Evaluation of *Cardiovirus* reservoir potential in *Mastomys* from the Kruger National Park against a backdrop of host phylogeography and historical outbreak strain characterization

By

James Henry du Toit van Sandwyk

Submitted in partial fulfilment of the requirements for the degree of

Philosophiæ Doctor

(Zoology)

in the Faculty of Natural and Agricultural Science

University of Pretoria

Pretoria

February 2014

Evaluation of *Cardiovirus* reservoir potential in *Mastomys* from the Kruger National Park against a backdrop of host phylogeography and historical outbreak strain characterization

By

J.H.d.T. van Sandwyk

J.H.d.T van Sandwyk: Mammal Research Institute
Department of Zoology and Entomology
University of Pretoria
Pretoria
0002
South Africa

Supervisors:

Prof A.D.S. Bastos: Mammal Research Institute
Department of Zoology and Entomology
University of Pretoria
Pretoria
0002
South Africa

Prof N.C. Bennett: Mammal Research Institute
Department of Zoology and Entomology
University of Pretoria
Pretoria
0002
South Africa

Email: jvsandwyk@zoology.up.ac.za

DECLARATION

I, James Henry du Toit van Sandwyk, declare that this thesis which I hereby submit for the degree of Philosophiæ Doctor at the University of Pretoria is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Sig 

08 February 2014

Date

ACKNOWLEDGEMENTS

I would like to express my gratitude towards my supervisors Nigel Bennett and Armanda Bastos for their help with the various aspects of my research project. I wish to thank Nigel for all his help with funding for the project, and Armanda for her help with the molecular component of the work and for providing me with the samples. Thank you to both of you for providing me with a solid foundation for my future pursuits in the field of science.

Babsie Potgieter is thanked for all her help with the technical aspects of my project.

Most of all I wish to especially thank my loving wife Elene van Sandwyk for all her help and support in my various endeavours and for her undying love.

"Omnia possum in eo qui me confortat"

TABLE OF CONTENTS

APPENDICIES

LIST OF FIGURES

LIST OF TABLES

DISCLAIMER

This thesis comprises a series of chapters, some of which have been prepared as stand alone manuscripts that will subsequently be submitted for publication purposes. As a consequence there may be some unavoidable repetition between chapters.

CHAPTER 1

General Introduction

1.1. Introduction

1.1.1. Discovery of Cardioviruses

The genus *Cardiovirus* comprises two currently recognized species, namely *Encephalomyocarditis virus (EMCV),* and *Theilovirus (THV)* (Knowles *et al.*, 2012). The first cardioviruses were discovered by Jungeblut and Sanders in the early 1940's when attempting to transfer the Yale-SK poliovirus from monkeys to mice, using cotton rats as an intermediate host and following the protocol developed by Armstrong (1939) for the transfer of Lansing strain (type 2) poliomyelitis virus. After it was discovered that the transfer had not been successful and a previously unknown pathogen had in fact been transferred, it was named Columbia-SK. Further attempts by Jungeblut and Dalldorf in 1943 produced similar results and direct transfer of cerebral tissue from infected mice to uninfected monkeys and hamsters proved to be fatal (Scraba & Palmenberg, 1999). Pleural fluids recovered from two species of primates at the Miami Zoo that had succumbed to an unknown disease in 1945, led to the discovery of the causative agent, a virus which was named Encephalomyocarditis virus by Helwig and Schmidt (1945). The Mengovirus was first isolated by Dick, Smithburn and Haddow in 1948 from the Mengo district of Uganda. The virus was isolated by inoculating mice inter-cerebrally with infected blood or spinal fluid from a rhesus monkey which had developed lower limb paralysis (Scraba & Palmenberg, 1999). Mengovirus is serologically indistinguishable from Encephalomyocarditis virus.

1.1.2. Classification

Cardioviruses are RNA viruses with a positive sense, single-stranded RNA (ssRNA) genome, that belong to the family *Picornaviridae*. The name comprises three words that are also characteristic of the family, namely "pico" (referring to the small size of the viral particles), "RNA" (referring to the ribonucleic acid genome),

and “virus”, in conjunction literally meaning small RNA virus. There are currently 17 officially designated genera belonging to the family, with a number under revision (Adams *et al.*, 2013). The current taxonomy is summarised in Appendix 1.1.

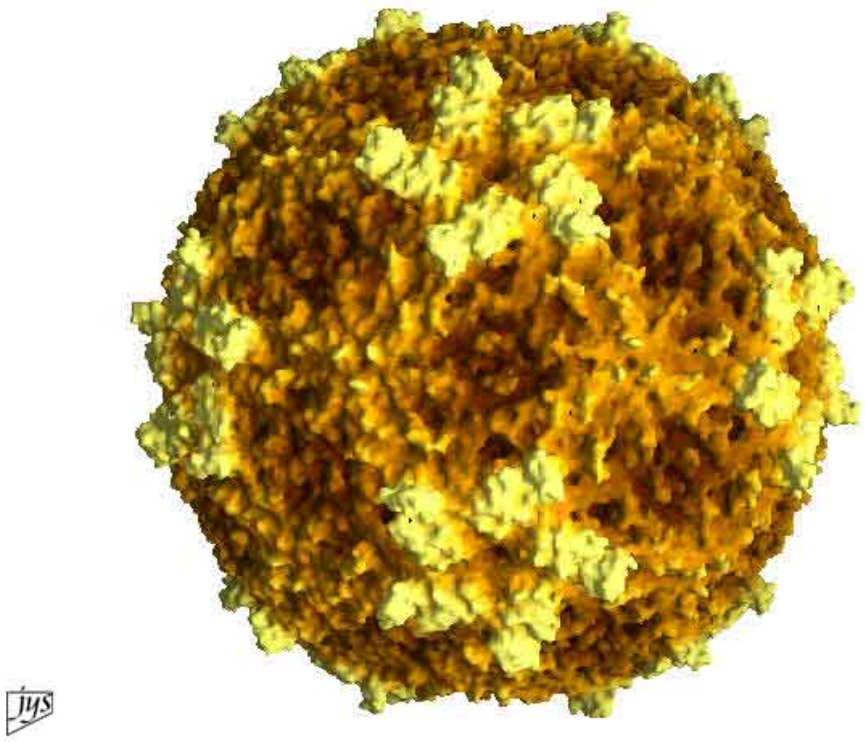

Figure 1.1 Computer-generated image of the protein coat of the Mengo virus (PBD ID: 2MEV, Krishnaswamy & Rossman, 1990), image courtesy of Dr. J.-Y.Sgro, UW-Madison, USA. http://virology.wisc.edu/virusworld.

1.1.3. Genome composition and virion structure

The *Picornaviridae* are known to be responsible for a variety of diseases affecting both humans and domesticated animals (Hughes, 2004). The most notorious of these infecting humans is poliovirus (species *Enterovirus C*, genus *Enterovirus*) causing the debilitating disease poliomyelitis, which was first identified in 1908 by Karl Landsteiner (Paul, 1971). Within domesticated and wild ungulates the most problematic virus is the "foot-and-mouth disease virus" (*Aphthovirus*), which is responsible for major economic losses worldwide (Hughes, 2004; Martinez-Salas *et al.*, 2008). Most picornaviral genomes consist of a single open reading frame of roughly 7000-8800 nt (complete genome sizes of 6400-8400 nt). One picornavirus,

canine picodicistrovirus (genus *Dicipivirus*, species *Cadicivirus A*), is an exception, consisting of two open reading frames, one encoding the capsid and the other encoding the non-structural polypeptides, separated by an intergenic region containing a second IRES (Molla *et al.,* 1992; Woo *et al.,* 2012). This single reading frame encodes a polyprotein that is then subjected to post translational cleaving resulting in between 11 and 14 distinctive proteins, depending on virus species (Hughes, 2004; Simmonds, 2006). The polyprotein is divided into four structural regions; L, P1, P2, and P3, although some viruses lack the leader polypeptide, L (Figure 1.2).

P1 encodes structural proteins VP4, VP2, VP3 and VP1, which combine to form a non-enveloped, icosahedral virus outer shell of approximately 30-nm (Figure 1.3) (Simmonds, 2006), with VP1-3 being surface-exposed. The VP1, VP2 and VP3 proteins form anti-parallel β-sheet barrels during folding, illustrated in Figure 1.4, using VP2 as an example (Arnold *et al.,* 1987). Post-translational modification of the precursor polyprotein in RNA viruses is usually facilitated by four different enzymatic reactions namely, cysteine proteases, acid proteases, serine proteases and metalloproteases (Arnold *et al.,* 1987). However, the mechanism of cleavage of VP0 (the precursor of VP4 and VP2) is not known and does not occur in some picornaviruses, e.g. parechoviruses.

Picornaviruses enter susceptible cells by binding to specific surface exposed receptors and are internalized by various mechanisms resulting in uncoating of the virus particles within vesicles and release of the RNA into the cytoplasm. The virus polyprotein in synthesized by translation of the ORF via a cap-independent mechanism involving the binding of cellular ribosomes to an internal ribosome entry site (IRES) located in the 5′ untranslated region of the genome. Following processing of the polyprotein (normally by virus-encoded proteinases) replication takes place using a virus encoded primer, 3B or VPg (viral protein genome-linked), which is covalently linked to two uracils via an AA template within a virus RNA secondary structure known as a *cre* (*cis*-acting replication element) to form VPg-pU-pU-OH. Thus every copy of the virus RNA (both positive and negative strands) is covalently linked to a VPg molecule. RNA replication proceeds on membranes and involves many of the virus non-structural polypeptides, particularly the RNA- dependant RNA-polymerase known as $3D^{pol}$ (Mettenleiter & Sobrino, 2008).

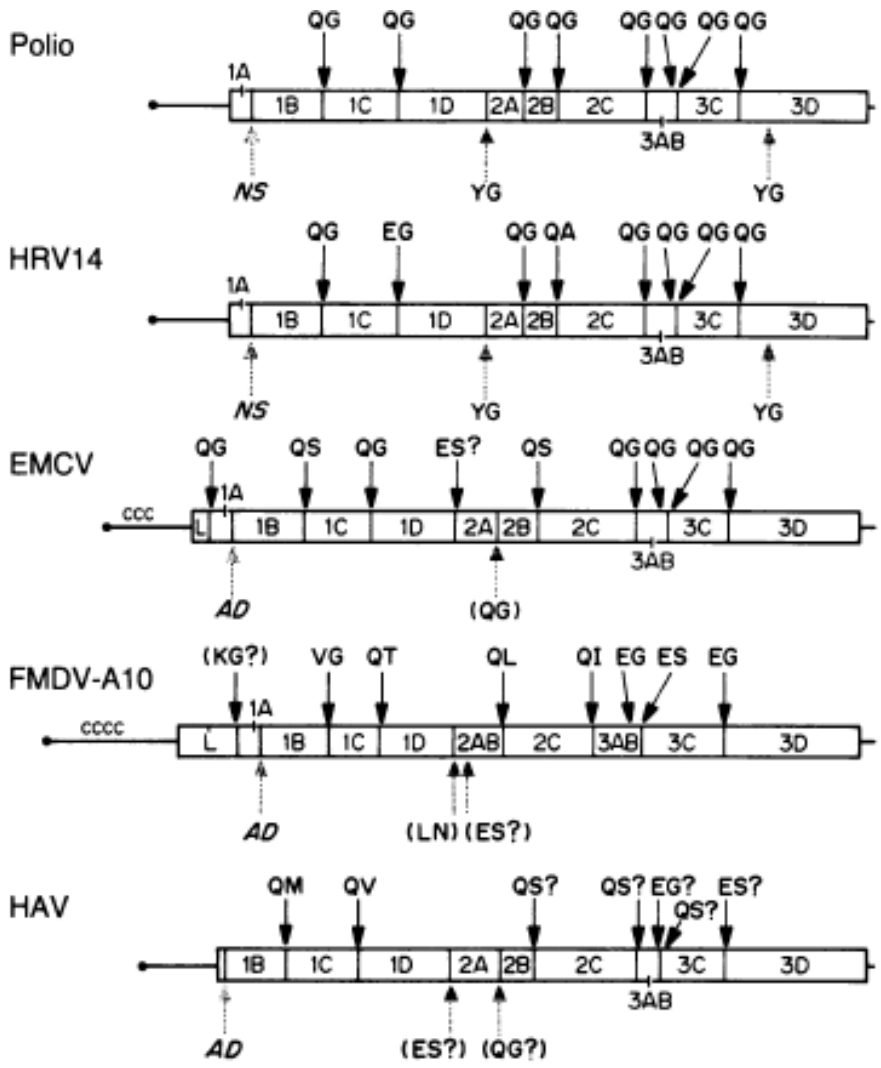

Figure 1.2 Structural representation of the picornaviral genome depicting the 3 major structural regions. Cleavage sites produced by protease 3C are shown above the genome and those for protease 3A below (Polio - Poliovirus; HRV - Human rhinovirus 14; EMCV - Encephalomyocarditis virus; FMDV- A10 - foot-and-mouth disease virus A10; HAV - Hepatitis A virus). Those in parentheses are indicative of sites where the proteolytic enzymes have not been identified. The proteolytic cleavage sites occur between amino acid pairs that are indicated by the standard single letter codes; question marks indicate sites where the sequence has not been previously determined. Figure taken from Arnold *et al.* (1987), with permission granted under PNAS copyright notice (http://www.pnas.org/site/misc/rightperm.shtml).

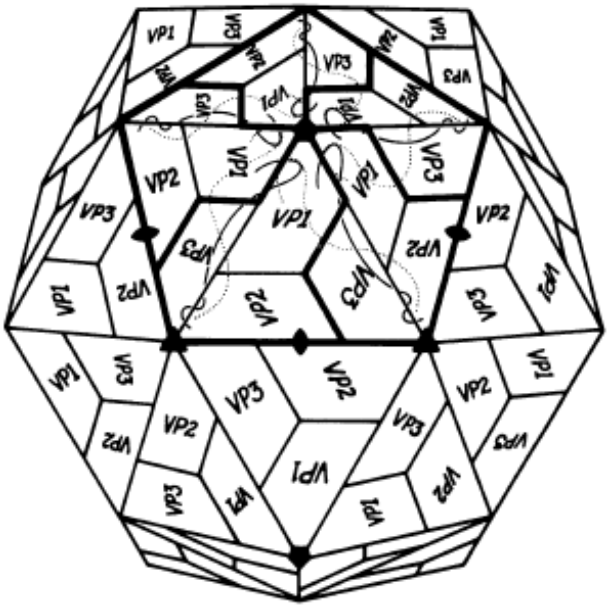

Figure 1.3 Diagrammatic representation of the capsid proteins VP1, VP2 and VP3 of a picornavirus forming an isochedral outer shell. Figure from Arnold *et al.* (1987). Copyright held by the original authors. Permission granted under PNAS copyright notice (http://www.pnas.org/site/misc/rightperm.shtml).

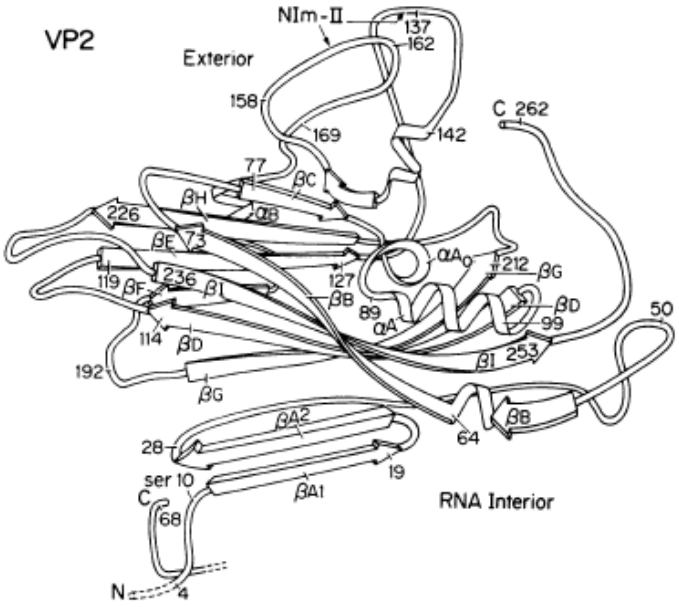

Figure 1.4 Ribbon drawing of VP2 obtained from human rhinovirus 14. The serine is conserved in *Cardiovirus*, *Enterovirus* and rhinovirus genomes. Figure from Arnold *et al.* (1987). Copyright held by the original authors. Permission granted under PNAS copyright notice: (http://www.pnas.org/site/misc/rightperm.shtml).

1.1.4. Specific properties of Cardioviruses

The *Cardiovirus* virion is made up of a single stranded molecule of RNA contained within an icosahedral protein capsid. Instead of a methylated nucleotide cap, the VPg structural protein, 20 amino acids in length, is covalently bonded to the 5' end of the RNA molecule by means of a phosphotyrosine bond. This protein is encoded by a section of the P3 region of the genome (Scraba & Palmenberg, 1999). As with other members of the *Picornaviridae* the capsid is composed of four structural proteins, VP1-VP4. Sixty copies of each protein are needed to form a single virion (Figure 1.2). The molecular weights of these proteins can differ between serotypes, and for Mengo viruses the weights for VP1, VP2, VP3 and VP4 are 31700kDa, 29000kDa, 25100kDa and 7200kDa respectively. Approximately 75 % of the capsid is usually occupied by the RNA the remaining volume is composed of water, polyamines and salts. To date no mRNA or tRNA, viral or cellular proteins have been observed within a *Cardiovirus* virion (Scraba & Palmenberg, 1999).

Cardiovirus genomes are approximately 7,600 to 8,000 nucleotides in length, with a polyadenylated 3' end and an untranslated 5' region attached to the VPg protein. EMC viruses have an unusual polypyrimidine tract that can vary from 60 to 420 bases in length. The polypyrimidine tract of Mengo and EMC are $C_{44}UC_{10}$ and $C_{125}UCU_3UC_{10}$, respectively. The 3' UTR's of cardioviruses are of intermediate size amoung Picornaviruses, varying between 124 and 126 nucleotides in length. This region is followed by a poly-A tail containing between 20 and 70 adenosine repeats. The poly-A tail is an integral component when it comes to the RNA binding to the 3D polymerase, as tracts of less than 15 repeats are not recognised by cellular enzymes (Scraba & Palmenberg, 1999).

1.1.5. Infection and replication mechanisms of EMCV

The virions are incorporated into the cell by means of endocytic vesicles. Acidification of the virion is needed before the viral genomic material can be released into the cytoplasm of the cell (Scraba & Palmenberg, 1999; Illustrated in Figure 1.5). Once incorporated into the cell the viral RNA acts as mRNA and hijacks the cellular mechanism to translate the polyprotein. The 3D region of the genome contains the coding region for an RNA polymerase enzyme which facilitates further replication of the viral genome (Scraba & Palmenberg, 1999). Replication takes place in the

cytoplasm of an infected cell, with all necessary processes such as translation, RNA synthesis and protein encapsulation being completed in the absence of a nucleus (Lidsky *et al.,* 2006). Pollack & Goldman (1973) demonstrated that poliovirus was able to successfully replicate in enucleated African green monkey kidney (BSC1) cells. This finding was later substantiated by Follett *et al.,* (1975) who were able to show the replication of echovirus, poliovirus, pseudorabies virus, reovirus, respiratory syncytial virus and Semliki Forest virus in nucleus-free BSC-1 cells. The yield of Semliki Forest virus in this study was comparable to that in normal cells, but the replication rates of the other previously-mentioned viruses where much lower than those of nucleated cells.

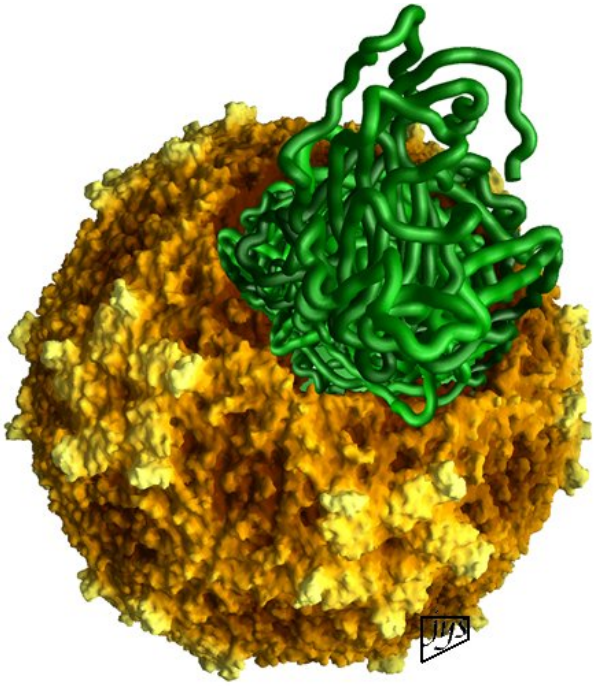

Figure 1.5 Computer-generated image of the Mengo virion rupturing and releasing genetic material (PBD ID: 2MEV, Krishnaswamy & Rossman, 1990). Image courtesy of Dr. J.-Y.Sgro, UW-Madison, USA. http://virology.wisc.edu/virusworld.

Though non-essential, the role of the nucleus in the infection has been demonstrated by the detection of viral proteins in the nucleus of EMCV-infected cells (Aminev *et al.,* 2003 *a*&*b*; Lidsky *et al.,* 2006). The role of viral components on the nucleus of an infected cell varies in an EMCV infection. The 2A protein enters the nucleoli and alters host cell translation mechanisms (Medvedkina *et al.,* 1974; Akiyama *et al.,* 1987). Lobert *et al.,* (1999) identified a 130 bp sequence contained within the VP2 region of the viral genome which was involved in the initiation of

replication within both Theiler's virus and Mengovirus isolates. Further studies have shown that on infection with a picornavirus a number of the host cells proteins are translocated from the nucleus into the cytoplasm where they may have a role on viral genomic translation (Hellen *et al.,* 1993; Meerovitch *et al.,* 1993; Blyn *et al.,* 1997; Izumi *et al.,* 2001; Svitkin *et al.,* 2001) and replication (McBride *et al.,* 1996; Waggoner & Sarnow, 1998; Lidsky *et al.,* 2006). The incorporated viral genome finally acts as a medium to facilitate viral propagation by interacting with newly formed capsid components to form new virions (Lobert *et al.,* 1999). As with other picornaviruses, it has been shown that full-length *Cardiovirus* RNA is able to infect both animal and cultured cells and that the mRNA molecule does not have to be capped or attached to VPg protein molecules to effectively infect the cells (Scraba & Palmenberg, 1999).

1.1.6. Epidemiology

Members of the order Rodentia are generally considered to be *Cardiovirus* reservoir hosts (Zimmerman, 1994; Thomson *et al.,* 2001). However, there have been insufficient studies conducted on the epidemiology of this virus to be absolutely certain. Outbreaks of the disease can be devastating especially to swine, with many of the isolates characterized to date being of suid origin (Knowles *et al.,* 1998; Koenen *et al.,* 1999; Maurice *et al.,* 2005; Denis *et al.,* 2006; Maurice *et al.,* 2007). The disease has remarkable epizootic potential causing outbreaks in domestic pigs, a range of wildlife species and even humans (Table 1.1). In sporadic outbreaks in Taronga Zoo, Sydney from the period 1987-1995 the disease was responsible for the deaths of a number of species namely; Common Squirrel Monkey (*Saimiri sciureus),* Mandril (*Mandrillus sphinx),* Chimpanzee (*Pan troglodytes),* pygmy hippopotamus (*Choeropsis liberiensis),* Goodfellow's Tree-kangaroo (*Dendrolagus goodfellowi)* and Ring-tailed Lemur (*Lemur catta),* (Reddacliff *et al*., 1997). More recently there have been reports from Asia of fatal EMCV infections in tigers and orangutans from China and Singapore, respectively and of the likely re-emergence of EMCV in rhesus macaques in the Caribbean following a period of quiescence of more than three decades (Kessler *et al.,* 1982; Masek-Hammerman *et al.,* 2012; Kessler, 2013).

Outbreaks of the disease are usually restricted to swine but as demonstrated in Table 1.1 below, have been shown to affect an extremely wide range of species.

Table 1.1 Summary of some of the major outbreaks of EMC and Mengo viruses reported globally, to date:

Date of the Outbreak	**Location**	**Species infected**	**Source**
1960-1966	Florida, USA	*Sus scrofa domestica*	Gainer, (1967)
1970	New South Wales, Australia	*Sus scrofa domestica*	Acland *et al.*, (1970) Acland & Littlejohns, (1975)
1975-1981	Cuba	*Sus scrofa domestica*	Ramos *et al.*, (1983)
			Kessler *et al.*, (1982)
1976	Northland, New Zealand	*Sus scrofa domestica*	Sutherland *et al.*, (1977)
1979	Natal-midlands, South Africa	*Sus scrofa domestica*	Williams, (1981)
1984	New South Wales, Australia	*Sus scrofa domestica*	Seaman *et al.*, (1986)
Early 1980's	State of Rio Grande do Sul, Brazil.	*Sus scrofa domestica*	Roehe *et al.*, (1985)
1986-1998	Italy	*Sus scrofa domestica*	Gualandi *et al.*, (1989) Koenen *et al.*, (1999)
1986-1997	Greece	*Sus scrofa domestica*	Tsangaris *et al.*, (1989) Paschaleri-Papadopoulou *et al.*, (1990) Koenen & Vanderhallen, (1997) Koenen *et al.*, (1997 & 1999)
1987 -1989	Minnesota, USA	*Sus scrofa domestica*	Kim *et al.*, (1989*a* & b, 1991) Christianson *et al.*, (1990)
1990-1991	San Antonio, Texas	*Papio* species, *Mandrillus leucopheus.*	Hubbard *et al.*, (1992)
1990	Puerto Rico	*Sus scrofa domestica*	Unpublished*
1990	Quebec, Canada	*Sus scrofa domestica*	Dea *et al.*, (1991*a*) Dea *et al.*, (1991*b*)
1990-1997	Belgium	*Sus scrofa domestica*	Koenen *et al.*, (1991, 1994, 1996, 1997) Koenen, (1994) Castryck *et al.*, (1996) Koenen & Vanderhallen, (1997) Vanderhallen & Koenen, (1997*a* & b)

Date of the Outbreak	Location	Species infected	Source
Early 1990's	Taiwan	*Sus scrofa domestica*	Hu *et al.*, (1993)
1992	South Korea	*Sus scrofa domestica*	Park *et al.*, (1992)
1992-1993	Côte d'Ivoire	*Sus scrofa domestica*	Unpublished*
1993-1994	Kruger National Park, South Africa	*Loxodonta Africana*	Grobler *et al.*, (1995)
1994-1997	Cyprus	*Suss crofa domestica*	Koenen *et al.*, (1999)
1995	France	*Sus scrofa domestica*	Koenen *et al.*, (1999)
1996	Audubon Park Zoo, New Orleans, USA	*Lama guanicoe, Alouatta* sp, *Miopithecus talapoin*	Wells *et al.*, (1989)
1997	New South Wales, Australia	*Saimiri sciureus, Mandrillus sphinx, Pan troglodytes, Choeropsis liberiensis, Dendrolagus goodfellowi, Lemur catta*	Reddacliff *et al*., (1997)
1988	Hawaii, USA	*Sus scrofa domestica*	Unpublished*
2001-2002	Singapore	*Pongo pygmaeus, Pongo abelii*	Yeo *et al.*, 2013
2001-2004	Democratic Republic of the Congo	*Pan paniscus*	Jones *et al.*, (2004) Jones *et al.*, (2011)
2004	Peru	*Homo sapiens*	Oberste *et al.*, (2009)
2006-7	Italy	*Eulemur macaco macaco, Lemur catta, Varecia variegata, Rubra, Eulemur albifrons, Macaca sylvanus, Callithrix jacchus*	Canelli *et al.*, 2010
2008-2009	Sabana Seca, Puerto Rico),	*Macaca mulatta*	Masek-Hammerman *et al.*, 2012
2009	Sierra Leone	*Pan troglodytes*	Unpublished* **
2012	Fujian province, southern China	*Panthera tigris amoyensis*	Liu *et al.*, 2013

*http://www.*Picornaviridae*.com/cardiovirus/emcv/emcv_outbreaks.htm

**http://tacugama.wildlifedirect.org/2009/04/29/emcv-strikes-again-with-tragic-consequences/

1.1.7. South African outbreaks

Two major outbreaks of EMCV have been recorded in South Africa. The first occurred in the Natal Midlands of the Kwa-Zulu Natal Province in 1979 and was restricted to swine (Williams, 1981). The second outbreak occurred between December 1993 and November 1994 in the Kruger National Park, South Africa, a conservation area which is situated along the north-eastern border between South Africa and Mozambique, between latitudes of 22°20'S and 25°32'S. The latter outbreak occurred primarily in African Elephants (*Loxodonta africana*), at the time the park had a population of 7,500 animals. At its peak, in January 1994 the virus was responsible for the mortalities of 32 elephants in a single month. By the end of the outbreak the virus had resulted in mortalities in 64 African elephants, 53 (83%) of which were adult bulls. The outbreak was believed to have been as a result of an associated rodent irruption which saw rodent trapping success increase from values of between 3 % and 12 % reported in the years preceding the onset of the outbreak to trapping successes of between 54 % and 56 % in 1993 and 1994, respectively (Grobler *et al*., 1995). Prior to the onset of this outbreak, 48 previously unreported elephant mortalities had been noted in the park spanning the years 1987 to 1993. Reanalysis of these deaths revealed male-biased mortality with 33 of 48 carcasses (69 %) being adult elephant bulls that most likely succumbed to EMCV infection. Retrospective serological screening of elephant sera collected from 1984 onwards confirmed that the virus had been present since 1987. The outbreak that occurred in the mid-1990s coincided with a rodent outbreak in the park and in neighbouring Mozambique. Members of the order Rodentia are considered to be the likely reservoir hosts of the virus and the increase in rodent numbers may have contributed to the spread of the virus (Grobler *et al*., 1995). Transmission of EMCV occurs *via* the faecal-oral route, with transmission most likely occurring following ingestion of food contaminated with either faeces or urine. In addition to affecting elephants, the virus was suspected as the likely cause for the decline in population numbers of myomorph rodents (Suborder Myomorpha) in the park from mid-1994 onwards (Grobler *et al.*, 1995). Carcasses of small or intermediate size mammals, such as species of antelope, did not increase notably during the outbreak. This may be because carcasses are readily scavenged by predators. However, as this is not the case during anthrax outbreaks or droughts, where ungulate carcasses are commonly found it was

speculated that the virus did not cause high mortalities in mammalian species other than elephants, with the possible exception of small rodents (Grobler *et al*., 1995).

1.1.8. Molecular epidemiology

Genetic characterization of phylogenetically informative virus genome regions makes it possible to determine the origin/source of an outbreak and to trace the spread as viruses show distinct temporal and/or geographical clustering patterns (Holmes, 1998). Previous studies have shown that picornaviruses generally cluster according to geographical origin (Knowles & Samuel, 2003), therefore if a comprehensive study of *Cardiovirus* diversity is performed in the KNP, it is likely that a valuable baseline dataset will become available that may prove valuable for determining both the geographical source as well as the likely murid rodent reservoir host, in future outbreaks in this region.

1.1.9. Aspects of the Disease

Mengovirus and EMCV are serologically indistinguishable. Transmission and maintenance of the viruses is believed to be facilitated by murid rodents, however, some arthropod vectors have been implicated (Wells *et al.,* 1989; Zimmerman, 1994; Backues *et al.,* 1999). Death as a result of EMCV infection occurs rapidly with anorexia and depression being observed in infected animals between 12 and 24 hours prior to death. The disease is characterized by the presence of lesions usually limited to the cardio-pulmonary system. Other clinical symptoms of the disease include pulmonary edema and streaks encompassing the entire myocardium. Death is usually as a result of acute cardiac arrest (Backues *et al.,* 1999).

1.1.10. Murid phylogeny

Members of the family Muridae are the most diverse group of small mammals present on the Earth today, not only on a purely numerical basis (with close on 1300 recognised species having been recorded to date), but also because they have a world-wide distribution, occurring in every conceivable habitat type (Jansa & Weksler, 2004). The family is divided into five sub-families, namely (i) Deomyinae, comprising the genera *Acomys*, *Deomys*, *Lophuromys* and *Uranomys* (Michaux *et al.,* 2001; Steppan *et al.,* 2004), (ii) Leimacomyinae comprising a single species

Leimacomysbuettneri, (iii) Gerbillinae comprising the genera *Ammodillus, Brachiones, Desmodilliscus, Desmodillus, Gerbillus, Gerbillurus, Microdillus, Meriones, Pachyuromys, Psammomys, Rhombomys, Sekeetamys, Tatera* and *Taterillus*, (iv) Lophiomyinae consisting of a single species the Manned rat *Lophiomysimhausi* and (v) the sub-family Murinae which is by far the largest, comprising 129 genera with over 500 species. This latter sub-family has a worldwide distribution (Michaux *et al.,* 2001; Jansa & Weksler, 2004; Steppan *et al.,* 2004).

One member of the Murinae, in particular, was of special interest to the current study, the genus *Mastomys* (Smith 1834), as it is considered to be the putative reservoir host of the virus. These rodents are one of the most prolific breeding murid species in Africa (Leirs *et al.,* 1996) occurring over much of the continent in a range of habitats from grassy thickets, to areas modified for human habitation (Mohr *et al.,* 2003). The genus is considered to be taxonomically confusing, since some species are morphologically indistinguishable from each other. All *Mastomys* species have been shown to be agricultural pests (Mohr *et al.,* 2003) and to be reservoir hosts for a myriad of pathogenic organisms. As early as the 1970's the reservoir host of Lassa fever was identified as the Natal multi-mammate mouse, *M. natalensis* (Monath *et al.,* 1974), however with advances in technology, two further species were put forward as possible hosts of this arenavirus. In a study using the then newly-developed technology of haemoglobin electrophoresis, McCormick and co-workers (1987) identified *M. huberti* and *M. erythroleucus* as possible hosts (Lecompte *et al.,* 2006). Later this view was brought into contention by a study conducted by Lecompte *et al.* (2006), in which the authors sampled approximately 1500 murids from 13 localities in Guinea in West Africa. The results of their study showed that only molecularly typed *M. natalensis* were in fact positive for the virus (Lecompte *et al.,* 2006). *Mastomys* has also been shown to host *Yersinia pestis* the causative agent of the plague. Isaäcson *et al.,* (1981) elucidated the susceptibility of the two species, *M. coucha* and *M. natalensis*, that make up the *M. natalensis* species complex. Their research showed that while *M. coucha* was highly susceptible to infection with the bacteria, *M. natalensis* displayed no obvious symptoms.

1.1.11. Aims of the study

The aim of the current study was to assess the diversity of extant cardioviruses in an endemic locality in South Africa, the Kruger National Park, against a backdrop of historic outbreak strain characterization, in an attempt to assess the diversity and origins of present day cardioviruses, through the use of two historic isolates from the KNP outbreak and through sampling and isolation of extant cardioviruses. The study also aimed to evaluate the taxonomically complex murid rodent genus, *Mastomys*, implicated as the reservoir of infection, by means of a mitochondrial DNA-based phylogeographical study. The host rodent data generated in the course of the study was subsequently used to develop a molecular diagnostic method for rapid differentiation of the two cryptic species that occur in South Africa, *viz. M. natalensis* and *M. coucha*.

