# Variations in the S segment of Rift Valley fever virus WITH SPECIAL REFERENCE TO THE NONSTRUCTURAL NSS CODING REGION 

By
Susan Claire Aitken

# DISSERTATION SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE 

Magister Scientiae (Virology)

# Department of Virology, School of Pathology Faculty of Health Sciences, University of the Witwatersrand 

Supervisor: Dr hab. Janusz T. Paweska<br>Special Pathogens Unit, National Institute for Communicable Diseases of the National Health<br>Laboratory Service

JuLY 2008

## Dedication

I dedicate this work to my late grandmother who was always so proud of every one of my achievements, but who sadly did not see the completion of this project. She was an incredible lady that lived every day of her life to the fullest up until the last. She is sadly missed but will live on forever in our hearts.

## Neva Aitken

12/07/1919 - 18/07/2007

## Acknowledgements

I would like to extend my gratitude to my supervisor Dr hab. Janusz T. Paweska and Mrs. Antoinette Grobbelaar for their unconditional support, advice and proficient insight during the course of my research. Their guidance through numerous challenges associated with certain aspects of my project have aided me in expanding my academic knowledge and propensity for troubleshooting, subsequently enabling me to achieve a thoroughly successful and enjoyable project.

Furthermore, I would like to thank the members of the National Institute for Communicable Diseases, Special Pathogens Unit (NICD-SPU) for their continuous technical support throughout this project. I also wish to express my sincere thanks to the NICD-SPU for financial support, providing laboratory equipment and instructions, and for supplying isolates of Rift Valley fever virus for this project.

I would also like to thank the CLS team at the Medical School of the University of the Witwatersrand and Mr. Peter Coetzee from the Polio Research Laboratory at the National Institute for Communicable Diseases for their assistance in sequencing.

## Declaration

I, Susan Claire Aitken (0200154E), am a student registered for the degree of Master of Science, in the year 2008. On submittal of my MSc thesis, I do hereby declare the following:

- I confirm that the MSc dissertation submitted for assessment for the above course is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the findings and ideas of others.
- I have not submitted the MSc work for any degree or examination at any other university.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the published findings or ideas in my writing.


#### Abstract

Rift Valley fever virus (RVFV) is a Phlebovirus member of the Bunyaviridae family and it is the causative agent of Rift Valley fever (RVF), a mosquito-borne viral zoonotic disease that poses a significant threat to domestic ruminants and human health in Africa. The RVFV is an encapsulated, negative-sense, singlestranded RNA virus with a tripartite segmented genome, containing L (large), M (medium) and S (small) segments. The S segment codes for two proteins, namely the nucleocapsid ( N ) protein and non-structural protein (NSs). There is evidence that the NSs protein is involved in virulence by blocking the expression of the interferon beta (IFN- $\beta$ ) promoter. It has been recently demonstrated that the SAP30-NSs-YY1 multiprotein complex represses the IFN- $\beta$ promoter. Consequently, the interferon expression is blocked, allowing virus to replicate.


A total of 45 isolates of RVFV recovered over a period of 53 years in 14 African countries, Madagascar and Saudi Arabia were characterized by full sequencing of the S segment of the virus. This data was added to another 27 strains of RVFV available on GenBank for phylogenetic analysis using MEGA4, giving a total of 72 strains analyzed. Alignments were made of the entire $S$ segment, the NSs gene, the N gene, and their deduced amino acid sequences. The laboratory strains, clone 13, MP12 and Smithburn, were also included in the alignments.

Two isolates were passaged ten times through two different amplification systems to asses the potential for sequence variation to occur in the original material through routine laboratory manipulations. Sequencing data was generated from the virus RNA present in the original clinical specimens and from the extracted RNA from the tenth passage of virus in each amplification system. The results showed $100 \%$ homology for each respective isolate, demonstrating that the RVFV S segment remained stable during ten serial passages in different propagation systems.

Phylogenetic analysis was conducted on the naturally occurring RVFV strains (n $=72$ ) and the findings indicate that circulating strains are compartmentalized and belong to one of three major lineages, namely Egyptian, western African, and central, eastern and southern African. The strains clustered in the Egyptian lineage had an average p-distance of $1.0 \%$, the western African strains $0.9 \%$, and the central, southern and eastern African strains 2.0\%. The overall average p-distance was $2.5 \%$, with a range from 0 to $4.1 \%$. For the N gene, the range was from 0 to $4.2 \%$, with an average of $2.2 \%$. For the N protein, the range was from 0 to $2 \%$, with an average of $0.2 \%$. The NSs gene had a range of 0 to $4.6 \%$, with an average of $2.4 \%$. The NSs protein had a range of 0 to $3.8 \%$, with an average of $1.7 \%$. The intergenic region (IGR) had a range of 0 to $9.2 \%$, with an average of $4.8 \%$.

Results of the study suggest that RVF outbreaks can result from either the rapid spread of a single strain over vast distances or from an increased activity of a strain circulating at an endemic level within an area/region during prolonged dry periods.

Sequencing alignment showed that the length of the $S$ segment ranged from 1690 to 1692 nucleotides. This difference in length was due to insertions and deletions found in the IGR, which is also the region with the most sequence divergence (4.8\%). Both the NSs and N genes had neither insertions nor deletions, and were both found to be stable, though the NSs gene was slightly more variable than the N gene ( $2.5 \%$ versus $2.2 \%$ )

The deduced amino acid sequences of the NSs protein were considerably more variable than that of the N protein ( $1.7 \%$ versus $0.2 \%$ ). Alignment of the NSs protein demonstrated that the 5 cysteine residues at positions 39, 40, 150, 179 and 195, are highly conserved among the isolates analyzed. These residues are important for conservation of the three-dimensional structure of the protein and the formation of filamentous structures observed in cells infected with natural strains of RVFV. The NSs protein is now implicated as the major factor of virulence and that its pathogenicity is associated with the blocking of interferon production. Therefore, any amino acid changes that result in changes to the filamentous structure of the NSs protein might impact on the binding kinetics between the NSs protein, SAP30
(Sin3A Associated Protein 30) and YY1 (Yin Yang-1). There were 6 amino acid changes in the NSs-SAP30 binding domain, with one being unique to the liveattenuated Smithburn vaccine strain.

Generated sequencing data contributes to global phylogenetic characterization of RVFV isolates and molecular epidemiology of the virus. In addition, findings of this study will further aid investigation on reassortment events occurring between strains of RVFV and genetically related viruses, the role of the NSs protein in the replicative cycle of the virus, the pathogenic effects of the NSs protein within the RVFV-infected host cells, and might help to identify molecular basis of RVFV virulence.

## Table of Contents

Dedication ..... ii
Acknowledgements ..... iii
Declaration ..... iv
Abstract ..... V
Table of contents ..... vii
List of figures ..... xi
List of tables ..... xii
Abbreviations ..... xiii
Chapter 1: Introduction ..... 1
1.1. Family Bunyaviridae ..... 1
1.1.1 Genus Phlebovirus ..... 1
1.1.1.1 The Rift Valley fever virus .....  2
1.2. Vectors and transmission .....  3
1.3. The outbreak history of Rift Valley fever ..... 5
1.4. The Rift Valley fever virus genome ..... 12
1.5. Virus replication and transcription ..... 14
1.5.1 Replication ..... 14
1.5.2 Transcription. ..... 16
1.6. The nucleocapsid gene and protein ..... 17
1.7. The non-structural gene and protein. ..... 17
1.8. Strains of Rift Valley fever virus ..... 20
1.9. Aims and objectives ..... 23
Chapter 2: Materials and Methods ..... 24
2.1. Culture of Vero cells ..... 24
2.2. Propagation of Rift Valley fever viral isolates in Vero cells ..... 25
2.2.1. Passaging of Rift Valley fever viral isolates in Vero cells ..... 26
2.2.2. Passaging of Rift Valley fever viral isolates in mice ..... 26
2.3. Extraction of RNA from infected tissue culture ..... 28
2.4. RT-PCR of extracted RNA using primers targeting the S RNA segment. ..... 28
2.4.1. Primer design ..... 28
2.4.2. Reverse transcription ..... 29
2.4.3. Polymerase chain reaction ..... 30
2.5. Automated sequencing of the $S$ segment RT-PCR product ..... 31
2.5.1. Primer design. ..... 31
2.5.2. Cycle sequencing. ..... 32
2.5.3. Sequencing ..... 32
2.6. Analysis of the sequencing data ..... 33
Chapter 3: Results ..... 34
3.1. Sequencing analysis of the Rift Valley fever virus $S$ segment ..... 34
3.1.1. 5' and $3^{\prime}$ non-coding regions...... ..... 36
3.1.2. The intergenic region ..... 37
3.1.3. The NSs gene and protein ..... 38
3.1.4. The N gene and protein. ..... 39
3.2. Phylogeny ..... 41
3.2.1. The S RNA segment phylogeny ..... 41
3.2.2. The N gene and protein phylogeny ..... 45
3.2.3. The NSs gene and protein phylogeny ..... 49
Chapter 4: Discussion ..... 54
Chapter 5: References ..... 59
Chapter 6: Appendix. ..... 72
6.1. Preparation of reagents ..... 72
6.1.1. EMEM ..... 72
6.1.2. PBS ..... 72
6.1.3. TBE buffer ..... 72
6.2. Routine laboratory methodologies ..... 73
6.2.1. RNA extraction using QIAamp kit. ..... 73
6.2.2. 1.2\% Agarose gel. ..... 74
6.2.3. PCR product clean-up using the Wizard ${ }^{\circledR}$ SV gel and PCR clean-up system ..... 74
6.2.4. Isopropanol purification of cycle sequencing product ..... 75
6.3. Full S RNA sequencing data ..... 76
6.4. Protein amino acid alignment ..... 119
6.4.1. N protein amino acid alignment ..... 119
6.4.2. NSs protein amino acid alignment ..... 125
6.5. Supplementary strain data ..... 131
6.6. Evolutionary relatedness of two passaged strains of Rift Valley fever virus ..... 132
6.7. Ethics clearance certificate ..... 133

## List of Figures

Figure 1. Virions of Rift Valley fever virus .....  3
Figure 2. Abortions in sheep caused by Rift Valley fever virus infection ..... 5
Figure 3. Haemorrhagic state in human, cause by Rift Valley fever virus infection ........................................................................................ 6 ..... 6
Figure 4. Settlement of nomadic pastoralist in Kenya .....  6
Figure 5. Flooded areas in Garissa district in North-eastern Provine of Kenya, an epicentre of Rift Valley fever outbreak in Kenya, 2006-2007 .....  9
Figure 6. A map showing the distribution of the 2007 Rift Valley fever outbreak in Kenya ..... 10
Figure 7. A diagrammatic cross-section representing a virus belonging to the Bunyaviridae family ..... 13
Figure 8. A schematic representation of the Rift Valley fever virus $S$ segment ambisense coding strategy ..... 15
Figure 9. The consensus sequence for the Rift Valley fever virus S segment in the viral sense ..... 35
Figure 10. The consensus amino acid sequence for the Rift Valley fever virus NSs protein ..... 38
Figure 11. The consensus amino acid sequence for the Rift Valley fever virus N protein ..... 39
Figure 12. Evolutionary relationship of S RNA segment of 72 strains of Rift Valley fever virus ..... 42
Figure 13. Evolutionary relationship of the N gene of 72 strains of Rift Valley fever virus ..... 46
Figure 14. Evolutionary relationship of the N protein of 72 strains of Rift Valley fever virus ..... 48
Figure 15. Evolutionary relationship of the NSs gene of 72 strains of Rift Valley fever virus ..... 50
Figure 16. Evolutionary relationship of the NSs protein of 72 strains of RiftValley fever virus................................................................ 52

Figure 17. The complete nucleotide sequence alignment of the $S$ segment of the 45 strains of Rift Valley fever virus sequenced in this study .76

# Figure 18. The deduced N protein amino acid alignment of the 45 strains of Rift Valley fever virus generated in this study <br> 119 

Figure 19. The deduced NSs protein amino acid alignment of the 45 strains of Rift Valley fever virus generated in this study 125

Figure 20. Evolutionary relationship of 2 passaged strains of Rift Valley fever virus

## List of Tables

Table 1. Identification, origin, year of isolation, source and accession numbers of 45 Rift Valley fever virus strains analyzed in the study27
Table 2. Primer sequences for the amplification of Rift Valley fever virus S RNA segment ..... 28

Table 3. Primer sequences for the sequencing of the Rift Valley fever virus $S$ segment .31

Table 4. Average p-distances within the 16 genotypes A-P............................... 41
Table 5. Supplementary strain data........................................................ 129

## Abbreviations

| ATF2 | Activating transcription factor 2 |
| :---: | :---: |
| BHK | Baby hamster kidney |
| BUNV | Bunyamwera virus |
| BSL | Biosafety level |
| CAR | Central African Republic |
| CDC | Centre for Disease Control |
| CPE | Cytopathic effect |
| DNA | Deoxyribonucleic acid |
| DNase | Deoxyribonuclease |
| dNTP | Deoxyribonucleotide triphosphate |
| dsRNA | Double-stranded RNA |
| DTT | 1,4-Dithiothreitol |
| EDTA | Ethylenediaminetetracetic acid |
| ELISA | Enzyme-linked immunosorbent assay |
| EM | Electron micrograph |
| EMEM | Eagles minimum essential medium |
| FCS | Foetal calf serum |
| G1 | Glycoprotein 1 |
| G2 | Glycoprotein 2 |
| HDAC-3 | Histone deacetylase 3 |
| IFN | Interferon |
| IgM | Immunoglobulin M |
| IGR | Intergenic region |
| IRF3 | Interferon regulatory factor 3 |
| LACV | La Crosse virus |
| L RNA | Large RNA segment |
| ME | Minimum evolution |
| MEGA | Molecular Evolutionary Genetics Analysis (software) |
| M-MuLV-RT | Moloney murine leukaemia virus reverse transcriptase |
| M RNA | Medium RNA segment |


| mRNA | Messenger RNA |
| :--- | :--- |
| NCoR | Nuclear receptor corepressor |
| NCR | Non-coding region |
| NICD | National Institute for Communicable Diseases |
| N protein | Nucleocapsid protein |
| NJ | Neighbourhood joining |
| NK-кB | Nuclear factor кB |
| NSm | Non-structural protein/gene of the M RNA segment |
| NSs | Non-structural protein/gene of the S RNA segment |
| PBS | Phosphate buffered saline |
| PTV | Runta Toro virus |
| RVF | Rift Valley fever |
| RVFV | Ribonucleic acid |
| RNA | RNA polymerase II |
| RNA pol II | Ribonuclease |
| RNase | Transition |
| RNP | Transversion |
| RT-PCR | Reverse transcription polymerase chain reaction |
| SAP30 | Sin3A associated protein 30 |
| SFSV | Sandfly fever Sicillian virus |
| SMB | Suckling mouse brain |
| SNS | Smithburn neurotropic strain |
| SPU | Special Pathogens Unit |
| S RNA | Small RNA segment |
| SSHV | Snowshoe Hare virus |
| Taq | Thermus aquaticus |
| TBE | Tris-borate-ethylenediaminetetracetic acid EDTA buffer |
| TS | RNA polymerase II transcription factor H |
| TF | TiIH |


| UUKV | Uukuniemi virus |
| :--- | :--- |
| UV-vis | Ultraviolet-visible |
| YY1 | Ying Yang |

## Chapter 1: Introduction

### 1.1. Family Bunyaviridae

The family Bunyaviridae is comprised of more than 300 viruses, which are categorized into five genera: Orthobunyavirus, Hantavirus, Nairovirus, Tospovirus and Phlebovirus (Nichol et al. 2005). Members of the Bunyaviridae family are enveloped and have a tripartite segmented, single-stranded RNA genome of negative or ambisense polarity (Brigden et al. 2001; Swanepoel and Coetzer 2004). Within the Bunyaviridae family there are viruses which are associated with severe human diseases, including Crimean-Congo haemorrhagic fever virus (Nairovirus), California encephalitis virus (Orthobunyavirus), Sin Nombre Hanta virus (Hantavirus) and Rift Valley fever virus (Phlebovirus) (Bishop and Shope 1979; Bi et al. 2008).

### 1.1.1. Genus Phlebovirus

The genus Phlebovirus is composed of two serogroups, the Sandfly fever viruses and the tick-transmitted uukuviruses (Murphy et al. 1995). The name Phlebovirus originates from the disease known as phlebotomous fever, which is caused by members of the Sandfly fever Sicilian (SFS) viruses transmitted by phlebotomous insects (Bishop and Shope 1979). Rift Valley fever virus (RVFV) (Figure 1), a member of the Sandfly fever group of Phleboviruses, is the causative agent of Rift Valley fever (RVF), an acute febrile disease predominantly affecting domestic ruminants in Africa and Madagascar, and is characterized by an increased mortality among newborn animals, abortion in pregnant animals, necrotic hepatitis and a hemorrhagic state (Swanepoel and Coetzer 2004).

Large outbreaks of the disease in sheep, cattle and goats occur at irregular intervals of several years when exceptionally heavy rains favour the breeding cycle of mosquitoes, the vectors of RVFV transmission. Transmission to humans usually occurs through contact with tissues and bodily fluids of infected and aborted animals, and less frequently through mosquito bites. Infected individuals often develop only mild febrile illness, but complications do occur in a small proportion of patients and
include ocular sequelae, encephalitis and fatal haemorrhagic disease (Swanepoel and Coetzer 2004; Ikegami et al. 2005b).

RVFV has a broad host range. The hosts most susceptible to disease are young domestic ruminants, including lambs, sheep, calves, some species of wild animals e.g. camels and buffalo, and various species of mice (Swanepoel and Coetzer 2004). Studies with different species of laboratory rodents have shown that the susceptibility to infection and the severity of clinical symptoms vary between the different species (Peters and Slone 1982; Anderson and Peters 1988; Ritter et al. 2000). There are a number of animals which can become infected with RVFV without developing clinical symptoms of the disease, including a wide range of wildlife species which might play a role in inter-epidemic maintenance of RVFV (Findlay 1931; Evans et al. 2007).

### 1.1.1.1. The Rift Valley fever virus

RVFV has the morphological and physiochemical properties typical of a member of the family Bunyaviridae. It is a negative-sense, single-stranded RNA virus with a tripartite segmented genome. The three segments are identified as large (L), medium $(\mathrm{M})$ and small (S). Each genome segment is contained within a separate nucleocapsid and is associated with a transcriptase enzyme. The genome segments are concealed within a spherical virus particle measuring 90 to 100 nm in diameter, which is surrounded by an envelope containing glycoprotein surface projections (Von Bonsdorff and Pettersson 1975; Bishop et al. 1980). The L, M and S RNA segments have molecular weights of 2.7, 1.7 and $0.6 \times 10^{6}$ respectively, which forms 1 to $2 \%$ of the weight of the virus particle (Rice et al. 1980).


Figure 1: Virions of Rift Valley fever virus. Picture courtesy of Mike Lecatsas and Monica Birkhead, National Institute for Communicable disease, Electron Micrograph Unit (NICD-EMU).

### 1.2. Vectors and transmission

The Aedes, Anopheles, Culex, Eretmapoites and Mansonia mosquitoes have been shown to be capable of transmitting RVFV under laboratory conditions (Hoogstraal et al. 1979; Zeller et al. 1997; Mellor and Leake 2000). RVFV can persist in nature for several years in infected mosquito eggs and then re-emerge following rainfall, which facilitates hatching of mosquito eggs (Le May et al. 2004). Environmental conditions that support massive proliferation of mosquitoes increase the possibility of a RVF outbreak to occur. It has been shown that the flooding of isolated grassland depressions, called dambos, plays a key role in the generation of epizootics of RVF in Kenya (Davies 1975). This is further supported by the finding that floodwaterbreeding aedine mosquitoes, which breed in dambos, are capable of transmitting RVFV transovarially (Davies 1975; Linthicum et al. 1985). During winter the eggs of aedine mosquitoes lie dormant, along with the RVFV which was transovarially passed on from the parent mosquito. These eggs are capable of surviving prolonged time
periods in dried mud. When the dambos flood during heavy rainfalls, transovarially infected Aedes mosquitoes hatch and the adult mosquitoes transmit the virus to livestock, where the virus amplifies to high titres and serves as a source of infection for other culicines and anopheline mosquitoes, which act as epizootic vectors. Biting flies such as midges, phlebotomids, stomoxids and simulids could possibly serve as mechanical transmitters of infection (Swanepoel and Coetzer 2004). Outbreaks occurring in arid countries in north and west Africa are independent of rainfall, and are dependent on mosquito vectors which breed in dams and rivers, but not the floodwater Aedes mosquitoes (Swanepoel and Coetzer 2004).

Activities that involve contact with infected animals are associated with an increased risk of human infection and include herding, milking, slaughtering, skinning, and the sheltering of animals within the home (Jouan et al. 1989; Woods et al. 2002). These activities are not limited to the rural setting, but also affect people in professional environments such as abattoir workers and veterinarians. Infected meat sold wholesale is unlikely to be a source of infection for humans as it has been shown that the virus cannot survive below pH 6.00 ; and meat generally falls to pH 6.00 or less during processing. It has, however, been postulated that unpasteurized milk from an infected animal might serve as a source of infection (Swanepoel and Coetzer 2004). Furthermore, RVFV is present in the nasopharyngeal secretions of infected humans (Abdel-Wahab et al. 1978) and can be transmitted by the respiratory tract under experimental conditions (Francis and Magill 1935; Miller et al. 1963; Brown et al. 1981).

### 1.3. The outbreak history of Rift Valley fever

The initial sign of an impending outbreak is observed as an increase in the incidences of mortality and abortions in domestic and wild ruminants such as sheep, goats, cattle, buffalo and camels (Figure 2). The demise of domestic animals is generally followed by increased reports of acute febrile illness and sporadically fatal haemorrhagic-like disease in humans (Figure 3) (Meegan 1979). Exact sources of outbreaks in humans are often difficult to establish, whilst some infections are not even reported. This is often attributed to the nomadic lifestyles (Figure 4) of infected individuals in some African countries (Jouan et al. 1989), cultural factors which make collection of specimens for laboratory confirmation difficult, as well as limited access to medical facilities (Siam et al. 1980).


Figure 2. Abortions in sheep caused by Rift Valley fever virus infection. Picture courtesy of Dr JT Paweska NICD-SPU.


Figure 3. Haemorrhagic state in a human, caused by Rift Valley fever virus infection. A Rift Valley fever patient with petechial rash and ecchymoses. Picture courtesy of Dr JT Paweska NICD-SPU.


Figure 4. Settlement of nomadic pastoralist in Kenya. Rift Valley fever is often diagnosed in pastoralists who have handled infected animals. Picture courtesy of Dr JT Paweska NICD-SPU.

The disease now known as Rift Valley fever was first recognized in sheep in the Rift Valley in Kenya at the turn of the $19^{\text {th }}$ century, but the causative agent was only isolated in 1930 following an outbreak of "enzootic hepatitis" on a sheep farm near Naivisha in the Rift Valley region in Kenya. The outbreak was severe and within seven weeks 3,500 lambs and 1,200 ewes were reported to have died of acute necrosis of the liver (Daubney et al. 1931). Over the next four decades epizootics of the disease were recorded only in eastern and southern Africa. These epizootics tended to occur in association with population-explosions of floodwater-breeding aedine mosquitoes following heavy rains (Swanepoel and Coetzer 2004).

Large outbreaks affecting livestock occurred in Kenya in 1930-31, 1968 and 197879, and lesser outbreaks at irregular intervals in the intervening years. A major epizootic, which caused an estimated 500,000 abortions and 100,000 deaths of sheep, occurred in South Africa in 1950-51; a second major, and more widespread outbreak caused extensive losses of sheep and cattle in 1974-76, while lesser outbreaks were recorded in 1952-53, 1955-59, 1969-71 and 1981 (Swanepoel and Coetzer 2004). Severe outbreaks occurred predominantly in the sheep farming areas of southern Namibia during 1955 and 1974-76. Extensive outbreaks of the disease in southern Africa occurred in areas dominated by cattle farming, in Zimbabwe during 1955, 1957, 1969-70 and 1978, in Mozambique during 1969, and in Zambia during 1973-74, 1978 and 1985. In addition, evidence of the occurrence of the infection was recorded in many other southern and east African countries (Swanepoel and Coetzer 2004).

It was noted from the original investigations in Kenya that febrile illness in humans accompanied outbreaks of disease in livestock, and that some patients experienced transient loss of visual acuity, but the occurrence of serious ocular sequelae was first recognized during the 1950-51 epizootic in South Africa (Weiss 1957). The potential lethality of the virus for man was recognized after seven deaths from encephalitis and/or haemorrhagic fever with necrotic hepatitis occurred during the 1974-76 RVF epizootic in South Africa (van Velden et al. 1977).

During 1973 and 1976, outbreaks of RVF affecting livestock were reported in Sudan (Swanepoel and Coetzer 2004). These epizootics were followed in 1977-78 by a major outbreak which occurred along the Nile delta and valley in Egypt, causing an unprecedented number of human infections and deaths, as well as numerous deaths and abortions in livestock, water buffalo and camels. Estimates of the number of human infections ranged from 18,000 to more than 200,000 , with at least 598 mortalities resulting from encephalitis and/or haemorrhagic fever (Meegan 1979). Compared to other outbreaks, the Egyptian epidemic was characterized by an increase in the occurrence of encephalitic, ocular and fatal haemorrhagic-like diseases (Meegan et al. 1979). The subsequent Egyptian outbreak occurred during 1993, and resulted in an estimated 600 to 1,500 human infections occurring within the Aswan governate (CDC 1994). During this outbreak an increase in ocular disease was noted (Arthur et al. 1993).

Prior to the 1970s, the presence of the virus circulation in the west African countries was known from serosurveillance studies. In 1987 a severe epizootic was reported in the Senegal River basin of southern Mauritania and northern Senegal. In Mauritania alone an estimated 224 human patients died of the disease, and there was a high rate of abortion in sheep and goats (Digoutte and Peters 1989; Jouan et al. 1989; Karlen 1996). The virus has remained endemic to the region and was responsible for a substantial outbreak in 1998, as well as several smaller outbreaks (Zeller et al. 1997; Nabeth et al. 2001).

From October 1997 to February 1998, a large outbreak of RVF occurred in northeastern Kenya and Somalia; outbreaks of the disease also occurred in Tanzania (CDC 1998; Woods et al. 2002). There were heavy losses of livestock and an estimated 500 human deaths. This particular outbreak was one of the largest in eastern Africa following the discovery of the virus in 1930, with approximately 89,000 infected individuals and 250 deaths (Porter 1999). Several ruthless outbreaks were recognised again in East Africa following heavy rainfall and flooding in 2006-07 (Figure 5), first in the North Eastern Province of Kenya in November 2006, and by the end of January 2007 in the coastal, central, Rift Valley and eastern Provinces (Figure 6). A total of 684 human cases of the disease were reported, with a recorded death rate of $20 \%$. In Somalia, Tanzania and Sudan the disease resulted in a total of 756 cases and a reported
$34 \%$ death rate. The high death rate recorded in the recent outbreaks in East Africa were, however, estimated only from a number of cases which were diagnosed mainly based on clinical symptoms (Anonymous 2007a; Anonymous 2007b). Small outbreaks of RVF were recognized in north-eastern South Africa during 1999 and 2008 (Paweska et al. 2008), and in central Madagascar and the Comoros Islands during 2008 (Eloit 2008; Rakotosamimanan 2008).


Figure 5. Flooded areas in Garissa district in North-eastern Province of Kenya, an epicentre of RVF outbreak in Kenya, 2006-2007. Picture courtesy of Dr JT Paweska NICD-SPU.


* As of January 25, 2007.

Figure 6. A map showing the distribution of the 2007 Rift Valley fever outbreak in Kenya. Included are the number and percentage of reported RVF cases $(\mathrm{n}=404)$ by district which occurred between November 2006 and January 2007 (CDC 2007).

The first documentation of RVF outside the African mainland occurred in Madagascar during 1979, where the virus was isolated from mosquitoes captured in the primary rainforest of Perinet (Mathiot et al. 1984). The virus was found to have genetic similarity to an isolated virus of Egyptian lineage. Although there was no outbreak of RVF in humans, antibodies to the virus were found in some of the Madagascan inhabitants (Saluzzo et al. 1989). The virus re-emerged in Madagascar during 1991 when the first documented infection of human and livestock occurred. Genetic typing showed that the 1991 isolate differed from the 1979 virus strain and that it was not related to the Egyptian lineage, indicating the presence of at least two separate strains of RVFV in Madagascar (Morvan et al. 1991; Morvan et al. 1992) .

