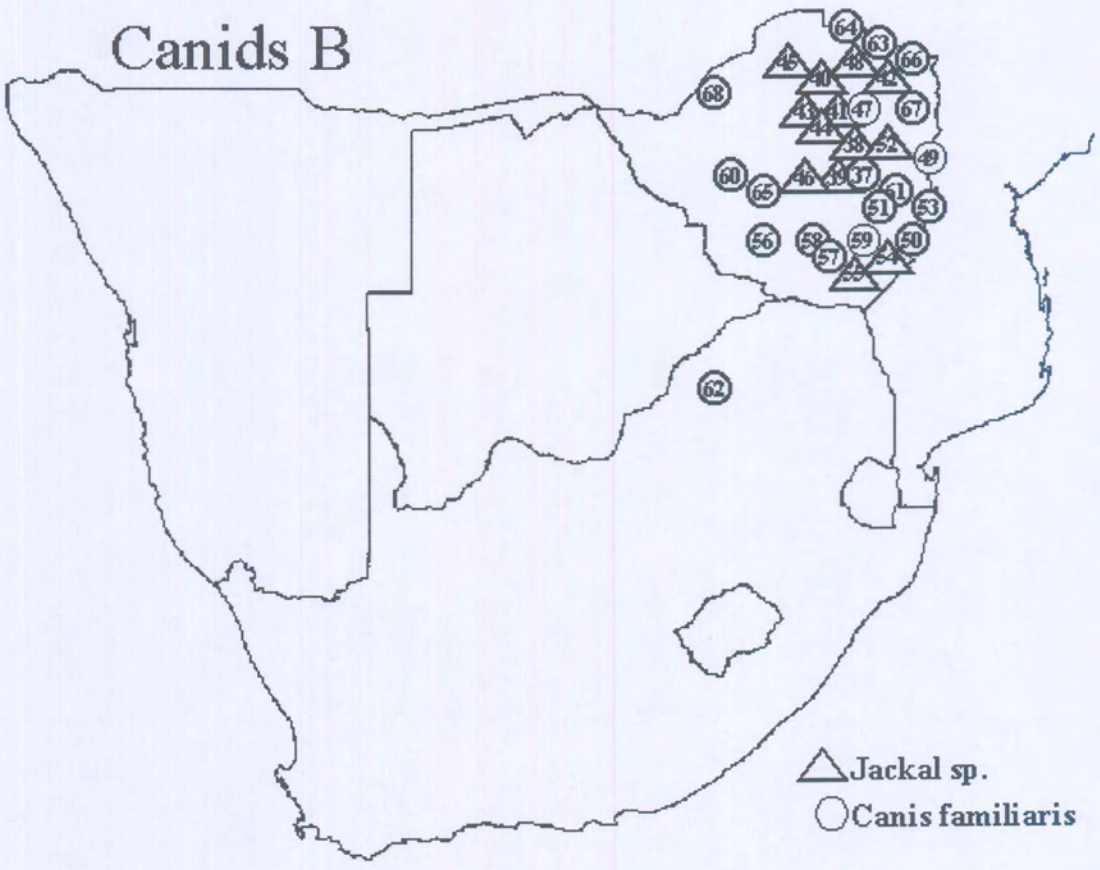

Figure 9b: Map of Zimbabwe and South Africa showing the geographical location of virus isolates in group B.

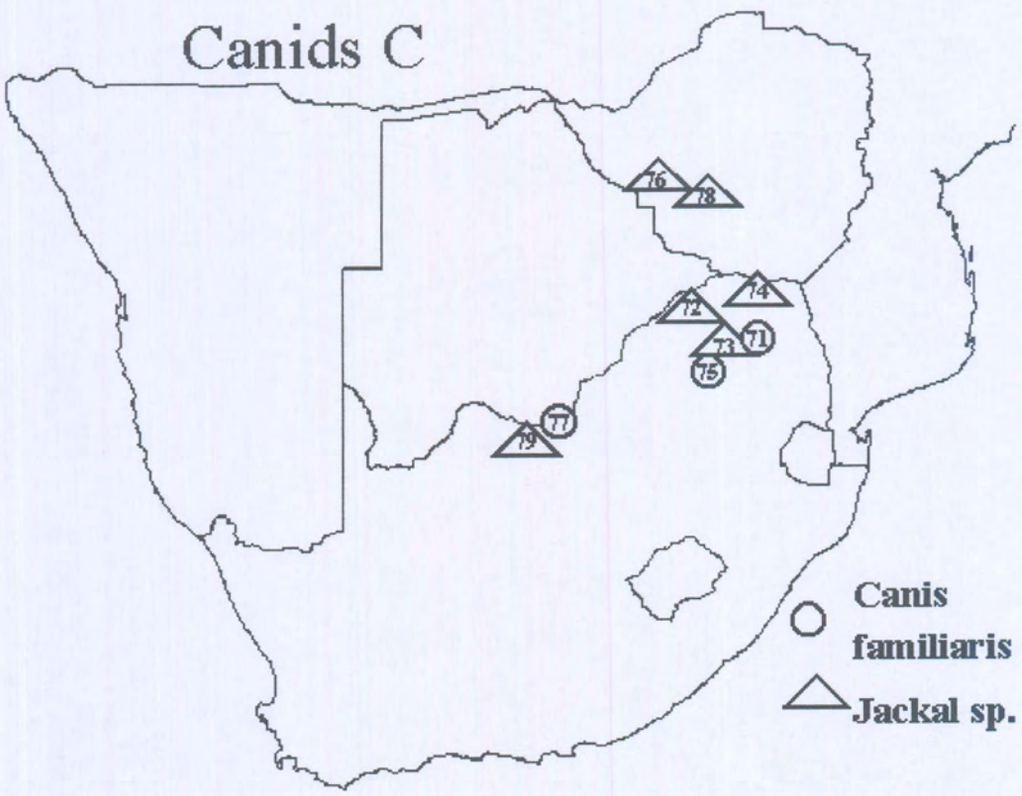

Figure 9c: Map of Zimbabwe and South Africa showing the geographical location of virus isolates in group C.

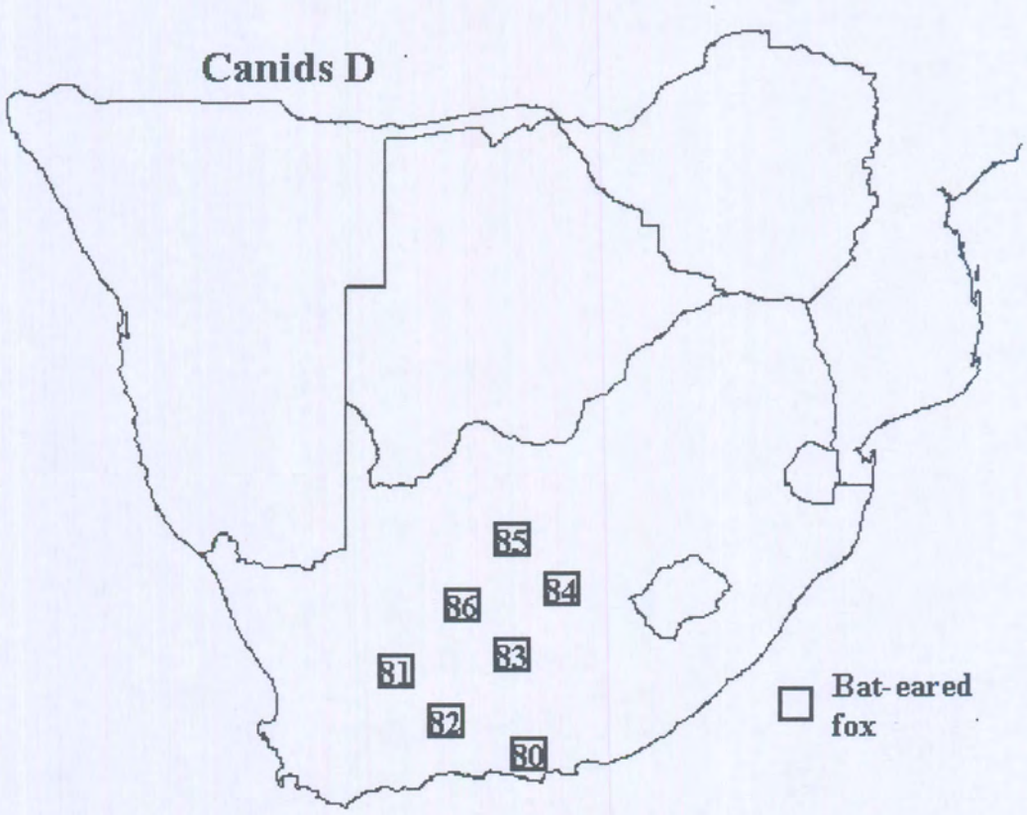

Figure 9d: Map of Zimbabwe and South Africa showing the geographical origin of viruses contained in group D.

2.5 Discussion

In this study, a collection of viruses from bat-eared foxes (9), jackal species (41) and dogs (39) in Zimbabwe and South Africa was assembled and genetic relatedness and phylogeny among the different isolates investigated.

All the viruses analysed in this investigation were found to be closely related suggesting a common and recent origin. The fact that there is minor genetic distinction amongst the southern African canid viruses is consistent with the relatively recent historical emergence of these cycles (Swanepoel *et al.*, 1993; Bingham *et al.*, 1999a; Bingham *et al.*, 1999b). Our data support findings of previous studies in which rabies virus isolates from both domestic dogs and jackals were shown to have no antigenic distinctions, implying that they are closely related (King et al., 1994; Bingham *et al.*, 1999b). As the southern African canid rabies virus lineage belongs to the cosmopolitan group it is likely that it descended from a relatively recent arrival on the African continent (Smith *et al.*, 1992; Nel *et al.*, 1993; Nel *et al.*, 1998).

Canid rabies in southern Africa is a recent phenomenon, having been introduced in the 1940s from Angola by dogs (Swanepoel *et al.*, 1993). It spread throughout Zimbabwe and northern South Africa in the dog population during the 1950s, and in the 1960s it spread into the KwaZulu-Natal region. The epidemiological evidence suggests that the disease established cycles in wild canid populations wherever wild populations capable of supporting rabies cycles occurred. In Zimbabwe rabies in jackals (*C. adustus*) appeared soon after the introduction of the disease in dogs, but jackal cycles were initially limited. Cycles in *C. mesomelas* also appeared after the introduction of dog rabies, in

Namibia during the late 1940s (von Maltitz, 1950), in South Africa in the 1950s (Swanepoel *et al.*, 1993) and in Zimbabwe (Bingham *et al.*, 1999d) and Botswana in the 1970s (Anon, 1973). *C. mesomelas* cycles have persisted ever since in these countries. Bat-eared fox rabies in South Africa was first reported in the 1960s and was probably derived from a *C. mesomelas* cycle as this was the only maintenance host cycle from which a bat-eared fox cycle could be established (Swanepoel *et al.*, 1993).

The jackal sub-group labelled as B1, with a bootstrap value of 93%, appears to be the major *Canis adustus* epidemic cycle that appeared in the early 1990s in the Mashonaland East area (Bingham *et al.*, 1999b). Group D is the bat-eared fox cycle of the Cape province (Figure 1, Thomson and Meredith, 1993). Two other bat-eared fox isolates (#'s 69 and 87) fell outside this group and neither was strongly associated with any specific group or cluster, although some relationship to an ungrouped jackal isolate was evident. Group (D) was close to, but separable from C, which was composed of viruses from dogs and jackals further to the north of South Africa and areas bordering southern Zimbabwe. Other trends were apparent. For example, isolates which clustered into Group C represent a cycle associated with *Canis mesomelas* in the area of northern South Africa and southern and western Zimbabwe. The group labelled B4 was composed of isolates from domestic dogs in northeastern Zimbabwe.