CHAPTER 2

A retrospective genetic analysis of *Cardiovirus* outbreaks in South Africa

J. H. d. T. van Sandwyk[1], N. C. Bennett[1], R. Swanepoel[2]& A. D. S. Bastos[1]

[1]*Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa*

[2]*National Institute for Communicable Diseases, National Health Laboratory Service, Private Bag X4, Sandringham, Johannesburg, 2131, South Africa*

Chapter Summary

Encephalomyocarditis virus (EMCV) outbreaks are rare in southern Africa, with only two being reported to date from South Africa and both coinciding with rodent irruptions. The first outbreak of acute myocarditis occurred in swine in the Natal Midlands in 1979, whilst the second, which resulted in the deaths of 64 African elephants (*Loxodonta africana*), occurred between 1993 and 1994 in the Kruger National Park (KNP). Three virus isolates, one from the swine outbreak and two from the elephant outbreak, were genetically characterised in this study. The P1 genome region, inclusive of the flanking leader (L) and 2A genes, of three historical South African viruses, one from swine and two from elephant, was characterised by PCR amplification and sequencing of up to 11 overlapping fragments. In addition to the resulting 3329 nucleotide dataset, the 3D region, widely used in molecular epidemiological studies, was characterised, and three datasets (P1, VP1/3 and 3D), complemented with available homologous EMCV data, were ultimately compiled for analyses. Phylogenetic inferences revealed the near-identical elephant outbreak strains to be most closely related to a Mengovirus from Rhesus Macaques (*Macaca mulatta*) in Uganda, differing from the latter by between 11 % (3D) and 15 % (VP3/1). The historical South African pig virus differed by 4 % (3D) and 11 % (VP3/1) from available European and Asian pig outbreak strain sequences. This study confirms the presence of two genetically distinct EMCV lineages from wild and domestic hosts in

southern Africa, and provides valuable, novel baseline data for future outbreaks in the sub-region.

2.1. Introduction

Picornaviral genomes are relatively small, consisting of a single open reading frame of roughly 7800-8200 nucleotides, which in turn encodes a single polyprotein that is then subjected to post-translational cleaving resulting in 11 distinct proteins (Hughes, 2004; Simmonds, 2006). This single polyprotein is divided into four regions, namely L, P1, P2, and P3. The latter three each encode five (1A-1D), three (2A-2C) and four (3A-3D) non-overlapping genes (Hughes, 2004; Simmonds, 2006). The P1 region is responsible for encoding the virion structural proteins VP4 (1A), VP2 (1B), VP3 (1C) and VP1 (1D), which combine to form a non-enveloped, icosahedral shell of approximately 30nm (Simmonds, 2006), with VP1-3 being surface-exposed. This is expected as the three surface exposed proteins (VP1-3) are under the greatest degree of selective pressure by the host's immune system (Urwin *et al.*, 2002). The VP3/1 junction has proved valuable for determining relatedness of outbreak strains (Knowles *et al.*, 1998). Past studies on EMCV have identified immune-dominant and neutralising epitopes in structural proteins, VP2, VP3 and VP1 (summarised in Table 2.1.).

The P2 region is divided into three peptide regions, namely 2A, 2B and 2C. Although it is known that P2 is involved in RNA synthesis and encoding the primary cleavage peptide, the exact roles of the 2A and 2B regions are unknown. The 2C region plays a role in the initiation of RNA synthesis by encoding the guanidine resistance marker. The final region of the polypeptide, P3, consists of four gene regions, 3A, 3B, 3C and 3D. The protein VPg is encoded by the 3B polypeptide. This protein attaches to the 5' end of the viral genome and plays a role in the initiation of RNA synthesis. The 3C region forms a viral protease$3C^{pro}$ responsible for secondary cleavage of the polypeptide, while the 3D region encodes the RNA-dependent RNA polymerase (RNA replicase), responsible for RNA replication (Scraba & Palmenberg, 1999). This gene region is often used for molecular epidemiology studies (Vanderhallen & Koenen 1998; Koenen *et al.*, 1999; Denis *et al.*, 2006; Jones *et al.*,

2011), and is the best-represented gene in the GenBank database, with sequences for more than 60 strains, predominantly from swine, presently available.

Table 2.1 Summary of the neutralizing epitope sites of the EMCV-1 serotype identified in various studies.

Nucleotide Position[a]	**Amino acid or recorded change**[b]		
	VP1	**VP2**	**VP3**
2775 A -> C[d]	1049 Lys ->Gln[d]		
2814 A ->G[d] A -> C[d]	1062 Asn ->Asp[d] Asn ->His[d]		
2928 A ->G[d]	1100 Thr ->Ala[d, e]		
2929 C ->U[d]	1100 Thr - >Ile[d, e]		
1217 A -> C[d]		2016 Gln ->His[d]	
1395-1397		2075 Pro[c,e]	
1599 A ->G[d]		2144Asn ->Asp[c, d, e]	
1600 C ->G[d]		2144 Asn ->Ser[c, d, e]	
1602-1604		2145 Arg[c, e]	
1608-1610		2147 Ser[c, e]	
1611-1613		2148 Lys[c, e]	
2106-2108			3057 Lys[c, e]
2118-2120			3061 Ala[c]
2139-2141			3068 Ser[c, e]
2154 A -> C[d]			3073 Lys ->Gln[d]
2158 C ->U[d]			3074 Thr ->Ile[d]
2160 C ->G[d]			3075 Gln ->Glu[d]
2347 C ->U[d]			3137 Thr ->Ile[d]

a. Nucleotide position based on Mengo isolate M (L220891).
b. Position indicated according to the Rossman system, in which the first digit refers to the capsid polypeptide (1, 2 or 3) and the next three identifies its position in the chain from the N-terminus.
c. Determinate amino acid residue of epitopes identified by Boege *et al.*, (1991).
d. Mutations and determinate amino acid residue of epitopes identified by Kobasa *et al.*, (1994).
e. Determinate amino acid residue of epitopes listed in Scraba & Palmenberg (1999)

Replication in lytic viruses takes place several times in a single infectious cycle, resulting in the proliferation of viral particles containing several mutations from a single infected cell. Many of these resulting mutations may in fact be deleterious to the virions, resulting in the characteristic low specificity recorded within infections (Drake *et al.,* 1998). These viruses are susceptible to variations in the mutation rate with extinction within a population resulting from as little as three-fold increase in mutation rate (Holland *et al.,* 1990; Drake *et al.,* 1998). Mutation rates in picornaviruses have been recorded at 0.8 and 1.0 per replicate for poliovirus and influenza viruses respectively (Drake *et al.,* 1998).These high mutation rates lead to

the formation of so called “quasispecies” or “swarm of mutants” in which a ‘master sequence’ predominates (Eigen, 1996), the sequence of which can be recovered by PCR amplification and sequencing of all variants present within the “dynamic cloud of mutants” (Díaz Arenas & Lehman, 2010).

The genus *Cardiovirus* comprises two currently recognised species, namely *Encephalomyocarditis virus (EMCV),* consisting of two serotypes and *Theilovirus (THV)* (Fenner *et al.,* 1993; Knowles *et al*., 2012; Philipps *et al.,* 2012) containing as many as 12 distinct serotypes, *viz.* Theiler's murine encephalomyelitis virus (TMEV), Thera virus (TRV), Vilyuisk human encephalomyelitis virus (VHEV), Saffold virus (SAFV) consisting of the genetically distinct isolates SAFV-1, SAFV-2, SAFV-3, SAFV-4, SAFV-5, SAFV-6, SAFV-7, SAFV-8 and SAFV-9 (http://www.*Picornaviridae*.com/).

Cardioviruses were first discovered in the early 1940’s when attempting to transfer the Yale-SK poliovirus from monkeys to mice, the unknown pathogen at the time was dubbed the Columbia-SK virus (Scraba & Palmenberg, 1999). This virus named Encephalomyocarditis virus by Helwig and Schmidt (1945). A strain of EMCV, the Mengo virus was first isolated by Dick *et al.,* (1948). The virus was isolated by inoculating mice inter-cerebrally with infected blood or spinal fluid obtained from a rhesus monkey which had developed lower limb paralysis. The monkey originated from the Mengo district of Entebee in Uganda (Scraba & Palmenberg, 1999). Initially, the degree of relatedness of diverse poliomyelitis and mouse encephalomyelitis viruses was not clear (Dick *et al.,* 1948; Bautista *et al.,* 1957), but the original Columbia-SK, EMC and Mengo isolates, as well as viruses from diverse hosts, sampled across a broad geographical range, are now all recognised as members of a single EMCV species and two serotypes (Craighead, 1965; Zimmerman, 1994; Philipps *et al.,* 2012).

The virus has an extremely broad host range, occurring in diverse natural murid rodent populations, in domestic and wild pigs (Maurice *et al.,* 2005), in a wide range of captive species (Thomson *et al.,* 2001) and even in humans (Dick *et al.,* 1948; Tesh & Wallace, 1978; Obertse *et al.,* 2009). Clinical symptoms are equally broad ranging from asymptomatic, to acute and highly lethal (Thomson *et al.,* 2001). Aside from outbreaks in swine (Knowles *et al.,* 1998; Maurice *et al.,* 2005; Denis *et al.,* 2006) there have been outbreaks in exotic animals as well. In sporadic outbreaks in Taronga Zoo, Sydney from 1987-1995, the disease was responsible for the deaths

of a number of species namely; Common Squirrel Monkey (*Saimiri sciureus)*, Mandril (*Mandrillus sphinx)*, Chimpanzee (*Pan troglodytes)*, pygmy hippopotamus (*Choeropsis liberiensis)*, Goodfellow's Tree-kangaroo (*Dendrolagus goodfellowi)* and Ring-tailed Lemur (*Lemur catta*) (Reddacliff *et al*., 1997). A review of the published literature however, shows that primates and elephants appear to be particularly susceptible to the disease (Simpson *et al.,* 1977; Tesh & Wallace, 1978; Seaman & Finnie, 1987; Jones *et al.*, 2011; Masek-Hammerman *et al.*, 2012). Murid rodents are generally accepted to be the carriers of the disease, but, there have been insufficient epidemiological studies conducted to be absolutely certain (Zimmerman, 1994; Thomson *et al*., 2001), however, a possible role for wild boars, in addition to rodents, in the epidemiology of the disease in domestic pigs in Europe, has been suggested (Billinis, 2009). Although EMC has essentially a worldwide distribution, the molecular characterisation of viruses primarily corresponds to swine outbreak strains of European, Asian and American origin. EMC virus data from Africa is presently limited to two historical primate strains isolated from Rhesus macaques in Uganda in the 1940s and to a contemporary strain isolated from a semi-wild bonobo in the Democratic Republic of Congo in 2009 (Jones *et al.,* 2011). Though outbreaks of EMCV in South Africa are rare, two previous outbreaks of the disease have occurred. The most recently documented outbreak occurred between December 1993 and November 1994. The virus caused mortalities in 64 African elephants (*Loxodonta africana*), of which 83 % were adult males, in the Kruger National Park (KNP), South Africa (Grobler *et al*., 1995). The earlier outbreak was restricted to swine and occurred in the Natal Midlands of the Kwa-Zulu Natal Province in 1979 (Williams, 1981).

In this study we genetically characterise one historical swine isolate and two elephant isolates in order to determine relatedness and possible origins of the virus and to establish a regional reference dataset for future outbreaks. This was achieved by amplification of the P1 and flanking regions, as well as the RNA polymerase gene (3D) region of the viral genome for which numerous reference sequences are available and due to the epidemiological utility of this gene.

2.2. Materials & Methods

2.2.1 RNA extraction and cDNA synthesis

Three historical viruses were characterised in this study, *viz.* a strain recovered from the 1979 swine outbreak and designated SAR/1979 and two viruses from African elephants that succumbed during the 1993/4 outbreak in the Kruger National Park, designated KNP/17/94 and KNP/19/94. Total RNA was extracted using the Zymo Research Mini RNA Isolation II kit, based on the RNA extraction method of Chomczynski & Sacchi (1987). Subsequent cDNA synthesis was achieved using a random hexanucleotide approach which generates first strand cDNA from all available RNA present within the sample and which has previously been used for characterisation of picornaviral genes (Meyer *et al.,* 1991; Bastos, 1998). RNA was reverse-transcribed in the presence of 1 × AMV Reverse Transcriptase buffer (Fermentas), 5 µM random hexamerprimers (IDT), 0.2 µM dNTPs (Inqaba) and 10 U Ribolock^{TM}RNase inhibitor (Fermentas). Reactions were performed at 42°C for 1 hour in a final reaction volume of 10 µl, containing 5 % DMSO and 10U of AMV Reverse Transcriptase (Fermentas). Following this, samples were heat-inactivated at 80°C for 1 minute and stored at -20°C.

2.2.2. PCR amplification

Primers, designed specifically for this study and summarised in Table 2.1, were used to generate overlapping genome fragments (371 to 1500 nt in length) spanning the entire P1 region. In addition, the epidemiologically informative 3D region was targeted with primers previously used to screen for Cardiovirus genome presence in invasive *Mus* from sub-Antarctic Marion Island (de Bruyn *et al.,* 2008). All genomic amplifications followed the same general protocol, in which a master mix was prepared which contained 1 × Dream*Taq*^{TM} Buffer (Fermentas), 0.2 µM dNTPs (Inqaba Biotech), and 0.4 µM of each primer (Inqaba Biotech). Each reaction contained 1.25 U of Dream*Taq*^{TM} DNA polymerase (Fermentas) and 100-200ng of cDNA template and was performed in a final volume of 50µl. Annealing temperature and extension times varied between reactions and were guided by the primer with the lowest calculated Tm of each primer pair (Table 2.1) and by the length of the fragment being targeted respectively. The size of the amplified products was assessed by agarose gel electrophoresis against the GeneRuler^{TM} 100 nt Plus DNA Ladder (Fermentas).

Table 2.2 Summary of *Cardiovirus* primers designed to amplify overlapping fragments of the P1 and flanking genome regions, and the partial 3D gene region of the three historical South African outbreak strains characterised in this study.

Primer pair (name, sequence and orientation)	Primer binding site*	Virus strain/s amplified	Expected size of amplicon	Sequence span generated*	Genome region targeted	Touch down PCR annealing temperatures (No of cycles)
EMC-AB1 5'- GGCCGAAGCCGCTTGGAATA - 3'(F) EMC-AB2 5' - ACGTGGCTTTTGGCCGCAGA - 3'(R)	213-232 478-497	KNP17/94 KNP/19/94SAR/1979	285nt	233-477	5'-UTR	62°C (2x) , 60.5°C (8x) , 59°C (35x)
EMC- AB1 5'- GGCCGAAGCCGCTTGGAATA - 3'(F) 1356R-AB 5'- TGGGTGTTTGTGACCGTGTT- 3'(R)	213-232 1224-1243	KNP17/94 KNP/19/94SAR/1979	1031nt	233-1223	5'-UTR- 1B	61°C (2x), 59°C (8x), 57°C (35x)
386F-JvS 5' - GGTCTGTTGAATGTCGTGAA - 3' (F) 884R-AB5' - GTTCCATGGTTGTAGCCAT- 3'(R)	366-385 753-771	KNP17/94 KNP/19/94SAR/1979	406nt	386-752	5'-UTR- L	55°C (2x), 53°C (8x), 51°C (35x)
687F-AB 5' - GCTCTCCTCAAGCGTATTCA - 3' (F) 1356R-AB 5'- TGGGTGTTTGTGACCGTGTT - 3'(R)	575-594 1224-1243	KNP17/94 KNP/19/94SAR/1979	669nt	595-1223	L - 1B	50°C (2x), 48°C (8x), 46°C (35x)
687F-AB 5' - GCTCTCCTCAAGCGTATTCA - 3' (F) 2868R-AB 5' - CCAATGGGACTAGACCTATCATA - 3'(R)	575-594 2733-2755	SAR/1979	2181nt	595-2732	L-1C	52°C (2x), 51°C (8x) 49°C (35x)
1156EPF-JvS 5' - GGAGAACTTGTCTGATCGAGT- 3' (F) MasRAB-2392R5' - AGGCACAGTGAAGGAGTAAG- 3'(R)	1184-1204 2401-2420	SAR/1979	1237nt	800-900	1B-1D	57°C (2x), 56°C (8x), 55°C (35x)
1307F-AB 5' - TCAGACCGAGTGTCTCAAGA - 3' (F) 1720R-AB 5' - CTTGGACCATCTGTTGTCCAT - 3'(R)	1194-1213 1587-1607	KNP/17/94 KNP/19/94	414nt	1214-1586	1B	55°C (2x), 53°C (8x), 51°C (35x)
1307F-AB 5' - TCAGACCGAGTGTCTCAAGA - 3' (F) 2113R-AB 5' - CAGTTCCCACCATGCGGAAGTG - 3' (R)	1194-1213 2436-2457	KNP/17/94 KNP/19/94	1264nt	1214-2435	1B-1C	59°C (2x), 57°C (8x), 55°C (35x)
1347F-AB 5' - CGGTCACAAACACCCARTCAAC - 3' (F) 2113R-AB 5' - CAGTTCCCACCATGCGGAAGTG - 3' (R)	1228-1249 2436-2457	KNP/17/94 KNP/19/94	1230nt	1250-2435	1B-1C	57°C (2x), 55°C (8x), 52°C (35x)
1904F-AB 5' - GCTTCCTGGACTTTGGTGAT - 3' (F) 2868R-AB 5' - CCAATGGGACTAGACCTATCATA - 3'(R)	1791-1810 2733-2755	SAR/1979	965nt	1811-2732	1B-1D	51°C (2x), 50°C (8x), 49°C (35x)
2246MF-JvS 5' - GTACCGTGGATCACTAGTCTA - 3' (F) 3204MR-JvS 5' - GACAGCAGGTAGGACAGACAA - 3'(R)	2249-2269 3213-3233	KNP/17/94 KNP/19/94	985nt	2270-3212	1C-1D	58°C (2x), 57°C (8x), 56°C (35x)
VP3-FM 5' - CACTTCCGCATGGTGGGAACTG - 3' (F) 2B-RM 5' - CGGCAGTAGGGTTTGAGCCATT - 3'(R)	2736-2457 3947-3968	KNP/17/94 KNP/19/94 SAR/1979	1233nt	2458-3946	1D-2B	60°C (2x), 57°C (8x), 55°C (35x)
EMC-3DfAB 5' - TCAGGTTGTGCAGCGACCTC - 3' (F) EMC-3DrAB 5' - CTTACCGGGTAACGCGTTGT - 3'(R)	7128-7147 7651-7670	KNP/17/94 KNP/19/94 SAR/1979	318nt	7313-7630	3D	55°C (2x), 53°C (8x), 51°C (35x)

*Positions based on Mengovirus isolate Rz-pMwt, GenBank accession number DQ294633.

2.2.3. Nucleotide sequencing and analysis

PCR products of the expected size were purified using a High Pure PCR Product Purification Kit (Roche Applied Science). Dye-terminator cycle sequencing was performed using the ABI PRISM Big Dye™ Terminator version 3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems), with each of the PCR primers at supplier-prescribed reaction conditions. Reactions were run on an ABI 3130 automated sequencer (Applied Biosystems), viewed with CHROMAS v. 1.43 and aligned to a reference sequence DQ294633, a cloned Mengovirus, termed Rz-pMwt derived from the 'M' strain (Palmenberg, pers. comm.) in MEGA v.5 (Tamura *et al.*, 2011). Alignment of the overlapping fragments resulted in a contiguous full-length P1 gene sequence inclusive of the flanking leader (L) protein and 2A coding regions. Each of the resulting South African EMCV sequences, ranging from 3,371 nucleotides (for the historical swine virus) to 3688 nucleotides (for the two elephant strains) in length, were deposited in the GenBank database under accession numbers, JN800421-3 and used in nucleotide Blast searches (www.ncbi.nlm.nih.gov/blast) to identify all available closely related sequences (Altschul *et al*., 1997). The 3D sequences generated for the South African isolates, were submitted to GenBank under accession numbers JX102656-8.

Three datasets were compiled for phylogenetic analysis, comprising of all available, homologous EMCV data in the GenBank database, corresponding to the P1 and 3D genome regions characterised in this study, and to the VP3/1 junction. The final, aligned P1 dataset used for phylogenetic analysis and comprising the three historical South African viruses and 21 virus sequences available in GenBank (www.ncbi.nlm.nih.giv),was 3329 nucleotides in length, and had an average empirically-determined nucleotide base frequency of A = 25.2 %,C = 26.1 %, G = 22.6 % and T = 26.2 %. The transition:transversion ratio (R) was 2.0 and 993 of the 1143 variable sites were parsimony informative (Pi). The VP3/1 junction dataset, which was 715 nucleotides in length, comprised 42 taxa, had an average base frequency of A = 25.2 %, C = 26.0 %, G = 21.4 % and T = 27.4 %, an R of 1.71 and of the 307 variable sites, 274 were parsimony informative. The 248 nucleotide long 3D dataset consisted of 70 taxa, had an average empirically-determined nucleotide frequency of A = 27.4 %, C = 24.9 %, G = 20.6 % and T = 27.1 %, an R of 2.3 and 82 of the 103 variable sites were parsimony informative. The best-fit model selected

under the Akaike Information Criterion (AIC) in MEGA v. 5, for each of the datasets was: the General Time Reversible (GTR) model, with a Gamma distribution shape parameter (G) of 1.78 and proportion of invariant sites (I) of 0.58 for the P1 dataset; the Tamura 3-parameter (T92) model, with a G of 3.08 and I of 0.526 for the VP3/1 dataset; and the T92 model with a G of 0.29 for the 3D dataset. The best-fit model identified for each of the datasets guided parameter selection for the Minimum evolution (ME) and Maximum Likelihood (ML) analyses performed in MEGA v. 5 (Tamura *et al.*, 2011) and PhyML v 3.0 (Guindon & Gascuel, 2003), respectively and for the selection of priors for the Bayesian inference (BI) performed with Mr Bayes v 3.1 (Huelsenbeck & Ronquist, 2001). Nodal support for ME and ML was assessed by 10000 and 5000 nonparametric bootstrap replications, respectively. For the BI analyses four chains were run for 10×10^6 generations using random starting trees with default heating and swap parameters. Resulting trees and parameters were recorded and split frequencies were compared every 1000^{th} generation to ensure convergence. Tracer plots were inspected to ensure that trees were sampled from the region of stationarity, and two independent runs were performed for each dataset.

2.3. Results

Phylogenetic inferences of all three genome regions confirmed the sister relationship between the 1994 KNP elephant outbreak strains and the Mengovirus isolates of the EMCV-1 serotype of EMCV (Figures 2.1-2.3). High levels of support were obtained for this grouping for all three gene regions and across all methods of phylogenetic inference. In contrast, although the P1 analysis (Figure 2.1) confirmed the sister relationship of the 1979 swine outbreak strain from South Africa (SAR/1979) to a lineage that is comprised primarily of isolates derived from global (China, Korea, Belgium and Panama) swine EMCV outbreaks and of rodent strains from the USA (100 % for node J), nodal support for this relationship was ≤ 75 % and ≤ 57 % in the VP3/1 (Figure 2.2, lineage I) and 3D (Figure 2.3, lineage IV) phylogenies, respectively.

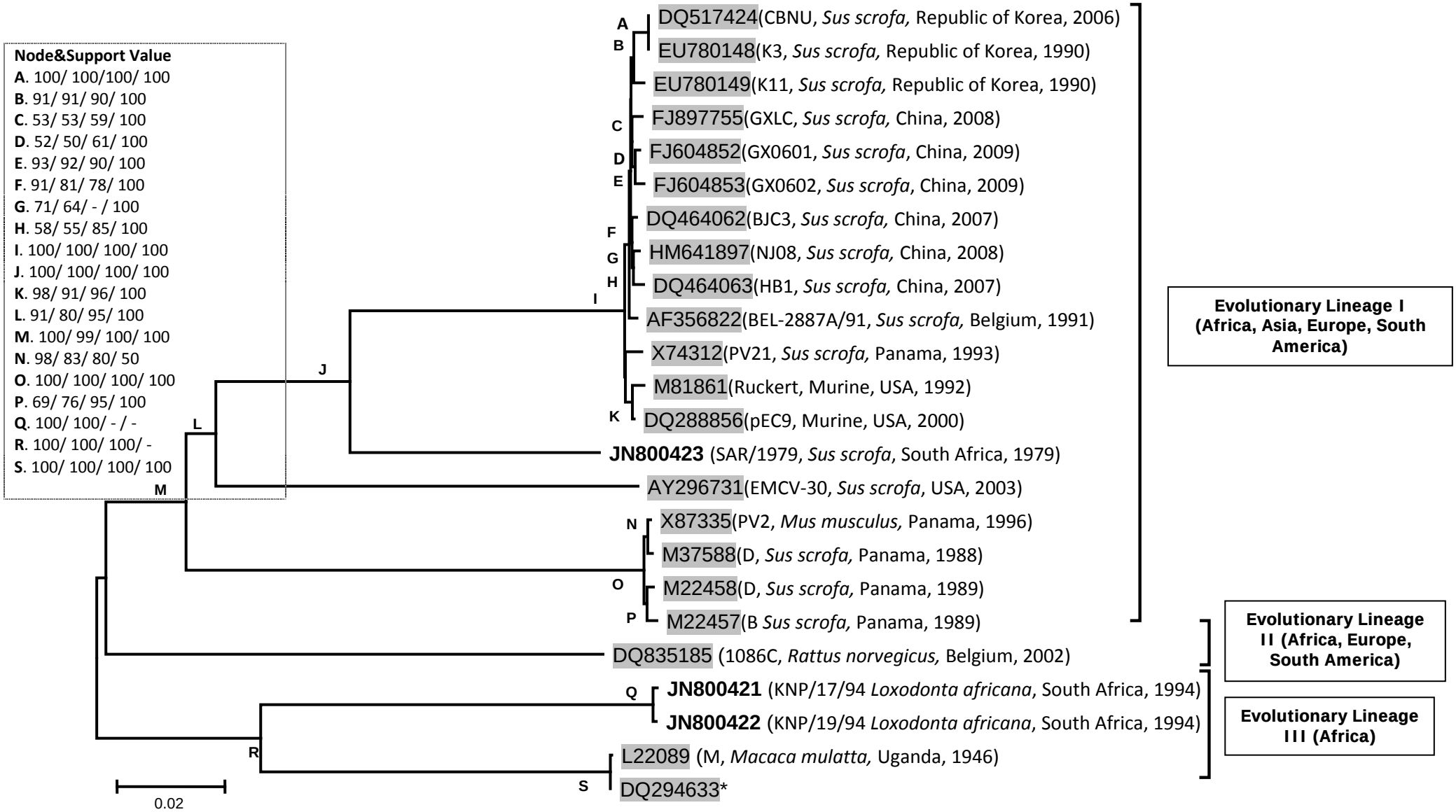

Figure 2.1 Neighbour Joining (NJ) tree indicating *Encephalomyocarditis virus* relationships based on complete P1 and flanking L and 2A nucleotide data. Nodal support values ≥ 50 % from Minimum Evolution (ME), Maximum likelihood (ML) and Bayesian inference (BI) analyses are summarized for each of the nodes (labeled A-S) as follows: NJ / ML / ME / BI. Sequences highlighted in grey correspond to those used in both datasets, whilst those generated in this study are indicated in bold. Taxon labels comprise the GenBank accession number, followed in brackets by the isolate name, host species, country of origin and year of isolation. A summary of the isolates included may be found in Appendix 2.1. The geographical distribution of each lineage is summarized behind the lineage name.

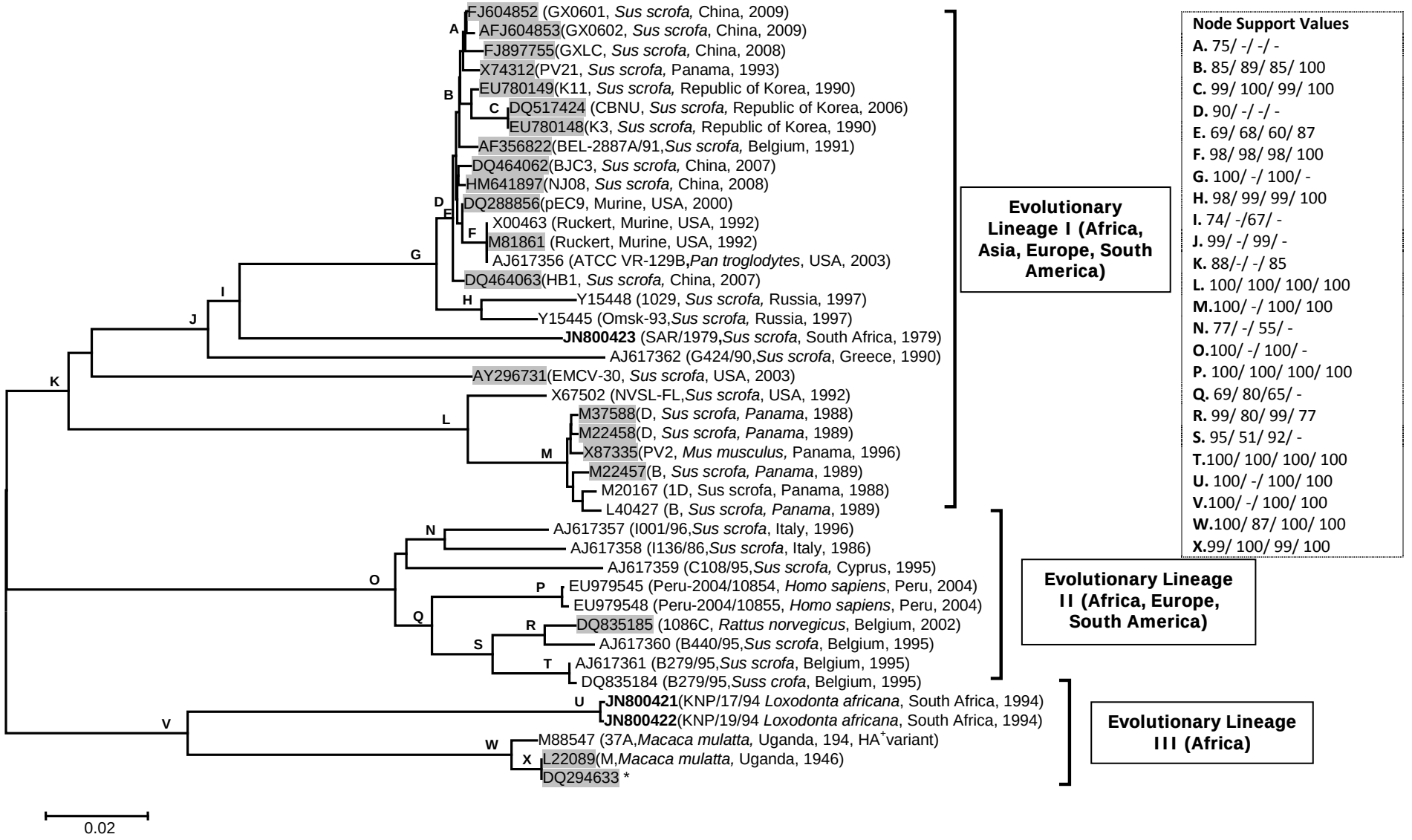

Figure 2.2 Neighbour Joining (NJ) tree inferred using an 831 nt region of the VP3/1 junction of 42 encephalomyocarditis viruses. Nodal support values ≥ 50 % from Minimum Evolution (ME), Maximum likelihood (ML) and Bayesian inference (BI) analyses are summarized for each of the nodes (labeled A-X) as follows: NJ / ML / ME / BI. Sequences highlighted in grey correspond to those used in both datasets, whilst those generated in this study are indicated in bold. Taxon labels comprise the GenBank accession number, followed in brackets by the isolate name, host species, country of origin and year of isolation. A summary of the isolates included may be found in Appendix 2.1. The geographical distribution of each lineage is summarized behind the lineage name.

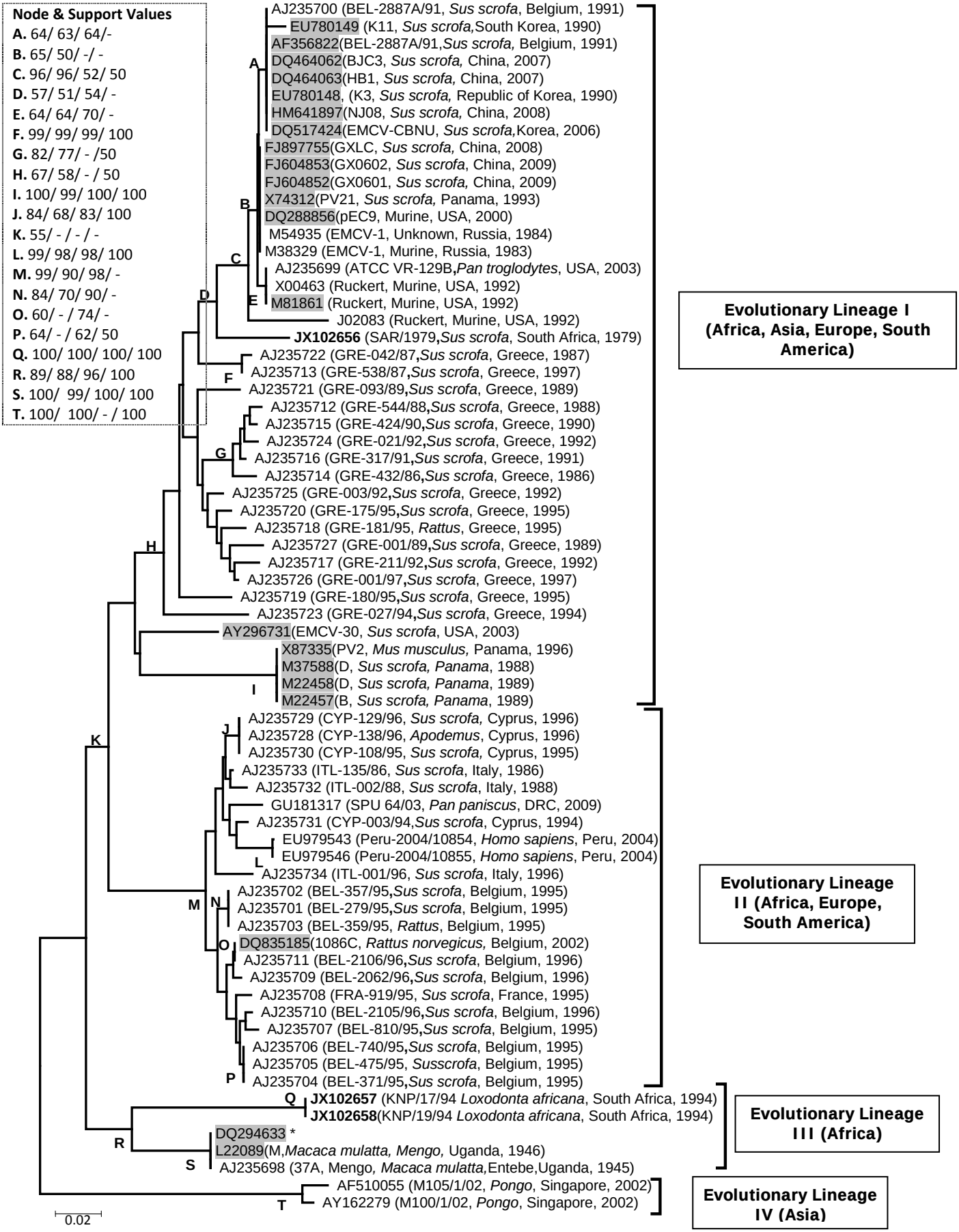

Figure 2.3 Neighbour Joining (NJ) tree indicating encephalomyocarditis virus relationships based on partial 3D genome region data. Nodal support values and taxon labels as per Figs 2.1 and 2.2. Refer to Appendix 2.1 for a complete isolate summary.

