UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Molecular epidemiology of African mongoose rabies and Mokola virus

by

Nicolette van Zyl

Submitted in partial fulfillment of the requirements for the degree

MAGISTER SCIENTIAE (MICROBIOLOGY)

in the

Department of Microbiology and Plant Pathology

Faculty of Natural and Agricultural Science

University of Pretoria

Pretoria, South Africa

Supervisor: Prof. Louis H. Nel

Co-Supervisor: Dr. Wanda Markotter

December 2008

I declare that the thesis, which I hereby submit for the degree M.Sc (Microbiology) at the University of Pretoria, South Africa, is my own work and has not been submitted by me for a degree at another university

_____

Nicolette van Zyl

Acknowledgements

My deepest gratitude to the following people for their guidance and academic input

Prof. Louis H. Nel (University of Pretoria)

Dr. Wanda Markotter (University of Pretoria)

Dr. E.C. Holmes (Center for Infectious Disease Dynamics, Pennsylvania State University)

A special word of thanks to

Dr. Jacqueline Weyer (formerly part of the University of Pretoria) for assistance and guidance with various techniques;

Mr. Albe van der Merwe (University of Pretoria) for assistance with the Beast software package;

Dr. Claude T. Sabeta (Agricultural Research Council, Onderstepoort Veterinary Institute, Rabies Unit, South Africa) for kindly allowing the use their facilities as well as providing viral isolates from the ARC-OVI archive;

Mr. Joe Kgaladi for providing sequence data;

The National Research foundation (NRF) for financial support;

My friends and colleagues in the Virology lab at the University of Pretoria, for their advice, encouragement, support and love during my studies;

Boeta for granting me this opportunity to pursue my tertiary education;

Mamma, Pappa, Ouma and Boeta for all their love, interest, support and encouragement throughout my studies, I couldn't have done it without you;

Louis for his love and understanding, you encouraged me when I needed it most;

The Lord for all His blessings.

Summary

Molecular epidemiology of African mongoose rabies and Mokola virus

by

Nicolette van Zyl

Supervisor: Prof. Louis H. Nel
Co-supervisor: Dr. Wanda Markotter
Department of Microbiology and Plant Pathology
University of Pretoria
For the degree MSc (Microbiology)

The African continent sustains a variety of lyssaviruses and this study focused on two of these lyssaviruses that are unique to the continent namely rabies virus mongoose biotype and Mokola virus (MOKV). Rabies virus (RABV) belongs to genotype (gt) 1 of the *Lyssavirus* genus in the family *Rhabdoviridae*, order *Mononegavirales,* while Mokola virus belongs to gt3 of this genus. Both these viruses cause fatal rabies encephalitis in vertebrate animals. Genotype 1 (rabies virus) isolates from southern African countries display great genetic diversity and are grouped into two main biotypes i.e. canid and mongoose biotypes. Due to the difference in the epidemiology and pathogenesis of these biotypes, it has been hypothesized that the two biotypes were introduced into Africa at different times. The objective was to study the molecular phylogeny of representative rabies virus isolates of the mongoose biotype, isolated in South Africa and Zimbabwe over a period of 27 years, towards a better understanding of the origin of this group. In this study the complete nucleoprotein (1353 nucleotides) and glycoprotein (1575 nucleotides) genes were sequenced. The evolutionary dynamics of this virus variant was investigated using Bayesian methodology, allowing for rate variation among

the different viral lineages. The phylogenetic analysis of this dataset confirms previous findings of extended evolutionary adaptation of isolates in specific geographic areas. Furthermore when these isolates are analyzed together with rabies virus isolates from across the world, they still form an independent cluster separate from any other African rabies virus isolates, thereby hinting towards a separate introduction to the continent before that of canid rabies. Molecular clock analysis estimates the age of the mongoose biotype to be approximately 200 years old, which is in concurrence with literature describing rabies in mongooses since the early 1800's.

In addition, a phylogenetic analysis of Mokola virus isolates (gt3) from South Africa, Zimbabwe, Cameroon and Central African Republic is described. All the South African isolates before 2008, as well as most of the Zimbabwean isolates (except isolate 21846) were included in this analysis. The complete nucleoprotein gene (1353nt) was amplified and sequenced. Phylogenetic analysis showed virus grouping to correspond to their geographic location. Further analysis showed Mokola virus isolates to display genetic diversity similar to that found in representative gt1 isolates.

Table of Contents

List of Abbreviations

µl	microliter
aa	amino acid
ABLV	Australian bat lyssavirus
BC	Before Christ
bp	basepairs
C	Celsius
cDNA	Complementary Deoxyribose nucleic acid
DBLV	Dakar bat lyssavirus
DUVV	Duvenhage virus
EBLV1	European bat lyssavirus type1
EBLV2	European bat lyssavirus type2
EDTA	ethylene diamine tetra acetic acid
ESS	Effective sample size
FAT	Fluorescent antibody test
G	glycoprotein
gt	genotype
HPP	Higher posterior probability
K	Lysine
L	polymerase protein
LBV	Lagos bat virus
M	Molar
M	matrix protein
ML	Maximum likelihood
mM	millimolar
MOKV	Mokola virus

MP	Maximum parsimony
mrca	most recent common ancestor
mRNA	messenger ribonucleic acid
N	nucleoprotein
ng	nanogram
NJ	Neighbor joining
nt	nucleotides
OTU	Operational taxonomic units
P	phosphoprotein
PBS	phosphate buffered saline
PCR	Polymerase chain reaction
pmole	picomoles
PV	Pasteur virus
R	Arginine
RABV	Rabies virus
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RT-PCR	Reverse transcription polymerase chain reaction
s	seconds
S.A.	South Africa
spp.	species
U.S.A	United States of America
UK	United Kingdom
VNAb	Virus neutralizing antibodies
WCBV	West Caucasian bat virus
WHO	World Health Organization

Chapter 1: Literature review

1.1. History of rabies

The history of rabies virus can be traced back to antiquity. The word “rabies” is a Latin word that originated from a Sanskrit word “rabhas” which when translated means “to do violence” (Steele and Fernandez, 1991). It is believed that Democritus was the first to make a recorded description of rabies in 500 years B.C. and even in ancient times it was already understood that the bite of a rabid dog will transmit the disease to humans or other animals (Steele and Fernandez, 1991). In Mesopotamia, the physician Aetious gave a very accurate description of manifestation of rabies in dogs. Symptoms include dogs becoming mute, delirious and unable to recognize their owners or surroundings, the animal refuses food or water, the mouth is usually open with the tongue hanging out, there is an abundance of saliva and the dogs tend to move around slowly (Steele and Fernandez, 1991).

The association between the bite of an infected animal and disease development was already well understood in the 4th century BC, when Aristotle wrote of “dogs suffering from madness, and that all animals they bite become diseased”. Many writers, especially the Romans, spoke of the infectivity of the saliva of rabid dogs in ancient writings (reviewed: Smithcors, 1958). In 1804, long before Pasteur began work on rabies, Zinke experimentally demonstrated the transmission of rabies virus through saliva. Krugelstein reached the conclusion in 1826 that rabies was a disease that affected the nervous system, and between 1852 and 1854, Bouchardat began conducting experiments to develop vaccines against rabies. In 1881 Pasteur published his first writing on rabies, and stated that “the central nervous system and especially the bulb which joins the spinal cord to the brain are particularly concerned and active in the development of the disease.” A few years later, he stated that the causative agent was not a bacterium, but a virus, and that clinical signs were dependant on the dose of infective material (Pasteur *et al.,* 1884). Pasteur reported on the attenuation of a street virus strain in rabbits in 1885, and thereby laid the groundwork for the development of the first neural tissue vaccine in 1886 (Pasteur, 1886). In the decades following Pasteur’s breakthrough numerous scientists studied rabies. In 1903, Remlinger discovered the filterability of the disease agent and Aldechi Negri (1903) described “protozoa” that would later be identified as neural viral cytoplasmic inclusions called

Negri-bodies. Fermi (1908) improved the rabies vaccine by treating nerve tissue with carbolic acid, and Semple (1919) produced the first inactivated neural tissue vaccine in 1919. Other notable advances of the 20th century include the development of vaccines that were prepared from virus grown in suckling mouse brain, as well as embryonated chicken and duck eggs (Steele and Fernandez, 1991).

The first large outbreak of rabies occurred in Franconia in 1271. It was reported that 30 people died following the bites of rabid wolves that invaded the towns and villages. From here onwards the spread of rabies across most of the world can be traced through history (Fleming, 1872). According to Irvin (1970) rabies was often reported in Europe during the first half of the 19th century. These epidemics were mostly controlled through host extermination, and apart from dogs, common wildlife vectors included foxes, badgers, bears and wolves. It is speculated that the transport of dogs on ships attributed to the spread of rabies to other parts of the world. In 1770, rabies was recorded in Boston, in 1741 in India and 1813 in Mauritius (Irvin, 1970). Rabies became well-established in wolves and foxes in North-America at the beginning of the 19th century, and by the end of this century, rabies in skunks was also a common occurrence (Irvin, 1970). The distribution of rabies in South America during this time was rather patchy, with rabies being reported in Chile, Peru, Argentina, and sometimes in Brazil and Mexico (Fleming, 1872). In Asia rabies was mostly found in hyenas, jackals, wild dogs and wolves as well as domestic dogs (Irvin, 1970). It has been speculated that rabies had been present in northern Africa for hundreds of years, but the cosmopolitan variant has only recently spread to southern Africa due to colonization that reached a peak in the early 1900's (Nel and Rupprecht, 2007). The Sahara Desert most probably hampered overland spread of rabies southward from the palearctic region through Africa. But in spite of this natural barrier, the cosmopolitan dog-associated rabies virus has spread across the African continent (Nel and Rupprecht, 2007). Before the presence of the cosmopolitan lineage in Africa, Hudson (1944) described the presence of rabies in dogs and jackals in East Africa, a dog virus has been circulating in West Africa (Bouffard, 1912) and the mongoose variant was present in southern Africa (Snyman, 1940). The variant of West Africa was known as "Oulou fato" and generally caused paralytic rabies, with infected animals showing no tendency

towards aggression and biting (Bouffard, 1912; Nicolau *et al.,* 1933). These virus variants will be discussed in more detail in section 1.5.1.5

1.2. Classification of lyssaviruses

The order *Mononegavirales* is made up of all the single-stranded, negative-sense RNA viruses belonging to the families *Rhabdoviridae, Bornaviridae, Filoviridae* and *Paramyxoviridae*. Rabies belongs to the *Lyssavirus* genus, that together with the genera *Vesiculovirus, Ephemerovirus, Novirhabdovirus, Cytorhabdovirus* and *Nucleorhabdovirus*, form the *Rhabdoviridae* family (Tordo *et al.,* 2005). The lyssaviruses are currently classified into seven genotypes (Figure 1.1) based on genotypic characteristics (Bourhy *et al*., 1993; Gould *et al*., 1998), but with the addition of new isolates from Eurasia as well as Africa this number may increase. Phylogenetic analysis of these viruses revealed that LBV forms three distinct lineages and it has been proposed that viruses from lineage C form a new genotype named Dakar bat lyssavirus (DBLV), while Irkut, Aravan, Khujand and WBCV all constitute recently acknowledged genotypes (Kuzmin *et al.,* 2003; Kuzmin et al., 2005; Markotter *et al.,* 2008a).

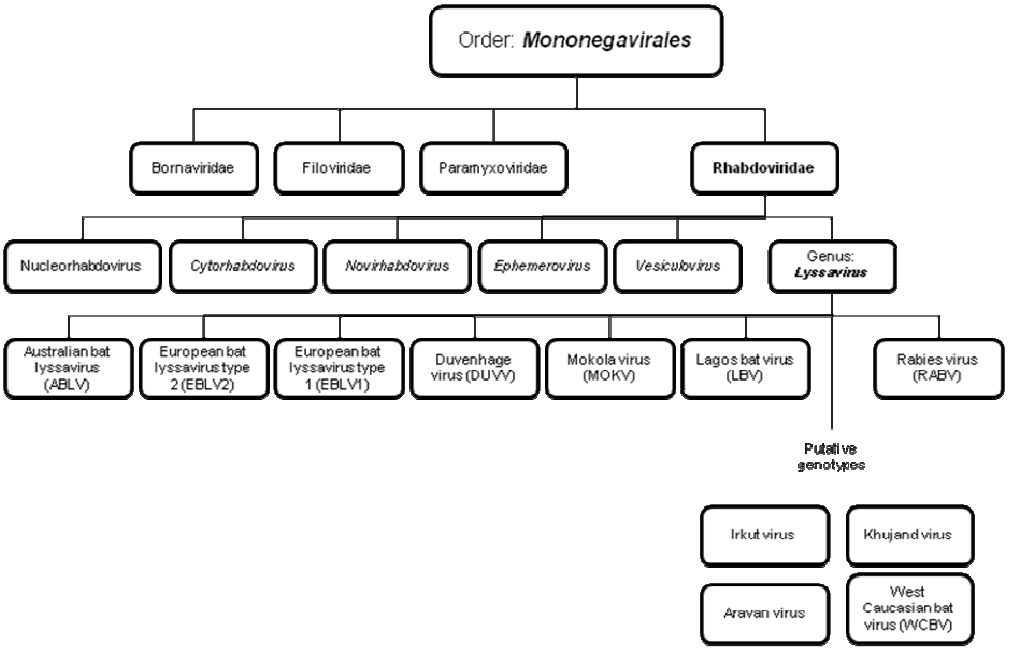

Figure 1.1. Diagram depicting current lyssavirus classification.

1.3 Global distribution of lyssaviruses

Table 1.1 provides an overview of the geographic distribution of the 7 genotypes as well as the putative genotypes contained in the *Lyssavirus* genus. It is estimated that 55 000 people die annually from rabies and that at least 44% of these deaths occur in Africa alone (WHO, 2004). All lyssaviruses are transmitted through the bite of a rabid animal. Most of these viruses affect warm-blooded animals and humans and almost invariably cause fatal disease. Lagos bat virus is the only lyssavirus that has not yet been associated with human deaths (Markotter, 2007).

Table 1.1. Geographic distribution and host species associated with lyssavirus genotypes.

Genotype	***Virus***	***Host***	***Distribution***
1	RABV	Dog, cat, insectivorous bat, fox, raccoon, jackal, mongoose, raccoon dog, skunk, coyote	Worldwide except for Australia, Antarctica, parts of Scandinavia, United Kingdom, several Western European countries and several islands.
2	LBV	Frugivorous bats (*Epomophorus, Rousettus* and *Eidolon* sp.), cat, dog, insectivorous bat, mongoose	Africa
3	MOKV	Shrew, cat, dog, rodent	Africa
4	DUVV	Insectivorous bat (*Nycteris and* possibly *Miniopterus* spp.)	Africa
5	EBLV-1	Insectivorous bat (*Eptesicus* spp.)	Europe
6	EBLV2	Insectivorous bat (*Myotis* spp.)	Europe
7	ABLV	Frugivorous bat (*Pteropus* spp.) and other insectivorous bats	Australia
Putative genotype	Irkut	Insectivorous bat (*Murina leucogaster*)	Central Asia
Putative genotype	Aravan	Insectivorous bat (*Myotis blythi*)	Central Asia
Putative genotype	Khujand	Insectivorous bat (*Myotis mystacinus*)	East Siberia
Putative genotype	WCBV	Insectivorous bat (*Miniopterus schreibersii*)	Caucasus mountain region

Nel and Markotter, 2007

The major reservoir for RABV in most developing countries is the domestic dog (*Canis familiaris*) and isolates have been obtained from numerous continents including Africa, Asia, Europe, North and South America (Childs, 2002). Other vectors of RABV include wild terrestrial animals such as red and arctic foxes, mongooses, jackals, skunks, raccoons as well as insectivorous and haematophagous bats from North and Latin America (Nel and Markotter, 2007). Rabies virus is considered the prototype member of the *Lyssavirus* genus and constitutes genotype 1 (gt1). The group of viruses that infects mostly members of the Canidae family and was disseminated worldwide between the 1500's and 1900's is often referred to as the cosmopolitan strain of RABV (Kissi *et al.,* 1995). RABV circulating in members of the Herpestidae in southern Africa is referred to as the mongoose biotype (Nel *et al.,* 2005). Gt2-4 viruses are found exclusively on the African continent. Lagos bat virus (gt2) was first isolated in 1956 from a fruit bat (*Eidolon helvum*) on Lagos Island, Nigeria (Boulger and Potterfield, 1958). Hereafter the majority of isolations were from frugivorous bats in Central African Republic (*Microteropus pusilus) (*Sureau *et al.,* 1977), South Africa (*Epomophorus wahlbergi*) (Meredith and Standing, 1988; Crick *et al*., 1982, Markotter *et al.,* 2006a), Dakar (Senegal) (*Eidolon helvum*) (Institute Pasteur, 1985) and France (*Rousettus aegyptiacus*). Additional isolations have been made from domestic cats (Shope, 1982; Crick *et al.,* 1982), an insectivorous bat, *Nycteris gambiensis* (Institut Pasteur, 1985), dogs (Mebatsion *et al*., 1992) and a water mongoose (Markotter, 2006b). Mokola virus (MOKV) constitutes gt3 of the *Lyssavirus* genus, has only been found on the African continent and will be discussed in detail in section 1.6. Duvenhage virus (gt4) has been isolated in South Africa (Paweska *et al.,* 2006), Zimbabwe (Foggin, 1988) and Kenya (Van Thiel *et al.,* 2008). This virus is associated with insectivorous bats, and of the 5 infections to date, three proved to be fatal to humans. When European bat lyssavirus type 1 (gt5) and European bat lyssavirus type 2 (gt6) were discovered it was initially thought that these viruses were "Duvenhage-like" viruses based on monoclonal antibody reaction patterns (King *et al.,* 1990). Later they were characterized as two distinct genotypes (Bourhy *et al.,* 1992). EBLV1 occurs predominantly in *Eptesicus serotinus* (Fooks *et al.,* 2003), but has been found in animals that are not indigenous to Europe i.e. captive fruit bats (Ronsholt *et al.,* 1998), in sheep (Ronsholt, 2002) and in a

stone marten (*Martes foina*) (Muller *et al.,* 2001). EBLV2 is mostly associated with insectivorous bats such as *M. daubentonii* and *M. dasycneme (*Harris *et al.,* 2006*).* Australian bat lyssavirus (gt7) was first isolated in 1996 from a female black flying fox (*Pteropus alecto*) (Fraser *et al.,* 1996). However, ABLV has now been found to occur in all 4 species of flying fox (*Pteropus alecto, P. poliocephalus, P. scapulatus and P. conspicullatus*) as well as an insectivorous bat species, *Saccolaimus flaviventris* (the yellow-bellied sheath-tailed bat) (Hooper *et al.,* 1997). Several viruses belonging to the *Lyssavirus* genus have not been classified as belonging to a existing or forming a new genotype. These include Irkut, Aravan, Khujand and West Caucasian bat virus (WCBV). Aravan virus was isolated in Kyrgyzstan in 1991 from *Myotis blythi* (Arai *et al.,* 2003), Khujand virus in northern Tajikistan in 2001 from *Myotis mystacinus* (Kuzmin *et al.,* 2003), WCBV from *Miniopterus schreibersi* in Caucasus (Botvinkin *et al.,* 2003) and Irkut in Siberia from *Murina leucogaster* in 2002 (Botvinkin *et al.,* 2003).

1.4 The virus

All lyssaviruses share features that classify them with other rhabdoviruses. These include the typical bullet-shaped virion, the helical ribonucleoprotein (RNP) core, the structural proteins of the virus and the RNA genome organization (Tordo *et al.,* 1986a). The bullet shaped rabies virion is approximately 130 - 200nm in length with an average diameter of 60 -110nm (Wunner, 2002). The viral negative sense single-stranded RNA genome bound by molecules of the nucleoprotein (N), phosphoprotein (P) and polymerase protein (L) all make up the helical ribonucleocapsid (RNP) of the virion. An envelope consisting of a layer of matrix protein surrounded by a lipid membrane derived from the host cell membrane during budding surrounds the RNP. The glycoprotein spikes form trimers that are embedded within the lipid membrane of the virus (Wunner, 2002). The rabies virus genome (PV strain) is approximately 11 932 nucleotides (nt) in length and encodes 5 proteins namely the nucleoprotein (N), the phosphoprotein (P), matrix protein (M), glycoprotein (G) and the viral RNA polymerase (L) (Fig.1.2). These coding regions are separated by short non-coding intergenic regions of 2-23nt in length, with the exception of the highly variable G-L intergenic region that is 423nt in length in the Pasteur virus (Tordo *et al.,* 1986b; Wunner, 2002) However some slight differences

in the length of genes and intergenic regions are found between different genotypes (Bourhy *et al.,* 1989, Bourhy *et al.,* 1993; Delmas *et al*., 2008; Kuzmin *et al.,* 2008).

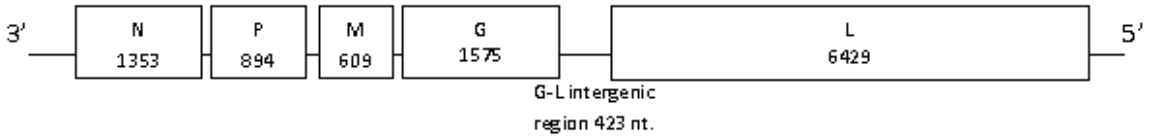

Figure 1.2. Schematic representation illustrating the order and length of genes (in nucleotides) on the RABV genome (PV strain).

- Nucleoprotein

The nucleoprotein is the major protein of the RNP core. It is 450-451 amino acids (aa) in length and is the most conserved gene when comparing amino acid sequence similarity between genotypes, despite short regions of high genetic diversity in the N gene (Kissi *et al.,* 1995, Wunner, 2002). The major function of the nucleoprotein is encapsidation of the RNA genome and viral replicative intermediates (RI) to prevent digestion of these components by endogenous cellular ribonucleases (Tordo and Poch, 1988; Wunner, 2002). The nucleoprotein also plays a role in viral transcription and replication through its ability to promote read-through of the termination signals. The N protein contains unique genotype specific epitopes that allow distinction between viruses of different genotypes on the basis of their reactivity pattern with anti-N monoclonal antibody (Mabs) panels (Schneider *et al.,* 1973; Smith, 1989). More recently sequencing targeting the nucleoprotein has been used to elucidate the phylogenetic relationship between different lyssaviruses (Kissi *et al.,* 1995; Kuzmin *et al.,* 2003).

- Phosphosprotein

The RABV phosphoprotein is phosphorylated, 297-302 aa in length and is a key component of the RNA polymerase complex where it acts as regulatory protein in viral genome replication (Chenik *et al.,* 1994). The P protein furthermore acts as a chaperone of soluble nascent N, thereby preventing its self-assembly and non-specific binding to cellular RNA (Wunner, 2002). When complexed with the N protein the phosphoprotein directs the encapsidation of the viral RNA, and when it is part of the RNA polymerase complex, functions as co-factor during transcription and replication of

the viral genome by stabilizing the L protein and placing the polymerase complex onto the RNA template (Chenik *et al.,* 1998).

- Matrix protein

With a length of only 202 aa the matrix protein is the smallest of the *Lyssavirus* proteins (Tordo *et al.,* 1986b). The multiple functions of this protein include sheathing the ribonucleoprotein-core during virion assembly, inhibiting viral RNA transcription, condensing the helical nucleoprotein cores into tight coils and associating with the membrane bilayers. The interaction of the M protein with the cytoplasmic domain of the G protein facilitates budding of new virions from an infected cell (Finke and Conzelmann, 2003).

- Glycoprotein

The mature glycoprotein is 503-524 aa in length and contains the following domains: a 19 aa signal peptide domain, a 439 aa ecto-domain, a 22 transmembrane domain and a 44 aa cytoplasmic domain (Wiktor *et al.,* 1973; Tordo and Poch, 1988, Wunner, 2002). The signal peptide provides the signal that ensures the insertion of the G protein into the membrane of the rough endoplasmic reticulum-Golgi plasma membrane pathway, and this peptide is cleaved off from the 524 aa glycoprotein after insertion into the Golgi apparatus membrane, thereby yielding the mature 505 aa glycoprotein. Various antigenic domains have been identified on the glycoprotein (Table 1.2).

Table 1.2. Antigenic domains present on the ectodomain of the glycoprotein of lyssaviruses.

Amino acid position	***Reference***
14-19	Mansfield *et al.,* 2004
231(site I)	Lafon *et al.,* 1983
34-42 and 198-200 (site II)	Lafon *et al*., 1983; Prehaud *et al.,* 1988
330-338 (site III), LYS^{330} and ARG^{333} involved in pathogenesis	Lafon *et al.,* 1983; Dietzschold et al., 1983; Seif *et al*., 1985; Tuffereau *et al.,* 1989
264 (site IV)	Dietzschold *et al*., 1990
342-343 (site V)	Benmansour *et al*., 1991

Of these domains Arg333 is probably the best studied and it has been shown that mutations in this area can affect virus pathogenicity (Dietzschold *et al.,* 1983; Seif *et al.,* 1985). Effective uptake of the virus into cells through interaction of the G protein (ectodomain) with cellular receptors can result in increased virulence as it has been reported that non-pathogenic strains of RABV (often strains with a mutation in position 333) use different receptors and routes of entry into cells than more pathogenic RABV strains (Dietzschold *et al.,* 1985; Jackson, 1991). The G protein also mediates pH-dependant fusion of the virus with the endosomal membrane after the virus is internalized by the host cell. Recently Faber *et al.* (2004) showed that in pathogenic strains of RABV the glycoprotein accelerates virus internalization. Furthermore the G protein is involved in the transsynaptic spread of the virus and promotes viral entry into the nervous system from the peripheral site of entry. This protein is the major target for virus specific T-helper cells and cytotoxic T cells, and it contains the major antigenic sites against which virus neutralizing antibodies (VNAb) are directed (Wiktor *et al.,* 1973). Pulmanausahakul *et al.* (2008) emphasized the importance of the interaction between the G and M protein in budding and cell-to-cell spread of rabies virus. All of these functions indicate why the G protein plays an important role in determining virulence and is one of the best studied RABV proteins (Wunner, 2002).

- Polymerase protein

The RABV polymerase protein, at 2127-2142 aa (Tordo *et al.,* 1988; Le Mercier *et al.,* 1997), comprises more than half of the genome's encoding potential. Together with the P protein the L protein (which is the catalytic component of the polymerase complex) is responsible for most of the enzymatic activities during transcription and replication of the viral genome. The polymerase is not only responsible for primary transcription of the RNA genome, but also performs co-transcriptional modifications such as 5' capping, methylation and 3' polyadenylation of the viral mRNA's (Tordo *et al.,* 1988; Poch *et al.,* 1990; Wunner, 2002).

- Non-coding regions

The non-coding regions (intergenic sequences) are located between the 3' end of one mRNA (nine nucleotide stop codon, WGA_7) and the 5' start of the following mRNA (start

codon with nine nucleotide consensus sequence, AACAYYHCT for RABV) (Tordo *et al.,* 1986b, Bourhy *et al.,* 1989, Gould *et al.,* 1998). The intergenic regions of RABV are 2 (N-P), 5 (P-M), 5 (M-G) and 423nt (G-L) in length. The G-L intergenic region is of specific interest as it contains two sequences resembling RABV transcription consensus sequences. Transcription of the G-L intergenic region has not yet been reported, and it is unlikely that the G-L region, as it is seen in present RABV, encoded a structural protein, since this transcript will only yield an 18 aa peptide that will most probably be too small to be of any structural value. Therefore Tordo *et al.* (1986b) proposed that this intergenic region is a remnant open reading frame (ORF). Phylogenetic studies based on the sequence of the G-L intergenic region have shown it to be the most variable region of the lyssavirus genome, as it is not subjected to the selection pressure of coding regions that is brought about by structural requirements of the encoded proteins (Bourhy *et al.,* 1995).

1.5 Genotype 1: Rabies virus (RABV)

Rabies due to infection with gt1 viruses is maintained almost world-wide through the interaction of two ecological cycles. The one cycle is primarily linked to domestic dogs (urban rabies) while the other cycle is maintained by a variety of wildlife vectors (sylvatic rabies). Humans and livestock are usually victims of a spill-over infection from the primary hosts and do not contribute to the maintenance of rabies (Childs, 2002).