Dogs and the wild canids support rabies cycles independently of each other, but cross-introductions of infection occur between species cycles. Groups B1, C and D are examples of clusters belonging to wild canid cycles that are maintained independently of dogs. Similar phylogenetic patterns may have emerged with other local jackal epidemics

if a larger number of appropriate samples had been analysed. Our data show that the phylogenetic groupings have loose associations with geographical locality and host species. This would indicate that the canid rabies virus lineage is opportunistic within whatever canid host population is available and ecologically capable of sustaining prolonged cycles. As canid hosts tend to be, to greater or lesser degree, ecologically exclusive (in other words, any single region will have one dominant canid species) this will give the effect of phylogenetic clusters that appear to have species or geographical associations. The data indicates that although species associations with phylogenetic clustering are apparent, such associations are loose, with a species occurring in more than one group or several species present within a group. This supports the findings of a previous study (Bingham *et al.*, 1999a; 1999b), which indicates that although canid species support separate rabies cycles, the virus may jump across species barriers relatively easily. In the case of dog-to-*Canis adustus* cycle transfers such species jumps probably occurred a total of 8 times between 1950 and 1996 (Bingham *et al.*, 1999b). The domestic dog thus remains a major reservoir of rabies in southern Africa and on the African continent, unlike in North America where it has been eliminated as a reservoir for rabies (see Smith, 1996 for review). In North America, rabies cases in dogs declined in the past 50 years or so through vaccination programs, although cases in wild carnivore species underwent a dramatic increase (Krebs *et al.*, 1995). A situation in the US whereby a large number of rabies cases were reported in one major host species in certain areas, resulting in the compartmentalisation of the disease in a single host species namely the raccoons, skunk, arctic fox and red fox (Smith and Baer, 1988). In these areas, the

involvement of single-species is universal, but there were accidental infection (known as spillover) in other animals.

The alternative hypothesis, as suggested by other researchers to account for rabies prevalence in jackals (Cumming, 1982; Cleaveland and Dye, 1995; Rhodes *et al.*, 1998), is that all wild host infections are spillover from dog rabies, with wild hosts having no independent involvement in rabies cycles. The phylogenetic pattern does not support this, as one would not expect the formation of phylogenetic groups, such as B1, C and D that are composed almost exclusively of isolates from wild canid hosts. In addition, our finding discredit the hypothesis that viruses responsible for jackal rabies belong to a separate lineage that is indigenous to southern Africa (Smith *et al.*, 1993).

This investigation describes the first genetic analysis of rabies virus isolates from Zimbabwe, and further clarifies the molecular epidemiology of canine rabies in South Africa. The close relationship of South African viruses is of particular interest where rabies cycles in dogs from KwaZulu/Natal and those from jackals in the *C. mesomelas* zone (northern South Africa), is concerned. The ongoing epidemic of dog rabies in KwaZulu/Natal, might have suggested an independent rabies cycle but our results indicate that the viruses involved in these cycles are epidemiologically linked. Furthermore, the close genetic relatedness of virus isolates from jackals from southern Zimbabwe and those from northern bordering areas of South Africa (Figures 9a and 9c), illustrates the influence of geographical location on rabies virus evolution and may resemble the dog-coyote-dog cycle that occurred at the Mexico-Texas border in 1988

(Clark *et al.*, 1994). The rabies outbreak in coyotes in south Texas is thought to have resulted from a spillover of infection from domestic dogs.

In conclusion, this study confirms the historical data, which suggests that the canid rabies variants of southern Africa have a recent common progenitor and that this virus was introduced with dogs and then spread into wild canid populations. This lineage is now well-established throughout the sub-continent, testifying to the highly efficient vector capabilities of the various canid hosts. The widespread nature and high prevalence of this virus makes it the most threatening variant for humans and domestic animals. For the same reasons it may also be the most likely to establish future cycles in new host species. This study demonstrates the value of multinational surveillance investigations in understanding the epidemiology of rabies in southern Africa.

CHAPTER 3

MOLECULAR EPIDEMIOLOGY OF RABIES ASSOCIATED WITH VIVERRID SPECIES IN ZIMBABWE AND SOUTH AFRICA.

3.1 Summary

Little is known about the genetic variation of viverrid rabies viruses from Zimbabwe. We thus set out to analyse rabies virus isolates obtained from non-canine species (*viz* the slender mongoose, African wild civet and honey badger) in Zimbabwe. Through the years, many parts of the rabies genome, notably the nucleoprotein, glycoprotein, matrix and phosphoprotein genes have been used to infer epidemiological relationships of rabies viruses (Sacramento *et al.*, 1992; Von Teichman *et al.*, 1995; Nadin-Davis *et al.*, 1997; 1999; this study). As described in Chapter 2, the G-L intergenic region is highly variable and an appropriate sensitive indicator of short-term genetic evolution in this group of viruses.

From this study, it was shown that viverrid viruses form a heterogeneous population of diverse host species. Furthermore, viverrid viruses from Zimbabwe were shown to be different from those arising from host species in South Africa, an indication of a pattern of evolution different from that observed in the canid viruses (Chapter 2). Whereas canid viruses from host species in Zimbabwe and South Africa are closely related, the viverrid viruses were markedly less so.

Data from this chapter is being prepared for publication.

3.2 Introduction

Phylogenetic analyses of nucleic acid sequences of the cytoplasmic domain of the glycoprotein and the intergenic pseudogene region of rabies viruses have confirmed the occurrence of two very distinct virus groups that dominate enzootic rabies in South Africa (Von Teichman *et al.*, 1995). The two rabies lineages were separated by a large genetic divergence possibly indicating diverse origins and separate evolutionary pathways of canid and viverrid viruses. It has also been shown that viverrid viruses from South Africa are highly divergent, based on nucleic acid sequences of their nucleoprotein and glycoprotein genes (Nel *et al.*, 1993; Von Teichman *et al.*, 1995).

The large genetic diversity observed amongst the viverrid rabies viruses was initially attributed to have arisen through discontinuous mongoose populations, brought about via physical barriers such as mountains and rivers to animal movement and/or habitat isolation (Smith *et al.*, 1993). The argument put forward by Smith was later supported by genetic analysis of mongoose haplotypes (Van Vuuren and Robinson, 1997). In these genetic studies by Van Vuuren and Robinson (1997), there was little evidence for gene flow between mongooses, implying that barriers are responsible for the separation of the mongoose populations. Epidemiology of viverrid rabies thus probably reflects specific socialising habits of viverrid species (Taylor, 1993).

Nucleic acid sequence divergence amongst the viverrid rabies viruses of *ca* 20% (Von Teichman *et al.*, 1995, this study) lends support to results of antigenic studies (Smith *et al.*, 1993; King *et al.*, 1994; Tremlett *et al.*, 1994; Foggin, 1988) that indicated two epidemiologically separate lineages of rabies virus in southern Africa. Both antigenic and genetic studies have shown that rabies virus strains circulating in particular

host species within defined geographical zones, tend to undergo genetic adaptation, and evolve into distinct biotypes that differ in antigenicity and pathogenicity (Tuffereau *et al.*, 1989; Wunner, 1991; Smith *et al.*, 1995). The virus biotypes are generally adapted in such a way that they are maintained in specific reservoir hosts that are responsible for the maintenance and spread of the disease. It has been shown that spillover can occur, although this spillover is thought to be infrequent (Chaparro and Esterhuysen, 1997; Nel *et al.*, 1997). It is generally accepted that the biotype involving viverrid host species was established much earlier than the canid biotype (Nel *et al.*, 1993; Von Teichman *et al.*, 1995).

The ecology of the slender mongoose (*Galerella sanguinea*) has been reviewed by Smithers (1983). This brown viverrid occurs over the whole of Zimbabwe and much of Africa. Mongoose rabies has also been reported from many other parts of the world (Foggin, 1988) but it is only in South Africa and parts of the Caribbean where it appears to be the main manifestation of enzootic disease (Zumpt and Hassel, 1982). The first case of mongoose rabies in Zimbabwe was reported in 1966, with subsequent ones occurring annually from 1969 to 1981 (Foggin, 1988). Mab studies on viruses from slender mongooses gave a reaction pattern different from the typical reaction found for canid viruses and from this observation, Foggin concluded that rabies epidemics in slender mongooses probably represented a separate cycle from that in domestic dogs and jackal species (Foggin, 1988). The results from Foggin's antigenic studies were subsequently confirmed (Smith *et al.*, 1993; King *et al.*, 1994). In Zimbabwe, canine rabies is mainly associated with human rabies cases and viverrid rabies with wild animal rabies cases (Foggin, 1988; Bingham, 1999d).

In South Africa, viverrid rabies has been present for at least the past 150-250 years (Snyman, 1940) and was established long before the introduction of canine rabies to the region (King *et al.*, 1994). Viverrid rabies is primarily confined to the central highveld plateau of South Africa (Figure 4) and contributes significantly to rabies statistics. For example, viverrid rabies contributes approximately 44% of rabies statistics in South Africa with 37% attributable to *Cynictis penicillata* (Barnard, 1979). In the Caribbean islands of Grenada, viverrid rabies contributed approximately 57% of all rabies cases, whereas it contributed only 2% in Zimbabwe (Foggin, 1988). The relatively low figure for viverrid rabies in Zimbabwe and other parts of Africa may indicate the lack of research efforts being directed towards surveillance for this rabies variant.

The epidemiology of viverrid rabies in Africa is not clear. Apart from the genetic studies to distinguish canid and viverrid viruses from South Africa (Von Teichman *et al.*, 1995; Jaftha, 1997), not much is known about the epidemiology of viverrid viruses from other parts of Africa. We therefore included viverrid viruses from host species from Zimbabwe in order to facilitate and improve our knowledge of the epidemiological cycles involving viverrid viruses in the sub-continent.

In this investigation of viverrid viruses from Zimbabwe and South Africa, we used nucleic acid sequence analysis of the G-L intergenic region of genomes of these isolates in order to gain a better understanding of the molecular epidemiology of viverrid viruses in southern Africa. Phylogenetic analysis of nucleic acid sequences of the viverrid rabies showed viruses from Zimbabwe to form a distinct group from those composed of viruses isolated from host species in South Africa. These groups were a result of geographical and not species compartmentalisation.

3.3 MATERIALS AND METHODS

3.3.1 Rabies viruses and virus isolation

Virus isolates from non-canine wild reservoir species from different areas of Zimbabwe and South Africa were included for analysis in this study. The virus isolates included all the viruses with recorded history from host species in Zimbabwe and were isolated from the slender mongoose (*Galerella sanguinea*), African wild civet (*Civettictis civetta*) and the honey badger (*Mellivora capensis*). Viverrid viruses were isolated from the following host species in South Africa; the African wild cat (*Felis lybica*), unidentified feline species, Suricates (*Suricata suricatta*), small-spotted genet (*Genetta genetta*), bovine (*Bos taurus*), ovine (*Ovis ories*), canine (*Canis familiaris*), ground squirrel (*Xerus inauris*), water-mongoose (*Atilax paludinosis*) and bat-eared foxes (*Otocyon megalotis*). All the rabies samples were submitted to the Rabies Unit of the Veterinary Research Institute, Onderstepoort (Pretoria) and diagnosed as rabies positive by FAT as described in Methods of Chapter 2.

The rabies positive isolates from Zimbabwe were differentiated using a panel of monoclonal antibodies to establish their biotype (OVI report, 2000). Fifty-eight viruses from non-canine wild carnivore species were typed with Mabs provided by Dr Alex Wandeler (Bingham *et al.*, 2001). The viverid virus isolates were mainly isolated from wild carnivores and domestic cats that had been made by Chris Foggin and John Bingham from the 1980s to 1997 (Bingham, 1999c). The viruses were isolated as described in Chapter 2 of this thesis. The host species of origin, year of origin and geographical origin of the virus isolates are shown in Table 4.