In September 2000, RVF broke out simultaneously in south-west Saudi Arabia and adjoining Yemen following heavy rains on the inland mountain range (Jupp et al. 2002; Al-Hazmi et al. 2003; Madani et al. 2003; Abdo-Salem et al. 2006). This was the first known occurrence of a RVF outbreak outside of the African region. The outbreaks lasted until early 2001, and resulted in 245 human deaths and the loss of thousands of sheep and goats. There were approximately 882 infected human cases, with an unusually high mortality rate of $14 \%$, though this was probably due to underreporting of mild and sub-clinical cases. There was speculation that the virus may have been imported from the Horn of Africa with infected animals, possibly during the 1997-98 epidemic in East Africa, as the strain of the virus responsible for that outbreak had a high sequence similarity with the virus isolated from the Saudi Arabian outbreak. Consequently, importation of livestock from the Horn of Africa was banned, which caused substantial economic losses for east African countries, where the main source of income is generated from trading of livestock. (CDC 2000; Bird et al. 2007). Subsequent detection of the immunoglobulin M (IgM) antibody in sentinel sheep suggests that the virus may have become endemic to the Arabian Peninsula (Elfadil et al. 2006). The outbreak highlighted the major socio-economic risks associated with the transmission of RVFV outside its traditional endemic areas and the possibility of a further inter-continental spread (Balkhy and Memish 2003). The prospect of a global threat is further exacerbated by the capability of RVFV to utilise a wide range of mosquito vectors (Hoogstraal et al. 1979; Meegan 1979; Wilson et al. 1994; Turell et al. 2008), and current trends in global climate change which facilitate the spread of vector-borne viruses (Purse et al. 2005).

### 1.4. The Rift Valley fever virus genome

The L RNA segment is 6404 nucleotides (nt) in length and consists of a single open reading frame which codes for a 243.6 kDa viral polymerase called the L protein (Figure 7). The L protein gene exhibits some sequence homologies to other viral RNA polymerases (Muller et al. 1991).

The M segment is 3885 nt in length and codes for the two envelope glycoproteins G1 and G2 (Figure 7), as well as for two proteins of unknown function, a $14-\mathrm{kDa}$ nonstructural protein (NSm) and a $78-\mathrm{kDa}$ protein (Gentsch and Bishop 1979; Collett 1986). The G1 and G2 glycoproteins are responsible for eliciting and interacting with neutralising antibodies (Rozhon et al. 1981), and are the major determinants of virus virulence (Shope et al. 1981). The L and M segments are of negative polarity (Collett 1986; Muller et al. 1991).

The $S$ segment of RVFV is 1690 nt in length and has two open reading frames encoding the non-structural (NSs) and nucleocapsid (N) proteins. In the majority of the members of the Bunyaviridae, the genes encoding these two proteins overlap and are translated in different reading frames (Akashi and Bishop 1983). The phleboviruses are unique in that they have an ambisense S RNA segment and translate the N and NSs genes in opposite polarities. The N protein is translated from the $3^{\prime}$ half of the negative sense RNA and the NSs protein from the 3 ' half of the positive sense RNA (Ihara et al. 1984; Marriot et al. 1989; Billecocq et al. 2004; Ikegami et al. 2005a). A schematic representation of the coding strategies of the RVFV S RNA segment is shown in Figure 8. Both the N and NSs genes are found in all Bunyaviridae S RNA segments, with exception of some Hantaviruses, which lack the NSs gene (Plyusnin 2002; Jaaskelainen et al. 2007).


Figure 7. A diagrammatic cross-section representing a virus belonging to the Bunyaviridae family. Adapted from Whitehouse 2004 .

### 1.5. Virus replication and transcription

### 1.5.1. Replication

Viruses belonging to the Bunyaviridae family have negative-sense RNA genomes. Each segment of the RVFV genome is closely associated with nucleoproteins and RNA polymerase to form ribonucleoproteins (RNPs) (Figure 7). The viral RNA is not functional unless in the RNP form, which acts as the template for transcription and replication (Emerson and Wagner 1972).

Bunyaviridae viruses tend to replicate within the host cell cytoplasm and mature by budding into smooth-surface vesicles in or near the Golgi region (Murphy et al. 1973; Swanepoel and Blackburn 1977; Kuismanen et al. 1982; Smith and Pifat 1982). These mature virions are then transported within these vesicles to the plasma membrane where, the membranes fuse and the mature virions are released (Smith and Pifat 1982).


Figure 8. A schematic representation of the Rift Valley fever virus $S$ segment ambisense coding strategy. Adapted from Simons et al. (1990).

### 1.5.2. Transcription

The encapsidated negative-sense RNA is transcribed by the RNA polymerase into a positive sense messenger RNA (mRNA), followed by the translation of the primary transcripts, then replication of the genome, and finally the translation of secondary transcripts (Dimmock et al. 2001). As the S RNA segment is ambisense, the NSs gene is translated from the full length, viral complementary RNA (vcRNA) (Figure 8). Translation is primed by a 'cap-snatching' mechanism similar to that used by the influenza virus (Bouloy et al. 1978; Krug 1981). Briefly, the 5 ' region of the mRNA is capped with non-viral sequences that act as primers (Bishop et al. 1983; Patterson et al. 1984; Jin and Elliott 1993). The "cap-snatching" mechanism appears to be well conserved amongst members of the Bunyaviridae family and the presence of these 5'caps have been identified on the mRNAs of bunyaviruses (Bishop et al. 1983; Patterson et al. 1984; Eshita et al. 1985; Bouloy et al. 1990; Jin and Elliott 1993), phleboviruses (Ihara et al. 1985; Collett 1986; Simons and Pettersson 1991), nairoviruses (Jin and Elliott 1993), and tospoviruses (Kormelink et al. 1992). The 3' end of the mRNA is not polyadenylated and is shorter than the template because the viral RNA polymerase terminates before the $5^{\text {' }}$ end of the template (Elliott 1996; Schmaljohn and Hooper 2001). Within the Bunyaviridae family the only virus which possesses a polyadenylated 3 '-end is the hantavirus Sin Nombre, which has a polyadenylated M segment mRNA (Hutchinson et al. 1996).

It has been shown that transcription of the N and NSs mRNAs is terminated in the intergenic region (IGR) between the two genes (Schmaljohn and Hooper 2001). Albarino et al. (2007) have recently shown that the exact termini of the N, NSs and M mRNAs of RVFV, Toscana virus (TOSV) and SFSV are found immediately upstream of the conserved sequence $3^{\prime} \mathrm{C}_{1-3}$ GUCG/A-5'. The same pentanucleotide sequence has also been found within the ten nucleotides of the transcriptional termination sites of the N mRNA of La Crosse virus (LACV), Snowshoe Hare virus (SSHV) and Germiston virus, as well as the M mRNA of Germiston virus (Ikegami et al. 2007). However, this pentanucleotide sequence is not found in Punto Toro (PT) and Uukuniemi (UUK) viruses (Giorgi et al. 1991).

### 1.6. The nucleocapsid gene and protein

A mutational analysis has mapped the interacting domain of the N protein to the N terminal 71 residues. It has been predicted that the N-terminal region contains alphahelices and that the conserved amino acid residues at the N -terminus, tyrosine (Y) 4, phenylalanine (F) 11, aspartic acid (D) 17, and tryptophan (W) 24, are important in the formation of homodimers. Base changes affecting these amino acids could ultimately affect the function of the N protein (Le May et al. 2005).

Studies using mini genomes and other systems revealed that the co-expressions of N and L proteins were required for the synthesis of RNA in viruses belonging to the Bunyaviridae family (Dunn et al. 1995; Lopez et al. 1995; Accardi et al. 2001; Flick and Pettersson 2001; Blakqori et al. 2003; Flick et al. 2003).

### 1.7. The non-structural gene and protein

The NSs gene codes for a nonstructural protein (NSs protein) of 265 amino acids that is abundantly expressed in infected cells. Viral NSs protein is produced shortly after infection (Ikegami et al. 2005b) and can be detected in the nucleus as early as 3-5 hours post infection (Le May et al. 2004). In RVFV, the NSs protein is found in the nucleus of infected cells, where it forms large filamentous structures composed of thick bundles of fibrils, 50 nm in diameter, that lie parallel to one another (Daubney et al. 1931; Swanepoel and Blackburn 1977; Yadani et al. 1999; Bouloy et al. 2001), whereas in UUKV, the NSs protein has been found to be dispersed throughout the cytoplasm and associated with 40S ribosomal unit (Simons et al. 1992).

The NSs protein is the most variable protein among phleboviruses (Giorgi et al. 1991), and therefore it has been used as a marker of variability among natural strains of RVFV (Sall et al. 1997). The NSs protein is phosphorylated by casein kinase II at two serine residues located in the C-terminus, at positions 252 and 256 (Yadani et al. 1999; Billecocq et al. 2004). Unlike RVFV, there is no evidence that the NSs of UUKV is phosphorylated. (Simons et al. 1992). The acidic amino acids representing the C-terminal domain of the NSs protein are required for self-association and are
essential for filament formation, but not for NSs transport into the nucleus (Yadani et al. 1999).

The exact function of the NSs protein is not fully understood. The NSs protein is probably not essential for replication, as clone 13, an isolate of RVFV that lacks the NSs gene, is still capable of infecting cells and replicating (Muller et al. 1995). However, it has been shown that co-expression of NSs protein with L and N proteins substantially enhanced minigenome replication and transcription, suggesting that RVFV NSs protein plays some role in RVFV RNA synthesis (Ikegami et al. 2005a).

Recent studies have shown that the NSs gene of RVFV is a major virulence factor preventing early induction of interferon- $\beta$ (IFN- $\beta$ ) during the course of infection (Goodbourn et al. 2000; Bouloy et al. 2001; Billecocq et al. 2004). The IFNs are a large family of multifunctional proteins that function as the first line of innate defence against viral infection (Goodbourn et al. 2000; Billecocq et al. 2004). RVFV has a negative-sense RNA genome and thus it has to produce a double-stranded RNA (dsRNA) during its life cycle in order to replicate. The dsRNA is a strong inducer of IFN- $\alpha$ and IFN- $\beta$, which in turn activate an innate immune response (Goodbourn et al. 2000; Billecocq et al. 2004). Hence, for RVFV to efficiently infect cells, it needs to prevent activation of an immune response. This is achieved by the shut-down of host protein production through inhibition of host's mRNA synthesis. Inhibition is accomplished by the binding of the NSs protein to the RNA polymerase II transcription factor, TFIIH (Billecocq et al. 2004; Ikegami and Makino 2004). TFIIH is a multisubunit protein complex which is important in the transcription of proteincoding genes, ribosomal RNA synthesis in vivo and in vitro, as well as nucleotide excision repair (NER) (Iben et al. 2002).

The NSs protein interacts with the p44 and XPB subunits of the host TFIIH transcription factor. By sequestering these subunits within the NSs protein filaments, it prevents proper assembly of these subunits into a mature complex. This results in a lowered concentration of cellular TFIIH and consequently decreases the level of phosphorylated RNA polymerase II (RNA Pol II), which in turn lowers the rate of RNA synthesis (Le May et al. 2004). This finding has led to the conclusion that the virulence of a given RVFV strain is not dictated by its IFN sensitivity but resides in
its capacity to efficiently block the production of IFN- $\alpha$ and IFN- $\beta$ (Bouloy et al. 2001). Furthermore, by inhibiting host transcription, RVFV increases virus propagation by making free host ribonucleotides available for viral RNA replication (Dasgupta 2004). These studies have shown that NSs inhibits IFN- $\beta$ expression immediately post infection, though this has been shown to occur without inhibiting IFN- $\beta$-specific transcription factors such as IRF3 (interferon regulatory factor 3), NF$\kappa \mathrm{B}$ (nuclear factor- $\kappa \mathrm{B}$ ) and ATF2 (activating transcription factor 2) (Billecocq et al. 2004).

The mechanism by which the RVFV NSs protein antagonizes IFN- $\beta$ gene expression has recently been demonstrated by Le May et al. (2008). These authors demonstrated that SAP30 (Sin3A Associated Protein 30), a subunit of several corepressor complexes associated to NCoR (nuclear receptor corepressor) and/or Sin3A, binds to NSs as well as YY1 (Yin Yang-1), the activator/repressor of interferon transcription. After RVFV infection, a multiprotein complex containing viral NSs protein and host factors YY1/SAP30/NCoR/Sin3A/HDAC-3 (histone deacetylase 3) is recruited on the IFN- $\beta$ promoter. The binding of NSs stabilizes the multiprotein complex, maintaining the promoter in a silent repressed state. Consequently, IFN- $\beta$ expression is blocked, allowing the virus to invade the host. The importance of the NSs-SAP30 interaction was confirmed by showing that its disruption negatively affects the interaction of NSs with the IFN- $\beta$ promoter, and diminishes the ability of NSs to inhibit IFN- $\beta$ expression, and hence exert its pathogenic effect. The interacting domain in NSs was found to be in the C-terminal region, between 2 proline ( P ) residues from amino acids 210 to 230. The authors suggest that all genes whose promoters interact with SAP30 and/or YY1 could theoretically be a target for NSs/SAP30-dependant abnormal transcriptional regulation, which could possibly explain some of the pathogenic effects due to virus such as abortion, hemorrhagic fever, hepatitis or encephalitis (Le May et al. 2008).

The function of the RVFV NSs protein is similar in other Bunyaviridae viruses. The PTV NSs protein has also been shown to function as a type I IFN antagonist by inhibiting IFN- $\alpha$ and IFN- $\beta$ in vivo and in vitro (Perrone et al. 2007). The results of a study by Weber et al. (2002) suggest that the NSs of Bunyamwera virus (BUNV) is an IFN induction antagonist that blocks the transcriptional activation
of IFN- $\alpha$ and IFN- $\beta$, which results in increased virulence of BUNV (Brigden et al. 2001; Weber et al. 2002). The NSs protein of BUNV shuts down host transcription by inhibiting phosphorylation of serine (S) residues within the C-terminal domain of the cellular RNA pol II (Thomas et al. 2004). In contrast to RVFV, the NSs protein of the Tomato spotted wilt virus (TSWV), genus Tospovirus, has been found to suppress post-transcriptional gene silencing (Takeda et al. 2002), and more recently Soldan et al. (2005) have described a similar role for the NSs protein in the LACV virus.

Sall et al. (1997) conducted a study of the variability of the RVFV NSs protein and included 18 isolates from different geographic regions in Africa recovered over a period of 38 years from a variety of host species, including arthropods, cattle and humans. Included in this study were the clone 13 isolate, Smithburn neurotropic (SNS) and MP12 strains. The results showed that there were a variety of nucleotide substitutions throughout the entire NSs gene, but no base insertions or deletions were found, except in clone 13. Pairwise comparison among these strains showed percentage divergence ranging from $0 \%$ to $9.6 \%$ at the nucleotide level which corresponded to $0 \%$ to $9.5 \%$ at the amino acid level. Phylogenetic analysis based on the NSs gene revealed existence of two major lineages, Egyptian and sub-Saharan. Amino acid sequence alignment showed that the five cysteine (C) residues at positions $39,40,150,179$ and 195 are conserved.

### 1.8. Strains of Rift Valley fever virus

The members of the Bunyaviridae family have been found to be genetically stable (Bishop and Shope 1979; Bilsel et al. 1988), but point mutations do occur and are most likely due to the high error rate of RNA polymerase (Dimmock et al. 2001). In the case of RVFV, point mutations, deletions, and the reassortment between two or more strains have been shown to occur which resulted in the subsequent generation of new strains (Gentsch and Bishop 1976; Turell et al. 1990; Sall et al. 1999).

A virus isolated in Lunyo, Uganda, was found to be symptomatically similar to RVFV, but produced different histological lesions in mice compared to other virus strains. Attempts to neutralize the virus in infected mice with potent immune sera to RVFV were unsuccessful, suggesting that Lunyo virus might be a variant of RVFV (Weinbren et al. 1957).

Zinga virus, first isolated in the Central African Republic (CAR) from Mansonia africana in 1969 (Digoutte et al. 1974a), has subsequently been classified as a strain of RVFV with a similar pathogenicity for laboratory animals and humans when compared to alternative RVFV strains (Meegan et al. 1983). This strain has since been isolated from mosquitoes in the CAR, Senegal, and Madagascar and from naturally infected humans in Senegal and the CAR (Digoutte et al. 1974a; Digoutte et al. 1974b; Digoutte 1981; Georges et al. 1983).

Two isolates were selected from the 1977-78 outbreak in Egypt as prototypes of RVFV, namely the Zagazig 548 (ZH548) strain which was isolated from a 52 year old male patient with an acute febrile illness, and the Zagazig 501 (ZH501) strain which was isolated from a 12 year old Zagazig female who died of a haemorrhagic-like illness. The ZH501 strain has been shown to be significantly more pathogenic for certain laboratory animals than previous isolates of RVFV from Sub-Saharan Africa (Meegan 1979).

The MP12 strain of RVFV was obtained by twelve serial passages in mice of the virulent ZH548 strain in the presence of 5-fluorouracil (Caplen et al. 1985). Virulence has been shown to be under polygenic control and, in the case of the MP12 genome, at least one mutation capable of independently attenuating the virus exists on each segment of the genome. Multiple attenuating mutations significantly reduce the possibility of reversion to virulence (Saluzzo and Smith 1990). The MP12 strain is temperature sensitive, attenuated for mice, and compared to its parent ZH548 strain, forms small plaques. When the genome of the MP12 isolate was compared to ZH548, no mutations were noted in the N protein gene or the $3^{\prime}$ and $5^{\prime}$ non-coding regions (NCRs), however there were three mutations in the NSs gene and one in the IGR of the MP12 strain. Since there is only one amino acid change, occurring at position 513 in the NSs gene, it is possible that this is an important amino acid responsible for
attenuation of RVFV. The MP12 strain has been successfully tested in human volunteers as a potential vaccine (Peters and Linthicum 1994).

The Smithburn neurotropic strain (SNS) of RVFV, isolated from mosquitoes in Uganda during 1944 and passaged intracerebrally in mice, was subjected to further passaging in embryonated chicken eggs and mice in South Africa. The virus was issued in the form of freeze-dried infected mouse brain for use as a partiallyattenuated vaccine for livestock from 1951 onwards. In 1958, reversion was made to the use of a lower mouse passage level of the virus, and since 1971 the virus has been grown in cell cultures for the preparation of freeze-dried vaccine, recommended particularly for use in non-pregnant sheep, as the virus retains some abortogenic and teratogenic properties for a proportion of pregnant ewes (Coetzer and Barnard 1977; Swanepoel and Coetzer 2004). The use of live-attenuated vaccine is restricted to countries where RVFV is endemic and only inactivated virus is recommended for use in RVFV-free regions (Assaad et al. 1983). Immunization of livestock is an important measure of controlling outbreaks of RVF. Annual vaccination of domestic ruminants in endemic countries not only protects these animals against clinical disease, but also prevents the virus from spreading into the human population by lowering the level of virus amplification in animal hosts.

The 74H39 strain of RVF isolated from a human in the CAR was shown to be composed of a heterogeneous population of viruses. When plaque-purified, the clone 13 isolate was shown to be naturally attenuated by having a large deletion in the nonstructural (NSs) gene of the S segment (Muller et al. 1995; Vialat et al. 2000). Clone 13 is a strong IFN inducer and only capable of growing in mice with a compromised type 1 IFN system (Weber and Elliot 2002). It has been shown that the truncated NSs protein of clone 13 is expressed and remains in the cytoplasm, where it is degraded rapidly by the proteasome (Vialat et al. 2000). When the S RNA of clone 13 was compared to the ZH548 strain, besides the large deletion in the NSs gene, only one sequence variation in the N protein gene and 6 nucleotide changes in the intergenic region were found. There were no sequence variations within the $5^{\prime}$ and 3 ' NCRs (Muller et al. 1995; Vialat et al. 1997).

### 1.9. Aims and objectives

The present study was undertaken to extend the currently available full sequencing data on the S RNA segment of RVFV. Generating full sequencing data on the S segment of RVFV isolates from disparate historic and geographic origins will further contribute to phylogenetic characterization and molecular epidemiology of the virus, including tracking its movement, identification of the sources of RVF outbreaks, and aid investigations of reassortment events.

In addition, a passage series was performed in commonly used in vitro and in vivo propagation systems to determine the stability of the viral genome during routine laboratory procedures.

A total of 45 isolates of RVFV recovered throughout Africa, Madagascar, and Saudi Arabia over the past 53 years were used. This study undertakes a comprehensive analysis of the NSs gene at the nucleotide and amino acid levels of these isolates. Analysis of the genetic variability within the NSs genes and their products might help in identifying genetic markers that play a role in RVFV virulence.

Specific aims of this project:

- Produce stocks of RVFV isolates available from NICD-SPU by a single passage in Vero cells, for sequencing studies.
- Conduct ten serial passages of two selected RVFV isolates in commonly used in vitro and in vivo isolation/propagation systems to determine genetic stability and reproducibility of results.
- Design primers for reverse transcriptase polymerase chain reaction (RT-PCR) and sequencing of the RVFV S segment.
- Perform RT-PCR on RNA extracted from stocks of RVFV isolates.
- Perform automated sequencing of the $S$ segment RT-PCR products.
- Analyze and discuss sequencing data.
- Publish the sequencing data in GenBank.


## Chapter 2: Methods and Materials

### 2.1. Culture of Vero cells

An established cell line of green monkey kidney cells (Vero), obtained from the Special Pathogens Unit of the National Institute for Communicable Diseases (SPUNICD), was maintained in a monolayer at $37^{\circ} \mathrm{C}$ in $75-\mathrm{cm}^{2}$ tissue culture flasks supplemented with Eagles Minimal Essential Medium (EMEM) (BioWhittakker, MD, USA) containing $10 \%$ foetal calf serum (FCS) (Delta Bioproducts, SA) and antibiotics (100 IU penicillin, $100 \mu \mathrm{~g}$ streptomycin, and $0.25 \mu \mathrm{~g}$ Amphotericin B) (BioWhittaker, MD, USA) (Appendix 6.1.1.).

Confluent Vero monolayers were trypsinised and subcultured at a $1: 5$ ratio weekly. Briefly, the media was removed and cell monolayer rinsed twice with sterile phosphate buffered saline (PBS) (Appendix 6.1.2.). Trypsin-EDTA (ethylenediaminetetracetic acid) solution, at the concentration of $0.25 \%$ trypsin and $0.05 \%$ EDTA respectively, free of $\mathrm{Ca}^{++}$and $\mathrm{Mg}^{++}$, was then added, and the flask incubated at $37{ }^{\circ} \mathrm{C}$ for 10 minutes, allowing the cells to dislodge. Then, 10 ml of EMEM was added to the trypsinised cells and the mixture aspirated with a sterile syringe and needle to ensure complete dislodgment and thorough mixing of the media and cells. A 2 ml aliquot of cell suspension was transferred into a flask containing 18 ml of EMEM, supplemented with $10 \%$ FCS and antibiotics. When the cell monolayers reached $100 \%$ confluency, the flasks were incubated at $28^{\circ} \mathrm{C}$ or used immediately for inoculation with RVFV.

### 2.2. Propagation of Rift Valley fever viral isolates in Vero cells

RVFV isolates $(\mathrm{n}=45)$ recovered throughout Africa, Madagascar, and Saudi Arabia over the past 60 years (Table 1), were obtained from the NICD-SPU. All isolates were originally recovered from specimens, and most of them were then inoculated intracerebrally in 1-2 day old suckling mice. Viral stocks were prepared either by an additional passage in mice or by 1-2 passages in baby hamster kidney (BHK) cells and stored as freeze-dried or wet preparations at $-70^{\circ} \mathrm{C}$ (Paweska, personal communication). Before RNA extractions were carried out, each isolate was propagated once in Vero cells under standard sterile conditions. Briefly, a 1:10 dilution of each banked virus isolate in EMEM was first centrifuged at $3000 \mathrm{x} g$ and then filtered through a $0.22 \mu \mathrm{~m}$ filter to remove any possible bacterial or fungal contaminations. Then, 5 ml of the resulting supernatant was inoculated into $75-\mathrm{cm}^{2}$ tissue culture flask containing 48-72 hours old monolayers of Vero cells. After one hour incubation at $37^{\circ} \mathrm{C}$, inoculated cells were rinsed once with sterile PBS, overlaid with 20 ml EMEM containing standard concentration of antibiotics (Appendix 6.1.1.), and then incubated at $37^{\circ} \mathrm{C}$ in a $5 \% \mathrm{CO}_{2}$ incubator. Inoculated cells were examined microscopically daily for cytopathic effects (CPE). In the presence of advanced CPE (at least 70\%), infective culture medium was centrifuged at $2000 \mathrm{x} g$, and then 0.5 ml aliquots of the supernatant were prepared in 1.8 ml cryotubes and stored at $-70^{\circ} \mathrm{C}$ until used. A summary of the RVFV isolates used in this study is given in Table 1.

### 2.2.1. Passaging of Rift Valley fever viral isolates in Vero cells

Two isolates, SPU2223KEN07 and SPU2207KEN07, recovered during the 2006-7 outbreak in Kenya, were passaged through Vero cells. Briefly, 1:10 dilution of each clinical specimen (serum) was first centrifuged at $3000 \times g$ and then filtered through a $0.22 \mu \mathrm{~m}$ filter to remove any possible bacterial or fungal contaminations. Five ml of the resulting supernatant was inoculated into $75-\mathrm{cm}^{2}$ tissue culture flask containing 48-72 hours old monolayers of Vero cells. After one hour incubation at $37^{\circ} \mathrm{C}$, inoculated cells were rinsed once with sterile PBS, overlaid with 20 ml EMEM containing standard concentration of antibiotics, and then incubated at $37^{\circ} \mathrm{C}$ in a $5 \%$ $\mathrm{CO}_{2}$ incubator. Inoculated cells were examined microscopically daily for cytopathic effects (CPE). In the presence of advanced CPE (at least 70\%), infective culture medium was used to make another 1:10 dilution for the next passage. These viruses were passaged through Vero cells 10 times, and RNA extraction was done from the clinical specimen (serum) and from passage 10 in Vero cells (Vero\#10) for each isolate.

### 2.2.2. Passaging of Rift Valley fever viral isolates in mice

The same two Kenyan isolates were also passaged through mice. This was done by a series of 10 serial intracerebral inoculations of 1-2 day old suckling mice. As for the Vero cell passages, RNA was extracted from suckling mouse brain (SMB) passage number 10 (SMB\#10) for each isolate. This part of the project was done by SPU staff members under Animal Ethics Code clearance number 107/2006.