2.3.1. Nucleotide sequence analysis

The P1 and VP3/1 phylogenies revealed the presence of three major EMCV lineages, denoted evolutionary lineage I to III (Figures 2.1 & 2.2) and corresponding to previously defined groups (Koenen *et al.*, 1999). The 3D gene tree recovered four lineages (denoted I-IV in Figure 2.3), of which three correspond to those recovered in the P1 and VP3/1 phylogenies. Based on available 3D data, the lineages vary in their geographical distribution, with lineage I occurring on five continents, lineage II on three continents and lineages III and IV each occurring on a single continent. Lineages I and II both have wide host ranges that include domestic pigs, murid rodents and primates, whereas lineage III (the Mengovirus lineage), has only been found in monkeys from Uganda and in African elephants from South Africa. Lineage IV is only known from two unpublished GenBank records that list orangutan as the host species (Figure 2.3). Average within-lineage nucleotide distances for lineages I-III ranged from 5.5 % (within lineage II) to 9.8 % (within lineage I) for the VP3/1 junction region, and from 2.2 % (within lineage II) to 6.6 % (within lineage III), for the 3D region. Between-lineage nucleotide distances ranged from 20.6 % (between lineage I and II) to 22.8 % (between lineages II and III) for VP3/1, and from 12.4 % (between lineage I and II) to 15.5 % (between lineages I and III) for the 3D region.

Comparison of the two elephant viruses, designated KNP/17/94 and KNP/19/94, revealed five variable nucleotide sites across the P1 and flanking genome region characterised, corresponding to an overall nucleotide sequence identity of 99.8 % across the ~3.3 kbp region. The five variable sites all occurred at third base positions and corresponded to silent mutations at amino acid position 55 of the Leader (L) protein; amino acid positions 54, 157 and 256 of the VP2 (1B) gene, and at amino acid position 158 of the VP1 (1D) gene. For the VP3/1 and 3D regions analysed, the pairwise nucleotide sequence identity was 99.86 % and 100 %, respectively for the two elephant isolates.

The two near-identical elephant-derived isolates, KNP/17/94 and KNP/19/94, which form a sister clade to Mengovirus isolates obtained from infected Rhesus macaques (*Macaca mulatta*) in Uganda differed by 13.3 % from the 'M' strain (L22089), across the P1 genome region characterised. Pairwise nucleotide sequence differences between each of the elephant strains and EMC viruses falling outside of lineage III (the elephant-Mengovirus lineage) was ≥ 19 % across the P1 region.

Nucleotide sequence differences between the elephant and Mengovirus strains were 15 % and 11 % across the VP3/1 and 3D regions respectively.

The strains showed variation at 337 of 2292 nucleotides of the VP3 - VP1 gene coding region, resulting in 25 of the 764 coded amino acid residues differing from Mengovirus isolate Rz-pMwt (DQ294633) and Mengovirus isolate M (L22089), summarised in the amino acid alignment in Figure 2.4. Genetic distances across the surface exposed VP3-VP1 region between both KNP strains and their two closest GenBank matches, the two wild type Mengovirus isolates Rz-pMwt (DQ294633) and M (L22089), were calculated applying the Tamura-Nei model (Tamura & Nei, 1993). The two wild type Mengoviruses differed by a value of 0.202 between the published isolates and KNP/17/94 and by 0.201 to KNP/19/94. Across this same region the two KNP isolates differed by a genetic distance of 0.002 and contained 4 nucleotide mutations which resulted in synonymous amino acid substitutions. Across the entire P1/partial-P2-amplicon, isolates KNP/17/94 and KNP/19/94 both showed the shortest genetic distance when compared to DQ294633 (0.188 and 0.187 respectively), KNP/19/94 showed an equal genetic distance to L22089 while KNP/17/94 showed a fractionally larger distance of 0.189. Similarly the results of the 3D analysis once again showed that the two elephant derived strains are closest to Mengo isolates, DQ294633 and L22089 (Genetic distance between KNP strains and Mengo isolates is calculated as 0.140 in all cases). Genetic distance calculated in MEGA v. 5 (Tamura *et al.*, 2011) for both the entire P1 and flanking coding regions as well as the 3D gene regions are summarised in Tables 2.4 and 2.5.

Genetic distance between isolates used in the study and their closest matches on GenBank calculated using MEGA v. 5 (Tamura *et al.*, 2011) are summarised in tables 2.3 and 2.4. For the historical swine virus, SAR/1979, pairwise nucleotide sequence similarity was highest (>89.9%) across the ~3.3 kbp (P1/flanking gene) region, to three lineage I EMCV strains (DQ288856, DQ717424, AF356822) of diverse geographical origin (USA, Korea and Belgium). For the VP1/3 and 3D genome regions for which substantially more data were available, SAR/1979 was most closely related (89 % sequence identity) to four isolates from China and Korea (HM641897, FJ604852, DQ517424, DQ464063) across the VP3/1 genome region, and had the highest nucleotide sequence identity (96 %) across the 3D gene region, to

a domestic pig virus from an outbreak that occurred in Greece in 1997 (AJ235726). Genetic distances calculated across the entire P1 and flaking regions, as well the 3D gene region are summarised in Table 2.3.

```
Mengo_RzpMwt_DQ294633 DQNTEEMENLSDRVSQDTAGNTVTNTQSTVGRLVGYGTVHDGEHPASCADTASEKILAVERYYTFKVNDWTSTQKPFEYI [ 80]
Mengo_L220891         ................................................................................ [ 80]
KNP/17/94             ..D............................................................................. [ 80]
KNP/19/94             ..D............................................................................. [ 80]

Mengo_RzpMwt_DQ294633 RIPLPHVLSGEDGGVFGATLRRHYLVKTGWRVQVQCNASQFHAGSLLVFMAPEYPTLDVFAMDNRWSKDNLPNGTRTQTN [160]
Mengo_L220891         ................................................................................ [160]
KNP/17/94             .V..............A.....................................................T..... [160]
KNP/19/94             .V..............A.....................................................T..... [160]

Mengo_RzpMwt_DQ294633 RKGPFAMDHQNFWQWTLYPHQFLNLRTNTTVDLEVPYVNIAPTSSWTQHASWTLVIAVVAPLTYSTGASTSLDITASIQP [240]
Mengo_L220891         ................................................................................ [240]
KNP/17/94             K............................................................................... [240]
KNP/19/94             K............................................................................... [240]

Mengo_RzpMwt_DQ294633 VRPVFNGLRHEVLSRQSPIPVTIREHAGTWYSTLPDSTVPIYGKTPVAPANYMVGEYKDFLEIAQIPTFIGNKVPNAVPY [320]
Mengo_L220891         ................................................................................ [320]
KNP/17/94             ...........T.................................................................... [320]
KNP/19/94             ...........T.................................................................... [320]

Mengo_RzpMwt_DQ294633 IEASNTAVKTQPLAVYQVTLSCSCLANTFLAALSRNFAQYRGSLVYTFVFTGTAMMKGKFLIAYTPPGAGKPTSRDQAMQ [400]
Mengo_L220891         ................................................................................ [400]
KNP/17/94             ............T.......T............................................................ [400]
KNP/19/94             ............T.......T............................................................ [400]

Mengo_RzpMwt_DQ294633 ATYAIWDLGLNSSYSFTVPFISPTHFRMVGTDQANITNVDGWVTVWQLTPLTYPPGCPTSAKILTMVSAGKDFSLKMPIS [480]
Mengo_L220891         ................................................................................ [480]
KNP/17/94             ..............................VD................................................ [480]
KNP/19/94             ..............................VD................................................ [480]

Mengo_RzpMwt_DQ294633 PAPWSPQGVENAEKGVTENTDATADFVAQPVYLPENQTKVAFFYDRSSPIGAFAVKSGSLESGFAPFSNKACPNSVILTP [560]
Mengo_L220891         ................................................................................ [560]
KNP/17/94             ................A.............................T..............ET......... [560]
KNP/19/94             ................A.............................T..............ET......... [560]

Mengo_RzpMwt_DQ294633 GPQFDPAYDQLRPQRLTEIWGNGNEETSEVFPLKTKQDYSFCLFSPFVYYKCDLEVTLSPHTSGAHGLLVRWCPTGTPTK [640]
Mengo_L220891         ................................................................................ [640]
KNP/17/94             .......R.............D..K.....S.............................N.............. [640]
KNP/19/94             .......R.............D..K.....S.............................N.............. [640]


Mengo_RzpMwt_DQ294633 PTTQVLHEVSSLSEGRTPQVYSAGPGTSNQISFVVPYNSPLSVLPAVWYNGHKRFDNTGDLGIAPNSDFGTLFFAGTKPD [720]
Mengo_L220891         ................................................................................ [720]
KNP/17/94             ........................I.....................................S.....H.............. [720]
KNP/19/94             ........................I.....................................S.....H.............. [720]

Mengo_RzpMwt_DQ294633 IKFTVYLRYKNMRVFCPRPTVFFPWPTSGDKIDMTPRAGVLMLE [764]
Mengo_L220891         ........................................... [764]
KNP/17/94             .....F...................S.............R.. [764]
KNP/19/94             .....F...................S.............R.. [764]
```

Figure 2.4 An amino acid alignment indicating the 25 variable sites out of 764 coded amino acids that differed between the KNP derived isolates, KNP/17/94 & KNP/19/94, and the two wild type Mengovirus isolates Rz-pMwt (DQ294633) and Mengovirus isolate M (L22089) across the 2296 nt VP3 – VP1 surface exposed regions of the P1 gene region characterised in this study.

```
EMCVDQ288856 DQNTEEMENL SDRVSQDTAG NTVTNTQSTV GRLVGYGTVH DGEHPASCAD TASEKILAVE RYYTFKVNDW TSTQKPFEYI  [ 80]
SAR/1979     .......... .......... .......... .......A.. .......... .......... .......... ..........  [ 80]

EMCVDQ288856 RIPLPHVLSG EDGGVFGAAL RRHYLVKTGW RVQVQCNASQ FHAGSLLVFM APEYPTLDAF AMDNRWSKDN LPNGTRTQTN  [160]
SAR/1979     ......I... .......... .....M.... ...G...... .......... .......... .......... .....K....  [160]

EMCVDQ288856 KKGPFAMDHQ NFWQWTLYPH QFLNLRTNTT VDLEVPYVNI APTSSWTQHA SWTLVIAVVA PLTYSTGAST SLDITASIQP  [240]
SAR/1979     .......... .......... .......... .......... .......... .......... .......... ..........  [240]

EMCVDQ288856 VRPVFNGLRH ETLSRQSPIP VTIREHAGTW YSTLPDSTVP IYGKTPVAPS NYMVGEYKDF LEIAQIPTFI GNKIPNAVPY  [320]
SAR/1979     .......... .......... .......... .......... .......... .......... .......... ..........  [320]

EMCVDQ288856 IEASNTAVKT QPLATYQVTL SCSCLANTFL AALSRNFAQY RGSLVYTFVF TGTAMMKGKF LIAYTPPGAG KPTSRDQAMQ  [400]
SAR/1979     .......... .......... ..A....S.. .......... .......... .......... .......... ..........  [400]

EMCVDQ288856 ATYAIWDLGL NSSYSFTVPF ISPTHFRMVG TDQVNITNAD GWVTVWQLTP LTYPPGCPTS AKILTMVSAG KDFSLKMPIS  [480]
SAR/1979     .......... .......... .......... ........V. .......... .......... .......... ..........  [480]

EMCVDQ288856 PAPWSPQGVE NAEKGVTENT DATADFVAQP VYLPENQTKV AFFYDRSSPI GAFTVKSGSL ESGFAPFSNG TCPNSVILTP  [560]
SAR/1979     .......... .........A .......... .......... .......... .......... .........E V.........  [560]

EMCVDQ288856 GPQFDPAYDQ LRPQRLTEIW GNGNEETSKV FPLKSKQDYS FCLFSPFVYY KCDLEVTLSP HTSGNHGLLV RWCPTGTPTK  [640]
SAR/1979     ......S... .......... ......A... .......... .......... .......... .......... ..........  [640]

EMCVDQ288856 PTTQVLHEVS SLSEGRTPQV YSAGPGISNQ ISFVVPYNSP LSVLPAVWYN GHKRFDNTGS LGIAPNSDFG TLFFAGTKPD  [720]
SAR/1979     .......... .......... .......... .......... .......... .......... .......... ..........  [720]

EMCVDQ288856 IKFTVYLRYK NMRVFCPRPT VFFPWPTSGD KIDMTPRAGV LMLE [764]
SAR/1979     .......... .......... .......... .......... .... [764]
```

Figure 2.5 An amino acid alignment indicating the 13 variable sites out of 764 coded amino acids that differed between the SAR/1979 and closest GenBank match EMCV strain pEC9 (DQ288856)across the 2296 nt VP3 – VP1 surface exposed regions of the P1 gene region characterised in this study.

Table 2.3 Estimates of nucleotide pairwise-divergence between P1 sequences are shown in the top right of the table, while those between 3D Sequences are shown in the bottom left of the table. Analyses were conducted using the Tamura-Nei model (Tamura &Nei, 1993). There were a total of 3692 and 248 positions in the final P1 and 3D dataset set respectively. Evolutionary analyses were conducted in MEGA v. 5 (Tamura *et al.,* 2011).

	KNP/17/94	**KNP/19/94**	**SAR/1979**	**DQ288856**	**DQ294633**	**L220891**	**M81861**	**J02083**
KNP/17/94		0.001	0.265	0.299	0.188	0.189	0.302	
KNP/19/94	0.000		0.266	0.299	0.187	0.187	0.302	
SAR/1979	0.231	0.231		0.125	0.250	0.251	0.128	
DQ288856	0.267	0.267	0.249		0.277	0.277	0.002	
DQ294633	0.140	0.140	0.166	0.156		0.000	0.278	
L220891	0.140	0.140	0.166	0.156	0.000		0.279	
M81861	0.273	0.273	0.066	0.004	0.161	0.161		
J02083	0.312	0.312	0.097	0.044	0.196	0.196	0.043	
AJ235726	0.225	0.225	0.044	0.069	0.161	0.161	0.074	0.114

Sequence variation, in the characterised Mengovirus neutralization epitopes (Table 2.1) between Mengo strains M (L220891)/RzpMwt (DQ294633) and the two elephant derived KNP isolates is summarised in Table 2.4. Likewise, the sequence variation between SAR/1979 and closest GenBank match EMCV strain pEC9 (DQ288856) at the epitope sites identified for Mengo isolates (Table 2.1) is summarised in Table 2.5.

Table 2.4 Changes in nucleotide and amino acid sequence between elephant strains (KNP/17/94 &KNP/19/94) and wild type Mengo strains M (L220891)/RzpMwt (DQ294633) at identified epitope sites.

Nucleotide Position[a]	**Amino acid change**[b]		
	VP1	**VP2**	**VP3**
1601 C -> T (3rd)		2144Asn ->Asn [c, d, e]	
1608 T -> A (1st) 1610 C -> T (3rd)		2147 Ser ->Thr [c, e]	
2141 C -> T (3rd)			3068 Ser ->Ser [c, e]
2162 A -> G (3rd)			3075 Gln ->Gln[d]

a. Nucleotide position based on Mengo isolate M (L220891)
b. Position indicated in the conventional way according to the Rossman system, first digit refers to the capsid polypeptide (1, 2 or 3) and the next three identifies its position in the chain from the N-terminus.
c. Determinate amino acid residue of epitopes identified by Boege *et al.*, (1991)
d. Mutations and determinate amino acid residue of epitopes identified by Kobasa *et al.*, (1994)
e. Determinate amino acid residue of epitopes listed in Scraba & Palmenberg (1999)

Table 2.5 Changes in nucleotide and amino acid sequence between SAR/1979 and closest GenBank match EMCV strain pEC9 (DQ288856) at the epitope sites identified for Mengo isolates

Nucleotide Position[a]	**Amino acid change**[b]		
	VP1	**VP2**	**VP3**
2928 A -> G(1st)[d]	1100 Thr ->Ala[d, e]		
1217 A -> G(3rd)[d]		2016 Gln ->Gln [d]	
1604 T -> C (3rd)		2145 Arg->Arg[c, e]	
2141 C -> T (3rd)			3068 Ser ->Ser [c, e]
2160 A -> G (3rd)[d]			3075 Gln ->Gln[d]

a. Nucleotide position based on Mengo isolate M (L220891)
b. Position indicated in the conventional way according to the Rossman system, first digit refers to the capsid polypeptide (1, 2 or 3) and the next three identifies its position in the chain from the N-terminus.
c. Determinate amino acid residue of epitopes identified by Boege *et al.* (1991)
d. Mutations and determinate amino acid residue of epitopes identified by Kobasa *et al.* (1994)
e. Determinate amino acid residue of epitopes listed in Scraba & Palmenberg (1999)

2.4. Discussion

The EMCV isolates used for phylogenetic inference showed the presence of three main lineages, corresponding to the previously defined groups for Koenen *et al.*, 1999. Addition a fourth lineage (IV) was shown to occur exclusively in Asia. Reports of EMCV in wild &/ exotic animals are limited and sporadic in origin, generally occurring in captive populations. These outbreaks have been recorded globally in North America, Australia, Asia and Europe as was the case with porcine epidemics the disease was either found to be present in the associated rodent populations or associated with a rodent irruption (Wells *et al.,* 1989; Reddacliff *et al.,* 1997; Canelli *et al.*, 2010; Liu *et al.,* 2013, Yeo *et al.,* 2013). Africa is the exception to this general trend being the only continent in which outbreaks of the EMCV have occurred with a native population of animals, these being reported from a population of semi-wild primates (Jones *et al.,* 2011) and from a natural population of elephants (Grobler *et al.,* 1996). The continent was also the origin of one of the first EMC virus isolates to be recorded, *viz.* the Mengovirus 'M' strain isolated from Rhesus macaques in Uganda (Dick *et al.,* 1949).

The origin of the two nearly identical elephant viruses is unclear, but given that the outbreak occurred in a natural, free-ranging population of elephants, and was

shown to be most closely related to a strain of African origin, it is surmised that the sequences generated in this study represent a regionally endemic virus lineage, since to date no strains belonging to this lineage have been recovered from non-African outbreaks. EMCV has shown great variation in the diversity of species it is able to infect (Reddacliff *et al*., 1997). Despite this, no carcasses of other species present in the KNP reserve at the time of the EMCV outbreak in elephants were linked to the virus. Smaller carcasses, are readily scavenged and thus do not provide an accurate indication of all species affected by the outbreak. However, a lack of larger antelope carcasses was also recorded. As larger mammal carcasses remain intact and therefore detectable during anthrax outbreaks, it would appear that seemingly few of the larger mammalian species were affected by EMCV. Although speculative, it has been proposed that the disease may have played a part in the drop in myomorphic rodent populations recorded in the park after 1994 (Grobler *et al.*, 1995).

Members of the order Rodentia are seen as the likely potential reservoir hosts of the virus, with the disease being spread through the faecal-oral route and infection occurring following the ingestion of food stuffs contaminated with either faeces or urine (Gainer *et al.*, 1968; Grobler *et al.*, 1995). From 1993-1994, Grobler and co-workers (1995) reported a trapping success rate to 54 % and 56 % respectively, compared to just 4% and 3 % in 1989 and 1990 respectively. This drastic increase in trapping success coincided with an increase in infection rate of EMCV within elephants, 8.3 % in 1993 and 25.2 % in 1994 (Grobler *et al.*, 1995), as determined by retrospective serological screening. It is believed that this increase in potential reservoir hosts numbers likely contributed to the spread of the virus in the park (Grobler *et al*., 1995). Serological screening of a range of rodents caught during the elephant outbreak showed that *Mastomys* individuals recorded the highest seroprevalence (38 %), making members of this rodent genus the most likely candidate to recover the virus from (Grobler *et al.*, 1995). In addition, unpublished records at the National Institute for Communicable Diseases (NICD) indicate that 26 EMCV isolates were made from rodents, mainly *Mastomys*, collected in surveys at four widely separated locations in South Africa, between 1961 and 1968 (personal communication). Although the rodent isolates were not available for inclusion in the present study, the implication is clear that there is endemic circulation of virus in the country.

Differences between the two isolates derived from elephant carcasses, KNP/17/94 and KNP/19/94 differed primarily in the capsid coding regions of P1, i.e. those coding the structural proteins VP4, VP2, VP3 and VP1, which combine to form a non-enveloped, icosahedral virus outer shell (Simmonds, 2006). The mutations recorded between the two elephant isolates did not occur at the neutralization epitopes identified in previous studies (Table 2.1). There were, however, mutations in nucleotides that resulted in both synonomous and in 25 non-synonomous amino acid substitutions between the KNP elephant derived isolates and wild type Mengo isolates, Mengo strains M (L220891)/RzpMwt (DQ294633). A similar trend was observed between the SAR/1979 strain and its closest GenBank match EMCV strain pEC9 (DQ288856). Sequence analysis of these regions is important to infer a more accurate representation of the virus' phylogeny and for determining its pathogenicity (Zang *et al.,* 1997; Nelsen-Salz *et al.,* 1996; Knowles *et al.,* 1998; Koenen *et al.,* 1999; Denis *et al.,* 2006;). The two elephant viruses grouped with mengoviruses within lineage III. As members of the family *Picornaviridae* generally cluster according to geographic origin (Cummings *et al.,* 2007), it is likely that the elephant viruses, which were most closely related to an isolate of East African origin (Figures 2.3-2.5), represents a regionally endemic lineage.

Despite a long presence on the continent, isolated reports of the disease in pig are confined to Cote d'Ivoire and South Africa. As previously stated the only other outbreak in South Africa prior to the KNP outbreak was restricted to swine, with the most prominent incident occurring in the Natal Midlands of Kwa-Zulu Natal Province in 1979, severely affecting the local population (Williams, 1981). The isolate obtained from the swine outbreak clustered more closely with isolates obtained from global swine EMCV. Based on the genetic distance derived from the P1 gene region, it was found that the SAR/1979 strain most similar to the encephalomyocarditis virus strain pEC9 (DQ288856), (90 % sequence similarity match on GenBank), originally submitted by Martin *et al.,* (2000), (Figure 2.3) and clustered within lineage I which is comprised of viruses of a wide geographical and host range. Based on 3D gene region analysis the SAR/1979 strain is once again most closely related to encephalomyocarditis virus strain GRE-001/97 (AJ235726, genetic distance 0.044), isolated from swine originating from Greece. At the time of the 1979 outbreak import of swine into South Africa was restricted from the USA due to an outbreak of

Aujeszky's Disease and importation from European countries was suspended due to possible outbreaks of Classical Swine fever and Swine Vesicular Disease. The probable geographical origin of the strain imported into South Africa therefore remains unclear.

Future studies should therefore be directed at determining the rodent reservoir host species, using a combination or serology and virus isolation in order to determine the prevalence and diversity of EMC viruses in endemic South African rodents. The data generated in this study provides valuable regional reference sequences, which together with intensified sampling and characterisation of contemporary viruses and additional retrospective studies, should improve the accuracy with which the geographical origin of outbreaks can be determined.

2.5. Acknowledgements

We would like to thank Dr. Gerdes of the ARC-OVI for providing an EMCV RNA extract for positive control purposes and Prof. Ann Palmenberg for information on the Rz-pMwt isolate. The project was funded by a DST-NRF South African Research Chair of Mammalian Behavioural Ecology and Physiology awarded to NCB.

CHAPTER 3

Standardisation and application of a duplex PCR for molecular assessment of *Cardiovirus* prevalence in murid rodents from the Kruger National Park, South Africa

J. H. d. T. van Sandwyk, N. C. Bennett & A. D. S. Bastos

Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa

Chapter Summary

Two outbreaks of encephalomyocarditis virus (EMCV) have occurred in South Africa. The first in the Natal Midlands of Kwa-Zulu Natal Province, South Africa in 1979 was restricted to swine. The second outbreak, which occurred from December 1993 to November 1994, in the Kruger National Park (KNP), primarily affected African elephants (*Loxodonta africana*). The latter coincided with a marked increase in murid rodent trap captures and serological and PCR screening indicated that the highest EMCV prevalence was in rodents of the genus *Mastomys*. In the absence of definitive confirmation that this rodent genus was the reservoir of the virus causing mortalities in elephants, the results remain speculative. The focus of this study was to assess EMCV presence and diversity in a range of murid rodent species occurring within the historical elephant outbreak locality. The increased availability of EMCV genome data, since the report of the elephant outbreak in 1995, and the demonstrated presence of two discrete lineages in South Africa, necessitated a re-evaluation of available diagnostic PCR assays. A duplex PCR format in which host and virus genomes are targeted was selected for optimisation and application to a range of clinical murid rodent samples, including heart, gastrointestinal tract (GIT) and kidney. The latter two sample types permit evaluation of the faecal-oral route of infection and possibility of urinary viral excretion, respectively. To this end, heart, kidney, and GI

tract tissues from 266 rodents from the KNP, corresponding to a total sample size of 798 were screened for EMCV genome presence using the duplex PCR assay. The alpha tubulin gene transcript served as an internal control and was combined with primers targeting the 5'UTR region. In addition, each tissue type was also screened with primers targeting the 3D gene region. The rodent species assessed varied in percentage composition with *Mastomys natalensis* being the most abundant and making up 48.5 % of the total. In conjunction with *M. coucha*, rodents of the genus *Mastomys* represented 70.3 % of the murid rodents captured. No infected rodents were identified for this or any of the three other rodent genera (*viz. Lemniscomys* (6.8 %), *Gerbilliscus* (7.5 %) and *Saccostomus* (15.4 %)), screened. These results suggest that if *Mastomys* is indeed the reservoir of infection, then inter-epidemic infection rates are extremely low. The establishment of a rapid, multi-gene, molecular screening method holds value for molecular assessment in murid rodents and other species, in the event of an outbreak.

3.1. Introduction

The family *Picornaviridae* contain a number of viral species that have demonstrated the ability to infect a variety of mammalian hosts, including humans, some of which are of public health concern. The genus *Enterovirus* contains a number of viruses of medical importance including serotypes of *Rhinovirus*, the causative agent of many upper and lower respiratory tract infections (Palmenberg *et al*, 2009), and the poliovirus agent, that has been responsible for a great deal of human suffering (De Jesus, 2007). Genera of veterinary importance include *Aphthovirus*, containing foot-and-mouth disease virus (FMDV) and *Cardiovirus* (Fenner *et al.*, 1993; Knowles *et al.*, 2012; Simmonds, 2006). The genus *Cardiovirus* comprises two currently recognized species, namely *Encephalomyocarditis virus* (EMCV), and *Theilovirus* (Fenner *et al.*, 1993; Knowles *et al*., 2012; Chapter 1).

Clinically EMCV infection presents as encephalitis and myocarditis. As mentioned previously (Chapter 2), two outbreaks have occurred in South Africa, the first in the Natal Midlands (1979) (Williams, 1981), and the second, within the Kruger National Park (KNP). The KNP outbreak was initially detected when a cluster of four elephant carcasses were observed late in December 1993, in the far north of the

reserve and reached its peak in January 1994 when 32 deaths were recorded. Retrospective analysis of outbreak data revealed that the elephant mortalities formed two main clusters. The first was localized in the southern and central regions, south of the Tropic of Capricorn, from the Satara ranger station in the north to the Crocodile Bridge ranger station on the southern border of the park (Figure 3.1). The second cluster of mortalities was located in the far north of the park, from the Shingwedzi ranger station and primarily along the Shingwedzi River extending northwards to the Punda Maria ranger station near the border with neighbouring Zimbabwe (Grobler *et al*., 1995; Figure 3.1).

Serological screening of captured rodents during this period revealed that 37.9 % (100/264) of *Mastomys natalensis* sera samples were positive for antibodies against EMCV, suggesting that it may be a possible reservoir host for the virus (Grobler *et al.*, 1995). However, since this report, it is now known that two cryptic species of *Mastomys* occur in KNP (Venturi *et al.*, 2004). It is therefore unclear whether EMCV prevalence is similar in these cryptic con-generics or whether just one species is the likely reservoir of infection. Also, as very few species of genera other than *Mastomys* were evaluated, it is possible that the true reservoir was overlooked due to under-sampling.

Cardiovirus detection and characterisation is primarily achieved by targeting two discrete regions of the viral genome, namely the 5'UTR and 3D gene, respectively. The latter target because of its demonstrated value for determining the origin and course of outbreaks (Koenen *et al.*, 1999) is well-represented in public databases such as Genbank. Whilst substantial data are available for northern hemisphere countries, there is a paucity of reference data for Africa, and southern Africa in particular. Therefore if a comprehensive study is performed in the KNP, where the virus occurs naturally, valuable baseline data will become available which will provide crucial information for determining the likely geographical source of regional outbreaks and provide insight into the probable murid rodent reservoir host.

Members of the order Rodentia are considered the most likely potential reservoir hosts of the virus and the increase in rodent numbers is believed to have contributed to the spread of the virus in the park in the mid-1990's (Grobler *et al.*, 1995). As inter-species transmission of EMCV likely occurs *via* the faecal-oral route, screening of gastrointestinal tract (GIT) samples is important, as is assessment of

kidney samples as transmission may result from ingestion of food contaminated with either faecal material or urine (Roy Bengis, pers. comm.).

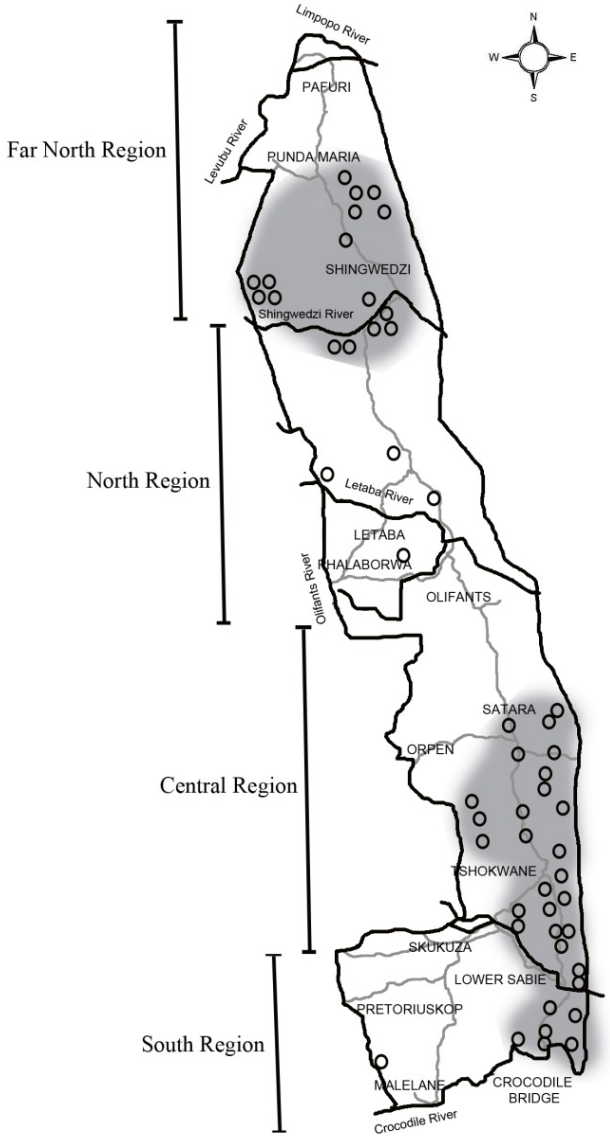

Figure 3.1 A figure showing the location of the two main elephant mortality clusters (the shaded areas) which occurred during the 1993/94 KNP outbreak (Adapted from Grobler *et al.*,1995). The circles represent the localities individual carcasses discovered during the outbreak. The rodent trapping sites were located within the southern cluster.

The rapid diagnosis of EMCV is clearly of great importance in epidemiological studies where the causative agent of an outbreak has to be rapidly and confidently identified. Research has shown that diagnosis of EMCV is most

accurate when screening from heart tissue (Vanderhallen & Koenen, 1998). Direct analysis for EMCV from gastrointestinal tract or kidneys may reveal the presence of the virus in reservoir hosts/carriers. However, a problem arises when screening these organs as they may contain substances that inhibit PCR amplification which may lead to false negative results. The aim of the current study was to overcome these difficulties by targeting a region of the 5' UTR of the EMCV genome (Chapter 1) as an amplicon for the accurate determination of the presence/absence of the virus in combination with an internal control confirming host RNA integrity. The increase in the number of cardiovirus sequences and serotypes since the initial report of a diagnostic PCR in 1992 (Kyu *et al.,* 1992), substantiated a redesign of the viral primers, and these have been reported previously (Chapter 2, van Sandwyk *et al*. 2012). As with the original published PCR format, we opted to include an internal control that ensures RNA integrity. For this purpose, primers targeting the alpha-tubulin II transcribed region of the rodent host, that have a similar annealing temperature to the virus primers were included in our assay. Given the high levels of murid rodent diversity in South Africa (Skinner & Chimimba, 2005), internal primer control specificity needed to be confirmed prior to large-scale application of the method. The aim of the current study was to provide an evaluation of the prevalence of EMCV field strains across a broad range of murid rodent hosts and to establish a rapid PCR-based diagnostic tool.

3.2. Materials & Methods

3.2.1. Sample collection and species identification

Animals were collected from the Kruger National Park , South Africa, from October 2008 to April 2010, along a longitudinal gradient between Tshokwane and Lower Sabie (co-ordinates S24-48.025', E31-52.456' to S24-58.043', E31-55.596'; Figure 3.1). Rodent trapping took place in October 2008 and 2009 at a time that coincides with the dry season and in April 2009 and 2010 to coincide with the wet season, with the explicit purpose of assessing whether there was any seasonal variation in the potential reservoir host infection rate. Animals were captured using

Sherman traps (H.B. Sherman Traps Inc. Florida, U.S.A.) and handled under the guidelines of the American Society of Mammalogists (ASM; http://www.mammalogy.org/committees/index.asp; Animal Care and Use Committee, 1998) as approved by the Animal Ethics Committee of the University of Pretoria. The traps were baited with a mixture of oats and fish paste. The captured specimens were euthanised by halothane overdose under the supervision of a State Veterinarian and subsequently dissected *in situ*. Heart, kidney and gastrointestinal tract samples were collected from each animal, placed in separate tubes and immediately snap-frozen and stored in liquid nitrogen.

Rodents captured were initially identified, based on morphology, to the genus level. As the genus *Mastomys* contains cryptic species which cannot be morphologically distinguished, all rodents classified within this genus were identified using a molecular approach. Rodent species were assigned on the basis of mitochondrial cytochrome *b* gene region characterisation (see Chapter 5 for a complete overview). A tabulated summary of captured specimens, including sex, trapping location and season is provided in Table 3.1.

3.2.2. cDNA synthesis and PCR optimisation

RNA was extracted from macerated clinical specimens using the Zymo Research Mini RNA Isolation II kit, which is based on the Chomczynski extraction method (Chomczynski & Sacchi, 1987), as well as from a single viral isolate originating from the 1979 Natal swine outbreak and from two isolates obtained from elephants during the 1994 KNP outbreak (Chapter 2). The RNA from these isolates was combined with that of a negative clinical sample to serve as a positive control. cDNA was synthesised from all available RNA within the sample utilizing a random hexanucleotide approach (Meyer *et al.*, 1991; Bastos, 1998; Chapter 2), i.e. from both the rodent host and virus. RNA was reverse-transcribed using a reaction mixture of 1 × AMV Reverse Transcriptase buffer (Fermentas), 5 µM random hexamer primer (IDT), 0.2 µM dNTPs (Fermentas) and 10 U Ribolock^{TM}RNase inhibitor (Fermentas). Reactions were performed in a final volume of 10 µl, containing 5 % DMSO and 10 U of AMV Reverse Transcriptase (Fermentas), and heat inactivated at 80°C for 1 minute, following one hour of incubation at 42°C.