Table 4: Virus isolates included in the study

Virus No.	Lab. ref #.	Species of Origin	Locality	Country of Origin	Yr isol.	Map reference long-lat	Accession no.
1	197/97	Herpestidae	Richmond	South Africa	1997	23°57'-31°25'	AF304166
2	516/97	Feline sp.	Clesburg	South Africa	1997	25°03'-30°43'	AF304167
3	167/97	*Galerella sanguinea*	De AAR	South Africa	1997	24°08'-30°50'	AF304168
4	221/98	*Surtcata suricatta*	Venterstad	South Africa	1998	25°29'-30°43'	AF304170
5	718/98	*Genetta genetta*	Gordonia	South Africa	1998	20°30'-27°58'	AF304169
6	401/98	Feline sp.	Kimberley	South Africa	1998	24°39'-28°50'	AF304171
7	328/98	*Otocyon megalotis*	Hay	South Africa	1998	22°55'-29°15'	AF304172
8	257/98	*Cynictis penicillata*	Victoria West	South Africa	1998	22°53'-30°35'	AF304173
9	211/98	Canis familiaris	Kuruman	South Africa	1998	22°33'-26°28'	AF304165
10	138/98	*Felis carac*	Kimberley	South Africa	1998	24°39'-28°51'	AF304174
11	89/98	*Otocyon megalotis*	Douglas	South Africa	1998	23°42'-28°38'	AF304175
12	85/98	Feline sp	Kuruman	South Africa	1998	23°26'-27°27'	AF304176
13	650/96	Feline sp	Postmasburg	South Africa	1996	22°40'-28°18'	AF304177
14	476/96	*Suricata suricatta*	Hopetown	South Africa	1996	24°02'-29°37'	AF304178
15	410/96	*Felis lybica*	Noupoort	South Africa	1996	24°48'-31°04'	AF304179
16	28/96	*Cynictis penicillata*	Kimberley	South Africa	1996	29°40'-29°03'	AF304180
17	1020/94	ovine	Ermelo	South Africa	1994	29°50'-26°43'	AF304181
18	23/98	Canis familiaris	Vryburg	South Africa	1998	24°01'-25°53'	AF304182
19	22574	*Civettictis civetta*	Matsine, Wedza	Zimbabwe	1994	31°43'-18°40'	AF304183
20	19571	*Mellivora capensis*	Bulawayo	Zimbabwe	1991	29°52'-19°26'	AF304184
21	22107	*Galerella sanguinea*	Rusape	Zimbabwe	1994	32°08'-18°32'	AF304185
22	21179	*Civettictis civetta*	Penhalonga	Zimbabwe	1992	32°43'-18°43'	AF304186
23	19518	*Galerella sanguinea*	Fort Rixon	Zimbabwe	1991	29°08'-19°51'	AF304187
24	19671	*Civettictis civetta*	Rusape	Zimbabwe	1991	32°10'-18°37'	AF304188
25	20948	*Mellivora capensis*	Bulawayo	Zimbabwe	1992	28°40'-19°59'	AF304189
26	30/00	Bovine	Brandfort	South Africa	2000	26°30'-28°42'	AF304190
27	669/90	*Cynictis penicillata*	Ermelo	South Africa	1990	29°52'-26°42'	AF079907
28	1236/80	*Genetta genetta*	Vryburg	South Africa	1980	22°45'-26°09'	AF079916
29	475/95	*Felis lybica*	Hay	South Africa	1995	22°55'-29°14'	AF079910

Table 4 (continued)

30	446/92	*Genetta genetta*	Postmasburg	South Africa	1992	22°30'-28°12'	AF079909
31	380/90	Felis sp.	Kuruman	South Africa	1990	21°41'-26°53'	AF079914
32	611/92	Genet	Postmasburg	South Africa	1992	22°53'-27°43'	AF079912
33	113/91	*Atilax paludinosus*	Beaufort West	South Africa	1991	23°02'-32°13'	AF079911
34	127/91	*Otocyon megalotis*	Malmesbury	South Africa	1991	18°25'-33°25'	AF079915
35	926/93	*Suricata suricatta*	Carolina	South Africa	1993	30°16'-26°04'	AF079908
36	19/92	*Suricata suricatta*	Cradock	South Africa	1992	25°37'-32°10'	AF070031
37	636/90	*Galerella pulverulenta*	Cradock	South Africa	1990	25°37'-32°10'	AF079930
38	637/90	*Xerus inauris*	Malmesbury	South Africa	1990	25°07'-31°30'	AF079933
39	668/92	*Atilax paludinosus*	Albert	South Africa	1992	26°07'-30°36'	AF079934
40	421/92	*Canis familiaris*	Albany	South Africa	1992	26°30'-33°20'	AF079929
41	701/92	*Canis mesomelas*	Albany	South Africa	1992	26°10'-33°13'	AF079932
42	298/90	*Felis lybica*	Carnavon	South Africa	1990	21°46'-30°13'	AF079928
43	5/91	*Canis mesomelas*	Rouxville	South Africa	1991	26°28'-30°34'	AF079927
44	710/90	*Cynictis penicillata*	Fauresmith	South Africa	1990	25°15'-29°54'	AF079925
45	732/90	*Cynictis penicillata*	Bultfontein	South Africa	1990	26°08'-28°18'	AF079926
46	1049/92	*Genetta*	Postmasburg	South Africa	1992	22°46'-27°55'	AF079913
47	480/90	*Cynictis penicillata*	Bloemhof	South Africa	1990	25°32'-27°25'	AF079923
48	466/90	*Cynictis penicillata*	Bothaville	South Africa	1990	26°37'-27°23'	AF079922
49	420/90	*Cynictis penicillata*	Wolmaranstad	South Africa	1990	26°14'-27°13'	AF079921
50	970/93	*Suricata suricatta*	Venterdorp	South Africa	1993	26°32'-26°23'	AF079924
51	256/90	*Otocyon megalotis*	Hay	South Africa	1990	22°55'-29°14'	AF079920
52	158/91	*Canis mesomelas*	Vetersdorp	South Africa	1991	27°10'-28°10'	AF079919
53	878/92	*Atilax paludinosus*	Harrismith	South Africa	1992	28°59'-28°08'	AF079918
54	485/94	*Suricata suricatta*	Standerton	South Africa	1994	29°14'-26°56'	AF079917

3.3.2 Primer selection

Oligonucleotide primers flanking the cytoplasmic domain and the G-L intergenic region and other internal primers were synthesised by Life Technologies (U.K.) and used without further purification. The list and details of the primers are shown in Table 3 on page 46.

3.3.3 Total RNA Extraction, cDNA synthesis and RT-PCR

Total viral RNA was extracted from the 20% lyophilised brain material (or brain-infected material) using TRIReagent according to the manufacturer's instructions (Molecular Research Centre, Cincinatti, U.S.A). Details of RNA extraction, cDNA synthesis, amplification and gel purification were performed as described in Chapter 2, pages 47-48.

3.3.4 Nucleic acid sequencing and phylogenetic analysis

Approximately 50 ng of purified PCR products were sequenced as described in Chapter 2, page 50 and the nucleic acid sequences obtained analysed in order to establish the phylogenetic relationships amongst these viruses as described in the methods section in Chapter 2, page 51. Each nucleic acid template was sequenced using both the G(+) and L (-) primers (for PCR products generated with the G-L primer-pair) and the G(+) and P5(-) and L(-) and P1(+) (for nested PCR products).

3.3.5 Nucleotide sequence accession numbers

All the nucleic acid sequences generated during the course of this investigation have been submitted to Genbank and assigned accession numbers (Table 4).

3.4 Results

3.4.1 Virus isolates, RNA preparation, cDNA synthesis and amplification.

Viral RNA extractions and RT-PCR were performed as described previously in Chapter 2. The G-L primer pair yielded the expected amplicon of approximately 850 bp for each of the virus isolates investigated (not shown). In the case of a nested PCR, products of approximately 400 bp in size were generated (not shown). All the FAT rabies positive samples were amplified with the set of primers used in this study.

3.4.2 Nucleotide sequence determination

On average, approximately 740 bases of nucleic acid sequence of the G-L region under investigation were obtained. The nucleic acid sequences are shown in the form of a multiple alignment (see Appendix 3) generated by ClustalW (Higgins and Sharp, 1988). Pairwise comparison of the 592 bp nucleic acid sequences showed the virus isolates to be heterogeneous, with variations between 1.9% and 30.3% for the southern African viverrid viruses. A sequence divergence of up to 30.3% was observed for the South African viruses and up to 13.9% for the Zimbabwean viruses (pairwise matrix shown, Appendix 4). On average, viverrid viruses from southern Africa had a sequence dissimilarity of

33.1% with PV. In addition, all the viverrid viruses from southern Africa lacked one TTP (Appendix 3) as previously shown for the canid viruses (see results section, Chapter 2). This signal has been shown to be lacking in street viruses, but present in laboratory strains such as the Pasteur virus (Sacramento *et al.*, 1992).

3.5 Phylogenetic analyses

The phylogenetic clustering of the viverrid virus isolates described in this chapter, were delineated into 5 distinct genetic groups (1-5), which were supported by high bootstrap values (Figure 10). As described in the methods section, the neighbour-joining (NJ) method of Saitou and Nei (1987) was used to generate the phylogenetic tree shown in Figure 10. In order to improve the clarity and separation of the genetic groups for purposes of interpretation, the phylogenetic tree was illustrated as a radial tree (Figure 11). The topology of the NJ tree was similar to that obtained with the Kitsch output algorithm (Figure 12) (Felsenstein, 1993). In general, the topology of the phylogenetic tree tended to reflect the geographical origin of the virus isolates described here (Figures 13a-e) as virus isolates originating from the same geographical region tended to cluster together.

All the viruses in Group 1 (Figures 10, 11 and 12) were isolated from host species from Zimbabwe. The viruses were isolated from host species that included African wild civets (isolate #'s 19, 22 and 24) from eastern Zimbabwe, slender mongooses (#'s 21 and 23) from eastern and western Zimbabwe and honey badgers (#'s 20 and 25) from southwestern Zimbabwe (Figure 13a). The sequence divergence within this group ranged from 0.7-15.5%. Viverrid viruses from Zimbabwe were distinguishable (group 1) from

those from South Africa (groups 2, 3, 4 and 5), separated into distinct groups with a high statistical support of 83%.

Group 2 was made up of viruses isolated from genets (isolates #'s 5, 30, 32), feline sp. (#'s 12, 13, 29, 31), and one each from a dog (# 9), bat-eared fox (# 34) and a water-mongoose (# 33) (Figure 13b). Virus isolates from group 3 all originated from host species from the central highveld plateau of South Africa. These viruses were isolated from suricates (#'s 50 and 54), yellow mongoose (#'s 47, 48, 49), bat-eared foxes (7, 11, 51) and one each from a jackal (52) and a water mongoose (53) (Figure 13c).

Group 4 comprised viruses from feline species (#'s 2, 6, 10, 15, 42), yellow mongooses (#'s 8, 16, 44, 45), suricates (#'s 4, 14 and 36), slender mongooses (#'s 3, 37), Xerus (# 28), jackal sp. (#'s 41, 43), domestic dogs (#'s 18, 40), and one each from a bovine (# 26), water mongoose (# 39), ground squirrel (# 38), genet (# 46) and an unidentified mongoose species (# 1). Three viruses, isolated from an ovine (# 17), a yellow mongoose (# 27) and a suricate (# 35), from northeastern South Africa, were the sole representatives of cluster 5 (Figure 13e). The largest sequence divergence was displayed by viruses from groups 3 and 5 (max 24.7%), with those from group 4 with a maximum sequence divergence of 23.6%. All the genetic groups made up of viruses from South Africa that have been described here were found to be composed of isolates from more than one host species, with group 4 being made up of the most diverse host species.