Table 1. Identification, origin, year of isolation, source and accession numbers of 45 Rift Valley fever virus strains analyzed in this study.

| Strain | Location of Origin | Year | Source | Accession Numbers |
| :---: | :---: | :---: | :---: | :---: |
| 56KEN65 | Kenya | 1965 | Bovine | EU312103 |
| 900085MAU88 | Mauritania | 1988 | Human | EU312104 |
| 214445KEN83 | Kenya | 1983 | Aedes mcintoshi | EU312105 |
| An991MAD91 | Madagascar | 1991 | Bovine | EU312106 |
| An999MAD91 | Madagascar | 1991 | Bovine | EU312107 |
| An1000MAD91 | Madagascar | 1991 | Bovine | EU312108 |
| AnK6087GUI84 | Guinea | 1984 | Bat (Micropteropus pusillus) | EU312109 |
| AnTAMBULEGY94 | Egypt | 1994 | Animal | EU312110 |
| Ar74RSA55 | South Africa | 1955 | Aedes circumluteolus | EU312111 |
| Ar811MAD79 | Madagascar | 1979 | Mosquito | EU312112 |
| Ar12568RSA71 | South Africa | 1971 | Eratmopodites quinquivittatus | EU312113 |
| Ar20368RSA81 | South Africa | 1981 | Culex zombaeusis | EU312114 |
| Ar21229SA00 | Saudi Arabia | 2000 | Mosquito | EU312115 |
| ArD38661SEN81 | Senegal | 1983 | Aedes dalzieli | EU312117 |
| B314KEN62 | Kenya | 1962 | Unknown | EU312118 |
| B1143KEN77 | Kenya | 1977 | Unknown | EU312119 |
| H1825RSA75 | South Africa | 1975 | Human | EU312120 |
| LunyoUGA55 | Uganda | 1955 | Mosquito | EU312121 |
| R1662CAR85 | CAR | 1985 | Human | EU312122 |
| SPU45ZAMB85 | Zambia | 1985 | Human | EU312123 |
| SPU77NAMB04 | Namibia | 2004 | Human | EU312124 |
| SPU204ANGL85 | Angola | 1985 | Human | EU312125 |
| SPU12002SOM98 | Somalia | 1998 | Caprine | EU312126 |
| SPU52001RSA99 | South Africa | 1999 | Buffalo | EU312127 |
| SPU384001KEN97 | Kenya | 1997 | Human | EU312128 |
| VRL763ZIM70 | Zimbabwe | 1970 | Bovine foetus | EU312130 |
| VRL825ZIM79 | Zimbabwe | 1979 | Bovine foetus | EU312131 |
| VRL1032ZIM78 | Zimbabwe | 1978 | Bovine | EU312132 |
| VRL1290ZIM78 | Zimbabwe | 1978 | Bovine foetus | EU312133 |
| VRL1516ZIM78 | Zimbabwe | 1978 | Ovine | EU312134 |
| VRL1887ZIM78 | Zimbabwe | 1978 | Bovine calf | EU312135 |
| VRL2230ZIM78 | Zimbabwe | 1978 | Bovine foetus | EU312136 |
| ZH501EGY77 | Egypt | 1977 | Human | EU312137 |
| ZH548EGY77 | Egypt | 1977 | Human | EU312138 |
| SPU2207KEN07 | Kenya | 2006 | Human | EU312139 |
| SPU2223KEN07 | Kenya | 2006 | Human | EU312140 |
| SPU2214KEN07 | Kenya | 2006 | Human | EU312141 |
| SPU2215KEN07 | Kenya | 2006 | Human | EU312142 |
| SPU2220KEN07 | Kenya | 2006 | Human | EU312143 |
| SPU10301KEN07 | Kenya | 2007 | Human | EU312144 |
| SPU10302KEN07 | Kenya | 2007 | Human | EU312145 |
| SPU10307KEN07 | Kenya | 2007 | Human | EU312146 |
| SPU10315KEN07 | Kenya | 2007 | Human | EU312147 |
| SPU77RSA08 | South Africa | 2008 | Buffalo | EU709747 |
| SPU152RSA08 | South Africa | 2008 | Human | EU709748 |

### 2.4. Extraction of RNA from infected tissue culture

The viral RNA was extracted from infected tissue culture supernatants as per the manufacturer's instructions using the QIAamp Viral RNA kit (Qiagen, CA, USA) (Appendix 6.2.1).

### 2.4. RT-PCR of extracted RNA using primers targeting the $S$ RNA segment

### 2.4.1. Primer design

Forward and reverse primers (F1 and R4, Table 2) were designed based on the S RNA sequence data of the MM12 strain of RVFV (GenBank accession number X53771). Vector NTI Advance 9 sequence analysis software (Invitrogen) was used to design oligonucleotides of 20 (forward primer F1) and 22 nucleotides (reverse primer R4) in length (Table 2). Primers have a base composition of $50-60 \%(\mathrm{G}+\mathrm{C})$ and end in C or GC to increase efficiency of priming, with melting temperatures (Tm) between 55$80^{\circ} \mathrm{C}$, and with little or no complementarity. The primers were synthesized by the Molecular and Cell Biology Department of the University of Cape Town (Cape Town, South Africa).

Table 2. Primer sequences for the amplification of the Rift Valley fever virus S RNA segment

| Primer | Position | Length (nt) | Sequence |
| :---: | :---: | :---: | :---: |
| F1 | $1-20$ | 20 | $5^{\prime}-$ ACA CAA AGA CCC CCT AGT GC - 3' |
| R4 | $1690-1669$ | 22 | $5^{\prime}-$ ACA CAA AGC TCC CTA GAG ATA C - 3' |

### 2.4.2. Reverse Transcription

The extracted RNA was reverse transcribed using the Expand Reverse Transcriptase (Roche, Germany). The Expand reverse transcriptase is an RNA directed DNA polymerase, which is a genetically engineered version of the Moloney Murine Leukaemia virus reverse transcriptase (M-MuLV-RT) that produces an increased amount of full-length cDNA transcripts.

The RNA was reverse transcribed in a final reaction volume of $10.5 \mu 1$, containing $1 \mu \mathrm{~g}$ of total RNA, and 20 pmol R4 primer (Table 2) in RNase- and DNase-free 0.2 ml thin-walled PCR tubes. The reaction mixture was incubated in a GeneAmp 2700 thermal cycler (Applied Biosystems, CA, USA) at $60^{\circ} \mathrm{C}$ for 10 minutes and then chilled on ice. To each reaction, 1x Expand reverse transcriptase buffer, 10 mM DTT, 1 mM dNTP mix, 20 U RNase inhibitor and 50 U Expand Reverse Transcriptase were added to give a final reaction volume of $20 \mu$. The reactions were incubated at $43^{\circ} \mathrm{C}$ for 1 hour. The reverse transcriptase products were stored at $4^{\circ} \mathrm{C}$ until use.

### 2.4.3. Polymerase Chain Reaction

The reverse transcriptase products were amplified, according to manufacturer's instructions, using the Expand High Fidelity Kit (Roche, Germany). The Expand High Fidelity PCR system is optimised to efficiently amplify DNA fragments up to 5 kb . It is composed of a unique enzyme mix containing thermostable Taq DNA polymerase and Tgo DNA polymerase, a thermostable DNA polymerase with proofreading activity. This polymerase mixture is designed to generate PCR products of high yield, high fidelity and high specificity from all types of DNA (Barnes 1994). The Expand High Fidelity PCR system shows a 3 -fold increase in the fidelity of DNA synthesis compared to Taq DNA polymerase due to the inherent 3' to $5^{\prime}$ exonuclease or "proofreading" activity of Tgo DNA polymerase.

Reactions were prepared in sterile RNase- and DNase-free 0.2 ml thin-walled PCR tubes. The PCR master mix consisted of 20 pmol F1 primer, 20 pmol R4 primer (Table 2), $2 \times$ Expand High Fidelity buffer, 10 mM dNTP mix, 2.6 U Expand High Fidelity enzyme mix. Per sample, $5 \mu \mathrm{l}$ of cDNA was added with $79.25 \mu \mathrm{l}$ nuclease free water to give a final volume of $100 \mu 1$. The PCR was performed using a GeneAmp 2700 thermal cycler (Applied Biosystems, CA, USA), with an initial denaturation step at $94^{\circ} \mathrm{C}$ for 2 minutes, followed by 35 cycles of denaturation at $94^{\circ} \mathrm{C}$ for 30 seconds, elongation at $55^{\circ} \mathrm{C}$ for 30 seconds, and extension at $72^{\circ} \mathrm{C}$ for 1.5 minute, and then followed by a final elongation step at $72^{\circ} \mathrm{C}$ for 7 minutes. The amplicons were analyzed with agarose gel electrophoresis using $1.2 \%$ agarose gel (Appendix 6.2.2) with a 0.5 x Tris-borate-ethylenediaminetetracetic acid EDTA (TBE) buffer (Appendix 6.1.3.) at 100 V for 1 hour. The DNA fragments were stained with $0.5 \mu \mathrm{~g} / \mathrm{ml}$ ethidium bromide $(\mathrm{EtBr})$ and then visualized using a UV transilluminator and the GeneTool software (SynGene, MD, USA).

The PCR products were purified as per manufacturer's instructions using the Wizard® SV Gel and PCR Clean-Up System (Promega, WI, USA) (Appendix 6.1.3), and stored at $-20^{\circ} \mathrm{C}$ until used.

### 2.5. Automated sequencing of the $S$ segment RT-PCR products

### 2.5.1. Primer design

The double-stranded DNA of the S segment generated by RT-PCR was sequenced using primers designed during the study (Table 3). Vector NTI Advance 9 sequence analysis software (Invitrogen) was used to design six forward and six reverse primers to meet the following conditions: about 17-22 nucleotides in length, a base composition of $50-60 \%(\mathrm{G}+\mathrm{C})$, end in C or GC to increase efficiency of priming, melting temperatures $(\mathrm{Tm})$ between $55-80^{\circ} \mathrm{C}$, and little or no complementarity to avoid primer dimer formation self complementarity. The primers were synthesized by the Molecular and Cell Biology Department of the University of Cape Town (Cape Town, South Africa).

Table 3. Primer sequences for the sequencing of the Rift Valley fever virus $S$ segment.

| Forward Primers |  |  |  |
| :---: | :---: | :---: | :---: |
| Primer | Position | Length (nt) | Sequence |
| F1 | 1-20 | 20 | 5' - ACA CAA AGA CCC CCT AGT GC - 3' |
| F2a | 461-480 | 20 | 5' - TGA GGA GTT GCA AGA TCA CC - ${ }^{\prime}$ ' |
| F2b (2F) | 514-536 | 22 | $5^{\prime}$ - TGG CTT ACA CAG GAT GAT AGC G-3, |
| F3a | 909-927 | 19 | 5' - AGC CAC TTA GGC TGC TGT C - ${ }^{\prime}$ ' |
| F3b (3F) | 1047-1063 | 17 | $5^{\prime}$ - TGC ATT CAT TGG CTG CG - ${ }^{\prime}$, |
| F4 | 1394-1412 | 19 | 5' - TGA GAG CCT CCA CAG TTG C - ${ }^{\prime}$ ' |
| Reverse Primers |  |  |  |
| Primer | Position | Length (nt) | Sequence |
| R1a | 548-531 | 18 | 5' - GGA TGG CCT CAG TCG CTA - ${ }^{\prime}$ |
| R1b (1R) | 394-373 | 21 | 5' - AGT GAG GGT TCT CCA AGA GGC - ${ }^{\prime}$, |
| R2a | 995-977 | 19 | 5' - GCC TTT GGA CTT GTG GAT T - 3' |
| R2b (2R) | 934-910 | 24 | 5, -TTA CAA GAC AGC AGC CTA AGT GGC- ${ }^{\text {, }}$, |
| R3 | 1458-1440 | 18 | 5, - AAC AAG CCC AGG AGG ATG - 3' |
| R4 | 1690-1669 | 22 | $5^{\prime}$ - ACA CAA AGC TCC CTA GAG ATA C - ${ }^{\prime}$, |

### 2.5.2. Cycle sequencing

The nucleotide sequences of the amplicons were determined using the BigDye ${ }^{\circledR}$ Terminator v3.1 Cycle Sequencing Ready Reaction kit (Applied Biosystems, Warrington, Great Britain), according to the manufacturer's instructions. The BigDye® Terminator set of dyes are labelled with novel, high-sensitivity dyes by Applied Biosystems using the discoveries of Rosenblum et al. (1997).

The cycle sequencing master mix for each primer consisted of: 0.8 pmol primer (Table 3), 5 x BigDye sequencing buffer, $10-40 \mathrm{ng}$ of cDNA , and nuclease free water to give a final volume of $20 \mu \mathrm{l}$. The cycle sequencing reaction was performed using a GeneAmp 2700 thermal cycler (Applied Biosystems, CA, USA), starting with an initial denaturation step at $96^{\circ} \mathrm{C}$ for 1 minute, followed by 25 cycles as follows: $96^{\circ} \mathrm{C}$ for 10 seconds, $50^{\circ} \mathrm{C}$ for 5 seconds, and $60^{\circ} \mathrm{C}$ for 4 minutes. The sample was then held at $4^{\circ} \mathrm{C}$ until ready to purify. Products were purified by isopropanol (Appendix 6.2.4).

### 2.5.3. Sequencing

The cycle sequencing product was analyzed with an ABI Prism® 3100 Genetic Analyzer (Applied Biosystems, Warrington, Great Britain) - a fully automated, fluorescence-based capillary electrophoresis platform.

The Applied Biosystems DNA sequencer detects fluorescence from the four different dyes that are used to identify the A, C, G and T extension reactions. Each dye emits light at a different wavelength when excited by an argon ion laser. All four colours, and therefore all four bases, can be detected and distinguished in a single gel lane or capillary injection.

The complete nucleotide sequence data of the RVFV S RNA segments generated in this study were submitted to GenBank and have been assigned GenBank accession numbers given in Table 1.

### 2.6. Analysis of the sequencing data

The data generated from sequencing was used for nucleotide and amino acid sequence alignment. Sequences were aligned using Sequencher (Gene Codes Corporation). Phylogenetic and molecular evolutionary analysis were conducted using Molecular Evolutionary Genetics Analysis software, version 4 (MEGA4) (Tamura et al. 2007). The Neighbour Joining (NJ) method was used to construct a phylogenetic tree. The NJ method is a simplified version of the Minimum Evolution (ME) method, which uses distance measures to correct for multiple hits at the same sites, and chooses a topology showing the smallest value of the sum of all branches as an estimate of the correct tree. Sequence divergence was done by calculating the average p-distances within and between groups.

The bootstrap test was used to ascertain the reliability of a given branch pattern by examining the frequency of its occurrence in a large number of trees, each based on the re-sampled dataset. MEGA4 used Felsenstein's bootstrap test (Felsenstein 1985), which was evaluated using Efron's bootstrap re-sampling technique.

A total of 75 isolates of RVFV were analyzed, of which 45 were originally sequenced in this study, and the remaining sequencing data for the other 30 isolates was obtained from GenBank. Sequencing data of 72 isolates, representing naturally circulating RVFV from Africa, Madagascar and Saudi Arabia, was used to generate phylogenetic trees to establish the relatedness of isolates based on their S RNA segments, and N and NSs proteins. Artificially generated strains, including the vaccine strains, were not included in the phylogenetic analyses.

## Chapter 3: Results

### 3.1. Sequencing analysis of the Rift Valley fever virus $S$ segment

Of the 45 strains sequenced in this study (Appendix 6.3.), six were recently sequenced by Bird et al., (2007), and are currently available in GenBank. For a more comprehensive analysis of the S RNA segment, the sequencing data from the 45 strains sequenced in this study were analyzed together with the 30 strains available in GenBank, 3 of which are laboratory strains. Only the naturally occurring strains ( $\mathrm{n}=$ 72) were used for phylogenetic analysis, and aligned to infer a consensus sequence (Figure 9). The three laboratory strains, clone 13, MP12 and SNS, were also included in sequence analysis $(\mathrm{n}=75)$, but not in the phylogenetic analysis.

Phylogenetic analysis indicates that circulating strains are compartmentalized and belong to one of three lineages, namely Egyptian, western African, and central, eastern and southern African. The strains clustered in the Egyptian lineage had an average p-distance of $1.0 \%$, the western African strains $0.9 \%$, and the central, southern and eastern African strains 2.0\%.

## 1 <br> 50

(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGATTACTTTCCTG (51) TGATATCTGTTGATTTGCAGAGTGGTCGTCGTGTTGTGTCAGTGGAGTAC (101) ATTAGAGGTGATGGTCCTCCCAGGATACCTTATTCTATGGTTGGGCCCTG (151) TTGTGTCTTTCTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGTCGGACTTGGA (251) GACTTTGCATCAAACGTTGCACCTCCACCAGCAAAGCCTTTTCAGAGACT (301) TATTGATCTAATAGGCCATATGACTCTTAGTGATTTCACAAGGTTCCCCA (351) ATCTGAAAGAAGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACATTAGAAGGGA (451) TCAGATTGCCACTCTAGCAATGAGGAGCTGCAAGATTACCAATGATCTAG (501) AGGACTCCTTTGTTGGCTTACACAGGATGATAGTGACCGAGGCTATCCTC (551) AGAGGGATTGACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (601) TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAAGAGGACATTT (651) CTAATGCTGTAGTTCCAAACTCAGCCCTCATTGCTCTTATGGAGGAGAGC (701) CTGATGCTGCGCTCATCACTTCCTAGCATGATGGGGAGAAACAACTGGAT (751) TCCAGTTGTTCCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (801) GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAGGCTGCCCCAC
(851) CCCCCACCCCCCAATCCCGACCGTAACCCCAACCCCCCTTTTCCCCCAAC
(901) CCCCTGGGCAGCCACTTAGGCTGCTGTCTTGTAGCCTGAGCAGCTGCCAT (951) GACAGCAGCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTTTCA (1001) AGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCTATTCACTGCTGCA
(1051) TTCATTGGCTGCGTGAACGTTGCAGCAACCTCCTCCTTTGTTCTACCTCG (1101) GAGGTTTGGGTTGATGACCCGGGAGAACTGCAGCAGATACAGAGAGTGAG (1151) CATCTAATATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACCATG (1201) CCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATGCAGGGGATAGGCC (1251) GTCCATGGTAGTCCCAGTGACAGGAAGCCACTCACTCAAGACGACCAAAG (1301) CCTGGCATGTCCAGCCAGCCAGGGCGGCAGCAACTCGTGATAGAGTCAAC (1351) TCATCCCGGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAGAGC (1401) CTCCACAGTTGCTTTGCCTTCTTTCGACATTTTCATCATCATCCTCCGGG (1451) GCTTGTTGCCACGAGTTAGAGCCAGAACAATCATTTTCTTGGCATCCTTC (1501) TCCCAGTCAGCCCCACCATACTGCTTTAAGAGTTCGATAACCCTACGGGC (1551) ATCAAACCCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCATTGC (1601) GGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTCTTGATAGTTGTCC (1651) ATTATTGTAATAGTGTTTGTATCTCTAGGGAGCTTTGTGT (1690)

Figure 9. The consensus sequence for the Rift Valley fever virus $S$ segment in the viral sense. The consensus sequence was generated after aligning the 45 strains sequenced in this study and 27 strains available from GenBank. Highlighted yellow region represents the NSs gene and the highlighted green region represents the N gene. The highlighted blue text is the proposed transcription termination motif, and the highlighted red text is the associated C/G-repeats.

To assess the potential for sequence variation to occur in the original material through routine laboratory manipulations, including primary virus isolation and/or subsequent virus propagation to prepare viral stocks, a passage series was done. Two isolates recovered from RVF patients during the 2007 outbreak in Kenya, SPU2223KEN07 and SPU2207KEN07, were passaged in two different amplification systems for RVFV, namely ten times in Vero cells and ten times in mice. Sequencing data generated from the virus RNA present in the original clinical specimens and from RNA extracted from Vero\#10 and SMB\#10 showed 100\% homology for each respective isolate (Appendix 6.6.). These results demonstrate that the RVFV S segment remained stable during ten serial passages in different propagation systems. This finding is important in terms of having confidence regarding originality of the virus genome of the strains selected for this study and their natural genetic relatedness. In addition, obtained results demonstrated the accuracy and reproducibility of the techniques used.

### 3.1.1. 5' and $3^{\prime}$ non -coding regions

The non-coding regions (NCRs) found at the extreme terminal ends of the S RNA segment were found to be highly conserved in the 75 isolates. Only two sequence variations were found in the 5 'NCR amongst all of them. The first was a 7A>C transversion in the ArD38388BF83 (Table 5) strain, isolated in Burkina Faso in 1983, and the SPU2214KEN07 strain, isolated in Kenya in 2007 (Table 1). The second sequence variation was a $26 \mathrm{~A}>\mathrm{T}$ transversion which occurred in the SPU52001RSA99 strain isolated in 1999 in South Africa.

Five sequence variations were found in the $3^{\prime}$ NCR. The first was a $1658 \mathrm{~T}>\mathrm{C}$ transition in the B314KEN62 strain, isolated in Kenya in 1962, and the SPU77NAMB04 strain, isolated in Namibia in 2004. The second variation was a $1663 \mathrm{~A}>\mathrm{G}$ transition in the SPU2214KEN07 strain, isolated in 2007 in Kenya (Table 1). The third variation was a $1666 \mathrm{~A}>\mathrm{T}$ transversion in the Ar20368RSA81 strain, isolated in 1981 in South Africa (Table 1), and the fourth was a $1671 \mathrm{~T}>\mathrm{A}$ transversion in the ZM657 Egyptian strain, isolated in this country in 1977 (Table 5). The fifth variation was 1697delT in the South African strain SPU77RSA08. This is the only deletion found that is not in the IGR.

### 3.1.2. The intergenic region

The intergenic region (IGR), a region between the two open reading frames (ORF) encoding the N protein and the NSs protein, was found to be variable. It has a p-distance range of 0 to $9.2 \%$, with an average of $4.8 \%$. This region of the S RNA segment is the only region where insertions and deletions were found.

A feature that makes the IGR very different to the rest of the S RNA segment is the C/G rich repeats. Such repeats, found downstream of the $5^{\prime}$ '-GGCTGCCC-3' transcription termination motif, are thought to be involved in transcription termination (Hutchinson et al. 1996; Ikegami et al. 2007). These motifs were found to be highly conserved in all 75 strains of RVFV analyzed in this study. The motif for NSs transcription termination was found at position 840-847, with the C rich repeat from 850-845/5, and for the N gene transcription termination, the motif was found at position $905 / 6-912 / 3$, with the $G$ rich repeats (when in the complementary sense) from position 899/900-903/4.

The attenuated MP12 strain has four nucleotide variations compared to its parent ZH548 strain, one of which is a transition found within the IGR, 862T>C (Vialat et al. 1997). This same transition was found in two of the Kenyan strains, B1143KEN77 and B314KEN62. Another Kenyan strain, SPU10307KEN07, also showed variation at this position, though in the form of a deletion of the nucleotide.

The IGR of clone 13 was found to have six nucleotide variations when compared to the ZH548 strain (Muller et al. 1995). When compared to the other strains, besides the section between position 879 and 884 , the only other unique difference was a transition, 860C> T.

### 3.1.3. The NSs gene and protein

The NSs open reading frame (ORF) is 798 nucleotides in length. It begins with an ATG start codon at position 35 and ends with a TAG stop codon at position 832. The ORF encodes a protein of 266 amino acid residues. A total of 167 sequence variations were found in this gene, 65 of which were unique. In the resulting protein a total of 51 amino acid changes were found, 27 of which were unique. Subsequent to the alignment of the amino acid sequences of the strains incorporated into this study, an NSs protein amino acid consensus sequence was defined (Figure 10).

```
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMHHRPSH
(51) EVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLIDLIGHMTLSD
(101) FTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDDIRRDQIATLAMRSCK
(151) ITNDLEDSFVGLHRMIVTEAILRGIDLCLLPGFDLMYEVAHVQCVRLLQA
(201) AKEDISNAVVPNSALIALMEESLMLRSSLPSMMGRNNWIPVVPPIPDVEM
(251) ESEEESDDDGFVEVD
```

Figure 10. The consensus amino acid sequence for the Rift Valley fever virus NSs protein.

As shown previously by Vialet et al. (1997), the MP12 strain varies from its parent strain, ZH548, by three nucleotides within the NSs ORF. These variations were all $\mathrm{T}>\mathrm{C}$ transitions and were found at positions 208, 403 and 513. The 208T>C transition also occurred in the Ar74RSA55 strain (Table 1), but the remaining two transitions were restricted to the MP12 strain. Of the three variations, only one, $513 \mathrm{~T}>\mathrm{C}$, resulted in an amino acid change, V160N (Vialat et al. 1997).

Amino acid changes that could affect the structure of the protein are of importance. One such change, I23F, occurred in twelve strains, and was due to the $101 \mathrm{~A}>\mathrm{T}$ transversion. A 159T>A transversion unique to the SPU52001RSA99 strain (Table 1), isolated in South Africa in 1999, resulted in the amino acid change I42Y. Another unique variation was found in the Ar20368RSA81 strain (Table 1), isolated in South Africa in 1981, and was due to the 182C>T transition, which resulted in the amino acid change H50Y. There was only one substitution that involved a cysteine, Y67C, which occurred in the SPU45ZAMB85 strain (Table 1), isolated in Zambia in 1985, and the Smithburn strain due to the $234 \mathrm{~A}>\mathrm{G}$ transition.

As defined by Sall et al. (2008), the interacting domain of NSs-SAP30 is found between amino acid residues 210 and 230. An analysis of this region found 6 amino acid variations. The first change, V210A, was found in the South African strain Ar20369RSA81 (Table 1). The second change, A214G, occurred in the 2 strains from Guinea, Ank6087GUI84 and ANK3837GUI81, and the Zinga strains from CAR (Table 5). The third variation, A217V, was common to all Egyptian lineage strains. The fourth variation, S222N, was restricted to the SNS vaccine strain. The fifth variation, S227L, occurred in the SPU10301KEN07 and SPU10307KEN07 Kenyan strains (Table 1). The sixth variation, S227T, occurred in the SPU2214KEN07 strain (Table 1). Besides the substitution in the SNS strain, there were no other notable changes.

### 3.1.4. The $\mathbf{N}$ gene and protein

The N gene is read in the complementary sense and starts at position $1652 / 4$ with an ATG start codon. It is 783 nucleotides in length and terminates at position 915/7 with a TAA stop codon. An exception to the stop codon was the LunyoUGA55 strain which contained a TGA stop codon. The N gene encodes a protein made up of 246 amino acid residues. Within the 783 nucleotides there were 134 substitutions, 53 of which were unique. Only 43 of the sequence variations resulted in an amino acid change, 7 of which where unique. The N gene shows almost the same number of nucleotide substitutions as the NSs gene, but fewer of them resulted in unique amino acid changes. The N protein amino acid consensus sequence inferred from the amino acid sequence alignment of the naturally occurring strains included in this study is shown in Figure 11.

$$
\begin{array}{ll}
\hline(1) & \text { MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQYGGADW } \\
(51) & \text { EKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKYKLKEGNPSRD } \\
(101) & \text { ELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDGLSPAYPRHMMHPSFA } \\
(151) & \text { GMVDPSLPEDYLRAILDAHSLYLLQFSRVINPNLRGRTKEEVAATFTQPM } \\
(201) & \text { NAAVNSNFISHEKRREFLKAFGLVDSNGKPSAAVMAAAQAYKTAA }
\end{array}
$$

Figure 11. The consensus amino acid sequence for the Rift Valley fever virus N protein.

It has been shown that, besides that large deletion in the NSs gene of clone 13, there was only one other nucleotide substitution which resulted in an amino acid change, an E159G substitution (Muller et al. 1995). However, this is not unique to clone 13 , since this variation occurred in twelve of the 75 strains analyzed, and is due to the $1180 \mathrm{~T}>\mathrm{C}$ transition.

Of the seven unique sequence variations that resulted in amino acid changes, three where found in the Entebbe strain (Table 5). These three variations were $100 \mathrm{~T}>\mathrm{C}, 1036 \mathrm{~T}>\mathrm{C}$ and $1096 \mathrm{C}>\mathrm{T}$, which resulted in amino acid changes K219R, S207N and R187K, respectively. The attenuated SNS strain had several unique sequence variations in the N gene, though only one, $1528 \mathrm{~T}>\mathrm{C}$, resulted in an amino acid change, K43R. Unique sequence variations, $931 \mathrm{~T}>\mathrm{C}$ and $1524 \mathrm{C}>\mathrm{A}$, were also found in two Madagascan strains, An999MAD91 and An1000MAD91 (Table 1), and resulted in the amino acid changes K242R and Q44H, respectively. An Egyptian isolate, ZH1776EGY78 (Table 5), had a variation, 1274C>T, that resulted in the amino acid change V128I.

### 3.2. Phylogeny

### 3.2.1. The $S$ RNA segment phylogeny

The evolutionary history was inferred using the Neighbour-Joining method (Saitou and Nei 1987). The phylogenetic tree is drawn to scale, with branch lengths in the identical units as those of the evolutionary distances used to infer the tree. In general, isolates from the same geographic locations tend to cluster together (Figure 11), which supports previous phylogenetic analyses that have defined three broad lineages, namely Egyptian, western African, and central, eastern and southern African (Sall et al. 1999). These clusters are supported by high bootstrapping values and correlate with the geographic origin of the isolates (Figure 11).

The three main phylogenetic lineages, shown in Figure 11, can be further divided into 16 genotypes (A-P). The chosen parameters of a genotype were groupings that had a bootstrap value of $70-100 \%$, and a p-distance (the proportion of distances, i.e. the number of pairwise nucleotide differences divided by the total number of nucleotides in the sequenced region, calculated with MEGA4) of less than 0.016 to all other members of the phylogenetic cluster, and were based on research by Venter et al (2001). These genotypes are shown in Figure 11 and summarized in Table 4.

Table 4. Average p-distances within the 16 genotypes A-P

| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 0.010 | - | 0.010 | - | 0.000 | 0.002 | 0.010 | 0.005 | - | - | - | - | - | - | - | 0.009 |

(-) Genotypes with non calculable p-distances due to only one representative

The average p-distance for the entire S segment was only $2.5 \%$, with a range from 0 to $4.1 \%$. With such low genetic diversity amongst strains analysed a low average p-distance value must be used for defining a genotype, which is consequently also supported by a low bootstrap value.