Table 3.1 List of murid rodent species captured within the Kruger National Park, South Africa, during the 2008, 2009 and 2010 wet and dry seasons.

Species captured	**Number**		**Location**	Season
	Male	**Female**		
Mastomys natalensis	6	0	KNP, Tshokwane	Dry 2008
Mastomys coucha	5	0	KNP, Tshokwane	Dry 2008
Lemniscomys rosalia	4	0	KNP, Tshokwane	Dry 2008
Saccostomus campestris	0	1	KNP, Tshokwane	Dry 2008
Gerbilliscus leucogaster	0	1	KNP, Tshokwane	Dry 2008
Mastomys natalensis	33	8	KNP, Tshokwane – Lower Sabie	Dry 2009
Mastomys coucha	15	16	KNP, Tshokwane – Lower Sabie	Dry 2009
Lemniscomys rosalia	8	2	KNP, Tshokwane – Lower Sabie	Dry 2009
Saccostomus campestris	5	1	KNP, Tshokwane – Lower Sabie	Dry 2009
Gerbilliscus leucogaster	3	1	KNP, Tshokwane – Lower Sabie	Dry 2009
Mastomys natalensis	14	10	KNP, Tshokwane	Wet 2009
Mastomys coucha	1	13	KNP, Tshokwane	Wet 2009
Lemniscomys rosalia	0	1	KNP, Tshokwane	Wet 2009
Gerbilliscus leucogaster	5	3	KNP, Tshokwane	Wet 2009
Saccostomus campestris	5	6	KNP, Tshokwane	Wet 2009
Mastomys natalensis	28	30	KNP, Tshokwane – Lower Sabie	Wet 2010
Mastomys coucha	3	5	KNP, Tshokwane – Lower Sabie	Wet 2010
Lemniscomys rosalia	2	1	KNP, Tshokwane – Lower Sabie	Wet 2010
Saccostomus campestris	16	7	KNP, Tshokwane – Lower Sabie	Wet 2010
Gerbilliscus leucogaster	1	6	KNP, Tshokwane – Lower Sabie	Wet 2010
Total Murid Rodents	**154**	**112**		

All available EMCV whole genome sequences were downloaded from GenBank and used to redesign the gene region targeted by the primer combinations of Kyu *et al.,* (1992), but with the aim of ensuring that the newly designed primers recognize and target the widest possible range of EMCV strains currently known to occur in mammals. The internal control was designed to be homologous in a large variety of species to ensure accurate amplification within any rodent sample screened. The 290 nt virus genome region targeted corresponded to the 5'-UTR of the EMCV genome and was amplified with primers EMC-AB1 (5'-GGC CGA AGC CGC TTG GAA TA -3') and EMC-AB2 (5'-ACG TGG CTT TTG GCC GCA GA-3'). In addition, a 318 nt target corresponding to the 3D region of EMCV was amplified

using EMC-3DfAB (5'-TCA GGT TGT GCA GCG ACC TC- 3') and EMC-3DrAB (5'-CTT ACC GGG TAA CGC GTT GT-3'). Both primer sets readily amplify strains belonging to the two distinct virus lineages identified from historical outbreaks in South Africa (Chapter 2). To act as an internal control, a 373 nt target corresponding to the translated region of the Chinese hamster alpha-tubulin II mRNA was amplified with primers α- tub89F (5'- GCA GCA ACC ATG CGT GAG T -3') and α- tub461R (5'- GGT CAG CCA GCT TGC GAA T -3') which were designed based on α-tubulin gene sequence. This exon 4 gene target corresponds to the same region amplified in the original method developed by Kyu *et al.* (1992). All primers and targets are summarised in Table 3.2.

The primers targeting the host genome were newly designed for two reasons, the first being that a two base pair difference between the viral amplicon (290 nt) and the host target of 288 bp included in the original Kyu *et al.* (1992) approach would be too small to allow visual confirmation of the distinct bands, and the host primer sequences were not provided in that report. A larger amplicon for the internal control was therefore developed to more readily distinguish between the host RNA fragment and the 290 bp EMCV target. The second reason was to ensure accurate amplification of the host target across a wide range of rodent species. As the α-tubulin gene is conserved across a number of mammalian orders it is an ideal target. Tubulin, of which α-tubulin and β-tubulin are the most common forms, is a globular protein responsible for the formation of microtubules (Dutcher, 2001). The gene region is readily transcribed in a variety of tissues in the host therefore ensuring the presence of the specific mRNA during screening from a variety of different starting materials. The internal PCR control primers were designed to ensure amplification of a size-distinct band and to have similar annealing temperatures to the primers targeting the EMCV genome (Table 3.1).

Results were readily interpreted following agarose gel electrophoresis; if both amplicons amplify EMCV genome presence is confirmed, if only the α-tubulin target amplifies then this would confirm the sample as being a true negative while if there is no amplification then the sample is excluded due to assignment of a 'false negative' status most likely attributable to degradation of the RNA or PCR inhibitory substances. To verify that a RNA transcript was being amplified by the newly developed tubulin primer set, murid host DNA extractions were subjected to DNase

(Fermentas) treatment at 37°C for 45 minutes followed by incubation at 65°C for 10 minutes, cDNA synthesis and subsequent PCR amplification.

Table 3.2 A summary of the primers evaluated in the current study and those previously used an assay similar to that of Kyu *et al.* (1992).

Primer Name	**Primer Sequence** (Sense sequence of the reverse primer)	**GC/mer**	**Tm (**°C)	**Gene target**	**#EMC virus genome position**	**Reference**
EMC-AB1*	5'-GGCCGAAGCCGCTTGGAATA-3' (F)	12/20	61	5'-UTR	213-232	van Sandwyk *et al.*, 2013
EMC-AB2*	5'ACGTGGCTTTTGGCCGCAGA -3' (R)	12/20	61	5'-UTR	478-497	van Sandwyk *et al.*, 2013
EMC-1*	5'-GTCGTGAAGGAAGCAGTTCC-3' (F)	11/20	59.4	5'-UTR	378-397	Kyu *et al.*, (1992)
EMC-2*	5'- CGGAGACGCCGGTTTTCGGTGCAC-3' (R)	16/24	69.5	5'-UTR	806-829	Kyu *et al.*, (1992)
EMC-3D FAB*	5' - TCAGGTTGTGCAGCGACCTC - 3' (F)	12/20	61	3D	7128-7147	van Sandwyk *et al.*, 2013
EMC-3D RAB*	5' - CTTACCGGGTAACGCGTTGT - 3'(R)	11/20	59	3D	7651-7670	van Sandwyk *et al.*, 2013
89F**	5'-GCAGCAACCATGCGTGAGT-3' (F)	11/19	59	Alpha-tubulin	89-107	This study
461R**	5'-GGTCAGCCAGCTTGCGAAT -3 (R)'	11/19	59	Alpha-tubulin	461-479	This study

*Positions based on Mengo virus isolate Rz-pMwt, GenBank accession number DQ294633.
** Positions based on Chinese hamster alpha-tubulinII mRNA, complete cds,GenBank accession number M12253.

In the optimisation of the 5'UTR and α-tubulin duplex PCR, each primer pair was initially amplified individually over a temperature range to determine the optimal annealing temperature, and subsequently combined in a single reaction.

Initially the EMCV primers were kept constant at 0.2 µM and the α-tubulin primers were ranged from 0.1 µM to 0.4 µM, in 0.1 µM increments. The same was then done by maintaining the α-tubulin primer concentration at a constant 0.2 µM and varying the EMCV primer concentration in 0.1 µM increments. It was found at an annealing temperature of 59°C that a final primer concentration of 0.4 µM for the α-tubulin-targeting primers and 0.2 µM for the EMCV 5'-UTR primers resulted in the best levels of co-amplification (results not shown). All clinical samples, and the positive control were subsequently screened by PCR in reactions containing 1 x DreamTaq^{TM} Buffer (Fermentas), 0.2 µM dNTPs (Inqaba Biotech), 1.25 U of DreamTaq^{TM} DNA polymerase (Fermentas) and 100-200ng of cDNA template in a final reaction volume of 50 µl.Touch-down PCRs comprising of 2, 8 and 35 cycles of annealing at 62°C, 60.5°C and 59°C each were preceded by denaturation at 96°C for

12s, and followed by elongation at 70°C when co-amplifying the virus 5'UTR genome region and host internal control target.

The 3D region was amplified as described in Chapter 2 and all resulting PCR products were visualized by 1.5 % agarose gel electrophoresis against the GeneRuler^{TM} 100 nt Plus DNA Ladder (Fermentas).

3.3. Results

3.3.1. Trap capture rates

Species composition varied between seasons with *M. coucha* making up a higher percentage of the total *Mastomys* trapped during the dry seasons. Overall *M. natalensis* was the most commonly trapped rodent species making up 48.5 % of the total murid rodents sampled. Analysis of identification of *Mastomys* species collected in October 2009 and in April 2010 showed that during spring, October 2009 *M. coucha* comprised 42.3 % while *M. coucha* made up only 12.1 % of the total captured during autumn, April 2010 (Table 3.3).

Table 3.3 Seasonal distribution and trapping rates of murid rodents from the Kruger National Park.

Species	**Oct 2008 (Dry)**	**April 2009 (Wet)**	**Oct 2009 (Dry)**	**April 2010 (Wet)**	
Animals caught/traps placed	13/300	59/100	92/700	99/700	
% Trapping success	4.3	59.0	13.1	14.1	
					Percentage of total trapped
Mastomys natalensis	6	24	41	58	48.5 % (129/266)
Mastomys coucha	5	14	31	8	21.8 % (58/266)
Lemniscomys rosalia	4	1	10	3	6.8 % (18/266)
Saccostomus campestris	1	11	6	23	15.4 % (41/266)
Gerbilliscus leucogaster	1	8	4	7	7.5 % (20/266)
Total	**17**	**58**	**92**	**99**	**266**

3.3.2. Optimised duplex PCR conditions

Following optimisation of the PCR using clinical samples combined with cDNA prepared from elephant *Cardiovirus* isolates, the duplex PCR method was applied to all clinical samples collected from all murid rodent individuals sampled over a two year period. None of the KNP murid rodent host samples (detailed in Appendix 3.1) amplified the cardiovirus genome target, but all amplified the internal rodent host gene target. All three sample types (heart, kidney, GIT) of the 266 specimens collected (Table 3.3), were screened using RT-PCR assays, targeting the 5'UTR and 3D regions in separate reactions. No evidence of EMC virus genome presence was found in any of the specimens (Table 3.3).

Figure 3.2 illustrates the range of results that can be anticipated when using the duplex host-virus PCR assay to screen clinical samples.

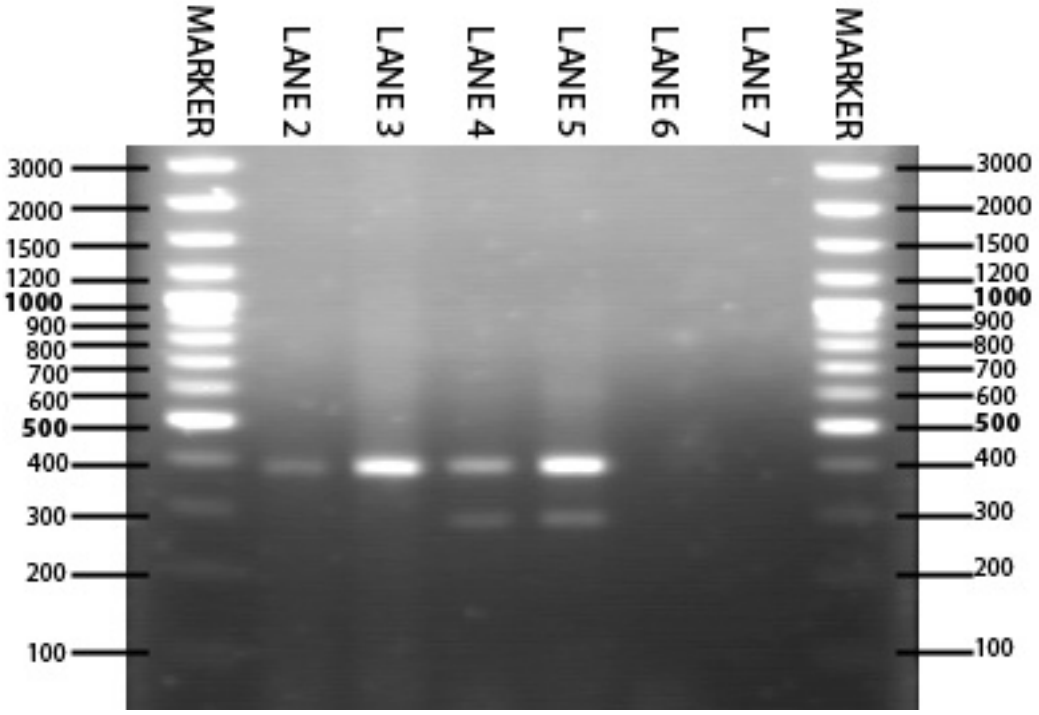

Figure 3.2 Agarose gel electrophoresis of murid clinical samples combined with EMCV RNA, prior to cDNA synthesis and PCR to confirm co-amplification of the virus genome and host α-tubulin target. Lanes 1 & 8 contain the 100 bp ladder (Fermentas), lanes 2 & 3 contain rodent samples showing positive amplification of the internal control and no amplification of the virus. Lanes 4 & 5 show amplification of internal control and EMCV target with which the rodent samples in lanes 2 & 3 were combined with the EMCV RNA, lane 6 is a false negative showing no internal control or virus amplification while lane 7 is a negative control containing no template cDNA.

3.4. Discussion

The ability of PCR to detect minute quantities of nucleic acid material has made it a popular tool within the field of diagnostics with many epidemiological and clinical studies employing it (White *et al.,* 1992; Greenfield & White, 1993; Dale & Dragon, 1994; Didomenico *et al.,* 1996; Rosenstraus *et al.,* 1998 reviewed by Elnifro *et al.,* 2000). The extraction of relatively pure nucleic acids from specimens has become routine without the need for extensive manipulation or the use of hazardous chemicals (Rosenstraus *et al.,* 1998). These samples may, however, still contain substances that inhibit the PCR reaction and provide misleading results. This inhibition can be monitored by including a second target nucleic acid amplicon to serve as an internal control. If amplification of the internal control occurs a negative result of the primary target can be seen as valid (Rosenstraus *et al.,* 1998). In viral studies where the target virus has an RNA genome the inclusion of a host-translated gene target allows for the accurate determination of RNA template sample quality. Thus the inclusion of the transcribed mammalian gene region further validates the results in that a true positive or negative can be easily identified, providing further assurance of sensitivity and specificity. Finally the use of a PCR-based test for the detection of viruses has several advantages over more traditional methods. It is far less time-consuming than viral isolation and has the added advantage of being usable over a longer time frame and not just at the onset of acute illness. Nucleic acid detection is more specific than serological techniques such as complement fixation determination and viral-neutralising antibody fixation due to the relative insensitivity of these tests arising from the high diversity of serologically distinct viral species (Kyu *et al.,* 1992). In their study Kyu *et al.* (1992) demonstrated the sensitivity of their approach in the ability to detect a single EMCV infected cell using PCR, though the technique had variable success owing to the loss of nucleic acids during extraction, additionally EMCV infection in clinically infected mice hearts could be detected up to 3 months following inoculation with the virus.

The persistence of the virus after infection allows for the possibility of detection in clinical samples long after an outbreak has occurred. The duplex PCR assay therefore has the potential to accurately identify EMCV in clinical samples and addresses concerns of false negatives through amplification of an internal control that

confirms the integrity of the RNA within the samples. The use of duplex PCR based techniques in diagnostic laboratories has some marked advantages over traditional PCR. The inclusion of more than one primer set allows for the amplification of multiple gene targets, thus providing a more cost-effective method for rapid identification of infection (Elnifro *et al.,* 2000) as an alternative to serology-based methods. Analogous with the development of traditional, uniplex PCR's, the optimization of annealing temperature, reaction constituents and amplification times is an important component which is best achieved through trial and error to optimise the amplification of the desired genomic target (Elnifro *et al.,* 2000). The optimisation of the developed duplex PCR was achieved by varying the primer concentrations of each of the primer pairs over a range of temperatures to ensure with absolute certainty that the inclusion of more than one primer pair within the PCR did not result in an increased chance of non-specific amplification or in the 'dampening' of amplification of either target. These measures therefore ensure the validity of the obtained results and justify the conclusion of EMCV genome absence in the individuals screened.

However, inadequate sampling cannot be excluded as a reason for the negative results as previous serological and PCR screening of rodents at the height of the KNP outbreak (1994) indicated a prevalence of 37.8 % (100/264) in the target genus (Grobler *et al.,* 1995). With the exception of the 59 % trapping success recorded in April 2009 wet season, trapping success ranged from 4 to 14 %. These rodent trap rates are similar to the values of between 3 % and 12 % reported by the Grobler *et al.* (1995) in the years preceding the onset of the 1993/1994 Kruger National Park outbreak. Grobler *et al.* (1995) preported a 0 % EMCV prevalence for the specimens from the low trapping rate captures. However, during the years of the outbreak, 1993 and 1994, these authors reported an increase in the trapping success rate which coincided with an overall increase in EMCV prevalence in rodent sera, determined on the basis of antibody neutralization tests (Grobler *et al.,* 1995, see Chapter 2). Grobler *et al.* (1995) demonstrated that elephants were shown to be seropositive during the years of the outbreak and the two years immediately preceding it. In 1991 a single elephant tested positive for neutralizing antibodies against EMCV. In 1992 when 15 elephants were tested, this increased to 30 % and reached 53.4 % in 1993 (N=116) before dropping to 14.5 % (N=172) in 1994. It was found that between 13 % and 53 % of all the elephants tested in 1994 were infected (Grobler *et al.,* 1995).

EMCV displays similar clinical symptoms to that of foot-and-mouth disease (FMDV) in young piglets (Acland & Littlejohns, 1986) but, the latter has far greater consequences in terms of the economic losses incurred by control measures and eradication of infected livestock. These losses are further compounded by trade restrictions imposed following a confirmed outbreak (Sellers & Daggupaty, 1990; Vanderhallen & Koenen, 1998). Although the origin of EMCV in the country cannot be deduced, the first recorded incidence of the virus in South Africa was found by the Arbovirus Unit at the National Institute for Virology (NIV). The NIV tested approximately 6000 rodents between 1965 and 1984 and demonstrated the first isolations of the virus in three *Mastomys* complex individuals trapped on the grounds of the Rietfontein Hospital, Edenvale, in the Gauteng Province, South Africa. Subsequent to this, further isolations were made from other rodents, primarily species of *Mastomys* in the Aliwal North, Eastern Cape Province, and Prieska, Northern Cape Province, of South Africa (Grobler *et al.,* 1995). An interesting anomaly is observed when the species composition of rodents trapped in the current study is compared to that of Grobler *et al.,* 1995, *viz.* the notable lack of any *M. coucha* in their study. This is likely because these individuals were incorrectly identified as *M. natalensis* and lumped together as such in their dataset. The South African reference strain of EMCV (AN7402) was originally isolated from a *Mastomys* rodent, which was suspected to be *M. coucha* (Gordon & Watson, 1986; Grobler *et al.,* 1995), once again calling into question the conclusions of studies in which species of *Mastomys* have been identified on the basis of morphology and distribution records alone (discussed in greater detail in Chapter 4).

The exact mode of transmission of the virus by the host is currently unknown, but it has been suggested that if rodents are the reservoir host, excretion of the virus in faecal material may be limited to a period of 24 hours following infection, although environmental persistence may be longer (Gainer *et al.,* 1968; Tesh& Wallace, 1978; Grobler *et al.,* 1995). The outbreak of EMCV in KNP coincided with a regional rodent irruption and a surge in the prevalence of antibodies for EMCV in rodents tested, as well as the high trapping success of rodents in close proximity to artificial drinking troughs (Grobler *et al.,* 1995). Hubbard *et al.* (1992) presented data demonstrating that the control of feral rats had a positive effect on limiting the spread of the disease. The correlation between rodents and the spread of EMCV is further

strengthened by findings presented by Reddacliff *et al.,* (1997) in which a chimpanzee was shown to be infected with *Capillaria hepatica*, a parasitic nematode causing hepatic capillariasis in rodents, in addition to EMCV, presumably contracted from a rodent source. In this particular case, infection was most likely as a result of ingestion of rodent carcasses. Reddacliff *et al.,* (1997) went on to state that infection of the nematode has been found in other primates and macropod species further supporting the direct or indirect role of rodents in the spread of EMCV, either through actual ingestion or through ingestion of contaminated foods. Numerous strains of EMCV have been isolated and characterised globally from suspected murid rodent hosts found in close proximity to outbreaks. These studies have shown that the interaction between native wildlife, rodents and domestic livestock plays a crucial role in the epidemiology of EMCV (Billinis, 2009). Reddacliff *et al.,* (1997) suggest that animals suspected of succumbing to infection should be routinely autopsied and tissue samples collected and stored, as the virus can be isolated from frozen tissue samples with relative ease. The preferred tissues for isolation of EMCV are heart, brain and spleen. EMCV infection has been shown to persist for the longest period in the heart tissue (Kyu *et al.,* 1992) and was only isolated from the heart muscle of infected elephants during the KNP outbreak (Grobler *et al.,* 1995). Our recommendation is that the active monitoring of mortalities should continue and in the case where the cause of death is suspected to be EMCV, that collection of rodents should commence immediately at the localities at which carcasses were found and should include GIT and kidney samples, as was done in this study, in addition to the afore-mentioned organs of choice for virus isolation. This approach would be the most effective means for obtaining an accurate indication of which rodent species are potentially contributing to the maintenance and spread of the virus, and for determining the importance of environmental contamination.

The apparent lack of EMCV infection in rodents likely reflects under-sampling rather than disappearance of the virus from the Kruger National Park. As no repeat outbreak has occurred in any large mammals it is likely that the virus has not been present at a sufficiently high level to cause it to spill over to other mammalian species. The park is one of the major tourist attractions within South Africa and is actively monitored by the State Veterinarian Office (Skukuza), with routine autopsies being performed on animals dying as a result of factors other than predation. It is therefore

unlikely that cases have gone undetected. However, the negative results over a two-year period suggests that inter-epidemic levels are likely very low and that intensified sampling over a longer duration is needed to identify the reservoir host and fluctuations in virus prevalence.

3.5. Acknowledgements

We would like to thank Dr Robert Swanepoel, National Institute for Communicable Diseases and Dr Truuske Gerdes, Onderstepoort Veterinary Institute for providing the virus isolates that were used as positive controls in the study. Dr. Roy Bengis, Office of the State Veterinarian, Skukuza is thanked for his help with the euthanasia of animals trapped in the Kruger National Park. Elene van Sandwyk is thanked for her help with the rodent trapping. Furthermore we would like to thank Helene Brettschneider for her help with the various aspects of the project. The research was approved by the Animal Ethics Committee at the University of Pretoria EC023–08. The project was funded by the DST-NRF South African Research Chair of Behavioural Ecology and Physiology held by NCB and by a NRF focus area awarded to ADSB.

CHAPTER 4

Molecular phylogeny of cryptic *Mastomys* congenerics from southern Africa

J. H. d. T. van Sandwyk, N. C. Bennett, C. T. Chimimba, & A. D. S. Bastos

Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa

Chapter Summary

The Rodentia constitute the most species-rich mammalian order representing approximately half of all mammal species. Within this order, the super-family Muroidea is the most diverse and speciose and has the broadest geographic distribution of any mammalian family, occurring on all continents except Antarctica. The genus *Mastomys* (Smith 1834) is one of the most prolific murid rodent species in Africa, being widely distributed across the continent in habitats ranging from grassy thickets, to areas modified for human habitation. Two morphologically indistinguishable cryptic species occur naturally in South Africa, *Mastomys natalensis* (Smith 1834) and *Mastomys coucha* (Smith 1836). The species vary in their susceptibility to a variety of pathogenic organisms making correct identification essential in epidemiological studies. To better understand their distribution and genetic diversity, samples of the species complex were collected from various provinces throughout South Africa, and in neighbouring Botswana. Species identification based on cytochrome *b* (cyt*b*) and D-loop mitochondrial gene sequencing revealed that *M. coucha* was predominantly restricted to the central plateau of South Africa, while *M. natalensis* were recovered from eastern provinces. The two species showed a great degree of sympatry over most of the eastern escarpment. Speciation between the two *Mastomys* species most likely occurred during the Pliocene as a result of climatic changes which occurred during this period. Sequencing and phylogenetic analysis of the nuclear interphotoreceptor retinoid-binding protein (IRBP) nuclear gene region did not recover the monophyletic lineages

identified with the mitochondrial sequence data, and may be indicative or contemporary or ancient hybridization.

4.1. Introduction

The Rodentia constitutes the most species-rich mammalian order, representing approximately half of all extant mammal species (Wilson & Reeder, 2005; Lecompte *et al.,* 2008). Rodents occur across the greatest geographic distribution of any mammalian order, occurring naturally on all continents, except Antarctica and the most isolated islands (Lecompte *et al.*, 2008). The family Muridae (Illiger 1811) is the most diverse family within the Rodentia comprising 730 currently recognised species from 150 genera (Steppan *et al*., 2004; Musser & Carleton, 2005; Lecompte *et al*., 2008).

Consensus on the exact number and composition of subfamilies within the Muridae has yet to be reached. Musser and Carleton (2005) divided the family into five subfamilies; namely (i) Deomyinae, comprising the genera *Acomys*, *Deomys*, *Lophuromys* and *Uranomys* (Michaux *et al.,* 2001; Steppan *et al.*, 2004), (ii) Leimacomyinae comprising a single species *Leimacomys buettneri*, (iii) Gerbillinae comprising the genera *Ammodillus, Brachiones, Desmodilliscus, Desmodillus, Gerbillus, Gerbillurus, Microdillus, Meriones, Pachyuromys, Psammomys, Rhombomys, Sekeetamys, Tatera* and *Taterillus*, (iv) Lophiomyinae consisting of a single species the Manned rat *Lophiomys imhausi* and (v) the sub-family Murinae (Michaux *et al.*, 2001; Steppan *et al.,* 2004). There is strong molecular support for three of the five subfamilies, *viz.* the Deomyinae, Gerbillinae, Leimacomyinae, Lophiomyinae and the speciose Murinae, which includes the otomyines, proposed by Musser & Carleton (2005). The species assemblage of the Leimacomyinae, characterised by Musser & Carleton (2005) has not yet been completely established (Lecompte *et al.,* 2008). However, as many as seven subfamilies have been recognised in southern Africa by Skinner & Chimimba (2005) namely the Deomyinae, Murinae, Gerbillinae, Mystromyinae, Cricetomyinae, Dendromurinae and the Petromyscinae. The most diverse subfamily, the Murinae (Illiger 1811) comprises approximately 561 species classified within 126 genera (Musser &

Carleton, 2005; Lecompte *et al.,* 2008), and includes Old world mice and rats (Steppan *et al.,* 2004).

The natural distribution of the Murinae encompasses the entire Old world with species endemic to the continents of Africa, Europe, Asia and continuing into the islands of Southeast Asia and New Guinea to Australia. Two genera within the family, *Mus* and *Rattus*, have also spread to the Americas and the remaining oceanic islands through human-mediated means. Studies based on a combination of nuclear and mitochondrial data have shown Africa, Southeast Asia, and the Australo-Papuan region to be centres of endemism for the Old world Murinae (Misonne, 1969; Watts & Baverstock, 1995; Lecompte *et al.,* 2008; Russo *et al.,* 2010). The major species radiations amongst African murids most likely occurred during the Miocene between 7 and 8 MYA. It has been suggested that the major contributing factor to this was changes in vegetation and landscape composition as a result of the climatic changes that occurred in the Miocene and Pliocene (deMenocal, 2004; Lecompte *et al.,* 2005*a*; Nicolas *et al.,* 2006; Lecompte *et al.,* 2008; Russo *et al.,* 2010).

Morphological classifications of the Muridae are plagued by problems of co-evolution with many common morphological characteristics evolving independently in geographically isolated landmasses. Molecular-based studies are not constrained by this dilemma and many studies have made use of slowly evolving nuclear genes to resolve species assemblages within the family (Michaux *et al.,* 2001; Jansa & Weksler, 2004; Lecompte *et al.,* 2008).

The genus *Mastomys* (Smith 1834), (Family Muridae, Subfamily Murinae) is one of the most ubiquitous murid rodent species occurring in Africa (Leirs *et al.,* 1996) and one of the most confusing, with heterospecifics being for the most part morphologically indistinguishable from one another. The genus currently contains eight species, namely Awash Mastomys (*Mastomys awashensis*), Southern Multimammate Mouse (*Mastomys coucha*), Guinea Multimammate Mouse (*Mastomys erythroleucus*), Hubert's Multimammate Mouse (*Mastomys huberti*), Verheyen's Multimammate Mouse (*Mastomys kollmannspergeri*), Natal Multimammate Mouse (*Mastomys natalensis*), Dwarf Multimammate Mouse (*Mastomys pernanus*) and Shortridge's Multimammate Mouse (*Mastomys shortridgei*) (Musser & Carleton, 2005). Two species of multimammate mouse, *M. coucha* and *M. natalensis* occur throughout South Africa and were originally classified as a single

species *Mastomys natalensis (sensu lato)* (Smith 1834). In both *M. natalensis* and *M. coucha* sexual dimorphism has been shown to be statistically insignificant (Bronner *et al.*, 2007), see Chapter 5 for an overview. The species falling within this genus have a wide distribution occurring all over the continent, in a range of habitats varying from grassy thickets, croplands to areas modified for human habitation (Mohr *et al.*, 2003). All species have been shown to be agricultural pests (Mohr *et al.*, 2003) and to be potential reservoir hosts for a variety of pathogenic organisms. Studies conducted on the genus have shown *M. natalensis* from Guinea in West Africa to be carriers of Lassa Fever (Lecompte *et al.*, 2006). Isaäcson *et al.*, (1981) elucidated the susceptibility of the different species making up the *M. natalensis* species complex to the bacterium *Yersinia pestis*: whereas *M. coucha*, identified on the basis of morphology, was highly susceptible to inoculation with the bacteria, *M. natalensis* displayed no detrimental symptoms.

The varying agricultural and medical importance of the different species of *Mastomys* provides the rationale for attempting to obtain a more accurate distributional range map for both species, through genetic identification and mapping of species, which will serve as a valuable reference when attempting to determine which of the two species is involved in outbreaks of disease.

4.2. Materials & Methods

4.2.1. Sample collection

Out of a total of 187 samples belonging to the *Mastomys* species complex collected from the Kruger National Park, South Africa (along a longitudinal gradient between Tshokwane and Lower Sabie, co-ordinates: S24-48.025', E31-52.456' to S24-58.043', E31-55.596'), 48 individuals were selected for genetic characterisation. From these, 16 unique haplotypes were identified (6 *M. coucha* and 10 *M. natalensis*) and used in further analyses. In addition 42 individuals sampled from four South African provinces, and from Moremi in Botswana, were selected for characterisation which had been sampled during previous studies (Table 4.1 & Figure 4.1).

Table 4.1 Summary of sampling localities and haplotypes identified on the basis of the mtDNA genome region sequenced of 58 *Mastomys*-complex individuals in this study

Country	**Locality**	**GPS**	**N (per species)**	**Haplotypes**
South Africa	Tshokwane, Kruger National Park, Mpumalanga	S24.800333° E031.874167° to S24-58.043', E31-55.596'	*M. coucha* (6) *M. natalensis* (10)	MCH 4, 6-10 MNH1-10
South Africa	Verena, Mpumalanga	S25.117333° E028.928500°	*M. coucha* (2)	MCH 1 & 3
South Africa	Thembisa, Gauteng	S26.016667° E028.233333°	*M. coucha* (1)	MCH 2
South Africa	Roodeplaat Nature Reserve, Gauteng	S25.637440° E028.356776°	*M. coucha* (1)	MCH 13
South Africa	Pretoria, Gauteng	S25.751202° E028.259733°	*M. coucha* (3)	MCH 11-12, 25-26
South Africa	Hammanskraal, Gauteng	S25.244329° E028.160191°	*M. coucha* (8)	MCH 20-24
South Africa	Tshilimbali, Limpopo	S22.766667° E030.200000°	*M. coucha* (1) *M. natalensis* (1)	MCH 5 MNH11
South Africa	Ga-Phaahla, Limpopo	S24.695389° E029.733911°	*M. coucha* (2)	MCH 27, 31
South Africa	Nkomo, Limpopo	S23.416000° E030.785500°	*M. coucha* (1)	MCH 33
South Africa	Mapate, Limpopo	S22.985866° E030.356469°	*M. natalensis* (1)	MNH20
South Africa	Bloublommetjieskloof, Limpopo	S24.311000° E029.769500°	*M. coucha* (5)	MCH 28-30, 32
South Africa	Beitbridge, Limpopo	S22.234783° E029.985037°	*M. coucha* (6) *M. natalensis* (5)	MCH 14-19 MNH13-17
South Africa	Richards Bay, Kwa-Zulu Natal	S28.748817° E032.050552°	*M. natalensis* (3)	MNH12
Botswana	Moremi, North-West District	S19.285383° E023.210700°	*M. natalensis* (2)	MNH18-19

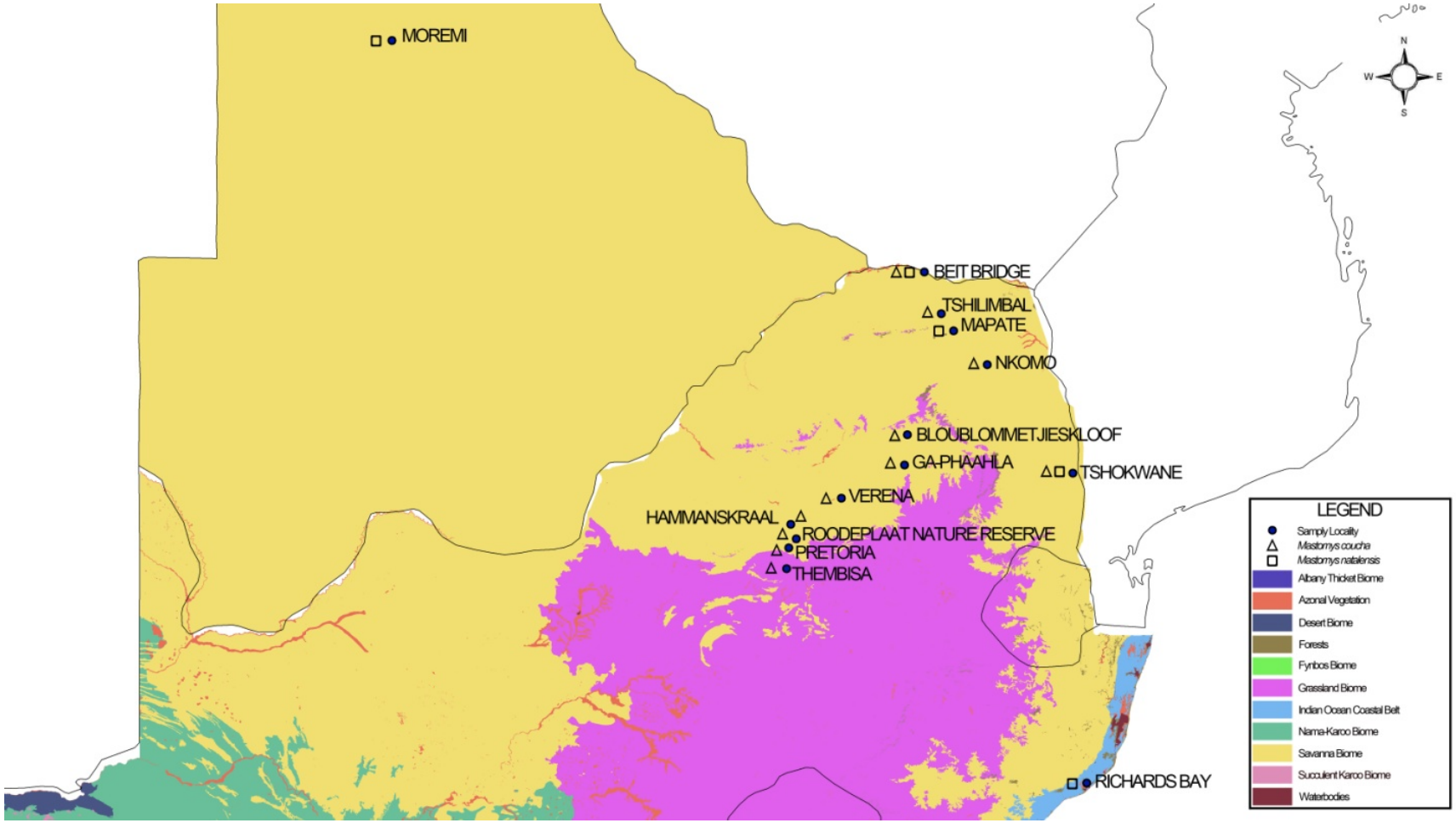

Figure 4.1 Map indicating the localities where *M. coucha* represented by (Δ) and *M. natalensis* represented by (•) specimens were collected. The insert shows a regional map of southern Africa, with the broader study area being highlighted in red.