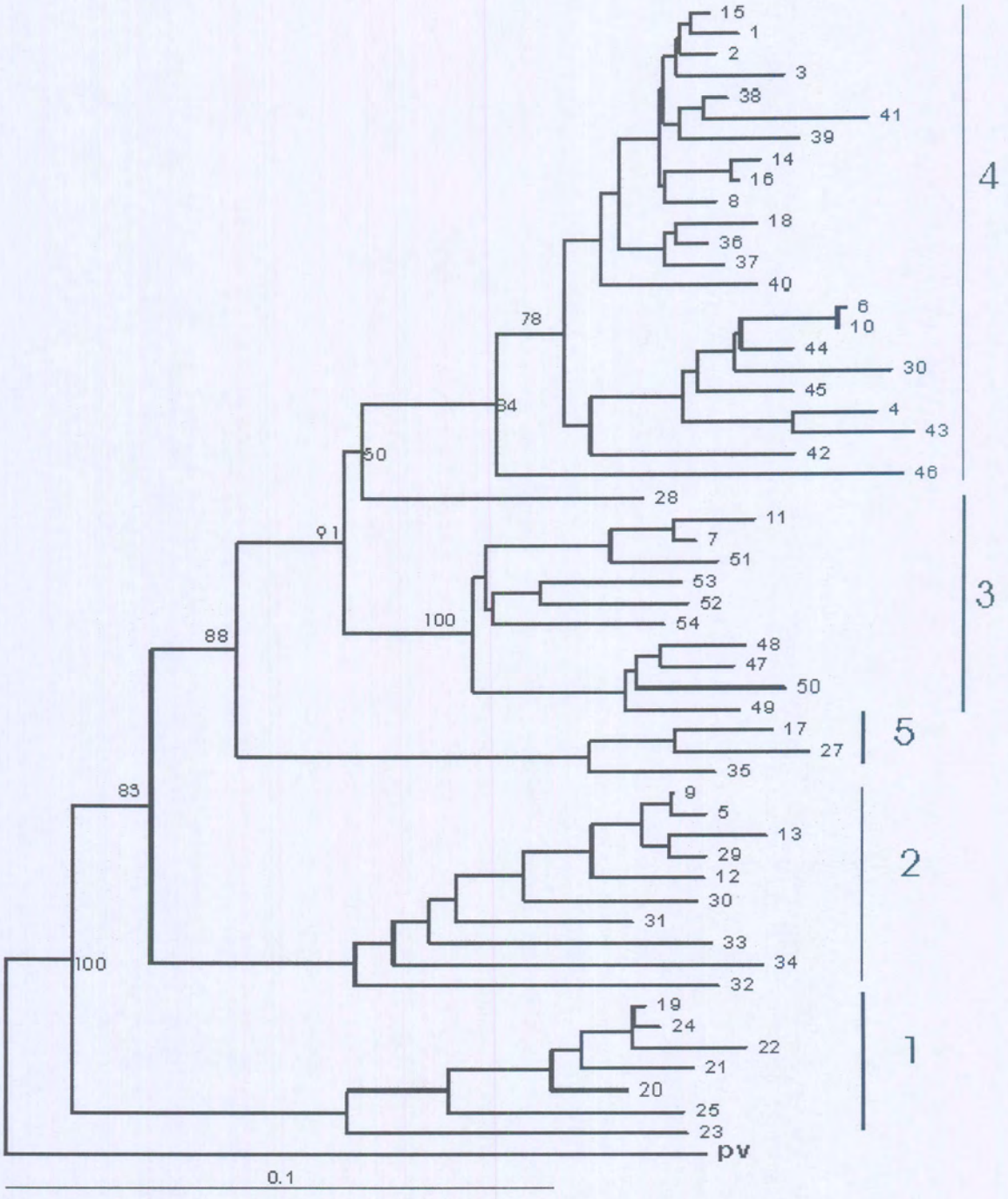

Figure 10: A phylogenetic tree showing the relationships amongst the viverrid viruses from southern Africa constructed with the neighbour-joining method of Saitou and Nei (1987) as described in the methods section. The tree was generated by comparing a 592 bp nucleotide sequence of the cytoplasmic domain and the G-L intergenic region of the glycoprotein gene of the viverrid rabies viruses from southern Africa. Bootstrap values are shown at the nodes. PV (Tordo *et al.*, 1986a) was included as an outgroup.

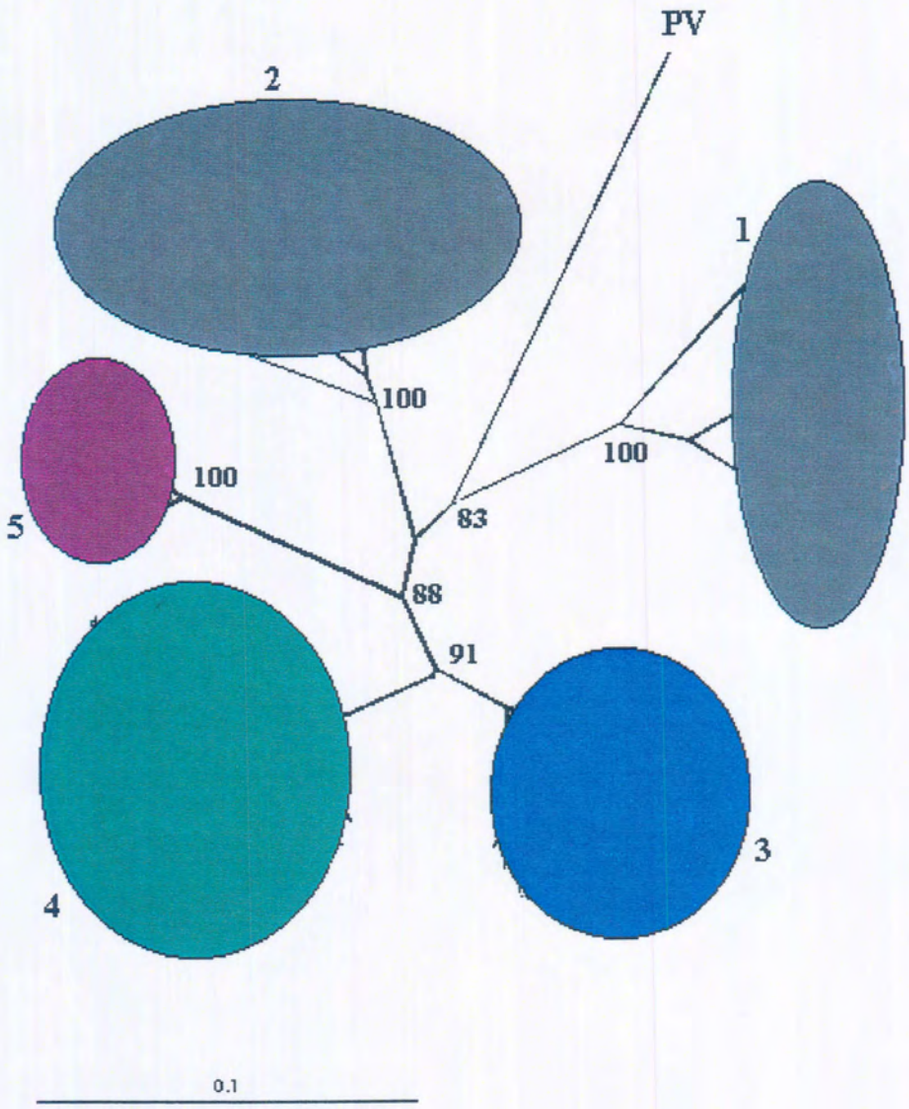
PV
2
1
100
100
100
5
83
88
91
4
3
0.1

Figure 11: A radial tree showing the phylogenetic relationships of viverrid viruses from Zimbabwe and South Africa generated using nucleic acid sequences of the cytoplasmic domain of the glycoprotein and the G-L intergenic region as described previously (Felsenstein, 1993). Nucleic acid sequences were aligned using CLUSTALW (Higgins and Sharp, 1988; Thompson *et al.*, 1994) and analysed by the Kitsch algorithm (Felsenstein, 1993). The sequence of the Pasteur virus (PV) was included as an outgroup.

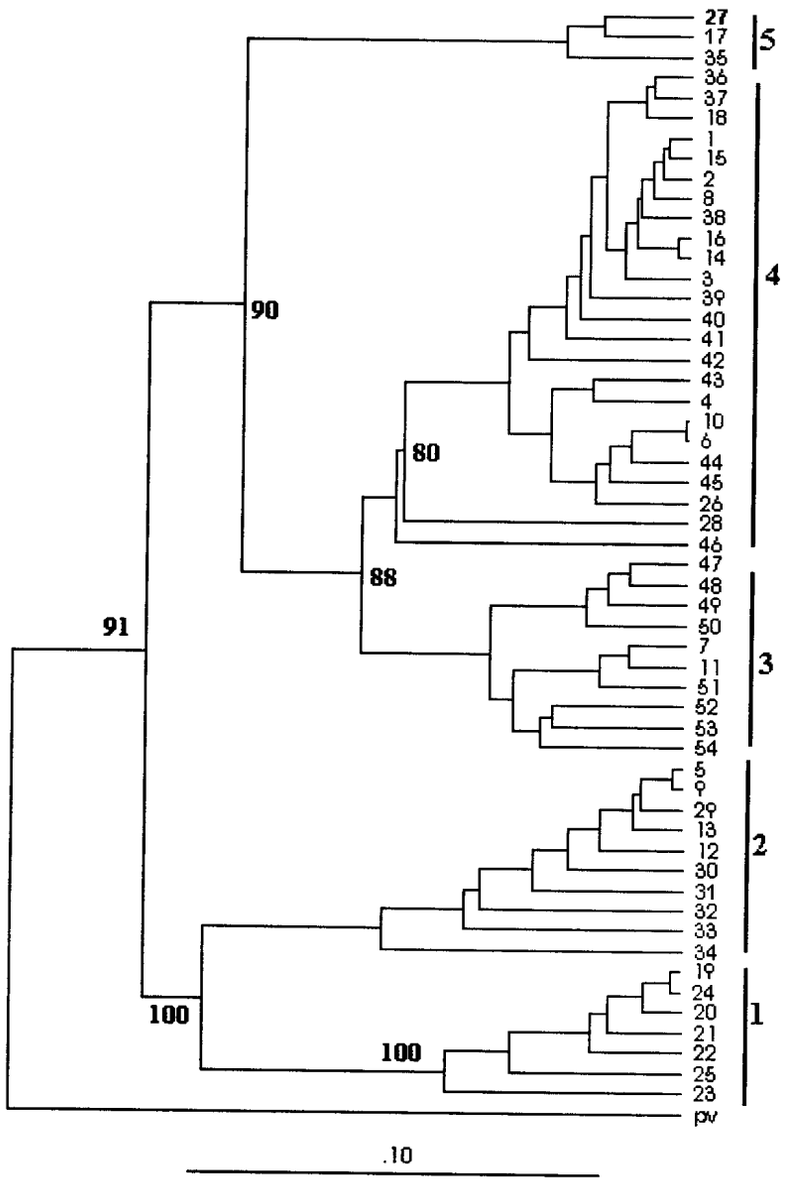

Figure 12: A phylogenetic tree showing the relationships of viverrid viruses from Zimbabwe and South Africa generated using nucleic acid sequences of the cytoplasmic domain of the glycoprotein and the G-L intergenic region with the Kitsch output algorithm. The Pasteur virus (PV) was included as an outgroup.

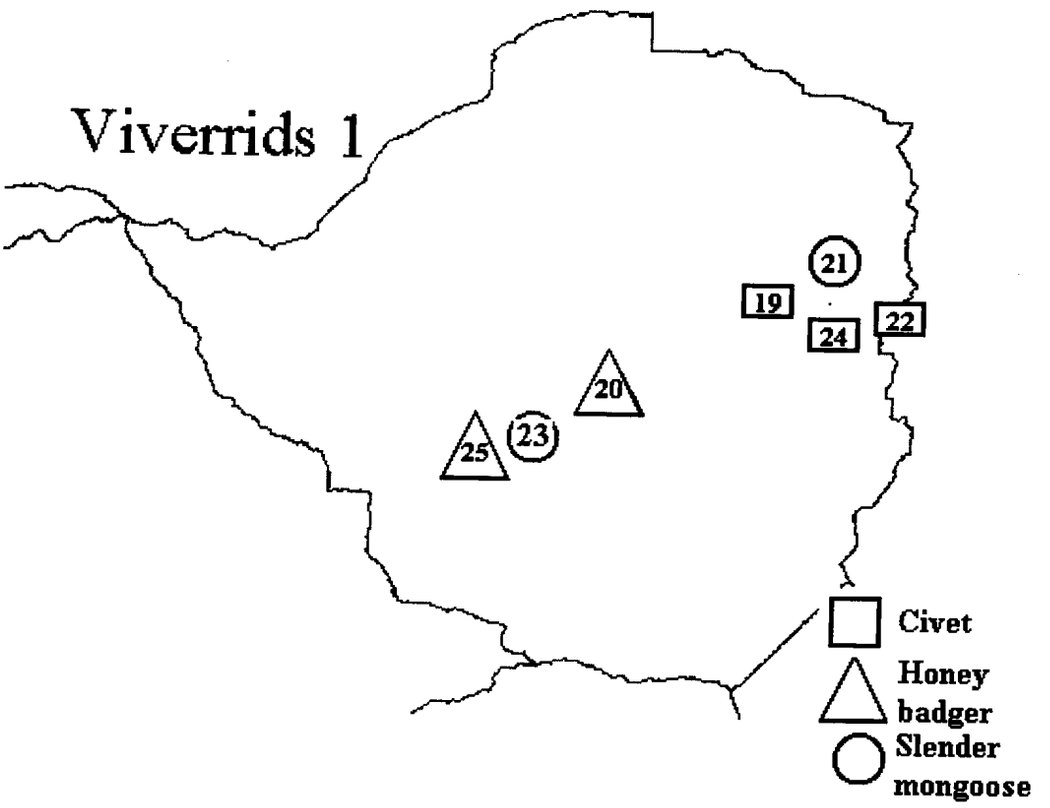

Figure 13a: Map of Zimbabwe showing the geographical location of the virus isolates in group 1.