Figure 12. Evolutionary relationships of the S RNA segment of 72 strains of Rift Valley fever virus. The evolutionary history was inferred using the Neighbour-Joining method (Saitou and Nei 1987). The bootstrap consensus tree inferred from 1000 replicates (Felsenstein 1985) is taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than $50 \%$ bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test ( 1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method (Kimura 1980) and are in the units of the number of base substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). There were a total of 1682 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007). The letter designations represent the genotype grouping of the strains. The tree is also divided into the three broad lineages, namely the Egyptian lineage (red), the western African lineage (blue) and the central, eastern and southern African lineages (green). Bold and underlined names represent strains sequenced in this study.

The phylogenetic tree shows the two possible modes of circulation, namely distant spread from one region to another, and local circulation in an enzootic/endemic area. Virus isolates found within the Egyptian lineage are the RVFV strains isolated during the 1977-79 Egyptian outbreak, the Tambul strain, as well as two Zimbabwean strains isolated in 1974 and 1978, and two Madagascar strains isolated in 1979. The strains from the 1977-79 outbreak are all closely related, which is an example of local circulation. Within the Egyptian lineage there is also an example of distance spread, namely the close relatedness of the Zimbabwean and Madagascan strains to the Egyptian strains. The original Madagascan isolate is Ar811MAD79, and the other isolate, MgH824MAD79 was from a laboratory infection whilst handling the Ar811MAD79 strain. The Zimbabwean and Madagascan strains show a close relationship, which suggests that the Ar811MAD79 strain was introduced to the island from the eastern coast of Africa. Furthermore, there is a close relationship to a Namibian strain from 2004, and a South African strain from 1981, which could be due to spread of the Zimbabwean isolates, or due to a reassortment event.

In the West African lineage there are strains from the Mauritanian outbreak in 1987-88, a strain from Burkina Faso in 1983, and a strain from Senegal in 1981. The RVF viruses isolated in Guinea in 1981 and 1984 should, according to geography, be included in this lineage, but have closer homology to members of the central, eastern and southern African lineage. Both strains are closely related to the Zinga strain isolated in CAR in 1969.

The majority of the RVFV strains are part of the central, eastern and southern African lineage. Represented in this lineage are isolates from South Africa, Zimbabwe, Zambia, Angola, CAR, Uganda, Kenya, Somalia, Saudi Arabia, Madagascar, and Guinea, which were collected over a period of 63 years, from 19442008. The most recent outbreaks to occur were in Kenya in 2007 and in South Africa in 2008. Viral isolates from these separate outbreaks are genetically very similar. There seems to be a pattern of Kenyan isolates being related to Southern African isolates. A Kenyan isolate from 1965 forms part of the same genotype group (H) as South African isolates from 1970-71 and 1975. Another example is the Kenyan isolates from the 1997-98 outbreak which form part of the same genotype group (G) as a South African isolate from 1999. The majority of the Kenyan isolates are found in the G genotype group and represent strains of RVFV isolated during several different outbreaks, suggesting that these outbreaks could be due to the re-emergence of epizootic viruses.

The South African isolates appear to be distantly related. Isolates from the 1971 and 1975 outbreaks show close homology, whilst isolates from the 1951, 1955 and 1999 outbreaks are genetically more distant. This suggests that more than one strain of RVFV was circulating in the country. This is also notable amongst the Zimbabwean strains. Isolates from the 1978-79 outbreak are more closely related than those from the 1970 and 1974 outbreaks. Isolates from the 1974 outbreak are particularly distinct, with at least two different strains of the virus circulating during the outbreak.

### 3.2.2. The $\mathbf{N}$ gene and protein phylogeny

The N gene was found to have virtually the same degree of sequence variation as the NSs gene. For the N gene, the p-distance ranged from 0 to $4.2 \%$, with an average of $2.2 \%$, and for the N protein, the range was from 0 to $2 \%$, with an average of $0.2 \%$.

When the N gene region is used to generate a phylogenetic tree (Figure 13), the isolates cluster in the same way as is seen for the entire $S$ segment (Figure 12). When the translated proteins of the N gene are inferred in a tree, the tree is dramatically different. The phylogenetic tree inferred for the N gene alignment shows the three distinct geographical divisions (Figure 13), however the phylogenetic tree inferred for N protein alignment only shows two lineages, the Egyptian cluster and central, eastern and southern African cluster, which also contains the western African strains (Figure 14). This is not surprising as 43 of the 134 sequence variations resulted in amino acid changes, and only 7 of those variations were unique to single isolates.


Figure 13. Evolutionary relationships of the N gene of 72 strains of Rift Valley fever virus. The evolutionary history was inferred using the Neighbour-Joining method (Saitou and Nei 1987). The bootstrap consensus tree inferred from 1000 replicates (Felsenstein 1985) is taken to represent the evolutionary history of the strains analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than $50 \%$ bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method (Kimura 1980) and are in the units of the number of base substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). There were a total of 738 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007). The tree is also divided into the three broad lineages, namely the Egyptian lineage (red), the western African lineage (blue) and the central, eastern and southern African lineages (green). Bold and underlined names represent strains sequenced in this study.


Figure 14. Evolutionary relationships of the N protein of 72 strains of Rift Valley fever virus. The evolutionary history was inferred using the Neighbour-Joining method (Saitou and Nei 1987). The bootstrap consensus tree inferred from 1000 replicates (Felsenstein 1985) is taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than $50 \%$ bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). There were a total of 245 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007). Unlike the other trees, the N protein phylogeny only shows two lineages, the Egyptian lineage (red) and the central, eastern and southern African lineages (green). Bold and underlined names represent strains sequenced in this study.

### 3.2.3. The NSs gene and protein phylogeny

The NSs gene has almost the same degree sequence variation as the N gene, but more of them result in amino acid changes; 167 variations resulted in 51 amino acid changes of which 27 were unique to single isolates. This is reflected in the phylogenetic trees inferred from the alignment of the NSs gene sequences and the alignment of the NSs protein amino acid sequences. Both phylogenetic trees (Figure 15 and Figure 16) still show many clusters and the division of the isolates into the three broad lineages, showing that the NSs protein has a higher degree of divergence compared to the N protein. The NSs gene has a p-distance range of 0 to $4.6 \%$, with an average of $2.4 \%$, and the NSs protein has a range of $0-3.8 \%$, with an average of $1.7 \%$.


Figure 15. Evolutionary relationships of the NSs gene of 72 strains of Rift Valley fever virus. The evolutionary history was inferred using the Neighbour-Joining method (Saitou and Nei 1987). The bootstrap consensus tree inferred from 1000 replicates (Felsenstein 1985) is taken to represent the evolutionary history of the strains analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than $50 \%$ bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method (Kimura 1980) and are in the units of the number of base substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). There were a total of 797 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007). The tree is also divided into the three broad lineages, namely the Egyptian lineage (red), the western African lineage (blue) and the central, eastern and southern African lineages (green). Bold and underlined names represent strains sequenced in this study.


Figure 16. Evolutionary relationships of the NSs protein of 72 strains of Rift Valley fever virus. The evolutionary history was inferred using the Neighbour-Joining method (Saitou and Nei 1987). The bootstrap consensus tree inferred from 1000 replicates (Felsenstein 1985) is taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than $50 \%$ bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). There were a total of 264 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007). The tree also shows the three broad lineages, namely the Egyptian lineage (red), the western African lineage (blue) and the central, eastern and southern African lineages (green). Bold and underlined names represent strains sequenced in this study.

## Chapter 4: Discussion

In this study, all 45 RVFV S segments were found to be highly conserved. Within the S segment, the N gene was shown to be more genetically stable than the NSs gene. The regions with the most genetic stability were the 5 ' and 3 ' terminal NCRs, while the greatest level of genetic variability was found in the IGR. Similar results were obtained by Bird et al. (2007), who demonstrated low virus genetic diversity among 33 RVFV S segments representing the virus strains recovered from throughout Africa and Saudi Arabia from 1944 to 2000. Moreover, these authors also showed a higher sequence divergence in the IGR between the N and NSs genes, than in the N and NSs genes. One possible reason for the low genetic diversity is viral maintenance through endemic or enzootic cycles involving little viral activity and hence less possible for mutations to occur (Sall et al. 1998).

The segmented genome structure of the members of the Bunyaviridae family carries the implication for reassortments to occur in nature, which might result in the appearance of new strains. New strains can also emerge as a result of point mutations, deletions, insertions, and recombination when co-infection of a susceptible host occurs with two or more strains of the virus (Gentsch and Bishop 1976; Turell et al. 1990; Sall et al. 1999), but recent studies indicate that this does not commonly occur in RVFV (Bird et al. 2007).

A part of this study involved serial passaging of two strains of RVFV through two different propagation systems in order to investigate if these routine laboratory procedures would induce any sequence variations within the $S$ segment of the virus. Comparison of sequencing data for each strain analyzed before passaging and after 10 passages in Vero cells and 10 passages in mice did not reveal any mutations. These results are of great importance with regards to having confidence in the originality of virus genomes of the strains analyzed in this study, since all of them were subjected to low passage either in in vivo and/or in in vitro amplification systems, before RNA was extracted for sequence analysis.

The NCRs of RVFV have been shown to be the most conserved. The $5^{\prime}$ and $3^{\prime}$ NCRs had only two and five mutations, respectively, among the 75 strain sequences. However, these few mutations could still be of importance. The 5' and 3' NCRs of phleboviruses are complementary, which can lead to the formation of circular or panhandle form of RNA (Bouloy et al. 1973; Samso et al. 1976; Hewlett et al. 1977; Pardigon et al. 1982), and it has previously been shown that a point mutation in the BUNV S segment at nucleotide 16 resulted in the formation of a loop structure in the panhandle, which caused a decrease in the free energy resulting in a strong increase in N mRNA transcription (Kohl et al. 2003). The possible biological significance of mutations in the NCRs found in this study requires further investigation.

The IGR of the S segment was found to be the most variable region, with a maximum pairwise difference of $4.8 \%$. Like the rest of the S segment there were several transitions and transversions. Apart from a deletion found in the 3 'NCR of the South African strain SPU77RSA08 (Table 1), the IGR is the only region with insertions and deletions. In several other Bunyaviridae the IGR has been found to form stem-loops structures, playing an important role in gene folding (Freier et al. 1986; Emery and Bishop 1987), but it is not certain whether the IGR of the RVFV S segment forms loops as the IGR of SFSV, which is most closely related to RVFV, does not form any stem-loop structures (Marriot et al. 1989).

The IGR is also the site of transcription termination, thus sequence variations of these termination sites could affect transcriptional termination. The pentanucleotide motifs reported by Albarino et al. (2007) were found to be highly conserved in the IGRs of all 75 RVFV strains. These were followed by the predicted C/G homopolymeric repeats, although some strains had longer C/G repeats than others. Ikegami et al. (2007) have recently shown that these polymeric C or G sequences found upstream of the CGUCG motif are also involved in efficient transcriptional termination. It has been shown that DNA-dependant RNA polymerases of prokaryotes and eukaryotes undergo slippage during transcription of homopolymeric sequences (Uptain et al. 1997) and that RNA-dependant RNA polymerases of negative-strand RNA viruses experience a similar slippage during transcription (Jacques et al. 1994; Barr et al. 1997). The termination signal motifs were found to function independently of these C/G repeats, and only deletion of the transcription termination motif disrupts
specific mRNA termination. These findings seem to rule out a secondary structure involving this region or some other role for this C/G-rich sequence in mRNA transcription (Albarino et al. 2007).

The N gene was found to have nearly the same number of mutations as the NSs gene, but the majority were silent mutations, which is not surprising as the nucleocapsid protein is vital to RVFV replication (Emerson and Wagner 1972) and any amino acid change that would affect its function might be lethal to the virus. It is possible that such mutations have occurred but were so detrimental that the virus was incapable of replication. The N protein has been found to be more conserved within the Bunyaviridae than the NSs protein (Giorgi et al. 1991). Studies comparing the amino acid sequences of the UUKV N protein with PTV and SFSV showed their homology to be 35 and $32 \%$ respectively. When the NSs protein was compared, the relatedness was much lower, with only 15\% homology (Simons et al. 1990).

In this study, the NSs gene was found to be highly variable in terms of nucleotide substitutions, though these did not all result in amino acid changes. As the NSs protein function depends on its ability to bind the p44 and XPB of the TFIIH transcription factor (Le May et al. 2004), any amino acid changes that would affect its secondary structure could have either a negative or positive effect. Cysteine is a hydrophobic amino acid and is highly reactive, capable of reacting with another cysteine to form a disulphide bond, a covalent bond that functions to stabilize the three-dimensional structure of a protein (Hames and Hooper 2000). Therefore, mutations within the cysteine residues of the NSs protein might influence its ability to sequester the p44 and XPB subunits of the TFIIH transcription factor, and subsequently influence its ability to inhibit the innate immune system and affect the virulence of RVFV strains. In a previous study, an amino acid sequence alignment of the NSs protein of several RVFV strains showed that the five cysteine residues at positions 39, 40, 150, 179 and 195 are conserved (Sall et al. 1997). These cysteine residues were found to be also conserved in all 75 strains analyzed in this study. In addition, it was found that at position 67 there was a missense mutation that resulted in a cysteine in two strains, the SNS strain and the Zambian strain SPU45ZAMB85. The SNS strain is a vaccine virus, and even though this amino acid change alone my not be responsible for its attenuation, as the Zambian strain is not attenuated, it might
be part of multiple variations that contribute to its attenuation. Bird et al. (2008) have recently shown that a genetically-engineered RVFV which lacks non-structural protein genes, the NSs gene alone or the $\mathrm{NSs} / \mathrm{NSm}$ genes in combination, was shown to be highly attenuated, with no detectable viremia and clinical illness observed after lethal challenge of rats (Bird et al. 2008).

Results of this study clearly demonstrate that, in general, isolates of RVFV from similar historic and geographic origins tend to cluster together (Figure 12). Similar to previous reports (Sall et al. 1999), the present phylogenetic analysis defined three broad genetic lineages, namely western African, Egyptian and central, eastern and southern African. The Egyptian and western African lineages show the least degree of evolution, whilst the central, eastern and southern African strains show a higher degree of genetic diversity. One possible explanation for the lower genetic diversity within the Egyptian and western African lineages is the relatively recent introduction of RVFV to these regions, and possibly also due to differences in natural mechanisms involved in RVFV maintenance between the various geographic parts of Africa. It is not fully understood how RVFV survives during inter-epidemic periods, but transovarial transmission in mosquitoes and the ability of the virus to persist in transovarial-infected mosquito eggs over long periods of dry seasons is most likely to play an important role (Davies 1975; Linthicum et al. 1985). The dynamics of RVFV transmission within the cryptic cycle, between and within the wildlife-livestockhuman interface, which is likely driven by various competent vectors and genetic susceptibility of host vertebrates, might also greatly influence the evolution and ecology of the virus. Other possible factors preventing high genetic divergence amongst ecologically and biologically diverse strains of RVFV were recently discussed by Bird et al. (2007), and include a "double filter" system. This concept involves maintaining the number of persistently infected mosquitoes above a critical population size, which is achieved by RVFVs ability to infect suitable mammalian amplification hosts easily, which in turn serve as reservoirs for the infection of large numbers of naïve mosquitoes (Bird et al. 2007). Any genetic variations that would limit the host range of RVFV could be detrimental to its survival.

The clustering of RVFV strains within the phylogenetic tree derived from the alignment of the S segment sequences suggests that several different strains of RVFV can co-circulate or be co-dormant within an area, but also that closely related strains are maintained within an area for a long time. For example, among the South African strains, the Ar20368RSA81 strains seemed to be very distantly related to the Ar74RSA55, H1825RSA75 and Ar12568RSA71 strains, and the H1825RSA75 and Ar12568RSA71 strains are very closely related. Another example are the Kenyan isolates B1143KEN77 and 214445KEN83, which were very closely related even though they were isolated years apart.

With a vector such as the mosquito it is easy for RVFV to travel vast distances. Infected mosquitos can transmit the virus to livestock, which in turn act as reservoirs that move to various regions. Infected moquitos can also move great distances by high winds and then infect livestock in the area they end up in. Another possibility it that transovarial-infected mosquito eggs can be carried by animals and even migrating birds to different areas. A good example of how one strain has been spread over vast distances is the 1979 Madagascar strain Ar811MAD79, of the Egyptian lineage, which was found to be closely related to the Zimbabwean strain VRL2250ZIM74, isolated in this country in 1974. Another example is the two West African isolates, ANK3837GUI81 and ANK6087GUI84, which were found to belong to the central, eastern and southern African cluster. These results demonstrate a widespread movement of different RVFV strains throughout Africa, and highlight the possibility of the virus spreading to the rest of the world. Furthermore, the recent and first occurrence of severe outbreaks of RVF in humans and livestock on the Saudi Arabian Peninsula in 2000 (Shoemaker et al. 2002) confirms the virus's ability to unexpectedly extend its distribution outside its traditional geographical confines.

Generated sequencing data contributes to global phylogenetic characterization of RVFV isolates and molecular epidemiology of the virus. In addition, findings of this study will further aid investigation on reassortment events occurring between strains of RVFV and genetically related viruses, the role of NSs protein in the replicative cycle of the virus, the pathogenic effects of NSs within the RVFVinfected host cells, and might help to identify molecular basis of RVFV virulence.

## Chapter 5: References

Abdel-Wahab, K. S., L. M. El Baz, E. M. El-Tayeb, H. Omar, M. A. Ossman and W. Yasin (1978). "Rift Valley Fever virus infections in Egypt: Pathological and virological findings in man." Trans R Soc Trop Med Hyg 72(4): 392-6.

Abdo-Salem, S., G. Gerbier, P. Bonnet, M. Al-Qadasi, A. Tran, E. Thiry, G. Al-Eryni and F. Roger (2006). "Descriptive and spatial epidemiology of Rift valley fever outbreak in Yemen 2000-2001." Ann N Y Acad Sci 1081: 240-2.

Accardi, L., C. Prehaud, P. Di Bonito, S. Mochi, M. Bouloy and C. Giorgi (2001). "Activity of Toscana and Rift Valley fever virus transcription complexes on heterologous templates." J Gen Virol 82(Pt 4): 781-5.

Akashi, H. and D. H. Bishop (1983). "Comparison of the sequences and coding of La Crosse and snowshoe hare bunyavirus S RNA species." J Virol 45(3): 1155-8.

Al-Hazmi, M., E. A. Ayoola, M. Abdurahman, S. Banzal, J. Ashraf, A. El-Bushra, A. Hazmi, M. Abdullah, H. Abbo, A. Elamin, T. Al-Sammani el, M. Gadour, C. Menon, M. Hamza, I. Rahim, M. Hafez, M. Jambavalikar, H. Arishi and A. Aqeel (2003). "Epidemic Rift Valley fever in Saudi Arabia: a clinical study of severe illness in humans." Clin Infect Dis 36(3): 245-52.

Albarino, C. G., B. H. Bird and S. T. Nichol (2007). "A shared transcription termination signal on negative and ambisense RNA genome segments of Rift Valley fever, sandfly fever Sicilian, and Toscana viruses." J Virol 81(10): 5246-56.

Anderson, G. W., Jr. and C. J. Peters (1988). "Viral determinants of virulence for Rift Valley fever (RVF) in rats." Microb Pathog 5(4): 241-50.

Anonymous (2007a). "Outbreaks of Rift Valley fever in Kenya, Somalia and United Republic of Tanzania, December 2006-April 2007." Wkly Epidemiol Rec 82(20): 169-78.

Anonymous (2007b). "Outbreak news. Rift Valley fever, Sudan." Wkly Epidemiol Rec 82(46): 401-2.

Arthur, R. R., M. S. el-Sharkawy, S. E. Cope, B. A. Botros, S. Oun, J. C. Morrill, R. E. Shope, R. G. Hibbs, M. A. Darwish and I. Z. Imam (1993). "Recurrence of Rift Valley fever in Egypt." Lancet 342(8880): 1149-50.

Assaad, F., F. G. Davies, G. A. Eddy, R. El Karamany, J. M. Meegan, Y. Ozawa, A. Shimshony, R. E. Shope, J. Walker and R. J. Yedloutschnig (1983). "The use of veterinary vaccines for prevention and control of Rift Valley fever: memorandum from WHO/FAO meeting." Bull World Health Organ 61: 261.

Balkhy, H. H. and Z. A. Memish (2003). "Rift Valley fever: an uninvited zoonosis in the Arabian peninsula." Int J Antimicrob Agents 21(2): 153-7.

Barnes, W. M. (1994). "PCR amplification of up to $35-\mathrm{kb}$ DNA with high fidelity and high yield from lambda bacteriophage templates." Proc Natl Acad Sci U S A 91(6): 2216-20.

Barr, J. N., S. P. Whelan and G. W. Wertz (1997). "cis-Acting signals involved in termination of vesicular stomatitis virus mRNA synthesis include the conserved AUAC and the U7 signal for polyadenylation." J Virol 71(11): 8718-25.

Bi, Z., P. B. H. Formenty and C. E. Roth (2008). "Hantavirus infection: a review and global update." J Infect Dev Countr 2(1): 3-23.

Billecocq, A., M. Spiegel, P. Vialat, A. Kohl, F. Weber, M. Bouloy and O. Haller (2004). "NSs protein of Rift Valley fever virus blocks interferon production by inhibiting host gene transcription." J Virol 78(18): 9798-806.

Bilsel, P. A., R. B. Tesh and S. T. Nichol (1988). "RNA genome stability of Toscana virus during serial transovarial transmission in the sandfly Phlebotomus perniciosus." Virus Res 11(1): 87-94.

Bird, B. H., C. G. Albarino, A. L. Hartman, B. R. Erickson, T. G. Ksiazek and S. T. Nichol (2008). "Rift valley fever virus lacking the NSs and NSm genes is highly attenuated, confers protective immunity from virulent virus challenge, and allows for differential identification of infected and vaccinated animals." $\mathbf{J}$ Virol 82(6): 2681-91.

Bird, B. H., M. L. Khristova, P. E. Rollin, T. G. Ksiazek and S. T. Nichol (2007). "Complete genome analysis of 33 ecologically and biologically diverse Rift Valley fever virus strains reveals widespread virus movement and low genetic diversity due to recent common ancestry." J Virol 81(6): 2805-16.

Bishop, D. H., C. H. Calisher, J. Casals, M. P. Chumakov, S. Y. Gaidamovich, C. Hannoun, D. K. Lvov, I. D. Marshall, N. Oker-Blom, R. F. Pettersson, J. S. Porterfield, P. K. Russell, R. E. Shope and E. G. Westaway (1980). "Bunyaviridae." Intervirology 14(3-4): 125-43.

Bishop, D. H. L., M. E. Gay and Y. Matsuoko (1983). "Nonviral heterogeneous sequences are present at the $5^{\prime}$ ends of snowshoe hare bunyavirus S complementary RNA." Nucleic Acids Res 11: 6409-6418.

Bishop, D. H. L. and R. E. Shope (1979). Bunyaviridae Comprehensive Virology. H. F.-C. R. Wagner. New York, Plenum Press. 14: 1-156.

Blakqori, G., G. Kochs, O. Haller and F. Weber (2003). "Functional L polymerase of La Crosse virus allows in vivo reconstitution of recombinant nucleocapsids." $\underline{\mathbf{J}}$ Gen Virol 84(Pt 5): 1207-14.

Bouloy, M., C. Janzen, P. Vialat, H. Khun, J. Pavlovic, M. Huerre and O. Haller (2001). "Genetic evidence for an interferon-antagonistic function of rift valley fever virus nonstructural protein NSs." J Virol 75(3): 1371-7.

Bouloy, M., S. Krams-Ozden, F. Horodniceanu and C. Hannoun (1973). "Three segment RNA genome of Lumbo virus (bunyavirus)." Intervirology 2(3): 173180.

Bouloy, M., N. Pardigon, P. Vialat, S. Gerbaud and M. Girard (1990). 'Characterization of the 5' and 3' ends of viral messenger RNAs isolated from BHK21 cells infected with Germiston virus (Bunyavirus)." Virology 175(1): 50-8.

Bouloy, M., S. J. Plotch and R. M. Krug (1978). "Globin mRNAs are primers for the transcription of influenza viral RNA in vitro." Proc Natl Acad Sci U S A 75(10): 4886-90.

Brigden, A., F. Weber, J. F. Fazakerley and R. M. Elliott (2001). "Bunyamwera Bunyavirus Nonstructural Protein NSs is a Nonessential Gene Product that Contributes to Viral Pathogenesis." Proc Natl Acad Sci U S A 98(2): 664-669.

Brown, J. L., J. W. Dominik and R. L. Morrissey (1981). "Respiratory infectivity of a recently isolated Egyptian strain of Rift Valley fever virus." Infect Immun 33(3): 848-53.

Caplen, H., C. J. Peters and D. H. Bishop (1985). "Mutagen-directed attenuation of Rift Valley fever virus as a method for vaccine development." J Gen Virol 66 ( Pt 10): 2271-7.

CDC (1994). "Rift Valley fever--Egypt, 1993." MMWR Morb Mortal Wkly Rep 43(38): 693, 699-700.

CDC (1998). "Rift Valley Fever--East Africa, 1997-1998." MMWR Morb Mortal Wkly Rep 47(13): 261-4.

CDC (2000). "Outbreak of Rift Valley fever--Saudi Arabia, August-October, 2000." MMWR Morb Mortal Wkly Rep 49(40): 905-8.

CDC (2007). "Rift Valley fever outbreak--Kenya, November 2006-January 2007." MMWR Morb Mortal Wkly Rep 56(4): 73-6.

Coetzer, J. A. and B. J. Barnard (1977). "Hydrops amnii in sheep associated with hydranencephaly and arthrogryposis with wesselsbron disease and rift valley fever viruses as aetiological agents." Onderstepoort J Vet Res 44(2): 119-26.

Collett, M. S. (1986). "Messenger RNA of the M segment RNA of Rift Valley fever virus." Virology 151(1): 151-6.

Dasgupta, A. (2004). "Targeting TFIIH to inhibit host cell transcription by Rift Valley Fever Virus." Mol Cell 13(4): 456-8.

Daubney, R., J. R. Hudson and P. C. Garnham (1931). "Enzootic hepatitis or Rift Valley fever: an undescribed virus disease of sheep, catle and man from East Africa." J Pathol Bacteriol 34: 545-579.

Davies, F. G. (1975). "Observations on the epidemiology of Rift Valley fever in Kenya." J Hyg (Lond) 75(2): 219-30.

Digoutte, J. P. (1981). "Viruses identified from 1963 to 1981." Rapport sur le fonctionement technique de l'Institut Pasteur de Dakar

Digoutte, J. P., R. Cordellier, Y. Robin, F. X. Pajot and B. Geoffroy (1974a). "[Zinga virus (Ar B 1976), a new arbovirus isolated in central africa (author's transl)]." Ann Microbiol (Paris) 125B(1): 107-18.

Digoutte, J. P., J. C. Jacobi, Y. Robin and V. J. Gagnard (1974b). "[Zinga virus infection in man]." Bull Soc Pathol Exot Filiales 67(5): 451-7.

Digoutte, J. P. and C. J. Peters (1989). "General aspects of the 1987 Rift Valley fever epidemic in Mauritania." Res Virol 140(1): 27-30.

Dimmock, N. J., A. J. Easton and K. N. Leppard (2001). Introduction to Modern Virology, Blackwell Science

Dunn, E. F., D. C. Pritlove, H. Jin and R. M. Elliott (1995). "Transcription of a recombinant bunyavirus RNA template by transiently expressed bunyavirus proteins." Virology 211(1): 133-43.

Elfadil, A. A., K. A. Hasab-Allah, O. M. Dafa-Allah and A. A. Elmanea (2006). "The persistence of rift valley fever in the Jazan region of Saudi Arabia." Rev Sci Tech 25(3): 1131-6.

Elliott, R. M. (1996). The Bunyaviridae. New York, Plenum Press.
Eloit, M. (2008). Rift Valley fever, Mayotte (Comoros - France), WAHID.
Emerson, S. U. and R. R. Wagner (1972). "Dissociation and reconstitution of the transcriptase and template activities of vesicular stomatitis B and T virions." $\underline{\mathbf{J}}$ Virol 10(2): 297-309.

Emery, V. C. and D. H. L. Bishop (1987). "Characterization of Punto Toro S mRNA species and identification of an inverted complementary sequence in the intergenic region of Punto Toro phlebovirus ambisense S RNA that is involved in mRNA transcription termination " Virology 156: 1-11.

Eshita, Y., B. Ericson, V. Romanowski and D. H. Bishop (1985). "Analyses of the mRNA transcription processes of snowshoe hare bunyavirus S and M RNA species." J Virol 55(3): 681-9.