Animals were captured using Sherman livetraps (H.B. Sherman Traps Inc. Florida, U.S.A.) and handled under the guidelines of the American Society of Mammalogists (ASM; http://www.mammalogy.org/committees/index.asp; Animal Care and Use Committee, 1998) as approved by the Animal Ethics Committee of the University of Pretoria. Refer to Chapter 2 for more detailed information on the collection of tissue samples.

4.2.2. DNA extraction

Total genomic DNA was extracted from 50 mg of heart tissue by means of the High Pure PCR Template Preparation Kit (Roche Applied Science) following the protocol prescribed by the supplier for mammalian tissue extraction. The DNA was then eluted in a final volume of 200 µl and stored at –20°C until further use.

4.2.3. Mitochondrial and nuclear gene amplification

A region of approximately 1.8 kbp of the cyt *b* gene and the non-coding D loop region of the mitochondrial genome was amplified with primers L14724-Rod (5'-TGA YAT GAA AAA YCA TCG TTG-3') and H16499 (5'-CTT GAA GTA GGA ACC AGA T-3') as using thermal cycling and reaction conditions described previously (Bastos *et al.*, 2011).

In addition, a nuclear gene region of approximately 1.2 kbp of the interphotoreceptor retinoid-binding protein (IRBP) gene of the *Mastomys* nuclear genome was amplified with primers IRBP-A (5'- ATG GCC AAG GTC CTC TTG GAT AAC TAC TGC TT-3') and IRBP-B (5'- CGC AGG TCC ATG ATG AGG TGC TCC GTG TCC TG-3') as described by Jansa & Weksler (2004). PCR reactions containing 1 x Buffer (Fermentas), 0.2 mM dNTPs (Fermentas), 0.4 µM of each primer (Inqaba Biotech), 1.25 U of Dream*Taq*^{TM} DNA polymerase (Fermentas) and 100-200ng of DNA template were performed in a reaction final volume of 50 µl. The amplification of the mitochondrial region was accomplished by touch-down PCR comprising of two, eight and 35 cycles of annealing at 49°C, 48°C and 47°C, respectively, with each being preceded by denaturation at 96°C for 12s, and followed by elongation at 70°C.

Amplification of the nuclear gene region was achieved using the same procedure, but with the annealing temperatures for the touch-down PCR being 67°C, 66°C and 65°C.

4.2.4. Nucleotide sequencing and phylogenetic analyses

Amplification of the PCR products was assessed by electrophoresis on a 1.5 % agarose gel, with product size being estimated against a Fermentas GeneRuler^{TM} 100 bp Plus DNA Ladder. Amplicons were purified directly from the tube using a Roche Applied Science High Pure PCR Product Purification Kit. This purified product was cycle sequenced in a final volume of 10 µl with each of the external PCR primers, in separate reactions. Sodium acetate precipitated products were run on an ABI3130 automated sequencer (Applied Biosystems, California), the results of which were viewed with CHROMAS (v. 1.43). Individual forward and reverse sequences were aligned using ClustalW included in the MEGA v. 5 programme (Tamura *et al.*, 2011), and similarity searches were performed with the resulting full-length sequences against the Genbank database (www.ncbi.nlm.nih.gov/blast) in order to identify all closely related sequences that have been deposited in this public database (Altschul *et al.*,1997). Four datasets (i-iv), summarised in Table 4.2 were compiled for analysis. Datasets (i) and (iv) were used to infer mitochondrial and nuclear gene trees, respectively and dataset (ii), which was restricted to the cyt *b* gene region, was used to date lineage divergence. An initial p-distance neighbour-joining tree was inferred with the *Mastomys*-only dataset (iii) in order to identify unique haplotypes and this dataset was subsequently used to compile *coucha*-only and *natalensis*-only datasets for haplotype network analyses. The best-fit model of sequence evolution was selected for datasets (i), (ii) and (iv) under the Akaike Information Criterion (AIC) in jModeltest 3.8 (Posada, 2008). The General Time Reversible (GRT) model (Gu & Zhang, 1997) incorporating invariant sites (I) and a gamma distribution shape parameter (G) was selected for the mitochondrial dataset (i), the GTR+G model was selected for cyt *b* dataset (ii) and for the IRBP dataset (iv), analysis was thus restricted to maximum likelihood (ML) incorporating the T92+G model. Phylogenic analysis of the sequences for the IRBP nuclear gene region showed no clear resolution amongst members of the *Mastomys* species complex, for the purposes of the study the analysis was thus restricted to maximum likelihood as this analysis.

Minimum evolution and maximum likelihood phylogenies were inferred in MEGA v. 5 (Tamura *et al.,* 2011) and PhyML v. 3.0 (Guindon & Gascuel, 2003), respectively. Nodal support values were obtained from 10 000 and 5000 bootstrap pseudo replications, respectively. Mr Bayes v. 3.1 (Huelsenbeck & Ronquist, 2001) was used for Bayesian inference with the model selected under the AIC for each dataset guiding the selection of priors. Four chains were run for 10×10^6 generations, using random starting trees and the default heating and swap parameters. Resulting trees and parameters were recorded and split frequencies compared every 1000^{th} generation to ensure convergence. The first 25 % of the run was discarded as 'burn-in' to ensure that trees were sampled from the region of stationarity. DnaSP v 5 (Librado & Rozas, 2009) was used to determine the haplotype (Nei & Tajima, 1981) and nucleotide diversity (Nei, 1987). A minimum-spanning network summarising haplotype connections and the minimum number of mutational steps between the mitochondrial haplotypes was inferred in TCS (Clement *et al.,* 2000) using dataset (iii).

A likelihood ratio test, with and without the molecular clock constraint was performed with nucleotide dataset (ii) under the GTR+G model of sequence evolution in MEGA 5 (Tamura *et al.,* 2011). Log likelihood scores of -3988.073 and -3849.020 with and without the molecular clock constraint, respectively, were recovered, resulting in rejection of the null hypothesis of equal evolutionary rate throughout the tree, at a 5 % significance level. A relaxed molecular clock was therefore imposed using BEAST version 1.6.1 (Drummond & Rambaut, 2007). An uncorrelated lognormal strict-clock analysis incorporating the Bayesin Skyline prior and the Tamura-Nei (1993) model (+G) of sequence evolution was implemented. Two independent runs of 1×10^8 generations were performed and the 12 MYA (LeCompte *et al.,* 2002; Russo *et al.,* 2010) divergence estimate for *R. rattus* and *M. musculus* split was used as a calibration point. Although divergence dates ranging between 10 and 41 MYA have been put forward as the divergence dates of *Rattus* and *Mus* (Kumar & Hedges, 1998; Smith & Patton, 1999; Adkins *et al.,* 2001; Benton & Donoghue, 2007) the date of 12 MYA was selected as it falls within the minimum (11 MYA) and maximum (12.3 MYA) age constraints which are supported by evidence in the fossil record (Benton & Donoghue, 2007). Markov Chain Monte Carlo was used to determine the posterior distributions of parameters over a

total of 1×10^8generations with samples drawn every 1000^{th} generation. A 20% burnin was applied. The two independent runs were combined using Log-Combiner version 1.6.1 (Drummond & Rambaut, 2007) and phylogenies inferred using TreeAnnotator1.6.1 (Drummond & Rambaut, 2007). Tracer version 1.5. (Rambaut & Drummond, 2007), was used to evaluate adequate mixing and convergence (Drummond *et al.,* 2005). FigTree, version 1.3.1 (Drummond & Rambaut, 2007) was used to visualize the inferred phylogeny, with interpretation of the divergence dates estimated in BEAST being guided by the timescale chart published by International Commissionon Stratigraphy (ICS) (http://www.stratigraphy.org).

4.3. Results

4.3.1 Nucleotide sequence analysis

The cyt*b* gene dataset comprised a partial region of 1131 bp in which the majority of the 173 mutations were located at the 3rd base position (71.1 % at the 3rd base position, 19.1 % at the 1st base position, 9.8 % at the 2nd base position). For the cyt*b*/ D-loop gene region the General Time Reversible model (Tavaré, 1986) incorporating invariant sites and a gamma distribution (GTR+I+G) was selected under the AIC criterion (AICc), (cyt *b*/D-loop: AICc = 12813.2578, Akaike weight = 0.3898, Gamma distribution shape parameter = 0.6497), while the GTR+G model was selected for the cyt *b* dataset used for dating (cyt *b*: AICc = 8298.2984, Akaike weight = 0.5121, Gamma distribution shape parameter = 0.3210). For the full length IRBP dataset the Tamura 3-parameter incorpotrating a gamma distribution (T92+G) was selected utilizing the AIC criterion (IRBP: AICc = 6910.212, *ln*L = -3330.893, Gamma distribution shape parameter = 0.39).

As suspected the phylogeny generated (Figure 4.2) using the full length cyt *b* and D-loop region data (dataset i), and rooted with out-group taxa belonging to the genera *Rattus* and *Mus* clearly indicates a clear split between *M. coucha* and *M. natalensis* supported by bootstrap values of 100% for most analysis. However the phylogeny (Figure 4.3) inferred using the IRBP region data (dataset iv), once again rooted with *Rattus* and *Mus,* showed no clear resolution between the two species.

Table 4.2 Summary of the sequence statistics and features of the four datasets (i – iv) compiled for analysis. (i) represents the cytochrome *b* / D-loop dataset, (ii) the partial cds of cyt*b* dataset used for dating, (iii) the partial cds of cyt *b* of *M. coucha* and *M. natalensis* haplotypes used to infer the haplotype networks, and (iv) the IRBP dataset.

Dataset number	**Nucleotide length**	**Gene Target (Genome position*)**	**Number of Taxa**	**Va sites (% of total)**	**Pi sites (% of total)**	**% Va sites that are Pi**	**R (Ti/Tv)**	**% AT**	**Percentage T/C/A/G**
(i)	1622	Cyt *b*/ D-loop (14169-15783)	63	600 (37.0%)	471 (29.0%)	78.5%	1.8	50.6	38.9/26.7/ 32.7/11.7
(ii)	1131	Cyt *b* Dating (14157-15287)	63	331 (29.3%)	187 (16.5%)	56.5%	3.0	60.5	27.8/ 27.1/ 32.7/12.4
(iii)	1131	Cyt *b* *Mastomys*-only (14169-15299)	53	173 (15.3%)	123 (10.9%)	71.1%	4.1	60.6	27.8/ 27.1/ 32.8/12.4
(iv)	1241	IRBP (8785-10114)	64	265 (21.3%)	68 (5.5%)	25.7%	1.4	41.2	20.2/ 29.1/ 21.0/29.7

*Positions based on *Mus musculus musculus*, GenBank accession number DQ874614.
Va, variable; Pi, parsimony informative; R, transition/transversion (Ti/Tv) ratio.

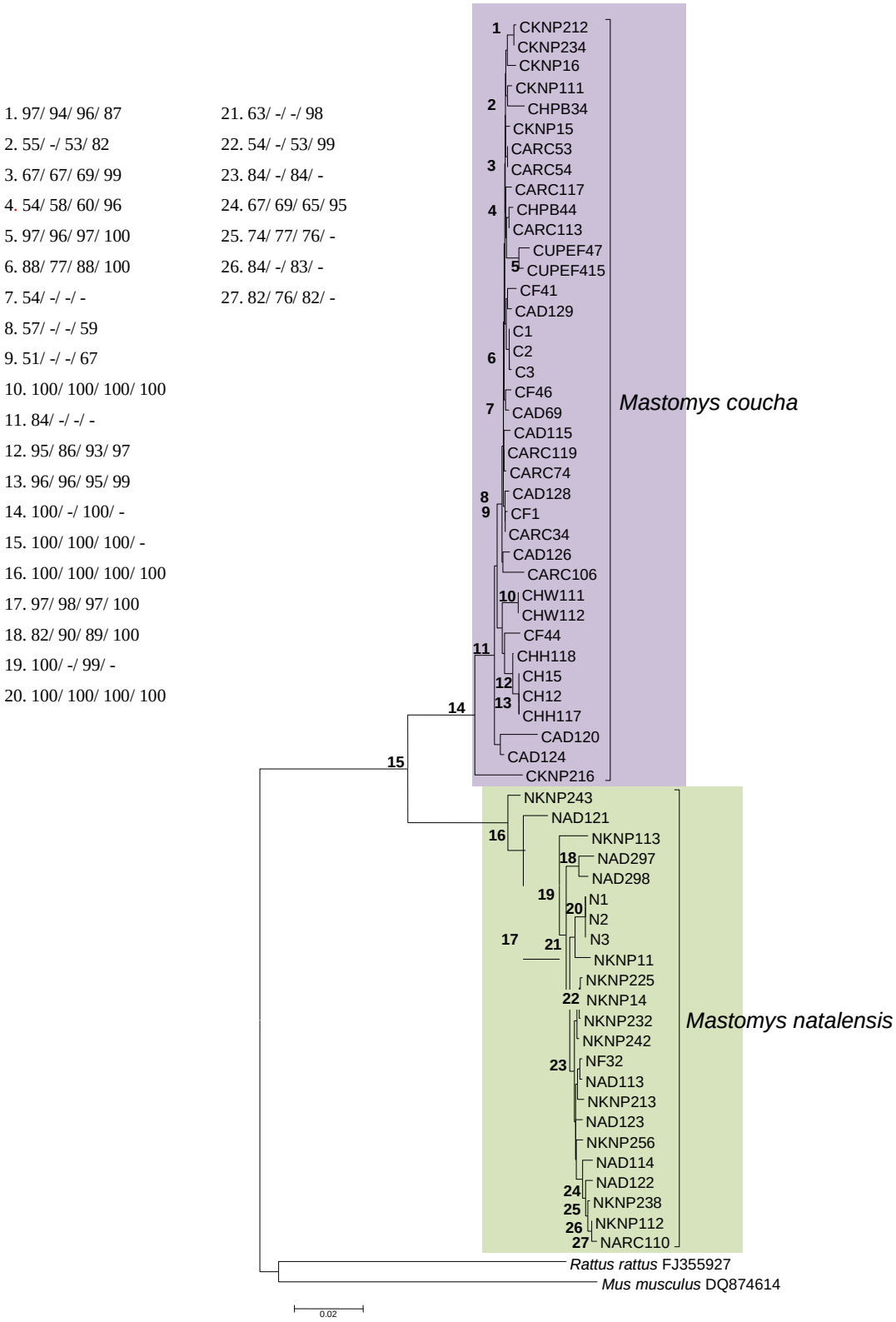

Figure 4.2 Neighbour Joining (NJ) tree indicating the *M. coucha* and *M. natalensis* relationships based on the full length cyt *b* and D-loop region data (dataset i), rooted with out-group taxa belonging to the genera *Rattus* and *Mus*. Bootstrap values > 50 % from Minimum Evolution (ME), maximum likelihood (ML) and posterior probabilities expressed as percentages from Bayesian inference (BI) analyses are summarized for each of the numbered nodes (1-27) as follows: NJ/ ML / ME/ BI

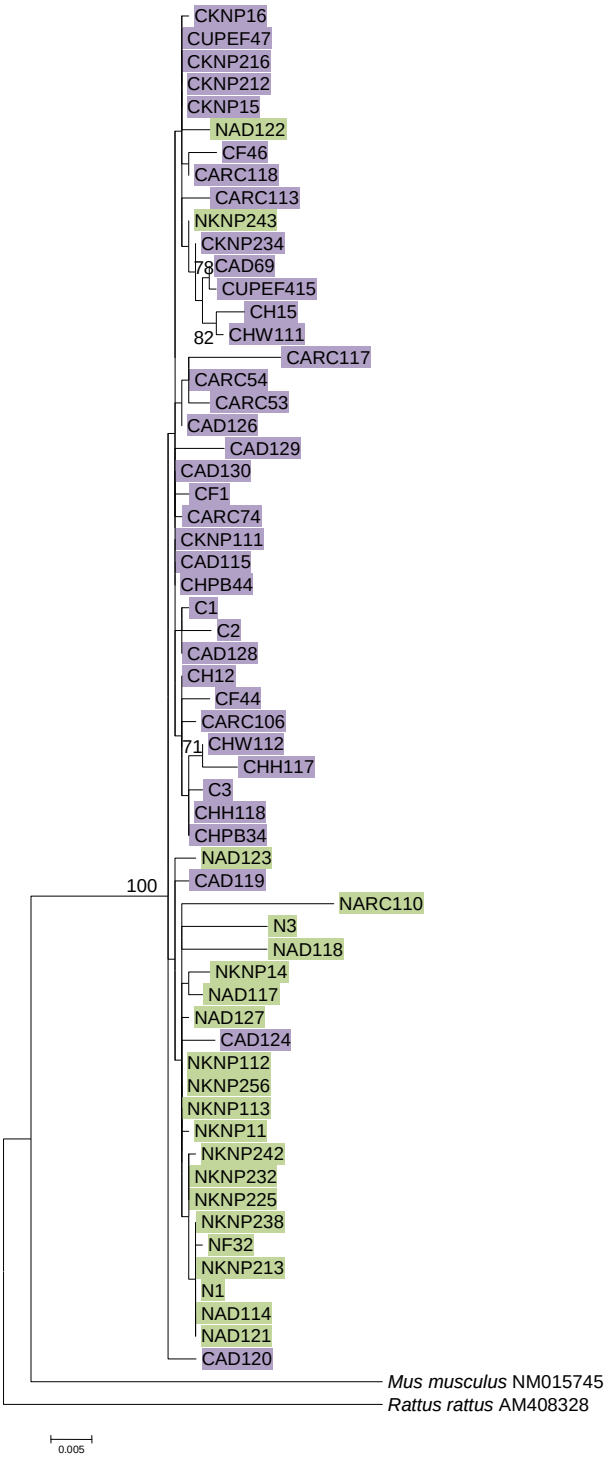

Figure 4.3 Maximum likelihood (ML) tree incorporating the T92+G evolutionary model indicating the *M. coucha* and *M. natalensis* relationships based on the full length IRBP region data (dataset iv), rooted with out-group taxa belonging to the genera *Rattus* (AM408328) and *Mus* (NM015745). Only those nodal support values > 50 % from maximum likelihood (ML) are shown.

Using DnaSP v 5 the combined *M. coucha/M. natalensis* cyt*b*/D-loop dataset (i) contained 127 polymorphic sites and 131 mutations. For the individual species datasets, *M. natalensis* had 86 polymorphic sites and 92 mutations, while and *M. coucha* had 127 polymorphic sites and 131 mutations. Haplotype diversity (Hd) and nucleotide diversity (π) was determined in DnaSP for the entire mitochondrial region characterised (dataset i). The *M. coucha* dataset had 33 unique haplotypes from the 38 sequences generated, Hd = 0.991 and π = 0.00826 and for the *M. natalensis* dataset 20 unique haplotypes were recovered from 23 sequences, Hd = 0.984 and π = 0.01092.

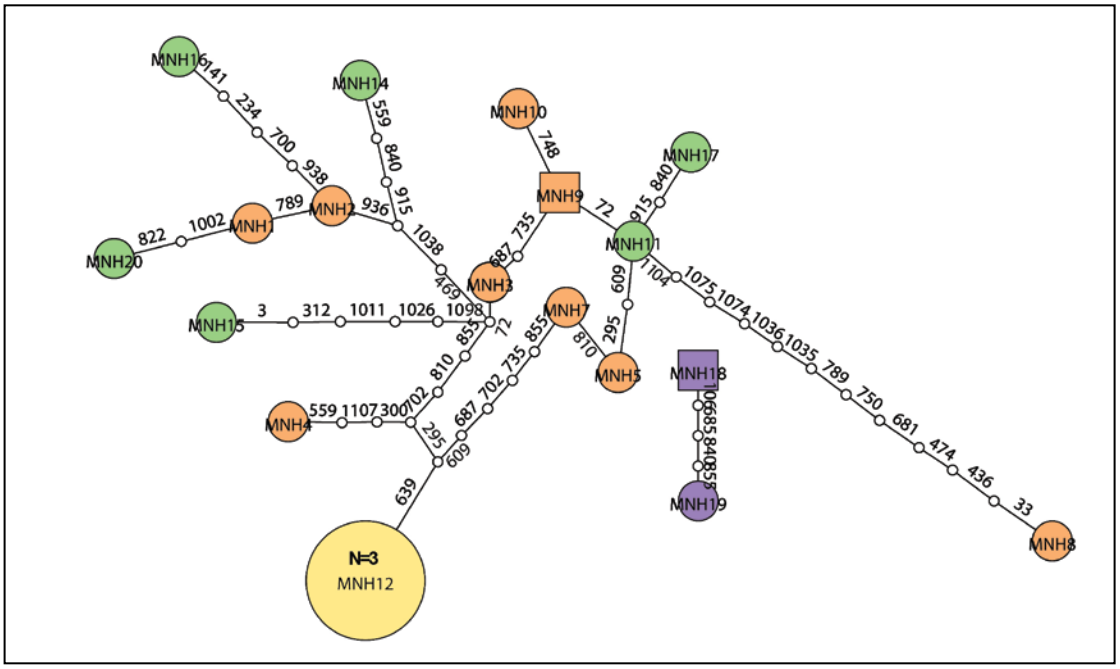

Figure 4.4 Minimum-spanning and TCS networks indicating the least number of mutational steps between composite mtDNA cyt *b* within southern African *Mastomys natalensis*. The respective sizes of the circles and squares representing allele frequencies (squares represent potential ancestral haplotypes as identified in TCS); connecting lines represent individual mutational steps, un-sampled or extinct alleles represented by small circles (Alleles connected in TCS 95 % confidence limit, i.e. 10 steps). Haplotype colour coding is representative of the provincial sample locality, *viz.* Orange = Mpumalanga, Green = Limpopo, Yellow = KwaZulu Natal and Botswana = Purple.

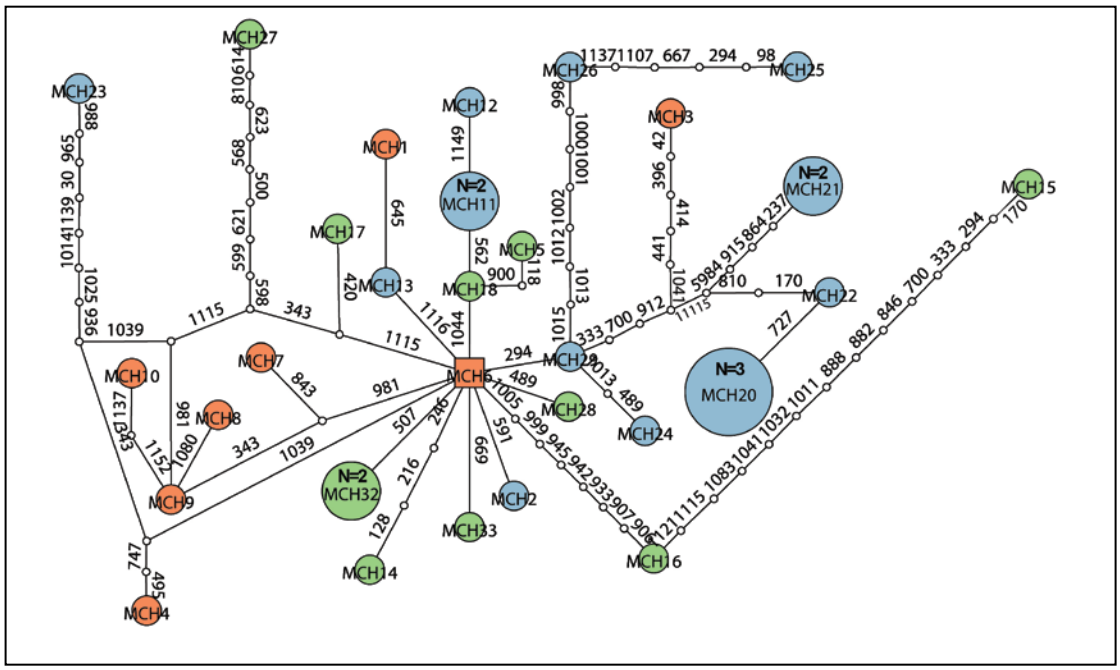

Figure 4.5 Minimum-spanning and TCS networks indicating the least number of mutational steps between composite mtDNA cyt *b* within southern African *Mastomys coucha*. The respective sizes of the circles and squares representing allele frequencies (squares represent potential ancestral haplotypes as identified in TCS); connecting lines represent individual mutational steps, un-sampled or extinct alleles represented by small circles (Alleles connected in TCS 95 % confidence limit, i.e., 10 steps). Haplotype colour coding is representative of the provincial sample locality, *viz.* Orange = Mpumalanga, Green = Limpopo and Blue = Gauteng.

The phylogeny inferred (Figure 4.6) by BEAST analysis of the mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) alleles of *Mastomys coucha* and *M. natalensis*, calibrated by a *Mus/Rattus* divergence of 12 MYA (dataset ii), cleary showed a divergence time between the two *Mastomys* species of 3.8 MYA during the Pliocene.

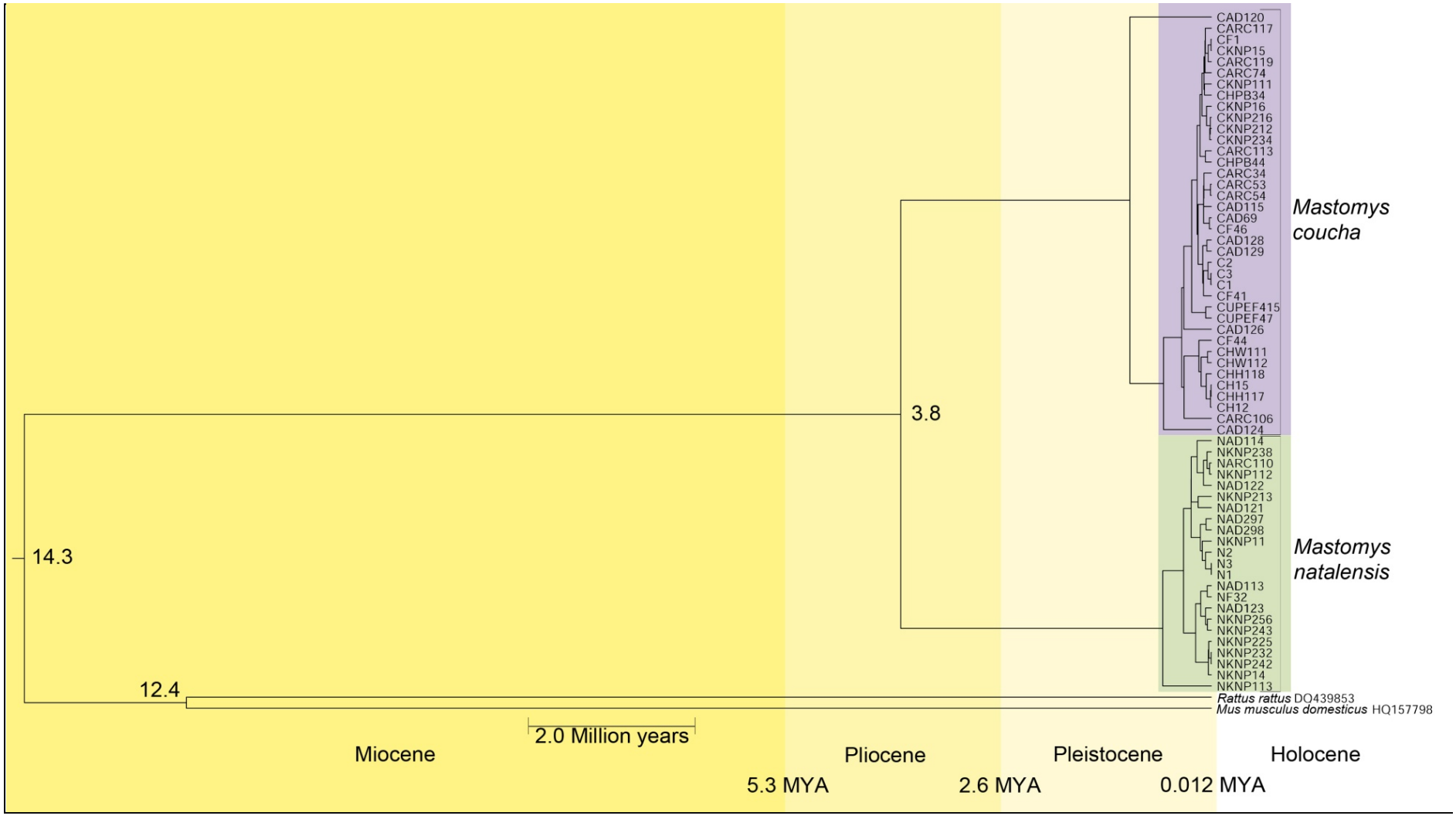

Figure 4.6. A relaxed molecular clock tree of the mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) alleles of *Mastomys coucha* and *M. natalensis* from southern Africa as obtained by BEAST analysis of 1×10^8 generations to determine divergence of the species, analysis calibrated by a *Mus / Rattus* divergence of 12 MYA, using dataset ii.

4.4. Discussion

The findings of the present study clearly show the existence of two species based on data derived from mitochondrial sequencing, however, this is not corroborated by the nuclear IRBP gene sequences. This same gene region (approximately 1200 bp) was successfully amplified and used by Jansa & Weksler (2004) to infer sub-familial relationships amongst different murid genera. This gene region has also been extensively used in higher-level relationships, interordinal relationships (DeBry & Sagel, 2001). Though it can be argued that the relatively recent divergence estimate of 3-4 MYA (LeCompte *et al.,* 2002; Colangelo *et al.,* 2013; the current study) may render this nuclear gene of little utility for resolving intra-generic relationships of *Mastomys* this is at odds with the findings of Robins *et al.,* (2008) who successfully used the IRBP gene region to determine divergence dates of between 0.4 and 3.5 MYA for species of the *Rattus* genus. The inability to recover monophyletic lineages with IRBP data, that are consistent with those recovered for *Mastomys* with the cyt *b* data is therefore unlikely to be due to incomplete lineage sorting or limited resolution power of the IRBP nuclear gene target selected for characterisation in this study. Instead, the results are suggestive of hybridisation in areas in which the two mitochondrially-distinct lineages occur.

The possibility of hybridization is supported by a study published by Duplantier *et al,* (1990) which reported the presence of intermediate karyotypes in wild-caught *Mastomys* in Senegal, suggesting ancient or contemporary hybridization events may occur between the two species. This together with the finding of the current study contrasts with results from a previous study by Gordon and Watson, (1986) in which no hybrid individuals were identified using karyotyping and penile/spermatozoon morphology for *M. natalensis* and *M. coucha* populations sampled from within the Kruger National Park.

The karyotypic diversity present in small mammal species hints that together with adaptive behavioural response stimuli and/or physiological changes (Baker & Wohlenberg, 1971; Axelrod & Raven, 1975; Vbra, 1985; Hewitt, 2004), karyotypic change may be but one of the factors driving speciation (Taylor, 2000; Russo *et al.,* 2010). LeCompte *et al.,* (2002) showed that based on karotyping and DNA/DNA

hybridisation data, *M. coucha* was the first species to diverge (Chevret *et al.,* 1994; Britton-Davidian *et al.,* 1995) from the other *Mastomys* species and that it forms a sister group to *M. natalensis, M. erythroleucus* and *M. huberti*.

It has been suggested that the use of cyt *b* data alone may be inappropriate for molecular clock dating at higher taxonomic levels (Robins *et al.,* 2008) as divergence time estimates can show marked variation when using a single mitochondrial or nuclear gene target, such as cyt *b* and IRBP, compared to a combined dataset. In particular, these authors found that divergence time estimates obtained with a combined dataset and cyt *b* alone differed 3-fold and 4-fold, respectively from the IRBP-based estimates of lineage divergence within the *Rattus* genus. The constrasting topologies recovered with the mitochondrial and nuclear gene targets precluded combining these datasets for the purpose of dating divergence of *M. coucha* and *M. natalensis*. With these constraints only the cyt *b* gene region was used to estimate a divergence date between the two species, a calibration point 12 MYA was imposed for the *Mus*/*Rattus* split as suggested by Lecompte *et al.,* (2002) and Russo *et al.,* 2010.. Using this estimate, the divergence of *M. coucha* and *M. natalensis* was determined to have taken place approximately 3.8 MYA, during the Pliocene. This divergence date is well supported by the literature which predictes the divergence of *M. coucha* from congeneric lineages to have taken place during Plio-Pleistocene (Lecompte *et al.,* 2002; Lecompte *et al.,* 2005*a*; Colangelo *et al.,* 2013). Additionally Colangelo *et al.,* (2013) clearly showed the differentiation of six well defined groups within *M. natalensis* to have occurred during the Pleistocene period. The divergence between *M. natalensis*, *M. erythroleucus*, *M. awashensis* and *M. huberti* has been more difficult to resolve using traditional methods such as karotyping. DNA/DNA hybridisation data alludes to a closer relationship between *M. natalensis* and *M. huberti* (Britton-Davidian *et al.,* 1995), while protein electrophoresis shows a closer link between *M. huberti* and *M. erythroleucus* (Duplantier *et al.,* 1990a; LeCompte *et al.,* 2002).