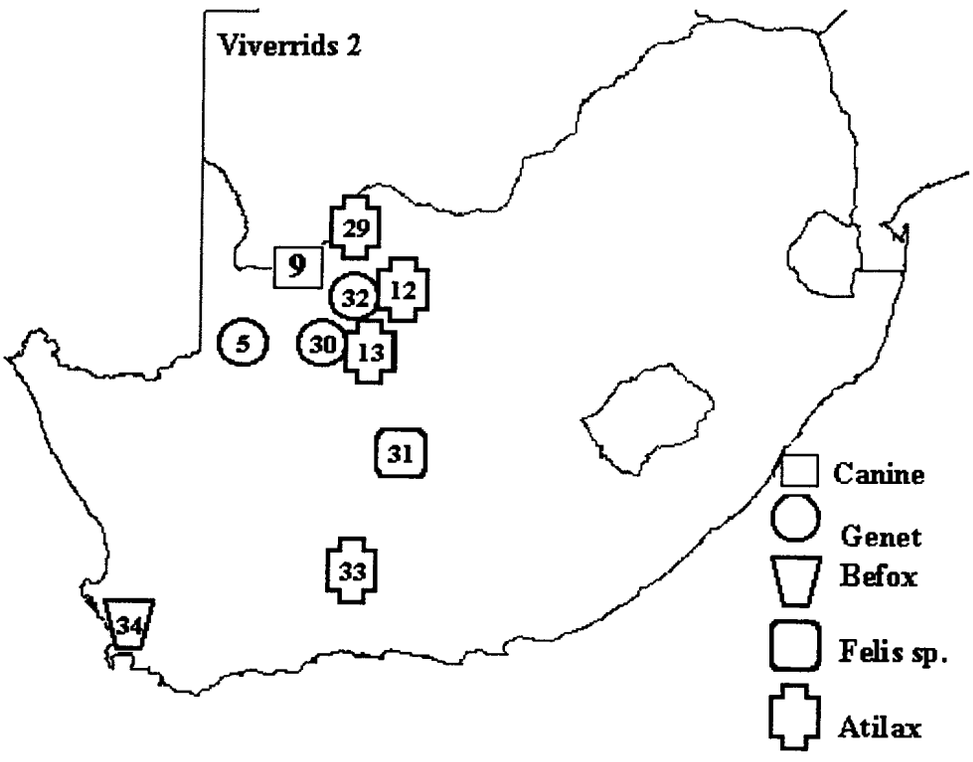

Figure 13b: A map of South Africa showing the geographical location of viruses in group 2.

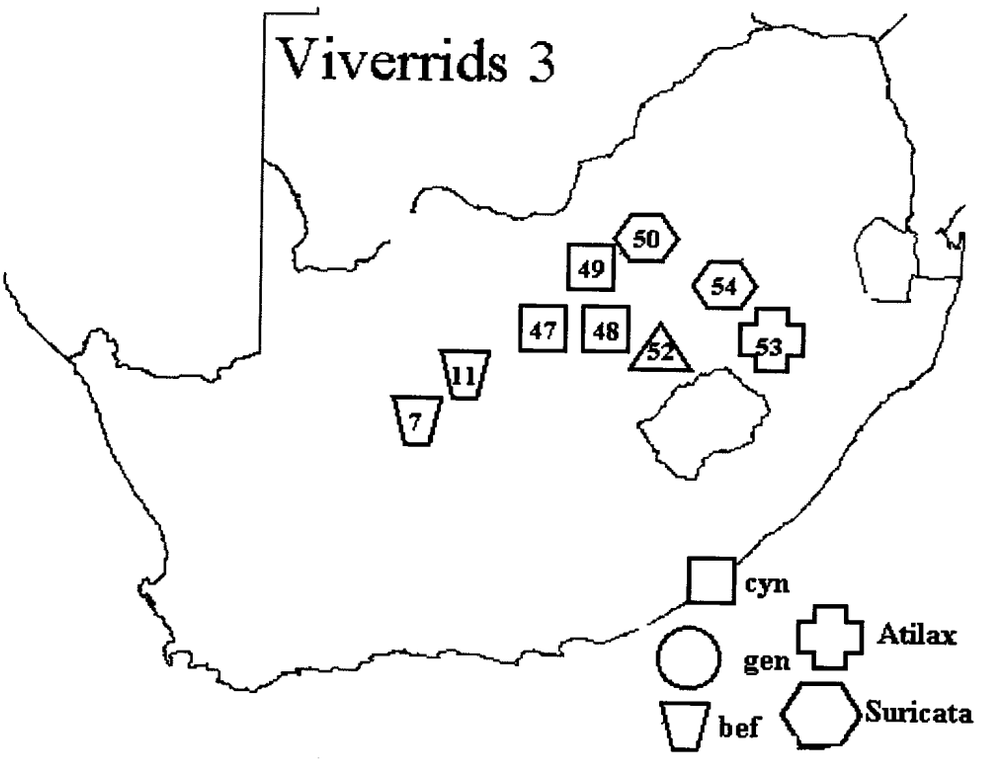

Figure 13c: A map of South Africa showing the geographical location of virus isolates in group 3. Key; cyn = *Cynictis penicillata*, gen = *Genetta genetta*, bef = bat-eared fox.

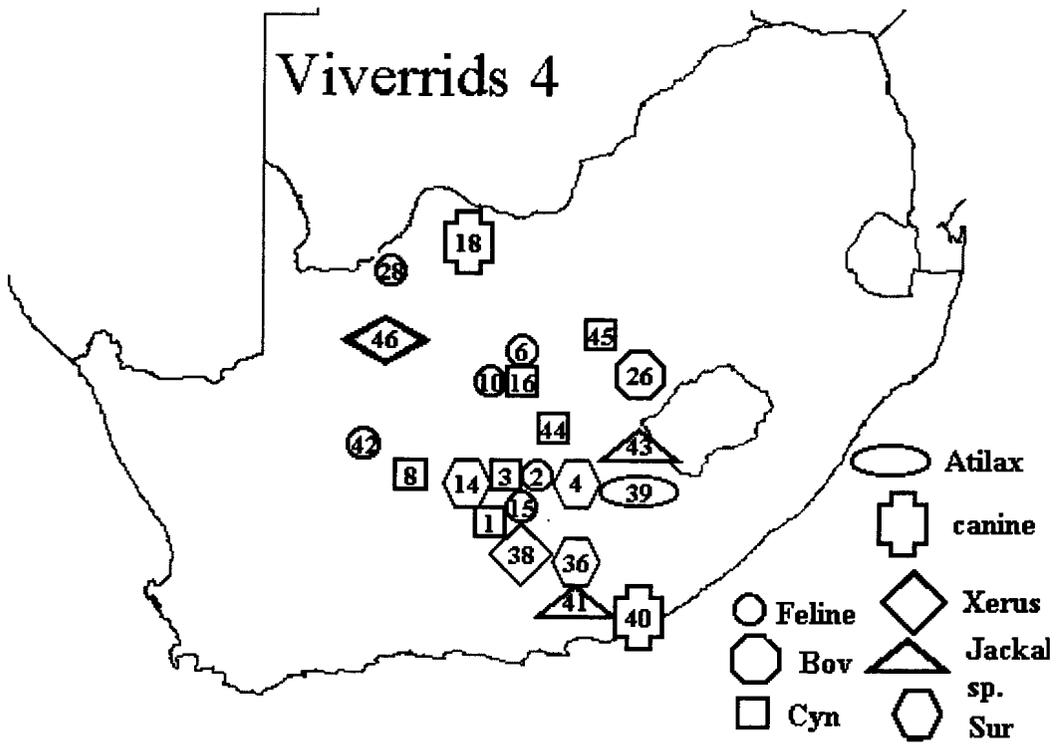

Figure 13d: A map of South Africa showing the geographical origin of the virus isolates in group 4. Key; bov=bovine, Cyn=*Cynictis penicillata*, sur=suricate.

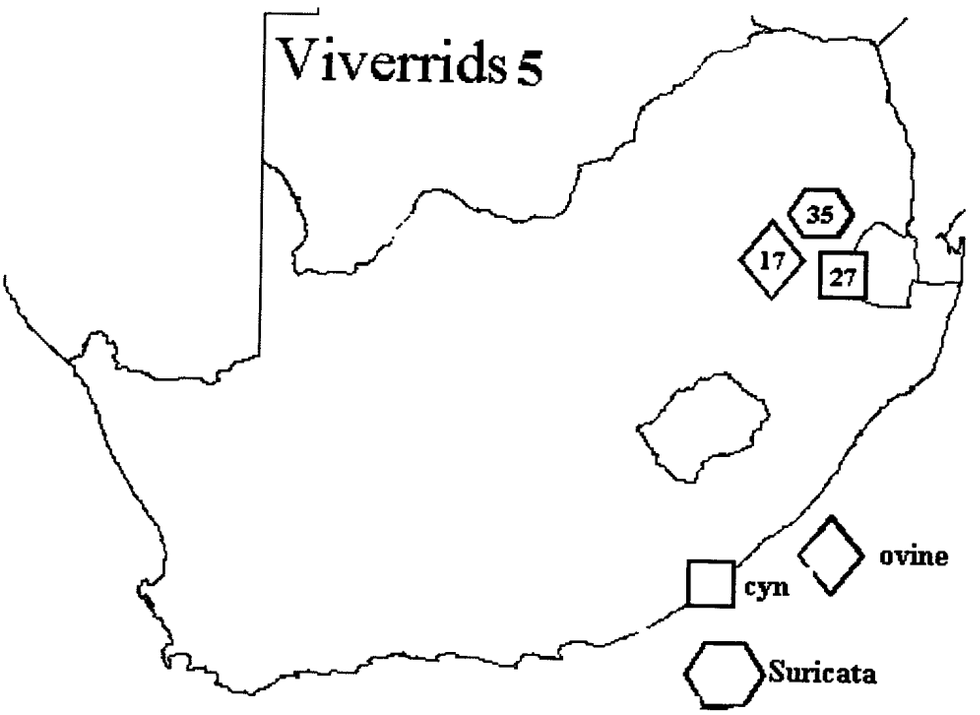

Figure 13e: A map of South Africa showing the geographical origin of the virus isolates in group 5. Key; bov=bovine, Cyn=*Cynictis penicillata*, sur=suricate.

3.6 Discussion

This investigation of phylogenetic relationships of viverrid rabies virus isolates in southern Africa (Zimbabwe and South Africa) has demonstrated that viverrid rabies virus isolates from Zimbabwe and South Africa are different. This study is the first comparative and genetic investigation of these lyssaviruses from Zimbabwe and South Africa. Previous studies have focused on the genetic distinction of canid and viverrid rabies viruses from South Africa (Von Teichman *et al.*, 1995; Jaftha, 1997). These studies included a limited number of virus isolates and the viverrid rabies viruses analysed were mainly from *Cynictis penicillata*. Recent studies extended analysis of viverrid rabies viruses from infected mongoose host species, suricates, felis sp., genets and canine species from South Africa (Jaftha, 1997). This study therefore included viruses from host species from South Africa other than the yellow mongoose and all those with recorded history from Zimbabwe.

The viverrid group of rabies virus isolates included in this study can be clearly distinguished from the rabies virus variant, represented by PV. Substantial divergence amongst members of this group is evident, an observation that supports results from antigenic studies conducted on southern African rabies viruses (King *et al.*, 1993; Wiktor *et al.*, 1984a; Foggin, 1988; Smith *et al.*, 1993). The genetic evidence obtained from this study further substantiates the postulation by Foggin regarding the existence of a rabies cycle in slender mongooses and Herpestidae species, different to the rabies cycle that commonly occurs in domestic dogs and jackal species (Foggin, 1988).