Evans, A., F. Gakuya, J. T. Paweska, M. Rostal, L. Akoolo, V. A. N. V. PJ, T. Manyibe, J. M. Macharia, T. G. Ksiazek, D. R. Feikin, R. F. Breiman and M. Kariuki Njenga (2007). "Prevalence of antibodies against Rift Valley fever virus in Kenyan wildlife." Epidemiol Infect: 1-9.

Felsenstein, J. (1985). "Confidence limits on phylogenies: An approach using the bootstrap." Evolution 39: 783-791.

Findlay, G. M. (1931). "Rift Valley fever or enzootic hepatitis " Trans R Soc Trop Med Hyg 25(4): 229-262.

Flick, K., J. W. Hooper, C. S. Schmaljohn, R. F. Pettersson, H. Feldmann and R. Flick (2003). "Rescue of Hantaan virus minigenomes." Virology 306(2): 219-24.

Flick, R. and R. F. Pettersson (2001). "Reverse genetics system for Uukuniemi virus (Bunyaviridae): RNA polymerase I-catalyzed expression of chimeric viral RNAs." J Virol 75(4): 1643-55.

Francis, T. and T. P. Magill (1935). "Rift Valley fever: A report of 3 cases of laboratory infection and the experimental transmission of the disease in ferrets." J Exp Med 62: 433-447.

Freier, S. M., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson and D. H. Turner (1986). "Improved free-energy parameters for predictions of RNA duplex stability." Proc Natl Acad Sci U S A 83(24): 9373-7.

Gentsch, J. and D. H. Bishop (1976). "Recombination and complementation between temperature-sensitive mutants of a Bunyavirus, snowshoe hare virus." $\underline{\text { J Virol }}$ 20(1): 351-4.

Gentsch, J. R. and D. L. Bishop (1979). "M viral RNA segment of bunyaviruses codes for two glycoproteins, G1 and G2." J Virol 30(3): 767-70.

Georges, A. J., S. A. Wahid, D. Y. Meunier, M. C. Georges, J. F. Saluzzo, C. J. Peters, J. B. McCormick and J. P. Gonzalez (1983). "Serological evidence of endemic Zinga virus and Rift Valley Fever virus in Central African Republic." Lancet 1(8337): 1338.

Giorgi, C., L. Accardi, L. Nicoletti, M. C. Gro, K. Takehara, C. Hilditch, S. Morikawa and D. H. Bishop (1991). "Sequences and coding strategies of the S RNAs of Toscana and Rift Valley fever viruses compared to those of Punta Toro, Sicilian Sandfly fever, and Uukuniemi viruses." Virology 180(2): 738-53.

Goodbourn, S., L. Didcock and R. E. Randall (2000). "Interferons: cell signalling, immune modulation, antiviral response and virus countermeasures." J Gen Virol 81(Pt 10): 2341-64.

Hames, B. D. and N. M. Hooper (2000). Amino acids and proteins. Instant notes: Biochemistry. Oxford, BIOS: 19-23.

Hewlett, M. J., R. F. Pettersson and D. Baltimore (1977). "Circular forms of Uukuniemi virion RNA: an electron microscopic study." J Virol 21(3): 108593.

Hoogstraal, H., J. M. Meegan, G. M. Khalil and F. K. Adham (1979). "The Rift Valley fever epizootic in Egypt 1977-78. 2. Ecological and entomological studies." Trans R Soc Trop Med Hyg 73(6): 624-9.

Hutchinson, K. L., C. J. Peters and S. T. Nichol (1996). "Sin Nombre virus mRNA synthesis." Virology 224(1): 139-49.

Iben, S., H. Tschochner, M. Bier, D. Hoogstraten, P. Hozak, J. M. Egly and I. Grummt (2002). "TFIIH plays an essential role in RNA polymerase I transcription." Cell 109(3): 297-306.

Ihara, T., H. Akashi and D. H. Bishop (1984). "Novel coding strategy (ambisense genomic RNA) revealed by sequence analyses of Punta Toro Phlebovirus S RNA." Virology 136(2): 293-306.

Ihara, T., Y. Matsuura and D. H. Bishop (1985). "Analyses of the mRNA transcription processes of Punta Toro phlebovirus (Bunyaviridae)." Virology 147(2): 31725.

Ikegami, T. and S. Makino (2004). "[Rift Valley fever virus]." Uirusu 54(2): 229-35.
Ikegami, T., C. J. Peters and S. Makino (2005a). "Rift valley fever virus nonstructural protein NSs promotes viral RNA replication and transcription in a minigenome system." J Virol 79(9): 5606-15.

Ikegami, T., S. Won, C. J. Peters and S. Makino (2005b). "Rift Valley fever virus NSs mRNA is transcribed from an incoming anti-viral-sense S RNA segment." $\underline{J}$ Virol 79(18): 12106-11.

Ikegami, T., S. Won, C. J. Peters and S. Makino (2007). "Characterization of Rift Valley fever virus transcriptional terminations." J Virol 81(16): 8421-38.

Jaaskelainen, K. M., P. Kaukinen, E. S. Minskaya, A. Plyusnina, O. Vapalahti, R. M. Elliott, F. Weber, A. Vaheri and A. Plyusnin (2007). "Tula and Puumala hantavirus NSs ORFs are functional and the products inhibit activation of the interferon-beta promoter." J Med Virol 79(10): 1527-36.

Jacques, J. P., S. Hausmann and D. Kolakofsky (1994). "Paramyxovirus mRNA editing leads to G deletions as well as insertions." EMBO J 13(22): 5496-503.

Jin, H. and R. M. Elliott (1993). "Non-viral sequences at the 5' ends of Dugbe nairovirus S mRNAs." J Gen Virol 74 ( Pt 10): 2293-7.

Jouan, A., I. Coulibaly, F. Adam, B. Philippe, O. Riou, B. Leguenno, R. Christie, N. Ould Merzoug, T. Ksiazek and J. P. Digoutte (1989). "Analytical study of a Rift Valley fever epidemic." Res Virol 140(2): 175-86.

Jupp, P. G., A. Kemp, A. Grobbelaar, P. Lema, F. J. Burt, A. M. Alahmed, D. Al Mujalli, M. Al Khamees and R. Swanepoel (2002). "The 2000 epidemic of Rift Valley fever in Saudi Arabia: mosquito vector studies." Med Vet Entomol 16(3): 245-52.

Karlen, A. (1996). A garden of germs Plague's progress: a social history of man and disease. London, Cassell Group: 149-173.

Kimura, M. (1980). "A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences " J Mol Evol 16: 111-120.

Kohl, A., A. Bridgen, E. Dunn, J. N. Barr and R. M. Elliott (2003). "Effects of a point mutation in the $3^{\prime}$ end of the $S$ genome segment of naturally occuring and engineered Bunyamwera viruses." J Gen Virol 84: 789-793.

Kormelink, R., F. van Poelwijk, D. Peters and R. Goldbach (1992). "Non-viral heterogeneous sequences at the 5 ' ends of tomato spotted wilt virus mRNAs." J Gen Virol 73 ( Pt 8): 2125-8.

Krug, R. M. (1981). "Priming of influenza viral RNA transcription by capped heterologous RNAs." Curr Top Microbiol Immunol 93: 125-49.

Kuismanen, E., K. Hedman, J. Saraste and R. F. Pettersson (1982). "Uukuniemi virus maturation: accumulation of virus particles and viral antigens in the Golgi complex." Mol Cell Biol 2(11): 1444-58.

Le May, N., S. Dubaele, L. Proietti De Santis, A. Billecocq, M. Bouloy and J. M. Egly (2004). "TFIIH transcription factor, a target for the Rift Valley hemorrhagic fever virus." Cell 116(4): 541-50.

Le May, N., N. Gauliard, A. Billecocq and M. Bouloy (2005). "The N terminus of Rift Valley fever virus nucleoprotein is essential for dimerization." J Virol 79(18): 11974-80.

Le May, N., Z. Mansuroglu, P. Leger, T. Josse, G. Blot, A. Billecocq, R. Flick, Y. Jacob, E. Bonnefoy and M. Bouloy (2008). "A SAP30 complex inhibits IFNbeta expression in Rift Valley fever virus infected cells." PLoS Pathog 4(1): e13.

Linthicum, K. J., F. G. Davies, A. Kairo and C. L. Bailey (1985). "Rift Valley fever virus (family Bunyaviridae, genus Phlebovirus). Isolations from Diptera collected during an inter-epizootic period in Kenya." J Hyg (Lond) 95(1): 197209.

Lopez, N., R. Muller, C. Prehaud and M. Bouloy (1995). "The L protein of Rift Valley fever virus can rescue viral ribonucleoproteins and transcribe synthetic genome-like RNA molecules." J Virol 69(7): 3972-9.

Madani, T. A., Y. Y. Al-Mazrou, M. H. Al-Jeffri, A. A. Mishkhas, A. M. Al-Rabeah, A. M. Turkistani, M. O. Al-Sayed, A. A. Abodahish, A. S. Khan, T. G. Ksiazek and O. Shobokshi (2003). "Rift Valley fever epidemic in Saudi Arabia: epidemiological, clinical, and laboratory characteristics." Clin Infect Dis 37(8): 1084-92.

Marriot, A. C., V. K. Ward and P. A. Nuttal (1989). "The S RNA segment of Sandfly fever Sicilian virus:evidence for an amibisense genome." Virology 169: 341345.

Mathiot, C., J. J. Ribot, Y. Clerc, P. Coulanges and N. Rasolofonirina (1984). "[Rift valley fever and Zinga virus: a pathogenic arbovirus in man and animal new for Madagascar]." Arch Inst Pasteur Madagascar 51(1): 125-33.

Meegan, J. M. (1979). "The Rift Valley fever epizootic in Egypt 1977-78. 1. Description of the epizzotic and virological studies." Trans R Soc Trop Med Hyg 73(6): 618-23.

Meegan, J. M., J. P. Digoutte, C. J. Peters and R. E. Shope (1983). "Monoclonal antibodies to identify Zinga virus as Rift Valley Fever virus." Lancet 1(8325): 641.

Meegan, J. M., H. Hoogstraal and M. I. Moussa (1979). "An epizootic of Rift Valley fever in Egypt in 1977." Vet Rec 105(6): 124-5.

Mellor, P. S. and C. J. Leake (2000). "Climatic and geographic influences on arboviral infections and vectors." Rev Sci Tech 19(1): 41-54.

Miller, W. S., P. Demchak, C. R. Rosenberger, J. W. Dominik and J. L. Bradshaw (1963). "Stability and infectivity of airborne yellow fever and Rift Valley fever viruses " Am J Hyg 77: 114-121.

Morvan, J., D. Fontenille, J. F. Saluzzo and P. Coulanges (1991). "Possible Rift Valley fever outbreak in man and cattle in Madagascar." Trans R Soc Trop Med Hyg 85(1): 108.

Morvan, J., J. L. Lesbordes, P. E. Rollin, J. C. Mouden and J. Roux (1992). "First fatal human case of Rift Valley fever in Madagascar." Trans R Soc Trop Med Hyg 86(3): 320.

Muller, R., C. Argentini, M. Bouloy, C. Prehaud and D. H. Bishop (1991). "Completion of the genome sequence of Rift Valley fever phlebovirus indicates that the L RNA is negative sense and codes for a putative transcriptase-replicase [corrected]." Nucleic Acids Res 19(19): 5433.

Muller, R., J. F. Saluzzo, N. Lopez, T. Dreier, M. Turell, J. Smith and M. Bouloy (1995). "Characterization of clone 13, a naturally attenuated avirulent isolate of Rift Valley fever virus, which is altered in the small segment." Am J Trop Med Hyg 53(4): 405-11.

Murphy, F. A., C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo and M. D. Summers (1995). Family Bunyaviridae. Virus Taxonomy: Sixth Report of the International Commitee on Taxonomy of Viruses. Vienna and New York, Springer-Verlag: 300-315.

Murphy, F. A., A. K. Harrison and S. G. Whitfield (1973). Bunyaviridae: morphologic and morphogenetic similarities of Bunyamwera supergroup viruses and several other arthropod-borne viruses.

Nabeth, P., Y. Kane, M. O. Abdalahi, M. Diallo, K. Ndiaye, K. Ba, F. Schneegans, A. A. Sall and C. Mathiot (2001). "Rift Valley fever outbreak, Mauritania, 1998: seroepidemiologic, virologic, entomologic, and zoologic investigations." Emerg Infect Dis 7(6): 1052-4.

Nichol, S. T., B. J. Beaty, R. M. Elliott, R. Goldbach, A. Plyusnin, C. S. Schmaljohn and R. B. Tesh (2005). Bunyaviridae. Virus Taxonomy: 8th Report of the International Commitee on Taxonomy of Viruses. C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger and L. A. Ball. Oxford, Academic Press: 695-716.

Pardigon, N., P. Vialat, M. Girard and M. Bouloy (1982). "Panhandles and hairpin structures at the termini of germiston virus RNAs (Bunyavirus)." Virology 122(1): 191-7.

Patterson, J. L., B. Holloway and D. Kolakofsky (1984). "La Crosse virions contain a primer-stimulated RNA polymerase and a methylated cap-dependent endonuclease." J Virol 52(1): 215-22.

Paweska, J., L. Blumberg, J. Weyer, A. Kemp, P. Leman, B. Archer, D. Nkosi and R. Swanepoel (2008). "Rift Valley fever outbreak in South Africa, 2008." Communicable Diseases Surveillance Bulletin 6(2): 1-2.

Perrone, L. A., K. Narayanan, M. Worthy and C. J. Peters (2007). "The S segment of Punta Toro virus (Bunyaviridae, Phlebovirus) is a major determinant of lethality in the Syrian hamster and codes for a type I interferon antagonist." $\underline{\mathbf{J}}$ Virol 81(2): 884-92.

Peters, C. J. and K. J. Linthicum (1994). Rift Valley fever. Handbook of zoonoses. Section B: viral zoonoses, 2nd edition. J. S. GW Beran. Boca Raton, CRC Press: 25-138.

Peters, C. J. and T. W. Slone (1982). "Inbred rat strains mimic the disparate human response to Rift Valley fever virus infection." J Med Virol 10(1): 45-54.

Plyusnin, A. (2002). "Genetics of hantaviruses: implications to taxonomy." Arch Virol 147(4): 665-82.

Porter, H. F. (1999). Onward to Africa. Forecast: disaster. The future of El Nino. New York, Dell Publishing: 127-144.

Purse, B. V., P. S. Mellor, D. J. Rogers, A. R. Samuel, P. P. Mertens and M. Baylis (2005). "Climate change and the recent emergence of bluetongue in Europe." Nat Rev Microbiol 3(2): 171-81.

Rakotosamimanan, J. L. H. (2008). Rift Valley fever, Madagascar, WAHID.
Rice, R. M., B. J. Erlick, R. R. Rosato, G. A. Eddy and S. B. Mohanty (1980). "Biochemical characterization of Rift Valley fever virus." Virology 105(1): 256-60.

Ritter, M., M. Bouloy, P. Vialat, C. Janzen, O. Haller and M. Frese (2000). "Resistance to Rift Valley fever virus in Rattus norvegicus: genetic variability within certain 'inbred' strains." $\underline{\underline{\mathrm{J}} \text { Gen Virol 81(Pt 11): 2683-8. }}$

Rosenblum, B. B., L. G. Lee, S. L. Spurgeon, S. H. Khan, S. M. Menchen, C. R. Heiner and S. M. Chen (1997). "New dye-labeled terminators for improved DNA sequencing patterns." Nucleic Acids Res 25(22): 4500-4.

Rozhon, E. J., P. Gensemer, R. E. Shope and D. H. Bishop (1981). "Attenuation of virulence of a bunyavirus involving an L RNA defect and isolation of LAC/SSH/LAC and LAC/SSH/SSH reassortants." Virology 111(1): 125-38.

Saitou, N. and M. Nei (1987). "The neighbor-joining method: a new method for reconstructing phylogenetic trees." Mol Biol Evol 4(4): 406-25.

Sall, A. A., A. Z. P. M. de, H. G. Zeller, J. P. Digoutte, Y. Thiongane and M. Bouloy (1997). "Variability of the NS(S) protein among Rift Valley fever virus isolates." J Gen Virol 78 ( Pt 11): 2853-8.

Sall, A. A., P. M. Zanotto, O. K. Sene, H. G. Zeller, J. P. Digoutte, Y. Thiongane and M. Bouloy (1999). "Genetic reassortment of Rift Valley fever virus in nature." J Virol 73(10): 8196-200.

Sall, A. A., P. M. Zanotto, P. Vialat, O. K. Sene and M. Bouloy (1998). "Molecular epidemiology and emergence of Rift Valley fever." Mem Inst Oswaldo Cruz 93(5): 609-14.

Saluzzo, J. F., G. W. Anderson, Jr., J. F. Smith, D. Fontenille and P. Coulanges (1989). "Biological and antigenic relationship between Rift Valley fever virus strains isolated in Egypt and Madagascar." Trans R Soc Trop Med Hyg 83(5): 701.

Saluzzo, J. F. and J. F. Smith (1990). "Use of reassortant viruses to map attenuating and temperature-sensitive mutations of the Rift Valley fever virus MP-12 vaccine." Vaccine 8(4): 369-75.

Samso, A., M. Bouloy and C. Hannoun (1976). "[Demonstration of circular ribonucleic acid in the Lumbo virus (Bunyavirus)]." C R Acad Sci Hebd Seances Acad Sci D 282(17): 1653-5.

Schmaljohn, C. and J. W. Hooper (2001). Bunyaviridae: the viruses and their replication. Fields Virology. D. K. BN Fields. Philadelphia, Lippincott Williams \& Williams. 2.2.

Shoemaker, T., C. Boulianne, M. J. Vincent, L. Pezzanite, M. M. Al-Qahtani, Y. AlMazrou, A. S. Khan, P. E. Rollin, R. Swanepoel, T. G. Ksiazek and S. T. Nichol (2002). "Genetic analysis of viruses associated with emergence of Rift Valley fever in Saudi Arabia and Yemen, 2000-01." Emerg Infect Dis 8(12): 1415-20.

Shope, R. E., G. H. Tignor, E. J. Rozhon and D. H. L. Bishop (1981). The association of the bunyavirus middle-sized RNA segment with mouse pathogenicity Replication of negative strand viruses D. B. R. Compans. New York, Elsevier/North Holland Publishing Co. : 146-152.

Siam, A. L., J. M. Meegan and K. F. Gharbawi (1980). "Rift Valley fever ocular manifestations: observations during the 1977 epidemic in Egypt." $\underline{\mathrm{Br} \mathrm{J}}$ Ophthalmol 64(5): 366-74.

Simons, J. F., U. Hellman and R. F. Pettersson (1990). "Uukuniemi virus S RNA segment: ambisense coding strategy, packaging of complementary strands into virions, and homology to members of the genus Phlebovirus." J Virol 64(1): 247-55.

Simons, J. F., R. Persson and R. F. Pettersson (1992). "Association of the nonstructural protein NSs of Uukuniemi virus with the 40 S ribosomal subunit." J Virol 66(7): 4233-41.

Simons, J. F. and R. F. Pettersson (1991). "Host-derived 5' ends and overlapping complementary 3 ' ends of the two mRNAs transcribed from the ambisense S segment of Uukuniemi virus." J Virol 65(9): 4741-8.

Smith, J. F. and D. Y. Pifat (1982). "Morphogenesis of sandfly viruses (Bunyaviridae family)." Virology 121(1): 61-81.

Swanepoel, R. and N. K. Blackburn (1977). "Demonstration of nuclear immunofluorescence in Rift Valley fever infected cells." J Gen Virol 34(3): 557-61.

Swanepoel, R. and J. A. W. Coetzer (2004). Rift Valley Fever. Infectious diseases of livestock with special reference to Southern Africa
J. A. W. Coetzer and R. C. Tustin, Oxford Univeristy Press. 2: 1037-1070.

Takeda, A., K. Sugiyama, H. Nagano, M. Mori, M. Kaido, K. Mise, S. Tsuda and T. Okuno (2002). "Identification of a novel RNA silencing suppressor, NSs protein of Tomato spotted wilt virus." FEBS Lett 532(1-2): 75-9.

Tamura, K., J. Dudley, M. Nei and S. Kumar (2007). "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0." Mol Biol Evol 24(8): 1596-9.

Thomas, D., G. Blakqori, V. Wagner, M. Banholzer, N. Kessler, R. M. Elliott, O. Haller and F. Weber (2004). "Inhibition of RNA polymerase II phosphorylation by a viral interferon antagonist." J Biol Chem 279(30): 31471-7.

Turell, M. J., K. J. Linthicum, L. A. Patrican, F. G. Davies, A. Kairo and C. L. Bailey (2008). "Vector competence of selected African mosquito (Diptera: Culicidae) species for Rift Valley fever virus." J Med Entomol 45(1): 102-8.

Turell, M. J., J. F. Saluzzo, R. F. Tammariello and J. F. Smith (1990). "Generation and transmission of Rift Valley fever viral reassortants by the mosquito Culex pipiens." J Gen Virol 71 ( Pt 10): 2307-12.

Uptain, S. M., C. M. Kane and M. J. Chamberlin (1997). "Basic mechanisms of transcript elongation and its regulation." Annu Rev Biochem 66: 117-72.
van Velden, D. J., J. D. Meyer, J. Olivier, J. H. Gear and B. McIntosh (1977). "Rift Valley fever affecting humans in South Africa: a clinicopathological study." $\underline{S}$ Afr Med J 51(24): 867-71.

Venter, M., S. A. Madhi, C. T. Tiemessen and B. D. Schoub (2001). "Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes." J Gen Virol 82(Pt 9): 2117-24.

Vialat, P., A. Billecocq, A. Kohl and M. Bouloy (2000). "The S segment of rift valley fever phlebovirus (Bunyaviridae) carries determinants for attenuation and virulence in mice." J Virol 74(3): 1538-43.

Vialat, P., R. Muller, T. H. Vu, C. Prehaud and M. Bouloy (1997). "Mapping of the mutations present in the genome of the Rift Valley fever virus attenuated MP12 strain and their putative role in attenuation." Virus Res 52(1): 43-50.

Von Bonsdorff, C. H. and R. F. Pettersson (1975). "Surface structure of Uukiniemi virus." J Virol 16(5): 1296-1307.

Weber, F., A. Bridgen, J. K. Fazakerley, H. Streitenfeld, N. Kessler, R. E. Randall and R. M. Elliott (2002). "Bunyamwera bunyavirus nonstructural protein NSs counteracts the induction of alpha/beta interferon." J Virol 76(16): 7949-55.

Weber, F. and R. M. Elliot (2002). "Antigenic drift, antigenic shift and interferon antagonists: how bunyaviruses counteract the immune system." Virus Res 88 : 129-136.

Weinbren, M. P., M. C. Williams and A. J. Haddow (1957). "A variant of Rift Valley fever virus." S Afr Med J 31(38): 951-7.

Weiss, K. (1957). "Rift Valley fever - a review." Bulletin of Epizootic Diseases in Africa 5: 431-458.

Whitehouse, C. (2004). "Crimean-Congo hemorrhagic fever." Antiviral Research 64: 145-160.

Wilson, M. L., L. E. Chapman, D. B. Hall, E. A. Dykstra, K. Ba, H. G. Zeller, M. Traore-Lamizana, J. P. Hervy, K. J. Linthicum and C. J. Peters (1994). "Rift Valley fever in rural northern Senegal: human risk factors and potential vectors." Am J Trop Med Hyg 50(6): 663-75.

Woods, C. W., A. M. Karpati, T. Grein, N. McCarthy, P. Gaturuku, E. Muchiri, L. Dunster, A. Henderson, A. S. Khan, R. Swanepoel, I. Bonmarin, L. Martin, P. Mann, B. L. Smoak, M. Ryan, T. G. Ksiazek, R. R. Arthur, A. Ndikuyeze, N. N. Agata and C. J. Peters (2002). "An outbreak of Rift Valley fever in Northeastern Kenya, 1997-98." Emerg Infect Dis 8(2): 138-44.

Yadani, F. Z., A. Kohl, C. Prehaud, A. Billecocq and M. Bouloy (1999). "The carboxy-terminal acidic domain of Rift Valley Fever virus NSs protein is essential for the formation of filamentous structures but not for the nuclear localization of the protein." J Virol 73(6): 5018-25.

Zeller, H. G., D. Fontenille, M. Traore-Lamizana, Y. Thiongane and J. P. Digoutte (1997). "Enzootic activity of Rift Valley fever virus in Senegal." Am J Trop Med Hyg 56(3): 265-72.

## Chapter 6: Appendix

### 6.1. Preparation of reagents

### 6.1.1. EMEM

To 500 ml EMEM add $1 \%$ L-Glutamine ( 5 ml ), $1 \%$ Non-essential amino acids ( 5 ml ), $1 \%$ Pen/Strep antibiotics ( 5 ml ) and $2.5 \%$ Amphotericin B ( 12.5 ml ). For the preparation of growth media, $10 \%$ foetal calf serum was added, and for maintenance medium, $2 \%$ foetal calf serum was added.

### 6.1.2. PBS

To 500 ml PBS add 5 ml Pen/Strep antibiotics

### 6.1.3. TBE Buffer

To make a 10X TBE buffer, 108 g of Tris base and 55 g of Boric acid were dissolved in one litre of distilled water $\left(\mathrm{dH}_{2} \mathrm{O}\right)$. To this 40 ml of 0.5 M EDTA ( pH 8.0 ) was added. The buffer was transferred to a 1 litre Schott bottle, autoclaved and stored at room temperature.

### 6.2. Routine laboratory methodologies

### 6.2.1. RNA extraction using QIAamp kit

Viral RNA was extracted from infective tissue culture using a QIAamp Viral RNA kit (Qiagen, CA, USA), which combines the selective binding properties of a silica-gel-based membrane with the speed of microspin. The infective tissue culture supernatant was added to $560 \mu 1$ of AVL buffer (provided in kit), containing $5.6 \mu 1$ carrier RNA (provided in kit), in a 1.5 ml microcentrifuge tube. The mixture was then mixed by pulse vortexing and incubated for 10 minutes at room temperature to allow for lysis to occur. The buffer AVL provides highly denaturing conditions to inactivate RNases and to ensure isolation of intact viral RNA. After ten minutes the mixture is briefly centrifuged to remove any droplets from the sides of the tube. A dilution of $96 \%$ ethanol in then added to the sample ( $560 \mu \mathrm{l}$ ) and then mixed by pulse-vortexing and centrifuged to remove droplets. Half of this mixute, $630 \mu 1$, was then applied to the QIAamp spin column in a 2 ml collection tube and centrifuge at $6000 \mathrm{x} g$ for 1 minute. This step was then repeated with the remaining mixture and a clean 2 ml collection tube. After transferring the QIAamp spin column to another new collection tube, $500 \mu 1$ of Buffer AW1 (provided in kit) was added and the tube centrifuged at $6000 \mathrm{x} g$ for 1 minute. This was followed by adding $500 \mu \mathrm{l}$ Buffer AW2 (provided in kit) to the QIAamp spin column in a new 2 ml collection tube and centrifugation at $20000 \mathrm{x} g$ for 3 minutes. The QIAamp spin column was then transferred to a clean 1.5 ml microcentrifuge tube. The the QIAamp spin column, 60 $\mu 1$ Buffer AVE (provided in kit) was added, incubated at room temperature for 1 minute and then centrifuged at $6000 \times g$ for 1 minute. The QIAamp spin column was then discarded and the eluted viral RNA was then stored at $-20^{\circ} \mathrm{C}$ or used immediately.

### 6.2.2. 1.2\% Agarose gel

For a $1.2 \%$ agarose gel, 1.2 g of agarose was weighed into a clean 500 ml Schott bottle and constituted with $100 \mathrm{ml} 1 \times$ TBE buffer. The mixture was boiled in the microwave to completely dissolve the agarose. After allowing the mixture to cool slightly, $0.5 \mu \mathrm{~g} / \mathrm{ml}$ of ethidium bromide ( EtBr ) was aliquoted into the mixture. EtBr is an intercalating agent that can insert into the DNA double helix and will fluoresce when exposed to ultraviolet (UV) wavelengths. When mixed, the liquid was poured into a casting tray and allowed to set until firm. The gel was then transferred into an electrophoresis chamber filled with $1 \times$ TBE buffer. The samples to be subjected to electrophoresis were mixed with loading dye at a ratio of $1: 5$ and loaded into the wells of the gel. In the first lane of the gel, $2 \mu \mathrm{l}$ of 1 kb molecular weight marker was run to aid in determining the length of the DNA fragments. The DNA fragments were viewed under UV light and the resultant image captured.