1.5.1 Worldwide epidemiology of rabies virus (RABV)

1.5.1.1 America

- ***North America***

Canine rabies in the USA has been described in historical reviews since 1703 (Baer, 1994), but only became widely distributed in the late 19th century (Nadin-Davis and Bingham, 2004). In the archives of the State of Virginia, reference is made to rabies in dogs in 1753 (Fleming, 1872). Various epizootics were reported and by 1860, canine rabies had spread across North America (Billings, 1884). Before 1945, domestic dogs were the most common reservoirs of RABV in North America, but due to effective vaccination campaigns, canine rabies was brought under control in the USA (Steele,

1988) and Canada in the 1940's and 1950's (Finnegan *et al.,* 2002). The vaccination efforts however did not curb the spread of sylvatic rabies in these areas. Antigenic and genetic characterization of RABV isolates originating from North-America has enabled the identification of the major reservoir hosts involved in the different wildlife variants of RABV. In the USA the grey fox (*Urocyon cinereoargenteus*), the striped skunk (*Mephitis mephitis*), raccoons (*Procyon lotor*) and coyotes (*Canis latrans*) all maintain at least one distinct variant of RABV in a specific geographic area (Smith, 1996; Belotto *et al.,* 2005). In Canada the red fox (*Vulpes vulpes*) is an important vector while the arctic fox (*Alopex lagopus*) maintains the endemic arctic lineage of RABV in the northern regions of the country. Recently it was reported that certain areas have managed to eliminate red fox rabies through oral vaccination campaigns (Rosatte *et al.,* 2007). Apart from the disease in numerous terrestrial animals, a number of insectivorous bat species function as hosts of RABV (Smith, 1996; Belotto *et al.,* 2005). Species often associated with rabies include *Eptesicus fuscus*, *Lasionycteris noctivagans, Myotis* spp. and *Lasiuris* spp. (especially *L. cinereus*), but is possible that up to 30 different lineages of RABV might exist in bat species of North America (Nadin-Davis *et al.,* 2001; Davis *et al*., 2006). In Arizona, bat rabies variants were transmitted to skunks, and this variant managed to establish skunk-to skunk transmission in a number of cases. However the last cases were observed in 2006 and it seems that an independent cycle of RABV in skunks is not being maintained (Leslie *et al.,* 2006).

- ***South and Central America***

Records of canine rabies in South America date back to the early 19th century (Steele, 1975). Since then dog rabies has been endemic in Central and South America and is maintained by domestic dogs (urban cycle), vampire bats (*Desmodus rotundus*), insectivorous bats in sylvatic cycles (Favi *et al*., 2002; Kobayashi *et al*., 2005). After dogs, bats are most frequently responsible for transmission of RABV to humans (Kobayashi *et al.,* 2005). Vampire bats are also primarily responsible for transmission of RABV to livestock such as cattle (Belotto *et al*., 2005). Canine rabies remains a problem in South American cities such as Mexico City, Sao Paulo and Buenos Aires and wildlife rabies is also reported in monkeys, wolves, coyotes, skunks and foxes (Belotto *et al.,* 2005). In Cuba, Puerto Rico and other Caribbean islands, small Indian mongooses

(*Herpestes auropunctatus*) that were imported from India to control rat populations in sugar cane fields, are now also important vectors of RABV (Everard and Everard, 1988).

1.5.1.2 Australia

After the last case of human rabies due to the bite of a rabid canine was reported in 1867 (Beran, 1981), Australia was considered "rabies-free" until 1996 when ABLV was discovered. Occasional cases of dog rabies due to importation are reported (Warrilow, 2005).

1.5.1.3 Europe

After rabies in canids were brought under control through mass vaccination in the 1940's, rabies in red foxes escalated in Europe, and the red fox remains the most important vector of rabies virus in Europe today (Bourhy *et al.,*1999; Finnegan *et al.,* 2002). After vaccination campaigns targeting foxes, numerous countries in Europe (i.e. United Kingdom, Finland, The Netherlands and others) brought fox rabies under control and are considered terrestrial rabies free (Bourhy *et al.,* 2005). Raccoon dogs (*Nyctereutes procynoides*) that were imported into Eastern Europe for fur production and the Arctic fox also plays a role in continuing the rabies infection cycle in eastern Europe, Norway and the former USSR respectively (Smith, 1996; Finnigan *et al.,* 2002). The rabies-related lyssaviruses (EBLV1 and EBLV2) have also been identified in Europe (see section 2.5 and 2.6).

1.5.1.4 Asia and Middle East

Rabies in Middle Eastern countries such as Turkey, Jordan and Iran have reported significant cycles of canine rabies, with spillovers into wildlife species like jackals, squirrels, foxes, monkeys and wolves also being regularly reported (Nadin-Davis *et al.,* 2003; Johnson *et al.,* 2003).

In Asia, canine rabies is responsible for at least 31 000 human deaths per annum (WHO, 2004). Japan is the only rabies-free country in Asia today, and only a few countries (Indonesia, Malaysia, China, the Philippines, Sri Lanka, Thailand and Vietnam) have rabies surveillance programs in place (WHO, 2004). Rabies has been

described in China for more than 2000 years (Wang and Huang, 2001) and since 1950 three major epidemics have occurred, with the third still continuing (Zhang et al., 2006).The majority of human rabies cases in China can be attributed to canine rabies (Zhang et al., 2006). In Korea rabies occurs in dogs as well as red foxes and raccoon dogs, even though extensive vaccination campaigns are conducted to try and curb rabies in this area (Hyun *et al.,* 2005). In India, Sri Lanka, Thailand and Indonesia, significant numbers of human rabies cases are reported. In India, up to 95% of the 20 000 human deaths caused by rabies can be attributed to canine rabies (Sudershan *et al.,* 2007). Wildlife rabies is sporadically reported in Asia, with red foxes, arctic foxes, mongooses and jackals being implicated (Nadin-Davis *et al.,* 2007). Generally the poor canine rabies control and inadequate post- exposure prophylaxis (PEP) that gives rise to the high incidence of human rabies cases in Asia, is driven by economic considerations.

1.5.1.5 Africa

The history of rabies in Africa prior to the 20th century is not well-documented. But, in spite of this fragmented knowledge it appears that rabies in sub-Saharan Africa only became epizootic in the middle 1900's (Nel and Rupprecht, 2007). When looking at the emergence of rabies in Africa, North Africa and sub-Saharan Africa should be regarded separately. It has been hypothesized that rabies in Northern Africa was primarily an urban disease, closely associated with rabies cycles in the Middle East, similar to the current situation is these areas, and that canine rabies (cosmopolitan lineage) spread to sub-Saharan Africa from the Palaearctic region (Europe, Middle East and North Africa) due to European colonization (Badrane and Tordo, 2001; Nel and Rupprecht, 2007). This cosmopolitan lineage became well-established in domestic dogs and wildlife after its dissemination across the continent, and emerged as a serious disease in the 1950's. Today, urban rabies is a major concern throughout Africa, with the domestic dog being the main vector of the disease. Three rabies-related viruses (Lagos bat virus, Mokola virus and Duvenhage virus) also occur exclusively on the African continent.

- ***North Africa***

Rabies is predominantly an urban disease in northern Africa (e.g. Algeria, Morocco, and Egypt) and has most probably been endemic for hundreds of years (Nel and Rupprecht, 2007). These countries annually report several hundred cases of canine rabies, most often in stray dogs. In the sparsely populated countries spanning the Sahara desert, rabies occurs only as scattered foci, and is primarily spread by dogs belonging to nomads or refugees, although occasional cases of rabies in camels have been recorded (Swanepoel, 2004).

- ***West and Central Africa***

It seems that rabies was present in West Africa before the arrival of the cosmopolitan lineage in the 20th century. Historical accounts from this region reveal that this lineage (known by locals as "Oulou fato") is considered to be less virulent than the cosmopolitan lineage, and that dogs can as such appear to be in a "carrier state", thereby transmitting virus without being taken ill themselves (Irvin, 1970; Nel and Rupprecht, 2007). Furthermore if dogs did develop disease it usually gave rise to paralytic rabies, with dogs showing no tendency towards aggression. This form of rabies has been recognized in Niger and Senegal, but appears to also have extended over many countries to the horn of Africa as well as eastward into Sudan and Ethiopia (Snyman, 1940). However, no description of this disease has been made recently (Irvin, 1970; Nel and Rupprecht, 2007). Furthermore Badrane and Tordo (2001) as well as Kissi *et al.* (1995) presented evidence that another rabies virus lineage currently circulating in West Africa can be distinguished from the current cosmopolitan variant. In addition, the cosmopolitan variant of rabies is dominant and widespread across this region today and primarily affects dogs and occasionally wildlife, livestock and humans in the West African countries of Nigeria, Senegal, Niger, Benin, Guinea, the Ivory Coast, Ghana, Central African Republic and Cameroon (Kissi *et al.,* 1995) as well as Ethiopia and Sudan (Johnson *et al.,* 2004a). This variant is also of conservation concern as the continued existence of an endangered canid species (the Ethiopian Wolf) is under threat due to large numbers of animals succumbing to rabies (Randall *et al.,* 2004).

- ***East Africa***

Similar to other regions of Africa, it would appear that rabies was familiar to indigenous people long before the arrival of Europeans. However, due to diagnostic techniques developed in the early 1900's, rabies was diagnosed for the first time in 1912 in Kenya (Nel and Rupprecht, 2007). A serious rabies epidemic spread from southern Tanzania across the entire country in the 1950's, and this epidemic has persisted in Tanzania and neighbouring Kenya, where it still remains a problem today (Nel and Rupprecht, 2007). African wild dogs in Kenya are also under threat since breeding pairs have died of rabies (East and Hofer, 1996) and other wildlife affected include the black-backed jackal and bat-eared foxes. In Malawi and Uganda rabies is wide-spread and just as big a concern as in other countries of East Africa (Nel and Rupprecht, 2007).

- ***Southern Africa***

As is the case in other African regions, the documentation of the history of rabies before the 20th century in southern Africa is somewhat sparse. Cluver (1927) however reports several outbreaks of rabies in South Africa between 1827 and 1890. The first official diagnosis of rabies was made in 1893 by Hutcheon in Port Elizabeth from a dog imported the previous September. A number of local dogs, cats and ruminants were affected, but the outbreak was brought under control by 1894 (Hutcheon, 1894). In 1902 an outbreak of rabies was reported in Bulawayo, Zimbabwe, and the disease spread across the country. Rigid measures were put in place to curb the disease and by 1914 no further cases were reported (Adamson, 1954; Bishop *et al.,* 2003; Dept of Agriculture archives, 2007). Rabies had been present in Zambia since before European influx, and it is believed that this outbreak in Zimbabwe was introduced from the Zimbabwean border with Zambia in 1901 (Swanepoel *et al.,* 2004).

Canine rabies only became endemic in southern Africa around 1947 as a result of an epizootic that originated in the border region between Namibia, Zambia and Angola. Rabies then spread across northern and eastern Botswana into Zimbabwe, reaching the northern parts of South Africa in the 1950's. From here, rabies spread to Mozambique in 1952 and followed the eastern coastal region of South Africa to reach Swaziland in 1954. Canine rabies spread from Mozambique into KwaZulu-Natal (KZN) and caused a

major epidemic in 1961. This epidemic was brought under control by 1968, but by 1976 rabies had re-entered KZN from Mozambique and has since then proven very difficult to control in the densely populated rural and peri-urban settlements of this province. The disease also entered the Free State, Lesotho and reached the Eastern Cape by 1987 (Swanepoel *et al.,* 1993; Coetzee and Nel, 2007). The spread of canid rabies gave rise to the emergence of rabies in wildlife throughout Namibia, Zimbabwe, Botswana and South Africa. From 1967 onwards, cases of rabies in bat-eared foxes (*Otocyon megalotis)* were reported in Namibia and the Northern Cape, and in 1977 to 1985, a unique outbreak of rabies was observed in kudu (*Tragelaphus strepsiceros*) in Namibia (Swanepoel *et al.,* 1993). However in 2002-2003 another outbreak of rabies in these antelope was reported (Mansfield *et al.,* 2006). The black-backed jackal also acts as a vector of rabies in the Limpopo province since the early 1950's, and still remains a problem today. Other vectors of rabies in southern Africa include suricates (*Suricata suricatta*), yellow mongooses (*Cynictis penicillata*) and other members of the Viverridae like genets (Bishop *et al.,* 2003).

In southern Africa we find the unique situation of two distinct biotypes of gt1 that circulate independently in different hosts. The canid biotype circulates primarily in canid hosts with occasional spill-over into other warm-blooded animals like ruminants, cats and humans, while the mongoose biotype primarily circulates in members of the Herpestidae and Viverridae (Swanepoel *et al.,* 1993; Johnson *et al*., 2004b). In the past members of the Herpestidae (i.e. mongooses) and Viverridae (i.e. genets and civets) were all grouped in the same family as Viverridae (Estes, 1995). Thus initial studies of rabies viruses associated with feliform carnivores (suborder *Feliformia*) included some viruses from genets but primarily focused on viruses from mongooses. Due to the previous classification of the host species these viruses were characterized as belonging to a "viverrid" biotype (King *et al.,* 1993a; Nel *et al*., 1993). The mongooses have since been reassigned to their own family (i.e. the Herpestidae) due to morphological differences with viverrids (Gregory and Hellman, 1939; Wozencraft, 1993) and thus the viverrid biotype became known as herpestid rabies, but Nel *et al.* (2005) have proposed that this virus biotype now be known as the "RABV mongoose biotype" due to its predominant association with these animals.

Mongoose rabies

Mongoose rabies has been present in South Africa since before the introduction of the canid rabies lineage. Snyman (1940) reports that according to local inhabitants, it was believed that the genet had been responsible for rabies outbreaks since 1885. Cluver also reported in 1927 that Europeans and natives in certain regions believed the bite of the genet to invariably provoke madness and death. Since 1916, several cases of suspected rabies had been reported from different areas of South Africa with strong evidence pointing to the involvement of yellow mongooses and genets as vectors of disease (Cluver, 1927). However, the first laboratory confirmed diagnosis of rabies caused by the bite of a mongoose was only made in 1928, when two children developed rabies after being bitten by a mongoose (Hertzenberg, 1928). It was believed the current infection of certain wild carnivores like mongooses and genets did not originate from the canine rabies in Port Elizabeth or Zimbabwe, but had been present in this area for a very long time, and had successfully adapted itself to members of the Viverridae (i.e. genets and civits) and Herpestidae (i.e. mongooses and suricates) families, which were then exclusively responsible for maintaining and propagating the infection. When this virus was transmitted to dogs or humans, the infection cycle invariably reached a dead end (Chaparro and Esterhuysen, 1993; Bishop *et al.,* 2003; Dept of Agriculture archives, 2007). Historical records also tell of rabies in Herpestidae in other parts of Africa. In the 1970's in Zimbabwe, mongoose rabies was primarily isolated from the slender mongoose (*Galerella sanguinea*) and this led to the assumption that this mongoose species acted as reservoir for this particular rabies virus variant in Zimbabwe (Foggin, 1988). Due to the increasing isolation of mongoose RABV from other animals such as the African civet and the honey badger (*Mellivora capensis*) (Bingham *et al.,* 2001) it is thought that the slender mongoose is only an indicator host, unlike the yellow mongoose which has been demonstrated to be an reservoir host in South Africa (Nel *et al.,* 2005). Recently Johnson *et al.* (2004b) also reported on the presence of rabies virus mongoose biotype in Botswana.

There was also a confirmed infection of an Egyptian mongoose (*Herpestes ichneumon*) in Israel (Irvin, 1970). However, this virus was never subjected to typing to confirm whether it was a canid virus or not. Even though 12 species of mongoose occur in

southern Africa, the yellow mongoose is most often implicated in rabies transmission, together with suricates and the slender mongoose (Chaparro and Esterhuysen, 1993).

It had already been recognized by Snyman in 1940 that the rabies of canids differed from rabies in mongooses in terms of host pathogenicity and epidemiology and he coined the terms European type and mongoose type. King *et al.* (1993b) demonstrated the different reaction patterns of these viruses against a panel of monoclonal antibodies directed against the nucleoprotein, with the canid biotype showing a constant reaction pattern for all isolates, while the mongoose biotype displayed some variance in pattern between isolates. This led researchers to believe that there was significant heterogeneity between isolates from this biotype. Nel *et al.* (1993) and Von Teichman *et al.* (1995) confirmed this finding through phylogenetic analysis of the nucleoprotein gene and the G-L intergenic region. These studies again demonstrated that the canid viruses were closely related to the European vaccine strains and each other, but were distant from the mongoose viruses, which clearly formed a separate group. Spillover of the mongoose biotype into other animals such as ruminants and even canids has been observed, but these are invariably dead-end infections (Nel *et al.,* 1998). This again clearly displays how well this virus variant has adapted to its host species. The reason for this lack of sustained transmission is unknown. It is certainly possible that there may be a difference in the behavior of animals when infected with the mongoose virus as opposed to their behavior when infected with a canid virus. Also few in-depth comparative pathogenesis studies have been conducted to determine if there is a difference on a cellular level between the progressions of infection of the two virus types. Chapparo and Esterhuysen (1993) did demonstrate that mongooses are less susceptible to the canid biotype virus, with less animals succumbing to infection and very little virus being shed in saliva when compared to mongooses infected by the mongoose biotype rabies virus. Further study conducted by Jaftha *et al.* (1997) and Nel *et al.* (2005) investigated the genetic diversity displayed by the mongoose viruses and found that these viruses formed a number of distinct phylogenetic clusters that corresponded with the geographic region of isolation of the virus isolates. This suggests that these virus groups have been evolving independently within different subsets of the mongoose population in the central plateau and other regions of South Africa. This

hypothesis was further supported by the study of Taylor in 1993 that showed a strong correlation between different mongoose subpopulations and the highly localized outbreaks of rabies in these animals. A recent study by Davis *et al.* (2007) again confirmed that mongoose rabies was present in South Africa before the arrival of canid rabies, lending further support to the view that mongoose rabies, unlike the canid variant, is indigenous to southern Africa. However, the origin of this virus is still unknown.

1.6 Genotype 3: Mokola virus (MOKV)

Mokola virus is a rabies-related lyssavirus and belongs to gt3 of the *Lyssavirus* genus (Bourhy *et al.,* 1989). Previously when comparing viruses from all 7 genotypes, Mokola virus was the most genetically divergent from gt1 viruses (Bourhy *et al.,* 1993) and has only been isolated on the African continent. However, it has recently been determined that West Caucasian bat virus is currently the most divergent lyssavirus known (Kuzmin *et al.,* 2005).

1.6.1 MOKV isolates

The isolation of MOKV has been reported very infrequently from only a few countries where surveillance and investigations were conducted and up to date there have been 25 isolations of this virus from across Africa (Table 1.3).

Table 1.3. Mokola virus reports.

Year of isolation	***Location***	***Species of origin***	***Reference***
1968/1969	Ibadan,Nigeria	Shrews (3)	Shope, 1970, Kemp *et al.,* 1972
1968/1971	Nigeria	Human (2)	Familusi *et al.,* 1972; Familusi and Moore, 1972
1970	Umhlanga Rocks, KwaZulu Natal province, South Africa	Cat	Meredith *et al.*, 1996
1974	Cameroon	Shrew	Le Gonidec *et al.,* 1978
1981/1982	Bulawayo, Zimbabwe	Dog, cat (4)	Foggin, 1982, 1983
1981	Central African Republic	Rodent	Saluzzo *et al.,* 1984
1990	Ethiopia	Cat	Mebatsion *et al.,* 1992
1993	Selous, Zimbabwe	Cat	Bingham *et al.,* 2001
1995	Mdantsane, Eastern Cape province, South Africa	Cat	Meredith *et al.,* 1996
1996	East London, Eastern Cape Province, South Africa	Cat	Von Teichman *et al.,* 1998
1996	Yellow Sands, Eastern Cape Province, South Africa	Cat	Von Teichman *et al.,* 1998
1997	Pinetown, Kwazulu Natal Province, South Africa	Cat (2 isolates)	Von Teichman *et al.,* 1998
1998	Pietermaritzburg, Kwazulu Natal Province, South Africa	Cat	Von Teichman *et al.,* 1998
2005	Nkomazi, Mpumalanga Province, South Africa	Dog	Sabeta *et al.,* 2007
2006	East London, Eastern Cape Province, South Africa	Cat	Sabeta *et al.*, 2007
2008	Eastern Cape province, South Africa	Cat	Personal communication (C.T. Sabeta)

Most of the Mokola virus isolates were obtained after samples had been submitted for routine rabies diagnosis. Many laboratories in Africa use conjugates that are immunologically bound to rabies antinucleocapsid antibodies that cannot distinguish between different lyssavirus genotypes when conducting a fluorescent antibody test (FAT) (Mebatsion *et al.,* 1992; Bingham *et al.,* 2001). Thus the conjugate will only confirm the presence or absence of a lyssavirus, but it will not distinguish which genotype of the lyssavirus is responsible for the infection. Therefore it is possible that other cases of Mokola virus have been misdiagnosed as rabies virus (gt1) infection as further analysis cannot be performed due to lack of widespread availability of a distinguishing MAb panel in laboratories in Africa.

MOKV was first isolated in 1968 in Nigeria from the pooled organs of shrews (*Crocidura* spp.) (Shope *et al.,*1970). In 1970, Shope *et al*. demonstrated the serological relationship between rabies virus and the newly isolated Mokola virus from Nigeria through complement fixation and other serological testing was performed by Kemp *et al*. in 1972. Two suspected human cases of MOKV infection were also reported (Familusi and Moore, 1972; Familusi *et al.,* 1972). However these cases are regarded as highly suspect as only a single virus isolation was made from one of the patients while the other patient didn't succumb to the infection. This virus isolate was analyzed in the same laboratory that was conducting studies on the first MOKV isolate from the shrews which casts some doubt over the authenticity of this diagnosis. Thereafter isolations were made in Cameroon (Le Gonidec *et al.,* 1978), Central African Republic (Saluzzo *et al.,* 1984), Ethiopia (Mebatsion *et al.,* 1992), Zimbabwe (Foggin, 1982; Foggin, 1983; Bingham *et al.,* 2001) and South Africa (Meredith *et al.,* 1996; Von Teichman *et al.,* 1998; Sabeta *et al.,* 2007). As with rabies virus (gt1), clinical symptoms of animals infected with Mokola virus may include acute encephalomyelitis, fever, salivation, aggression and progressive unusual behavior and disorientation, but will inevitably progress to fatal rabies (Von Teichman *et al.,* 1998; Nel *et al.,* 2000, Bingham *et al.,* 2001).

Mokola virus is currently the only lyssavirus that has not been isolated from bats. In the majority of the Mokola cases from South Africa, as well as the cases from Zimbabwe,

cats were the infected species. This is however a dead-end infection as felines succumb to infection with this virus. It is therefore easy to speculate that the reservoir hosts for Mokola virus may be some prey species of domestic cats. However, attempts to isolate the virus from various species (i.e. shrews and rodents) that may be preyed upon by cats has failed, and this again underlines the little knowledge we have about the epidemiology of this virus (Meredith *et al.,* 1996; Nel *et al.,* 2000).

1.6.2 Rabies virus vaccines and MOKV

The glycoprotein of lyssaviruses contains several antigenic sites of which antigenic domain III is especially important in viral pathogenicity. Positively charged residues K330 and R333 play an important role in the ability of the virus to produce a lethal infection. Rabies virus isolates contain an arginine in position 333 while Mokola virus isolates contain an aspartic acid residue in position 333, and also contain other amino acids than rabies virus in the positions surrounding position 333 (Tordo *et al.,* 1993; Badrane *et al.,* 2001). Mokola virus together with Lagos bat virus make up phylogroup II viruses, while gt1 together with the other lyssavirus genotypes (gts 4,5,6,7) form phylogroup I (Badrane *et al.,* 2001). Current commercially available vaccines are all based on gt1 virus strains. It has been shown that viruses in these two phylogroups display differences in pathogenicity (Badrane *et al.,* 2001), and due to the great diversity between viruses of the two phylogroups, it is not surprising that traditional rabies vaccines do not protect against challenge with Mokola virus (Bahloul *et al.,* 1998). At least four of the cats that were infected with Mokola virus in South African had been vaccinated with standard rabies vaccines 7 months, 12 months, 15 months and 36 months before the animals were diagnosed (Nel *et al*., 2000). Challenge experiments conducted by Mebatsion *et al.* (1992) tested the efficacy of gt1 based vaccines to protect mice against infection with Mokola virus. It was found that these vaccines do not protect mice against challenge with Mokola virus. These experiments, the cases of the vaccinated cats mentioned previously, together with the case of a Mokola virus infected dog that had been vaccinated against rabies six months before it was diagnosed with rabies and died in 1982 in Zimbabwe, demonstrate the lack of cross protection by rabies vaccines against infection with Mokola virus (Nel *et al.,* 2000). Due to the lack of protection by commercial gt1 based vaccines and the implied public health risk,

research is being conducted on the development of a DNA vaccine that will protect against infection with Mokola virus (Nel et al., 2003) as well as a chimeric vaccine that will protect against gt1 rabies virus infection but also against infection with Mokola virus (Bahloul *et al.,* 1998). It is hoped that by affording protection against such divergent genotypes, that the vaccine will also protect against infection with gt 2,4,5,6 and 7 viruses.

1.6.3 Possible origin of MOKV

There are several features that are shared by members of the *Mononegavirales.* These include similar gene order, complimentarity of the 3' and 5' termini, a presumptive single 3' terminal promoter, transcription by sequential interrupted synthesis, replication by synthesis of a complete positive sense transcript, the virion-associated RNA-dependant RNA polymerase and maturation by budding (Pringle and Easton, 1997). Despite the diversity observed between the different families of the *Mononegavirales*, these viruses all seem to have originated from a common ancestor (Conzelmann, 1998). When studying the viruses of the *Rhabdoviridae* family, it appears that all rhabdoviruses may share a monophyletic origin as five of the *Rhabdoviridae* genera contain viruses that are transmitted by insect or have insect hosts in nature. Thus it seems likely that the primary host of the rhabdovirus ancestor was an insect (Hogenhout *et al.,* 2003). Prior to 2006, it was thought that two unclassified arthropod-borne viruses (Obodhiang virus and Kotonkan virus) belong to the rabies virus related serogroup after some cross complement fixation experiments showed some cross reactivity with Mokola virus (Buckley, 1973; Bauer *et al.,* 1975; Shope, 1982; Smith, 1996). This was thought to indicate a possible ancestral link between the rabies virus serogoup and arthropod borne viruses. However Kuzmin *et al*. (2006) showed that these viruses should rather be classified as members of the *Ephemerovirus* genus based on sequencing data of the partial nucleoprotein gene. Even though all members of the *Lyssavirus* genus are currently only adapted to infect mammals, Mokola virus has been shown to readily infect and multiply in the *Aedes albopictus* mosquito cell line as well as in mosquitoes themselves and *Aedes albopictus* cells were susceptible to rabies virus infection in vitro (Reagan and Wunner, 1985; Seganti *et al*., 1990). This ability to multiply in insect hosts or cells is speculated to hint at a possible insect progenitor virus that may have

given rise to the current lyssaviruses. The true host of Mokola virus in nature is still unknown, and it has been speculated that it may be a species which possibly plays a role as a prey species for small animals such as shrews and cats. The mode of transmission of Mokola virus is suspected to be through biting by an infected animal. Studies have shown that subcutaneously and orally-infected shrews are capable of transmitting the virus by biting adult mice. The ability of Mokola virus to orally infect shrews possibly suggests that a lower life-form such as a subterranean vertebrate or non-vertebrate might be able to support virus replication (Aitken *et al.,* 1984). However, the ability to orally infect animals with a lyssavirus has not been proven beyond any doubt. Because bats are hosts to most lyssaviruses there has been speculation that lyssaviruses originated from an insect rhabdovirus, which in turn was then contracted by insectivorous bats from these insect hosts. However in the case of Mokola virus it may be that an insectivorous reservoir species contracted the virus from an insect host (Badrane *et al.,* 2001). The evolutionary link of lyssaviruses with other rhabdoviruses now seems less clear after it has been shown that all "insect lyssaviruses" were not lyssaviruses after all (Bourhy *et al*., 2005; Kuzmin *et al.,* 2006). Thus future characterization of other unclassified rhabdoviruses and discovery of new viruses may provide more information as to the true evolutionary link.

1.7. Lyssavirus evolution and epidemiology

The advent of molecular techniques such as PCR and nucleotide sequencing empowered scientists to conduct research on the taxonomy, epidemiology and possible evolution of organisms. The generation of sequence information led to the development of phylogenetic analysis software packages that enable researchers to now analyze organisms on a molecular level, and this analysis is becoming increasingly important as a tool of investigation around the world (Strimmer and Robertson, 2005). The use of phylogenetic studies to establish speciation, geographical links and possible common ancestry has made this an important method to study the *Lyssavirus* genus (Kissi *et al.,* 1995; Bingham *et al.,* 1999; Paez *et al.,* 2003; Velasco-Villa *et al.,* 2005; Kuzmin *et al.,* 2005; Real *et al.,* 2005).

The majority of phylogenetic studies conducted on lyssaviruses have utilized the N and G proteins. The fact that the N gene is well conserved allows the comparison of isolates that are suspected to have undergone long-term evolution. Furthermore the abundance of the N mRNA in infected cells makes this an ideal target for DNA sequencing (Bourhy *et al.,* 1992). The glycoprotein is involved in host cell receptor recognition and membrane fusion, which makes it the major elicitor of host neutralizing antibodies. Because of the function of glycoprotein, the G gene undergoes some mutation in order to evade the host's immune system; therefore it has generally been believed that when analyzing isolates that are suspected of being closely related the G gene is a better choice due to its increased variability when compared to the N gene (Badrane *et al.,* 2001).

Wu *et al.* (2007) recently proposed that, in viruses belonging to the *Lyssavirus* genus, all genes when analyzed individually should generate the same tree topology when the same method is applied to each gene. This is possible because no evidence of recombination has been found in lyssaviruses. Therefore, in spite of two redundant areas in the lyssavirus genome (Psi-region between G and L protein and the area from the first stop codon of the M to the next transcription initiation signal), all genes of the lyssavirus genome should be suitable for use in a phylogenetic analysis. However Markotter *et al.* (2008a) showed that all genes are not suitable for use in classification of lyssaviruses. Therefore it appears that full genome analysis might provide the best phylogenetic resolution.