The phylogenetic analysis of nucleic acid sequences obtained here and sequence information contributed by other workers showed distinct virus groups. Four genetic

groups could be discerned from the South African viruses; these having been isolated from host species in western South Africa, eastern and then eastern to western South Africa and the central highveld plateau. The genetic groups described here are composed of viruses from more than one host, and their formation was found to be associated with geographical location of the virus rather than host-specificity factors. The amount of spillover from the viverrid species to canids and other non-viverrid hosts appears minimal. The high genetic divergence observed amongst members of the viverrid viruses is the opposite to results obtained in the case of canid viruses, which showed very little genetic variation amongst the viruses from both Zimbabwe and South Africa. The postulation by Swanepoel (1994) that rabies entered southern Africa from Angola in the 1950s and spread to Zimbabwe and South Africa simultaneously, could explain the close relatedness of canid viruses from Zimbabwe and South Africa.

The large sequence divergence amongst the viverrid clusters (2 and 3) could be an indication of separate evolutionary pathways amongst the different viverrid virus populations in geographically distinct areas. In addition, the large genetic diversity displayed by these viruses also indicates distinct limitation on virus movement and of the likely evolvement of different virus variants that occurred over a long period of time. Evolution of viverrid viruses is different to that of canid viruses (discussed in Chapter 2 of this thesis), which are closely related because of recent introduction into the sub-continent. Our data thus show and support that viverrid rabies is indigenous to southern Africa, and discredit the hypothesis that jackal rabies is indigenous to the region (Smith *et al.*, 1993). The phylogenetic information that we obtained from the 592-bp G-L intergenic region shows that the use of one gene gives meaningful epidemiological

information. This is particularly so since there are no recombination events in lyssavirus replication (Badrane and Tordo, 2001). Based on the molecular epidemiological information obtained here and the high degree of genetic variability amongst the viverrid rabies viruses, future epidemiological studies involving viruses of this rabies variant may be conducted on shorter nucleotide stretches say 200-300 nucleotides per virus sample, and should be sufficient for meaningful epidemiological information.

All the viruses analysed in this investigation, whether from a canid or viverrid host, were found to lack one TTP, similar to observations reported elsewhere (Von Teichman *et al*, 1995). In other studies, an analysis of the G-Psi gene transcription and stop signals (TTPs) of fixed rabies strains (PV and ERA) showed the presence of 2 TTPs, different from another group of street viruses that produced only one TTP (Bourhy *et al.*, 1989; Morimoto *et al.*, 1989; Sacramento *et al.*, 1992). The one TTP arose from the deletion of the G stop signal from the genome. The G stop signal is therefore more of transcriptional, rather than evolutionary significance (Sacramento *et al.*, 1992; Tordo *et al.*, 1986a).

Rabies virus strains have been shown to circulate in particular host species tend to undergo some adaptation (Smith, 1996; Nadin-Davis *et al.*, 1993). These virus strains would at times infect accidental or "wrong" species and this is referred to as spillover. In this investigation, we were unable to establish a single dominant reservoir host species for the viverrid rabies virus variant in Zimbabwe from our genetic studies. The viverrid rabies variant had been isolated frequently in the 1970s from slender mongooses from the Bulawayo area, and the slender mongoose had been assumed to be the reservoir host species of the virus (Foggin, 1988). The viverrid rabies variant was isolated many years

later from other host species such as the honey badger and African wild civet (Bingham *et al.*, 2001). It is therefore possible that other mammalian species that interact with the honey badger, slender mongoose or the African wild civet could also be the reservoir host species. This study has demonstrated that viverrid viruses from southern Africa are highly divergent, with their phylogenetic relationships associated with geographical, rather than host factors.

CHAPTER 4

GENOTYPIC ANALYSES OF MOKOLA VIRUSES FROM ZIMBABWE.

4.1 Summary

In this chapter, a genotypic and phylogenetic analyses of Mokola viruses (genotype 3 of the lyssaviruses) from mammalian species from Zimbabwe are described, and a comparison made with other Mokola viruses from host species in South Africa. All the recorded Mokola isolates from Zimbabwe were obtained for these analyses, and with the exception of isolate number 21846, which was obtained in 1993, all the isolations were made in the early 1980s. A 456 bp region of the nucleoprotein gene was amplified and sequenced. Phylogenetic analysis demonstrated viruses from Zimbabwe to be distinct from those isolated from South Africa. In general, the virus groups obtained here could be correlated with their geographical origin. Mokola viruses from southern Africa were more closely related to Lagos bat (genotype 2) than viverrid rabies viruses (genotype 1).

Part of the data from this chapter has been published in the Journal of South African Veterinary Association. Bingham, J., Javangwe, S., Sabeta, C.T., Wandeler, A.I. and Nel, L.H. 2001. Report of isolations of unusual lyssaviruses (rabies and Mokola virus) retrospectively from Zimbabwe. *Journal of the South African Veterinary Association* **72**(2): 92-4.

4.2 Introduction

In the early 1970s, it became apparent that rabies-related viruses could also cause rabies-like illness in man and other mammalian species (Meredith, 1971; Bourhy *et al.*, 1990). As described in Chapter 1, Mokola virus is a rabies-related virus, and is placed in genotype 3 of the Lyssaviruses. The Mokola virus shares the same genomic organisation with the PV strain of rabies described in Chapter 1 (Pringle, 1995; Mercier *et al.*, 1997; Tordo *et al.*, 1986b; Bourhy *et al.*, 1989). Amongst the Lyssaviruses, Mokola is the most genetically distant virus from rabies virus (genotype 1) (Bourhy *et al.*, 1993b).

Mokola viruses are exclusive to the African continent. The Mokola virus was first isolated from shrews (*Crocidura spp*) from Nigeria and Cameroon, a rodent (*Lophuromys sikapusi* or rusty-bellied forest rat) from the Central African Republic and a cat from Ethiopia (Kempt *et al.*, 1972; Le Gonidec *et al.*, 1978; Saluzzo *et al.*, 1984; Mebatsion *et al.*, 1992). The virus was also isolated from humans in Nigeria (Familusi and Moore, 1972; Shope *et al.*, 1970). The first isolation of Mokola virus in South Africa was made in 1970 and then further isolations made on a number of occasions between 1995 and 1998 (Von Teichman *et al.*, 1998; Meredith *et al.*, 1996) from domestic mammals. In Zimbabwe, Mokola viruses were isolated during surveys for rabies-related viruses in the early 1980s (Foggin, 1982, 1983, 1988) from domestic cats and a dog. No further Mokola isolates were reported in Zimbabwe until 1993 when another Mokola virus was isolated from a domestic cat (Bingham *et al.*, 2001).

Chris Foggin (1988) conducted a study in which mice that had previously been vaccinated with animal (Rabisin, Institut Pasteur) and human vaccines (Human Diploid Cell Vaccine) were challenged with Mokola and the CVS strain of rabies respectively.

These studies showed that rabies vaccines used could offer little protection to Mokola virus infection (Foggin, 1988). To date, anti-rabies vaccines are still not protective against Mokola virus infection (Mercier *et al.*, 1997). The fact that anti-rabies vaccines are not protective against Mokola is further evidenced by the death of domestic animals from Mokola-induced encephalopathy, despite having been previously vaccinated against rabies (Von Teichman *et al.*, 1998; Foggin, 1988). Other studies have shown that PV elicits virus neutralising antibodies (VNA) against genotypes 1, 4, 5 and 6, but failed to protect against genotypes 2 and 3 (Bahloul *et al.*, 1998; Fekadu *et al.*, 1988).

Here we report on the genetic analysis of Mokola viruses from Zimbabwe and a comparison with other Mokola virus nucleic acid sequences from South Africa (Nel *et al.*, 2000), in order to establish their genetic relatedness. Apart from the genetic analysis of Zim82 (Bourhy *et al*, 1993b), this report presents the first genetic data of Mokola viruses from Zimbabwe. The results show Mokola viruses from Zimbabwe to be different from those from South Africa.

4.3 MATERIALS AND METHODS

4.3.1 Viruses and virus isolation

Mokola virus isolates (Table 5) were stored in the form of 20% mouse brain material in the collection of the Onderstepoort Veterinary Institute (Pretoria). Mokola viruses from Zimbabwe had been previously typed using anti-lyssavirus nucleocapsid Mabs produced in mice by the Centre of Expertise for Rabies, Canada (Bingham *et al.*, 2001) and previously by Wiktor *et al.* (1980, 1984a). With the exception of one isolate (21846) made in 1993, all the Mokola viruses from Zimbabwe were isolated in the early

1980s. Other Mokola viruses from South Africa were also included and compared to those from Zimbabwe (Nel *et al.*, 2000). The year of isolation, host species of origin and geographical origin of the virus isolates are shown in Table 5 and Figure 17.

Table 5: Virus isolates included in the study

Virus No.	Lab Ref #	Genotype	Species of origin	Country of origin	Yr of isolation	long-lat	Accession no.
1	12341	3	Feline	Zimbabwe	1981	28°37'-20°11'	AF319515
2	12574	3	Feline	Zimbabwe	1981	28°35'-20°08'	AF319516
3	13270	3	Feline	Zimbabwe	1982	28°47'-20°05'	AF319514
4	15010	3	mongoose	Zimbabwe	1984	30°51'-20°04'	AF319518
5	21846	3	Feline	Zimbabwe	1993	30°26'-18°02'	AF319517
6	Zim82	3	Feline	Zimbabwe	1982	28°47'-20°05'	AF074812
7	252/97	3	Feline	South Africa	1997	30°48'-29°48'	AF074816
8	229/97	3	Feline	South Africa	1997	30°51'-29°48'	AF074815
9	071/98	3	Feline	South Africa	1998	29°36'-30°24'	AF074817
10	700/70	3	Feline	South Africa	1970	31°05'-29°45'	AF074811
11	112/96	3	Feline	South Africa	1996	27°55'-32°57'	AF074810
12	543/95	3	Feline	South Africa	1995	27°46'-32°56'	AF074814
13	322/96	3	Feline	South Africa	1996	32°48'-27°58'	AF074813
14	Lagos Bat	2	n/a	n/a	n/a	n/a	*
15	256/90	1	*Otocyon megalotis*	South Africa	1990	22°55'-29°14'	AF074820
16	810/95	1	*Canis mesomelas*	South Africa	1995	26°36'-26°36'	AF074821
17	420/90	1	*Cynictis penicillata*	South Africa	1990	26°14'-27°13'	AF074819
18	1236/80	1	*Genetta genetta*	South Africa	1980	22°45'-26°09'	**
19	710/90	1	*Cynictis penicillata*	South Africa	1990	25°15'-29°54'	AF074818

* (Bourhy *et al.*, 1995)
** (Jacobs, 1997)

4.3.2 Primer selection

It was decided to target the central part of the nucleoprotein gene (Figure 14) for amplification. An N(1)-N(2) primer pair was used to amplify a portion of the nucleoprotein gene. N(1) primes at position 587-605 and the reverse primer [N(2)-] at position 1029-1013 of the N gene. The numbering of the primer positions is according to the published PV sequence (Tordo, 1996, 1986b).

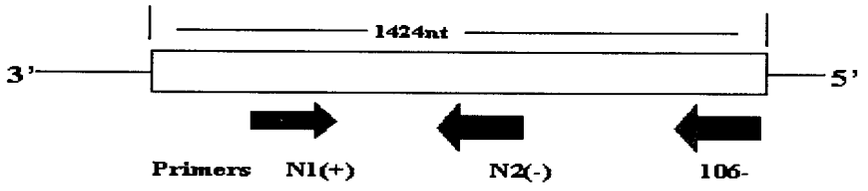

Figure 14: Oligonucleotide primers used for the amplification and sequencing of a portion of the nucleoprotein gene of the Mokola viruses (modified from Nel *et al.*, 1993).

4.3.3 Viral RNA extraction and cDNA synthesis.

Total viral RNA was extracted from the virus isolates as described in section 2.3.3 on page 47. For cDNA synthesis, 5 µl (approximately 0.8 µg) of total viral RNA was denatured with 100 ng of N1 (+) primer at 65°C, and immediately cooled on ice. The rest of the reaction conditions were as described in section 2.3.4 in Chapter 2 page 47.