The Pleistocene and Pliocene are well documented as periods of major climatic, geographic and vegetative change which would have impacted on various animal species (deMenocal, 2004; Hewitt, 2004; Russo *et al.,* 2010). In South Africa this time period was characterised by major geological changes in which the interior plateaus were raised by more than 1800m above the level they occupied in the Miocene period (Baker &

Wohlenberg, 1971; Russo *et al.,* 2010). At the end of the Miocene (5-6.5 MYA) a catastrophic drop in temperature resulting in a global recession of coastlines by more than 100m occurred resulting in diversification in African Murinae, and in particular within the Praomyini tribe (Baker & Wohlenberg, 1971, Russo *et al.*, 2010). Indeed, Lecompte *et al.,* (2008) records the genera *Colomys, Heimyscus, Hylomyscus, Mastomys, Myomyscus, Nilopegamys, Praomys, Stenocephalemys,* and *Zelotomys* to have speciated approximately 7.6 (± 0.6) MYA. This rapid speciation would most likely have been driven by changes in climate and vegetation, as has occurred in other southern African mammalian fauna (deMenocal, 2004; Axelrod & Raven, 1978; Russo *et al.,* 2010). The change from the hot, wet climate of the Piacenzian (3.6–2.5 MYA), (Willis *et al.,* 1999) to the cold, drier climate of the Gelasian (2.5– 1.8 MYA), (Webb & Bartlein, 1992) has already been shown to have caused major diversifications within the Palearctic species (Dubey *et al.,* 2006; Dubey *et al.,* 2007; Mouline *et al,* 2008; Colangelo *et al.,* 2013) and amongst the African *Praomys* group (Lecompte *et al.,* 2005*a*; Mouline *et al,* 2008). Southern Africa experienced periods of aridification (2.4 - 2.9 MYA; 1.6 - 1.8 MYA and 0.8 - 1.2 MYA), (deMenocal, 2004) followed by periods dominated by a more moist climate (Axelrod & Raven, 1978). These cycles may have caused either allopatric speciation, through physical isolation of populations in the remaining pockets of favourable habitats, or alternatively sympatric or ecological speciation driven by changes in habitat (Russo *et al.,* 2010). The variation in temperature between these periods led to expansion and subsequent contraction of the eastern and southern African savanna biomes perhaps leading to the recorded speciation in several rodent genera, *Hylomyscus* (Nicolas *et al.,* 2006), *Praomys* (Nicolas *et al.,* 2008), *Tatera* (Colangelo *et al.,* 2005) and *Micaelamys* (Russo *et al.,* 2010).

Venturi *et al.,* (2004) predicted that the distributions of *M. coucha* and *M. natalensis* were segregated along the eastern escarpment of South Africa, primarily determined by rainfall and altitude. The distribution of *M. natalensis* coincides with the low altitude, high rainfall (>700 mm per annum) savannah regions of the eastern and north eastern regions of the country (Smit *et al.,* 2001; Venturi *et al.,* 2004). In contrast, *M. coucha* appears to favour the higher altitude, low rainfall (<600 mm per annum) grassland areas of the central plateau extending into the central and northern highlands of

neighbouring Zimbabwe (Gordon, 1978; Venturi *et al.*, 2004). The overall results of Venturi *et al.*, (2004) are further supported by the results of the current study, although the degree of sympatry is far greater than previously postulated. The two species had previously been predicted to occur sympatrically around Pretoria (Gauteng Province), in the vicinity of Grahamstown and within the Addo Elephant Park (Eastern Cape Province) as well as around the Satara region located in the north of the Kruger National Park (Mpumalanga and Limpopo Provinces) (Venturi *et al.*, 2004). However, based on the results of the current study it appears that they occur synpatrically along the entire eastern escarpment. Outbreaks of plague in South Africa have in the past been restricted to areas of low rainfall (123-613 mm per annum), coinciding the distribution of *M. coucha* (Bronner *et al.*, 2007). The findings of the current study support this as phylogenic analyses of the cyt *b* and D-loop regions of the specimens revealed that *M. coucha* were predominantly collected from the central plateau, while those of *M. natalensis* showed distributions favouring the eastern provinces.

The present study expands knowledge of the relatedness of the *Mastomys* species in southern Africa and provides a stable foundation for future studies to build upon to further elucidate the phylogeography of these medically and agriculturally significant rodent species. It is however, meant as only the introductory portion to a much larger study in which multiple genes of *Mastomys* sampled from across South Africa are genetically characterised in order to more fully elucidate the diversity and relatedness of the two cryptic species within this genus.

4.5. Acknowledgements

We would like to thank Dr. Roy Bengis, Office of the State Veterinarian, Skukuza for his help with the euthanasia of animals trapped in the Kruger National Park. Elene van Sandwyk and Helene Brettschneider are thanked for their help with the various aspects of the project.The research was approved by the Animal Ethics Committee of the University of Pretoria EC023–08. The project was funded by a DST-NRF South African Research Chair of Behavioural Ecology and Physiology and a grant from the University of Pretoria to NCB.

CHAPTER 5

Development of a multiplex PCR for distinguishing between cryptic *Mastomys* species in South Africa

J. H. d. T. van Sandwyk, N. C. Bennett & A. D. S. Bastos

Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa

Chapter Summary

The genus *Mastomys* (Smith 1834) is one of the most abundant murid species occurring over a wide distributional range in Africa in habitats that include natural and human modified areas. Individual species within the *Mastomys* species complex are morphologically indistinguishable, however, advances in cytogenetic and electrophoretic techniques have shown that the complex comprises of eight separate cryptic species, differing in diploid chromosome number, haemoglobin electromorphic and allozyme properties. In the two species co-occurring in southern Africa, *Mastomys natalensis* and *Mastomys coucha*, sexual dimorphism is statistically insignificant, but both possess pronounced age-related variation. *Mastomys* species are important agricultural pests and are considered to be reservoir hosts for a variety of zoonoses. The aim of the present study was to develop a molecular diagnostic tool that can quickly and accurately distinguish between the two *Mastomys* species occurring sympatrically in southern Africa. We achieved this by designing species-specific primer sets that target a 700 bp and 968 bp fragment of the cyt *b* gene, of the *M. coucha* and *M. natalensis*, mitochondrial genomes, respectively. These primers are multiplexed in a single reaction, permitting one-step PCR identification. The approach provides a rapid and cost effective means to accurately identify the respective southern African *Mastomys* species and to retrospectively reassess the identity of specimens classified primarily on the basis of species distribution maps.

5.1. Introduction

The genus *Mastomys* (Smith 1834) is one of the most widely distributed murid species in Africa (Leirs *et al.,* 1996) with a ubiquitous occurrence over the continent in habitats that range from grassy thickets and croplands to areas modified for human habitation (Mohr *et al.,* 2003). Taxonomically the genus is complex with most species being for the most part morphologically indistinguishable from one another. The two species of multimammate mouse, *Mastomys natalensis* and *Mastomys coucha*, occurring throughout South Africa were originally described as a single species *Mastomys natalensis sensu lato* (along with other members of the genus) (Smith 1834), but subsequently divided into two species (Bronner *et al.,* 2007). Though individual species within the complex are morphologically indistinguishable, advances in cytogenetic and electrophoretic techniques have revealed that the species complex comprises of eight separate cryptic species that differ in diploid chromosome number, as well as various haemoglobin electromorphic and allozyme properties (Gordon, 1978; Green, *et al.,* 1980; Hallett, 1977; Duplantier, *et al.,* 1990*a*; Bronner *et al.,* 2007). In both *M. natalensis* and *M. coucha* sexual dimorphism has been shown to be statistically insignificant and studies attempting to identify species craniometrically are riddled with pronounced age-related variation amongst individual specimens (Bronner *et al.,* 2007).

The two South African species differ in chromosomal number (2N) and autosomal fundamental number (aFN); *Mastomys coucha* has a 2N of 36 and an aFN of 52-56, wherease *Mastomys natalensis* has a 2N of 32 and an aFN of 52-54. Karotypes have in the past been one of the most reliable methods to assign species identifications (Green *et al.,* 1980; Duplantier *et al*., 1990a; Granjon *et al.,* 1997; Lavrenchenko *et al.,* 1998; Lecompte *et al.,* 2005*b*). However kartyotping requires fresh or frozen samples with intact genetic material making the identification of ethanol or dried museum samples impossible using this approach (Lecompte *et al.,* 2005*b*). Other techniques such as assessing haemoglobin variability (Dobrokhotov, 1982; Robbins *et al.,* 1983) and the search for a diagnostic allele (Duplantier *et al.,* 1990*b*) have been developed, but these have been shown to have limited success in distinguishing between the two species of

Mastomys. In South Africa, individual species have been indentified within sympatric populations of *M. coucha* and *M. natalensis* by the use of haemoglobin patterns (Green *et al.,* 1980) and diagnostic alleles (Smit *et al.,* 2001), however as with karyotyping the need for fresh or frozen samples is a serious limiting factor (Lecompte *et al.,* 2005*b*). Bronner *et al.,* (2007) found that age-variation in the two South African cryptic species was pronounced and thus needed to be corrected for, further confounding accurate identification. With these limiting factors to accurate identification in mind, Lecompte *et al.,* (2005) designed a molecular diagnostic method for distinguishing between *M. coucha*, *M. erythroleucus*, *M. natalensis* and *M. huberti*. However the approach required that each specimen be independently screened with four different primer sets under different reaction conditions in order to correctly identify it, greatly increasing the costs involved. In areas such as South Africa in which only two sympatrically occuring species need to be distinguished, the approach is needlessly time as well as cost intensive.

Venturi *et al.,* (2004) used museum samples and other collected data to summarise the distribution and predict the extant distributions of both *M. coucha* and *M. natalensis*. The variation in distribution of the two species has been suggested to be as a result of rainfall patterns and altitude. *Mastomys natalensis* is shown to be distributed in wet, low lying, coastal regions of eastern and north-eastern South Africa (Figure 5.1), whereas *M. coucha* predominates the higher, moderate rainfall areas of central and north-eastern parts of South Africa (Figure 5.2).

All species of *Mastomys* are known to be agricultural pests (Leirs *et al.,* 1996; Mohr *et al.* 2003) and to be reservoir hosts for a variety of pathogenic organisms (Gratz, 1997). As early as the 1970's the reservoir host of Lassa fever was identified as the Natal multimammate mouse, *M. natalensis* (Monath *et al.,* 1974). However, with advances in technology, two further species were put forward as potential hosts of this arenavirus. In a study using, hemoglobin electrophorisis McCormick *et al.,* (1987), proposed *M. huberti* and *M. erythroleucus* as two other potential hosts of this disease.

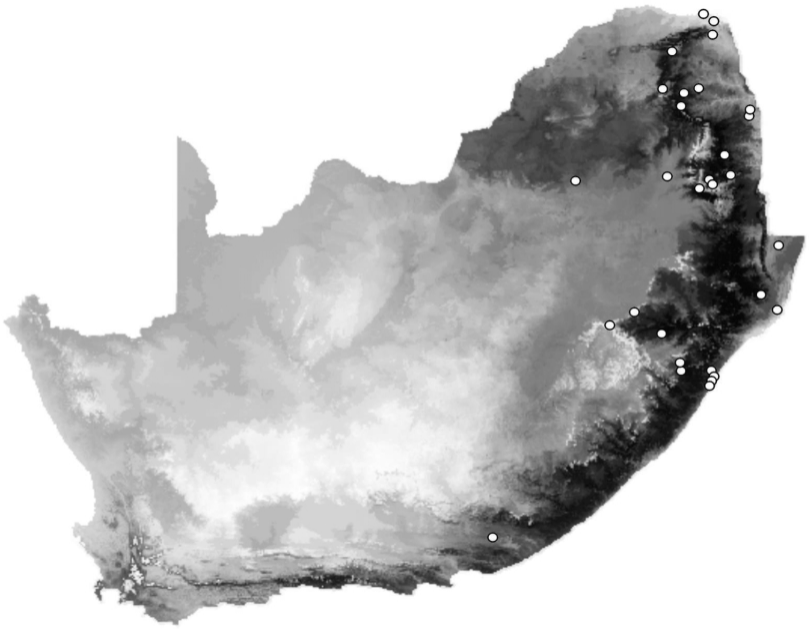

Figure 5.1 The proposed distribution of *M. natalensis* in South Africa, areas of darker shading indicate increased likelihood of occurrence. The localities of all positively identified specimens are indicated by a circle (Figure reprinted with permission courtesy of Prof. P. le F. N. Mouton, Department of Botany & Zoology, Stellenbosch University, South Africa. Co-editor: African Zoology. Originally published in Venturi *et al.,* 2004).

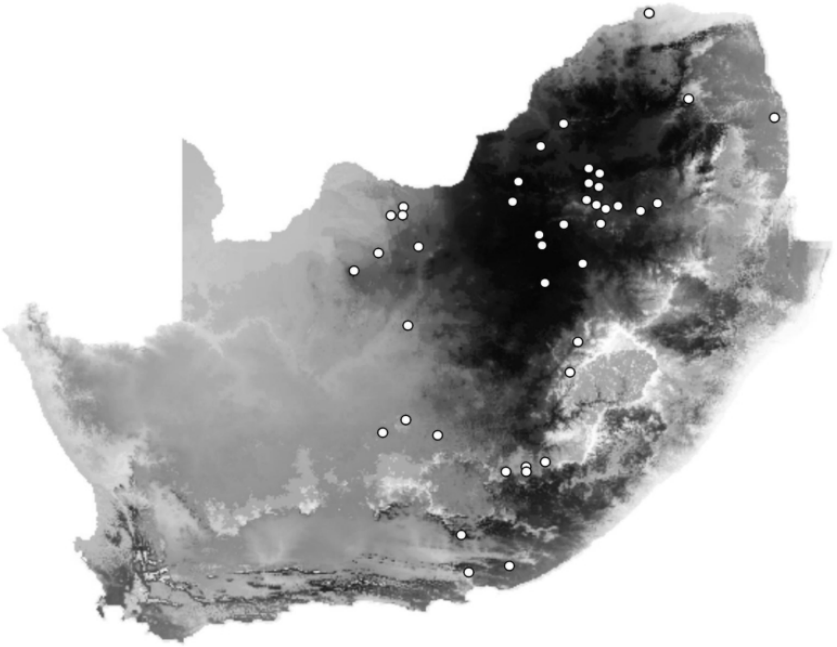

Figure 5.2 The proposed distribution of *M. coucha* in South Africa with areas of darker shading indicating higher likelihood of occurrence. Localities of positively identified specimens are indicated by a circle (Figure reprinted with permission courtesy of Prof. P. le F. N. Mouton, Department of Botany & Zoology, Stellenbosch University, South Africa. Co-editor: African Zoology. Originally published in Venturi *et al.,* 2004).

This view was later brought into contention from a study conducted by Lecompte, *et al.*, (2006), in which the authors sampled approximately 1500 murids from 13 genera in localities from Guinea, West Africa. The results of their study revealed that only molecularly typed *M. natalensis* were in fact positive for the virus (Lecompte *et al.*, 2006).

The finding that the genus *Mastomys* from southern Africa comprises two cryptic species has complicated the interpretation of the findings of many previous studies (Bronner *et al.*, 2007) and led to the development of a range of species delineating approaches (Green *et al.*, 1980; Dobrokhotov, 1982; Robbins *et al.*, 1983; Duplantier *et al.*, 1990*b*; Smit *et al.*, 2001). Many of these are labour intensive and have very stringent sample requirements. There is therefore a need to establish an approach by which the two species can be quickly and accurately distinguished and which can be readily applied to historical material such as museum specimens. The present study aimed to address this through the development of a molecular diagnostic test that can reliably and cost-effectively differentiate between the two *Mastomys* species, which occur sympatrically in southern Africa.

5.2. Materials & Methods

5.2.1. Sample collection

Samples were collected along a longitudinal gradient between S24-48.025', E31-52.456' and S24-58.043', E31-55.596' in the Kruger National Park, Mpumalanga Province, South Africa, a locality which coincides with the geographical ranges of both species. A total of 187 individuals belonging to the *Mastomys* species complex were collected (full details in Chapter 3). Additionally genetic material was incorporated into the study from *Mastomys* individuals which had been sampled from one locality in Botswana and from four provinces in the northern region of South Africa (see Chapter 4).

5.2.2. DNA extraction and amplification

DNA extracts of eight individuals, prepared as described in Chapter 4, were used as template for amplification of a 1.2 kbp region corresponding to the cyt *b* gene. The target was amplified with primers L14724-Rod (5' TGA YAT GAA AAA YCA TCG TTG-3') and H15915-Rod (5'- CAT TTC AGG TTT ACA AGA C-3') using Fermentas Dream*Taq*^{TM} DNA polymerase and previously described thermal cycling conditions (Bastos *et al.*, 2011). PCR products were purified using a High Pure PCR Product Purification Kit (Roche Applied Science). Dye-terminator cycle sequencing was performed using the ABI PRISM Big Dye™ Terminator version 3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems), with each of the PCR primers at supplier-prescribed reaction conditions. After the subsequent removal of unincorporated primers, ddNTPs and dNTPs by sodium acetate precipitation, the fluorescently labeled DNA products were run on an ABI3130 automated sequencer (Applied Biosystems, California), and viewed with CHROMAS v. 1.43.

5.2.3. Haplotype selection and primer design

Sequence data were combined with the *Mastomys*-only cytochrome *b* dataset (ii, Chapter 4) in MEGA v. 5 (Tamura *et al.*, 2011) and complemented with homologous data identified in blast searches against the GenBank database (www.ncbi.nlm.nih.gov/blast) (Altschul *et al.*, 1997). The final homologous dataset, 1140 nucleotides in length, consisted of 52 *M. natalensis* sequences, 36 *M. coucha* sequences and 10 sequences from *M. erythroleucus* and *M. huberti*. Two *Gerbilliscus leucogaster* specimens were included for out-group purposes. A p-distance tree was inferred using MEGA v. 5 (Tamura *et al.*, 2011) in order to identify unique haplotypes and compile a haplotype dataset.

Three primers were identified from the haplotype dataset, the first Muni-49F (5'-GAC TTA CCT GCC CCA TCC AAT AT-3') acts as a universal forward primer being able to bind to both species, while the species-specific *M. coucha* primer, termed 'Mcou-r749R' (5'-GGG TCT CCA TGT AAG TCT GAG-3') and species-specific *M. natalensis* primer, termed 'Mnat-1017R' (5'-TGG TTG GCC TCC GAT TCGA-3') in conjunction with the universal forward were designed to amplify a size-distinct fragment for each

species. This was achieved by ensuring specificity of the primer to the target species alone. The relevant species-specific primer ensured amplification of a 700 bp fragment for *M. coucha* and a 968 bp for *M. natalensis* under the same thermal cycling conditions, when the relevant target species nucleic acid was present. Sequences of *M. erythroleucus* and *M. huberti* were included within the alignment used for primer design to ensure the species-specific primers showed no cross reactivity with closely related sister species. Each primer pair was initially individually evaluated over a broad annealing temperature range (63°C-69°C degrees in 1°C increments), in order to identify an overlap in optimal annealing temperature. On confirming that a fragment of the correct size amplified for each of the sequence-verified target species, the primers were multiplexed at 65°C and their performance at the shared, optimal annealing temperature was evaluated over a primer concentration range of 5 to 20 pmol in 5 pmol increments in relation to the other primers to obtain the optimum concentration of each primer

5.2.4 Species-specific PCR and specimen identification

A master mix was prepared which contained 1 × buffer (Fermentas), 0.2 µM dNTPs (Inqaba Biotech), 0.6 µM of Muni-49F (Inqaba Biotech), 0.4 µM of Mcou-749R (Inqaba Biotech), 0.2 µM of Mnat-1020R (Inqaba Biotech), 1.25 U of Dream*Taq*^{TM} DNApolymerase (Fermentas). Each reaction was performed in the presence of 100-200ng of DNA template in a reaction final volume of 50µl.

The resulting PCR product was then visualized by electrophoresis on a 1.5 % agarose gel with product size being estimated against a GeneRuler^{TM} 100 bp Plus DNA Ladder (Fermentas).

5.3. Results

It was found that the addition of 0.6 µM of Muni-49F, 0.4 µM of Mcou-749R and 0.2 µM of Mnat-1017R to each 50µl PCR reaction resulted in the best amplification, however, it should be noted that primer concentrations may need to be optimised for problematic specimens. Optimal thermal cycling conditions were determined to be: an initial denaturation cycle of 96°C, for 8s and followed by PCR cycling conditions of:

96°C, 12s; 65°C, (30s); 72°C, (90s) for 35 cycles, followed by a final elongation step at 72°C for 60s and a hold at room temperature.

Sequences generated from animals characterized in this study and those from previous studies resulted in 29 and 21 unique haplotypes being identified for 51 *M. natalensis* and 36 *M. coucha*, respectively, the haplotype frequencies are summarised in Figures 5.3 and 5.4.

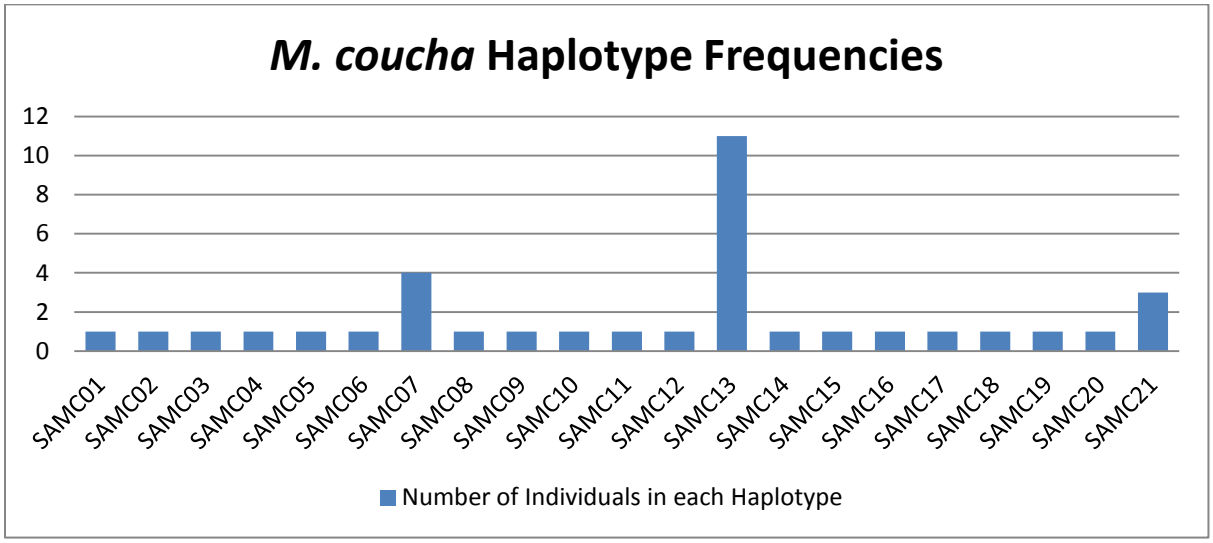

Figure 5.3 Number of *M. coucha* individuals belonging to each of the 21 assigned haplotypes identified in the original cyt *b* dataset used to design species specific primers.

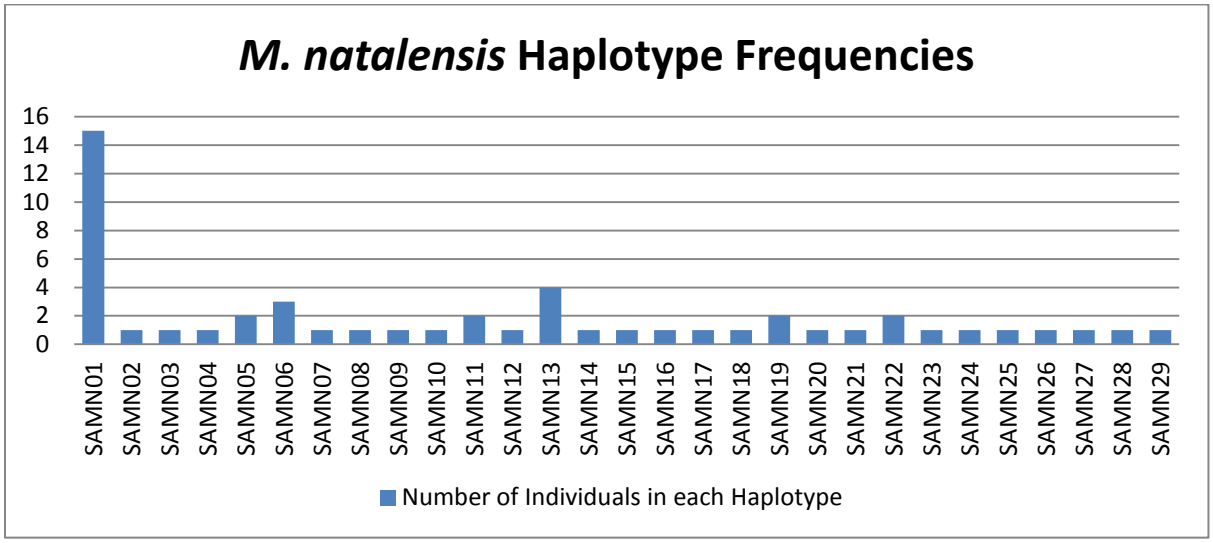

Figure 5.4 Number of *M. natalensis* individuals belonging to each of the 21 assigned haplotypes identified in the original cyt *b* dataset used to design species specific primers.

The 187 *Mastomys* specimens captured on two sampling occasions in the Kruger National Park were identified using the developed Muliplex PCR technique. It was found that they comprised of 58 *M. coucha* and 129 *M. natalensis* specimens (Table 5.1).

Table 5.1 Summary of *Mastomys* species assignment achieved using the duplex PCR approach, for 187 samples collected from the Kruger National Park (2008-2010).

Sampling season	**Number of *M. coucha* individuals (Number verified by sequencing and PCR)**	**Number of *M. natalensis* individuals (Number verified by PCR and sequencing)**
October 2008 (Dry season)	5 (5)	6 (6)
April 2009 (Wet season)	14 (14)	24 (23)
October 2009 (Dry season)	31	41
April 2010 (Wet season)	8	58
Total	**58**	**129**

Species identification was based on amplification of specific size fragments, 700 bp for *M. coucha* (lanes 3 & 5) and 968 bp for *M. natalensis* (lanes 2 & 4) and a representative sample of each species was always included as a positive control (Figure 5.5). Verification of the species assignment by PCR was verified for 48 of these samples by sequencing of cyt *b* mitochondrial gene region.

Sensitivity of a diagnostic is defined by number of true positives/(the sum of the number of true positives and the number of false negatives) and specificity is determined by calculating the number of true negatives/(the sum of the number of false positives and the number of true negatives) (Vanderhallen & Koenen, 1998). The sensitivity of the assay was validated by confirming species assessment of 48 of the 187 samples screened using the *Mastomys* diagnostic PCR through cyt *b* gene region sequencing and subsequent BLAST identification. The specificity was determined by including other

morphologically similar species that occur sympatrically, namely *Lemniscomys rosalia*, *Saccostomus campestris*, *Micaelamys namaquensis*, *Aethomys ineptus*, *Steatomys sp.* and *Gerbilliscus leucogaster*. In all non-target species cases this yielded a negative result.

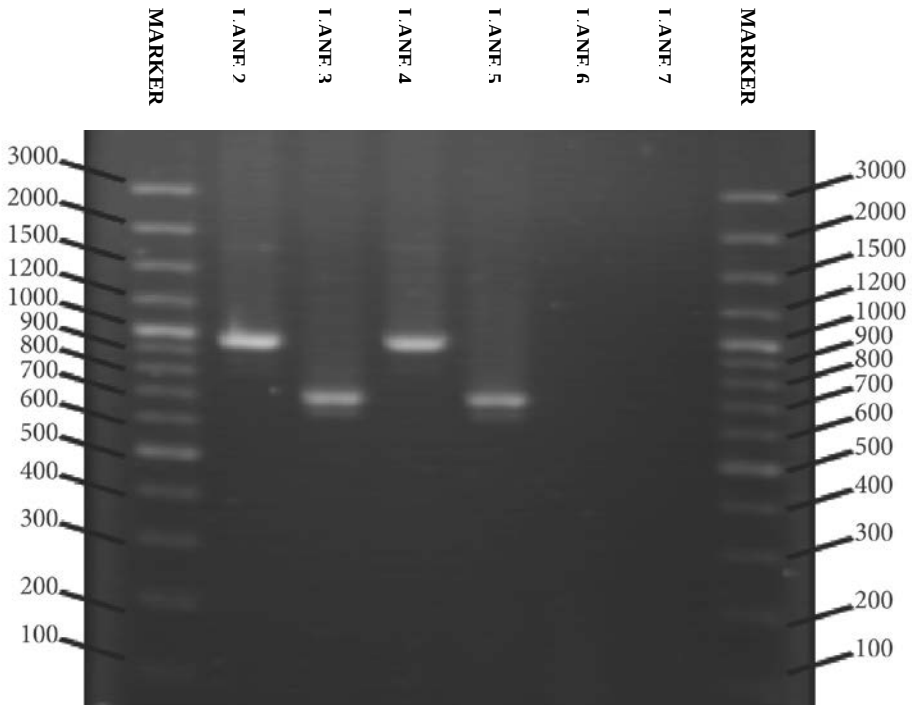

Figure 5.5 A pictorial representation of *Mastomys* species identification based on amplification of specific size fragments, 700 bp for *M. coucha* (lanes 3 & 5) and 968 bp for *M. natalensis* (lanes 2 & 4) respectively. Lanes 1 and 8 contain GeneRuler^{TM} 100 bp Plus DNA Ladder, lanes 6 and 7 are negative controls to which no template DNA was added.

5.4. Discussion

The present study provides a rapid and reliable molecular diagnostic approach for distinguishing between the two cryptic species of *Mastomys* occurring in southern Africa. Previous studies have shown that distributional separation along the eastern escarpment is influenced by altitude and rainfall. *Mastomys coucha* has a distribution encompassing the moderate rainfall (<600 mm per annum), high altitude grassland and arid savanna areas of central and northeastern South Africa, whereas *M. natalensis* occurs in high rainfall

(>700 mm per annum), low altitude savanna habitats along the eastern coast (Dippenaar *et al.*, 1993; Venturi *et al*, 2004; Bronner *et al.*, 2007). As a result of the variation in distribution, the outbreaks of various diseases which are harboured by the species may be restricted to the regions in which the relevant host *Mastomys* species predominates. Isaäcson *et al.*, (1981) showed that the two species of *Mastomys* differed in their susceptibility to infection with *Yersinia pestis*, the causative agent of plague. While *M. coucha* was highly susceptible to inoculation with the bacteria, *M. natalensis* displayed no detrimental symptoms. Outbreaks of the disease in South Africa have in the past been isolated to areas of low rainfall (123–613 mm per annum), coinciding the distribution of *M. coucha* (Bronner *et al.*, 2007). The outbreaks have, however, also coincided with the distributions of two gerbil species the Highveld gerbil, *Tatera brantsii* and the Namaqua gerbil, *Desmodillus auricularis*. *Mastomys* may play an important role in the transmission cycle of the disease, the genus has been shown to readily occur in close proximity to human habitation and this may provide a crucial link between gerbils and humans by transmitting fleas of the subfamily Xenopsyllinae, which are known to vector the causative agent (Davis, 1964; Venturi *et al.*, 2004). Outbreaks have also occurred in areas free of *M. coucha*, but which coincide with the distribution of various gerbil species, which in previous studies have shown to be positive for plague bacteria (Venturi *et al.*, 2004). The epidemiology of the disease may in fact be more complex than initially anticipated and a variety of taxa may be involved in the spread of the disease. It is also possible that resident *M. coucha* may not obtain a high enough level of infection in regions where outbreaks of the disease have been found to be absent (Venturi *et al.*, 2004). The two species of *Mastomys* have been shown to occur either sympatrically or in close proximity at various localities in South Africa, including Addo Elephant Park (Eastern Cape province), Kruger National Park (Mpumalanga & Limpopo Province) and in Pretoria (Gauteng Province) (Venturi *et al.*, 2004). It has been found from the cyt *b* sequence analysis that in the area between Tshokwane and Lower Sabie, in the Kruger National Park, there is a high degree of sympatry between the two species of *Mastomys*. Analysis of *Mastomys* species collected in October 2009 (Spring) and in April 2010 (Autumn) showed that during spring *M. coucha* comprised 42.3 % whereas *M. coucha*

made up only 12.1 % of the total captured during autumn. *Mastomys coucha* has a natural distribution coinciding with the drier regions of South Africa (Venturi *et al.*, 2004). It therefore follows that the species would be more abundant during periods more conducive to the conditions they inhabit.

As the two species of *Mastomys* that reside within the borders of South Africa were previously grouped under a single species, *M. natalensis sensu lato*, and as numerous subsequent studies have identified areas of sympatry (Bronner *et al.*, 2007, current study), studies lacking genetic verification of these two species need to be treated with caution since correct identification cannot be guaranteed. The duplex PCR developed in this study provides a rapid and practical means of conducting large-scale species identification surveys. As it is PCR-based, only a small piece of tissue, such as an ear or tail clip is required and animals need to be sacrificed in order to secure an accurate species assignment, and provides a more cost effective means of identification than sequence analysis or DNA barcoding.

5.5. Acknowledgements

We would like to thank Dr. Roy Bengis, Office of the State Veterinarian, Skukuza for his help with the euthanasia of animals trapped in the Kruger National Park. Elene van Sandwyk is thanked for her help with the rodent trapping. Furthermore we would like to thank Helene Brettschneider for her help with the various aspects of the project. The research was approved by the Animal Ethics Committee of the University of Pretoria EC023–08. The project was funded by a DST-NRF South African Research Chair of Behavioural Ecology and Physiology and a grant from the University of Pretoria to NCB.

CHAPTER 6

General Conclusions & Future Prospects

To better indentify the causative strains of EMCV responsible for these historic outbreaks (Chapter 1-3) in South Africa the phylogenetically informative P1 and flanking coding regions as well as RNA polymerase (replicase) gene (3D) region of the viral genome were amplified. One historical swine isolate and two elephant isolates were genetically characterised in order to determine the relatedness and possible origins of these viruses and to establish a regional reference dataset for future outbreaks. Genetic characterisation of two strains isolated from elephants revealed that the causative virus was most likely of African origin as it is closely related to a virus of East African origin, forming a sister clade to Mengovirus isolates from Rhesus Macaques in Uganda. The origin of the swine virus is less clear, as it was found on the basis of VP1/3 data to cluster with other EMC viruses from predominately swine populations in Europe, Asia, Central, South and North America, within one of the four genotypes detailed in this study, *viz.* genotype A. The 3D data revealed the South African pig virus to be most closely related to a 1997 domestic pig virus from Greece, but as the South African pig virus predates this European outbreak strain, no inferences on origin can be made (Chapter 2, van Sandwyk *et al.,* 2012). Although serologically all of the EMC viruses tested to date cross-react and are considered to represent a single serotype, the recovery of four monophyletic lineages that display high levels of within and between-lineage diversity is of importance from an evolutionary and molecular epidemiological perspective.

To assess the extant diversity of EMCV within South Africa, and unequivocally determine the identity of the rodent reservoir and its inter-epidemic levels of virus prevalence, tissue samples (heart, kidney and gastrointestinal tract) were collected from 266 murid rodents, over a two-year period from the Kruger National Park. The area in the national park that was targeted for cardiovirus detection in rodents corresponded to the locality in which the historical elephant mortalities occurred (Grobler *et al.,* 1995).

Although numerous samples types were screened by RT-PCR targeting two discrete genome regions, the 5'UTR and 3D genome regions, no infected rodents were identified. This is in spite of the extensive trapping period over several seasons in the locality of the 1993/4 outbreak and recovery of a high proportion of *Mastomys*, the purported reservoir host genus, on each sampling occasion.

If the data on rodent trapping versus virus detection, reported by Grobler *et al.*, (1995) for 1994, are considered, it is clear that there was an almost 9-fold increase in the rodent numbers in that year, compared to the average for the five preceding sampling years prior to the outbreak onset in 1993 (data summarised in Table 6.1).

Table 6.1 Summary of the number of rodents trapped during and preceding the 1993/ 1994 EMCV KNP outbreak which were shown to be seropositive (Grobler *et al.*, 1995).