1.7.1 Host origin of lyssaviruses.

Speculation about the origin of lyssaviruses has mostly centered on the "insect origin" hypothesis, where it is suggested that insect rhabdoviruses gave rise to lyssaviruses after transmission to insectivorous bats in the distant past (Shope, 1982; Badrane and Tordo, 2001). The fact that members from all genera in the Rhabdoviridae family have either been isolated from insects or are able to replicate in insect cells lends supporting evidence to this hypothesis (King and Crick, 1988). This could point towards rhabdoviruses, and specifically lyssaviruses having a possible progenitor insect virus. Badrane and Tordo (2001) dated the most recent common ancestor of all lyssaviruses

(possibly an insect virus) to have existed between 7080 and 11631 years ago. This study employed molecular clock analysis based on the G gene sequence data from chiropteran as well as carnivoran rabies isolates, and also included representative members of all the other lyssavirus genotypes (gt2-6). It seems plausible that ancient spillover events from insects to bats to carnivores, and subsequent host adaptation, gave rise to the diversity of gt1 subgroups we observe today (Badrane and Tordo, 2001).

1.8. Molecular phylogeny

Molecular phylogeny is a field of science that aims to decipher the evolutionary development of a specific organism, species or higher taxa (www.answers.com). Molecular phylogeny specifically employs molecular data such as DNA sequences to reconstruct this evolutionary link between organisms. This relationship is often shown in the form of a phylogenetic tree. Phylogenetic trees are graphs or mathematical structures that depict the evolutionary relationship between organisms. The actual pattern of ancestor-descendant relationships is referred to as the phylogeny of the sequences used in the study, the sequences used form the tips of the tree's branches and the branches connect these tips to their (unobservable) ancestral sequences (Holder and Lewis, 2003; Vandamme, 2003). Since the first investigations into evolutionary relationships using molecular data in the 1960's, there has been an dramatic increase in the understanding of the evolution of genetic material such as DNA and RNA, and this has led to the development of ever increasing methods to recover evolutionary information from sequence data using sophisticated software packages. The early stages of phylogenetics saw three basic assumptions being made namely 1.) sequence evolution was primarily a random event, 2.) calculating the number of differences between sequences was sufficient to estimate the nucleotide divergence between sequences and 3.) all genes evolve at a constant rate among lineages and that this allows evolutionary events to be placed within a time frame (i.e. the strict molecular clock) (Vandamme, 2003).

Table 1.4. Some common terms used in phylogenetics (Li and Graur, 2000, Holder and Lewis, 2003; Drummond *et al.,* 2006)

Branches	The lines that connect units in a phylogenetic tree. They represent the relationship between the units in terms of descent and ancestry.
Branch length	Represents the number of changes that have occurred in a branch.
External nodes	Represents the currently existing taxonomic units in the study, often referred to as Operational Taxonomic Units (OTUs).
Nodes	Represents taxonomic units in a phylogenetic tree e.g. species, populations, individuals or genes.
Phylogenetic tree	Graph illustrating the evolutionary relationships between organisms or gene sequences. Branches connect the tips (sequences) to other adjacent nodes.
Rooted tree	In this tree there is a specific node that is called the root. A unique path leads from this node to another node and the direction of each path corresponds to evolutionary time, where the root is the common ancestor of the OTU's under study or is distantly related to the OTU's in the study.
Scaled branches	The lengths of the branches are indicative of the number of changes.
Topology	The branching pattern of the tree.
True tree	This tree represents the true evolutionary history that can be built with the given number of OTU's.
Unrooted tree	This tree does not define the evolutionary path but only displays the relationship between the OTU's.
Unscaled branches	The length of the branch is not proportional to the number of changes.
Likelihood	The probability of the data when using a specific model and tree hypothesis. The likelihood is a measure of how well the data corresponds with the predictions made by the model and tree hypothesis.
Prior probability	Also referred to as the "prior". The probability of a hypothesis (or parameter value) without reference to the available data. Priors can be derived from first principles, but are often based on general knowledge or previous experiments.
Posterior probability	Product of the prior probability and the likelihood.
Effective sample size (ESS)	The number of independent samples that would correspond to the auto-correlated samples produced by the Markov chain Monte Carlo (MCMC). Provides a measure of whether the chain has been run for an adequate length (i.e. if all ESS values of all continuous parameters are greater than 200).

1.8.1 Multiple sequence alignment

A basic pre-requisite for phylogenetic analysis of any dataset is to carry out a sequence alignment. This enables phylogenetic inference by comparing homologous residues (i.e. residues that descend from a common ancestor), thereby determining rates or patterns of change in the DNA or protein sequences. The best way to achieve this is by aligning the sequences so that the homologous residues are organized in columns in the multiple sequence alignment. If all sequences are the same length this is easy to achieve, but unfortunately insertions and deletions of nucleotides (referred to as indels) are often part of viral evolution, and thereby complicate this alignment process (Higgins, 2003). A scoring system computes the "cost" of each alignment between two nucleotide sequences in order to obtain the optimal alignment. The alignment that achieves the highest "score" then represents the optimal alignment for those two sequences in the study (Higgins, 2003). By counting the number of matches between nucleotides, divided by the number of aligned positions (excluding gaps) within the sequence alignment, a simple measure of sequence identity is calculated. When sequences are sub-optimally aligned it will result in a lower alignment score when compared to a situation where more matches are found between nucleotides, a higher alignment score will be calculated (Higgins, 2003). However, the introduction of excess gaps in order to maximize the alignment score can lead to non-sensical alignments. To prevent this from happening, gap penalties (GP) are introduced that penalize the alignment score. When aligning amino acid sequences a similar process takes place, except that the biochemical properties of the amino acids are also taken into account. Scoring matrixes such as PAM and BLOSSUM allow for the fact that substitutions of amino acids remain relatively conserved to amino acids with similar physical or chemical properties (Higgins, 2003).

When more than two sequences are aligned, progressive alignment algorithms add similar sequences to the existing alignment, followed by more dissimilar sequences (Baldauf, 2003; Higgins, 2003). The process of multiple sequence alignment is typically conducted by the construction of a crude guide tree (usually a neighbour-joining tree) that utilizes calculated similarity scores from a pairwise alignment of all the sequences included in the multiple alignment. By following the order of the relationships between

the sequences as presented in the guide tree, a multiple alignment is built up (Baldauf, 2003).

1.8.2 Tree construction methods

1.8.2.1. Traditional approaches

When constructing phylogenetic trees we find that methods can be distinguished based on two criteria. The first differentiates between the use of discrete character states (discrete character methods) or a distance matrix of pairwise dissimilarities between sequences. The second criteria determines whether the method used clusters sequences (Operational Taxonomic Units-OTUs) in a stepwise manner (stepwise clustering methods) or whether it assumes all theoretical trees that can be constructed from this dataset (exhaustive search methods) (Vandamme, 2003). Table 1.5 provides a summary of the most commonly used phylogenetic methods.

Table 1.5. Phylogenetic analysis methods and the strategies they are based on.

	Exhaustive search	***Stepwise clustering***
Character state	Maximum Parsimony (MP) Maximum Likelihood (ML)	
Distance matrix		UPGMA Neighbour-joining (NJ)

Vandamme, 2003

Stepwise clustering methods

Stepwise clustering constructs a tree by sequentially adding OTU's to a growing tree. It first clusters the most closely related sequences together, which is then considered as a new OTU. The next most closely related sequence is then added to the newly created OTU, and this process is repeated until all OTU's have been clustered (Vandamme, 2003).

Distance methods

Distance methods make use of the evolutionary distances in a distance matrix to construct the tree. A pairwise comparison of all sequences included in the analysis is used to compute the dissimilarity between these sequences (roughly percentage nucleotide difference) (Felsenstein, 1984; Felsenstein, 1988; Vandamme, 2003). The values in the distance matrix (amount of dissimilarity) provide a measure of the evolutionary relatedness between sequences, and can be interpreted and visualized by construction of a phylogenetic tree (Felsenstein, 1988; Vandamme, 2003). When analyzing highly divergent sequences, the divergence between sequences based on the percentage difference can give an underestimation of the true genetic distance between these sequences. This is because consecutive mutations ("multiple hits") could have taken place at the same position within a sequence as it became more divergent from its ancestral sequence (Baldauf, 2003; Strimmer and Von Haeseler, 2003). To compensate for this problem, distance matrix methods take multiple hits into consideration when calculating a distance matrix by applying mathematical models (nucleotide substitution models) that correct the percentage difference between analyzed sequences. These corrected distances are referred to as the "evolutionary distance" and these values are larger than the percentage values calculated by a direct comparison between sequences (Vandamme, 2003).

There are various evolutionary models to choose from, and an inappropriate choice of model can influence the outcome of any phylogenetic analysis. There are two types of model namely codon-based models and standard nucleotide substitution models. Codon-based models make use of information from the genetic code by focusing on the codon instead of individual nucleotides while nucleotide substitution models look at individual nucleotides. Codon-based models are computationally very expensive when compared to standard nucleotide models and as such are not often used (Shapiro *et al.,* 2006). Commonly used nucleotide substitution models include the Jukes-Cantor one-parameter, Kimura two-parameter and the general time reversible with gamma distributed rate heterogeneity and a proportion of invariant sites (GTR+Γ+I) model (Jukes and Cantor, 1969; Kimura, 1980; Strimmer and Von Heaseler, 2003; Shapiro *et*

al., 2006). Specific examples of distance based methods include the unweighted pair group method with arithmetic mean (UPGMA) and the neighbour-joining method.

- **UPGMA**

The UPGMA approach employs a clustering algorithm whereby most-closely related sequences are sequentially added to the tree. A distance matrix is first constructed for a pairwise comparison of all the sequences under study. These OTUs are then clustered in the manner that was described in section 6.2.1.1(Vandamme, 2003).

- **Neighbour-Joining**

Usually a neighbour-joining (NJ) tree is an unrooted tree, where the distance between any two sequences represents the number of mutations that has occurred in the evolutionary process since the divergence from the common ancestor (Saitou and Nei, 1987; Holder and Lewis, 2003; Van de Peer, 2003). The construction of a NJ tree starts by creating a corrected distance matrix of the sequences being analyzed, and then represents the OTUs as a star tree in which one node connects an OTU to every other OTU. The algorithm then calculates the divergence from every OTU to every other OTU by adding the distances between them, and a new matrix is created that summarizes these values. The program then connects the two OTUs with the smallest distance between them with a bifurcating node, and calculates the branch lengths of the newly created OTU, as well as the length of the branches to the other unclustered OTUs. These new distances are again summarized in a new distance matrix. This process is repeated until all sequences have been clustered (Saitou and Nei, 1987; Van de Peer, 2003).

Character state methods

Character state methods use the individual characters of nucleotide or amino acid sequences to construct a phylogenetic tree. Every position in an alignment of nucleotide or amino acid sequences is considered to be a “character”, while the “state” of a character refers to the type of amino acid or nucleotide base at that specific position (Vandamme, 2003). In a multiple alignment, the state of characters within each position is examined separately and independently from each other, and the tree that is constructed best accommodates all this information (Baldauf, 2003; Vandamme, 2003).

Exhaustive search methods

Exhaustive search methods examine every possible tree topology for a given dataset. Probability scores are assigned to constructed trees on the basis of how well the data fits these trees under a specific model of evolution. The best tree is then selected based on an optimality criterion (which is determined by the method of tree construction)(Vandamme, 2003). As the number of sequences in a dataset is increased, the number of possible trees also increases, and thus this method is typically not practical to apply to more than 10 taxa (Vandamme, 2003). Search strategies (i.e. heuristic, branch etc.) have been developed in an attempt to search "tree space" for the optimal tree, without the need to examine every possible tree topology. However, these strategies cannot guarantee that the best tree had been examined during the analysis (Swofford and Sullivan, 2003). Examples of exhaustive search methods are Maximum parsimony (MP) and Maximum likelihood (ML).

- **Maximum Parsimony**

In an attempt to construct a tree that requires the least amount of evolutionary changes to explain the dataset, the maximum parsimony method examines every possible tree topology that can be constructed from a given dataset (Swofford and Sullivan, 2003). The nucleotide sequence data is examined directly, and only characters that contain at least two different states at a specific position where at least two of these occur with a minimum frequency of two (i.e. parsimony informative characters) from four or more aligned sequences are used in the analysis (Bourhy *et al.,* 1995; Li and Graur, 2000).

The ancestral sequence at each character site is inferred separately for a given topology, under the assumption that substitutions can occur in all directions. The least amount of nucleotide or amino acid changes that will explain the entire evolutionary process in computed. This computation is performed for all potentially correct tree topologies, with the tree requiring the least amount of substitutions being chosen as the optimal tree (Li and Graur, 2000; Swofford and Sullivan, 2003).

- **Maximum Likelihood**

Using a likelihood function the maximum likelihood method identifies the optimal tree from all possible tree topologies that can be constructed from a given dataset. Using a

chosen nucleotide substitution model, the program calculates the total likelihood for a specific tree topology by determining the likelihood scores for each column in the sequence alignment and then multiplying these individual column scores with each other. In this method, the tree topology that has the highest likelihood of all possible topologies constructed is chosen as the correct tree, because it maximizes the probability of observing the given data (Li and Graur, 2000; Holder and Lewis, 2003).

1.8.2.1.1 Bootstrap analysis

Bootstrap analysis was first introduced by Felsenstein in 1985. It is a statistical technique that generates confidence limits on phylogenies, where the underlying sample distribution of the dataset being analyzed is unknown (Felsenstein, 1985; Van de Peer, 2003). This method randomly resamples columns from a multiple alignment, until a new dataset equal in length to the original dataset has been constructed, thereby creating pseudoreplicate datasets. For each of these pseudoreplicate datasets, a phylogenetic tree is constructed, and the frequency with which specific clades (present in the original tree) are reproduced in these other trees in computed. The statistical support for the clustering of a specific clade is equal to the proportion at which this clade was reproduced from the total number of pseudoreplicate datasets. Usually 200-2000 resamplings are recommended for an analysis, while bootstrap support values of 75% to 100% are regarded as statistically significant (Baldauf, 2003; Van de Peer, 2003).

1.8.2.2. Bayesian approach

Bayesian analysis has generated some excitement among scientists interested in conducting a phylogenetic analysis since it generates both a tree estimate and measures of uncertainty in the primary analysis. Bayesian phylogenetics is closely related to ML approaches. Bayesian approaches assume that the optimal hypothesis for a given dataset is the one that maximizes the posterior probability (Holder and Lewis, 2003). The likelihood of a given hypothesis multiplied by the prior probability is the posterior probability of that hypothesis. The prior probability of a hypothesis conveys the belief of the researcher before he/she has seen the data. In most cases, the researcher specifies prior probability distributions that he/she considers to be largely uninformative, in order to obtain differences in the posterior probability that are due to differences in

the likelihood of hypotheses (Holder and Lewis, 2003). Bayesian analysis not only provides measures of support faster than bootstrapping, but they also allow the implementation of complex models of evolution. Because the likelihood functions for phylogenetic models are too complex to integrate analytically the Bayesian approach relies on MCMC (Markov chain Monte Carlo) an algorithm that is used to estimate the probability distributions in a wide variety of scenarios. MCMC takes a series of steps to form a conceptual chain. At each step the program proposes a new location in the parameter space as the next link in the chain. Usually this proposed location is similar to the present location because it is generated by the random perturbation of a few of the parameters in the present state of the chain. At this new location the relative posterior probability density is calculated. The move is accepted if the new position has a higher posterior probability density than that of the present location of the chain. This new location becomes the next link in the chain and the cycle is repeated (Holder and Lewis, 2003). When the proposed location has a lower posterior probability density than the current location the move will only be accepted a proportion (*p*) of the time. Proportion *p* is the ratio of the posterior of the proposed new location compared with the posterior of the current location. In essence this implies that short steps downward are accepted often, whereas large downward leaps are discouraged. Should the proposed location be rejected, the current location is added to the chain as the next link so that the last two links in the chain will be identical. The cycle is then repeated. If this procedure is repeated millions of times, the program creates a long chain of locations in parameter space. If the chain tends to stay in regions of high posterior probability, almost all moves are downhill and therefore rarely accepted. To estimate the posterior probability of a region we can use the time that the chain spends in that region of the parameter space. Therefore by creating long chains, this method of estimation can be made randomly accurate (Holder and Lewis, 2003). When looking at a phylogenetic application, the relevant location in the parameter space is not only a description of the tree, but also a specification of all the parameters in the model of sequence evolution that was chosen. During the analysis a chain is created that moves through different models and trees of evolution. When the analysis is complete, the researcher is given an estimate of the probability that any given tree is the “true” evolutionary tree from that given dataset.

However, this probability is dependent on the fact that the model of evolution that was chosen is suitable for the data and that prior distributions on the parameters were realistic, even though it still represents an instinctive measure of the amount of confidence that should be placed in that tree (Holder and Lewis, 2003). Table 1.6 is a summary of all the different methods with brief advantages and disadvantages.

Table 1.6. Comparison of phylogenetic methods.

Method	***Advantages***	***Disadvantages***	***Software***
Neighbour joining	Fast	Information may be lost due to compression of sequences into distances; reliable estimates of pairwise distances can be hard to obtain for divergent sequences.	PAUP MEGA PHYLIP
Parsimony	Fast (more than 100 sequences), robust if branches are short (closely related sequences or dense sampling).	Substantial variation in branch lengths can cause poor performance.	PAUP MEGA PHYLIP
Minimum evolution	Uses models to correct for unseen changes.	When distances are large, distance corrections can break down.	PAUP PHYLIP
Maximum likelihood	The likelihood captures what the data tells us about the phylogeny under a given model.	Can be prohibitively slow (depending on the thoroughness of the search and access to computational resources).	PAUP PHYLIP PAML
Bayesian	Strong connection to maximum likelihood method; might assess support for trees faster than using bootstrapping.	Can be difficult to determine whether the Markov chain Monte Carlo (MCMC) approximation has run for long enough; Prior distribution for parameters must be specified.	MrBayes BEAST

Holder and Lewis, 2003

1.8.3 Molecular clocks

Up to a few years ago, scientists wishing to infer divergence dates assumed a constant rate of evolution throughout the phylogenetic tree i.e. all branches in the tree evolved at the same rate. This is known as a strict molecular clock. However various datasets yielded results that challenged this assumption of clocklike evolution, and it was seen that rate variation did exist between branches. This could mislead not only the date of divergence but also phylogenetic inference (Drummond *et al.,* 2006). These problems led to the strict molecular clock approach being replaced by a model that assumes that every branch has a different rate of evolution. This method has the disadvantage that phylogenies can be inferred, but rates of evolution of divergence times cannot be determined. This is due to the fact that the individual contributions of rate and time to molecular evolution cannot be separated. This unrooted method was widely accepted as an alternative to the strict molecular clock and forms the basis of most modern phylogenetic programs such as PHYLIP and PAUP. In phylogenetics the strict molecular clock and the unrooted-tree models form the two extremities of a field. It has however become more obvious as more sequence data was analysed that both these assumptions are biologically unrealistic, with the true evolutionary process being located somewhere between the two extremes. The Bayesian approach allows rates of evolution to vary across the tree and therefore is an alternative method that is referred to as the relaxed clock (Drummond *et al.,* 2006). BEAST is a software application that employs a Bayesian Marcov chain Monte Carlo (MCMC) method for performing relaxed phylogenies. It is able to co-estimate phylogeny and divergence time under relaxed clock models (Drummond and Rambaut, 2007).

When testing different approaches available in the BEAST package (i.e. strict molecular clock, uncorrelated/autocorrelated relaxed clock with lognormal rate model and uncorrelated/autocorrelated relaxed clock with exponential rate model) using simulated as well as real sequence datasets, it was found that the lognormal rate models performed better than the strict clock or the exponential rate model relaxed clock. It was found that even if data is clocklike, the uncorrelated relaxed clock models perform well, where if data is not clocklike and a strict clock was enforced, very poor results were obtained. Drummond *et al.* (2006) advise the use of the uncorrelated relaxed clock with

lognormal rate model because it has comparable accuracy with the exponential rate model, but it gives much smaller higher posterior densities (HPD). Also due to the lognormal distribution of this model, it can better accommodate data that is close to being clocklike. Recent studies of viral datasets using this new approach include studies on dengue-4 virus (Bennett *et al.,* 2003; Drummond *et al.,* 2006), human influenza A (Rambaut *et al.,* 2008) and rabies virus (Davis *et al.,* 2007). This again indicates the suitability of this approach to various different types of data. The study of the mongoose and canid variant of the rabies virus in southern Africa by Davis *et al*. (2007) focused on the GL intergenic region. Davis and colleagues estimate the mean age of the mongoose variant to be 73 years whereas the canid variant is approximated to be 30 years old (Davis *et al.,* 2007). In the past few years, analysis of gene sequences and estimation of phylogenies has become established as part of most molecular studies. It is clear that the MCMC-based approach is allowing scientists to attempt to answer previous difficult questions. It is however too early to tell whether Bayesian approaches will transform tree estimation, but these new techniques are bringing us closer to understanding the tree of life and molecular genetics (Drummond *et al.,* 2006).

1.9. Aims of study

This study focused on two lyssaviruses that are exclusive to the African continent namely Mokola virus and rabies virus mongoose biotype. In the past 15 years serological and phylogenetic analysis has shown the presence of two distinct groups (i.e. canid biotype and mongoose biotype) of rabies virus circulating in southern Africa (Foggin, 1988; Nel *et al.,* 1993; Von Teichman, 1995; Sabeta, 2002; Nel *et al.,* 2005). The variable G-L intergenic region was employed in most of these phylogenetic analyses. Due to the absence of selection pressure, the G-L region displays high genetic variability and subsequently has been advocated as an ideal area to distinguish between closely related isolates. However, the nucleoprotein and the glycoprotein are both more conserved regions of the lyssavirus genome, and therefore have historically been proposed as ideal targets to distinguish between isolates that are suspected to share a distant relationship (Tordo *et al.,* 1986a).

A representative panel of isolates was selected based on previous phylogenetic groupings (Nel *et al.,* 2005). Isolates were chosen to represent all 5 geographic groups as well as being isolated over the past 20 years. In this study the emphasis was placed on using two full gene sequences namely the nucleoprotein and glycoprotein genes to characterize these isolates phylogenetically. As very few G-L intergenic region sequences of rabies virus isolates from other regions of the world are available in the public domain, utilization of these 2 gene sequences allowed the mongoose biotype to be placed within a worldwide phylogenetic perspective. Furthermore, this study aimed at placing this information in a historical context by determining the age of this variant using a molecular clocking technique. This study also highlighted the caution that should be applied when interpreting molecular clock data.

Before the discovery of WCBV, it was shown that Mokola virus is genetically the most distant from rabies virus (genotype 1) when using serological studies and analysis of specific genomic nucleotide sequences (Johnson *et al.,* 2002; Nel *et al.,* 2003). Furthermore it has been hypothesized that Mokola virus isolates display approximately the same amount of genetic diversity among a small number of isolates from a single continent as is displayed among genotype 1 (rabies) isolates from around the world (Badrane *et al.,* 2001), and even more notable is that these Mokola virus isolates cluster according to their geographical region (Bingham *et al.,* 2001). This diversity among the Mokola virus isolates suggests an even greater diversity in nature (Badrane *et al.,* 2001). Since these findings have been published, new isolates of Mokola virus have been made and more sequences of existing Mokola virus isolates have become available in the public domain. Thus the aim of this study was to include a larger number of Mokola virus isolates and reassess the phylogeny and genetic diversity of this genotype using the full length nucleoprotein gene sequence.

Specific objectives

- Sequence and phylogenetically analyze the full length nucleoprotein and glycoprotein gene sequences of 27 representative mongoose biotype isolates.

- Apply molecular clock to this data to place the time of the origin mongoose biotype into historical context.
- Sequence and phylogenetically analyze the full length nucleoprotein gene of all Mokola virus isolates that are available in South Africa at the Agricultural Research Council Onderstepoort Veterinary Institute archives.

Chapter 2: Evolutionary study of mongoose rabies in southern Africa

2.1. Introduction

In Africa two distinct variants of rabies virus is found namely the canid biotype which circulates primarily in canids, and the mongoose biotype, which was previously isolated mostly from members of the Viverridae, but is now more commonly isolated from members of the *Herpestidae* family, with the yellow mongoose and suricate being the principle hosts. The Viverridae family comprises genets and civets. These animals occur all over Africa, with some members occurring in Madagascar and Southeast Asia (Estes, 1995). Mongooses belong to the family *Herpestidae*, which is divided into two subfamilies namely the Herpestinae occurring on the African continent and Eurasia and the Galidiinae that occur in Madagascar (Veron *et al*., 2004).

As early as 1940, Snyman already observed differences in host pathogenicity and epidemiology between the canid and mongoose rabies virus biotype. He described the lack of dog-to-dog transmission when these animals are infected with the mongoose biotype virus (Snyman, 1940). Later various studies described the antigenic and genetic variation found among the mongoose viruses (King *et al.,* 1993a; Nel *et al.,* 1993; Smith *et al.,* 1993; Von Teichman *et al*., 1995). King *et al*. (1993a) and Smith *et al*. (1993) observed two distinctly different reaction patters when the two biotype viruses were tested against the same panel of monoclonal antibodies. They also recorded spillover of the two biotypes into different hosts (King *et al*., 1993a). Taylor (1993) demonstrated that the localized outbreaks of rabies in mongooses can be related to the well-developed social structure of these animals that maintains specific groups of animals in restricted territories. In studying the pathogenesis of the two distinct biotypes of rabies virus in southern Africa, Chapparo and Esterhuysen (1993) showed that mongooses are clearly more susceptible to infection with the mongoose biotype virus that infection with the canid biotype virus, and possibly transmit the mongoose biotype virus more efficiently due to higher viral loads present in the saliva of mongooses infected with the mongoose biotype when compared to animals infected with the canid biotype. Nel *et al.* (1993) and Von Teichman *et al.* (1995) used sequence information from different areas of the RABV genome to study the phylogenetic relationship between rabies viruses from South Africa. In both these studies the mongoose biotype isolates formed a clear separate phylogenetic group distinguishable from the canid biotype isolates. However

this analysis also demonstrated occasional spillover from mongoose to canid and vice versa (Nel *et al.,* 1993). In 1995, when analyzing rabies virus isolates from across the world, Kissi *et al.* designated viruses belonging to the mongoose biotype as Africa3 viruses. This nomenclature was again repeated in the recent article by Bourhy *et al.* (2008) and this nomenclature was followed in this study.

Further phylogenetic analysis reflected the geographic origin of the virus isolates, with tree topology corresponding to five distinct geographical regions of South Africa (Nel *et al.,* 2005). All these findings supported the view that mongoose rabies has been indigenous to southern Africa, has evolved differently from the canine (cosmopolitan) strain of rabies in Africa and has been present in the region for a number of centuries. When reviewing historical records it seems plausible that this virus variant first infected members of the Viverridae (genets and civets) and later became established in the herpestids. Because previous phylogenetic studies (Nel *et al*., 2005; Davis *et al*., 2007) focused on the GL intergenic region, this study will utilize sequence data of the full length nucleoprotein and glycoprotein genes. This would allow comparative phylogenetic analysis to determine whether these full length genes provide sufficient phylogenic resolution to demonstrate the phylogeny of this virus biotype. Because the exact time of origin of this virus variant was not known, this study further aimed at determining the age of the most recent common ancestor (mrca) of the mongoose RABV biotype to attempt at placing the origin of the mongoose RABV biotype into historical context.

2.2. Materials and Methods

2.2.1 Viruses and virus isolation

In this study 27 rabies virus isolates obtained from various animals from different regions across Zimbabwe and South Africa were analyzed. These isolates were obtained from the Onderstepoort Veterinary Institute, Pretoria and were chosen to represent the 5 sub-groups previously defined (Nel *et al.,* 2005) as well as span approximately the past 20 years (1980-2006). The year of isolation, specific host species and geographical origins of the viruses are shown in Table 2.1.

Table 2.1. Rabies virus mongoose biotype isolates included in this study.