4.3.4 PCR amplification

Ten microlitres of cDNA were added to a reaction mixture containing 5 µl of reaction buffer (5X), 5 µl of dNTPS (1 mM), 3.5 microlitres of N1(+) primer (100 ng/µl) (5'-TTTGAGACAGCCCCTTTTG-3'), 3 microlitres of N2(-) primer (100 ng/µl) (5'-CCCATATAGCATCCTAC-3') (Tordo, 1996; Nel *et al.*, 2000), 1.25 U of Taq DNA polymerase (Promega) and made up to 50 µl with sterile distilled water. This was subjected to amplification in a GeneAmp 2400 thermal cycler, with the following cycling parameters [94°C for 1 min, 50°C for 45 s, 72°C for 1 min] for 20 cycles. After electrophoretic separation in 1% agarose gels, the PCR amplicons were viewed with a UV transilluminator.

4.3.5 Purification of PCR amplicons and nucleic acid sequencing

The PCR amplicons (456 bp) were purified from the agarose gels and recovered using Geneclean™ as described in Chapter 2, page 50. Fifty nanograms of purified PCR products were sequenced on an ABI377 sequencer (PE Applied Biosystems) as described in Chapter 2, page 50.

4.3.6 Phylogenetic analysis

Phylogenetic trees (Figures 15 and 16) were constructed as described previously in Chapter 2 on page 51. The phylogenetic trees were displayed using TreeView (Page, 1996).

4.4 RESULTS

4.4.1. Mokola virus isolates, virus propagation and monoclonal antibody typing.

Mokola virus isolates from Zimbabwe were isolated from cats and an unidentified mongoose species. The viruses were identified as Mokola using a panel of Mabs (Dr Wandeler, Centre of Expertise for Rabies, Canadian Food Inspection Agency, Canada), which could differentiate rabies and rabies-related viruses (Bingham *et al.*, 2001).

4.4.2. Viral RNA extraction, cDNA synthesis and PCR.

Total viral RNA was extracted from 20% lyophilised mouse brain material and then used as templates for cDNA synthesis. A PCR amplicon of the expected size of 456 bp was generated with the use of the N1(+) and N2(–) primers for each of the virus isolates.

4.4.3 Purification of PCR amplicons, nucleic acid sequencing and phylogenetic analysis

PCR amplicons (456 bp) generated from viral specific cDNA templates were purified from agarose gels by Geneclean (Bio101, U.S.A.) and sequenced on an ABI377 sequencer as per the manufacturer's instructions. Nucleotide sequences of the nucleoprotein gene of the virus isolates are shown as a multiple alignment (Appendix 5). The nucleic acid sequences of the corresponding regions of viverrid rabies viruses, Mokola isolates from South Africa (Nel *et al.*, 2000; Jacobs, 1997) and Lagos bat (Kissi *et al.*, 1995) were also included for comparison.

Sequence alignment of the Mokola viruses from southern Africa show that South African viruses have a C at position 122, whereas those from Zimbabwe (except for 21846) have a T at the same position. Based on genetic distance matrix (Appendix 6), the three Mokola viruses (#'s 3, 4 and 6) from Zimbabwe are practically identical. These viruses were all isolated in the early 1980s. Mokola viruses isolated from Zimbabwe have a maximum sequence divergence of 10.4% (0.00-10.4%), whereas those from South Africa have a maximum sequence divergence of 9.2% (1.8-9.2%). The sequence difference between isolates from South Africa on the one hand and those from Zimbabwe on the other, were found to be as high as 14.5%.

Nucleic acid sequences of the nucleoprotein gene were aligned in ClustalW (Thompson, 1994; Higgins and Sharp, 1988), and their phylogenetic relationships derived using the NJ method of Saitou and Nei (1987) (Figure 15). Similar topologies were obtained with both the NJ methods and Kitsch output option (Figures 15 and 16, respectively). The Mokola virus isolates from Zimbabwe formed a distinct genetic group, different from the one comprising viruses from South Africa (Figure 15) with a statistically significant bootstrap value (100%).

4.4.4. Accession numbers

The nucleic acid sequences of the Mokola viruses analysed in this study have been submitted to Genbank (Table 5).

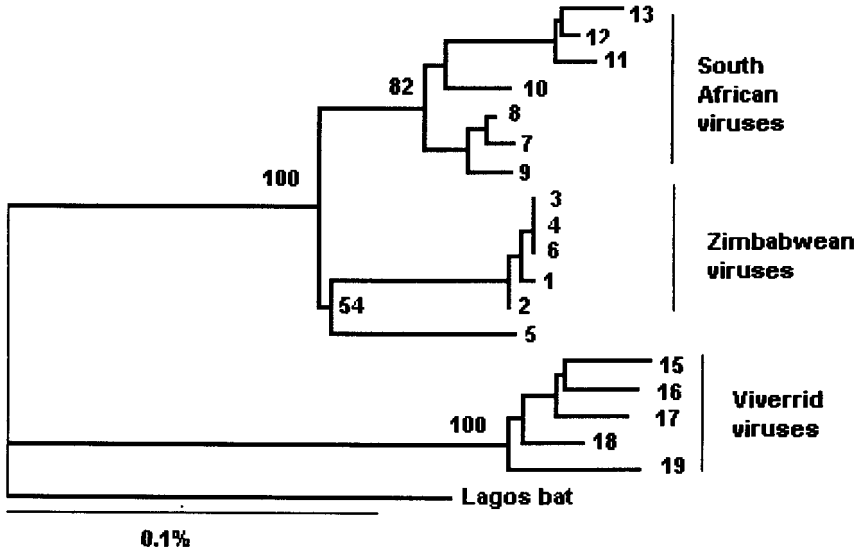

Figure 15: A neighbour-joining tree showing the relationships of Mokola viruses from Zimbabwe and South Africa generated according to the method of Saitou and Nei (1987). Viverrid rabies viruses from South Africa were included in the analysis to show the relationships of southern African Mokola viruses with those from genotype 1 (viverrid rabies viruses) of the lyssaviruses. Bootstrap analysis was undertaken with a seed number of 111 and a 1000 replicate sets.

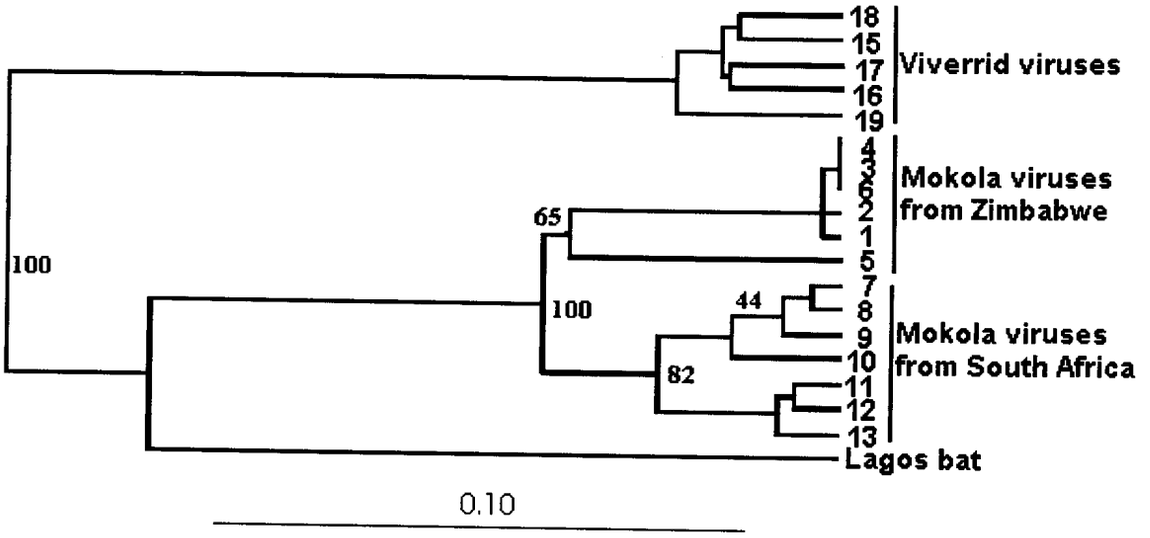

Figure 16: A phylogenetic tree showing the relationships of the Mokola viruses from Zimbabwe and South Africa generated with the Kitsch algorithm. The relationships are presented as an unrooted tree with branch lengths being proportional to the estimated genetic distances between the virus isolates. The scale bar represents 10 nucleotide substitutions per 100 positions. Viverrid viruses from South Africa were included in this analysis to show the relationships of Mokola viruses with viruses from genotype 1. Other viruses such as Lagos bat (genotype 2) (Bourhy *et al.*, 1995; Kissi *et al.*, 1995) were included for comparison with isolates from this study. The phylogeny is based on sequence alignment of the N1-N2 nucleoprotein gene sequences (Appendix 5) using the Kitsch option as described elsewhere (Felsenstein, 1993).

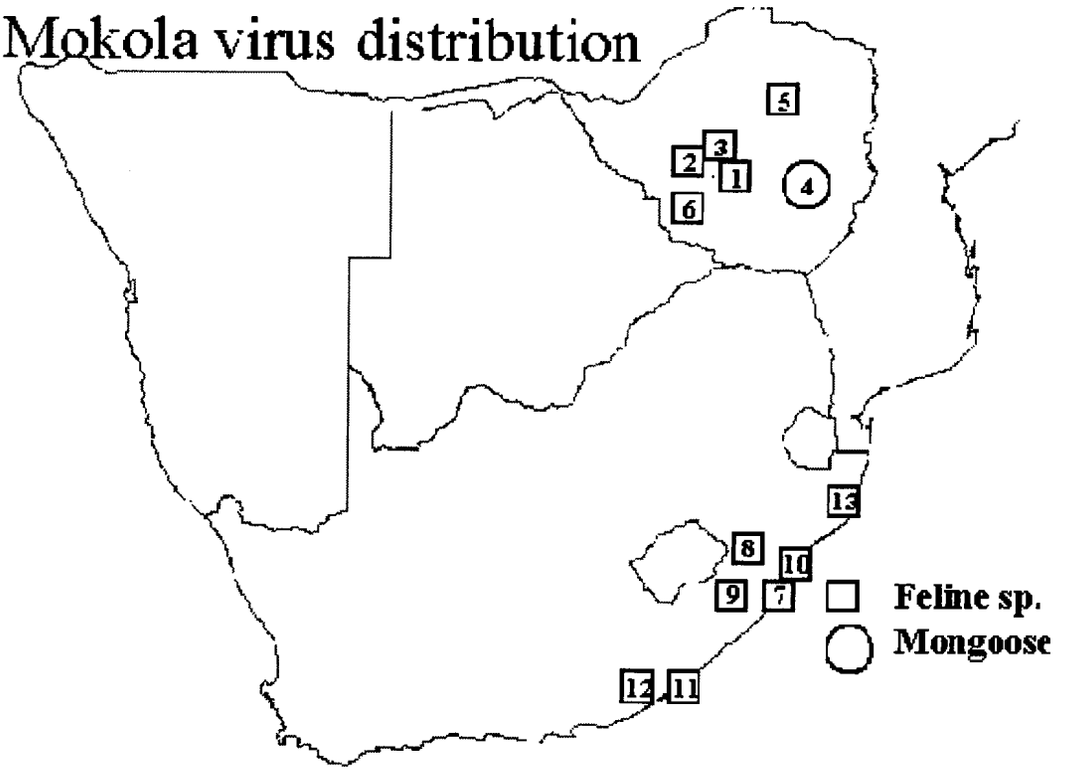

Figure 17: Map of Zimbabwe and South Africa showing the geographical location of the Mokola viruses included in this analysis.

4.5 Discussion

The genetic studies of Mokola viruses described here have demonstrated that viruses from Zimbabwe are different from those from South Africa. It was our aim to establish the genetic relatedness amongst the different Mokola virus isolates of the southern African countries of Zimbabwe and South Africa. Considering that Mokola viruses from these 2 adjacent southern African countries had such a large genetic diversity (of up to 14.5%), it would be interesting to find out to what extent viruses from this region would compare with Mokola viruses from other parts of Africa. Such a large genetic diversity amongst a very small number of viruses in southern Africa may indicate the long periods that these viruses have taken to evolve (Nel *et al.*, 2000; Bingham *et al.*, 2001). In addition, there could be other factors that result in reduced constraints on genetic variation of Mokola virus evolution in southern Africa. Changes in the host environment have generally been associated with a shift in the dominant variant in a quasi-species, a factor that may influence the large genetic diversity as observed here with Mokola. Such an enormous capacity of RNA viruses to adapt to changing host environments (Morimoto *et al.*, 1998) may partly explain the emergence of new pathogens, such as the silver-haired bat rabies variant, that has recently been associated with a proportion of recent human rabies cases in the United States (Rupprecht *et al.*, 1995).