### 6.2.3. PCR product clean-up using the Wizard ${ }^{\circledR}$ SV gel and PCR clean-up system

To each PCR reaction, an equal volume of Membrane Binding solution (provided in kit) is added and mixed. The mixture is then transferred into an SV minicolumn, which is in a collection tube, and incubate at room temperature for 1 minute. The SV minicolumn assembly is then centrifuged at $10000 \mathrm{x} g$ for 1 minute, after which the flow-through is discarded and the minicolumn re-inserted into the collection tube. To each minicolumn, $700 \mu \mathrm{l}$ Membrane Wash solution (provided in kit) is added, and then the minicolumn is centrifuged at $10000 \times g$ for 1 minute, the flowthrough discarded and the minicolumn is re-inserted into the collection tube. This is then repeated using $500 \mu \mathrm{l}$ of Membrane Wash solution. The minicolumn is then centrifuged for an extra minute at $10000 \times g$ to remove any excess solution. The minicolumn is then transferred into a sterile 1.5 ml centrifuge tube. To each minicolumn, $30-50 \mu \mathrm{l}$ of nuclease free water is added and allowed to incubate at room temperature for 1 minute, after which it is centrifuged at $10000 \times g$ for 1 minute. The minicolumn is then discarded and the DNA can be stored at $4^{\circ} \mathrm{C}$ or $20^{\circ} \mathrm{C}$, or alternatively used immediately.

### 6.2.4. Isopropanol purification of cycle sequencing product

To each well of the 96 -well plate containing cycle sequenced product, $80 \mu 1$ of $80 \%$ isopropanol is added and allowed to stand for 15 minutes. The 96 -well plate is then centrifuged for 45 minutes at $2000 \times g$. The 96 -well plate is then flipped to drain the isopropanol, taking care not to loose the cycle sequencing product. The plate is then centrifuged upside down for 1 minute at $750 \mathrm{x} g$ to drain excess isopropanol. The plate is then allowed to air-dry for 5 minutes to evaporate any remaining isopropanol. To the dried cycle sequencing product, $20 \mu 1$ of Hi-Di ${ }^{\text {TM }}$ Formamide (Applied Biosystems, USA) is added. Hi-Di ${ }^{\text {тм }}$ Formamide is a highly deionized formamide that contains a stabilizer and is used as an injection solvent in DNA analysis on the ABI Prism® DNA analyzer.

### 6.3. Full S RNA segment sequencing data

1
40
214445 KEN83
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT

56 KEN65
900085 MAU88 An1000MAD91
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAACGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCTAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT
(1) ACACAAAGACCCCCTAGTGCTTATCAAGTATATCATGGAT

Figure 17. The complete nucleotide sequence alignment of the $S$ RNA segment of the 45 strains of Rift
Valley fever virus sequenced in this study. Areas with black lettering and white background show identical nucleotides. Areas with black lettering and yellow background represent non-similar nucleotides.

214445 KEN 83 (41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC 56KEN65 (41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
900085MAU88
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC
(41) TACTTTCCTGTGATATCTGTTGATTTGCAGAGTGGTCGTC

Figure 17. Continued

```
                                    81
                                    120
    214445KEN83 (81) GTGTTGTGTCAGTGGAGTACATTATAGGTGATGGTCCTCC
        56KEN65 (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
    900085MAU88
    An1000MAD91
    An991MAD91
    An999MAD91
    AnK6087GUI84
AnTAMBULEGY94
    Ar12568RSA71
    Ar20368RSA81
        Ar21229SA00
            Ar74RSA55
        Ar811MAD79
ArD38661SEN81
        B1143KEN77
            B314KEN62
        H1825RSA75
        LunyoUGA55
        R1662CAR85
SPU10301KEN07
SPU10302KEN07
SPU10307KENO7
SPU10315KEN07
SPU12002SOM98
    SPU152RSA08
    SPU204ANGL85
    SPU2207KENO7
    SPU2214KEN07
    SPU2215KEN07
    SPU2220KEN07
    SPU2223KEN07
SPU384001KEN97
        SPU45ZAMB85
SPU52001RSA99
    SPU77NAMB04
        SPU77RSA08
    VRL1032ZIM78
    VRL1290ZIM78
    VRL1516ZIM78
    VRL1887ZIM78
    VRL2230ZIM78
        VRL763ZIM70
        VRL825ZIM79
        ZH501EGY77
        ZH548EGY77
            Consensus
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACTTTAGAGGAGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGCCCTCC
        (81) GTGTTGTGTCAGTGGAGTACTTTAGAGGAGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTATATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGCCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGAGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAAAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACTTTAGAGGAGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGCCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGCCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACTTTAGAGGAGATGGTCCTCC
        (81) GTGTTGTGTCAGTGGAGTACTTTAGAGGAGATGGTCCTCC
    (81) GTGTTGTGTCAGTGGAGTACATTAGAGGTGATGGTCCTCC
```

Figure 17. Continued

|  |  | 121160 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| $56 \mathrm{KEN65}$ | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| $900085 \mathrm{MAU8} 8$ | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| An1000MAD91 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| An991MAD91 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| An999MAD91 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| AnK6087GUI84 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| AnTAMBULEGY94 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| Ar12568RSA71 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| Ar20368RSA81 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| Ar21229SA00 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| Ar74RSA55 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| Ar811MAD79 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| ArD38661SEN81 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| B1143KEN77 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| B314KEN62 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| H1825RSA75 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| LunyoUGA55 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| R1662CAR85 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU10301KEN07 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU10302KEN0 7 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU10307KEN07 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU10315KEN07 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU12002SOM98 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU152RSA08 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU204ANGL85 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU2207KEN07 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU2214KEN07 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU2215KEN07 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU2220KEN07 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU2223KEN07 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU384001KEN97 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU45ZAMB85 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU52001RSA99 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTAT |
| SPU77NAMB0 4 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| SPU77RSA08 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| VRL1032 $21 M 78$ | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| VRL1290ZIM78 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| VRL1516ZIM78 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| VRL1887ZIM78 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| VRL2230ZIM78 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| VRL763ZIM70 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| VRL825ZIM79 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| ZH501EGY77 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| ZH548EGY77 | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |
| Consensus | (121) | CAGGATACCTTATTCTATGGTTGGGCCCTGTTGTGTCTTT |

Figure 17. Continued

|  |  | 161200 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| $56 \mathrm{KEN65}$ | (161) | CTCATGCGCCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| $900085 \mathrm{MAU8} 8$ | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| An1000MAD91 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| An991MAD91 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| An999MAD91 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| AnK6087GUI84 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| AnTAMBULEGY94 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| Ar12568RSA71 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| Ar20368RSA81 | (161) | CTCATGCACCATCGTCCTAGTTACGAGGTTCGCTTGCGAT |
| Ar21229SA00 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTCCGCTTGCGAT |
| Ar 74 RSA 55 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| Ar811MAD79 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| ArD38661SEN81 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| B1143KEN77 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| B314KEN62 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| H1825RSA75 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| LunyoUGA55 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| R1662CAR85 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU10301KEN07 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU10302KEN0 7 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU10307KEN07 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU10315KEN07 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU12002SOM98 | (161) | CTCATGCACCATCGTCCTAGTCATGAGGTTCGCTTGCGAT |
| SPU152RSA08 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU204ANGL85 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU2207KEN07 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU2214KEN07 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU2215KEN07 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU2220KEN07 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU2223KEN07 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU384001KEN97 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU45ZAMB85 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU52001RSA99 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU77NAMB0 4 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| SPU77RSA08 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| VRL1032 ${ }^{\text {LIM7 }} 8$ | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| VRL1290ZIM78 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| VRL1516ZIM78 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| VRL1887ZIM78 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| VRL2230 2 IM78 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| VRL763ZIM70 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| VRL825ZIM79 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| ZH501EGY77 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| ZH548EGY77 | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |
| Consensus | (161) | CTCATGCACCATCGTCCTAGTCACGAGGTTCGCTTGCGAT |

Figure 17. Continued

```
201
    240
```

214445 KEN83 56KEN65
900085 MAU88 An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94 Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar74RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA08 SPU204ANGL85 SPU2207KENO 7 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus

201
240
(201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCCTACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGCCGGAGAATTCCCATACCGAGT (201) TCTCTGACTTCTATAATGTCGGAGAATTTCCATACCGAGT (201) TCTCTGATTTCTACAATGTTGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCCTACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGGGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATGCCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTATAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTTGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT (201) TCTCTGATTTCTACAATGTCGGAGAATTCCCATACCGAGT

Figure 17. Continued

|  |  | 241280 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| $56 \mathrm{KEN65}$ | (241) | CGGACTTGGAGACTTTGCATCAAGCGTTGCACCTCCACCA |
| $900085 \mathrm{MAU8} 8$ | (241) | CGGGCTTGGAGACTTTGCATCAAATGTTGCACCTCCACCA |
| An1000MAD91 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| An991MAD91 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| An999MAD91 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| AnK6087GUI84 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| AnTAMBULEGY94 | (241) | TGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| Ar12568RSA71 | (241) | CGGACTTGGAGACTTTGTATCAAACGTTGCACCTCCACCA |
| Ar20368RSA81 | (241) | CGGACTTGGAGACTTTGCATCAAATGTTGCACCTCCACCA |
| Ar21229SA00 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| Ar 74 RSA55 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| Ar811MAD79 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| ArD38661SEN81 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| B1143KEN77 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| B314KEN62 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| H1825RSA75 | (241) | CGGACTTGGAGACTTTGTATCAAACGTTGCACCTCCACCA |
| LunyoUGA55 | (241) | TGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| R1662CAR85 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU10301KEN07 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU10302KEN07 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU10307KEN07 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU10315KEN07 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU12002SOM98 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU152RSA08 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU20 4ANGL85 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU2207KEN07 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU2214KEN07 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU2215KEN07 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU2220KEN07 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU2223KEN07 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU384001KEN97 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU45ZAMB85 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU52001RSA99 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU77NAMB04 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| SPU77RSA08 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| VRL1032 ZIM78 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| VRL1290ZIM78 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| VRL1516ZIM78 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| VRL1887ZIM78 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| VRL2230ZIM78 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| VRL763ZIM70 | (241) | CGGACTTGGAGACTTTGCGTCAAACGTTGCACCTCCACCA |
| VRL825ZIM79 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| ZH501EGY77 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| ZH548EGY77 | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |
| Consensus | (241) | CGGACTTGGAGACTTTGCATCAAACGTTGCACCTCCACCA |

Figure 17. Continued

|  |  | 281320 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| $56 \mathrm{KEN65}$ | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| $900085 \mathrm{MAU8} 8$ | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGTCATA |
| An1000MAD91 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| An9 91MAD91 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| An999MAD91 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| AnK6087GUI84 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| AnTAMBULEGY94 | (281) | GCGAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| Ar12568RSA71 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| Ar20368RSA81 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| Ar21229SA00 | (281) | GCAAAGCCTTTCCAGAGACTTATTGATCTAATAGGCCATA |
| Ar74RSA55 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTGATAGGCCATA |
| Ar811MAD79 | (281) | GCGAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| ArD38661SEN81 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| B1143KEN77 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| B314KEN62 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTGATAGGCCATA |
| H1825RSA75 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| LunyoUGA55 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCACA |
| R1662CAR85 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATTTAATAGGCCATA |
| SPU10301KEN07 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU10302KEN07 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU10307KEN07 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU10315KEN07 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU12002SOM98 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU152RSA08 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU204ANGL85 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU2207KEN07 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU2214KEN07 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU2215KEN07 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU2220KEN07 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU2223KEN07 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU384001KEN97 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATGGGCCATA |
| SPU45ZAMB85 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATTGGCCATA |
| SPU52001RSA99 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| SPU77NAMB04 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTGATAGGCCATA |
| SPU77RSA08 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| VRL1032 ZIM78 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATTTAATAGGCCATA |
| VRL1290ZIM78 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| VRL1516ZIM78 | (281) | GCGAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| VRL1887ZIM78 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| VRL2230ZIM78 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| VRL763ZIM70 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| VRL825ZIM79 | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| ZH501EGY77 | (281) | GCGAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| ZH548EGY77 | (281) | GCGAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |
| Consensus | (281) | GCAAAGCCTTTTCAGAGACTTATTGATCTAATAGGCCATA |

Figure 17. Continued

|  |  | 321360 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| $56 \mathrm{KEN65}$ | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| $900085 \mathrm{MAU8} 8$ | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| An1000MAD91 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| An991MAD91 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| An9 99MAD91 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| AnK6087GUI84 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| AnTAMBULEGY94 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| Ar12568RSA71 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| Ar20368RSA81 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| Ar21229SA00 | (321) | TGACTCTAAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| Ar74RSA55 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| Ar811MAD79 | (321) | TGACTCTTAGTGACTTCACAAGGTTCCCCAATCTGAAAGA |
| ArD38661SEN81 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAGAGA |
| B1143KEN77 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| B314KEN62 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| H1825RSA75 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| LunyoUGA55 | (321) | TGACTCTTAGTGATTTCATAAGGTTCCCCAATCTGAAAGA |
| R1662CAR85 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| SPU10301KEN07 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| SPU10302KEN0 7 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| SPU10307KEN07 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| SPU10315KEN07 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| SPU12002SOM98 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| SPU152RSA08 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| SPU204ANGL85 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| SPU2207KEN07 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| SPU2214KEN07 | (321) | TGACTCTCAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| SPU2215KEN07 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| SPU2220KEN07 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| SPU2223KEN07 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| SPU384001KEN97 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATTTAAAAGA |
| SPU45ZAMB85 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| SPU52001RSA99 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| SPU77NAMB0 4 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| SPU77RSA08 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTAAAAGA |
| VRL1032 ZIM78 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| VRL1290ZIM78 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| VRL1516ZIM78 | (321) | TGACTCTTAGTGACTTCACAAGGTTCCCCAATCTGAAAGA |
| VRL1887ZIM78 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| VRL2230ZIM78 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| VRL763ZIM70 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| VRL825ZIM79 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| ZH501EGY77 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| ZH548EGY77 | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |
| Consensus | (321) | TGACTCTTAGTGATTTCACAAGGTTCCCCAATCTGAAAGA |

Figure 17. Continued

|  |  | 361400 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| $56 \mathrm{KEN65}$ | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCCCTGGCTTTC |
| $900085 \mathrm{MAU8} 8$ | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCCCTGGCTTTC |
| An1000MAD91 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| An991MAD91 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| An999MAD91 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| AnK6087GUI84 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCATTGGCTTTC |
| AnTAMBULEGY94 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| Ar12568RSA71 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCCCTGGCTTTC |
| Ar20368RSA81 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| Ar21229SA00 | (361) | AGCCATATCCTGGCCTCTTGGAGAGCCCTCACTGGCTTTC |
| Ar 74 RSA 55 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTAGCTTTC |
| Ar811MAD79 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| ArD38661SEN81 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCTCTGGCTTTC |
| B1143KEN77 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| B314KEN62 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCCCTGGCTTTC |
| H1825RSA75 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCCCTGGCTTTC |
| LunyoUGA55 | (361) | AGCCATATCCTGGCCTCTTGGAGAGCCCTCCCTAGCTTTC |
| R1662CAR85 | (361) | AGCCATATCCTGGCCTCTTGGCGAACCCTCACTGGCTTTC |
| SPU10301KEN07 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTAGCTTTC |
| SPU10302KEN0 7 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTAGCTTTC |
| SPU10307KEN07 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTAGCTTTC |
| SPU10315KEN07 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTAGCTTTC |
| SPU12002SOM98 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| SPU152RSA08 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTAGCTTTC |
| SPU204ANGL85 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| SPU2207KEN07 | (361) | AGCCATATCCTGGCCTCTTGGGGAACCCTCACTGGCTTTC |
| SPU2214KEN07 | (361) | AGCTATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| SPU2215KEN07 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTAGCTTTC |
| SPU2220KEN07 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTAGCTTTC |
| SPU2223KEN07 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTAGCTTTC |
| SPU384001KEN97 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCTTCACTGGCTTTC |
| SPU45ZAMB85 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| SPU52001RSA99 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| SPU77NAMB0 4 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| SPU77RSA08 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTAGCTTTC |
| VRL1032 ${ }^{\text {LIM7 }} 8$ | (361) | AGCCATATCCTGGCCCCTTGGAGAACCCTCACTGGCTTTC |
| VRL1290ZIM78 | (361) | AGCCATGTCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| VRL1516ZIM78 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTAGCTTTC |
| VRL1887ZIM78 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| VRL2230ZIM78 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| VRL763ZIM70 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCCCTGGCTTTC |
| VRL825ZIM79 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| ZH501EGY77 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| ZH548EGY77 | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |
| Consensus | (361) | AGCCATATCCTGGCCTCTTGGAGAACCCTCACTGGCTTTC |

Figure 17. Continued

```
                    401
                    440
```

214445 KEN83 56KEN65
900085 MAU88 An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94 Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar74RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA08 SPU204ANGL85 SPU2207KENO 7 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07

SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus

401
440
(401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA
(401) TTTGACCTAAGCTCTACCAGAGTGCACAGGTCTGATGATA
(401) TTTGACCTGAGCTCCACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGATCTTAGCTCTACTAGAGTGCACAGGAATGATGATA (401) TTTGACCTAAGCTCTACTAGAGTGCATAGGAATGATGACA (401) TTTGACCTAAGCTCCACCAGAGTGCATAGGTCTGATGATA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTGAGCTCCACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTACACAGGAATGATGACA (401) TTTGACCTAAGCTCCACCAGAGTGCATAGGTCTGATGATA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGATCTGAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGACGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGACGACA (401) TTTGACCTAAGCTCCACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGATCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGATCTAAGCTCCACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCCACCAGAGTGCATAGGTCTGATGATA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCATAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCATAGGAATGATGACA (401) TTTGACCTAAGCTCTACTAGAGTGCACAGGAATGATGACA

Figure 17. Continued

```
                    441
                    480
```

$214445 \mathrm{KEN8} 3$ 56KEN65
900085 MAU88 An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94 Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar74RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA08 SPU204ANGL85 SPU2207KENO 7 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus
(441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGACCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGGAGGGACCAGATTGCCACTCTAGCAATGAGGAGTTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGTTG (441) TTAGAAGGGACCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGTTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGAGATCAGATTGCCACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGTTG (441) TTAGAAGGGACCAGATTGCCACTCTAGCAATGAGGAGTTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGACCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGGAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGGAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TCAGGAGGGATCAGATTGCCACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGTTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGACCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGTTG (441) TTAGAAGGGATCAGATTGCCACTCTAGCAATGAGGAGTTG (441) TTAGAAGGGATCAGATTGCTACTCTAGCAATGAGGAGCTG

Figure 17. Continued

```
                    41
                    5 2 0
```

214445 KEN83 56KEN65
900085MAU88 An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94 Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar74RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA08 SPU204ANGL85 SPU2207KEN07 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus

481
520
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGGTCACCAATGATCTTGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTGCCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTGCCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTGCCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATCACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTGGAGGACTCCTTTGTTGGCTTA
(481) CAAGATCACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATCACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATCACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATCACCAATGATCTACAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTGGAAGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTGGAGGACTCCTTTGTTGGCTTA (481) CAAGATTACCAATGATCTAGAAGACTCCTTTGTTGGCCTG (481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGATTA
(481) CAAGATTACCAATGATTTAGAAGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATTTAGAAGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATTTAGAAGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATTTAGAAGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATTTAGAAGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATTTAGAAGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATTTAGAAGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATTTAGAAGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATCACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATTTAGAAGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATCACCAATGATCTAGAGGACTCCTTTGTTGGCTTA (481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA (481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA (481) CAAGATTACCAATGATCTGGAGGACTCCTTTGTTGGCTTA (481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA (481) CAAGATCACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATCACCAATGATCTAGAGGACTCCTTTGTTGGCTTA
(481) CAAGATTACCAATGATCTAGAGGACTCCTTTGTTGGCTTA

Figure 17. Continued

|  |  | 521560 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| $56 \mathrm{KEN65}$ | (521) | CACAGGATGATAGTGACTGAGGCTATCCTCAGAGGGATTG |
| $900085 \mathrm{MAU8} 8$ | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| An1000MAD91 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| An991MAD91 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| An9 99MAD91 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| AnK6087GUI84 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| AnTAMBULEGY94 | (521) | CACAGGATGATAGCGACTGAGGCCATCCTCAGAGGGATTG |
| Ar12568RSA71 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| Ar20368RSA81 | (521) | CACAGGAAGATAGTGACCGAGGCCATCCTCAGAGGGATTG |
| Ar21229SA00 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| Ar74RSA55 | (521) | CACAGAATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| Ar811MAD79 | (521) | CACAGGATGATAGCGACCGAGGCCATCCTCAGAGGGATTG |
| ArD38661SEN81 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| B1143KEN77 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| B314KEN62 | (521) | CACAGGATGATAGTGACCGAGGCTATTCTCAGAGGGATTG |
| H1825RSA75 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| LunyoUGA55 | (521) | CACAGGATGATAGTGACCGAGGCTATTCTCAGAGGGGTTG |
| R1662CAR85 | (521) | CACAGGATGGTAGTGACTGAGGCCATCCTCAGAGGGATTG |
| SPU10301KEN07 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU10302KEN07 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU10307KEN07 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU10315KEN07 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU12002SOM98 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU152RSA08 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU204ANGL85 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU2207KEN07 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU2214KEN07 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU2215KEN07 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU2220KEN07 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU2223KENO7 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU384001KEN97 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU45ZAMB85 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGAATTG |
| SPU52001RSA99 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| SPU77NAMB0 4 | (521) | CACAGGATGATAGTGACCGAGGCCATCCTCAGAGGGATTG |
| SPU77RSA08 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| VRL1032 ZIM78 | (521) | CACAGGATAATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| VRL1290ZIM78 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| VRL1516ZIM78 | (521) | CACAGGATGATAGCGACCGAGGCCATCCTCAGAGGGATTG |
| VRL1887ZIM78 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| VRL2230ZIM78 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| VRL763ZIM70 | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| VRL825ZIM79 | (521) | CACAGGAAGATAGTGACCGAGGCTATCCTCAGAGGGATTG |
| ZH501EGY77 | (521) | CACAGGATGATAGCGACTGAGGCCATCCTCAGAGGGATTG |
| ZH548EGY77 | (521) | CACAGGATGATAGCGACTGAGGCCATCCTCAGAGGGATTG |
| Consensus | (521) | CACAGGATGATAGTGACCGAGGCTATCCTCAGAGGGATTG |

Figure 17. Continued

214445 KEN83 56KEN65
900085MAU88 An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84
AnTAMBULEGY94
Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar74RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA0 8 SPU204ANGL85 SPU2207KEN07 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07 SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus
(561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT
(561) ACCTGTGCCTGTTGCCAGGCTTTGATCTCATGTATGAGGT
(561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGATCTCATGTATGAGGT (561) ACTTGTGCCTGTTGCCAGGCTTTGATCTCATGTATGAGGT
(561) ATCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGCTGCCAGGCTTTGACCTCATGTATGAGAT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGATCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ATCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACTTGTGCCTGTTGCCAGGCTTTGATCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT
(561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCGTGTATGAGGT
(561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGATCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT
(561) ACCTGTGCCTGTTGCCAGGCTTTGATCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACTTGTGCCTGTTGCCAGGCTTTGATCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGATCTCATGTATGAGGT (561) ACCTGTGCCTGTTGCCAGGCTTTGATCTCATGTATGAGGT
(561) ACCTGTGCCTGTTGCCAGGCTTTGACCTCATGTATGAGGT

Figure 17. Continued

|  |  | 601640 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| $56 \mathrm{KEN65}$ | (601) | TGCTCACGTTCAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| $900085 \mathrm{MAU8} 8$ | (601) | TGCTCATGTTCAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| An1000MAD91 | (601) | TGCTCACGTACAGTGTGTTCGACTCCTGCAGGCAGCAAAA |
| An991MAD91 | (601) | TGCTCACGTACAGTGTGTTCGACTCCTGCAGGCAGCAAAA |
| An999MAD91 | (601) | TGCTCACGTACAGTGTGTTCGACTCCTGCAGGCAGCAAAA |
| AnK6087GUI84 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| AnTAMBULEGY94 | (601) | TGCTCACGTACAGTGCGTTCGGCTTCTGCAAGCAGCAAAA |
| Ar12568RSA71 | (601) | TGCTCACGTTCAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| Ar20368RSA81 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| Ar21229SA00 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| Ar74RSA55 | (601) | TGCTCATGTTCAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| Ar811MAD79 | (601) | TGCTCACGTACAGTGCGTTCGGCTTCTGCAAGCAGCAAAA |
| ArD38661SEN81 | (601) | TGCTCATGTTCAGTGTGTCCGGCTCCTGCAGGCAGCAAAA |
| B1143KEN77 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| B314KEN62 | (601) | TGCTCACGTCCAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| H1825RSA75 | (601) | TGCTCACGTTCAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| LunyoUGA55 | (601) | TGCTCACGTCCAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| R1662CAR85 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| SPU10301KEN07 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| SPU10302KEN07 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| SPU10307KEN07 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| SPU10315KEN07 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| SPU12002SOM98 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| SPU152RSA08 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| SPU204ANGL85 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| SPU2207KEN07 | (601) | TGCTCACGTGCAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| SPU2214KEN07 | (601) | TGCTCACGTGCAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| SPU2215KEN07 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| SPU2220KEN07 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| SPU2223KEN07 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| PU384001KEN97 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| SPU45ZAMB85 | (601) | TGCTCACGTGCAGTGTGTACGGCTCCTGCAGGCAGCAAAA |
| SPU52001RSA99 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| SPU77NAMB04 | (601) | TGCTCATGTACAGTGTGTTAGGCTCCTGCAAGCAGCAAAA |
| SPU77RSA08 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| VRL1032 ZIM78 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| VRL1290ZIM78 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| VRL1516ZIM78 | (601) | TGCTCACGTACAGTGCGTTCGGCTTCTGCAAGCAGCAAAA |
| VRL1887ZIM78 | (601) | TGCTCACGTACAGTGTGTACGGCTCCTGCAGGCAGCAAAA |
| VRL2230ZIM78 | (601) | TGCTCACGTACAGTGTGTACGGCTCCTGCAGGCAGCAAAA |
| VRL763ZIM70 | (601) | TGCTCACGTTCAGTGTGTTCGGCTCCTGCAGGCAGCAAGA |
| VRL825ZIM79 | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |
| ZH501EGY77 | (601) | TGCTCACGTACAGTGCGTTCGGCTTCTGCAAGCAGCAAAA |
| ZH548EGY77 | (601) | TGCTCACGTACAGTGCGTTCGGCTTCTGCAAGCAGCAAAA |
| Consensus | (601) | TGCTCACGTACAGTGTGTTCGGCTCCTGCAGGCAGCAAAA |

Figure 17. Continued

|  |  | 641680 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| $56 \mathrm{KEN65}$ | (641) | GAGGATATTTCTAATGCTGTAGTTCCAAACTCAGCTCTCA |
| $900085 \mathrm{MAU8} 8$ | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| An1000MAD91 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| An991MAD91 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| An9 99MAD91 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| AnK6087GUI84 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCTGGCCTCA |
| AnTAMBULEGY94 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| Ar12568RSA71 | (641) | GAGGATATTTCTAATGCTGTAGTTCCAAACTCAGCTCTCA |
| Ar20368RSA81 | (641) | GAGGACATTTCTAATGCTGTGGCTCCAAACTCAGCCCTCA |
| Ar21229SA00 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| Ar74RSA55 | (641) | GAGGACATCTCTAATGCTGTAGTCCCAAACTCAGCCCTTA |
| Ar811MAD79 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| ArD38661SEN81 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| B1143KEN77 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| B314KEN62 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| H1825RSA75 | (641) | GAGGATATTTCTAATGCTGTAGTTCCAAACTCAGCTCTCA |
| LunyoUGA55 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| R1662CAR85 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCTGCCCTCA |
| SPU10301KEN07 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU10302KEN0 7 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU10307KEN07 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU10315KEN07 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU12002SOM98 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU152RSA08 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU204ANGL85 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU2207KEN07 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAATTCAGCCCTCA |
| SPU2214KEN07 | (641) | GAAGATATTTCTAATGCTGTAGTTCCAAATTCAGCCCTCA |
| SPU2215KEN07 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU2220KEN07 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU2223KEN07 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU384001KEN97 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU45ZAMB85 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU52001RSA99 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU77NAMB0 4 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| SPU77RSA08 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| VRL1032 7 IM78 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCTGCCCTCA |
| VRL1290ZIM78 | (641) | GAGGACATTTCCAATGCTGTAGTTCCAAACTCAGCCCTCA |
| VRL1516ZIM78 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| VRL1887ZIM78 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| VRL2230ZIM78 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| VRL763ZIM70 | (641) | GAGGATATTTCTAATGCTGTAGTTCCAAACTCAGCTCTCA |
| VRL825ZIM79 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| ZH501EGY77 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| ZH548EGY77 | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |
| Consensus | (641) | GAGGACATTTCTAATGCTGTAGTTCCAAACTCAGCCCTCA |