Virus	*Lab ref nr*	*Species*	*Country*	*Year of isolation*	*Locality of isolation*	*GenBank accession number*
1	Sn0080	*Genetta genetta*	Namibia	1980	Mariental	(N) FJ392392 (G) FJ465409
2	669/90	*Cynictis penicillata*	South Africa	1990	Grootgewaagd	(N) FJ392385 (G) FJ465402
3	420/90	*Cynictis penicillata*	South Africa	1990	Wolmaranstad	(N) FJ392383 (G) FJ465401
4	113/91	*Atilax paludinosus*	South Africa	1991	Beaufort West	(N) FJ392383 (G) FJ465390
5	19/92	*Suricata suricatta*	South Africa	1992	Cradock	(N) FJ392372 (G) FJ465384
6	878/92	*Atilax paludinosus*	South Africa	1992	Harrismith	(N) FJ392366 (G) FJ465406
7	926/93	*Suricata suricatta*	South Africa	1993	Carolina	(N) FJ392389 (G) FJ465407
8	485/94	*Suricata suricatta*	South Africa	1994	Standerton	(N) FJ392384
9	22107	*Galerella sanguinea*	Zimbabwe	1994	Rusape	(N) FJ392391 (G) FJ465408
10	767/95	*Cynictis penicillata*	South Africa	1995	Kroonstad	(N) FJ392388 (G) FJ465405
11	364/96	*Cynictis penicillata*	South Africa	1996	Uitenhage	(N) FJ392379 (G) FJ465397
12	759/96	Feline	South Africa	1996	Belfast	(N) FJ392387 (G) FJ465404
13	211/98	Canine	South Africa	1998	Kuruman	(N) FJ392374 (G) FJ465392
14	221/98	*Suricata suricatta*	South Africa	1998	Venterstad	(N) FJ392375 (G) FJ465393
15	718/98	*Genetta genetta*	South Africa	1998	Gordonia	(N) FJ392386 (G) FJ465403
16	279/99	*Cynictis penicillata*	South Africa	1999	Potchefstroom	(N) FJ392377 (G) FJ465395
17	344/99	*Cynictis penicillata*	South Africa	1999	Bethlehem	(N) FJ392378 (G) FJ465396
18	28/00	*Galerella sanguinea*	South Africa	2000	Hoopstad	(N) FJ392369 (G) FJ465387
19	30/00	Bovine	South Africa	2000	Brandfort	(N) FJ392370 (G) FJ465388
20	228/01	Ovine	South Africa	2001	Bultfontein	(N) FJ392376 (G) FJ465394
21	22/01	Feline	South Africa	2001	Kroonstad	(N) FJ392367 (G) FJ465385
22	23/01	*Suricata suricatta*	South Africa	2001	Bothaville	(N) FJ392368 (G) FJ465386
23	389/02	*Felis nigripes*	South Africa	2002	Gordonia	(N) FJ392382 (G) FJ465400
24	32/02	*Cynictis penicillata*	South Africa	2002	Harrismith	(N) FJ392371 (G) FJ465389
25	155/03	*Galerella sanguinea*	South Africa	2003	Kroonstad	(N) FJ392373 (G) FJ465391
26	381/06	*Cynictis penicillata*	South Africa	2006	Hoopstad	(N) FJ392380 (G) FJ465398
27	385/06	Canine	South Africa	2006	Bethlehem	(N) FJ392381 (G) FJ465399

Where insufficient amounts of the original brain material were available, isolates were passaged in suckling mice (Koprowski, 1996). Lyophilized brain material was reconstituted in sterile PBS (0.138 *M* NaCl; 0.0027 *M* KCl, 0.43m*M* $Na_2HPO_4{\cdot}2H_2O$, 0.14m*M* KH_4PO_4, Sigma-Aldrich). Two to three day old suckling mice received 30µl of the reconstituted material intracranially. Animals were monitored and collected upon death, and brain material was removed aseptically. Infected whole brains were stored at -70℃ after confirmation of lyssavirus infection through FAT (Dean *et al.,* 1996). The FAT is the most widely used method to diagnose rabies infection in humans and animals. In brief, tissue smears were fixed in acetone, dried at 37℃ and then incubated for 30 minutes at 37℃, with a conjugate that is immunologically bound to rabies antinucleocapsid antibodies. A polyclonal fluorescein isothiocyanate conjugated immunoglobulin (Onderstepoort Veterinary Institute, Rabies Unit, South Africa) that is capable of detecting all *Lyssavirus* genotypes was used at a 1:300 dilution. Evans Blue counterstain (0.5% in PBS (0.01 M phosphate buffer, pH 7.4; 0.138 *M* NaCl; 0.0027 *M* KCl, Sigma-Aldrich) was added to the working dilution conjugate. These smears were then washed in phosphate buffered saline (pH7.2), air-dried, mounted with 20% glycerol solution (0.05 M Tris-buffered saline, pH 9.0 with 20% glycerol, Sigma-Aldrich) and examined under a fluorescent microscope. The presence of fluorescence indicates the presence of lyssavirus particles. Isolates that displayed fluorescence were subjected to antigenic typing using the FAT with a panel of sixteen antinucleocapsid monoclonal antibodies (N-MAbs) (Dr. Alex Wandeler, Centre of Expertise for Rabies, Canadian Food Inspection Agency, Nepean, Ontario, Canada) to further distinguish these isolates as belonging to the mongoose or canid biotype.

2.2.2 RNA extraction

Total RNA was extracted using TRIzol® reagent (Invitrogen) according to the manufacturer's instructions. Briefly, 50-100mg of brain tissue and 1ml of TRIzol® reagent were homogenized and incubated at room temperature for 5 minutes. Thereafter 200µl of chloroform was added to the homogenate, shaken vigorously for 15 seconds and incubated at room temperature for 2-3 minutes. This mixture was centrifuged at 10 000g for 15 minutes to allow phase separation and the aqueous phase was transferred to a sterile 1.5ml eppendorf tube. The total RNA was precipitated by

addition of 500µl isopropanol to the supernatant and incubated at room temperature for 10 minutes. RNA was pelleted by centrifugation at 10 000g for 10 minutes, the pellet was subsequently washed with 70% ethanol and followed by centrifugation at 7 500g. The supernatant was removed and the pellet was air-dried and dissolved in 50µl nuclease-free water (Promega). Ten microlitres of total RNA was used in the reverse transcription (RT) reaction.

2.2.3 Primers for RT-PCR

All oligonucleotides used for RT-PCR and sequencing were produced by Inqaba Biotechnical Industries (Pty) Ltd. (S.A.) and used without further purification. PCR amplification and sequencing were performed with the primer sets indicated in Table 2.2. The symbols (+) and (-) refer to messenger and genomic sense respectively.

Table 2.2. Oligonucleotide primers used for cDNA synthesis, PCR and sequencing of the nucleoprotein and glycoprotein genes of rabies virus isolates. Positions indicated relative to Pasteur virus genome (M13215).

Name	Sequence	Application	Target region
001lys (+) (Markotter *et al.,* 2006a)	5'-ACGCTTAACGAMAAA-3'	RT-PCR, sequencing	N(3' noncoding region -70 to -57)
VivN rev (-)	5'-GATGTCTGGCGTCTTGCC-3'	PCR, sequencing	N(1666-1684)
VivNFseq (+)	5'-GAAACCCGAAGCCCTGAAGC-3'	Sequencing	N(1262-1281)
VivMF (+)	5'-GATTCCTCTCTGCTTCTAG-3'	RT-PCR, sequencing	G(3080-3099)
L(-) (Sacramento *et al.,* 1991)	5'-CAAAGGAGAGTTGAGATTGTAGTC-3'	PCR, sequencing	G(5520-5543)
VivGFseq (+)	5'-GGATTCGTGGATGAAAGAGGC-3'	Sequencing	G(3996-4016)

2.2.4 cDNA synthesis

For complementary DNA (cDNA) synthesis of the N and G genes, 10µl total RNA and 20pmol of the positive sense primers, 001lys (N-gene) and VivMF (G-gene) (Table 2.2)

were denatured at 70℃ for 5 minutes. The reaction mixtures were then cooled on ice and the RNA was reverse transcribed at 42℃ for 60 minutes in a 20µl reaction mixture containing 12.5U Avian Myeloblastosis Virus reverse transcriptase (AMV-RT) (Roche Applied Science, Germany), 4µl 5X incubation buffer (250nM Tris-HCl, 40mM $MgCl_2$, 150mM KCL, 5mM dithiothreithol) (Roche Applied Science, Germany), 20mM of deoxynucleotide triphosphates (dNTPs) and 40U of RNase inhibitor (Roche Applied Science, Germany). For amplification during polymerase chain reaction (PCR) 5µl of cDNA was used.

2.2.5 Polymerase Chain Reaction (PCR)

Amplification was performed in a DNA thermal cycler (Perkin-Elmer Geneamp PCR 2400). Five µl of cDNA was added to a reaction mixture comprising of 10×reaction buffer (5µl) (50mM KCl, 10mM Tris-HCL, 0.1%Triton X-100), 50mM $MgCl_2$ solution (1.5µl), 2.5mM each of four dNTPS (Promega), 40pmol each of the primers (001lys and VivNrev, or VivMF and L(-)), 2.5U of BioTaq DNA polymerase (Bioline, U.K) and then made up to a final volume of 50µl with RNase/DNase-free water (Promega). The amplification involved an initial denaturation at 94℃ for 2 minutes, followed by 30 cycles of denaturation at 94℃, primer annealing at 37℃ for 30s (42℃ for 45s) and primer extension at 72℃ for 90s (2.5min). (Values in brackets indicate conditions for glycoprotein PCR). This was followed by a final extension of 7 min at 72℃.

2.2.6 Agarose gel electrophoresis

Once amplification was completed products were analyzed on 1%(w/v) agarose gels, prepared in 1X sodium boric acid (SB) electrophoresis buffer (5mM disodium borate decahydrate, adjusted to pH8.5 with boric acid). The PCR amplicons were resolved against a Lambda DNA/EcoRI+HindIII molecular marker (Promega). Five µl of the samples were loaded in loading dye (40% sucrose. 0.25% bromophenol blue). Gels were run at 120V in a horizontal gel tank system using a Biorad Wide Mini Sub electrophoresis cell. The gel was then stained in a 0.5µg/ml ethidium bromide solution and the bands were visualized using a UV transilluminator.

2.2.7 Purification and sequencing

PCR amplicons were gel-purified using the Wizard® SV Gel and PCR Clean-up system (Promega) according to the manufacturer's instructions. Briefly the excised agarose fragment was mixed with Membrane binding solution at a ratio of 10µl solution per 10mg of gel slice, and then melted at 60℃. The melted gel mixture was added to the Minicolumn assembly and incubated for 1 minute at room temperature, where after it was centrifuged at 12 000g for 1 minute. The column was washed with 700µl Membrane wash solution (to which ethanol has been added previously) and centrifuged for 1 minute at 12 000g. The wash step was repeated with 500µl Membrane wash solution followed by centrifugation at 12000g for 5 minutes. After this wash step the column assembly was again centrifuged at 12 000 g for 1 minute to allow evaporation of any residual ethanol. The DNA was then eluted from the column by addition of 50µl nuclease-free water, incubation for 1 minute and centrifugation for one minute at 12 000g. One µl of the eluted DNA was quantified using a 100bp DNA molecular weight marker (Fermentas) as reference on a 1% agarose gel before sequencing.

Automated sequencing was performed using the Bigdye Terminator 3.1 system (PE Applied Biosystems). A reaction mixture containing Terminator Reaction ready pre-mix (2µl), 3.2 picomoles (pmole) of the primer and 100ng of purified PCR product was made up to 20µl with nuclease-free water. These reactions were cycled in a Perkin Elmer Geneamp 2400, using the following cycle conditions: initial denaturation at 94℃ for 1 minute, followed by 25 cycles of 96℃ for 10s, 50℃ for 5 s and 60℃ for 4 minutes. The EDTA/NaOAc/EtOH method was used according to the BigDye Terminator v3.1 cycle sequencing protocol (Applied Biosystems, 2002). In brief, for each 10µl reaction, 1µl 125mM EDTA, 1µl 3M sodium acetate and 25µl of 100% ethanol were added to the reaction mixture. Tubes were vortexed and incubated at room temperature for 15 minutes. Tubes were subsequently centrifuged at 12 000g for 30 minutes and the supernatant was removed, 100µl 70% ethanol was added and tubes were again centrifuged at maximum speed for 15 minutes and supernatant was removed thereafter. The DNA pellets were air dried at room temperature for 20 minutes and stored at -20℃. The precipitated reactions were submitted to the sequencing facility of the Faculty of

Natural and Agricultural Sciences, University of Pretoria. These reactions were then analyzed on an ABI 3100 DNA sequencer (PE Applied Biosystems).

2.2.8 Phylogenetic analysis

Sequences were edited and consensus sequences assembled in Vector NTI 9.1.0 (Invitrogen Corporation, 2004) using the ContigExpress application. To compile datasets that are representative of all gt1 variants datasets from Kissi *et al.,* 1995 (N-gene) and Holmes *et al.,* 2002 (G-gene) were used (Table 1, appendix). Both these studies aimed at providing a comprehensive view of the phylogeny of all genotypes of the *Lyssavirus* genus. This resulted in two final datasets containing 98 sequences (N gene, 1353 nt in length) and 95 sequences (G gene, 1575 nt in length) respectively.

Alignments were constructed using the BioEdit sequence alignment editor v.7.0 (Hall, 1999) using the ClustalW subroutine (Thompson *et al.,* 1994; Hall, 1999). Calculation of genetic distances and construction of phylogenetic trees was carried out using MEGA3.1 software (Kumar *et al.,* 2004). Genetic distances between sequences were calculated using the p-distance model, and based on these distances a Neighbour-joining tree (NJ) was constructed (Saitou and Nei, 1987). The branching order of the trees was evaluated by bootstrap analysis of a 1000 pseudo-replicate datasets, with a random seed generator number of 64238 (Felsenstein, 1985). In general, bootstrap values of more than 70% were regarded as providing evidence for a phylogenetic grouping. Graphical output of phylogenetic trees was obtained using MEGA3.1 tree explorer (Kumar *et al.,* 2004).

2.2.9. Molecular clock analysis

Alignments generated in BioEdit were converted to Nexus format to be used as input files in Beauti (Bayesian Evolutionary Analysis Utility v. 1.4.6) (Drummond and Rambaut, 2007) which forms part of the BEAST software package. In all cases the general time reversible model with gamma distribution and proportion of invariable sites (GTR+G+Γ) was chosen as the model of nucleotide substitution. This model was selected as the best fit model by subjecting datasets to Modeltest v.3.7 (Posada and Crandall, 1998) in the PAUP software package (Swofford, 2003). This model was also

suggested as the best fit for RNA viruses (J.W.Tang and E.C. Holmes, personal communication, Drummond *et al*., 2006). Parameter values were estimated from the datasets, and clocks were calibrated using the dates of isolation of the isolates used in the datasets. Population histories were constructed using the Bayesian skyline plot (Drummond *et al.,* 2002) and the relaxed (uncorrelated lognormal) molecular clock was chosen as demographic model. This allowed estimation of nucleotide substitution rates for both genes as wells as an estimation of the age of the most recent common ancestor (mrca). Uncertainty in the data for each estimated parameter is reflected in the 95% high probability density (HPD) values. MCMC chains were run for 10^7 states to ensure adequate ESS (>100) for all parameters and runs were performed in duplicate. The Beast output files were analysed in Tracer, runs were combined using Logcombiner v1.4.6 (Drummond and Rambaut, 2007) to generate maximum clade credibility (MCC) trees, and trees were generated in Figtree v1.1.2 (Rambaut, 2008).

2.3 Results

2.3.1 Virus isolates, RNA extraction, cDNA synthesis and PCR

RNA extraction, cDNA synthesis and PCRs were performed as described in section 2.2 to 2.5. The 001lys and VIVNrev(-) primer pair (nucleoprotein) yielded a product of approx. 1600bp and the VivMF and L(-) primer set (glycoprotein) yielded a product of approx. 2.4kbp.

2.3.2 Purification of PCR amplicons and nucleotide sequence determination

The purification of PCR amplicons and nucleotide sequencing was conducted as described in section 2.6. Sequence files were manually inspected using the ContigExpress application in VectorNTI 9.1.0 (Invitrogen Corporation, 2004) in order to confirm that bases were accurately called and to correct bases where necessary. Nucleotide sequencing yielded the full length nucleoprotein and glycoprotein genes of all the selected isolates. The obtained sequences were trimmed using BioEdit software v 7.0 (Hall, 1999) to yield a final sequence fragment containing the open reading frame of both the nucleoprotein (1353nt) and glycoprotein (1575nt) genes respectively. All sequences generated in this study were submitted to GenBank and assigned accession numbers (Table 2.2).

2.3.3 Phylogenetic analysis

Phylogenetic trees were constructed as described in section 2.7. Trees constructed using the NJ method utilizing the complete gene sequences of the N and G gene respectively both displayed similar topology (Figure 2.1).

2.3.3.1 Phylogeny of mongoose viruses compared to other representative gt1 virus isolates from around the globe.

In order to determine the relationship between the mongoose biotype isolates and rabies virus isolates from across the globe, representative isolates were selected as described in section 2.7. A neighbor-joining (NJ) tree was constructed based on the alignment of the full N and G gene sequences respectively (Figure 2.1A and 2.1B). The analysis included representative isolates from all the recognized sub-groups of gt1 viruses including raccoon variant, bat variant (North and South America), skunk variant, fox variant and canid isolates from all continents where rabies virus gt1 is present (Table 1, appendix). The NJ method yielded high bootstrap support (≥79%) for all the major clusters found within gt1 viruses. Clusters defined on the phylogenetic trees are named in accordance with groups designated by Bourhy *et al.* (2008). The same tree topology was obtained when analyzing the full length nucleoprotein gene and full length glycoprotein gene sequences respectively. The 27 isolates of the mongoose biotype that were sequenced in this study grouped into the same cluster on the same branch as the viruses of the cosmopolitan group and arctic-related viruses, but still separate from any other cluster of gt1 viruses. It is interesting to note that in comparison to the nucleoprotein gene sequence tree, the glycoprotein gene sequence tree yielded a much lower bootstrap support value (56% vs. 92%) for the independent grouping of the mongoose biotype isolates.

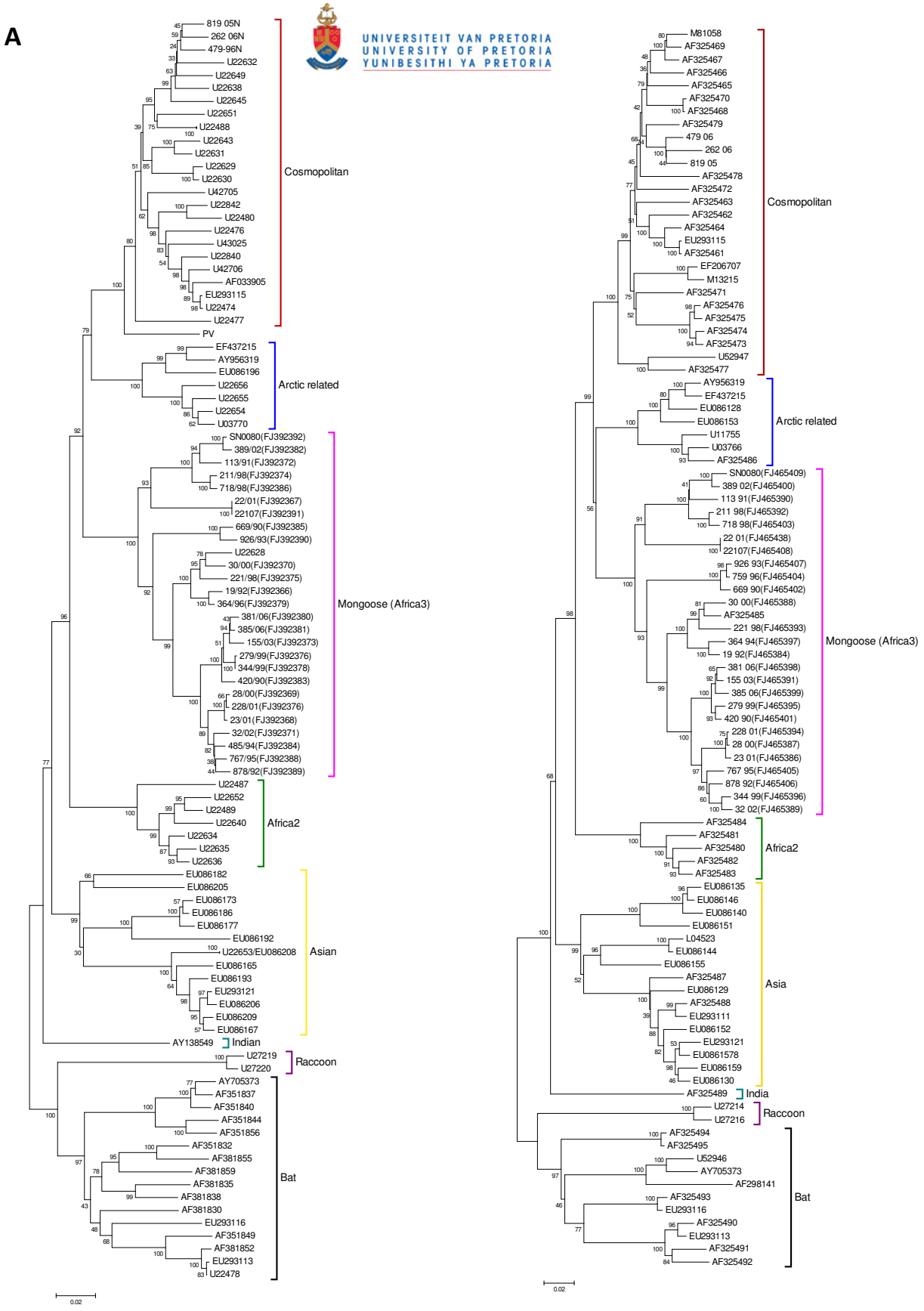

Figure 2.1 Neighbor-joining tree constructed from selected isolates of all subgroups of gt1 lyssaviruses using the full-length nucleoprotein gene (A) and glycoprotein gene (B) sequence. GenBank accession numbers are indicated in the figure. Bootstrap values are indicated at the nodes and branch lengths are drawn to scale.

When performing a pairwise comparison using the p-distance method in Mega3.1 (Kumar *et al.,* 2004) it was found that the mongoose biotype isolates displayed an average nucleotide sequence identity of 93.3% for the nucleoprotein gene and an average nucleotide sequence identity of 92% for the glycoprotein gene. The higher sequence identity values for the nucleoprotein gene dataset are due to the more conserved nature of this gene when compared to the glycoprotein gene. Distance matrixes are displayed in Figure1 and Figure2 (appendix). Results of relationships between gt1 mongoose biotype isolates and other gt1 rabies virus isolates are shown in Table 2.3.

Table 2.3 Pairwise comparison values for all gt1 isolates calculated using the p-distance method (Kumar *et al.,* 2004)

Gene analysed	***Average nucleotide sequence homology within mongoose biotype***	***Individual isolates displaying least amount of nucleotide sequence identity (most divergent of the isolates used in this study, when comparing the mongoose biotype to other gt 1 isolates.)***	***Individual isolates displaying largest amount of sequence identity (least divergent of the isolates used in this study, when comparing the mongoose biotype to other gt 1 isolates)***	***Average nucleotide sequence identity displayed by other gt1 isolates(excluding mongoose biotype)***
Nucleoprotein gene	93.3%	279/99 vs. NY771 (81.9%)[1]	364/96 vs. 9215HON (89.5%)[2]	87.4%
Glycoprotein gene	92%	1500AFS vs USA7-BT (77.3%)[3]	211/98 vs 9147FRA (88.3%)[4]	85.5%

[1] Isolate 279/99 is a gt1 mongoose biotype isolate from South Africa while isolate NY771 originates from a raccoon in the USA..

[2] Isolate 364/96 is a gt1 mongoose biotype isolate from South Africa and isolate 9215HON originated from a human in Hungary.

[3] Isolate 1500AFS is a gt1 mongoose biotype isolate from South Africa while isolate USA7-BT is a isolate from a bat (*Myotis* sp.) in the USA.

[4] Isolate 211/98 is a gt1 mongoose biotype isolate from South Africa while isolate 9147FRA is a fox isolate from France.

2.3.3.2 Phylogeny of mongoose biotype isolates.

Figure 2.2 shows the geographic distribution of the mongoose biotype isolates that were utilized in this study.

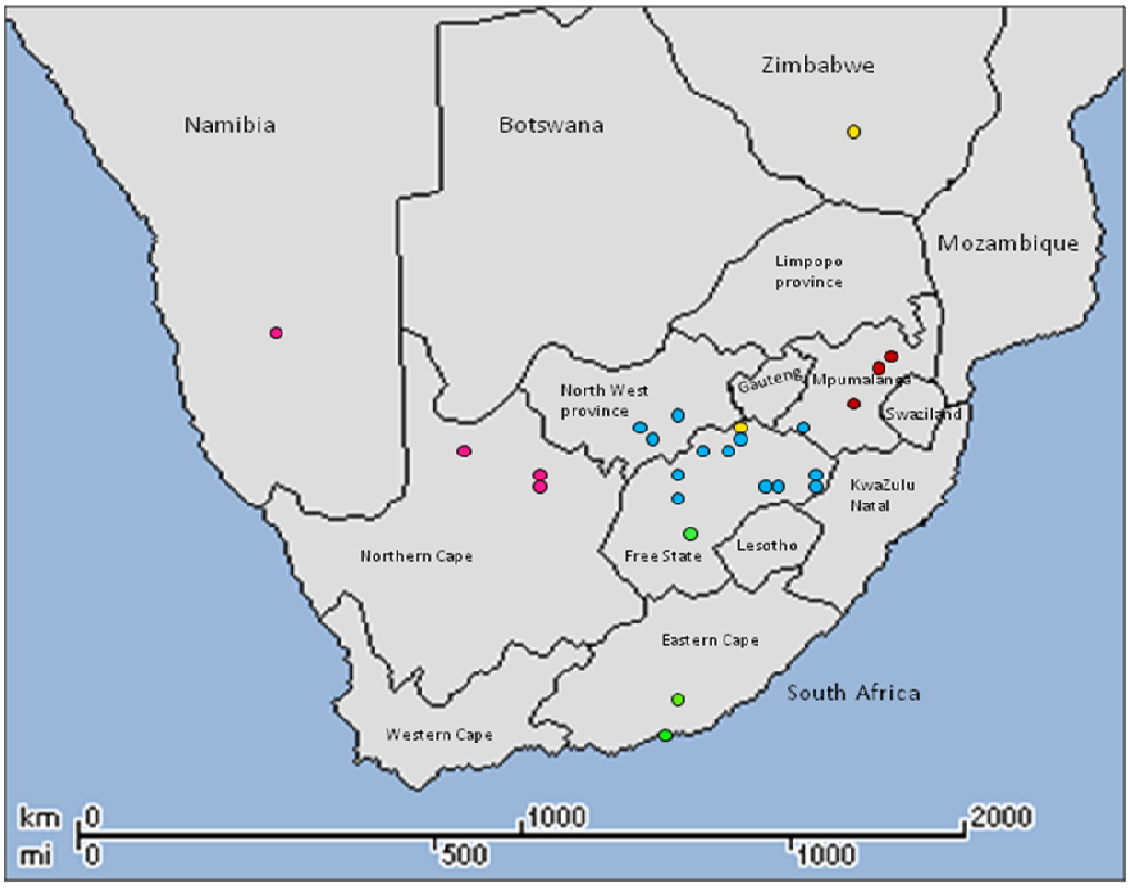

Figure 2.2 Map of southern Africa (http://pidwa.co.za/image/southern_africa) showing distribution of mongoose biotype sub-groups used in this study. Coloured dots represent approximate locality of isolation for isolates. Colours of dots correspond to colours indicating 5 previously defined sub-groups (Nel *et al.,* 2005) in Figures 2.3A and 2.3B.

When analyzing the phylogeny of the mongoose biotype isolates it is clear that both the nucleoprotein (Figure 2.3A) and glycoprotein (Figure2.3B) data supports the clustering of viral isolates according to geographic location (Figure 2.2) with group 1 isolates originating from Zimbabwe, group 2 isolates originating from the north-eastern side of South Africa, group 3 and 4 isolates forming independent cycles in the central plateau of South Africa and group 5 isolates occurring in the north-western region of the Cape.

The NJ method indicated high bootstrap support (99-100%) for all the major clusters representing the 5 subgroups of the mongoose biotype. This grouping corresponds to the geographic grouping into 5 main sub-groups that had been previously identified by Nel *et al.,* (2005) when utilizing the G-L intergenic region in a phylogenetic analysis. Group 1 comprised of only 2 isolates in this study and these isolates do not share close geographical locations. Isolate 22/01 was isolated from a feline in the Free State province, thus it is possible that this animal was imported from Zimbabwe before displaying signs of disease. Another more feasible possibility is that an error occurred during the labeling of lyophilized brain material stocks at Onderstepoort Veterinary Institute as the two isolates have very similar laboratory reference numbers. The fact that these 2 isolates are a 100% identical at nucleotide sequence level seems to indicate a labeling error that would explain the strange grouping of these 2 isolates.