Year	Number of animals trapped	Number of animals infected (% of total)
1984	85	0 (0 %)
1986	22	0 (0 %)
1987	55	0 (0 %)
1989	39	0 (0 %)
1990	32	0 (0 %)
1993	36	3 (8 %)
1994	416	150 (36%)

Over the three years that the present study ran, a total of 266 rodents were trapped *viz.* 17 in 2008, 150 in 2009 and 99 in 2010, of which none were shown to be positive. During the first year of the historical outbreak in the Kruger National Park, the seroprevalence rate within the rodent population was 8 % (just 3 positive individuals from the 36 sampled), whereas in 1994 it increased to 36 % for the 416 animal trapped at the peak of the rodent irruption and outbreak. The inability to detect any positive animals during the current study may have been a result of a combination of factors, including: (i) the infection rate is too low to have been detected from the current sample size, (ii) the primary reservoir host may not be *Mastomys* and (iii) the window of molecular virus detection is too short to have been detected. If we assume that molecular and serological prevalence are equal and match the seroprevalence estimate obtained in the first year of

the historical elephant outbreak, then based on the number of animals trapped in this study, we should have detected 12 positive animals in 2009 and 8 in 2010. As the infection status of individuals was verified in this study by using an internal RNA control to ensure RNA integrity, and as the virus had previously been shown to be detectable by RT-PCR for up to three months post-infection, providing a more accurate diagnosis than traditional serological methods which may result in a false negative after an extended period (Brewer *et al.,* 2001), the negative result was unexpected. This finding implies that the inter-epidemic infection rate must be extremely low since out of a total of 266 rodents trapped over these two years, none were shown to be positive. It can also be expected that the infection rate may be too low to maintain the virus within the rodent population. However, a negative result does not prove the lack of the virus within the population, this could only be conclusively said if the majority of the rodents present in the park had been sampled. To accurately assess the diversity of cardioviruses in South Africa, far more extensive sampling would be required from a variety of regions, not just those restricted to the areas where previous outbreaks have occurred. The lack of finding in the sampled rodents shown to be positive for EMCV, made it impossible to assess the diversity of extant EMC viruses in South Africa, hampering the original aim of the study to better understand the diversity and source of contemporary South African cardioviruses. However, the retrospective genetic analysis, indicated that the South African virus recovered from the pig outbreak in 1979 was most closely related to the virus recovered from an outbreak in Greece in 1997 and that the historical outbreak in South Africa was caused by a genotype A virus, implying that they may share a common ancestor. The route of the introduction into South Africa remains unresolved. At the time of the 1979 outbreak import of swine into South Africa was restricted from the USA due to an outbreak of Aujeszky's Disease, while at the same time an embargo was placed on the importation from Europe due to possible outbreaks of Classical Swine fever and Swine Vesicular Disease (Records of the South African Department of Agriculture). Insights into the possible origin and route of importation of this virus may become clearer once viruses of suid and reservoir rodent strains of African origin are genetically characterised. In contrast, the grouping of the two elephant viruses with Mengovirus strains isolated from Ugandan macaques in the 1940s is highly suggestive of the presence of an

'endemic' EMCV lineage. The EMCV outbreak coincided with a rodent outbreak, one of the most prevalent species during this outbreak was *Mastomys* which showed a 38 % seroprevalence implying that this may be the host or vector of the disease (Grobler *et al.*, 1995). However, as previously stated the lack of finding any rodents positive for EMCV, in this study, hampered efforts to investigate the role of this rodent in the epidemiology of the disease. Any future studies aiming to assess the extant diversity of EMCV during periods between outbreaks would need to make use of extensive trapping and sampling over an extended period of time.

During the 1993/4 outbreak in the Kruger National Park the seroprevalance in 1994 showed a marked increase in infected *Mastomys*, with 100/264 (38 %) of screened individuals testing positive for EMCV. Infection rates in other species of rodents were substantially lower with only 1/34 (3 %) *Aethomys chrysophilus*, 2/41 (5 %) *Saccostomus campestris* and 2/63 (3 %) *Gerbilliscus leucogaster* testing seropositive (Grobler *et al*, 1995). The 10-fold increase of positive *Mastomys* as well as the much higher trapping success rate at which these murids were captured when compared to other rodent species that cohabitate the area made them a prime reservoir host candidate.

To better understand the distribution of the two *Mastomys* species occurring sympatrically in southern Africa, we developed a molecular diagnostic tool that could quickly and accurately distinguish between them. This was achieved by amplifying species-specific sized fragments of the cytochrome *b* (cyt *b*) gene region for each respective species that could be multiplexed in a single reaction. To this end we designed three primers that could be used in combination with one another (see Chapter 5). Results revealed that contrary to available distribution maps (Venturi *et al.*, 2004) and predicted distribution (Venturi *et al.*, 2004), both species rather than just *M. natalensis* were found to occur sympatrically in the southern region of the Kruger National Park, where the samples were collected.

In order to determine whether the unexpected sympatric distribution of the two species was limited to the southern region of the Kruger National Park, we expanded the sampling of *Mastomys* from this region to other areas within South African and southern Africa. Species identification was initially determined on the basis of full length cyt *b* and D-loop mitochondrial gene region characterisation. The results concurred with the

findings of Venturi *et al.*, (2004) that the distribution of *M. coucha* was predominantly restricted to the central plateau of South Africa, while *M. natalensis* showed distributions favouring the eastern provinces. However, the two species showed a significant degree of sympatry, with both species occurring along most of the eastern escarpment and into neighbouring Zimbabwe. A nuclear gene region was subsequently amplified and sequenced in order to determine whether the mtDNA species delineation could be recovered with a nuclear marker. For this purpose, the widely used IRBP gene region, encoding the Interphotoreceptor Retinoid-Binding Protein, was amplified and sequenced (see Chapter 4). Unfortunately the results revealed no resolution between *M. coucha* and *M. natalensis*, implying that there may be hybridization occurring between the two species in the regions where they occur sympatrically, or alternatively that there has been incomplete lineage sorting. The former is supported by a study published by Duplaintier *et al*, (1990) that reported the presence of intermediate karyotypes in wild-caught *Mastomys* in Senegal, suggesting ancient or contemporary hybridization events occurring between the two species. LeCompte *et al.,* (2002) showed that, based on karyotyping and DNA/DNA hybridisation data, *M. coucha* was the first species to diverge (Chevret *et al.*, 1994; Britton-Davidian *et al.,* 1995) from the other *Mastomys* species and that it is sister to the *M. natalensis*, *M. erythroleucus* and *M. huberti* group. Previous studies using the dated divergence of the *Mus*/*Rattus* split of 12 million years ago (MYA), (Russo *et al.*, 2010) as a calibration point, estimated that the divergence of *Mastomys coucha* from its sister lineage took place approximately 3 MYA (LeCompte *et al.,* 2002, Colangelo *et al.*, 2013). Similar to these findings the BEAST analysis performed in the current study showed a divergent date between *M. coucha* and *M. natalensis* of approximately 3.8 MYA.

The retrospective analyses of the two historic South African strains provides and important baseline to studying EMCV in the sub-region, in addition intensified sampling of suspected hosts and characterisation of both contemporary viruses and retrospective studies, would allow for improved accuracy in determining the geographical origin of the causative agent during future outbreaks. The expansion of this project to include not only the epidemiology of the virus but also to include aspects of the host phylogeography provided the study with a more well-rounded epidemiological investigation as well as

establishing methodology which may prove useful in future studies that may build on initial findings. EMCV infection does not often manifest in severe clinical symptoms in humans and is therefore usually of little importance, however, its broad geographic distribution and diverse potential host range can make it of concern in specific circumstances, e.g. in immunocompromised individuals and those who have received porcine xenografts (Carocci & Bakkali-Kassimi, 2012).

Future studies investigating the diversity of Picornaviruses, or any other organisms, have been and will be greatly aided by advances in technology that allow for direct RNA sequencing, and rapid and cost-effective generation of whole genome sequence data. A wide range of commercially available platforms and advances in next generation and 3rd generation sequencing technologies have already impacted on epidemiological studies and, in particular, on refining transmission pathways (Etienne *et al.,* 2012; Köser *et al.,* 2012). Whereas we are rapidly reaching the point where extensive sequence data can be rapidly generated, these approaches still carry serious cost drawbacks that put them out of the reach of many laboratories. Until this changes, more traditional methods, such as those used and developed in the current study, will still have a place in the developing world.

Literature Cited

Acland, H.M., and Littlejohns, I.R. (1986). Encephalomyocarditis virus infection. In: *Diseases of swine* (6th ed); 399–402. Leman, A.D., Straw, B., Glock, R.D., Mengeling, W.L., Penny, R.H C., and Scholl, E. (Editors). Iowa State: University Press, Ames, Iowa.

Acland, H.M., Littlejohns, I.R. and Walker, R.I. (1970). Suspected encephalomyocarditis virus infection of pigs. *Australian Veterinary Journal*. **46**, 348.

Acland, H.M. and Littlejohns, I.R. (1975). Encephalomyocarditis virus infection of pigs. An outbreak in New South Wales. *Australian Veterinary Journal*. **51**, 409-415.

Adams, M.J., King, A.M.Q. and Carstens, E.B. (2013). Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2013). *Archives of Virology*. **158,** 2023–2030.

Adkins, R.M., Gelke, E.L., Rowe, D., and Honeycutt, R.L. (2001). Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. *Molecular Biology and Evolution*. **18**, 777-791.

Akiyama, T., Ishida, J., Nakagawa, S., Ogawara, H., Watanabe, S., Itoh, N., Shibuya, M. and Fukami, Y. (1987). Genistein, a specific inhibitor of tyrosine specific protein kinases. *Journal of Biological Chemistry*. **262,** 5592-5595.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research.* **25,** 3389-3402.

Amineva, S.P., Aminev, A.G., Andreev, V.G., Orljankin, B.G. and Gusev, A.A. (1997). Diagnostic porcine encephalomyocarditis virus by nested polymerase chain reaction. *Unpublished.*

Aminev, A.G., Amineva, S.P., and Palmenberg, A.C. (2003*a*). Encephalomyocarditis viral protein 2A localizes to nucleoli and inhibits cap-dependent mRNA translation. *Virus Research*. **95,** 45-57.

Aminev, A.G., Amineva, S.P. and Palmenberg, A.C. (2003*b*). Encephalomyocarditis virus (EMCV) proteins 2A and 3BCD localize to nuclei and inhibit cellular mRNA transcription but not rRNA transcription. *Virus Research*. **95**, 59-73.

An, D.J., Jeong, W., Jeoung, H.Y., Yoon, S.H., Kim, H.J., Choi, C.U. and Park, B.K. (2009). Encephalomyocarditis in Korea: serological survey in pigs and phylogenetic analysis of two historical isolates. *Veterinary Microbiology*. **137,** 37-44.

Armstrong, C. (1939). The experimental transmission of poliomyelitis to the eastern cotton rat, *Sigmodon hispidus hispidus*. *Public Health Reports*. **54**, 2302-2305.

Arnold, E., Luo, M., Vriend, G., Rossman, M.G., Palmenberg, A.C., Parks, G.D., Nicklin, M.J.H. and Wimmer, E. (1987). Implications of the picornavirus capsid structure for polyprotein processing. *Proceedings of the National Academy of Sciences*. **84**, 21-25.

Axelrod, D.I., and Raven, P.H. (1978). Late Cretaceous and Tertiary vegetation history of Africa. In *Biogeography and ecology of southern Africa*. Dr. Werger, M.J.A. (Editor). The Hague: Junk Publishers.

Backues, K.A., Hil, M., Palmenberg, C., Miller, C., Soike, K.F., and Aguilar, R. (1999). Genetically engineered Mengovirus vaccination of multiple captive wildlife species. *Journal of Wildlife Diseases.* **35,** 384-387.

Bae, Y.S., Eun, H.M. and Yoon, J.W. (1989). Genomic differences between the diabetogenic and non diabetogenic variants of encephalomyocarditis virus. *Virology*. **170,** 282-287.

Bai, J., Jiang, K., Zeshan, B., Wang, X., Li, Y. and Jiang, P. (2012). Inhibition of encephalomyocarditis virus replication by shRNA targeting 1D and 3AB genes in vitro and in vivo. *Virus Genes.* **44,** 183-190.

Baker, B.H., and Wohlenberg, J. (1971). Structure and evolution of the Kenya rift valley. *Nature*. **229**, 538-542.

Blyn, L.B., Towner, J.S., Semler, B.L., and Ehrenfeld, E. (1997).Requirement of poly(rC) binding protein 2 for translation of poliovirus RNA. *Journal of Virology*. **71**, 6243-6246.

Bastos, A.D.S. (1998). Detection and characterisation of foot-and-mouth disease virus in sub-Saharan Africa. *Onderstepoort Journal of Veterinary Research*. **65,** 37-47.

Bastos, A.D., Nair, D., Taylor, P.J., Brettschneider, H., Kirsten, F., Mostert, E., Von Maltitz, E., Lamb, J.M., Van Hooft, P., Belmain, S.R., Contrafatto, G., Downs, S., and Chimimba, C.T. (2011). Genetic monitoring detects an overlooked cryptic species and reveals the diversity and distribution of three invasive *Rattus* congeners in South Africa. *BMC Genetics*. **12**, 26-43.

Bautista, G., Jungeblut, C.W., and Kodza, H., 1957. Nonspecific cross protection in hamsters between Columbia SK virus and poliomyelitis virus type II (MEF) virus infection. *Journal of Infectious Diseases.* **100,** 268-272.

Benton, M.J. and Donoghue, P.C. (2007). Paleontological evidence to date the tree of life. *Molecular Biology and Evolution*. **24,** 26-53.

Billinis, C. (2009). Encephalomyocarditis virus infection in wildlife species in Greece. *Journal of Wildlife Diseases*. **45,** 522-526.

Boege, U., Kobasa, D., Onodera, S., Parks, G.D., Palmenberg, A. C. & Scraba, D.G. (1991). Characterization of Mengoviru sneutralization epitopes. *Virology*. **181**, 1-13.

Brewer, L. A., Lwamba, H. C. M., Murtaugh, M. P., Palmenberg, A. C., Brown, C. and Njenga, M. K. (2001). Porcine encephalomyocarditis virus persists in pig myocardium and infects human myocardial cells. *Journal of Virology*. **75,** 11621-11629.

Britton-Davidian, J., Catalan, J., Granjon, L., and Duplantier, J.-M. (1995). Chromosomal phylogeny and evolution in the genus *Mastomys* (Mammalia, Rodentia). *Jounal of Mammalogy*. **76,** 248–262.

Bronner, G.N., Van der Merwe, M. and Njobe, K. (2007). Nongeographic Cranial Variation in Two Medically Important Rodents from South Africa, *Mastomys natalensis* and *Mastomys coucha*. *Journal of Mammalogy*. **88**, 1179-1194.

Carocci, M. and Bakkali-Kassimi, L. (2012). The encephalomyocarditis virus. *Virulence*. **4,** 351-367.

Canelli, E., Luppi, A., Lavazza, A., Lelli, D., Sozzi, E., Moreno Martin, A. M., Gelmetti, D., Pascotto, E., Sandri, C., Magnone, W., and Cordioli, P. (2010). Encephalomyocarditis virus infection in an Italian zoo. *Virology Journal*. **7,** 64.

Castryck, F., Miry, C., Koenen, F. and Ducatelle, R. (1996). Clinical and pathological aspects of encephalomyocarditis virus infections in pigs in Belgium. *Proceedings of the 14th International Pig Veterinary Society Congress.* Italy: Bologna, 7-10 July, 132.

Chevret, P., Granjon, L., Duplantier, J.-M., Denys, C., and Catzeflis, F.M. (1994). Molecular phylogeny of the *Praomys* complex (Rodentia: Murinae): a studybased on DNA/DNA hybridization experiments. *Zoological Journal of the Linnaean Society.* **112,** 425–442.

Christianson, W.T., Kim, H.S., Joo, H.S., and Barnes, D.M. (1990). Reproductive and neonatal losses associated with possible encephalomyocarditis virus infection in pigs. *Veterinary Research*. **126**, 54-57.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidiniumthiocyanate–phenol–chloroform extraction. *Analytical Biochemistry.***162,** 156-159.

Clement, M., Posada, D., and Crandall, K. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*. **9,** 1657-1660.

Cohen, S.H., Naviaux, R.K., van den Brink, K.M. and Jordan, G.W. (1988). Comparison of the nucleotide sequences of diabetogenic and nondiabetogenic encephalomyocarditis virus. *Virology*. **166,** 603-607.

Colangelo, R., Corti, M., Verheyen, E., Anessi, F., Oguge, N., Makundi, R.H., Verheyen, W. (2005). Mitochondrial phylogeny reveals differential modes of chromosomal evolution in the genus *Tatera* (Rodentia: Gerbillinae) in Africa. *Molecular Phylogenetics and Evolution*. **35,** 556-568.

Colangelo, P., Verheyen, E., Leirs, H., Tatard, C., Denys, C., Dobigny, G., Duplantier, J-M.,Brouat, C., Granjon., L., and Lecompte, E. (2013). A mitochondrial phylogeographic scenario for the mostwidespread African rodent, *Mastomys natalensis*. *Biological Journal of the Linnean Society*. **108,** 901-916.

Craighead, J.E. (1965). Some Properties of the encephalomyocarditis, Columbia SK and Mengo Viruses. *Experimental Biology and Medicine*.**119,** 408-412.

Cummings, K.J., Cox-Ganser, J., Riggs, M.A., Edwards, N., Kreiss, K. (2007).Respirator donning in post-hurricane New Orleans.Emerging Infectious Diseases [serial on the Internet].Available from http://www.cdc.gov/EID/content/13/5/700.htm.

Dale, B., and Dragon, E.A. (1994). Polymerase chain reaction in infectious disease diagnosis. *Laboratory Medicine*. **25**, 637–641.

Davis, D.,H.,S. (1964). Ecology of wild rodent plague. In *Ecological Studies in Southern Africa*; 301. Davis, D., H., S. (editor). W. Junk, The Hague.

Dea, S.A., Bilodeau, R. and Martineau, G.P. (1991*a*). Isolation of encephalomyocarditis virus among stillborn and post-weaning pigs in Quebec. *Archives of Virology*. **117,** 121-128.

Dea, S., Bilodeau, R., Sauvageau, R. and Martineau, G.P. (1991*b*). Outbreaks in Quebec pig farms of respiratory and reproductive problems associated with encephalomyocarditis virus. *Journal of Veterinary Diagnostic Investigation*. **3,** 275-282.

De Bruyn, P.J.N., Bastos, A.D.S., Eadie, C., Tosh, C., and Bester, M.N. (2008). Mass mortality of adult male subantarctic fur seals: are alien mice the culprits? *PLoS ONE*. **3,** e3757.

DeBry, R. W., and Sagel, R.M. (2001). Phylogeny of Rodentia(Mammalia) inferred from the nuclear-encoded gene IRBP. *Molecular Phylogenetics and Evolution*. **19,** 290-301.

De Jesus, N.H. (2007). Epidemics to eradication: the modern history of poliomyelitis. *Virology Journal*. **4**, 70.

deMenocal, P.B. (2004).African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth Planet Science Letters.* **220,** 3-24.

Denis, P., Liebig, H.D., Nowotny, N., Billinis, C., Papadopoulos, O., O'Hara, R.S., Knowles, N.J. and Koenen, F. (2006). Genetic variability of encephalomyocarditis virus (EMCV) isolates. *Veterinary Microbiology.* **113**, 1-12.

Díaz Arenas, C. and Lehman, N. (2010). Quasispecies-like behavior observed in catalytic RNA populations evolving in a test tube. *BMC Evolutionary Biology*. **10**, 80.

Dick, G.W.A., Smithburn, K.C., and Haddow, A.J. (1948). Mengo encephalomyelitis virus: isolation and immunological properties. *British Journal of Experimental Pathology*. **29**, 547-558.

Didomenico, N., Link, H., Knobel, R., Caratsch, T., Weschler, W., Loewy, Z.G., and Rosenstraus, M. (1996). COBAS AMPLICOR™: a fully automated RNA and DNA amplification and detection system for routine diagnostic PCR. *Clinical Chemistry*. **42**, 1915-1923.

Dippenaar, N.J, Swanepoel, P. and Gordon, D.H. (1993). Diagnostic morphometrics of two medically important southern African rodents, *Mastomys natalensis* and *M. coucha* (Rodentia: Muridae). *South African Journal of Science.* **89**, 300–303.

Dobrokhotov, B.P. (1982.) The utilization of electrophoresis of haemoglobins for identification of species of the genus *Mastomys* from West Africa. *Zoologichesky Zhurnal*. **61**, 290-294.

Drake, J.W., Charlesworth, B., Charlesworth, D. and Crow, J.F. (1998). Rates of spontaneous mutation. *Genetics.* **148,** 1667-1686.

Dubey, S., Zaitsev, M., Cosson, J.F., Abdukadier, A., and Vogel, P. (2006). Pliocene and Pleistocene diversification and multiple refugia in a Eurasian shrew (Crocidura suaveolens group). *Molecular Phylogenetics and Evolution*. **38**, 625-647.

Dubey, S., Cosson, J.F., Vohralík, V., Kryštufek, B., Diker, E., and Vogel, P. (2007). Molecular evidence of Pleistocene bidirectional faunal exchange between Europe and the Near East: the case of the bicolored shrew (*Crocidura leucodon*, Soricidae). *Journal of Evolutionary Biology*. **20**, 1799–1808.

Duplantier, J.M., Britton-Davidian, J., and Granjon, L. (1990*a*). Chromosomal characterization of three species of the genus *Mastomys* in Senegal. *Zeitschrift fu¨ r Zoologister Systematik und Evolutionsforschung*. **28**, 289–298.

Duplantier, J.M., Granjon, L., Mathieu, E. and Bonhomme, F. (1990*b*) Structures ge´ne´tiques compare´es de trois espe`ces de rongeurs africains du genre *Mastomys* au Senegal. *Genetica.* **81**, 179-192.

Drummond, A.J., Rambaut, A., Shapiro, B., and Pybus, O.G. (2005). Bayesian coalescentinference of past population dynamics from molecular sequences. *Molecular Biology and Evolution.* **22,** 1185-1192.

Drummond, A.J. and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology.* **7**, 214.

Dutcher, S. (2001). The tubulin fraternity: alpha to eta. *Current Opinion in Cell Biology*. **13,** 49-54.

Duke, G.M., Hoffman, M.A. and Palmenberg, A.C. (1992).Sequence and structural elements that contribute to efficient encephalomyocarditis virus RNA translation. *Journal of Virology.* **66,** 1602-1609.

Eigen, M. (1996). On the nature of virus quasispecies. *Trends in Microbiology.* **4,** 216-218.

Elnifro, E.M., Ashshi, A.M., Cooper, R.J., and Klapper, P.E. (2000). Multiplex PCR: optimization and application in diagnostic virology. *Clinical Microbiology Reviews*. **13**, 559 – 570.

Etienne, K.A., Gillece, J., Hilsabeck, R., Schupp, J.M., Colman, R., Lockhart, S.R., Gade, L., Thompson, E.H., Sutton, D.A., Neblett-Fanfair, R., Park, B.J., Turabelidze, G., Keim, P., Brandt, M.E., Deak, E., and Engelthaler, D.M. (2012). Whole Genome Sequence Typing to Investigate the Apophysomyces Outbreak following a Tornado in Joplin, Missouri, 2011. *PLoS ONE*. **7,** e49989.

Eun, H.M., Bae,Y.S. and Yoon, J.W. (1988). Amino acid differences in capsid protein, VP1, between diabetogenicand nondiabetogenic variants of encephalomyocarditis virus. *Virology*. **163,** 369-373.

Fata-Hartley, C.L. and Palmenberg, A.C. (2005). Dipyridamole reversibly inhibits mengovirus RNA replication. *Journal of Virology*. **79,** 11062-11070.

Fenner, F.J., Gibbs, E.P.J., Murphy, F.A., Rott, R., Studdert, M.J. and White, D.O. (1993). *Veterinary Virology.* (2nd edition.) San Diego: Academic Press Inc.

Follett, E.A., Pringle, C.R. and Pennington, T.H. (1975). Virus development in enucleate cells: echovirus, poliovirus, pseudorabies virus, reovirus, respiratory syncytial virus and Semliki Forest virus. *Journal of General Virology*. **26**, 183-196.

Gainer, J.H., Sandefur, J.R., and Bigler, W.J. (1968). High mortality in a Florida swine herd infected with the encephalomyocarditis virus. An accompanying epizootiologic survey. *Cornell Veterinarian*. **58,** 31-47.

Gordon, D.H. (1978). Distribution of sibling species of the *Praomys (Mastomys) natalensis* group in Rhodesia (Mammalia: Rodentia). *Journal of Zoology*.**1 86**, 397-401.

Gordon, D.H. and Watson, C.R.R. (1986). Identification of cryptic species of rodent (*Mastomys, Aethomys, Saccostomus*) in the Kruger National Park. *South African Journal of Wildlife Research*. **21,** 95-99.

Grachev, S.A., Pletnev, A.G., Vasilenko, S.K., Petrov, N.A. and Chizhikov, V.E. (1983). The nucleotide sequence of the 3'-terminal region of Encephalomyocarditis virus RNA. *BioorgKhim*. **9,** 986-989.

Granjon, L., Duplantier, J.M., Catalan, J. and Britton-Davidian, J. (1997.) Systematics of the genus *Mastomys* (Thomas, 1915) (Rodentia: Muridae). A review. *Belgium Journal of Zoology*. **127**, 7-18.

Gratz, N. (1997). The burden of rodent-borne diseases in Africa south of the Sahara. *Belgian Journal of Zoology*. **127**, 71-84.

Green, C.A., Keogh, H., Gordon, D.H., Pinto, M., and Hartwig, E.K. (1980). The distribution, identification and naming of the *Mastomys natalensis* species complex in southern Africa (Rodentia: Muridae). *Journal of Zoology*. **192**, 17–23.

Greenfield, L., and White, T.J. (1993).Sample preparation methods. In *Diagnostic molecular microbiology: principles and applications;* pp. 122–137 Persing, D.H., Smith, T.F., Tenover, F.C. and White, T.J. (Editors), Washington DC: American Society for Microbiology.

Grobler, D.G., Raath, J.P., Braack, L.E.O., Keet, D.F., Gerdes, G.H., Barnard, B.J.H., Kriek, N.P.J., Jardine, J. and Swanepoel, R. (1995). An outbreak of encephalomyocarditis-virus infection in free-ranging African elephants in the Kruger National Park. *Onderstepoort Journal of Veterinary Research.* **62,** 97-108.

Gu, X, and Zhang, J. (1997). A simple method for estimating the parameters of substitution rate variation among sites. *Molecular Biology and Evolution.* **14.** 1106-1113.

Gualandi, G.L., Cammi, G. and Cardeti, G. (1989). A serologic survey of encephalomyocarditis virus infection in pigs in Italy. *Microbiologica.* **12**, 129-132.

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*. **52,** 696-704.

Guy, M., Chilmonczyk, S., Cruciere, C., Eloit, M. and Bakkali-Kassimi, L. (2009). Efficient infection of buffalo rat liver-resistant cells by encephalomyocarditis virus requires binding to cell surface sialicacids. *Journal of General Virology*. **90,** 187-196.

Hallett, J.M. (1997.) *Cytological and cytogenetical studies on the multimammate mouse Praomys (Mastomys) natalensis.* M.Sc. thesis. University of the Witwatersrand at Johannesburg, South Africa.

Hammoumi, S., Guy, M., Eloit, M. and Bakkali-Kassimi, L. (2012). Encephalomyocarditis virus may use different pathways to initiate infection of primary human cardiomyocytes. *Archives of Virology*. **157,** 43-52.

Helwig, F.C. and Schmidt, E.C.H. (1945). A filter-passing agent producing interstitial myocarditis in anthropoid apes and small animals. *Science*. **102**, 31-33.

Hellen, C.U., Witherell, G.W., Schmid, M., Shin, S.H., Pestova, T.V., Gil, A., andWimmer, E. (1993). A cytoplasmic 57-kDa protein that is required fortranslation of picornavirus RNA by internal ribosomal entry is identical to the nuclear pyrimidine tract-binding protein. *Proceedings of the National Academy of Science.* **90,** 7642-7646.

Hewitt, G.M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B.* **359**,183-195.

Holland, J J., Domingo, E., de la Torre, J.C. and Steinhauer, D.A. (1990). Mutation frequencies at defined single codon sites in vesic-stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. *Journal of Virology*. **64,** 3960-3962.

Holmes, E.C. (1998). Molecular epidemiology and evolution of emerging infectious diseases. *British Medical Bulletin*. **54**, 533-543.

Hu, D.G., Chan, C.H., Shieh, W.Y. and Chang, C.S. (1993). Epidemiological studies of swine encephalomyocarditis virus infection. *Research Report - Animal Industry Research Institute*. Taiwan Sugar Corporation 1993 (1992/93), 149-154.

Hubbard, G.B., Soike, K.F., Butler, T.M., Carey, K.D., Davis, H., Butcher, W.I. and Gauntt, C.J. (1992). An *encephalomyocarditis virus* epizootic in a baboon colony.*Laboratory Animal Science*. **42**, 233-239.

Huelsenbeck, J.P. and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics.* **17,** 754-755.

Hughes, A.L. (2004). Phylogeny of the *Picornaviridae* and differential evolutionary divergence of picornavirus proteins. *Infection, Genetics and Evolution*. **4**, 143-152.

Isaäcson, M., Arntzen, L. and Taylor, P. (1981).Susceptibility of Members of the *Mastomys natalensis* Species Complex to Experimental Infection with *Yersinia pestis*. *The Journal of Infectious Diseases.***144**, 80.

Izumi, R. E., Valdez, B., Banerjee, R., Srivastava, M. and Dasgupta, A. (2001). Nucleolin stimulates viral internal ribosome entry site-mediated translation. *Virus Research*. **76**, 17-29.

Jansa, S.A. and Weksler, M. (2004). Phylogeny of muroid rodents: Relationships within and among major lineages as determined by IRBP gene sequences. *Molecular Phylogenetics and Evolution*. **31**, 256-276.

Jones, P., Mahamba, C., Rest, J. and Andre, C. (2004). Fatal inflammatory heart disease in a bonobo (*Pan paniscus*). *Journal of Medical Primatology*. **34,** 45-49.

Jones, P., Cordonnier, N., Mahamba, C., Burt, F.J., Rakotovao, F., Swanepoel, R., André, C., Dauger, S. and BakkaliKassimi, L. (2011). Encephalomyocarditis virus mortality in semi-wild bonobos (*Pan panicus*). *Journal of Medical Primatology.* **40**, 157-163.

Kassimi, L.B., Boutrouille, A., Gonzague, M., Mbanda, A.L. and Cruciere, C. (2002). Nucleotide sequence and construction of an infectious cDNA clone of an EMCV strain isolated from aborted swine fetus. *Virus Research*. **83,** 71-87.

Kessler, M. J., Brown, R.J., Kalter, S. S., Altman, N. H.(1982). A fatal epizootic of undetermined etiology in new world monkeys. *American Journal of Primatology*. **3,** 253-261.

Kessler, M. J. (2013). Comment on Masek-Hammerman *et al.*, "Epizootic Myocarditis Associated With Encephalomyocarditis Virus in a Group of Rhesus Macaques (*Macaca mulatta*)," Vet Pathol. 2012;49:386–392. *Veterinary Pathology*. **50,** 209.

Kim, H.S., Christianson, W.T. and Joo, H.S. (1989*a*). Pathogenic properties of encephalomyocarditis virus isolates in swine fetuses. *Archives of Virology*. **109**, 51-57.

Kim, H.S., Joo, H.S., and Bergeland, M.E. (1989*b*). Serologic, virologic and histopathologic observations of encephalomyocarditis virus infection in mummified and stillborn pigs. *Journal of Veterinary Diagnostic Investigation*. **1,** 101-104.

Kim, H.S., Christianson, W.T. and Joo, H.S. (1991). Characterization of encephalomyocarditis virus isolated from aborted swine fetuses. *American Journal of Veterinary Research.* **52**, 1649-1652.

Knowles, N.J., Dickinson, N.D., Wilsden, G., Carra, E., Brocchi, E., and De Simone, F. (1998). Molecular analysis of Encephalomyocaridits viruses isolated from pigs and rodents in italy. *Virus Resarch*. **57**, 53-62.

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

Knowles, N.J., Hovi, T, Hyypiä, T., King, A.M.Q., Lindberg, A.M., Pallansch, M.A., Palmenberg, A.C., Simmonds, P., Skern, T., Stanway, G., Yamashita, T. and Zell, R. (2012). *Picornaviridae*. In Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses; pp 855-880. King, A.M.Q., Adams, M.J., Carstens, E.B. and Lefkowitz, E.J. (Editors). San Diego: Elsevier.

Kobasa, D., Mulvey, M. Lee, J.S. and Scraba, D.G. (1995). Characterization of Mengo neutralizing epitopes. *Virology*. **214,** 118-127.

Koenen, F., De Clercq, K., and Strobbe, R. (1991). Isolatie van het *encephalomyocarditis virus* bijbiggenafkomstig van zeugen met reproduktiestoornissen in België. *Vlaams Diergeneeskd Tijdschr.* **60,** 113-115.

Koenen, F. (1994). Clinical, serological, virological and histological observations of encephalomyocarditis virus infection in pigs. *Pig Journal*. **33,** 21-29.

Koenen, F., De Clercq, K., Lefebvre, J., and Strobbe, R. (1994). Reproductive failure in sows following experimental infection with a Belgian EMCV isolate. *Veterinary Microbiology*.**3 9**, 111-116.

Koenen, F., Vanderhallen, H., Castryck, F., Miry, C. and Fockedey, M. (1996). Epidemiologic, pathogenetic and molecular analysis of the first encephalomyocarditis outbreak in fattening pigs in Belgium. *Proceedings of the 14th International Pig Veterinary Society Congress*. Italy: Bologna, 7-10 July, 96.

Koenen, F. and Vanderhallen, H. (1997). Comparative study of the pathogenic properties of a Belgian and a Greek encephalomyocarditis virus (EMCV) isolate for sows in gestation. *Journal of Veterinary Medicine B.* **44**, 281-286.

Koenen, F., Vanderhallen, H., Papadopoulos, O., Billinis, C., Paschaleri-Papadopoulou, E., Brocchi, E., De Simone, F., Carra, E., and Knowles, N.J. (1997). Comparison of the pathogenic, antigenic and molecular characteristics of two encephalomyocarditis virus (EMCV) isolates from Belgium and Greece. *Research in Veterinary Science*. **62**, 239-244.

Koenen, F., Vanderhallen, H., Dickinson, N.D., and Knowles, N.J. (1999). Phylogenetic analysis of European encephalomyocarditis viruses: Comparison of two genomic regions. *Archives of Virology*. **144**, 893-903.

Köser C.U., Holden, M.T., Ellington, M.J., Cartwright, E.J., Brown, N.M., Ogilvy-Stuart, A.L., Hsu, L.Y., Chewapreecha, C., Croucher, N.J., Harris, S.R., Sanders, M., Enright, M.C., Dougan, G., Bentley, S.D., Parkhill, J., Fraser, L.J., Betley, J.R., Schulz-Trieglaff, O.B., Smith, G.P., and Peacock, S.J. (2012). Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *New England Journal of Medicine.* **366,** 2267-2275.

Krishnaswamy, S. and Rossman, M.G. (1990). Structural refinement and analysis of Mengo virus. *Journal of Molecular Biology*. **211**, 803-844.

Kumar, S. and Hedges, S.B. (1998). A molecular timescale for vertebrate evolution. *Nature*.**392**, 917-920.

Kyu, B., Matsumori, A., Sato, Y., Okada, I., Chapman, N.M., and Tracy, S. (1992). Cardiac persistence of cardioviral RNA detected by polymerase chain reaction in a murine model of dilated cardiomyopathy. *Circulation*. **86**, 522–530.