The virus isolates included in this study were obtained from Zimbabwe. However, Mokola viruses other than the ones described in this chapter and originally isolated and documented by Foggin (1988) could be available from elsewhere (Foggin, 1988; Bingham, personal communication). These Mokola viruses could be included for

further studies in future. Amongst the viruses from Zimbabwe, it could be shown that 3 isolates (#'s 3, 4 and 6) are practically identical, and differed in 1 or 2 positions over the entire 292-bp nucleoprotein region (comparing the Zimbabwean isolates only). The 122-nucleotide residue was shown to be a thymine for Zimbabwean viruses, and a cytosine for South African viruses, did not correspond to an antigenic site or a particular protein folding pattern. This site could have an implication in pathogenesis, but at the moment it is not very clear. Three of the Mokola viruses from Zimbabwe, isolated between 1981-1984, were isolated from host species from the same geographical location. The more recently identified Mokola isolate from Zimbabwe (5), made in a domestic cat 13 years after the initial Mokola outbreak in the Bulawayo area, had a larger sequence divergence in comparison to all the other Mokola viruses. From the nucleic acid comparisons of the Mokola viruses from Zimbabwe, it can be concluded that isolates #'s 3 and 6 are likely to be the same virus isolate. At the time this study was conducted, the only information available for number 3 was that it was isolated in 1982. Subsequent nucleic acid sequence analysis of the Mokola viruses showed isolate # 6 to be identical to isolate # 3.

The Mokola viruses from Zimbabwe formed one distinct genetic group and this is suggestive of the strong influence that geographical influence has on virus evolution. Such influence on viral evolution has been shown for canid and viverrid viruses (discussed in chapters 2 and 3 of this thesis) and for rabies viruses previously (Ito *et al.*, 2001a; Nadin-Davis *et al.*, 1999; Smith *et al.*, 1992).

In the 1980s, Chris Foggin conducted a series of experiments in which he inoculated Mokola viruses into experimental wild carnivore species including genets, mongooses and jackal species. During the course of these experiments, it is mentioned

that the genets, a jackal and a mongoose died of Mokola-induced infection, but we are not clear about the fate of the other experimental animals. Could this, for lack of other plausible explanations, suggest why a virus made in a mongoose (# 4) and a cat (# 3) are identical (Foggin, 1988). Records of the virus isolate (# 4) in question show that it was made in an unidentified mongoose species from Masvingo province. The fact that the mongoose isolate was made 300 km away from Bulawayo (where other Mokola isolates were made), we would not expect these viruses to cluster and therefore rules out the mongoose isolate to be a *bona fide* Mokola virus. It is therefore a strong possibility that the ampoule containing the virus could have been labelled wrongly. The virus isolate in question (# 4) was made available for genetic studies from Onderstepoort Veterinary Institute. During the investigation, there were several attempts to amplify this virus isolate. Only then (1999) was the virus confirmed by antigenic typing to be of the Mokola biotype. Antigenic typing of the Mokola isolate (# 4) was achieved by Mabs that could easily differentiate rabies and rabies-related viruses. Such a panel of Mabs facilitated the retrospective identification of another Mokola isolate (# 6) (Bingham *et al.*, 2001). For example, the Centocor conjugate that was used in the 1980s in Zimbabwe was unable to detect Mokola isolates such as isolate # 4 (OVI report, 2000). The confirmation of isolate # 4 by Mabs to be Mokola described above, ruled out possible errors involving amplification and sequencing.

Mokola viruses were also compared with virus isolates from GT1 (viverrid rabies) and GT2 (Lagos bat). Our data confirms that Mokola viruses from Zimbabwe and South Africa are more closely related to the Lagos bat than to viverrid rabies virus, as reported elsewhere (Nel *et al.*, 2000). Partial sequences of the nucleoprotein gene have been used

for epidemiological purposes. In these studies, it was demonstrated that analysis of a limited number of sequences of the N protein of lyssaviruses gave minor branching differences and lower bootstrap scores compared to an analysis of whole gene sequences (Bourhy *et al.*, 1995; Ito *et al.*, 2001a). In our analyses, we have obtained high bootstrap values, suggesting that the region of the genome targeted was appropriate and length of sequence used sufficient to give useful epidemiological information.

One of the most important and unanswered questions in Lyssavirus epidemiology is the reservoir host(s) for Mokola virus. Most of the Mokola viruses from southern Africa were made in domestic mammals (Von Teichman *et al.*, 1998; Nel *et al.*, 2000; Foggin, 1988), although previous isolations of Mokola elsewhere in Africa were from shrews, a rodent and humans (Familusi and Moore, 1972). Foggin (1988) suggested that Mokola viruses could be naturally confined to wild reservoir host species and that their presence was recognised during surveys or only on rare occasions, when they are transmitted to domestic animals and humans. It may therefore be possible that the candidate for the reservoir host species will be found amongst the shrews, bats and rodents, these being mammalian species that are preyed upon by cats. The presence of virus neutralising antibodies (VNA) to Mokola in sera of animals tested during surveys for rabies-related viruses could support such a postulation of a reservoir host being found amongst the rodents. Results of these surveys further showed that sera from some of these animals, mainly rodents and a side-striped jackal, had neutralising activity against Mokola but not against rabies virus indicating previous exposure to the Mokola virus (Foggin, 1988).

In South Africa, Mokola viruses were first isolated in the 1970s and repeatedly isolated thereafter between 1995 and 1998 (Von Teichman *et al.*, 1998). In Zimbabwe, Mokola viruses were isolated much later in the early 1980s and once again in 1993. The reason why Mokola viruses have been isolated only in a few African countries could be due to the lack of active surveillance for rabies-related viruses. The isolations of Mokola have occurred far apart on the African continent, which is likely to be an indication of the lack of research effort, rather than the true distribution of these viruses. This may imply that the incidence of Mokola and other rabies-related viruses could be much higher than is currently reported.

The isolation of a Mokola virus from a domestic cat in Zimbabwe in 1993 and other recent isolations from South Africa, indicate the continual presence of this virus (Bingham *et al.*, 2001; Von Teichman *et al.*, 1998). Research efforts should therefore be directed at establishing the true incidence of Mokola virus through active surveillance. Our understanding of the epidemiology of Mokola and other unusual lyssaviruses can be improved with active surveillance coupled by both antigenic and genetic typing.

Rabies vaccines are ineffective against the Mokola virus and this has been shown experimentally and in circumstantial situations (Foggin, 1988; Nel *et al*, 2000; Von Teichman *et al.*, 1998; Fekadu *et al.*, 1988). Challenge-protection tests conducted in mice (Foggin, 1988) showed little protection to a Zimbabwean Mokola virus in comparison to protection against CVS virus challenge, but the human scenario is not known. This makes Mokola a virus of public health importance. The fact that there is no vaccine protective against Mokola virus has generated substantial interest in that direction. For instance, recent studies that make use of DNA-mediated immunisation

(DMI) with glycoprotein segments from various genotypes, have demonstrated the potential for such vaccines to protect against the whole *Lyssavirus* genus (Bahloul *et al.*, 1998, Jallet *et al.*, 1999). This approach may increase the spectrum of efficacy of classical rabies vaccines. However, considering that Mokola is exclusive to Africa, the problems with these vaccines is that their market is limited and costs may be prohibitive. Despite the failure by many African countries to control rabies in dogs and its spread to humans, I strongly feel that public health hazards associated with the Mokola virus, be made known to laboratory workers, researchers, veterinarians, gamekeepers and pet owners.

CHAPTER 5

CONCLUDING REMARKS

The aims of the studies described in this thesis were to characterise rabies virus variants and Mokola viruses from southern Africa (Zimbabwe and South Africa specifically) in order to further our understanding of the epidemiology and diversity of these lyssaviruses. It was also our objective to further our knowledge on the role that these variants play in rabies epizootics in the southern African sub-continent, in order to institute proper and adequate control measures for this horrific disease.

The predominant lyssavirus genus responsible for rabies encephalitis in Africa is rabies virus. Zimbabwe and generally other African countries are the regions in the world where rabies infection offers a great diversity and this has increased our interest in rabies-related viruses (Swanepoel *et al.*, 1993; King and Crick, 1988; Foggin, 1988). Two aspects of epidemiology, viz surveillance and knowledge of the distribution of the antigenic and genetic virus variants, are essential components of an efficient and economical rabies control program. It is apparent that in other parts of the world notably the Americas, continental Europe and Latin America, both surveillance and virus typing have led to the identification of compartmentalisation of distinctive variants by either species-specific and/or geographical factors (Smith, 1996). Rabies is widespread in most African countries but not as properly and adequately controlled as in the developed world. Zimbabwe is one of the many African countries faced with significant socio-economic and political problems, such that previous gains in rabies control have come to nought.

Epidemiological trends in Zimbabwe show that dog rabies has generally been associated with communal rural areas and jackal rabies with the commercial farming areas. The recent land reform programme in Zimbabwe has resulted in the removal of some of the former commercial farming boundaries. These changes can be expected to contribute to an increase in rabies incidence as there are increased chances of domestic and wild animals interaction. The recent outbreak of foot and mouth disease in Matabeleland and Masvingo provinces of Zimbabwe can be attributed to the removal of such boundaries.

The WHO website (www.who.int) reports that 90% of the global rabies cases occurs in Asia but no data is given for Africa. The reasons for this may be that the disease is not being reported to WHO or that rabies does not feature as priority. In many African countries, AIDS/HIV and poverty are problems that most governments are struggling to alleviate/eradicate and this makes rabies, despite it being a horrific but manageable disease, a low priority. Most of the deaths from rabies are young children (under 15) (WHO 1989; 1991). The recent death of a nine-year old girl of rabies from Zimbabwe illustrates the plight that some people in the developing world are facing. Africa is poor and is less able to prevent such a preventable disease as rabies. In my opinion, there should be increased involvement of international organisations such as the UN and the WHO that should work more closely with rabies groups in Africa, such as the South Eastern African Rabies Group (SEARG). The WHO or UN could use such an infrastructure to disseminate information and provide an aid package that could include rabies vaccine and anti-rabies prophylaxis biologicals.

Genetic characterisation of lyssaviruses has facilitated our understanding of rabies epidemiology, prevention, treatment and development of laboratory techniques for diagnosis. However, despite the greater specificity of genetic typing methods, I still think that broad coverage is more readily accomplished using the simpler protocols of antigenic typing. In general, the phylogenetic relationships of lyssaviruses have been primarily influenced by geographical factors and by species factors only in a few instances. There was a larger heterogeneity amongst the viverrid viruses in comparison to the canid viruses. Our data further shows canid viruses to be of a recent and common genetic lineage in comparison to that of Mokola and viverrid rabies viruses.

The identification of animal reservoir(s) for Mokola virus is still a challenging and tall order for rabies researchers. In most areas of the world, which are endemic for rabies, reservoir status is generally inferred from case surveillance data. Now that we have biologicals to analyse samples, it will be useful to screen by antigenic typing methods, rabies samples that are part of repositories in central veterinary laboratories (of which they are very few) in Africa. I also think that in order to improve rabies surveillance there is need to set up study sites in rural areas and obtain samples directly from individuals living in those areas. The samples can then be sent to central laboratories, which in Africa are mainly found in the main cities.

The information that we obtained from this study has expanded our understanding of the molecular epidemiology of lyssaviruses (rabies and Mokola viruses) in Zimbabwe and South Africa. Our research efforts should be further directed towards active surveillance of lyssaviruses by both genetic and antigenic typing methodologies. In addition, we should strive to develop vaccines with a wider spectrum of efficacy against

lyssaviruses. Although a number of laboratories are currently working towards this direction, problems with these vaccines are that the market for this research is limited and the victims for which they are intended for, are mostly very poor.