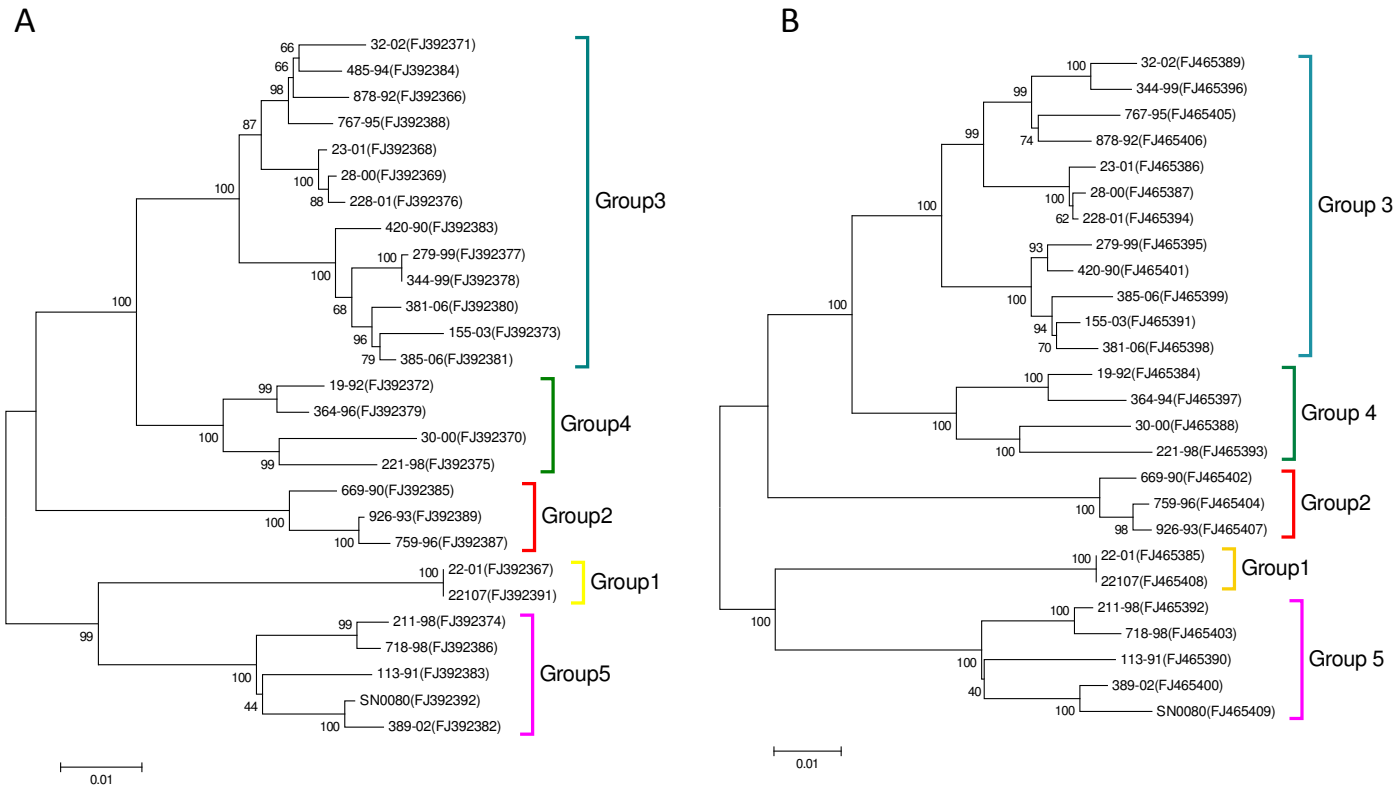

Figure 2.3. Neighbor-joining tree based on the 1353 nt of the full length nucleoprotein (A) and 1575 nt of the full length glycoprotein gene (B) of the 27 mongoose biotype isolates. Bootstrap values are indicated at the nodes and branch lengths are drawn to scale. Bracket colours and group numbers correspond to geographic locations indicated in Figure 2.2.

2.3.4 Molecular clock analysis

The molecular clock analysis was conducted as described in section 2.7.1. It was confirmed that sufficient sampling of the MCMC chain had taken place with all effective sample size values (ESS) for important parameters above 100. Using the date of isolation of all the sequences in the nucleoprotein gene dataset (Table1, appendix) (assuming an uncorrelated relaxed lognormal molecular clock and a constant population size) the rate of nucleotide substitution calculated in substitutions/site/year was estimated to be $1.363E^{-4}$ to $3.758E^{-4}$ (mean rate $2.495E^{-4}$). This is in concordance with previously published nucleotide substitution rate estimates (Badrane *et al.,* 2001; Badrane and Tordo, 2001; Holmes *et al.,* 2002). In this study it was determined that the mongoose variant is approx.135 to 360 (mean 229) years old when analyzing the nucleoprotein dataset. When conducting the analysis utilizing the glycoprotein dataset the rate of nucleotide substitution calculated in mutations/site/year was estimated to be $5.179E^{-4}$ to $8.514E^{-4}$ (mean rate $6.875E^{-4}$). It was estimated that the mongoose variant is approx 119 to 202 (mean 159) years old. Figure 2.4 and 2.5 display maximum clade credibility trees constructed using the output generated by Beast during the molecular clock analysis. Both trees display the same grouping as was observed in the NJ trees (Figure 2.1) with the mongoose biotype clade again sharing an ancestral node with the cosmopolitan clade.

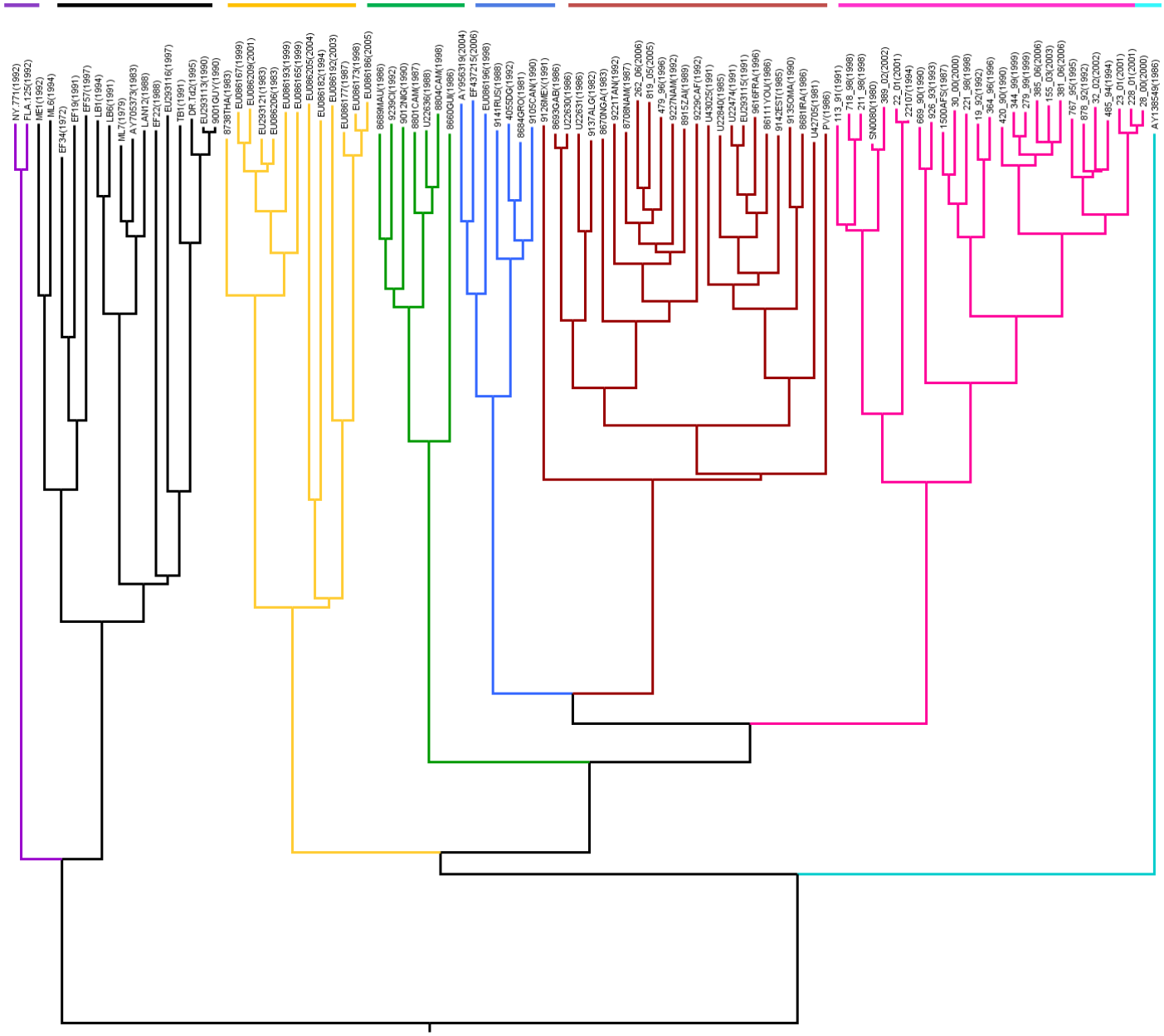

Figure 2.4 Maximum clade credibility tree constructed from a relaxed uncorrelated log-normal molecular clock using the nucleoprotein gene dataset, with branch lengths scaled to reflect evolutionary time in years.

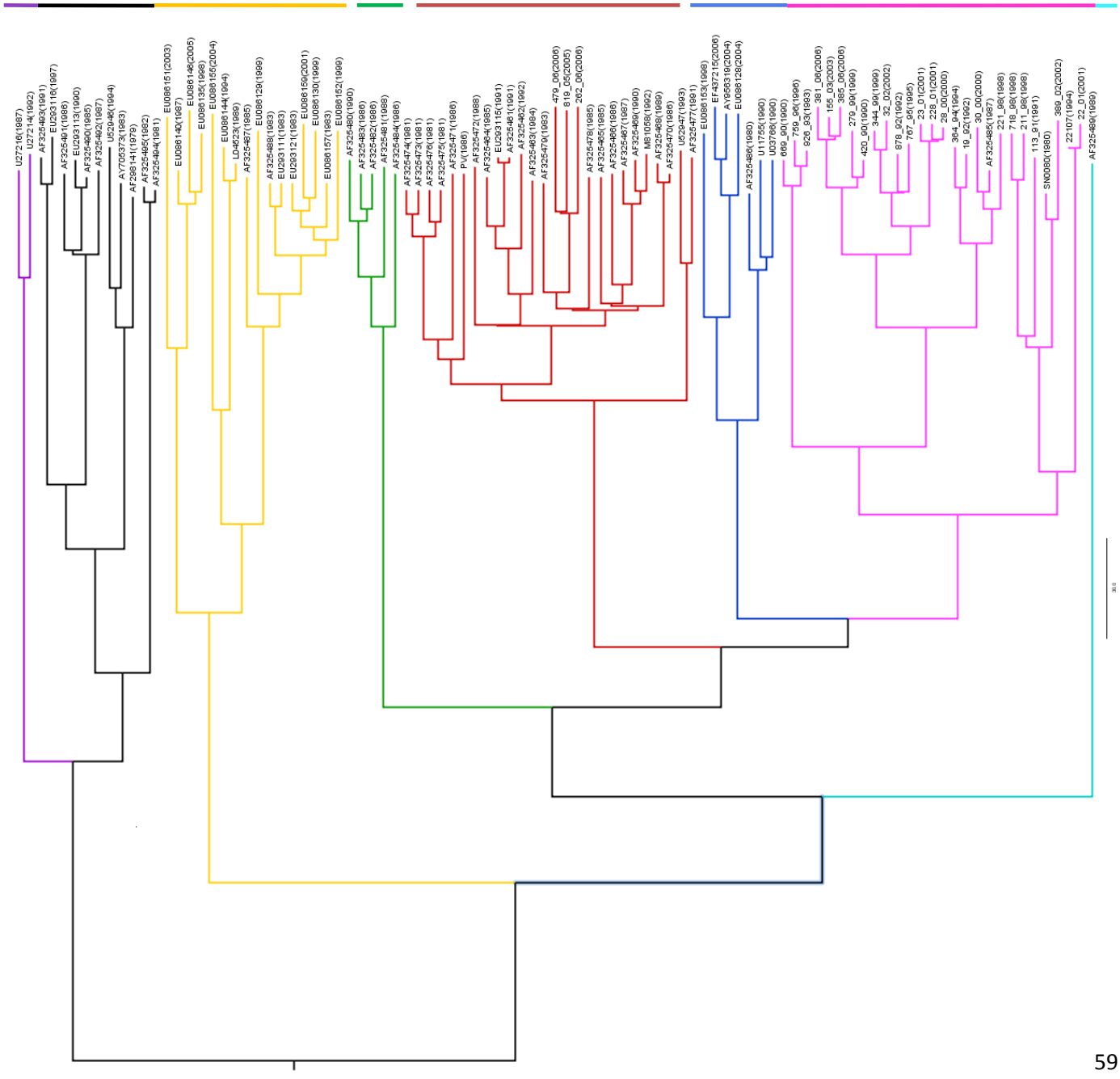

Figure 2.5 Maximum clade credibility tree constructed from a relaxed uncorrelated log-normal molecular clock using the glycoprotein gene dataset, with branch lengths scaled to reflect evolutionary time in years.

2.4 Discussion

The phylogenetic structure of RNA viruses often reflects their spatial and temporal dynamics (Biek *et al*., 2006; Grenfell *et al.,* 2004; Moya *et al*., 2004). Therefore detailed phylogenetic analysis often provides insight into the pattern and rate of geographical distribution of a given viral population. This is especially true for viruses like lyssaviruses that are subject to very little natural selection (Bourhy *et al*., 1999; Davis *et al*., 2005; Kissi *et al*., 1995). The aim of this study was to determine the broader phylogenetic context of the rabies virus mongoose biotype on a global scale, in order to attempt a reconstruction of the history of this virus group. Using the complete nucleoprotein and glycoprotein sequences that were obtained in this study as well as sequences obtained from the GenBank database, phylogenetic trees were constructed and molecular clocks were applied to these datasets. This was the first study of the rabies virus mongoose biotype that focused on two full length genes to apply molecular clocking techniques.

The use of a relaxed molecular clock instead of the traditional strict molecular clock is still in its initial stages. New developments are made regularly, with software packages being adjusted to accommodate biological data and needs of scientists. Recently Hughes (2008) suggested that previous findings on the most recent common ancestor (mrca) of EBLV should be revisited, since application of relaxed clocking techniques refutes initial estimates of the time of emergence of European bat lyssaviruses. Bourhy *et al*. (2008) also showed that a relaxed uncorrelated lognormal molecular clock is a suitable method to estimate the mrca of lyssavirus datasets. This together with other findings (Drummond and Rambaut, 2007) provide convincing support for the use of a relaxed molecular clock approach instead of the traditional strict clock method.

It has been postulated numerous times by various authors (Cluver, 1927; Snyman, 1940; Nel *et al*., 2005) that the mongoose variant has been present in southern Africa before the arrival of the canid variant. This was not only supported by anecdotal evidence, but also by the amount of heterogeneity that is displayed by mongoose variant isolates compared to canid variant isolates (Snyman, 1940; Nel *et al.,* 1993; Swanepoel, 1993; Nel *et al.,* 2005). It is interesting to note that both the full length nucleoprotein as well as the full length glycoprotein gene datasets supports the

phylogeny as was determined from the GL-intergenic region sequence. In analysis of the phylogeny of lyssaviruses associated with terrestrial animals, there is always a strong correlation between the phylogeny and geography of the isolates (Bourhy *et al.,* 1999; Kissi *et al*., 1995; Zhang *et al.,* 2006). Thus once again, as with other terrestrial gt1 virus variants, the clear correlation between geographic location and clustering of virus isolates was confirmed.

Analysis of the nucleoprotein dataset estimated the age of the mongoose variant to be approx 229 years old (95% HPD values=135-360 years), while the analysis of the glycoprotein data estimated a slightly more recent timescale (mean=158 years, 95% HPD values=159-202 years). The overlap of the 95% HPD values suggests that these estimates are robust. Davis *et al.* (2007) conducted a similar study using sequences from the G-L intergenic region and obtained an approximate age of 70 years old for the mongoose variant. It is not known why the different genes yield different results when molecular clocks are applied to the datasets, however many assumptions are made during the course of a molecular clock analysis. It seems logical that the large difference in substitution rates between the G-L intergenic region and the nucleoprotein and glycoprotein datasets ($0.826E^{-3}$ vs. mean rate $2.495E^{-4}$ to $6.875E^{-4}$) could affect the outcome of the molecular clock analysis. When taking a closer look at the nucleotide substitution rates of these three areas of the lyssavirus genome and the clock estimates that were obtained during analysis of these genes and pseudogene, a distinct trend is observed. The more conserved the gene under study, the older the mrca estimate appears to be. Thus we find that the G-L intergenic region yielded the youngest estimate (≈70 years) of the mrca of the mongoose biotype of rabies virus, with the glycoprotein gene yielding an older estimate (≈159 years) and the nucleoprotein gene, being the most conserved of these three areas, yielding the oldest estimate (≈229 years). It does not appear that analysis of only the mongoose virus isolates affects the outcome compared to analysis of all gt1 isolates.

When aiming to place the phylogeography of the mongoose variant in southern Africa in a global context, two hypotheses seem most likely when looking at the broader picture of *Lyssavirus* evolution. The first hypothesis states that the mongoose variant of rabies

virus first originated from a virus that was present in bats and then spilled over into new hosts such as viverrids and herpestids. This would be similar to the situation in the USA with bat rabies that spilled over into skunks and became established to form an independent transmission cycle in these animals (Leslie *et al*., 2006). However, the lack of any substantial evidence of the presence of rabies virus (gt1) in bats in Africa (Markotter *et al.,* 2006b; Kuzmin *et al.,* 2008) seems to argue against this possibility.

The second hypothesis is that the mongoose variant originated in another host such as civets or genets, and that this virus variant was introduced into southern Africa. There is evidence for the origin of Viverridae and Herpestidae to have been in Asia, and migration of these animals took place to Africa via the Arabian microplate a few million years ago (Gaubert and Cordeiro- Estrela, 2006). Furthermore, there is evidence of the colonization of southern Africa by agro-pastoralists since 700 A.D. It is speculated that migration of these people from the north downward into the rest of Africa could have disseminated rabies across Africa. However these populations experienced great decline in southern Africa and only flourished after 1500 A.D (King *et al.,* 1994). There is some speculation that these farmers might have owned dogs, and that these animals would have been susceptible to rabies virus infection, and thus would have been in sufficient contact with wildlife to transmit this virus to wild animals. From approx 700 A.D. there was trade between countries in the Middle–East and eastern Africa as well as Portugal (King *et al.,* 1994). It seems unlikely that the Muslim traders of the Middle East would have owned dogs that could have been harboring a rabies virus that then became established in southern Africa. All the above mentioned movement of animals and humans does not fall within the estimated timeframe of the origin of rabies virus mongoose biotype.

It still seems plausible that this virus variant was introduced into southern Africa from another country and it is almost instinctive to expect to see some strong relationship/similarity between the mongoose variant isolates and rabies virus isolates from another country e.g. India. This has not been observed when conducting phylogenetic analysis using lyssavirus isolates from across the globe. When reading historical writings about rabies in southern Africa, it seems that the mongoose virus was

first associated with members of the Viverridae such as genets, and then only later became established in the Herpestidae. There was a strong belief among local people even before the early 1800's that the bite of genets caused rabies (Cluver, 1927; Snyman, 1940). Olivier (1997) reported that certain sub-groups of the mongoose variant evolved after other groups. The cluster formed by isolates from *Cynictis penicillata* hosts were found to have evolved last of all the groups identified. This was suggested to indicate possible spill-over of this virus variant into mongooses from another viverrid host. This again strengthens the argument that the mongoose variant may first have been associated with members of the Viverridae such as genets or civets, and then became established in members of the Herpestidae such as mongooses and suricates. When we look at the phylogenetic clustering of the group of mongoose biotype isolates (Figures 2.1, 2.3, 2.4) it is seen that the mongoose viruses share a common ancestor with viruses from the cosmopolitan lineage. This particular group of viruses was spread worldwide through European colonization and therefore it seems likely that the mongoose variant of rabies virus was introduced by European travelers during the age of colonization of southern Africa, where it then established itself in viverrid and herpestid hosts due to lack of high numbers of canids sometime around the late 1700's or early 1800's.

For long it has been assumed that most, if not all, RNA viruses share a long evolutionary history, with possible co-evolution between virus and host taking place for millions of years. This has been a plausible explanation for the very specific host adaptation we find in these viruses. However, when the origin of RNA viruses are inferred from sequence data we find that this group of viruses could not have evolved more than 50 000 years ago. This paradox is observed in most cases where a molecular clock is applied to a RNA viral dataset (Holmes, 2003). An example of this paradox would be the evolution of lentiviruses such as HIV (Human Immunodeficiency virus) and SIV (Simian Immunodeficiency virus) where based on evidence of pathogenesis, it would appear that the range of SIV's have eco-evolved with their natural hosts. Not only are these viruses asymptomatic in the primates, but the phylogenies of the hosts and the viruses match. When compared to the high virulence of HIV in humans, this surely points to a extended period of host-virus co-evolution.

However, molecular clock calculations only date the origin of these primate viruses to a few thousand years ago (Holmes, 2003). This seems relatively young for a virus variant that is, on various levels, extremely well adapted to its specific hosts. Holmes (2003) proposes three main reasons for the observed discrepancy.

Firstly it is possible that the rate of evolution has fluctuated during the course of virus evolution, not just between viruses but also between lineages of the same virus. For example if some viruses or specific lineages of a virus has experienced periods where there was a reduced rate of nucleotide substitution, it would dramatically alter the divergence time of that virus or lineage compared to others. A second reason for the possible underestimation of divergence times is that current methods used to estimate evolutionary distance based on sequences are not adequate to model natural evolution. However, the development of methods that allow the variation of substitution rates to vary among sites in an alignment has been a large improvement. But it is possible that RNA viruses evolve in such a manner that there are great variations among sites. RNA genomes are kept relatively small due to the error threshold of the error prone viral polymerase, to avoid a large number of deleterious mutations. This implies that mutation can only occur in regions that are unrestricted by function, and since most of the viral genome encodes proteins, these areas of “high mutation” are limited to a small proportion of the genome. This in turn leads to large rate variation among sites in the genome (Holmes, 2003). A third explanation for the observed paradox is that it is not a paradox, but a reflection of the real situation. The RNA viruses that are under study have only been identified in the past 100 years. Thus it is possible that viral families like the flaviviruses or rhabdoviruses have histories dating back millions of years, but that the ancient members of these families have gone extinct and have been replaced by the viruses we sample today. It is interesting to note that when a “tree of life” is constructed for RNA virus families, they are only distantly related, almost to the extent of only being randomly connected to each other (Zanotto *et al.,* 1996). Thus it seems likely that lineages that would reflect the true path of evolution of these viruses have been replaced through continuous extinction and multiple substitutions (Holmes, 2003).

Similarly we find ourselves confronted with an mrca date of approx 200 years old for the mongoose variant. Therefore it seems most plausible that the estimated time of divergence is indeed not a paradox but a true reflection of the viral dataset.

Chapter 3: Molecular phylogeny of Mokola virus

3.1 Introduction

Mokola virus forms gt3 of the *Lyssavirus* genus and is one of four lyssaviruses that occur exclusively on the African continent. There have only been 25 isolations of Mokola virus in approx 40 years. This virus was first isolated from shrews (*Crocidura* sp.) that were captured in the Mokola forest, near Ibadan, Nigeria in 1968 (Shope, 1970). After this initial isolation further isolations were made from various hosts including humans (Familusi *et al.,* 1972), a rodent (*Lophuromys sikapusi)*, domestic cats and dogs (Foggin, 1983, Saluzzo *et al.,* 1984, Von Teichman *et al.,* 1998). MOKV was first isolated in South Africa from a cat in Umhlanga Rocks, KwaZulu-Natal province in 1971 (Schneider *et al.,* 1985). Other cases of MOKV infection in cats and a dog were observed in Zimbabwe during 1981 and 1982 (Foggin, 1982). In the period between 1995-1998 further isolations were made from cats in the regions surrounding East London and Pietermaritzburg (Von Teichman *et al.,* 1998) and most recently two isolations were made form a cat in East London and a dog in Nkomazi, Mpumalanga in 2004 and 2005 respectively (Sabeta *et al.,* 2007).

Until recently only a few sequences from a selected number of Mokola virus isolates were available on GenBank. The completion of the full genomes of the Central African Republic and Cameroon isolates have increased the amount of sequence information that is available in the public domain. Few evolutionary studies have focused on the epidemiology of the African lyssaviruses, in particular MOKV (Bourhy *et al.,* 1993; Von Teichman *et al.,* 1998; Nel *et al.,* 2000; Bingham *et al*., 2001; Badrane *et al*., 2001; Sabeta *et al.,* 2007). Bourhy *et al*. (1993) and Badrane *et al*. (2001) both focused on the elucidation of the phylogeny of the *Lyssavirus* genus, and as such the emphasis was not on the phylogeny of MOKV. Von Teichman *et al.* (1998) reported on the isolation of new MOKV isolates form South Africa, while Nel *et al*. (2000) and Bingham *et al.* (2001) described the phylogenetic clustering of viral isolates according to their geographic location as well as the apparent heterogeneity among MOKV isolates that led to the conclusion that MOKV isolates display greater genetic diversity that gt 1 isolates. The report of new MOKV cases in 2007,8 years after the last reported case of MOKV infection, and another report in 2008, again emphasized that there is still a gap in knowledge as to the true occurrence of MOKV (Sabeta *et al*., 2007).

Thus little is known about the phylogeny of MOKV when utilizing full length gene sequences of all the MOKV isolates up to date. As a result this study aimed at compiling a complete representative dataset of the full nucleoprotein gene sequences for all available Mokola virus isolates, with the intent of utilizing this data in a larger study that will aim at elucidating the Mokola virus phylogeny as well as the molecular diversity in this genotype by using full length gene sequences of all Mokola virus isolates up to date. Secondary objectives would include a preliminary analysis of the molecular diversity within the gt3 viruses and well as molecular diversity when compared to the intrinsic homology of gt1 viruses.

3.2 Materials and methods

3.2.1 Viruses and virus isolation

In this study 12 Mokola virus isolates obtained from various animals from different regions across Zimbabwe and South Africa were analysed. These isolates were obtained from the Onderstepoort Veterinary Institute, Pretoria, and constitute all the isolates of Mokola virus that were stored in the Onderstepoort Veterinary Institute archives at the time this study commenced. The year of isolation, specific host species and geographical origins of the viruses are shown in Table 3.1. Isolation of viruses was conducted using the mouse inoculation test and FAT as described in section 2.2.1.

Table 3.1 Mokola virus isolates used in this study.

Lab reference number	*Year of isolation*	*Location*	*Species of origin*	*GenBank accession number*
700/70	1970	Umhlanga Rocks, KwaZulu Natal, South Africa	Cat	FJ465416
12341, 12574	1981	Bulawayo, Zimbabwe	Cat (2)	FJ465417, FJ465418
13270	1982	Bulawayo, Zimbabwe	Cat	-
21846	1993	Zimbabwe	Cat	-
543/95	1995	Mdantsane, Eastern Cape province, South Africa	Cat	FJ465415
112/96	1996	East London, Eastern Cape Province, South Africa	Cat	FJ465411
322/96	1996	Yellow Sands, Eastern Cape Province, South Africa	Cat	FJ465414
252/97 and 229/97	1997	Pinetown, Kwazulu Natal Province, South Africa	Cat (2)	FJ465413
071/98	1998	Pietermaritzburg, Kwazulu Natal Province, South Africa	Cat	FJ465410
404/05	2005	Nkomazi, Mpumalanga Province, South Africa	Dog	-
173/06	2006	East London, Eastern Cape Province, South Africa	Cat	FJ465412

[1] Sequence for isolate 252/97 was obtained from Dr. Markotter and has not been submitted to GenBank.

3.2.2 RNA extraction

RNA extraction was performed as described in section 2.2.2.

3.2.3 Primers for RT-PCR

All oligonucleotides used for RT-PCR and sequencing were produced by Inqaba Biotechnical Industries (Pty) Ltd. (S.A.) and used without further purification. PCR amplification and sequencing were performed with the primer sets indicated in Table 3.2. The symbols (+) and (-) refer to messenger and genomic sense respectively.

Table 3.2. Oligonucleotide primers used for cDNA synthesis, PCR and sequencing of the nucleoprotein gene of Mokola virus isolates. Positions indicated relative to full genome NC_006429)

Name	Sequence	Application	Target
001lys (+) (Markotter *et al.*,2006a)	5'-ACGCTTAACGAMAAA-3'	RT-PCR, sequencing	N (Pos -56 to -70, 3' end of genome)
550B (-) (Markotter *et al.*,2006a)	5'-GTRCTCCARTTAGCRCACAT-3'	PCR, sequencing	N(Pos 577)
MokN450F (+)	5'-GAGCATGCTTCATTGGTCGG-3'	RT-PCR, sequencing	N (Pos 527)
N8(-) (Bourhy *et al.*, 1993)	5'-AGTTTCTTCAGCCATCTC-3'	PCR, sequencing	N(Pos 1610)

3.2.4 cDNA synthesis

cDNA synthesis was performed as described in section 2.2.4 with the following modifications. Positive sense primers, 001lys and MokN450F (Table 3.2) were used to prime the reverse transcription.

3.2.5 PCR

PCR was performed as described in section 2.2.5 with the following modifications. Primer set 001lys and 550B as well as primer set Mok450F and N8(-) were used respectively to amplify the complete nucleoprotein gene in 2 separate reactions. The amplification involved an initial denaturation at 94℃ for 2 minutes, followed by 30 cycles of denaturation at 94℃, primer annealing at 37℃ for 30s and primer extension at 72℃ for 90s. This was followed by a final extension of 7 min at 72℃. Once amplification was completed products were analyzed by ethidium-bromide stained agarose gel electrophoresis as described in section 2.2.6.

3.2.6 Purification and sequencing

PCR amplicons were gel-purified as described in section 2.2.7. Automated sequencing was performed using the BigDye Terminator 3.1 system (PE Applied Biosystems) as described in section 2.2.7.

3.2.7 Phylogenetic analysis

Sequences were edited and contigs assembled in Vector NTI 9.1.0 (Invitrogen Corporation, 2004) using the ContigExpress application. Additional sequences of other Mokola virus isolates were obtained from the GenBank database (GenBank numbers of isolates used indicated in appendix, Table 2). This resulted in two final datasets containing 17 complete nucleoprotein gene sequences (1353 bp in length) and 19 partial nucleoprotein gene sequences respectively. Other sequences of gt1 isolates, representative of all the virus subgroups, were obtained from the GenBank database for comparative studies (Table 1, appendix). Alignments and phylogenetic tree construction was conducted as described in section 2.2.8.