LaRue, R., Myers, S., Brewer, L., Shaw, D.P., Brown, C., Seal, B.S. and Njenga, M.K. (2003). A Wild-Type Porcine Encephalomyocarditis Virus Containing a Short Poly(C) Tract Is Pathogenic to Mice, Pigs, and *Cynomolgus Macaques*. *Journal of Virology*. **77,** 9136-9146.

Lavrenchenko, L.A., Likhnova, O.P., Baskevich, M.I., and Afework, B. (1998). Systematics and distribution of *Mastomys* (Muridae, Rodentia) from Ethiopia, with the description of a new species. *Zeitschrift fu¨ r Zoologister Systematik und Evolutionsforschung.* **63**, 37-51.

Lecompte, E., Granjon, L., Peterhans, J.K., and Denys, C. (2002). Cytochrome *b*-based phylogeny of the *Praomys* group (Rodentia, Murinae): a new African radiation? *Comptes Rendus Biologies.* **325,** 827–840.

Lecompte, E., Denys, C., and Granjon, L. (2005*a*). Confrontation of morphological and molecular data: the *Praomys* group (Rodentia: Murinae) as a case of adaptive convergences and morphological statis. *Molecular Phylogenetics and Evolution*. **37**, 899-919.

Lecompte, E., Brouat, C., Duplantier, J., Galan, M., Granjon, L., Loiseau, A., Mouline, K. and Cosson, J. (2005*b*). Molecular identification of four cryptic species of *Mastomys* (Rodentia, Murinae). *Biochemical Systematics and Ecology*. **33**, 681–689.

Lecompte, E., Fichet-Calvet, E., Daffis, S., Koulémou, K., Sylla, O., Kourouma, F., Doré, A., Soropogui, B., Aniskine, V., Allali, B., Kouassi Kan, S., Lalis, A., Koivogui, L., Günther, S., Denys, C., and ter Meulen, J. (2006). *Mastomys natalensis* and Lassa fever, West Africa. *Emerging Infectious Diseases*. **12**, 1971-1974.

Leirs, H., Verhagen, R., Verheyen, W., Mwanjabe, P., and Mbise, T. (1996). Forecasting rodent outbreaks in Africa; an ecological basis for *Mastomys* control in Tanzania. *Journal of Applied Ecology*. **33**, 937-943.

Librado, P. and Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. **25,** 1451-1452.

Lidsky, P.V., Hato, S., Bardina, M.V., Aminev, A.G., Palmenberg, A.C., Sheval, E.V., Polyakov, V.Y., Van Kuppeveld, F.J. and Agol, V.I. (2006). Nucleocytoplasmic traffic disorder induced by cardioviruses. *Journal of Virology*. **80**,2705-2717.

Liu, H., Yan, Q., Zhao, B., Luo, J., Wang, W., Du, Y., Yan, J., and He, H. (2013). Isolation, molecular characterization, and phylogenetic analysis of encephalomyocarditis virus from South China Tigers in China. *Infection, Genetics and Evolution*. **19**, 240-243.

Lobert, P.E., Escriou, N., Ruelle, J. and Michiels, T. (1999). A coding RNA sequence acts as a replication signal in cardioviruses. *Proceedings of the National Academy of Science of the United States of America.* **96**, 11560-11565.

Mann, L.M., Anderson, K., Luo, M. and Bond, C.W. (1992). Molecular and structural basis of hemagglutination in Mengovirus.*Virology*. **190,** 337-345.

Martin, L.R., Neal, Z.C., McBride, M.S. and Palmenberg, A.C. (2000). Mengovirus and encephalomyocarditis virus poly(C) tract lengths can affect virus growth in murine cell culture. *Journal of Virology*.**74**, 3074-3081.

Martinez-Salas, E., Saiz, M., and Sobrino, F. (2008).Foot-and-Mouth Disease Virus. In *Animal Viruses: Molecular Biology*; pp.1-38. Mettenleiter, T.C., and Sobrino, F. (Editors). Norwich: Caister Academic Press.

Masek-Hammerman, K., Miller, A.D., Lin, K.C., MacKey, J., Weissenbock, H., Gierbolini, L., Burgos, A., Perez, H., and Mansfield, K.G. (2012). Epizootic myocarditis associated with encephalomyocarditis virus in a group of rhesus macaques (*Macaca mulatta*). *Veterinary Pathology*. **49**, 386-392.

Maurice, H., Nielen, M., Brocchi, E., Nowotny, N., Bakkali-Kassimi, L., Billinis, C., Loukaides, P., O`Hara, R.S., and Koenen, F. (2005). The occurence of encephalomyocarditis virus (EMCV) in European pigs from 1990 to 2001. *Epidemiology & Infection*. **133**, 547-557.

Maurice, H., Nielen, M., Vyt, P., Frankena, K., and Koenen, F. (2007). Factors related to the incidence of clinical encephalomyocarditis virus (EMCV) infection on Belgian pig farms. *Preventive veterinary medicine*. **78,** 24-34.

McBride, A.E., Schlegel, A., and Kirkegaard, K. (1996). Human protein Sam68 relocalization and interaction with poliovirus RNA polymerase in infected cells. *Proceedings of the National Academy of Science of the United States of America*. **93**, 2296-2301.

McCormick, J.B., Webb, P.A., Krebs, J.W., Johnson, K.M. and Smith, E.S. (1987). A prospective study of epidemiology and ecology of Lassa Fever. *Journal of Infectious Diseases*. **155**, 437-444.

Medvedkina, O.A., Scarlat, I.V., Kalinina, N.O. and Agol, V.I. (1974).Virus-specific proteins associated with ribosomes of Krebs-II cells infectedwith encephalomyocarditis virus. *FEBS Letters*. **39**, 4-8.

Meerovitch, K., Svitkin, Y.V., Lee, H.S., Lejbkowicz, F., Kenan, D.J., Chan, E. K., Agol, V.I., Keene, J.D. and Sonenberg, N. (1993). La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. *Journal of Virology*. **67**, 3798-3807.

Meng, X.J., Paul, P. and Vaughn, E.M. Nucleotide sequence of the VP1 gene of swine encephalomyocarditis virus (EMCV) NVSL-FL strain. *Unpublished.*

Mettenleiter T.C., and Sobrino, F. (Editors). (2008). *Animal Viruses: Molecular Biology*. Norwich: Caister Academic Press.

Meyer, R.F., Brown, C.C., House, C., House, J.A., Molitor, T.W. (1991). Rapid and sensitive detection of foot-and-mouth disease virus in tissues by enzymatic RNA amplification of the polymerase gene. *Journal of Virological Methods*. **34,** 161-172.

Michaux, J., Reyes, A., and Catzeflis, F. (2001). Evolutionary history of the most speciose mammals: molecular phylogeny of muroid rodents. *Molecular Biology and Evolution.* **17**, 280-293.

Misonne, X. (1969). African and Indo-Australian Muridae. *Evolutionary trends.* Ann Royal Museum for Central Africa, Tervuuren. **172**, 1-219.

Mohr, K., Vibe-Petersen, S., Jeppesen, L.L., Bildsoe, M., and Leirs, H. (2003). Foraging of multimammate mice, *Mastomys natalensis*, under different predation pressure: cover, patch-dependent decisions and density-dependent GUDs. *Oikos.* **100,** 459-468.

Molla, A., Jang, S.K., Paul, A.V., Reuer, Q. and Wimmer, E. (1992). Cardioviral internal ribosomal entry site is functional in a genetically engineered dicistronic poliovirus. *Nature.* **356,** 255-257.

Monath, T.P., Newhouse, V.F., Kemp, G.E., Setzer, H.W. and Cacciapuoti, A. (1974). Lassa virus isolation from *Mastomys natalensis* rodents during an epidemic in Sierra Leone.*Science.***185,** 263-265.

Mouline, K., Granjon, L., Galan, M., Tatard, C., Abdoullaye, D., Atteyene, S.A.,Duplantier, J.P., Cosson, J-F. (2008). Phylogoegraphy of a Sahelain rodent species *Mastomys huberti: a* Plio-Pleistocene story of emergence andcolonization of humid habitats. *Molecular Ecology*. **17,** 1036-1053.

Musser, G.G., and Carleton, M.D. (2005).Superfamily Muroidea. In *Mammal species of the world A taxonomic and geographic reference Volume 2;* pp.894-1531.(3rd edition.) Wilson, D.E., Reeder, D.M. (Editors). Baltimore: Johns Hopkins University.

Nelsen-Salz, B., Zimmermann, A., Wickert, S.T., Arnold, G., Botta, A., Eggers, H.J. and Kruppenbacher, J.P. (1996). Analysis of sequence and pathogenic properties of two varients of *Encephalomycarditis virus* differing in a single amino acid in VP1. *Virus Research*. **41,** 109-122.

Nei, M. and Tajima, F. (1981).DNA polymorphism detectable by restriction endonucleases. *Genetics.* **97**, 145-163.

Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press: New York, USA.

Nicolas, V., Quérouil, S., Verheyen, E., Verheyen, W., Mboumba, J.F., Dillen, M., and Colyn, M. (2006). Mitochondrial phylogeny of African wood mice, genus *Hylomyscu*s (Rodentia, Muridae): Implications for their taxonomy and biogeography. *Molecular Phylogenetics and Evolution*. **38**, 779-793.

Nicolas, V., Bryja, J., Akpatou, B., Konecny, A., Lecompte, E., Colyn, M., Lalis, A., Couloux, A., Denys, C., and Granjon, L. (2008). Comparative phylogeography of two sibling species of forest-dwelling rodent (*Praomysrostratus* and *P. tullbergi*) in West Africa: different reactions to past forest fragmentation. *Molecular Ecology*. **17**, 5188-5134.

Nix, W.A., Oberste, M.S., Gotuzzo, E., Blair, P., Ksiazek, T.G., Comer, J.A., Goldsmith, C.S., Olson, J. and Kochel, T.J. Human encephalomyocarditis virus disease in Peru. *Unpublished.*

Oberste, M.S., Gotuzzo, E., Blair, P., Nix, W.A., Ksiazek, T.G., Comer, J.A., Rollin, P., Goldsmith, C.S., Olson, J., and Kochel, T.J. (2009). Human Febrile Illness Caused by Encephalomyocarditis virus Infection, Peru. *Emerging Infectious Diseases*. **15**, 640-646.

Palmenberg, A.C. and Duke, G.M.The genomic sequence of Mengovirus and its relationship to other cardioviruses. *Unpublished.*

Palmenberg, A.C., Kirby, E.M., Janda, M.R., Drake, N.L., Duke, G.M., Potratz, K.F. and Collett, M.S. (1984). The nucleotide and deduced amino acid sequences of the encephalomyocarditis viral polyprotein coding region. *Nucleic Acids Research*. **12,** 2969-2985.

Palmenberg, A.C., Spiro, D., Kuzmickas, R., Wang, S., Djikeng, A., Rathe, J.A., Fraser-Liggett, C.M., Liggett, S.B. (2009). Sequencing and analyses of all known human rhinovirus genomes reveals structure and evolution. *Science*. **324**, 55.

Park, N.Y., Ri, C.Y., Chung, C.Y., Kee, H.Y., and Bae, S.Y. (1992). Pathological findings on encephalomyocarditis virus infections of swine in Korea. *Korean Journal of Veterinary Research*. **32**, 99-109.

Paschaleri-Papadopoulou, E. Axiotis, I. and Laspidis, C. (1990). Encephalomycarditis of swine in Greece. *Veterinary Research*. **126**, 364-365.

Paul, J.R. (1971). *A History of Poliomyelitis, Yale studies in the history of science and medicine*; pp. 16–18. New Haven, Conn: Yale University Press.

Petrov, N.A., Chizhikov, V.E., Blinov, V.M., Karginov, V.A., Mikryukov, N.N., Gutorov, V.V., Grishaev, M.P., Beklemishev, A.B. and Vassilenko, S.K. (1984). Nucleotide sequence of the 3'-terminus of encephalomyocarditis virus RNA. *BioorgKhim*. **10,** 274-279.

Philipps, A., Dauber, M., Groth, M., Schirrmeier, H., Platzer, M., Krumbholz, A., Wutzler, P., and Zell, R. (2012). Isolation and molecular characterization of a second serotype of the encephalomyocarditis virus. *Veterinary Microbiology*. **161,** 49-57.

Pollack, R., and Goldman, R. (1973). Synthesis of infective poliovirus in BSC-1monkey cells enucleated with cytochalasin B. *Science*. **179**, 915-916.

Posada, D.I. (2008). jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*. **25,** 1253-1256.

Rambaut, A. and Drummond, A.J. (2007). Tracer v1.4. Institute of Evolutionary Biology, University of Edinburgh. [http://beast.bio.ed.ac.uk/Tracer].

Ramos, J.R., Gomez, L., Mayo, M., and Sanchez, G. (1983). Infections caused by *Encephalomyocarditis virus* in pigs and other species in Cuba between 1975 and 1981. *Revista Cubana de Cienc Veterinarias*. **14**, 71-77.

Reddacliff, L.A., Kirkland, P.D., Hartley, W.J. and Reece, R.L. (1997). Encephalomyocarditis virus infections in an Australian zoo. *Journal of Zoo and Wildlife Medicine*. **28**, 153-157.

Robbins, C.B., Krebs Jr., J.W., and Johnson, K.M. (1983). *Mastomys* (Rodentia: Muridae) species distinguished by hemoglobin pattern differences. *American Journal of Tropical Medicine and Hygiene*. **32,** 624-630.

Robins, J.H., McLenachan, P.A., Phillips, M.J., Craig, L., Ross, H.A. and Matisoo-Smith, E. (2008). Dating of divergences within the *Rattus* genus phylogeny using whole mitochondrial genomes. *MolecularPhylogenetics and Evolution*. **49,** 460-466.

Roehe, P.M., Rodrigues, N.C., Oliveira, S. J. de, Guizzardi, I.I., Barcellos, D.E.S.N. de, Vidor, T., Oliveira, L.G. and Bangel, E.V. (1985). Encephalomyocarditis virus (EMCV) in swine in the State of Rio Grande doSul, Brazil. *Revista de Microbiologia.* **16**, 117-120.

Rosenstraus, M., Wang, Z., Chang, S.-Y., Debonville, D. and Spadoro, J. (1998). An internal control for routine diagnostic PCR: design, properties, and effect on clinical performance. *Journal of Clinical Microbiology*. **36**, 191-197.

Russo, I.M., Chimimba, C.T., and Bloomer, P. (2010). Bioregion heterogeneity correlates with extensive mitochondrial DNA diversity in the Namaqua rock mouse, *Micaelamys namaquensis* (Rodentia: Muridae) from southern Africa - evidence for a species complex. *BMC Evolutionary Biology*. **10,** 307.

Scraba D.G., and Palmenberg, A.C. (1999).Cardioviruses (*Picornaviridae*).In *Encyclopedia of virology*; pp 1–10.2nd edition.Webster, R.G. and Granoff, A. (Editors). London: Academic Press.

Seaman, J.T., Boulton, J.G. and Carrigan, M.J. (1986). Encephalomyocarditis virus disease of pigs associated with a plague of rodents. *Australian Veterinary Journal*. **63**, 292-294.

Sellers, R. F., and Daggupaty, S. M. (1990). The epidemic of foot-and-mouth disease in Saskatchewan, Canada, 1951–1952. *Canadian Journal of Veterinary Research.* **54**, 457–464.

Shi, K., Chen, J., Qu, S. and Liu, Q. Separation, Identification and Genome Sequence Analysis of EMCV-GXLC Isolated in Southern China. *Unpublished.*

Simmonds, P. (2006). Recombination and Selection in the Evolution of Picornaviruses and Other Mammalian Positive-Stranded RNA Viruses. *Journal of Virology*. **80**, 11124-11140.

Simpson, C.F. Lewis, A.L. and Gaskin L.M. (1977). Encephalomyocarditis virus infection of captive elephants. *Journal of the American Veterinary Medical Association*. **171,** 902-905.

Skinner, J.D. and Chimimba, C. T. (2005). *The Mammals of the Southern African Subregion*; pp 814. Cambridge Univeristy Press, Cambridge U.K.

Smit, A., Van der Bank, H., Falk, T. and De Castro, A. (2001). Biochemical genetic markers to identify two morphologically similar South African *Mastomys* species (Rodentia: Muridae). *Biochemical Systematics and Ecology*. **29**, 21-30.

Smith, M.F. and Patton, J.L. (1999). Phylogenetic relationships and the radiation of Sigmodontine rodents in South America: evidence from *cytochrome b. Journal of Mammalian Evolution*. **6**, 89-128.

Steppan, S.J., Adkins, R.A. and Anderson, J. (2004). Phylogeny and divergence date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Systematic Biology*. **53**, 533-553.

Song, M.S., Lee, E.H., Shin, J.-Y., Joo, E. and Choi,Y.-K. Genetic characterization of encephalomyocarditis virus isolated from aborted swine fetus in Korea. *Unpublished*.

Sutherland, K.J., Horner, G.W., Hunter, R., and Fife, B.H. (1977). An outbreak of viral encephalomyocarditis in pigs. *New Zealand Veterinary Journal*. **25**, 225.

Svitkin, Y. V., Imataka, H., Khaleghpour, K., Kahvejian, A., Liebig, H.D. and Sonenberg, N. (2001). Poly(A)-binding protein interaction with elF4G stimulates picornavirus IRES-dependent translation. *RNA.* **7**, 1743-1752.

Tamura, K. and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution.***10,** 512-526.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution.* **28,** 2731-2739.

Tan, B.-H.and Yap, E.P.H. Genetic Identification of an encephalomyocarditis virus isolated from Orang Utan. *Unpublished.*

Taylor, P.J. (2000). Patterns of chromosomal variation in southern African rodents. *Journal of Mammalogy*. **81**, 317-331.

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. In: Some mathematical questions in biology - DNA sequence analysis; pp 57-86. Miura, R.M. (editor). American Mathematical Society: Providence, RI, USA.

Tesh, R.B. and Wallace, G.B. (1978). Observations on the natural history of encephalomyocarditis virus. *American Journal of Tropical Medicine and Hygiene*. **27,** 133-143.

Thomson, G.R., Bengis, R.G., and Brown, C.C.(2001). Picornavirus Infections. In, Infectious diseases of wild mammals; pp. 119-130. 3rd edition. Williams, E.S. and Barker I.K. (Editors). Iowa: Iowa State University Press.

Tsangaris, T., Lekkas, S., Laspidis, C., Paschaleri-Papadopoulou, E. and Axiotis, I. (1989). Encephalomyocarditis of swine in Greece. Pathologoanatomical lesions of the disease. *American Veterinary Medical Association.* **40**, 222-237.

Urwin, R., Holmes, E.C., Fox, A.J., Derrick, J.P. and Maiden, M.C. (2002). Phylogenetic evidence for frequent positive selection and recombination in the meningococcal surface antigen PorB. *Molecular Biology and Evolution.***19,** 1686-1694.

Vanderhallen, H. and Koenen, F. (1997*a*). Rapid diagnosis of encephalomyocarditis virus infections in pigs using a reverse transcription-polymerase chain reaction. *Journal of Virological Methods*. **66**, 83-89.

Vanderhallen, H. and Koenen, F. (1997*b*). Identification of EMCV in clinical samples by RT-PCR followed by genetic typing using sequence analysis. *Proceedings of the 4th International Congress of Veterinary Virology.* Scotland: Edinburgh. 24-27 August, 201.

Vanderhallen, H., and Koenen, F., (1998). Identification of encephalomyocarditis virus in clinical samples by reverse transcription-PCR followed by genetic typing using sequence analysis. *Journal of Clinical Microbiology*. **36,** 3463-3467.

van Sandwyk, J.H.D.T., Bennett, N.C., Swanepoel,R. and Bastos A.D.S. (2012). Retrospective genetic characterisation of Encephalomyocarditis viruses from African elephant and swine recovers two distinct lineages in South Africa. *Veterinary Microbiology*. **162,** 23-31.

Vbra, E.S. (1985). Ecological and adaptive changes associated with early hominid evolution. In *Ancestors: The Hard Evidence*, pp 63-71. Delson E., (Editor). New York: John Wiley & Sons.

Venturi, F.P., Chimimba, C.T., Van Aarde, R.J. and Fairall, N. (2004). The distribution of two medically and agriculturally important cryptic rodent species, *Mastomys natalensis* and *M. coucha* (Rodentia: Muridae) in South Africa. *African Zoology*. **39,** 235-245.

Waggoner, S., and Sarnow, P. (1998). Viral ribonucleoprotein complex formationand nucleolar-cytoplasmic relocalization of nucleolin in poliovirus infected cells. *Journal of Virology*. **72,** 6699-6709.

Watts, C.H.S., Baverstock, P.R. (1995). Evolution in the Murinae (Rodentia) Assessed by Microcomplement Fixation of Albumin. *Australian Journal of Zoology*. **43**,105-118.

Webb, T., and Bartlein, P.J. (1992). Global changes during the last 3 million years: climatic controls and biotic responses. *Annual Review of Ecology and Systematics*. **23**, 141-173.

Wells, S.K., Gutter, A.E., Soike, K.F. and Baskin, G.B. (1989). Encephalomyocarditis virus: epizootic in a zooloogical collection. *Journal of Zoo and Wildlife Medicine*. **20**, 291-296.

White, T.J., Madej, R., and Persing, D.H. (1992). The polymerase chain reaction: clinical applications. *Advances in Clinical Chemistry*. **29**, 161–196.

Williams, M.C. (1981). Encephalomyocarditis virus infection. *Journal of The South African Veterinary Association.* **52**, 76.

Willis, K.J., Kleczkowski, A., and Crowhurst, S.J. (1999). 124 000-year periodicity in terrestrial vegetation change during the late Pliocene epoch. *Nature*. **397**, 685-688.

Wilson, D.E., and Reeder, D.M. (2005). Mammal species of the world. A taxonomic and geographic reference. Baltimore: Johns Hopkins University Press.

Woo, P.C., Lau, S.K., Choi, G.K., Huang, Y., Teng, J.L., Tsoi, H.W., Tse, H., Yeung, M.L., Chan, K.H., Jin, D.Y., and Yuen, K.Y. (2012). Natural occurrence and characterization of two internal ribosome entry site elements in a novel virus, canine picodicistrovirus, in the picornavirus-like superfamily. *Journal of Virology*. **86,** 2797-2808.

Yeo, D.S., Lian, J.E., Fernandez, C.J., Lin, Y.-N., Liaw, J. C.-W., Soh, M.-L., Lim, E., A.-S., Chan, K.-P., Ng, M.-L., Tan, H.-C., Oh, S., Ooi, E.-E., and Tan, B.-H. (2013). A highly divergent encephalomyocarditis virus isolated from non-human primates in Singapore. *Virology Journal*. **10,** 248.

Zhang, G., Guo, X., and Yang, H. (1997). Molecular characterization of encephalomyocarditis virus isolated from China. Unpublished.

Zhao, T., Zhang, J., Ge, X., Guo, X. and Yang, H. Analysis of Genomic Molecular Characterization of encephalomyocarditis virus Isolated from Different Hosts. *Unpublished.*

Zimmerman, J.J. (1994). Encephalomyocarditis. In *Handbook of Zoonoses, Section B. Viral;* pp. 423-436. G. W. Beran (Editors). Boca Raton, Florida: CRC Press.

APPENDICES

Appendix 1.1 Summary of known *Picornaviridae* genera and assigned species as per the ICTV, Virus Taxonomy: 2012 Release (http://ictvonline.org/virusTaxonomy.asp?version=2012). Type species in each genus indicated by a "*".

Genus	No of assigned species	Virus Name
Aphthovirus	4	*Bovine rhinitis A virus* *Bovine rhinitis B virus* *Equine rhinitis A virus* *Foot-and-mouth disease virus**
Aquamavirus	1	*Aquamavirus A**
Avihepatovirus	1	*Duck hepatitis A virus**
Cardiovirus	2	*Encephalomyocarditis virus** *Theilovirus*
Cosavirus	1	*Cosavirus A**
Dicipivirus	1	*Cadicivirus A**
Enterovirus	12	*Enterovirus A* *Enterovirus B* *Enterovirus C** *Enterovirus D* *Enterovirus E* *Enterovirus F* *Enterovirus G* *Enterovirus H* *Enterovirus J* *Rhinovirus A* *Rhinovirus B* *Rhinovirus C*
Erbovirus	1	*Equine rhinitis B virus**
Hepatovirus	1	*Hepatitis A virus**
Kobuvirus	3	*Aichivirus A** *Aichivirus B** *Aichivirus C**
Megrivirus	1	*Melegrivirus**
Parechovirus	2	*Human parechovirus** *Ljungan virus*
Salivirus	1	*Salivirus A**
Sapelovirus	3	*Avian sapelovirus* *Porcine sapelovirus** *Simian sapelovirus*
Senecavirus	1	*Seneca Valley virus**
Teschovirus	1	*Porcine teschovirus**
Tremovirus	1	*Avian encephalomyelitis virus**

Appendix 2.1 Summary of EMCV isolates used in phylogenetic analyses, detailing Genbank accession number, year of isolation, host species, country of origin and reference.

Genbank	Serotype	Isolate	Year of Isolation	Host species	County of Origin	Reference
AF356822	EMCV-1	BEL-2887A/91	1991	*Sus scrofa*	Belgium	Kassimi *et al.*, 2002
AF510055	EMCV-1	M105/1/02	2002	*Pongo*	Singapore	Yeo *et al.*, 2013
AJ235698	EMCV-1	Mengo	1946	*Macaca mulatta*	Uganda	Dick *et al.*, 1948
AJ235700	EMCV-1	BEL-2887A/90	1991	*Sus scrofa*	Belgium	Koenen *et al.*, 1999
AJ235701	EMCV-1	BEL-279/95	1995	*Sus scrofa*	Belgium	Koenen *et al.*, 1999
AJ235702	EMCV-1	BEL-357/95	1995	*Sus scrofa*	Belgium	Koenen *et al.*, 1999
AJ235703	EMCV-1	BEL-359/95	1995	*Rattus*	Belgium	Koenen *et al.*, 1999
AJ235704	EMCV-1	BEL-371/95	1995	*Sus scrofa*	Belgium	Koenen *et al.*, 1999
AJ235705	EMCV-1	BEL-475/95	1995	*Sus scrofa*	Belgium	Koenen *et al.*, 1999
AJ235706	EMCV-1	BEL-740/95	1995	*Sus scrofa*	Belgium	Koenen *et al.*, 1999
AJ235707	EMCV-1	BEL-810/95	1995	*Sus scrofa*	Belgium	Koenen *et al.*, 1999
AJ235708	EMCV-1	FRA-919/95	1995	*Sus scrofa*	France	Koenen *et al.*, 1999
AJ235709	EMCV-1	BEL-2062/96	1996	*Sus scrofa*	Belgium	Koenen *et al.*, 1999
AJ235710	EMCV-1	BEL-2105/96	1996	*Sus scrofa*	Belgium	Koenen *et al.*, 1999
AJ235711	EMCV-1	BEL-2106/96	1995	*Sus scrofa*	Belgium	Koenen *et al.*, 1999
AJ235712	EMCV-1	GRE-544/88	1988	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235713	EMCV-1	GRE-538/87	1987	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235714	EMCV-1	GRE-432/86	1986	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235715	EMCV-1	GRE-424/90	1990	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235716	EMCV-1	GRE-317/91	1991	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235717	EMCV-1	GRE-211/92	1992	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235718	EMCV-1	GRE-181/95	1995	*Rattus norvegicus*	Greece	Koenen *et al.*, 1999
AJ235719	EMCV-1	GRE-180/95	1995	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235720	EMCV-1	GRE-175/95	1995	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235721	EMCV-1	GRE-093/89	1989	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235722	EMCV-1	GRE-042/87	1987	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235723	EMCV-1	GRE-027/94	1994	*Sus scrofa*	Greece	Koenen *et al.*, 1999

Genbank	Serotype	Isolate	Year of Isolation	Host species	County of Origin	Reference
AJ235724	EMCV-1	GRE-021/92	1992	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235725	EMCV-1	GRE-003/92	1992	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235726	EMCV-1	GRE-001/97	1997	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235727	EMCV-1	GRE-001/89	1989	*Sus scrofa*	Greece	Koenen *et al.*, 1999
AJ235728	EMCV-1	CYP-138/96	1996	*Apodemus*	Cyprus	Koenen *et al.*, 1999
AJ235729	EMCV-1	CYP-129/96	1996	*Sus scrofa*	Cyprus	Koenen *et al.*, 1999
AJ235730	EMCV-1	CYP-108/95	1995	*Sus scrofa*	Cyprus	Koenen *et al.*, 1999
AJ235731	EMCV-1	CYP-003/94	1994	*Sus scrofa*	Cyprus	Koenen *et al.*, 1999
AJ235732	EMCV-1	ITL-002/88	1988	*Sus scrofa*	Italy	Koenen*et al.*, 1999
AJ235733	EMCV-1	ITL-135/86	1986	*Sus scrofa*	Italy	Koenen *et al.*, 1999
AJ235734	EMCV-1	ITL-001/96	1996	*Sus scrofa*	Italy	Koenen *et al.*, 1999
AJ617356	EMCV-1	ATCC VR-129B	1944	*Pan troglodytes*	USA	Denis *et al.*, 2006
AJ617357	EMCV-1	I001/96	1996	*Sus scrofa*	Italy	Denis *et al.*, 2006
AJ617358	EMCV-1	I136/86	1986	*Sus scrofa*	Italy	Denis *et al.*, 2006
AJ617359	EMCV-1	C108/95	1995	*Sus scrofa*	Cyprus	Denis *et al.*, 2006
AJ617360	EMCV-1	B440/95	1995	*Sus scrofa*	Belgium	Denis *et al.*, 2006
AJ617361	EMCV-1	B279/95	1995	*Sus scrofa*	Belgium	Denis *et al.*, 2006
AJ617362	EMCV-1	G424/90	1990	*Sus scrofa*	Belgium	Denis *et al.*, 2006
AY162279	EMCV-1	M100/1/02	2002	*Pongo*	Singapore	Yeo *et al.*, 2013
AY296731	EMCV-1	EMCV-30	2003	*Sus scrofa*	USA	LaRue *et al.*, 2003
DQ288856	EMCV-1	pEC9	--	*Murine*	USA	Martin *et al.*, 2000
DQ294633	EMCV-1	Mengo-Rz-pMwt	1946	*Macaca mulatta*	Uganda	Fata-Hartley & Palmenberg, 2005
DQ464062	EMCV-1	BJC3	--	*Sus scrofa*	China	Zhang *et al.*, 2007
DQ464063	EMCV-1	HB1	--	*Sus scrofa*	China	Zhang *et al.*, 2007
DQ517424	EMCV-1	EMCV-CBNU	--	*Sus scrofa*	Republic of Korea	Song *et al.*,unpublished
DQ835184	EMCV-1	B279/95	1995	*Sus scrofa*	Belgium	Hammoumi *et al.*, 2012
DQ835185	EMCV-1	1086C	--	*Rattus norvegicus*	Belgium	Guy *et al.*, 2009
EU780148	EMCV-1	K3	1990	*Sus scrofa*	Republic of Korea	An *et al.*, 2009

Genbank	Serotype	Isolate	Year of Isolation	Host species	County of Origin	Reference
EU780149	EMCV-1	K11	1990	*Sus scrofa*	Republic of Korea	An *et al.,* 2009
EU979543	EMCV-1	Peru-2004/10854	2004	*Homo sapiens*	Peru	Oberste *et al.,* 2009
EU979545	EMCV-1	Peru-2004/10854	2004	*Homo sapiens*	Peru	Oberste *et al.,* 2009
EU979546	EMCV-1	Peru-2004/10855	2004	*Homo sapiens*	Peru	Oberste *et al.,* 2009
EU979548	EMCV-1	Peru-2004/10855	2004	*Homo sapiens*	Peru	Oberste *et al.,* 2009
FJ604852	EMCV-1	GX0601	--	*Sus scrofa*	China	Zhao *et al.*,unpublished
FJ604853	EMCV-1	GX0602	--	*Sus scrofa*	China	Zhao *et al.*,unpublished
FJ897755	EMCV-1	GXLC	2008	*Sus scrofa*	China	Shi *et al.*,unpublished
GU181317	EMCV-1	SPU 64/03	2003	*Pan paniscus*	DRC#	Jones *et al.,* 2011
HM641897	EMCV-1	NJ08	2008	*Sus scrofa*	China	Bai *et al.,* 2011
L220891	EMCV-1	Mengo-M	1946	*Macaca mulatta*	Uganda	Palmenberg & Duke, 1993
L40427	EMCV-1	B	1958	*Sus scrofa*	Panama	Eun *et al.,* 1988
M20167	EMCV-1	D	1958	*Sus scrofa*	Panama	Eun *et al.,* 1988
M22457	EMCV-1	B	1958	*Sus scrofa*	Panama	Bae *et al.,* 1989
M22458	EMCV-1	D	1958	*Sus scrofa*	Panama	Bae *et al.,* 1989
M37588	EMCV-1	D	1958	*Sus scrofa*	Panama	Cohen *et al.,* 1988
M38329	EMCV-1	unnamed	--n	--	--	Grachev *et al.,* 1983
M54935	EMCV-1	unnamed	--	--	--	Petrov *et al.,* 1984
M81861	EMCV-1	R (Ruckert)		*Murine*	USA	Duke *et al.,* 1992
M88547	EMCV-1 (Mengo)	37A	1946	*Macaca mulatta*	Uganda	Mann *et al.,* 1992
X00463	EMCV-1	R (Ruckert)		*Murine*	USA	Palmenberg *et al.,* 1984
X67502	EMCV-1	NVSL-FL	1961	*Sus scrofa*	USA	Meng *et al.,* unpublished
X74312	EMCV-1	PV21	1958	*Sus scrofa*	Panama	Zimmerman *et al.,* 1994
X87335	EMCV-1	PV2	1995	*Mus musculus*	Germany	Nelsen-Salz *et al.,* 1996
Y15445	EMCV-1	Omsk-93	--	*Sus scrofa*	Russia	Amineva *et al.*,unpublished
Y15448	EMCV-1	1029	--	*Sus scrofa*	Russia	Amineva *et al.*,unpublished

#DRC: Democratic Republic of the Congo; -- Not known

Appendix 3.1 Results of the screening for EMCV infection and internal RNA control amplification for each specimen and organ type using the developed diagnostic PCR. Results are shown as number of samples which amplified the EMCV target/number of samples which amplified the internal RNA control.

Species	**Oct 2008 (Dry)**			**April 2009 (Wet)**			**Oct 2009 (Dry)**			**April 2010 (Wet)**		
	Heart	Kidney	GIT	Heart	Kidney	GIT	Heart	Kidney	GIT	Heart	Kidney	GIT
Mastomys natalensis	0/6	0/6	0/6	0/24	0/24	0/24	0/41	0/41	0/41	0/58	0/58	0/58
Mastomys coucha	0/5	0/5	0/5	0/14	0/14	0/14	0/31	0/31	0/31	0/8	0/8	0/8
Lemniscomys rosalia	0/4	0/4	0/4	0/1	0/1	0/1	0/10	0/10	0/10	0/3	0/3	0/3
Saccostomus campestris	0/1	0/1	0/1	0/11	0/11	0/11	0/6	0/6	0/6	0/23	0/23	0/23
Tatera leucogaster	0/1	0/1	0/1	0/8	0/8	0/8	0/4	0/4	0/4	0/7	0/7	0/7
Mastomys natalensis	0/6	0/6	0/6	0/24	0/24	0/24	0/41	0/41	0/41	0/58	0/58	0/58
Total	**0/17**	**0/17**	**0/17**	**0/58**	**0/58**	**0/58**	**0/92**	**0/92**	**0/92**	**0/99**	**0/99**	**0/99**