Figure 17. Continued

```
6 8 1
7 2 0
```

$214445 K E N 83$ 56KEN65
900085 MAU88 An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94

Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar 74 RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN0 7 SPU10307KEN07 SPU10315KENO 7 SPU12002SOM98 SPU152RSA0 8 SPU204ANGL85 SPU2207KENO 7 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA0 8 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus

|  | 681 |
| :---: | :---: |
| 1) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| 681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGTTCATCACT |
| 681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| 681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| 681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| 681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| 681) | TTGTTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCTTGATGCTGCGCTCATCACT |
| 681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGTCTGATGCTGCGCTCATCACT |
| (681) | TTGTTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| 681) | TTGCTCTTATGGAGGAGAGCTTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTTATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTTATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCACATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTAATGCTGCGTTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGTTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCTTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCTTGATGCTGCGCTCATCACT |
| (681) | TTGTTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGTTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |
| (681) | TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT |

720
(681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT
(681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGTTCATCACT (681) (681) TTGCOCROATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGTTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCTTGATGCTGCGCTCATCACT 681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGTCTGATGCTGCGCTCATCACT (681) TTGTTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT
(681) IIGCICIIATGGAGGAGAGCCIGATGCIGCGCICAICACI (681) (681) TTGCTCTTATGGAGGAGACCTTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTTATCACT (681) TIGCTCIIATGGAGGAGAGCCTGATGCTGCGCTCATCACI (681)
(681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT
(681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT 681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCACATCACI (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) ITGCICIIATGGAGGAGAGCCIGATGCIGCGCICAICACI (681) (681) TTGCTCTOATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTAATGCTGCGTTCATCACT (681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TIGITCIIATGGAGGAGAGCCTGATGCTGCGCTCATCACI (681) (681) TTGCTCROATGGAGGAGCTTGATGCTGCGCTCATCACT (681) TTGCTCTTATGGAGGAGAGCTTGATGCTGCGCTCATCACT (681) TTGTTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT (681) TTGTTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT
(681) TTGCTCTTATGGAGGAGAGCCTGATGCTGCGCTCATCACT

Figure 17. Continued

214445 KEN83 56KEN65
900085MAU88 An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84
AnTAMBULEGY94
Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar74RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA0 8 SPU204ANGL85 SPU2207KEN07 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07 SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus
(721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT (721) CCCTAGCATGATGGGGAGAAACAACTGGGTTCCAGTTGTT (721) CCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT
(721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT (721) TCCCAGCATGATGGGGAGAAACAACTGGATTCCAGTTATT (721) CCCTAGCATGATGGGGAGAAACAACTGGGTTCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT
(721) TCCTAGCATGATGGGGAGAAACAACTGGGTCCCAGTTGTT
(721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT
(721) TCCTAGTATGATGGGGAGAAACAACTGGATTCCAGTTATT
(721) CCCTAGCATGATGGGAAGAAACAACTGGATTCCAGTTATT (721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT (721) CCCTAGCATGATGGGGAGAAACAACTGGGTTCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT
(721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT
(721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAATAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT
(721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT
(721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTC
(721) TCCTAGCATGATGGGGAGAAACAACTGGATCCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT (721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT
(721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTATT
(721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT
(721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT
(721) CCCTAGCATGATGGGGAGAAACAACTGGGTTCCAGTTGTT
(721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT
(721) TCCCAGCATGATGGGGAGAAACAACTGGATTCCAGTTATT
(721) TCCCAGCATGATGGGGAGAAACAACTGGATTCCAGTTATT
(721) TCCTAGCATGATGGGGAGAAACAACTGGATTCCAGTTGTT

Figure 17. Continued

214445 KEN83 56KEN65
900085MAU88 An1000MAD91 An991MAD91
An999MAD91
AnK6087GUI84
AnTAMBULEGY94
Ar12568RSA71
Ar20368RSA81
Ar21229SA00
Ar74RSA55
Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA0 8 SPU204ANGL85 SPU2207KEN07 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus
(761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) TCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGATTCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAGGAGA (761) CCTCCAATCCCAGATGTTGAGATAGAATCAGAGGAAGAGA (761) CСTCСAATTCCAGATGTTGAGATGGAATCAGAGGAGGAGA (761) CCTCCAATCCCAGATGTTGAGATAGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAGGAGA (761) CCTCCAATCCCAGATGTTGAGATGGATTCAGAGGAGGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATAGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAGGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CСTCСAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CСTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAGGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAGGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGGGGAAGAGA (761) CСTCСAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAGGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATAGAATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGGATCAGAGGAAGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAGGAGA (761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAGGAGA
(761) CCTCCAATCCCAGATGTTGAGATGGAATCAGAGGAAGAGA

Figure 17. Continued

|  |  | 801840 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| $56 \mathrm{KEN65}$ | (801) | GTGATGACGATGGATTTGTTGAGGTTGATTAGAGATTAAG |
| $900085 \mathrm{MAU8} 8$ | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| An1000MAD91 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAAGTTAAG |
| An991MAD91 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAAGTTAAG |
| An9 99MAD91 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAAGTTAAG |
| AnK6087GUI84 | (801) | GTGATGATGATGGATTTGTTGAAGTTGATTAGAGGTTAAG |
| AnTAMBULEGY94 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| Ar12568RSA71 | (801) | GTGATGACGATGGATTTGTTGAGGTTGATTAGAGATTAAG |
| Ar20368RSA81 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| Ar21229SA00 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| Ar74RSA55 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAGG |
| Ar811MAD79 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| ArD38661SEN81 | (801) | GTGATGATGATGGATTTGTTGAGGTTAATTAGAGGTTAAG |
| B1143KEN77 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| B314KEN62 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| H1825RSA75 | (801) | GTGATGACGATGGATTTGTTGAGGTTGATTAGAGATTAAG |
| LunyoUGA55 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTGAG |
| R1662CAR85 | (801) | GTGATGATGATGGATTTGTTGAAGTTGATTAGAGGTTATG |
| SPU10301KEN07 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU10302KEN07 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU10307KEN07 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU10315KEN07 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU12002SOM98 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU152RSA08 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU204ANGL85 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU2207KEN07 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU2214KEN07 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTGAG |
| SPU2215KEN07 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU2220KEN07 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU2223KENO7 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU384001KEN97 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU45ZAMB85 | (801) | GTGATGATGATGGATTTGTTGAAGTTGATTAGAGGTTAAG |
| SPU52001RSA99 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU77NAMB0 4 | (801) | GTGATGATGATGGGTTTGTTGAGGTTGATTAGAGGTTAAG |
| SPU77RSA08 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGCTAAG |
| VRL1032 ZIM78 | (801) | GTGATGATGATGGATTTGTTGAAGTTGATTAGAGGTTAAG |
| VRL1290ZIM78 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| VRL1516ZIM78 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| VRL1887ZIM78 | (801) | GTGATGATGATGGATTTGTTGAAGTTGATTAGAGGTTAAG |
| VRL2230ZIM78 | (801) | GTGATGATGATGGATTTGTTGAAGTTGATTAGAGGTTAAG |
| VRL763ZIM70 | (801) | GTGATGACGATGGATTTGTTGAGGTTGATTAGAGATTAAG |
| VRL825ZIM79 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| ZH501EGY77 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| ZH548EGY77 | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |
| Consensus | (801) | GTGATGATGATGGATTTGTTGAGGTTGATTAGAGGTTAAG |

Figure 17. Continued

214445 KEN83 56KEN65
900085MAU88 An1000MAD91 An991MAD91
An999MAD91
AnK6087GUI84
AnTAMBULEGY94
Ar12568RSA71
Ar20368RSA81
Ar21229SA00
Ar74RSA55
Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA0 8 SPU204ANGL85 SPU2207KEN07 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07 SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus

[^0]Figure 17. Continued

```
                    881
                920
```

214445 KEN83 5 6KEN65
900085 MAU88 An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94 Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar 74 RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA0 8 SPU204ANGL85 SPU2207KENO 7 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA0 8 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus

881
920
(880) CAACCCCCAT-TTCCCCCAACCCCCCTGGGCAGCCACTTA
(879) CAACCACCCT-TTTTCCCCAAACCCCTGGGCAGCCACTTA
(881) CAACCACCCCC-TTCCCCCAAACCCCTGGGCAGCCGCTTA (880) CAACCCCCAT-TTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-TTTCCCCAACCCCCCTGGGCAGCCACTTA
(880) CAACCCCCCT-TTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCCCCTT-TT-CCCCCAACCCCCTGGGCAGCCACTTA (880) CAACTCCCC--TTCCCCCCAACCCCCTGGGCAGCCACTTA (879) CAACCACCCC-TTTTCCCCAAACCCCTGGGCAGCCACTTA (880) CAACTCCCC--TTCCCCCCAACCCCTTGGGCAGCCACTTA (879) CAACCCCCAT-TTTTCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCACCCCT-TTCCCCCAAACCCCTGGGCAGCCACTTA (880) CAACTCCCCC-TTCCCCCCAACCCCCTGGGCAGCCACTTA (880) CAACCACCCCT-TTCCCCCAAACCCCTGGGCAGCCACTTA (880) CAACTCCCCT-TCCCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCACCCC-TTTCCCCCAAACCCCTGGGCAGCCACTTA (879) CAACCACCCCCTTTTCCCCAAACCCCTGGGCAGCCACTTA (880) CAACCACCCCCTTTCCCCCAAACCCCTGGGCAGCCACTCA (880) TAACCCCCCT-TT-CCCCCAACCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-CTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-CTTCCCCAACCCCCCTGGGCAGCCACTTA (879) CAACCCCCAT-CTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-CTTCCCCAACCCCCCTGGGCAGCCACTTA (879) CAACCCCCAT-TTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-CTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-TTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-TTTCCC-AACCCCCCTGGGCAGTCACTTA (880) CAACCCCCAT-TTTCCCCAACCCCCCTGGGCAGTCACTTA (880) CAACCCCCAT-CTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-CTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-СТTCCCCAACCCCCCTGGGCAGCCACTTA (879) CAACCCCCAT-TTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACCCCCTT-AT-CCCCCAACCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-TTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACTCCCC--TTTCCCCCAACCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-CTTCCCCAACCCCCCTGGGCAGCCACTTA (880) CAAACCCCCT-TTTCCCCCAACCCCCTGGGCAGCCACTTA (880) CAACCCCCAT-TTCCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACTCCCC--TTCCCCCCAACCCCCTGGGCAGCCACTTA (880) CAACCCCCCT-AT-CCCCCAACCCCCTGGGCAGCCACTTA (880) CTACCCCCCT-AT-CCCCCAACCCCCTGGGCAGCCACTTA (879) CAACCACCCC-TTTTCCCCAAACCCCTGGGCAGCCACTTA (880) CAACCCCCAT-TTCCCCCAACCCCCCTGGGCAGCCACTTA (880) CAACTCCCC--TTCCCCCCAACCCCCTGGGCAGCCACTTA (880) CAACTCCCC--TTCCCCCCAACCCCCTGGGCAGCCACTTA (881) CAACCCCCCT TTTCCCCCAACCCCCTGGGCAGCCACTTA

Figure 17. Continued

|  |  | 921960 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| 56 KEN 65 | (918) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCATGACAGCT |
| $900085 \mathrm{MAU8} 8$ | (920) | GGCTGCTGTCTTGTACGCCTGAGCAGCTGCCATGACAGCC |
| An1000MAD91 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| An991MAD91 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| An999MAD91 | (919) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCATGACAGCA |
| AnK6087GUI84 | (918) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCTTGACAGCA |
| AnTAMBULEGY94 | (918) | GGCTGCTGTCTTGTAAGCCTGAGCGGCTGCCATGACAGCA |
| Ar12568RSA71 | (918) | GGCTGCTGTCTTGTACGCCTGAGCAGCTGCCATGACAGCT |
| Ar20368RSA81 | (918) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCATGACAGCA |
| Ar21229SA00 | (918) | GGCTGCTGTTTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| Ar 74 RSA 55 | (919) | GGCTGCTGTCTTGTACGCCTGAGCAGCTGCCATGACAGCC |
| Ar811MAD79 | (919) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCATGACAGCA |
| ArD38661SEN81 | (919) | AGCTGCTGTCTTGTACGCCTGAGCAGCTGCCATGACAGCC |
| B1143KEN77 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| B314KEN62 | (919) | GGCTGCTGTCTTGTACGCCTGAGCAGCTGCCATGACAGCC |
| H1825RSA75 | (919) | GGCTGCTGTCTTGTACGCCTGAGCAGCTGCCATGACAGCT |
| LunyoUGA55 | (920) | GGCTGCTGTCTTGTACGCCTGAGCAGCTGCCATGACAGCT |
| R1662CAR85 | (918) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCATGACAGCA |
| SPU10301KEN07 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU10302KEN0 7 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU10307KEN07 | (918) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU10315KEN0 7 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU12002SOM98 | (918) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU152RSA08 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU204ANGL85 | (919) | GGCTGCTGTCTTGTAGGCCTGGGCAGCTGCCATGACAGCA |
| SPU2207KEN07 | (918) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU2214KEN07 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU2215KEN07 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU2220KEN07 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU2223KEN07 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU384001KEN97 | (918) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| SPU45ZAMB85 | (918) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCATGACAGCA |
| SPU52001RSA99 | (919) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCATGACAGCA |
| SPU77NAMB0 4 | (918) | GGCTGCTGTCTTGTAAGCCTGGGCAGCTGCCATGACAGCA |
| SPU77RSA08 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| VRL1032 ${ }^{\text {ZIM7 }} 8$ | (919) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCATGACAGCA |
| VRL1290ZIM78 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| VRL1516ZIM78 | (918) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCATGACAGCA |
| VRL1887ZIM78 | (918) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCATGACAGCA |
| VRL2230ZIM78 | (918) | GGCTGCTGTCTTGTAAGCCTGAGCAGCTGCCATGACAGCA |
| VRL763ZIM70 | (918) | GGCTGCTGTCTTGTACGCCTGAGCAGCTGCCATGACAGCT |
| VRL825ZIM79 | (919) | GGCTGCTGTCTTGTAGGCCTGAGCAGCTGCCATGACAGCA |
| ZH501EGY77 | (918) | GGCTGCTGTCTTGTAAGCCTGAGCGGCTGCCATGACAGCA |
| ZH548EGY77 | (918) | GGCTGCTGTCTTGTAAGCCTGAGCGGCTGCCATGACAGCA |
| Consensus | (921) | GGCTGCTGTCTTGTA GCCTGAGCAGCTGCCATGACAGCA |

Figure 17. Continued

|  |  | 9611000 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (959) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| 56KEN65 | (958) | GCTGACGGCTTCCCATTGGAATCCACAAGCCCAAAAGCTT |
| 900085MAU88 | (960) | GCTGACGGCTTCCCATTGGAATCCACAAGCCCAAAAGCTT |
| An1000MAD91 | (959) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| An991MAD91 | (959) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| An999MAD91 | (959) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAGGCTT |
| AnK6087GUI84 | (958) | GCTGATGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| AntAMBULEGY94 | (958) | GCTGACGGCTTCCCATTAGAATCCACAAGTCCAAAGGCTT |
| Ar12568RSA71 | (958) | GCTGACGGCTTCCCATTGGAATCCACAAGCCCAAAAGCTT |
| Ar20368RSA81 | (958) | GCTGACGGCTTCCCATTAGAGTCCACAAGCCCAAAAGCTT |
| Ar21229SA00 | (958) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| Ar74RSA55 | (959) | GCTGATGGCTTCCCATTGGAATCCACAAGCCCAAAAGCTT |
| Ar811MAD79 | (959) | GCTGACGGCTTCCCATTGGAATCCACAAGCCCAAAGGCTT |
| Ard38661SEN81 | (959) | GCTGACGGCTTCCCATTGGAATCCACAAGCCCAAAAGCTT |
| B1143KEN77 | (959) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| B314KEN62 | (959) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| H1825RSA75 | (959) | GCTGACGGCTTCCCATTGGAATCCACAAGCCCAAAAGCTT |
| LunyoUGA55 | (960) | GCTGACGGCTTTCCATTAGAATCCACAAGCCCAAAAGCTT |
| R1662CAR85 | (958) | GCTGATGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU10301KEN07 | (959) | GCTGACGGTTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU10302KEN07 | (959) | GCTGACGGTTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU10307KEN07 | (958) | GCTGACGGTTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU10315KEN07 | (959) | GCTGACGGTTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU12002SOM98 | (958) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU152RSA08 | (959) | GCTGACGGTTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU204ANGL85 | (959) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU2207KEN07 | (958) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU2214KEN07 | (959) | GCTGACGGCTTCCCATTGGAATCCACAAGCCCAAAAGCTT |
| SPU2215KEN07 | (959) | GCTGACGGTTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU2220KEN07 | (959) | GCTGACGGTTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU2223KEN07 | (959) | GCTGACGGTTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU384001KEN97 | (958) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU45ZAMB85 | (958) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU52001RSA99 | (959) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU77NAMB04 | (958) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| SPU77RSA08 | (959) | GCTGACGGTTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| VRL1032ZIM78 | (959) | GCTGATGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| VRL1290ZIM78 | (959) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| VRL1516ZIM78 | (958) | GCTGACGGCTTCCCATTGGAATCCACAAGCCCAAAGGCTT |
| VRL18872IM78 | (958) | GCTGACGGCTTCCCATTGGAATCCACAAGCCCAAAAGCTT |
| VRL2230ZIM78 | (958) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| VRL763ZIM70 | (958) | GCTGACGGCTTCCCATTGGAATCCACAAGCCCAAAAGCTT |
| VRL825ZIM79 | (959) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |
| ZH501EGY77 | (958) | GCTGACGGCTTCCCATTAGAATCCACAAGTCCAAAGGCTT |
| ZH548EGY77 | (958) | GCTGACGGCTTCCCATTGGAATCCACAAGTCCAAAGGCTT |
| Consensus | (961) | GCTGACGGCTTCCCATTAGAATCCACAAGCCCAAAAGCTT |

Figure 17. Continued

```
                                    1001
                                    1040
```

$214445 K E N 83$ 56KEN65
900085 MAU88 An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94

Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar 74 RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN0 7 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA0 8 SPU204ANGL85 SPU2207KENO 7 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA0 8 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus

1001
1040
(999) TCAAGAATTCCCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(1000) TCAAGAATTCTCTCCTCTTCTCATGGCTTATGAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TTAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATGAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCCCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(1000) TCAAGAATCCTCTCCTCTTCTCATGGCTTATGAAATTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTAATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTAATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(999) TCAAGAATTCCCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(998) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT
(1001) TCAAGAATTCTCTCCTCTTCTCATGGCTTATAAAGTTGCT

Figure 17. Continued

214445 KEN83 56KEN65
900085MAU88 An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94 Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar74RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA0 8 SPU204ANGL85 SPU2207KEN07 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77
Consensus
(1039) ATTTACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA
(1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA
(1040) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG
(1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG
(1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCGGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCGGCA (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1039) ATTTACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCGGCA (1040) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCG (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCGGCA (1039) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA (1038) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA
(1041) ATTCACTGCTGCATTCATTGGCTGCGTGAACGTTGCAGCA

Figure 17. Continued

|  |  | 1081 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| $56 \mathrm{KEN65}$ | (1078) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| $900085 \mathrm{MAU8} 8$ | (1080) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| An1000MAD91 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| An9 91MAD91 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| An999MAD91 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| AnK6087GUI84 | (1078) | ACCTCCTCCTTTGTTCTGCCTCGGAGGTTTGGGTTGATGA |
| AnTAMBULEGY94 | (1078) | ACCTCCTCTTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| Ar12568RSA71 | (1078) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| Ar20368RSA81 | (1078) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| Ar21229SA00 | (1078) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| Ar74RSA55 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATAA |
| Ar811MAD79 | (1079) | ACCTCCTCTTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| Ard38661SEN81 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| B1143KEN77 | (1079) | A |
| B314KEN62 | (1079) | ACCTCCTCCTTTGTTCTGCCTCGGAGGTTTGGGTTGATGA |
| H182 | (1079) | A |
| LunyoUGA55 | (1080) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| R1662CAR85 | (1078) | ACCTCCTCTTTTGTTCTGCCTCGGAGGTTTGGGTTGATGA |
| SPU10301KEN07 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU10302 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU10307KEN07 | (1078) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU10315KEN07 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU12002SOM98 | (1078) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU152RSA08 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU204ANGL85 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU2207KEN07 | (1078) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU2214KEN07 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGATTTGGGTTGATGA |
| SPU2215KEN07 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU2220KEN07 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU2223KEN07 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| PU384001KEN97 | (1078) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU45ZAMB85 | (1078) | ACCTCCTCCTTTGTTCTGCCTCGGAGGTTTGGGTTGATGA |
| SPU52001RSA99 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU77NAMB04 | (1078) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| SPU77RSA08 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| VRL1032 ZIM78 | (1079) | ACCTCCTCTTTTGTTCTGCCTCGGAGGTTTGGGTTGATGA |
| VRL1290ZIM78 | (1079) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| VRL1516ZIM78 | (1078) | ACCTCCTCTTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| VRL1887ZIM78 | (1078) | ACCTCCTCCTTTGTTCTGCCTCGGAGGTTTGGGTTGATGA |
| VRL2230ZIM78 | (1078) | ACCTCCTCCTTTGTTCTGCCTCGGAGGTTTGGGTTGATGA |
| VRL763ZIM70 | (1078) | ACCTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| VRL825ZIM79 | (1079) | ACCTCCTCCTTTGTTCTGCCTCGGAGGTTTGGGTTGATGA |
| ZH501EGY77 | (1078) | ACCTCCTCTTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| ZH548EGY77 | (1078) | ACCTCCTCTTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |
| Consensus | (1081) | ACcTCCTCCTTTGTTCTACCTCGGAGGTTTGGGTTGATGA |

Figure 17. Continued

|  |  | 11211160 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCCAA |
| $56 \mathrm{KEN65}$ | (1118) | CCCGAGAGAACTGCAGCAGATACAGAGAGTGAGCATCCAA |
| $900085 \mathrm{MAU8} 8$ | (1120) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCCAA |
| An1000MAD91 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| An9 91MAD91 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| 1 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| AnK6087GUI84 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| AnTAMBULEGY94 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| Ar12568RSA71 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCCAA |
| Ar20368RSA81 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| Ar21229SA00 | (1118) | CCCGGGAAAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| Ar74RSA55 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCCAA |
| Ar811MAD79 | (1119) | A |
| ArD38661SEN81 | (1119) | CCCGGGAGAACTGCAGCAGGTACAGAGAGTGAGCATCCAA |
| B1143KEN77 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCCAA |
| B314KEN62 | (1119) | CCCGAGAGAACTGCAGCAGATACAGAGAGTGAGCATCCAA |
| H1825RSA75 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCCAA |
| LunyoUGA55 | (1120) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCCAA |
| R1662CAR85 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU10301KEN07 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU10302KEN07 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU10307KEN07 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU10315KEN07 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU12002SOM98 | (1118) | CCCGGGAGAACTGTAGCAGATACAGAGAGTGAGCATCTAA |
| SPU152RSA08 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU204ANGL85 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU2207KEN07 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU2214KEN07 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU2215KEN07 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAG |
| SPU2220KEN07 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU2223KEN07 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU384001KEN97 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU45ZAMB85 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU52001RSA99 | (1119) | CCCGGGAGAACTGCAGCAGATACAAAGAGTGAGCATCTAA |
| SPU77NAMB0 4 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| SPU77RSA08 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| VRL1032 ZIM78 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| VRL1290ZIM78 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| VRL1516ZIM78 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| VRL1887ZIM78 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| VRL2230ZIM78 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| VRL763ZIM70 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCCAA |
| VRL825ZIM79 | (1119) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| ZH501EGY77 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| ZH548EGY77 | (1118) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |
| Consensus | (1121) | CCCGGGAGAACTGCAGCAGATACAGAGAGTGAGCATCTAA |

Figure 17. Continued
$214445 \mathrm{KEN83}$ (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC 56KEN65 (1158) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCTACC
900085MAU88 (1160) TATTGCCCTTAGGTAGTCTTCTGGTAGAGAAGGGTCCACC An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94 Ar12568RSA71
Ar20368RSA81 Ar21229SA00

Ar74RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA08 SPU204ANGL85 SPU2207KEN07 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77

Consensus
(1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC
(1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC
(1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC
(1158) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTCCTGGTAGAGAAGGATCCACC (1158) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTCCTGGTAGAGAAGGATCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCTACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1160) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGATCCACC (1158) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTCAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTTCTGGCAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTTCTGGCAGAGAAGGGTCCACC (1159) AATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1159) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTCCTGGTAGAGAAGGATCCACC (1158) TATTGCCCTTAGGTAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGGTAGTCTTCTGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCTACC (1159) TATTGCCCTCAGATAGTCTTCAGGTAGAGAAGGGTCCACC (1158) TATTGCCCTTAGATAGTCTCCTGGTAGAGAAGGATCCACC (1158) TATTGCCCTTAGATAGTCTCCTGGTAGAGAAGGATCCACC
(1161) TATTGCCCTTAGATAGTCTTCTGGTAGAGAAGGGTCCACC

Figure 17. Continued

|  |  | 12011240 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (1199) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| $56 \mathrm{KEN65}$ | (1198) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| $900085 \mathrm{MAU8} 8$ | (1200) | ATGCCAGCAAAGCTGGGGTGCATCATGTGCCTTGGGTATG |
| An1000MAD91 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTTGGGTATG |
| An991MAD91 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTTGGGTATG |
| An999MAD91 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTTGGGTATG |
| AnK6087GUI84 | (1198) | CTTGGGTATG |
| AnTAMBULEGY94 | (1198) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| Ar12568RSA71 | (1198) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| Ar20368RSA81 | (1198) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| Ar21229SA00 | (1198) | ATGCCAGCAAAGCTGGGATGCATCATGTGCCTTGGGTATG |
| Ar74RSA55 | (1199) | ATGCCAGCAAAACTGGGGTGCATCATATGCCTTGGGTATG |
| Ar811MAD79 | (1199) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| D38661SEN81 |  | TGCCAGCAAAGCTGGGATGCATCATATGCCTTGGGTATG |
| B1143KEN77 | (1199) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| B3 | (1199) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| H1825RSA75 | (1199) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| LunyoUGA55 | (1200) | G |
| R1662CAR85 | (1198) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| SPU10301KEN07 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTCGGGTATG |
| SPU10302KEN07 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTCGGGTATG |
| SPU10307KEN07 | (1198) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTCGGGTATG |
| SPU10315KEN07 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTCGGGTATG |
| SPU12002SOM98 | (1198) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTTGGGTATG |
| SPU152RSA08 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTCGGGTATG |
| SPU204ANGL85 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTTGGGTATG |
| SPU2207KEN07 | (1198) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTTGGGTATG |
| SPU2214KEN07 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTTGGGTATG |
| SPU2215KEN07 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTCGGGTATG |
| SPU2220KE | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTCGGGTATG |
| SPU2223KEN07 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTCGGGTATG |
| PU384001KEN97 | (1198) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTTGGGTATG |
| SPU45ZAMB85 | (1198) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTTGGGTATG |
| SPU52001RSA99 | (1199) | ATGCCAGCAAAGCTAGGATGCATCATATGCCTTGGGTATG |
| SPU77NAMB04 | (1198) | ATGCCAGCAAAGCTAGGGTGCATCATATGCCTTGGGTATG |
| SPU77RSA08 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTCGGGTATG |
| VRL1032 ZIM78 | (1199) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| VRL1290ZIM78 | (1199) | ATGCCAGCAAAGCTGGGATGCATCATATGCCTTGGGTATG |
| VRL1516ZIM78 | (1198) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| VRL1887ZIM78 | (1198) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| VRL2230ZIM78 | (1198) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| VRL763ZIM70 | (1198) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| VRL825ZIM79 | (1199) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| ZH501EGY77 | (1198) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |
| ZH548EGY77 | (1198) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTCGGGTATG |
| Consensus | (1201) | ATGCCAGCAAAGCTGGGGTGCATCATATGCCTTGGGTATG |