3.3 Results

3.3.1 Virus isolates, RNA extraction, cDNA synthesis and PCR

A final panel of 12 virus isolates was selected (Table 3.1). These isolates were chosen to represent all the available Mokola virus isolates from southern Africa. RNA extraction, cDNA synthesis and PCRs were performed as described in section 3.2.2 to 3.2.5. The 001lys and 550B(-) primer pair yielded a product of approx. 660bp, while the primer set MokN450F and N8(-) yielded a product of approx 1100bp. Unfortunately isolates 21846 and 13270 never yielded amplicons with primer pair MokN45F and N8(-). It was suspected that a lack of sufficient intact RNA could have been responsible for the lack of amplification, however attempts to grow these viruses in cell culture were unsuccessful and thus no further work was conducted on these isolates. Attempts to culture isolate 404/05 were unsuccessful and thus this isolate was not included in this study.

3.3.2 Purification of PCR amplicons and nucleotide sequence determination

The purification of PCR amplicons and nucleotide sequencing was conducted as described in section 3.2.6. Sequence files were manually inspected using the ContigExpress application in Vector NTI 9.1.0 (Invitrogen Corporation, 2004) in order to confirm that bases were accurately called and to correct bases where necessary. Nucleotide sequencing yielded the full length nucleoprotein gene of all the selected isolates. The obtained sequences were trimmed using BioEdit software v 7.0 (Hall, 1999) to yield a final sequence fragment containing the open reading frame of the nucleoprotein (1353nt). All sequences generated in this study were submitted to GenBank and assigned accession numbers (Table 3.2).

3.3.3 Phylogenetic analysis

All isolates sequenced in this study yielded a nucleoprotein gene of 1353nt. Phylogenetic trees were constructed as described in section 3.2.7. Trees constructed using the NJ method utilizing the partial (405nt) as well as complete gene sequence of the nucleoprotein gene both displayed similar topology (Figure 3.1 and 3.2) and also showed the previously described grouping of isolates according to geographic location (Nel *et al.,* 2000). However, bootstrap support for the clades are not very high (<70%) and must thus be considered with caution. As expected the full length nucleoprotein gene dataset (Figure 3.2) provides better resolution of the phylogeny (bootstrap support >70% for majority of clades) compared to the partial nucleoprotein gene sequence data. When looking at the phylogeny as constructed using the full length nucleoprotein gene sequence, it is of interest to note that the South African isolates appear to form a cluster together with isolates from the rest of Africa, but very distinctly separate from the isolates from Zimbabwe. This is unforeseen as it would be expected that the isolates from neighboring countries would have been more closely related. Furthermore the grouping of the isolate from Cameroon with the isolate from Ethiopia is reason for concern as these two countries are situated on opposite ends of the African continent, and due to the geographic grouping of isolates it would be expected that these two isolates should not group together.

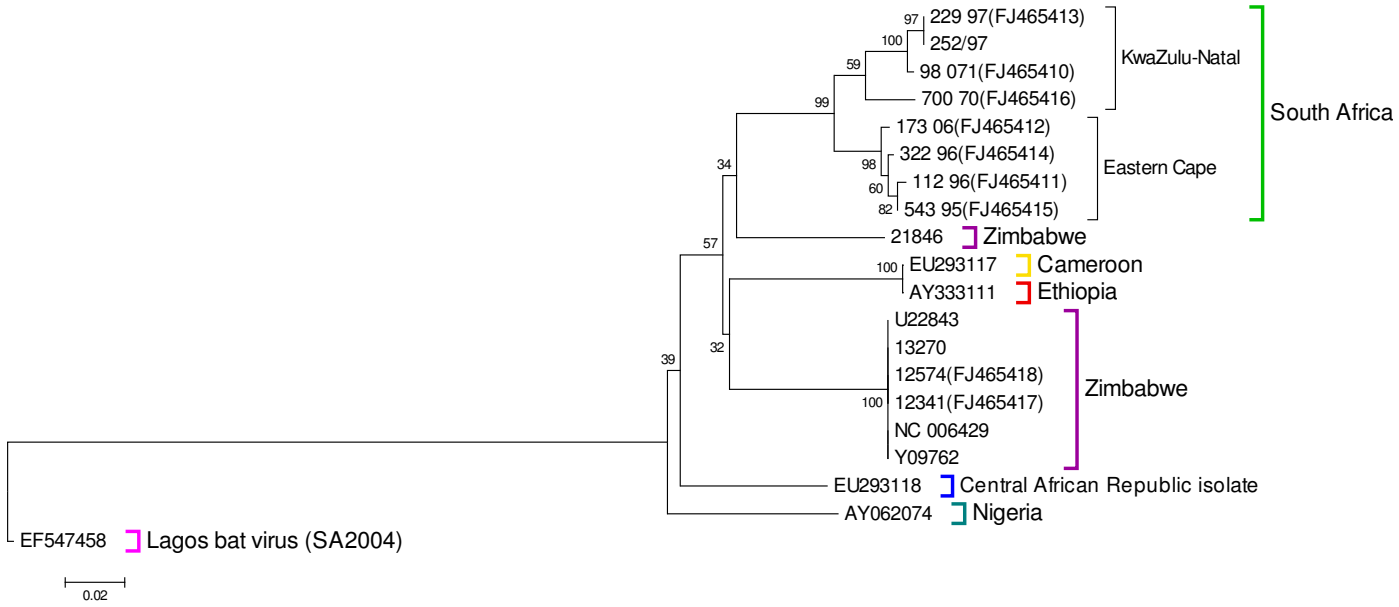

Figure 3.1 Neighbour-joining phylogenetic tree constructed from all available Mokola virus isolates using partial nucleoprotein gene sequences (nt1-405). A Lagos bat virus isolate from South Africa was used as outgroup. Bootstrap values are indicated at nodes and branches are drawn to scale.

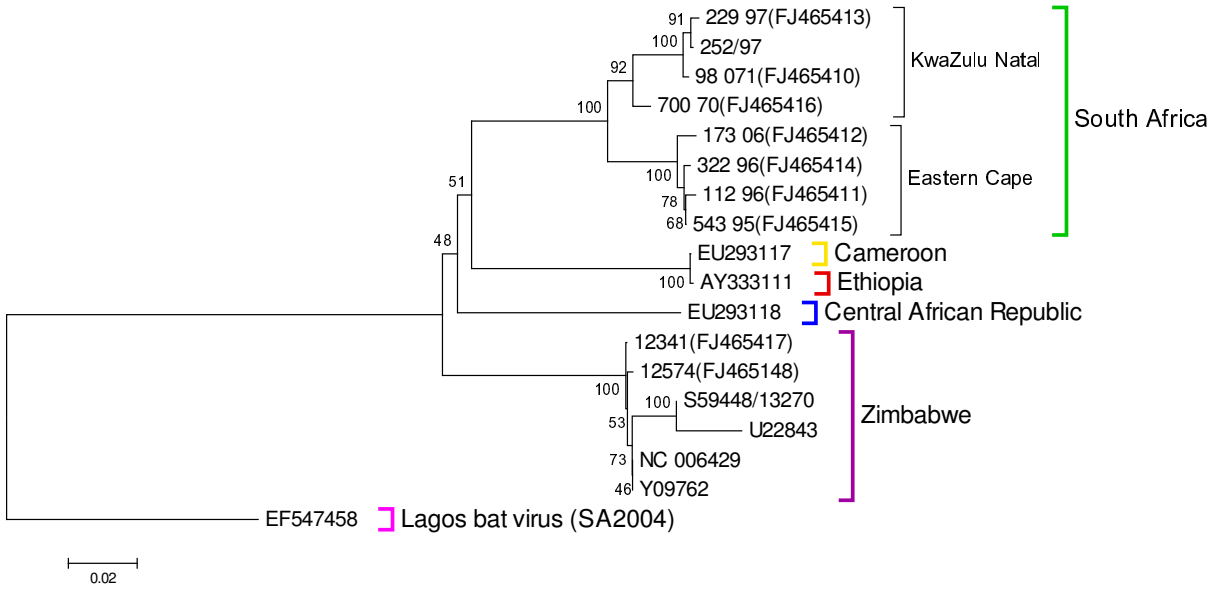

Figure 3.2 Neighbour-joining phylogenetic tree constructed from available Mokola virus isolates using full length nucleoprotein gene (1353nt) sequences. A Lagos bat virus isolate from South Africa was used as outgroup. Bootstrap values are indicated at nodes and branches are drawn to scale.

3.3.1 Phylogenetic comparison of Mokola virus and rabies virus (gt1) isolates.

In Figures 3.3 and 3.4 a distinct separation of the two different genotypes can be seen. Both gt1 and gt3 viruses form clades corresponding to the various subgroups corresponding to their geographic location. Once again the partial nucleoprotein gene sequence does not produce very high bootstrap support values for the subgouping of the viral isolates with in gt1 and gt3 respectively. The same tree topology was obtained for the deduced N amino acids tree with slightly higher bootstrap values (Figure 3.5).

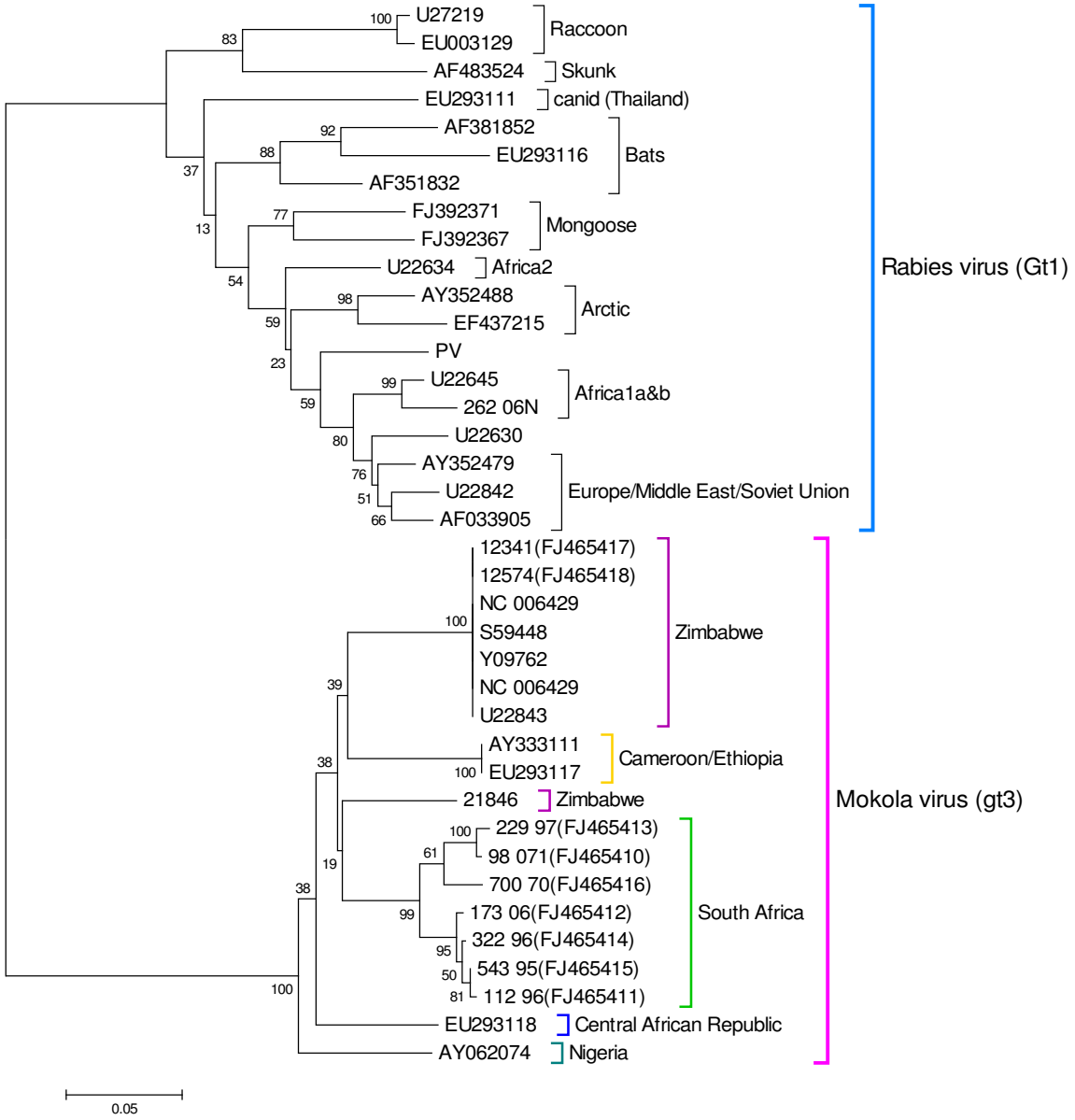

Figure 3.3 Neighbour-joining tree constructed using partial nucleoprotein gene (bp1-405) sequences of rabies virus (gt1) and Mokola virus (gt3) isolates. Gt1 isolates were selected to represent all major subgroups. Bootstrap values are indicated at nodes and branch lengths are drawn to scale.

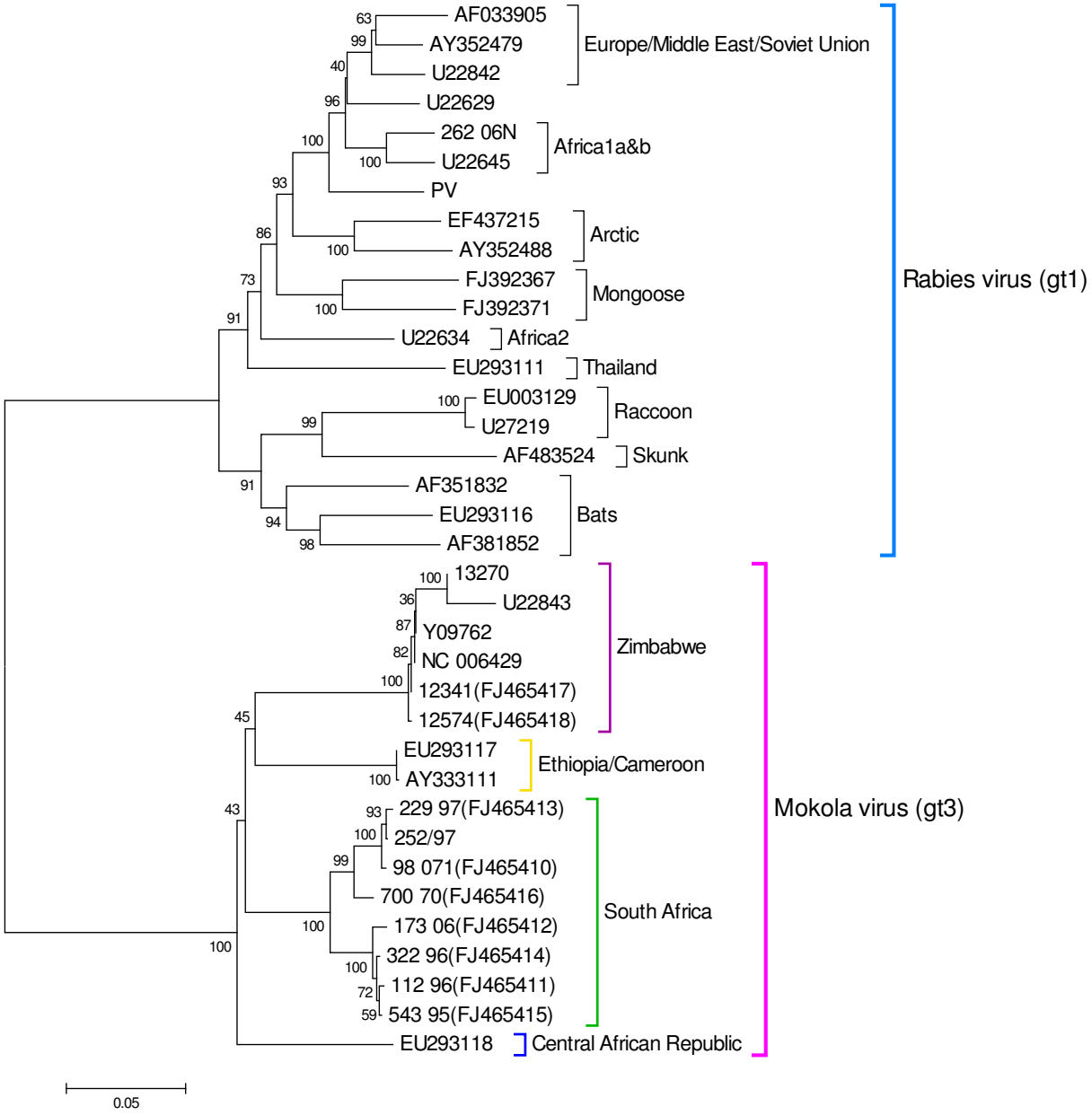

Figure 3.4 Neighbour-joining tree constructed using full-length nucleoprotein gene sequences of rabies virus (gt1) and Mokola virus (gt3) isolates. Gt1 isolates were selected to represent all major subgroups. GenBank accession numbers are indicated on the figure. Bootstrap values are indicated at nodes and branch lengths are drawn to scale.

Figure 3.5 Neighbour-joining tree constructed using full-length nucleoprotein sequences (450aa) of rabies virus (gt1) and Mokola virus (gt3) isolates. Gt1 isolates were selected to represent all major subgroups. GenBank accession numbers are indicated on the figure. Bootstrap values are indicated at nodes and branch lengths are drawn to scale.

It is interesting to note that in all cases, the tree drawn using the full-length nucleoprotein gene sequence gives better resolution of the topology of the tree when compared to the tree of the partial nucleoprotein gene sequence dataset. This is due to the fact that the first approximately 400nt of the nucleoprotein gene is very conserved and is therefore not suitable to distinguish between isolates that are closely related such as the Zimbabwe cluster. The partial nucleoprotein tree was constructed as this was the only part of the genome that had representative sequences for the majority of the Mokola virus isolates. Sequence homology values calculated through the p-distance method are shown in Table 3.3.

Table 3.3 Average nucleotide sequence similarity within gt1 and gt3 isolates used in this study calculated using the p-distance method (Kumar *et al*., 2004). Values in this table were calculated using the selected isolates indicated in Figure 3.3, Figure 3.4 and Figure 3.5.

	Average identity within gt based on partial nucleoprotein gene sequence	***Average identity within gt based on full length nucleoprotein gene sequence***
Mokola virus(gt3)	91.5%	91.3%
Rabies vius (gt1)	86.6%	86.7%

When analyzing the full length nucleoprotein gene sequence dataset it was found that individual Mokola virus isolates displayed up to 11.9% nucleotide difference (U22843 and 173/06), while individual gt1 isolates displayed up to 18.5% nucleotide difference (262/06 and AF483524). Individual Mokola virus isolates display up to 11.6% nucleotide sequence difference (EU293118 and 21846) based on the partial nucleoprotein sequence analysis, while gt1 isolates display up to 20.3% nucleotide

sequence difference (EU29311 and EU003129) based on the partial nucleoprotein sequence.

When comparing the average nucleotide sequence identity between sub-groups of Mokola virus, isolates from South Africa only display 88.2% homology with isolates from Zimbabwe and 87.9% homology with isolates from the other countries (Ethiopia, Cameroon and Central African Republic). When the selected isolates that were chosen to represent the gt1 subgroups (as seen in Figure 3.3) were analyzed together with the gt3 isolates, MOKV isolates display 92.1-98.6% homology within subgroups while gt1 isolates display 91.2-99.2% identity within different subgroups. These values are however dependant on the isolates that are used in the analysis, and as such must be viewed as estimates as not all gt1 isolates that currently have sequence information available on GenBank were included in this analysis.

3.4 Discussion

This study aimed at producing the first comprehensive phylogenetic analysis utilizing all the Mokola virus isolates up to date. This work will form part of a larger collaborative effort to sequence various regions of the genome of all the Mokola virus isolations up to date. This will allow a complete molecular phylogenetic study using various full-length genes. As has been observed with other lyssavirus phylogenetic studies, the Mokola virus isolates group according to the geographic location they were obtained from. Isolates form Ethiopia and Cameroon appear to form a cluster separate from the Zimbabwean and South African groups. It is possible that this is a skewed grouping due to the lack of any other isolates from surrounding countries. However, the fact that more sequence diversity is observed between isolates from a single country (i.e. South Africa) than between these isolates, as well as the fact that their localities of isolation are located on opposite ends of the African continent, implies that these isolates cannot be this closely related. This suspicious clustering calls for the elucidation of this grouping as both these sequences were obtained from GenBank. Furthermore some variation in the topology of phylogenetic trees is observed. It seems plausible that this variation is due to the small sample size utilized in this study. And if the low bootstrap values are viewed as insufficient support for certain groupings, therby allwing the

collapsing of certain branches, the topology does seem to correspond. The region of the nucleoprotein that was used to construct the partial nucleoprotein trees is considered a highly conserved region of the Lyssavirus genome, and consequently this area of the genome is generally not used in phylogenetic analysis as it does not provide sufficient phylogenic signal to provide clear elucidation. Therefore the low bootstrap values obtained could at least in part be attributed to the unfavorable phylogenetic data. Unfortunately this was the only part of the genome that had sequence data available in the public domain for all Mokola isolates to date. The lack of very clear histories of each of the isolates severely hampers the compiling of a thorough phylogenetic analysis. Foggin (1988) describes the collection of all the MOKV isolates obtained in Zimbabwe during 1981 and 1982. These isolations were all made from animals in a small geographic area. Thus it is not surprising that these isolates all appear very cellos related. The unusual separate clustering of isolate 21846 could possibly be partially explained by the fact that this sample was obtained from a cat nearly 100km form the site where previous cases were found. However it does seem geographic separation alone cannot account for the large difference observed between these isolates. Unfortunately nothing in the published case history of this isolate, points to any possible mislabeling or confusion that could clarify this matter (Bingham *et al.*, 2001).

The combination of the two abovementioned stumbling blocks most probably contributed to the vague interpretational value of this analysis.

It has been speculated that Mokola virus displays more genetic variation that gt1 viruses. However, in the last few years, numerous discoveries of new gt 1 variants have been made (Leslie *et al.,* 2006), and these new variants dramatically increased the genetic variation within this group. It would appear that, based on the preliminary findings of this study, it might not be the case that Mokola virus isolates display more genetic variability than gt1 isolates, but that the genetic variability within gt3 is similar to or less than the genetic variability found within gt1. It must however be taken into account that the Mokola virus isolates originate from only a limited number of countries, all located on the same continent, while their diversity is being compared to isolates that originate not only from various hosts, but also from various regions across the globe.

Thus when these factors are taken into account, the diversity that we observe between Mokola virus isolates is significant. It is very probable, that with increased surveillance, the number of isolates will increase significantly, which might in turn have an impact on the amount of genetic diversity we observe in future comparative studies.

Africa has been considered the birthplace of lyssaviruses as it harbors the greatest diversity of these viruses. Due to the very infrequent isolation of Mokola virus, one can only speculate as to its distribution and host in nature, as well as the age of this virus. When taking into account that Mokola virus is easily adaptable to grow in insect cell cultures and mosquitoes themselves (Buckley, 1975; Aitken, 1984) and has been isolated from animals that are at least partially insectivorous (e.g. shrews and cats) it seems that Mokola virus might have been one of the first lyssaviruses to have evolved from another rhabdovirus. The initial isolation of MOKV from shrews was a random event, and in years that followed active surveillance efforts did not yield any positive results. Of 408 serum samples of various sources in Nigeria that were tested for MOKV antibodies, only one each from cattle, goat, swine, birds and fruit bats (*Eidolon helvum*) tested positive. Sera from man, sheep, shrews and other bats were negative and three serum samples out of 500 samples taken from dogs in Nigeria were also seropositive for MOKV (Aghomo *et al*., 1990). Recent surveillance efforts yielded no positive results and It would appear that there is a decrease in MOKV cases. If this is the case it might very well be that Mokola virus is nearing the end of its existence.

Chapter 4: Concluding remarks

Africa has been considered the “birthplace” of lyssaviruses and four of these viruses are unique to this continent. Furthermore it has been confirmed by various studies and methods that the mongoose variant of rabies virus is distinctly different from the canid variant found in southern Africa (King *et al*., 1993; Nel *et al*., 1993; Nel *et al.,*2005). This study as well as the study by Davis *et al.,* (2007) provided supporting evidence to the belief that the mongoose variant has been present in southern Africa before the arrival of the canid variant in the early 1900’s. But perhaps the most important conclusion that can be drawn from these studies as well as the study by Hughes (2008) and Bourhy *et al.* (2008) is that when molecular clock studies are conducted, the results should be interpreted with care, and that caution should be applied to assume these outcomes as fact. At the very best, it would appear that with methods currently at our disposal, we can make very broad estimates and observations as to the true evolutionary pathway of biological organisms and viruses. The lack of fossil evidence to not only provide reliable calibration points, but also support estimates of viral evolutionary timeframes severely compromises the support for molecular clock outcomes. As methods are refined and our understanding of viral evolutionary mechanisms increase, we may reach a stage where molecular clock data can be interpreted with greater confidence. Very little is known about the pathogenicity and true distribution of the rabies virus mongoose biotype. Future studies may focus on elucidating the pathogenesis of this virus variant in canids, which may in turn answer why mongoose rabies invariably reaches a dead-end infection in canid hosts as this will contribute significantly towards our understanding of the spread of this rabies biotype in nature.

Only 25 isolates of Mokola virus have been made in 40 years. There is very little surveillance for Mokola viruses in Africa, and at best, the majority of cases are identified after sick animals were submitted for routine rabies diagnosis. It is most probably the case that the incidence of Mokola virus is underreported due to limited diagnostic capability of African laboratories as well as absence of any large scale surveillance for this virus. The host for MOKV is still unidentified. The involvement of domestic animals as well as the lack of vaccine protection against Mokola virus infection underlines the possible public health threat posed by this virus. At the commencement of this study, there were very few full length gene sequences available on GenBank for any of the

Mokola virus isolates. In the majority of the GenBank submissions, only a partial gene sequence of the virus isolate was obtained for diagnostic purposes. Therefore this study will form part of larger initiative to compile the first complete phylogenetic analysis of all Mokola virus isolates to date using full length gene sequences. On analysis of the partial as well as the full length nucleoprotein gene sequence phylogeny, it was found that, like all other lyssaviruses associated with terrestrial animals, Mokola virus isolates group according to their geographic location. However it was found that isolates from South Africa displayed the highest amount of homology with isolates from Zimbabwe. This is expected due to the close geographical location of these countries. Furthermore isolates from South Africa formed two distinct clusters corresponding to isolates form the Eastern Cape province and KwaZulu Natal province respectively. This is noteworthy as isolates from these adjoining provinces still display enough genetic diversity to form two distinguishable clades.

The natural host of Mokola virus is still unknown and it may prove worthwhile to initiate surveillance programs for this virus to determine the true occurrence in southern Africa. Not only will this improve our understanding of the epidemiology of Mokola virus, but it may just provide the discovery of a host or “missing link” that will illuminate some of the mysteries regarding this unique African lyssavirus.

Chapter 5: Appendix

Table 1. Rabies virus (gt1) isolates used in phylogenetic analysis.

Virus code	Host species	Year of isolation	Geographic location	Reference/ Source	GenBank accession number	
					N gene	G gene
ARC1-PF	Polar fox *(Alopex lagopus)*	1980	Arctic circle	Badrane and Tordo, 2001		AF325486
ARC2-DG	Polar fox *(Alopex lagopus)*	1990	Arctic circle	Badrane and Tordo, 2001		U03766
04029AFG	Dog (*Canis familiaris)*	2004	Afghanistan	Bourhy *et al.*, 2008		EU086128
9137ALG	Dog (*Canis familiaris)*	1982	Algeria	Kissi *et al.,* 1995	U22643	M81058
9704ARG	Bat *(Tadarida brasilliensis)*	1997	Argentina	Delmas *et al.*, 2008	EU293116	EU293116
ARG1-BT	Bat	1991	Argentina	Badrane and Tordo, 2001		AF325493
86111YOU	Red fox *(Vulpes vulpes)*	1986	Bosnia	Kissi *et a*l., 1995	U42706	
BRA2-BV	Bovine (*Bos Taurus)*	1986	Brazil	Badrane and Tordo, 2001		AF325491
9908CBG	Dog (*Canis familiaris)*	1999	Cambodia	Bourhy *et al.,* 2008	EU086167	EU086130
8801CAM	Dog (*Canis familiaris)*	1987	Cameroon	Kissi *et al.,* 1995	U22634	
8804CAM	Cat (*Felis domesticus)*	1988	Cameroon	Kissi *et al.,* 1995	U22635	
8805CAM	Unknown	1988	Cameroon	Kissi *et al*., 1995, Badrane and Tordo, 2001	U22636	AF325481
CAN3-SK	Skunk *(Mephitis mephitis)*	1990	Canada	Badrane and Tordo, 2001		U11755
9105CAN	Red fox *(Vulpes vulpes)*	1990	Canada	Kissi *et al*., 1995	U22655	
4055DG	Dog (*Canis familiaris)*	1992	Canada	Nadin-Davis *et al.,* 1994	U03770	U03767

ML7	Bat *(Lasionycteris noctivagans)*	1979	Canada	Nadin-Davis *et al.*, 2001	AF351837	
LAN12	Bat *(Lasionycteris noctivagans)*	1988	Canada	Nadin-Davis *et al.*, 2001	AF351840	
LB1	Bat *(Lasiurus borealis)*	1994	Canada	Nadin-Davis *et al.*, 2001	AF351844	
LB6	Bat *(Lasiurus borealis)*	1991	Canada	Nadin-Davis *et al.*, 2001	AF351856	
EF34	Bat *(Eptesicus fuscus)*	1972	Canada	Nadin-Davis *et al.*, 2001	AF351832	
EF19	Bat *(Eptesicus fuscus)*	1991	Canada	Nadin-Davis *et al.*, 2001	AF381855	
EF57	Bat *(Eptesicus fuscus)*	1997	Canada	Nadin-Davis *et al.*, 2001	AF381859	
ME1	Bat *(Myotis evotis)*	1992	Canada	Nadin-Davis *et al.*, 2001	AF381835	
ML6	Bat *(Myotis lucifugus)*	1994	Canada	Nadin-Davis *et al.*, 2001	AF381838	
EF22	*Eptesicus fuscus*	1988	Canada	Nadin-Davis *et al.*, 2001	AF381830	
9229CAF	Dog (*Canis familiaris)*	1992	Central African Republic	Kissi *et al.*, 1995	U22651	
CHI1-BK	Buck	1986	China	Badrane and Tordo, 2001		AF325471
CHI2-DG	Dog (*Canis familiaris)*	1989	China	Badrane and Tordo, 2001		L04523
9811CHI	Dog (*Canis familiaris)*	1998	China	Bourhy *et al.*, 2008	EU086173	EU086135
02041CHI	Dog (*Canis familiaris)*	1987	China	Bourhy *et al.*, 2008	EU086177	EU086140
02046CHI	Dog (*Canis familiaris)*	1994	China	Bourhy *et al.*, 2008	EU086182	EU086144
05005CHI	Dog (*Canis familiaris)*	2005	China	Bourhy *et al.*, 2008	EU086186	EU086146
9142EST	Raccoon dog (*Nyctereutes procyonoides)*	1985	Estonia	Kissi *et al.*, 1995	U22476	

9342EST	Raccoon dog (*Nyctereutes procyonoides)*	1985	Estonia	Kissi *et al.,* 1995	U43432	
8658YOU	Cattle (*Bos taurus)*	1981	Federal Republic of Yugoslavia	Kissi *et al.,* 1995	U42705	AF325463
9147FRA	Fox *(Vulpes vulpes)*	1991	France	Delmas *et al.,* 2008	EU293115	EU293115/ AF325461
9001FRA	Dog (*Canis familiaris)* bitten by bat	1990	France	Delmas *et al.,* 2008	EU293113	EU293113
9616FRA	Sheep (*Ovis* sp.)	1996	France	Bourhy *et al.,* 2008	AF033905	
GUY1-BV	Bovine *(Bos taurus)*	1985	French Guyana	Badrane and Tordo, 2001		AF325490
9001GUY	Dog (*Canis familiaris)*	1990	French Guyana	Kissi *et al.,* 1995	U22478	
8693GAB	Dog (*Canis familiaris)*	1986	Gabon	Kissi *et al.,* 1995	U22629	
8698GAB	Dog (*Canis familiaris)*	1986	Gabon	Kissi *et al.,* 1995; Badrane and Tordo, 2001	U22630	AF325470
8684GRO	Arctic fox *(Alopex lagopus)*	1981	Greenland	Kissi *et al.,* 1995	U22654	
8660GUI	Dog (*Canis familiaris)*	1986	Guinea	Kissi *et al.,* 1995	U22487	AF325484
HUN1-HM	Fox *(Vulpes vulpes)*	1992	Hungary	Badrane and Tordo, 2001		AF325462
9215HON	Human *(Homo sapiens)*	1991	Hungary	Kissi *et al.,* 1995	U43025	
AY956319	Human *(Homo sapiens)*	2004	India	Bourhy *et al.,* 2008	AY956319	AY956319
NNV-Rab-H	Human *(Homo sapiens)*	2006	India	Bourhy *et al.,* 2008	EF437215	EF437215
03003INDO	Dog (*Canis familiaris)*	2003	Indonesia	Bourhy *et al.,* 2008	EU086192	EU086151
8681IRA	Dog (*Canis familiaris)*	1986	Iran	Kissi *et al.,* 1995	U22842	
IRN1-HM	Human *(Homo sapiens)*	1988	Iran	Badrane and Tordo, 2001		AF325472

9239CI	Dog (*Canis familiaris)*	1992	Ivory Coast	Kissi *et al.,* 1995	U22652	
IVC1-UN	Unknown	1986	Ivory Coast	Badrane and Tordo, 2001		AF325482
9910LOA	Dog (*Canis familiaris)*	1999	Laos	Bourhy *et al.,* 2008	EU086193	EU086152
MAD1-DG	Dog (*Canis familiaris)*	1985	Madagascar	Badrane and Tordo, 2001		AF325478
8698MAU	Camel (*Camelus* sp.)	1985	Mauritania	Kissi *et al.,* 1995	U22489	AF325483
MAL1-HM	Human *(Homo sapiens)*	1985	Malaysia	Badrane and Tordo, 2001		AF325487
9126MEX	Dog (*Canis familiaris)*	1991	Mexico	Kissi *et al.,* 1995, Badrane and Tordo, 2001	U22477	AF325477
MEX2-BT	Bat (*Desmodus* spp.)	1987	Mexico	Badrane and Tordo, 2001		AF325492
87012MAR	Dog (*Canis familiaris)*	1986	Morocco	Kissi *et al.,* 1995	U22631	AF325467
MOR1-DG	Dog (*Canis familiaris)*	1989	Morocco	Badrane and Tordo, 2001		AF325468
MOR3-HM	Human *(Homo sapiens)*	1990	Morocco	Badrane and Tordo, 2001		AF325469
9913BIR	Dog (*Canis familiaris)*	1999	Myanmar	Bourhy *et al.,* 2008	EU086165	EU086129
8708NAM	Kudu (*Tragelaphus strepsiceros)*	1987	Namibia	Kissi *et al.,* 1995	U22632	
9227NAM	Jackal (*Canis* sp.)	1992	Namibia	Kissi *et al.,* 1995	U22649	
NEP1-DG	Dog (*Canis familiaris)*	1989	Nepal	Badrane and Tordo, 2001		AF325489
9901NEP	Dog (*Canis familiaris)*	1998	Nepal	Bourhy *et al.,* 2008	EU086196	EU086153
9012NIG	Dog (*Canis familiaris)*	1990	Niger	Kissi *et al.,* 1995	U22640	AF325480
8670NGA	Human *(Homo sapiens)*	1983	Nigeria	Kissi *et al.,* 1995	U22488	AF325479

04030PHI	Human *(Homo sapiens)*	2004	Phillipines	Bourhy *et al.,* 2008	EU086205	EU086155
8618POL	Raccoon dog (*Nyctereutes procyonoides)*	1985	Poland	Kissi *et al.,* 1995	U22840	AF325464
POL2-HM	Human *(Homo sapiens)*	1985	Poland	Badrane and Tordo, 2001		AF325465
9141RUS	Arctic fox *(Alopex lagopus)*	1988	Russia	Kissi *et al.,* 1995	U22656	
1500AFS	Yellow mongoose (*Cynictis penicilata)*	1987	South Africa	Kissi *et al.,* 1995	U22628	AF325485
2900AFS	Water mongoose (*Atilax paludinosus*)	1991	South Africa	Kissi *et al.,* 1995	U22861	
479/96	Dog (*Canis familiaris)*	1996	South Africa	Unpublished		
819/05	Dog (*Canis familiaris)*	2005	South Africa	Unpublished		
262/06	Dog (*Canis familiaris)*	2006	South Africa	Unpublished		
94257SRI	Dog (*Canis familiaris)*	1986	Sri Lanka	Badrane and Tordo, 2001		AF325489
1294	Dog (*Canis familiaris)*	1986	Sri Lanka	Bourhy *et al.,* 2008	AY138549	
9135OMA	Red fox *(Vulpes vulpes)*	1990	Sultanate of Oman	Kissi *et al.,* 1995	U22480	
9221TAN	Dog (*Canis familiaris)*	1992	Tanzania	Kissi *et al.,* 1995	U22645	
8738THA	Human *(Homo sapiens)*	1983	Thailand	Kissi *et al.,* 1995, Badrane and Tordo, 2001	U22653/ EU086208	AF325488
8743THA	Human *(Homo sapiens)*	1983	Thailand	Delmas *et al.,* 2008	EU293121/EU08 6207	EU293121/EU0861 58
8734THA	Human (*Homo sapiens)*	1983	Thailand	Bourhy *et al.,* 2008	EU086206	EU086157
8764THA	Human *(Homo sapiens)*	1983	Thailand	Delmas *et al.,* 2008	EU293111	EU293111
Dr. Td2	Cow *(Bos Taurus)*	1995	Trinidad	Nadin-Davis et al . 2001	AF381852	

TUN1-DG	Human *(Homo sapiens)*	1986	Tunisia	Badrane and Tordo, 2001		AF325466
USA1-DG	Dog (*Canis familiaris)*	1981	USA	Badrane and Tordo, 2001		AF325576
USA2-SP	Sheep (*Ovis* sp.)	1981	USA	Badrane and Tordo, 2001		AF325475
USA3-SK	Skunk *(Mephitis mephitis)*	1981	USA	Badrane and Tordo, 2001		AF325473
USA4-SK	Skunk *(Mephitis mephitis)*	1981	USA	Badrane and Tordo, 2001		AF325474
USA7-BT	Bat (*Myotis* spp.)	1979	USA	Badrane and Tordo, 2001		AF298141
USA8-BT	Bat (*Myotis* spp.)	1981	USA	Badrane and Tordo, 2001		AF325494
USA9-BT	Bat (*Myotis* spp.)	1982	USA	Badrane and Tordo, 2001		AF325495
USA10-BT	Human *(Homo sapiens)*	1994	USA	Badrane and Tordo, 2001		U52946
USA11-CO	Coyote *(Canis latrans)*	1993	USA	Badrane and Tordo, 2001		U52947
USA13-RC	Raccoon *(Procyon lotor)*	1987	USA	Badrane and Tordo, 2001		U27216
SHBRV-18	*Lasionycteris noctivagans*	1983	USA		AY705373	AY705373
TB1	*Tadarida brasiliensis*	Unknown	USA	Nadin-Davis *et al.,* 2001	AF351849	
NY 771	Raccoon *(Procyon lotor)*	1992	USA	Nadin-Davis *et al.,* 1996	U27219	U27215
FLA.125	Raccoon *(Procyon lotor)*	1992	USA	Nadin-Davis *et al.,* 1996	U27220	U27216
01016VNM	Dog (*Canis familiaris)*	2001	Vietnam	Bourhy *et al.,* 2008	EU086209	EU086159
8915ZAI	Dog (*Canis familiaris)*	1989	Zaire	Kissi *et al.,* 1995	U22638	
Pasteur virus				Tordo *et al.,* 1986	M13215	M13215

Table 2. Mokola virus isolates used in this study.

Year of isolation	***Location***	***Species of origin***	***Reference***	***Genbak Accession number***
1968/1969	Ibadan,Nigeria	Shrews (3)	Shope, 1970, Kemp *et al.*,, 1972	AY062074
1974	Cameroon	Shrew	Le Gonidec *et al.*,, 1978	EU293117
1981/1982	Bulawayo, Zimbabwe	Dog, cat (4)	Foggin, 1982, 1983	U22843; Y09762; S59448; NC_006429
1981	Central African Republic	Rodent	Saluzzo *et al.*,, 1984	EU293118
1990	Ethiopia	Cat	Mebatsion *et al.*,, 1992	AY333111

Chapter 6: References

- Adamson, J.S. 1954. Ecology of rabies in southern Rhodesia. *Bulletin World Health Organization.* 10:753-759.
- Aitken, T.H.G., Kowalski, R.W., Beaty, B.J., Buckley, S.M., Wright, J.D., Shope, R.E., Miller, B.R. 1984. Arthropod studies with rabies-related Mokola virus. *Am. J. Trop. Med. Hyg.* 33(5): 945-952.
- Arai, Y.T., Kuzmin, I.V., Kameoka, Y., Botvinkin, A.D. 2003. New Lyssavirus genotype from the lesser mouse-eared bat (*Myotis blythi*), Kyrghyzstan. *Emerg. Infect. Dis.*9:333-337.
- Badrane, H., Tordo, N. 2001. Host switching in the Lyssavirus history from the Chiroptera to Carnivora orders. *J. Virol.* 17: 8096-8014.
- Badrane, H., Bahloul, C., Perrin, P., Tordo, N. 2001. Evidence of two lyssavirus phylogroups with distinct pathogenicity and Immunogenicity. *J. Virol.*75:3268-3276.
- Baer, G.M. 1994. Rabies- a historical perspective. *Infect. Agents Dis*. 3:168-180.
- Bahloul, C., Jacob, Y., Tordo, N., Perrin, P. 1998. DNA-based immunization for exploring the enlargement of immunological cross-reactivity against the lyssaviruses. *Vaccine* 16(4):417-425.
- Baldauf, S.L. 2003. Phylogeny for the faint of heart: a tutorial. *Tren. Gen*. 19(6): 345-351.
- Bauer, S.P., Murphy, F.A. 1975. Relationship of two arthropod-borne rhabdoviruses (Obodhiang and Kotonkan) to the rabies serogroup. *Infect. Immun*. 12(5):1157-1172.
- Belotto, A., Leanes, L.F., Schneider, M.C., Tamayo, H., Correa, E. 2005. Overview of rabies in the Americas. *Virus Res*.111:5-12.
- Benmansour, A., Leblois, H., Coulon, P., Tuffereay, C., Gaudin, Y., Flamand, A. 1991. Antigenicity of rabies virus glycoprotein. *J Virol*. 65:4198-4203.
- Bennett, S.N., Holmes, E.C., Chirivella, M., Rodriguez, D.M., Beltran, M., Vorndam, V., Gubler, D.J., McMillan, W.O. 2003. Selection-driven evolution of emergent dengue virus. *Mol Biol Evol*. 20(10):1650-1658.
- Beran, G.W. 1981. Rabies and infections by rabies-related viruses. In: *Handbook series in Zoonosis, Section B: Viral Zoonoses.* Ed. Beran, G.W. pp.57-135. Boca Raton FL: CRC press.
- Biek, R., Drummond, A. J. & Poss, M. 2006. A virus reveals population structure and recent demographic history of its carnivore host. *Science* 311:538–541.
- Billings, F.S. 1884. The relation of Animal Diseases to the Health, The dog: rabies of the Dog, Phenomena of canine rabies, Prevention. Appleton, New York. p. 139
- Bingham, J., Foggin, C. M., Wandeler, A. I., Hill, F. W. G. 1999. The epidemiology of rabies in Zimbabwe 1. Rabies in dogs (*Canis familiaris*). *Onderstepoort J.Vet.Res.*66 (1):1-10
- Bingham, J., Javangwe, S., Sabeta, C.T., Wandeler, A.I., Nel, L.H. 2001. Report of isolations of unusual lyssaviruses (rabies and Mokola virus) identified retrospectively from Zimbabwe. *Tyskr. S. Afr. Vet. Ver*. 72(2): 92-94.
- Bishop, G.C., Durrheim, D.N., Kloeck, P.E., Godlonton, J.D., Bingham, J., Speare, J. 2003. Rabies: Guide for the medical, veterinary and allied professions. Rabies Advisory Group, South African Department of Agriculture and Health.
- Botvinkin, A.D., Poleschuk, E.M., Kuzmin, I.V., Borisova, T.I., Gazaryan, S.V., Yager, P., Rupprecht, C.E. 2003. Novel lyssavirus isolated from bats Russia. *Emerg. Infect. Dis*. 9:1623-1625.
- Bouchardat, A. 1852. Divers remendes sur la rage. *Bull. Acad. Med. Paris* 18:6.
- Bouchardat, A. 1854. Remendes proposes contre la rage. *Bull Acad. Med. Paris* 20:714.
- Bouffard, G. 1912. Sur l'existence de la rage dans le Haut-Senégal et le Niger. *Ann. Inst. Pasteur* 26:727–731.
- Boulger l.R., Potterfield, J.S. 1958. Isolation of a virus from Nigerian fruit bats. *Trans.R.Soc.Trop.Med.Hyg*. 52:421-424.

- Bourhy, H., Tordo, N., Lafon, M., Sureau, P. 1989. Complete cloning and organization of rabies-related virus, Mokola virus. *J. Gen.Virol.* 70:2063-2074.
- Bourhy H., Kissi B., Lafon M., Sacramento D. 1992. Antigenic and Molecular Characterization of Bat Rabies Virus in Europe. J. *Clin.Microbiol.* 30 : 2419-2426
- Bourhy, H., Kissi, B., Tordo, N. 1993. Molecular diversity of the lyssavirus genus. *Virology* 194: 70-81.
- Bourhy, H., Kissi, B., Tordo, N., Badrane, H., Sacramento, D. 1995. Molecular epidemiological tools and phylogenetic analysis of bacteria and viruses with special emphasis on lyssaviruses. *Prevent. Vet. Med*. 25: 161-181.
- Bourhy, H., Kissi, B., Audry, L., Smreczak, M., Madkowska-Todys, M., Kulonen, K., Tordo, N., Zmudzinski, J.F., Holmes, E.C. 1999. Ecology and evolution of rabies virus in Europe. *J Gen Virol*.80:2545-2557.
- Bourhy, H., Cowley, J.A., Larrous, F., Holmes, E.C., Walker, P.J. 2005. Phylogenetic relationships among rhabdoviruses inferred using the L polymerase gene. *J Gen Virol*.86:2849-2858.
- Bourhy, H., Reynes, J., Dunham, E.J., Dacheux, L., Larrous, F., Huong, V.T.Q., Xu, G., Yan, J., Mirandas, M.E.G., Holmes, E.C. 2008. The origin and phylogeography of dog rabies virus. J. Gen. Virol.89:2673-2681
- Buckley, S.M. 1973. Singh's *Aedes albopictus* cell cultures as helper cells for the adaptation of Obodhiang and Kotokan viruses of the rabies serogroup to some vertebrate cell cultures. *Appl. Microbiol*. 25(4): 695-696.
- Buckley, S.M. 1975. Arbovirus infection of vertebrate and insect cell cultures, with special emphasis on Mokola, Obodhiang and Kotonkan viruses of the rabies serogroup. *An. NY Acad.Sci*. 266:241-250.
- Chaparro, F., Esterhuysen, J.J. 1993. The role of the yellow mongoose (*Cynictis penicillata)* in the epidemiology of rabies in South Africa- preliminary results. *Onderstepoort J. Vet.Res*.60(4):373-377.
- Chenik, M., Schnell, M., Conzelmann, K.K., Blondel, D. 1998. Mapping the interacting domain between rabies virus polymerase and phosphoprotein. *J Virol*.72:1925-1930.
- Chenik, M., Chebli, K., Gaudin, Y., Blondel, D. 1994. In vivo interaction of rabies virus phosphoprotein (P) and nucleoprotein (N), existence of two N binding sites on P protein. *J. Gen. Virol*.75:2889-2896.
- Childs, E.C. 2002. 'Epidemiology'. In *Rabies.* Jackson, A.C., Wunner, W.H.(Eds).Elsevier Science. London pp. 113-162.
- Cluver, E. 1927. Rabies in South Africa. *J.Med.Assoc.S.A.* 1:247-253.
- Coetzee, P., Nel, L.H. 2007. Emerging epidemic dog rabies in coastal South Africa: A molecular epidemiological analysis. *Virus Res*. 126(1-2):186-195.
- Conzelmann, K. 1998. Nonsegmented Negative-strand RNA viruses: Genetics and manipulation of the viral genomes. *An Rev Gen*. 32:123-162
- Crick, J., Tignor, G.H., Moreno, K. 1982. A new isolate of Lagos Bat virus from the Republic of South Africa. *Trans.R.Soc.Trop.Med.Hyg*. 76: 211-213.
- Davis, P. L., Holmes, E.C., Larrous, F., Van der Poel, W. H., Tjornehoj, K., Alonso, W. J. & Bourhy, H. 2005. Phylogeography, population dynamics, and molecular evolution of European bat lyssaviruses. *J Virol* 79:10487–10497.
- Davis, P.L., Bourhy, H., Holmes, E.C. 2006. The evolutionary history and dynamics of bat rabies virus. *Infect. Genet. Evol*. 6:464-473.
- Davis, P.L., Rambaut, A., Bourhy, H., Holmes, E.C. 2007.The evolutionary dynamics of canid and mongoose rabies virus in southern Africa. Arch.Virology.152(7):1251-1258

- Dean, D.J., Abelseth, M.K., Atanasiu, P. 1996. The fluorescent antibody test. In: *Laboratory techniques in rabies.* F.X. Meslin, M.M. Kaplan and H. Koprowski (eds). World Health Organization, Geneva. p88-95.
- Dietzschold, B., Wunner, W.H., Wiktor, T.J., Lopes, A.D., Lafon, M., Smith, C.L., Koprowski, H. 1983.Characterization of an antigenic determinant of the glycoprotein that correlates with pathogenicity of rabies virus. *Proc Natl Acad Sci* USA. 80:70-74.
- Dietzschold, B., Wiktor, T.J., Trojanowski, J.Q., Macfarlan, R.I., Wunner, W.H., Torres-Anjel, M.J., Koprowski, H. 1985. Differences in cell-to-cell spread of pathogenic and apathogenic rabies virus in vivo and in vitro. *J. Virol*. 56:12-18.
- Dietzschold, B., Gore, M., Marchadier, D., Niu, H.S., Bunschoten, H.S., Otvos, L., Wunner, W.H., Ertl, H.J.C., Osterhaus, A.D.M.E., Koprowski, H. 1990. Structural and immunological characterization of a linear virus-neutralizing epitope of the rabies virus glycoprotein and its possible use in a synthetic vaccine. *J Virol*. 64:3804-3809.
- Drummond, A.J., Nicholis, G.K., Rodrigo, A.G., Solomon, W. 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* 161:1307-1320.
- Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A. 2006. Relaxed phyloegenetics and dating with confidence.*PLoS Biology* 4:e88
- Drummond, A.J., Rambaut, A. 2007. "BEAST: Bayesian evolutionary analysis by sampling trees." *BMC Evol.Biol*. **7**:214.
- East, M.L., Hofer, H. 1996. Wild dogs in the Serengeti. *Science* 271:275-276.
- Estes, R.D. 1995. The behavior guide to African mammals. Halfway house, Russel Freidman, South Africa.pp.278-322
- Everard, C.O., Everard, J.D. 1988. Mongoose rabies. *Rev Infect Dis*.10:610-614.
- Faber, M., Pulmanausahakul R., Nagao K., Prosniak M., Rice A.B., Koprowski H., Schnell M.J., Dietzschold, B. 2004. Identification of viral genomic elements responsible for rabies virus neuroinvasiveness. *Proc. Natl .Acad. Sci. USA* 101:16328–16332.
- Familusi, J.B., Osunkoya, B.O., Moore, D.L., Kemp, G.E., Fabiyi, A. 1972. A fatal infection with Mokola virus*. Am J Trop Med Hyg*. 21(6):959-963.
- Familusi, J.B., Moore, D.L. 1972. Isolation of a rabies related virus from the cerebrospinal fluid of a child with “aseptic meningitis”. *Afr. J. Med. Sci*.3:93-96.
- Favi, M., De Mattos, C.A., Yung, J., Chala, E., Lopez, L.R., De Mattos, C.C. 2002.Fisrt case of human rabies in Chile caused by an insectivorous bat virus variant. *Emerg. Infect. Dis*.8:79-81.
- Felsenstein, J. 1984. Distance methods for inferring phylogenies: a justification. *Evolution*, 39:783-791.
- Felsenstein, J. 1985. Confidence intervals on phylogenies: an approach using the bootstrap. *Evolution* 39(4):783-791.
- Felsenstein, J. 1988. Phylogenies from molecular sequences: inference and reliability. *An.Rev.Gen*. 22: 521-565.
- Fermi, C. 1908. Uber die Immunisierung gegen Wutkrankheit. Z Hyg Infektionskr.58:233.
- Finke, S., Conzelmann, K. 2003.Dissociation of rabies virus matrix protein functions in regulation of viral RNA synthesis and virus assembly. *J Virol*.22:12074-12082.
- Finnegan, C.J., Brookes, S.M., Johnson, N.J., Smith, J., Mansfield, K.L., Keene, V.L., McElhinney, L.M., Fooks, A.R. 2002. Rabies in North America and Europe. *J R Soc Med*.95:9-13.
- Flemming, G. 1872. *Rabies and hydrophobia*. Chapmann and Hall, London p. 7
- Foggin, C.M. 1982. Atypical rabies virus in cats and a dog in Zimbabwe. *Vet. Rec*.110: 338.

- Foggin, C.M. 1983. Mokola virus infection in cats and a dog in Zimbabwe. *Vet Rec.*113:115.
- Foggin, C.M. 1988.Rabies and rabies-related viruses in Zimbabwe: Historical, virological and ecological aspects. PhD thesis. University of Zimbabwe. Harare.
- Fooks, A.R., Brookes, S.M., Johnson, N., McElhinney, L.M., Hutson, A.M. 2003. European bat lyssaviruses: An emerging zoonosis. *Epidemiol Infect*.131:1029-1039.
- Fraser, G.C., Hooper, P.T., Lunt, R.A., Gould, A.R., Gleeson, L.J., Hyatt, A.D., Russell, G.M., Kattenbelt, J.A. 1996. Encephalitis caused by a Lyssavirus in fruit bats in Australia. *Emerg. Infect. Dis*.2(4):327-331.
- Gaubert, P., Cordeiro-Estrela, P. 2006. Phylogenetic systematics and tempo of evolution of the Viverrinae (Mammalia, Carnivora, Vivveridae) within feliformians: Implications for faunal exchange between Asia and Africa. *Mol.Phylo.Evol*.41:266-278.
- Gould, A.R., Hyatt, A.D., Lunt, R., Kattenbelt, J.A., Hengstberger, S., Blacksell, S.D. 1998.Characterisation of a novel lyssavirus isolated from pteropid bats in Australia. *Virus Res*.54:165-187.
- Gregory, W.K., Hellman, H. 1939. On the evolution and major classification of the civets (Viverridae) and allied fossil and recent Carnivora: a phylogenetic study of the skull and dentition. *Proc. Am. Phil. Soc.* 81:309–392.
- Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L., Daly, J. M., Mumford, J. A. & Holmes, E. C. 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303:327–332.
- Hall, T.A. 1999.BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids.Symp.Ser*.41:95-98.
- Harris, S.L., Brookes, S.M., Jones, G., Hutson, A.M., Raceya, P.A., Aegerter, J., Smith, G.C., McElhinney, L.M., Fooks, A.R. 2006. European bat lyssaviruses: Distribution, prevalence and implications for conservation. *Biol Conserv*.131:193-210.
- Hertzenberg, L. 1928. Two cases of hydrophobia. *J. Med. Assoc. South Afr*. 2:659–661
- Higgins, D. 2003. Multiple alignment. In: *The Phylogenetic Handbook: a practical approach top DNA and protein phylogeny*. M. Salemi and A .Vandamme(Eds), Cambridge University Press, London,pp.45-69.
- Hoffmeyer, M., Hoffmeyer, D., Nel, L., Bingham, J.2004. A second outbreak of rabies in African wild dogs (*Lycaon pictus*) in Madikwe Game Reserve, South Africa, demonstrating the efficacy of vaccination against natural rabies challenge, *Animal Conserv.* **7**:193–198.
- Hogenhout, S.A., Redinbaugh, M.G., Ammar, E. 2003. Plant and animal rhabdovirus host range: a bug's view. *Trends Microbiol*. 11(6):254-271.
- Holder,M. and Lewis, P.O. 2003. Phylogeny estimation: Traditional and Bayesian approaches. *Nature Rev.Gen*. 4:275-284.
- Holmes, E.C., Woelk, C.H., Kassis, R., Bourhy, H. 2002. Genetic Constraints and the Adaptive Evolution of Rabies Virus in Nature. *Virology* 292:247-257.
- Holmes, E.C. 2003. Molecular Clocks and the Puzzle of RNA Virus Origins. *J. Viro*l.77(7):3893-3897
- Hooper P.T., Lunt R.A., Gould A.R., Samaratunga H., Hyatt A.D., Gleeson L.J., Rodwell B.J., Rupprecht C.E., Smith J.S., Murray P.K. 1997. A new lyssavirus - the first endemic rabies-related virus recognized in Australia. *Bull. Inst. Pasteur* 95:209-218.
- Hudson, J.R. 1944. A short history of rabies in Kenya. *East Afr. Med. J.* 21:322-327.
- Hughes, G.J. 2008. A reassessment of the emergence time of European bat lyssavirus type1. *Infect. Genet. Evol.* In press.
- Hutcheon, D. 1894. Reports of the Colonial Veterinary Surgeon for the year 1893. Cape of Good Hope, Dept. of Agric. pp 7-10.

- Hyun, B.H., Lee, K.K., Kim, I.J., Lee, K.W., Park, H.J., Lee, O.S., An, S.H., Lee, J.B.2005. Molecular epidemiology of rabies virus isolates from South Korea. *Virus Res*.114:113–25
- Institute Pasteur. 1985.Isolations of Lagos bat virus in West Africa. *Internal reports of Centre Collaborateur OMS de Reference et Recherche Pour les Arbovirus*. Dakar (Senegal). Available from http://www.pasteur.fr/recherche/banques/CRORA/bibref/br01600.htm and http://www.pasteur.fr/recherche/banques/CRORA/bibref/br01610.
- Irvin, A.D. 1970. The epidemiology of wildlife rabies. *The Vet. Record: J.British Vet.Ass.* 87:338-348.
- Jackson, A.C. 1991. Biological basis of rabies virus neurovirulence in mice: Comparative pathogenesis study using the immunoperoxidase technique. *J.Virol*. 65:537-540.
- Jaftha, J.B. 1997. Molecular genetic analysis of some enzootic rabies viruses in Southern Africa. MSc. Thesis. University of Pretoria.
- Johnson, N., McElhinney, L.M., Smith, J., Lowings, A.R. 2002. Phyloegenetic comparison of the genus Lyssavirus using the distal coding sequences of the glycoprotein and nucleoprotein genes. *Arch. Virol.* 147(11):2111-2123.
- Johnson, D., Selden, D., Parson, G., Healy, D., Brookes, S.M., McElhinney, L.M., Hutson, A.M., Fooks, A.R. 2003 .Isolation of a European bat lyssavirus type 2 from a Daubenton's bat in the United Kingdom. *Vet Rec*.383-387.
- Johnson, N., McElhinney, L.M., Ali, Y.H., Saeed, I.K., Fooks, A.R. 2004a. Molecular epidemiology of canid rabies in Sudan: evidence for a common origin of rabies with Ethiopia. *Virus Res*. 104(2): 201-205.
- Johnson, N. Letshwenyo, M., Baipoledi, E.K., Thobokwe, G., Fooks, A.R. 2004b. Molecular epidemiology of rabies in Botswana: a comparison between antibody typing and nucleotide sequence phylogeny. *Vet. Microbiol*. 101(1):31-38.
- Jukes, T. and Cantor, C.R. 1969. Evolution of protein molecules. In: *Mammalian protein metabolism.* H.N. Munro (Eds) New York, Academic Press. pp. 21-132.
- Kemp, G.E., Causey, O.R., Moore, D.L., Odelola, A., Fabiyi, A. 1972. Mokola virus: Further studies on IbAn 27377, a new rabies-related etiological agent of zoonosis in Nigeria. *Am.J.Trop.Med.Hyg*.21:356-359.
- Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J.Mol.Evol*.15:111-120.
- King A.A, Crick J. Rabies-related viruses.1988. In: *Rabies*. Campbell, J.B. and Charlton, K.M., (eds). Boston: Kluwer academic Publishers.
- King, A.A., Davies, P., Lawrie, A. 1990. The rabies viruses of bats. *Vet. Microbiol*. 23:165-174.
- King, A.A., Meredith, C.D., Thomson, G.R. 1993a. Canid and viverrid viruses in South Africa. *Onderstepoort J.Vet.Res*. 60:295-299.
- King, A.A. 1993b. Monoclonal antibody studies on rabies-related viruses. *Onderstepoort J.Vet.Res*. 60:283-287.
- King, A.A., Meredith,C.D., Thomson, G.R. 1994. The biology of southern African lyssavirus variants. In *Lyssaviruses*. Rupprecht, C.E., Dietzchold, B., Koprowski, H. (Eds) Springer-Verlag, Berlin, Heidelberg. pp 267-295
- Kissi, B., Tordo, N., Bourhy, H. 1995. Genetic polymorphism in the rabies virus nucleoprotein gene. *Virology*.209:526-537.
- Kobayashi, Y., Sato, G., Shoji, Y., Sato, T., Itou, T., Cunha, E.M., Samara, S.I., Carvalho, A.A., Nociti, D.P., Ito, F.H., Sakai, T.2005. Molecular epidemiological analysis of bat rabies viruses in Brazil. *J. Vet. Med. Sci*.67:647-652.
- Koprowski, H. 1996. The mouse inoculation test. In: *Laboratory techniques in rabies*. F.X. Meslin, M.M. Kaplan and H.Koprowski (eds). World Health Organization, Geneva.p80-87.

- Krugelstein, F. 1826. Die Geschichte der Hundswuth und der Wasserscheu und deren Bahndlung. *Gotha.*
- Kumar, S., Tamura, K., Nei , M. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment, *Brief. Bioinform.* 5:150–163.
- Kuzmin, I.V., Orciari, L.A., Arai, Y.T., Smith, J.S., Hanlon, C.A., Kameoka, Y., Rupprecht, C.E. 2003.Bat lyssaviruses (Aravan and Khujand) from Central Asia: phylogenetic relationships according to N, P and G gene sequences. *Virus Res*.91: 65-79.
- Kuzmin, I.V., Hughes, G.J., Botvinkin, A.D., Orciari, L.A., Rupprecht, C.E. 2005. Phylogenetic relationships of Irkut and West Caucasian bat viruses within the lyssavirus genus and suggested quantitative criteria based on the N gene sequence for lyssavirus genotype definition. *Virus Res*.111:28-43.
- Kuzmin, I.V., Hughes, G.J., Rupprecht, C.E. 2006. Phylogenetic relationships of seven previously unclassified viruses within the family Rhabdoviridae using partial nucleoprotein gene sequences. *J Gen Virol*. 87:2323-2331
- Kuzmin, I.V., Niezgoda, M., Franka, R., Agwanda, B., Markotter, W., Beagley, J.C., Urazova,O.Y., Breiman, R.F.,Rupprecht, C.E. 2008. Lagos bat virus in Kenya. *J.Clin. Micro* 46(4):1451-1461
- Lafon, M., Wiktor, T.J., Macfarlan, R.I. 1983. Antigenic sites on the CVS rabies virus glycoprotein: Analysis with monoclonal antibodies. *J Gen Virol*.64:843-851.
- Le Gonidec, G.L., Rickenbach, A., Robin, Y., Heme, G. 1978. Isolement d'une souche de virus Mokola au Cameroun. *Annales Microbiologie (Institute Pasteur)* 129A:245-249.
- Le Mercier, P.L., Jacob, Y., Tordo, N. 1997. The complete Mokola virus genome sequence: structure of the RNA-dependant RNA polymerase. *J. Gen. Virol*. 78:1571-1576.
- Leslie, M.J., Messenger, S., Rohde, R.E., Smith, J., Cheshier, R., Hanlon, C., Rupprecht, C.E. 2006. Bat-associated rabies virus in Skunks. *Emerg Infect Dis*. 12:1274-1277.
- Li, W., Graur, D. 2000. *Fundamentals of Molecular Evolution*. Second Edition. Sinauer Associates Inc.
- Mansfield, K.L., Johnson, N., Fooks, A.R.2004. Identification of a conserved linear epitope at the N terminus of the rabies virus glycoprotein. *J Gen Virol*.85:3279–3283.
- Mansfield K., McElhinney L., Hübschle O., Mettler F., Sabeta C., Nel L.H., Fooks A.R. 2006. A molecular epidemiological study of rabies epizootics in kudu (*Tragelaphus strepsiceros*) in Namibia. *BMC Vet.Res* 2:2
- Markotter, W., Randles, J., Rupprecht, C.E., Sabeta, C.T., Wandeler, A.I., Taylor, P.J., Nel, L.H. 2006a. Recent Lagos bat virus isolations from bats (suborder Megachiroptera) in South Africa. *Emerg Infect Dis*.12:504-506.
- Markotter W, Kuzmin I, Rupprecht CE, Randles J, Sabeta CT, Wandeler AI, Nel LH. 2006b. Isolation of Lagos bat virus from water mongoose. *Emerg Infect Dis*.12:1913-1918.
- Markotter, W. 2007.Molecular epidemiology and pathogenesis of Lagos bat virus, a rabies-related virus specific to Africa. Ph. D. dissertation, University of Pretoria, Pretoria, South Africa.
- Markotter, W., Kuzmin, I., Rupprecht, C.E., Nel, L.H. 2008a.Phylogeny of Lagos bat virus: Challenges for lyssavirus taxonomy. *Virus Res*. 135:10-21
- Markotter, W., Van Eeden, C., Kuzmin, I., Rupprecht, C.E., Paeska, J.T., Swanepoel, R., Fooks, A.R., Sabeta, C.T., Cliquet, F., Nel, L.H. 2008b. Epidemiology and pathogenicity of African bat lyssaviruses. *Dev. Biol. Basel. Krager* 131:972-977.
- Mebatsion, T., Cox, J.H., Frost, J.W. 1992. Isolation and characterization of 115 street rabies virus isolates from Ethiopia by using monoclonal antibodies: identification of 2 isolates as Mokola and Lagos bat viruses. *J Infect Dis*.166:972–977.
- Meredith, C.D., Standing, E.1988. Lagos Bat virus in South Africa. *The Lancet*.11:832-833.

- Meredith, C.D., Nel, L.H., Von Teichman, B.F. 1996. Further isolation of Mokola virus in South Africa. *The Vet. Record*, February 3:119-120.
- Moya, A., Holmes, E.C., Gonzales-Candelas, F. 2004. The population genetics and evolutionary epidemiology of RNA viruses. *Nature Rev: Microbiol*. 2:279-288.
- Müller, T., Cox, J., Peter, W., Scafer, R., Bodamer, P., Wulle, U., Burow, J., Muller, W. 2001. Infection of a stone marten with European Bat Lyssavirus (EBL1). *Rabies Bull Eur*.25:9–11.
- Nadin-Davis, S.A., Huang, W., Armstrong, J., Casey, G.A., Bahloul, C., Tordo, N., Wandeler, A.I. 2001. Antigenic and genetic divergence of rabies viruses from bat species indigenous to Canada. *Virus Res*. 74:139-156.
- Nadin-Davis, S.A., Simani, S., Armstrong, J., Fayaz, A., Wandeler, A.I. 2003. Molecular and antigenic characterization of rabies viruses from Iran identifies variants with distinct epidemiological origins. *Epidemiol. Infections* 131:777-790.
- Nadin-Davis, S.A., Bingham, J. 2004. Europe as a source of rabies for the rest of the world. In *Historical perspectives of rabies in Europe and the Mediterranean basin*. Ed. King, A.A., Fooks, A.R., Aubert, M, Wandeler, A.I. OIE. Pp.259-280.
- Nadin-Davis,S.A., Turner, G., Paul, J.P.V., Madhusudana, S.N., Wandeler, A.I. 2007. Emergence of arctic-like rabies lineage in India. *Emerg. Infect. Dis*. 13(1):111-116.
- Negri, A. 1903. Zur Aetiologie der Tollwuth: Die Diagnose der Tollwuth auf Grund der neueren Befunde. Z Hyg Infektkrankh. 44:519-540.
- Nel, L.H., Thomson, G.R., Von Teichman, B.F. 1993. Molecular epidemiology of rabies virus in South Africa. *Onderstepoort J Vet Res*.60:301-306.
- Nel, L.H., Bingham, J., Jacobs, J.A., Jaftha. 1998. A nucleotide specific polymerase chain reaction assay to differentiate rabies virus biotype sin South Africa. *Onderstepoort J vet res*.65:297-303
- Nel, L.H., Jacobs, J., Jaftha, J., Von Teichman, B., Bingham, J. 2000. New cases of Mokola virus infection in South Africa: A genotypic comparison of southern African virus isolates. *Virus Genes* 20(2):103-106.
- Nel, L.H., Niezgoda, M., Hanlon, C.A., Morril, P.A., Yager, P.A., Rupprecht, C.E. 2003. A comparison of DNA vaccines for the rabies-related virus, Mokola. *Vaccine*. 21(19-20):2598-2606.
- Nel, L.H., Sabeta, C.T., Von teichman, B., Jaftha, J.B., Rupprecht, C.E., Bingham, J. 2005. Mongoose rabies in southern Africa: a re-evaluation based on molecular epidemiology. *Virus Res*. 109:165-173.
- Nel, L.H. and Rupprecht, C.E. 2007. The Emergence of Lyssaviruses in the Old World. In *Wildlife and Emerging Zoonotic disease: The biology, circumstances and consequences of cross- species transmission.* Childs J.E., Mackenzie, J.S. Richt J.A. (Eds) Springer. Pp 161-193.
- Nel, L.H., Markotter, W. 2007. Lyssaviruses. *Critic. Rev. Microbiol*. 33:1-24.
- Nicolau, S., Mathis, C., Constantinesco, V. 1933. La rage autochtone (maladie de chien fou) en Afrique Oceidentale Francaise. (Etude critique et experimentale.) *Ann. Inst. Pasteur*. Paris. 50,778-839.
- Olivier, M. 1997. Rabies associated with wildlife in the family Viverridae in South Africa: molecular analysis of the virus Nucleoprotein and Glycoprotein encoding gene. MSc. Thesis, University of Pretoria, South Africa.
- Paez, A., Nunez, C., Garcia, C., Boshell, J. 2003. Epidemiologia molecular de epizootias de rabia en Colombia, 1994-2002: evidencia de rabia humana y canina asociada a quiropteros. *Biomedica* 23:19-30.
- Pasteur, L. 1881. Sur ene maladie nouvelle, provoquee par la salive d'un enfant mort de la rage. *C.R. Acad. Sci* 92:159.

- Pasteur, L. , Chamberland, C.E., Roux, M. 1884. Nouvelle communication sur la rage. *C.R. Acad. Sci*. 98:457.
- Pasteur, L. 1885. Methode pour prevenir la rage après morsure. *C.R.Acad.Sci*. 101:765.
- Pasteur, L. 1886. Resultas de l'application de la method pour previnir la rage après morsure. *C.R.Acad.Sci*.102:459
- Paweska, J.T., Blumberg, L., Liebenberg, C., Hewlett, R.H., Grobbelaar, A.A., Leman, P.A., Croft, J.E., Nel, L.H., Nutt, L., Swanepoel, R. 2006. Fatal human infection with rabies-related Duvenhage virus, South Africa. *Emerg Infect Dis*.12:1965-1967.
- Poch, O., Blumberg, B.M., Bougueleret, L., Tordo, N. 1990. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: Theoretical assignment of functional domains. *J Gen Viro*l.71:1153-1162.
- Posada, D., Crandall, K.A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14 (9): 817-818.
- Prehaud, C., Coulon, P., Lafay, F., Thiers, C., Flamand, A.1988. Antigenic site II of the rabies virus glycoprotein: Structure and role in viral virulence. *J Virol*.62:1-7.
- Pringle, C.R., Easton, A.J. 1997. Monopartite negative strand RNA genomes. *Seminars Viro*l. 8(1):49-57.
- Pulmanausahakul, R., Li, J., Schnell, M.J., Dietzschold, B. 2008. The Glycoprotein and the matrix protein of rabies virus affect pathogenicity by regulating viral replication and facilitating cell-to-cell spread. *J. Virol*. 82(5): 2330–2338.
- Rambaut, A., Pybus, O.G., Nelson, M.I., Viboud, C., Taubenberger, J.K., Holmes, E.C. 2008. The genomic and epidemiological dynamics of human influenza A virus. *Nature* 453:615-619.
- Randall, D.A., Williams, S.D., Kuzmin, I.V., Rupprecht, C.E., Tallents, L.A., Tefera, Z., Argaw, K., Shiferaw, F., Knobel, D.L., Sillero- Zubiri, C., Laurenson, K. 2004.Rabies in endangered Ethiopian wolves. *Emerging Infect Dis*.10:2214-2217.
- Reagan, K.J., Wunner, W.H.1985. Rabies virus interaction with various cell lines is independent of the acetylcholine receptor. *Arch Virol*.84:277-282.
- Real, L.A., Henderson, J.C., Blek, R., Snaman, J., Jack, T.L., Childs, J.E., Stahl, E., Waller, L., Tinline, R., Nadin-Davis, S.A. 2005. Unifying the spatial population dynamics and molecular evolution of epidemic rabies virus. *PNAS*. 102(34):12107-12111.
- Remlinger, P. 1903. Le passage du virus rabique a travers les filters. *Ann Inst Pasteur Paris*.17:834-849.
- Rossatte, R.C., Power, M.J., Donovan, D., Davies, J.C., Allan, M. Bachmann, P., Stevenson, B., Wandeler, A., Muldoon, F. 2007. Elimination of arctic variant rabies in red foxes, Metropolitan Toronto. *Emerg. Infect. Dis*. 13:25-27.
- Ronsholt, L., Sorensen, K.J., Bruschke, C.J.M., Wellenberg, G.J., Van Oirschot, J.J., Johnstone, P., Whitby, J.E., Bourhy, H. 1998.Clinicilly silent rabies infection in (zoo) bats. *The Vet. Rec*. 142:519-520.
- Ronsholt L. A new case of European Bat Lyssavirus (EBL) infection in Danish sheep. 2002. *Rabies Bull Eur*.2:15.
- Sabeta, C.T. 2002. Aspects of the molecular epidemiology of rabies in Zimbabwe and South Africa. PhD. Dissertation, University of Pretoria, South Africa.
- Sabeta, C.T., Markotter, W., Mohale, D.K., Shumba, W., Wandeler, A.I., Nel, L.H. 2007. Mokola virus in domestic mammals, South Africa. *Emerg. Infect. Dis*. 13(9):1371-1373
- Sacramento, D., Bourhy, H., Tordo, N. 1991. PCR technique as an alternative method for diagnosis and molecular epidemiology of rabies virus. *Mol. Cell. Probes* 5:229-240.
- Saitou, N., Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol.Biol.Evol*. 4:406-425.

- Saluzzo JF, Rollin PE, Dauguet C, Digoutte JP, Georges AJ, Sureau P. 1984. Premier isolment du virus Mokola a partir D'un rongeur (*Lophuromys sikapusi*). *Ann Virol*.135:57-66.
- Schneider, L.G., Dietzschold, B., Dierks, R.E., Matthaeus, W., Enzmenn, P.J., Strohmaier, K. 1973. Rabies group-specific ribonucleoprotein antigen and a test system for grouping and typing of rhabdoviruses. *J. Virol*. 11:784-755.
- Schneider, L.G., Barnard, B.J.H., Schneider, H.P. 1985. Application of monoclonal antibodies for epidemiological investigations and oral vaccination studies. 1. African viruses. In.Kuwert, E., Merieux, C., Koprowski, H., Bogel, K. Springer-Verlag. Rabies in the Tropics pp. 47-59.
- Seganti, L., Superti, F., Bianchi, S., Orsi, N., Divizia, M., Pana, A. 1990. Susceptibility of mammalian, avian, fish, and mosquito cell lines to rabies infection. *Arch Virol*.34:155-163.
- Seif, I., Coulon, P., Rollin, P.E., Flamand, A. 1985. Rabies virus virulence: Effect on pathogenicity and sequence characterization of mutations affecting antigenic site III of the glycoprotein. *J Virol*.53:926-935.
- Semple D. 1919. On the nature of rabies and antirabic treatment. BR. Med. J.2:333.
- Shapiro, B., Rambaut, A., Drummond, A.J. 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol*. 23(1):7-9.
- Shope R.E.1982. Rabies-related viruses. *The Yale Journal of Biology and Medicine* 55:271-275.
- Shope, R.E., Murphy, F.A., Harrison, K.A., Causey, O.R., Kemp, G.E., Simpson, D.I.H., Moore, D.L. 1970. Two African viruses serologically and morphologically related to rabies virus. *J.Virol*. 6 (5):690-692.
- Smith, J.S. 1989. Rabies virus epitopic variation: use in ecological studies. Advances Virus Res. 36:215-253.
- Smith, J.S., Yager, P.A., Orciari, L.A. 1993. Rabies in wild and domestic carnivores of Africa: epidemiological and historical associations determined by limited sequence analysis. *Onderstepoort J.Vet. Res.* 60(4):307-314.
- Smith, J.S. 1996. New aspects of rabies with emphasis on epidemiology, diagnosis and prevention of the disease in the United States. *Clin Microbiol Rev*.9:166-176.
- Smithcors, J.F. 1958. The history of some current problems in animal disease. VII. Rabies (part 1). *Vet.Med.* 53:149
- Snyman, P.S. 1940. The study and control of vectors of rabies in South Africa. *Onderstepoort J. Vet.Sci. Animal Ind*. 15(1,2): 9-140.
- Steele, J.H. 1975. History of Rabies. In The Natural history of rabies. Ed. Baer, G.M. New York, Academic Press.pp.1-29.
- Steele, J.H. 1988. Rabies in the Americas and remarks on global aspects. *Rev. Infect. Dis* 10 Suppl 4:S585-597.
- Steele, J.H., Fernandez, P.J. 1991. History of rabies and global aspects in: *The natural history of rabies.* 2nd Edition.
- Strimmer, K., Von Haeseler, A. 2003. Nucleotide substitution models. In: *The Phylogenetic Handbook: A practical approach to DNA and protein phylogeny*. M. Salemi and A. Vandamme (Eds), Cambridge University Press, London, pp. 72-97.
- Strimmer, K., Robertson, D.L. 2005. Inference and Applications of Molecular Phylogenies: An Introductory Guide. In: Sansom, C., Horton, R.N. (eds.). *The Internet for Molecular Biologists (Practical Approach Series).* Oxford University Press, Oxford, UK.
- Sudershan, M., Madhusudana, S., Mahendra, B., Rao, N., Ashwath Narayana, D., Abdul Rahman, S., Meslin, F.,Lobo, D., Ravikumar, K. 2007. Assessing the burden of human rabies in India: results of a national multi-center epidemiological survey. *Internat. J. Infect. Dis*. 11(1): 29 – 35.

- Sureau, P., Germain, M., Herve, J.P,, Geoffrey, B., Cornet, J.P., Heme, G., Robin, Y. 1977.Isolement du virus Lagos bat en Empire Centrafricain. *Bull Soc Path Exo*t. 70:467-470.
- Swanepoel, R., Barnard, B.J.H., Meredith, C.D., Bishop, G.C., Bruckner, G.K., Foggin, C.M., Hubschle, O.J.B. 1993. Rabies in southern Africa. Onderstepoort J. Vet. Res.60:325-346.
- Swanepoel, R. 2004. Rabies. In: *Infectious diseases of livestock: with special reference to southern Africa*. Second edition / edited by J.A.W. Coetzer, and R.C. Tustin; Publisher Cape Town (South Africa). Oxford University Press.
- Swofford, D.L. and Sullivan, J. 2003. Phylogeny based on parsimony and other methods. In *The Phylogenetic Handbook: A practical approach to DNA and protein phylogeny*. M. Salemi and A. Vandamme (Eds), Cambridge University Press, London, pp. 160-206.
- Swofford, D.L. 2003. PAUP. Phylogenetic Analysis Using Parsimony (and other methods). Version 4. Sinauer Associates, Sunderland, Mass.
- Taylor, P.J. 1993. A systematic and population genetic approach to the rabies problem in the yellow mongoose (*Cynictis penicillata*). *Onderstepoort J.Vet.Res.*60(4):379-387.
- Thompson, J.D., Higgins, D.G., Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids. Res.* 22(22):4673-4680.
- Tordo, N., Poch, O., Ermine, A., Keith, G.1986a. Primary structure of leader RNA and nucleoprotein genes of the rabies genome: Segmented homology with VSV. *Nucleic Acids Res.*14:2671-2683.
- Tordo, N., Poch, O., Ermine, A., Keith, G., Rougeon, F. 1986b. Walking along the rabies genome: Is the large G-L intergenic region a remnant gene? *Proc. Natl. Acad. Sci. USA*.83:3914-3918.
- Tordo, N., Poch, O. 1988.Structure of rabies virus. In*: Rabies*. Campbell JB and Charlton KM, editors. Boston: Kluwer academic publishers.p25-45.
- Tordo, N., Poch, O., Ermine, A., Keith, G., Rougeon, F. 1988. Completion of the Rabies Virus genome sequence determination: Highly Conserved Domains among the L (polymerase) proteins of unsegmented negative-strand RNA viruses. *Virology*.165:565-576.
- Tordo, N., Bourhy, H., Sather, S., Ollo, R. 1993. Structure and Expression in Baculovirus of the Mokola Virus Glycoprotein: An Efficient Recombinant Vaccine. *Virology* 194(1):59-69.
- Tordo, N., Benmansour, A., Calisher, C., *et al.* 2005. Rhabdoviridae In: *Virus Taxonomy: The Classification and Nomenclature of Viruses.* The Eight Report of the International Committee on Taxonomy of Viruses. Edited by Fauquet C.M., Mayo, M.A., Maniloff, J., Desselberger, U. and Ball, L.A. London: Elsevier/ Academic Press :623-644.
- Tuffereau, C., Leblois, H., Benejean, J., Coulon, P., Lafay, F., Flamand, A.1989. Arginine or lysine in position 333 of ERA and CVS glycoprotein is necessary for rabies virulence in adult mice. *Virology*.172:206-212.
- Vandamme, A. 2003. Basic concepts in molecular evolution. In *The Phylogenetic Handbook: A practical approach to DNA and protein phylogeny.* M. Salemi and A. Vandamme (Eds), Cambridge Univerity Press, London. Pp.1-21.
- Van de Peer, Y. 1993. Phylogeny inference based on distance methods. In *The phylogenetic handbook : a practical approach to DNA and protein phylogeny.* M. Salemi and A. Vandamme (Eds), Cambridge Univerity Press, London.
- Van Thiel, P.P., van den Hoek, J.A., Eftimov, F., Tepaske, R., Zaaijer, H.J., Spanjaard, L., de Boer, H.E., van Doornum, G.J., Schutten, M., Osterhaus, A., Kager, P.A. 2008. Fatal case of human rabies (Duvenhage virus) from a bat in Kenya: The Netherlands, December 2007.*Euro. Surveill.*13(2)

- Velasco-Villa, A., Orciari, L.A., Souza, V., Juarez-Islas, V., Gomez-Sierra, M., Castillo, A., Flisser, A., Rupprecht, C.E. 2005 .Molecular epizootiology of rabies associated with terrestrial carnivores in Mexico. *Virus Res*.111:13-27.
- Veron, G, Colyn, M., Dunham, A.E., Taylor, P., Gaubert, P. 2004. Molecular systematics and origin of sociality in mongooses (Herpestidae, Carnivora). *Mol. Phylog. Evol.* 30:582-598.
- Von Teichman, B.F., De Koker, W.C., Bosch, S.J.E., Bishop, G.C., Meredith, C.D., Bingham, J. 1998.Mokola virus infection: description of recent South African cases and a review of the virus epidemiology. *J S Afr Vet Ass*.69:169-171.
- Von Teichman, B.F., Thomson, G.R., Meredith, C.D., Nel, L.H. 1995. Molecular epidemiology of rabies virus in South Africa: Evidence for two distinct virus groups. J Gen Virol.76:73–82.
- Warrilow, D. 2005. Australian bat lyssavirus: A recently discovered new rhabdovirus. *Curr. Top. Microbiol. Immunol*. 292:25-44.
- Wang, X.J., Huang, J.T., 2001. Epidemiology. In: Yu, Y.X. (Ed.), *Rabies and Rabies Vaccine. Chinese medicine technology*. Press, Beijing, pp. 127–144.
- WHO expert consultation on rabies. 2004. Technical report series 931. 1st report. World Health Organisation, Geneva.
- Wiktor, T.J., Clark, H.F. 1973. Application of the plaque assay technique to the study of rabies virus-neutralizing antibody interaction. *Ann Microbiol (Paris*).124:271-282.
- Wu, X., Franka, R., Velasco-Villa, A., Rupprecht, C.E. 2007. Are all lyssavirus genes equal for phylogenetic analyses? *Virus Res*. 129(2):91-103.
- Wozencraft, W.C. 1993. Carnivora: Herpestidae. In*: Mammal species of the world: a taxonomic and geographic reference*. 2nd Ed. Wilson, D.E., Reeder, D.M. Eds. Smithsonian Institution Press, Washington, London. P 1206.
- Wunner W H. 2002.Rabies Virus. In*, Rabies*. A. C. Jackson and W. H. Wunner (Eds), Elsevier Science, London. p23-77.
- Zanotto, P. M. D. A., M. J. Gibbs, E. A. Gould, and E. C. Holmes. 1996. A reassessment of the higher taxonomy of viruses based on RNA polymerases. *J. Virol*. 70:6083–6096.
- Zhang, Y.Z., Xiang, C., Ziao, Y., Wang, D., Jiang, R., Xiao, Q., Hao, Z.H., Zhang, L., Yu, Y., Fu, Z.F. 2006. Molecular characterization of rabies virus isolates in China during 2004. *Virus Res*. 121:179–188.
- Zinke, G.G. 1804. Neue Ansichten der Hundswuth ihrer Ursachen und Folgen, nebst einer sichern Behandlungsart der von tollen Tieren gebissenen Menschen, C.E. Gabler. *Jena Rev*. 16:212

Websites:

1. Department of Agriculture Archives , South African department of Agriculture, veterinary services www.nda.agric.za/vetweb/history
2. www.answers.com

Communications

Van Zyl, N., Markotter, W. and Nel, L.H. Molecular phylogeny of mongoose rabies in southern Africa. Molecular and Cell Biology Group symposium, Pretoria, South Africa, 17 October 2007.

Van Zyl, N., Markotter, W. and Nel, L.H. Evolutionary history of mongoose rabies in southern Africa. South and East Africa Rabies Group meeting, Gaborone, Botswana, 25-28 August 2008.