Figure 17. Continued

|  |  | 1241 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (1239) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| 56KEN65 | (1238) | CAGGGGATAGGCCATCCATGGTAGTCCCAGTGACAGGAAG |
| $900085 \mathrm{MAU8} 8$ | (1240) | CAGGGGATAAGCCATCCATGGTAGTCCCAGTGACAGGAAG |
| An1000MAD91 | (1239) | CAGGGGATAGGCCGTCCATGGTTGTCCCAGTGACAGGAAG |
| An991MAD91 | (1239) | CAGGGGATAGGCCGTCCATGGTTGTCCCAGTGACAGGAAG |
| - | (1239) | CAGGGGATAGGCCGTCCATGGTTGTCCCAGTGACAGGAAG |
| AnK6087GUI84 | (1238) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| AnTAMBULEGY94 | (1238) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| Ar12568RSA71 | (1238) | CAGGGGATAGGCCATCCATGGTGGTCCCAGTGACAGGAAG |
| Ar20368RSA81 | (1238) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| Ar21229SA00 | (1238) | CAGGGGATAGGCCGTCCATGGTTGTCCCAGTGACAGGAAG |
| Ar 74 RSA 55 | (1239) | CAGGGGATAGGCCATCCATGGTAGTCCCAGTGACAGGAAG |
| AL811MAD79 |  | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| rD38661SEN81 | (1239) | CAGGGGATAGGCCATCCATGGTAGTCCCAGTGACAGGAAG |
| B1143KEN77 | (1239) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| B314KEN62 | (1239) | CAGGGGATAGGCCATCCATGGTAGTCCCAGTGACAGGAAG |
| H1825RSA75 | (1239) | CAGGGGATAGGCCATCCATGGTAGTCCCAGTGACAGGAAG |
| LunyoUGA55 | (1240) | CAGGAGATAGGCCATCCATGGTAGTCCCAGTGACAGGAAG |
| R1662CAR85 | (1238) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| PU10301KEN07 | (1239) | CAGGGGATAGACCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU10302KEN07 | (1239) | CAGGGGATAGACCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU10307KEN07 | (1238) | CAGGGGATAGACCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU10315KEN07 | (1239) | CAGGGGATAGACCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU12002SOM98 | (1238) | CAGGGGATAGGCCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU152RSA08 | (1239) | CAGGGGATAGACCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU204ANGL85 | (1239) | CAGGGGATAGGCCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU2207KEN07 | (1238) | CAGGGGATAGACCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU2214KEN07 | (1239) | CAGGGGACAGGCCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU2215KEN07 | (1239) | CAGGGGATAGACCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU2220KEN07 | (1239) | CAGGGGATAGACCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU2223KEN07 | (1239) | CAGGGGATAGACCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU384001KEN97 | (1238) | CAGGGGATAGGCCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU45ZAMB85 | (1238) | CAGGGGATAGGCCATCCATGGTGGTCCCGGTGACAGGAAG |
| SPU52001RSA99 | (1239) | CAGGGGATAGGCCGTCCATGGTTGTCCCAGTGACAGGAAG |
| SPU77NAMB04 | (1238) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| SPU77RSA08 | (1239) | CAGGGGATAGACCGTCCATGGTTGTCCCAGTGACAGGAAG |
| VRL1032 2 IM78 | (1239) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| VRL1290ZIM78 | (1239) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| VRL1516ZIM78 | (1238) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| VRL1887ZIM78 | (1238) | CAGGGGATAGGCCATCCATGGTAGTCCCAGTGACAGGAAG |
| VRL2230ZIM78 | (1238) | CAGGGGATAGGCCATCCATGGTAGTCCCAGTGACAGGAAG |
| VRL763ZIM70 | (1238) | CAGGGGATAGGCCATCCATGGTGGTCCCAGTGACAGGAAG |
| VRL825ZIM79 | (1239) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| ZH501EGY77 | (1238) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| ZH548EGY77 | (1238) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |
| Consensus | (1241) | CAGGGGATAGGCCGTCCATGGTAGTCCCAGTGACAGGAAG |

Figure 17. Continued

```
                    1281
                    1320
    214445KEN83 (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        56KEN65 (1278) CCACTCACTCAAGACGACCAAAGCCTGGCAAGTCCAGCCA
    900085MAU88 (1280) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
    An1000MAD91
    An991MAD91
    An999MAD91
    AnK6087GUI84
AnTAMBULEGY94
    Ar12568RSA71
    Ar20368RSA81
        Ar21229SA00
            Ar74RSA55
        Ar811MAD79
ArD38661SEN81
        B1143KEN77
            B314KEN62
        H1825RSA75
        LunyoUGA55
        R1662CAR85
SPU10301KEN07
SPU10302KEN07
SPU10307KEN07
SPU10315KEN07
SPU12002SOM98
    SPU152RSA08
        SPU204ANGL85
        SPU2207KEN07
        SPU2214KEN07
        SPU2215KEN07
        SPU2220KEN07
        SPU2223KEN07
SPU384001KEN97
        SPU45ZAMB85
SPU52001RSA99
        SPU77NAMB04
        SPU77RSA08
    VRL1032ZIM78
    VRL1290ZIM78
    VRL1516ZIM78
    VRL1887ZIM78
    VRL2230ZIM78
        VRL763ZIM70
        VRL825ZIM79
        ZH501EGY77
        ZH548EGY77
            Consensus
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCAAGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCAAGTCCAGCCA
        (1280) CCACTCACTCAAGACGACCAGAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTTAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAGAGCCTGGCAAGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAGAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAGAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCAAGTCCAGCCA
        (1279) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
        (1278) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
    (1281) CCACTCACTCAAGACGACCAAAGCCTGGCATGTCCAGCCA
```

Figure 17. Continued

```
                                    1321
                                    1360
```

214445 KEN83 56KEN65
$900085 \mathrm{MAU8} 8$ An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94 Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar 74 RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN0 7 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA08 SPU204ANGL85 SPU2207KENO 7 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85
SPU52001RSA99 SPU77NAMB0 4 SPU77RSA0 8 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus
(1319) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC
(1318) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC
(1320) GCCAGGGCGGCAGCAACTCTTGATAGAGTCAACTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC
(1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCGGCAGCAACTCGTGATAGAGTTAACTCATCCC (1318) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC (1319) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1319) GCCAGGGCGGCAGCAACTCTTGATAGAGTCAACTCATCCC (1319) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1319) GCCAGGGCGGCAGCAACTCGTGATAGAGTTAACTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC (1320) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCTC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAGCTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAGCTCATCCC (1318) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAGCTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAGCTCATCCC (1318) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAGCTCATCCC (1319) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAGCTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAGCTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAGCTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAGCTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAGCTCATCCC (1318) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAAGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1319) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAGCTCATCCC (1319) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1319) GCCAGAGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCAGCAGCAACTCGTGATAGAGTCAACTCATCCC (1319) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC (1318) GCCAAGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC
(1321) GCCAGGGCGGCAGCAACTCGTGATAGAGTCAACTCATCCC

Figure 17. Continued

```
    1 3 6 1 ~ 1 4 0 0
    214445KEN83 (1359) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        56KEN65 (1358) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
    900085MAU88 (1360) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
    An1000MAD91
    An991MAD91
    An999MAD91
    AnK6087GUI84
AnTAMBULEGY94
    Ar12568RSA71
    Ar20368RSA81
        Ar21229SA00
            Ar74RSA55
        Ar811MAD79
ArD38661SEN81
        B1143KEN77
            B314KEN62
        H1825RSA75
        LunyoUGA55
        R1662CAR85
SPU10301KEN07
SPU10302KEN07
SPU10307KEN07
SPU10315KEN07
SPU12002SOM98
    SPU152RSA08
        SPU204ANGL85
        SPU2207KEN07
        SPU2214KEN07
        SPU2215KEN07
        SPU2220KEN07
        SPU2223KEN07
SPU384001KEN97
        SPU45ZAMB85
SPU52001RSA99
        SPU77NAMB04
            SPU77RSA08
    VRL1032ZIM78
    VRL1290ZIM78
    VRL1516ZIM78
    VRL1887ZIM78
    VRL2230ZIM78
        VRL763ZIM70
        VRL825ZIM79
        ZH501EGY77
        ZH548EGY77
            Consensus
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GAGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCTTCCTTTAGCTTGTACTTGTTGATGAG
        (1358) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GAGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAGGGATTTCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GAGAGGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1360) GGGAAGGATTCCCCTCCTTTAGCTTGTACTTGTTGATGAG
        (1358) GGGAAGGATTCCCTTCCTTTAGCTTATATTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCCTCCTTTAGCTTGTACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCTTCCTTTAACTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
        (1359) GGGAAGGATTCCCTTCCTTTAGCTTATATTTGTTGATGAG
        (1359) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCCTCCTTTAGCTTATACCTGTTGATGAG
        (1359) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
        (1358) GGGAAGGATTCCCTTCCTTTAGCTTATACTTGTTGATGAG
    (1361) GGGAAGGATTCCCCTCCTTTAGCTTATACTTGTTGATGAG
```

Figure 17. Continued

```
                                    1401
                                    1440
    214445KEN83 (1399) AGCCTCCACAGTTGCTTTGCCTTCTTTCGACATTTTCATC
        56KEN65 (1398) AGCCTCCACAGTTGCTTTGCCTTCTTTCGACATTTTCATC
        900085MAU88 (1400) AGCCTCCACAGTTGCCTTGCCTTCTTTCGACATTTTCATC
        An1000MAD91 (1399) AGCCTCCACAGTTGCTTTGCCTTCTTTCGACATTTTCATC
        An991MAD91
        An999MAD91
    AnK6087GUI84
AnTAMBULEGY94
    Ar12568RSA71
    Ar20368RSA81
        Ar21229SA00
            Ar74RSA55
        Ar811MAD79
ArD38661SEN81
        B1143KEN77
            B314KEN62
        H1825RSA75
        LunyoUGA55
        R1662CAR85
SPU10301KEN07
SPU10302KEN07
SPU10307KEN07
SPU10315KEN07
SPU12002SOM98
    SPU152RSA08
        SPU204ANGL85
        SPU2207KEN07
        SPU2214KEN07
        SPU2215KEN07
        SPU2220KEN07
        SPU2223KEN07
SPU384001KEN97
        SPU45ZAMB85
SPU52001RSA99
        SPU77NAMB04
            SPU77RSA08
    VRL1032ZIM78
    VRL1290ZIM78
    VRL1516ZIM78
    VRL1887ZIM78
    VRL2230ZIM78
        VRL763ZIM70
        VRL825ZIM79
        ZH501EGY77
        ZH548EGY77
            Consensus
\begin{tabular}{|c|c|}
\hline & 1401 \\
\hline (1399) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGACATTTTCATC \\
\hline 8) & A \\
\hline 1 & AGCCTCCACAGTTGCCTTGCCT \\
\hline 399) & A \\
\hline 9) & AGCCTCCACAGTTGCTTTGCCTTCTTTCG \\
\hline & \\
\hline 8) & AGCCTCCACAGTTGCTTTGCCTTCTTTTGAC \\
\hline 398) & AGCCTCCACAGTTGCTTTGCCTTCTTT \\
\hline 398) & AGCCTCCACAGTTGCT \\
\hline 3 & AGCCTCCACAGTTGCTTTGCCTTCTT \\
\hline & AGCCTCCACAGTTGCTTTGCCTTCTTTCGAC \\
\hline 399) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGACA \\
\hline 9) & A \\
\hline 9) & AGCCTCCACAGTTGCCTTGCCTTCT \\
\hline 399) & AGCCTCCACAGTTGCT \\
\hline & AGCCTCCACAGTTGCTTTGCCTTCTTTTGA \\
\hline & A \\
\hline & AGCCTCCACAGTTGCTTTGCCTTCCT \\
\hline 398) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGA \\
\hline 99) & A \\
\hline 399) & A \\
\hline 8) & AGCC \\
\hline 399 & AGCCTCCACAGTTGCTTTGCCTTCTTTCGA \\
\hline ) & A \\
\hline 399) & AGCCTCCACAGTTGCTTTGCCTTCT \\
\hline ) & AGCCT \\
\hline (1398) & AGCCTCTACAGTTGCTTTGCCTTCTT \\
\hline 399) & AGCCTCTACAGTTGCTTTGCCTTCTTT \\
\hline (1399) & AGCCTCCACAGTTGCTTTGCC \\
\hline (1) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGA \\
\hline (1399) & AGCCTCCACAGTTGCTTTGCC \\
\hline (1398 & AGCCTCCACAGTTGCTTTGCCTTCTTTCGA \\
\hline 398) & AGCCTCCACAGTTGCTTTGCCTTCTTT \\
\hline (1399) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGA \\
\hline (398) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGA \\
\hline 399) & AGCCTCCACAGTTGCTTTGCCTTCTTTCG \\
\hline 399) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGACATT \\
\hline (1399) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGA \\
\hline 398) & AGCCTCCACAGTTGCTTTTCCTTCTTTCGACATTI \\
\hline 398) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGACA \\
\hline 398) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGACAT \\
\hline 398) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGACATTT \\
\hline (1399) & AGCCTCCACAGTTGCTTTACCTTCTTTCGACATT \\
\hline (1398) & AGCCTCCACAGTTGCTTTGCCTTCTTTCGA \\
\hline & CCTCCACAGTTGCTTTGCC \\
\hline & \\
\hline
\end{tabular}
(1401) AGCCTCCACAGTTGCTTTGCCTTCTTTCGACATTTTCATC
```

Figure 17. Continued

|  |  | 1441 |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| $56 \mathrm{KEN65}$ | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| $900085 \mathrm{MAU8} 8$ | (1440) | ATCATCCTCCGGGGCTTATTGCCACGAGTTAGAGCCAGAA |
| An1000MAD91 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| An9 91MAD91 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| 1 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| AnK6087GUI84 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| AnTAMBULEGY94 | (1438) | ATCATCCTCCTGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| Ar12568RSA71 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| Ar20368RSA81 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| Ar21229SA00 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| Ar74RSA55 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| Ar811MAD7 | (1439) | ATCATCCTCCTGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| Ard38661SEN81 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| B1143KEN77 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| B314KEN62 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| H1825RSA75 | (1439) | CTTGTTGCCACGAGTCAGAGCCAGAA |
| LunyoUGA55 | (1440) | ATCATCCTCCGGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| R1662CAR85 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| SPU10301KEN07 |  | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU103 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU10307KEN07 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU10315KEN07 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU12002SOM98 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU152RSA08 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU204ANGL85 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAAAGCCAGAA |
| SPU2207KEN07 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU2214KEN07 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU2215KEN07 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU2220KEN0 7 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU2223KEN07 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| PU384001KEN97 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU45ZAMB85 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| SPU52001RSA99 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| SPU77NAMB04 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| SPU77RSA08 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| VRL1032 ZIM78 | (1439) | ATCATCCTCCTGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| VRL1290ZIM78 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| VRL1516ZIM78 | (1438) | ATCATCCTCCTGGGCTTGTTGCCACGAGTTAGAGCCAGGA |
| VRL1887ZIM78 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| VRL2230ZIM78 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| VRL763ZIM70 | (1438) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |
| VRL825ZIM79 | (1439) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGGA |
| ZH501EGY77 | (1438) | ATCATCCTCCTGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| ZH548EGY77 | (1438) | ATCATCCTCCTGGGCTTGTTGCCACGAGTTAGAGCCAGAA |
| Consensus | (1441) | ATCATCCTCCGGGGCTTGTTGCCACGAGTCAGAGCCAGAA |

Figure 17. Continued

```
                    1481
                            1520
    214445KEN83 (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        56KEN65 (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        900085MAU88
        An1000MAD91
        An991MAD91
        An999MAD91
    AnK6087GUI84
AnTAMBULEGY94
    Ar12568RSA71
    Ar20368RSA81
        Ar21229SA00
            Ar74RSA55
        Ar811MAD79
ArD38661SEN81
            B1143KEN77
            B314KEN62
            H1825RSA75
            LunyoUGA55
            R1662CAR85
SPU10301KEN07
SPU10302KEN07
SPU10307KEN07
SPU10315KEN07
SPU12002SOM98
    SPU152RSA08
        SPU204ANGL85
        SPU2207KEN07
        SPU2214KEN07
        SPU2215KEN07
        SPU2220KEN07
        SPU2223KEN07
SPU384001KEN97
        SPU45ZAMB85
SPU52001RSA99
        SPU77NAMB04
            SPU77RSA08
    VRL1032ZIM78
    VRL1290ZIM78
    VRL1516ZIM78
    VRL1887ZIM78
    VRL2230ZIM78
        VRL763ZIM70
        VRL825ZIM79
            ZH501EGY77
            ZH548EGY77
            Consensus
        (1480) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCTGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CGATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1480) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1479) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
        (1478) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
    (1481) CAATCATTTTCTTGGCATCCTTCTCCCAGTCAGCCCCACC
```

Figure 17. Continued

214445 KEN83 56KEN65
900085MAU88 An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94 Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar74RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN07 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA08 SPU204ANGL85 SPU2207KEN07 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77
Consensus
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAT
(1520) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAT
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACCCTGCGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACTCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAT
(1518) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAT
(1519) ATACTGCTTTAAGAGTTCGATAACTCTACGGGCATCAAAC
(1519) ATACTGCTTTAGGAGTTCGATAACCCTACGGGCATCAAAT
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAT
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAT
(1520) ATACTGCTTTAAGAGCTCGATAACCCTACGAGCATCAAAT
(1518) ATACTGCTTTAAGAGTTCAATAACCCTGCGAGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGAGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGAGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACCCTACGAGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGAGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGAGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGAGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGAGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGAGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATGACCCTACGGGCATCAAAT
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCAATAACTCTACGGGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGAGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACTCTACGGGCATCAAAC
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACTCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACTCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACTCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAT
(1519) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC (1518) ATACTGCTTTAAGAGTTCGATAACTCTACGGGCATCAAAC
(1518) ATACTGCTTTAAGAGTTCGATAACTCTACGGGCATCAAAC
(1521) ATACTGCTTTAAGAGTTCGATAACCCTACGGGCATCAAAC

Figure 17. Continued

```
                                    1561
                                    1600
```

214445 KEN83 5 6KEN65
$900085 \mathrm{MAU8} 8$ An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94 Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar 74 RSA55 Ar811MAD79 ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN0 7 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA0 8 SPU204ANGL85 SPU2207KENO 7 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA0 8 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus
(1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT
(1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT
(1560) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATTTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAGGCAAACTCTCGGACCCACTGTTCAATTTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1560) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGCTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAATTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCCTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1559) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT (1558) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT
(1561) CCTTGATAAGCAAACTCTCGGACCCACTGTTCAATCTCAT

Figure 17. Continued

```
                    1601
                    1640
```

214445 KEN83 5 6KEN65
$900085 \mathrm{MAU8} 8$ An1000MAD91
An991MAD91
An999MAD91
AnK6087GUI84 AnTAMBULEGY94 Ar12568RSA71
Ar20368RSA81 Ar21229SA00 Ar 74 RSA55 Ar811MAD79
ArD38661SEN81 B1143KEN77 B314KEN62 H1825RSA75 LunyoUGA55 R1662CAR85 SPU10301KEN07 SPU10302KEN0 7 SPU10307KEN07 SPU10315KEN07 SPU12002SOM98 SPU152RSA08 SPU204ANGL85 SPU2207KENO 7 SPU2214KEN07 SPU2215KEN07 SPU2220KEN07 SPU2223KEN07
SPU384001KEN97 SPU45ZAMB85 SPU52001RSA99 SPU77NAMB0 4 SPU77RSA0 8 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus
(1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGTTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1600) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC
(1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGTTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1600) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCGAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCGAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1599) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC (1598) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC
(1601) TGCGGTCCACTGCTTGAGCAGCAAACTGGATCGCAAGCTC

Figure 17. Continued

```
                    1641
                    1680
    214445KEN83 (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        56KEN65 (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
    900085MAU88 (1640) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
    An1000MAD91 (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
    An991MAD91 (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
    An999MAD91 (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
    AnK6087GUI84
AnTAMBULEGY94
    Ar12568RSA71
    Ar20368RSA81
        Ar21229SA00
            Ar74RSA55
        Ar811MAD79
ArD38661SEN81
        B1143KEN77
            B314KEN62
        H1825RSA75
        LunyoUGA55
        R1662CAR85
SPU10301KEN07
SPU10302KEN07
SPU10307KEN07
SPU10315KEN07
SPU12002SOM98
    SPU152RSA08
        SPU204ANGL85
        SPU2207KEN07
        SPU2214KEN07
        SPU2215KEN07
        SPU2220KEN07
        SPU2223KEN07
SPU384001KEN97
        SPU45ZAMB85
SPU52001RSA99
        SPU77NAMB04
            SPU77RSA08
    VRL1032ZIM78
    VRL1290ZIM78
    VRL1516ZIM78
    VRL1887ZIM78
    VRL2230ZIM78
        VRL763ZIM70
        VRL825ZIM79
        ZH501EGY77
        ZH548EGY77
            Consensus
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATTGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTATCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTACTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1640) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAGTAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTACTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTC-A
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1639) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
        (1638) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
    (1641) TTGATAGTTGTCCATTATTGTAATAGTGTTTGTATCTCTA
```

Figure 17. Continued

|  |  | $1681 \quad 1693$ |
| :---: | :---: | :---: |
| $214445 \mathrm{KEN8} 3$ | (1679) | GGGAGCTTTGTGT |
| $56 \mathrm{KEN65}$ | (1678) | GGGAGCTTTGTGT |
| $900085 \mathrm{MAU8} 8$ | (1680) | GGGAGCTTTGTGT |
| An1000MAD91 | (1679) | GGGAGCTTTGTGT |
| An991MAD91 | (1679) | GGGAGCTTTGTGT |
| 99MAD91 | (1679) | GGGAGCTTTGTGT |
| AnK6087GUI84 | (1678) | GGGAGCTTTGTGT |
| AnTAMBULEGY94 | (1678) | GGGAGCTTTGTGT |
| Ar12568RSA71 | (1678) | GGGAGCTTTGTGT |
| Ar20368RSA81 | (1678) | GGGAGCTTTGTGT |
| Ar21229SA00 | (1678) | GGGAGCTTTGTGT |
| Ar 74 RSA55 | (1679) | GGGAGCTTTGTGT |
| Ar811MAD79 | (1679) | GGGAGCTTTGTGT |
| ArD38661SEN81 | (1679) | GGGAGCTTTGTGT |
| B1143KEN77 | (1679) | GGGAGCTTTGTGT |
| B314KEN62 | (1679) | GGGAGCTTTGTGT |
| H1825RSA75 | (1679) | GGGAGCTTTGTGT |
| LunyoUGA55 | (1680) | GGGAGCTTTGTGT |
| R1662CAR85 | (1678) | GGGAGCTTTGTGT |
| PU10301KEN07 | (1679) | GGGAGCTTTGTGT |
| SPU10302KEN07 | (1679) | GGGAGCTTTGTGT |
| SPU10307KEN07 | (1678) | GGGAGCTTTGTGT |
| SPU10315KEN07 | (1679) | GGGAGCTTTGTGT |
| SPU12002SOM98 | (1678) | GGGAGCTTTGTGT |
| SPU152RSA08 | (1679) | GGGAGCTTTGTGT |
| SPU204ANGL85 | (1679) | GGGAGCTTTGTGT |
| SPU2207KEN07 | (1678) | GGGAGCTTTGTGT |
| SPU2214KEN07 | (1679) | GGGAGCTTTGTGT |
| SPU2215KEN07 | (1679) | GGGAGCTTTGTGT |
| SPU2220KEN07 | (1679) | GGGAGCTTTGTGT |
| SPU2223KEN07 | (1679) | GGGAGCTTTGTGT |
| SPU384001KEN97 | (1678) | GGGAGCTTTGTGT |
| SPU45ZAMB85 | (1678) | GGGAGCTTTGTGT |
| SPU52001RSA99 | (1679) | GGGAGCTTTGTGT |
| SPU77NAMB04 | (1678) | GGGAGCTTTGTGT |
| SPU77RSA08 | (1678) | GGGAGCTTTGTGT |
| VRL1032 1 IM78 | (1679) | GGGAGCTTTGTGT |
| VRL1290ZIM78 | (1679) | GGGAGCTTTGTGT |
| VRL1516ZIM78 | (1678) | GGGAGCTTTGTGT |
| VRL1887ZIM78 | (1678) | GGGAGCTTTGTGT |
| VRL2230ZIM78 | (1678) | GGGAGCTTTGTGT |
| VRL763ZIM70 | (1678) | GGGAGCTTTGTGT |
| VRL825ZIM79 | (1679) | GGGAGCTTTGTGT |
| ZH501EGY77 | (1678) | GGGAGCTTTGTGT |
| ZH548EGY77 | (1678) | GGGAGCTTTGTGT |
| Consensus | (1681) | GGGAGCTTTGTGT |

Figure 17. Continued

### 6.4. Protein amino acid alignment

### 6.4.1. N protein amino acid alignment

```
                    1
                            4 5
    214445KEN83 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
            56KEN65 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    900085MAU88 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    An1000MAD91 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
        An991MAD91 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
        An999MAD91 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    AnK6087GUI84 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    AnTAMBULEGY94 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    Ar12568RSA71 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    Ar20368RSA81 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
        Ar21229SA00 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
            Ar74RSA55 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
            Ar811MAD79 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
ArD38661SEN81 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
            B1143KEN77 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
            B314KEN62 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
            H1825RSA75 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
            LunyoUGA55 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
            R1662CAR85 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
SPU10301KEN07 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
SPU10302KEN07 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
SPU10307KEN07 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
SPU10315KEN07 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
SPU12002SOM98 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    SPU152RSA08 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    SPU204ANGL85 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    SPU2207KEN07 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    SPU2214KEN07 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    SPU2215KEN07 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    SPU2220KEN07 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    SPU2223KEN07 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
SPU384001KEN97 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
        SPU45ZAMB85 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
SPU52001RSA99 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
        SPU77NAMB04 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
            SPU77RSA08 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    VRL1032ZIM78 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    VRL1290ZIM78 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    VRL1516ZIM78 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    VRL1887ZIM78 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
    VRL2230ZIM78 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
        VRL763ZIM70 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
        VRL825ZIM79 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
        ZH501EGY77 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
        ZH548EGY77 (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
            Consensus (1) MDNYQELAIQFAAQAVDRNEIEQWVREFAYQGFDARRVIELLKQY
```

Figure 18. The deduced $N$ protein amino acid alignment of the 45 strains of Rift Valley fever virus generated in this study. Areas with black lettering and white background show identical nucleotides. Areas with black lettering and green background represent non-similar nucleotides.

214445 KEN83 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY 56KEN65 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY 900085MAU88 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY An1000MAD91 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY An991MAD91 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY An999MAD91 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY AnK6087GUI84 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY AnTAMBULEGY94 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY Ar12568RSA71 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY Ar20368RSA81 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY Ar21229SA00 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY Ar74RSA55 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY Ar811MAD79 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY ArD38661SEN81 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY B1143KEN77 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY B314KEN62 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY H1825RSA75 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY Lunyouga5s (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY R1662CAR85 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU10301KEN07 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU10302KEN07 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU10307KEN07 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU10315KEN07 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU12002SOM98 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU152RSA08 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU204ANGL85 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU2207KEN07 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU2214KEN07 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU2215KEN07 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU2220KEN07 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU2223KEN07 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY
SPU384001KEN97 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY SPU45ZAMB85 (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY
SPU52001RSA99 SPU77NAMB0 4 SPU77RSA08 VRL1032ZIM78 VRL1290ZIM78 VRL1516ZIM78 VRL1887ZIM78 VRL2230ZIM78 VRL763ZIM70 VRL825ZIM79 ZH501EGY77 ZH548EGY77 Consensus
(46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINRY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY (46) GGADWEKDAKKMIVLALTRGNKPRRMMMKMSKEGKATVEALINKY

Figure 18. Continued

214445 KEN83 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG 56KEN65 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
900085 MAU 88 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
An1000MAD91 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
An991MAD91 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
An999MAD91 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
AnK6087GUI84 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
AnTAMBULEGY94 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
Ar12568RSA71 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
Ar20368RSA81 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
Ar21229SA00 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
Ar74RSA55 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG Ar811MAD79 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
ArD38661SEN81 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
B1143KEN77 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
B314KEN62 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
H1825RSA75 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
LunyoUGA55 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
R1662CAR85 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
SPU10301KEN07 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU10302KEN07 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU10307KEN07 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU10315KEN07 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU12002SOM98 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU152RSA08 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU204ANGL85 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU2207KEN07 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU2214KEN07 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU2215KEN07 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU2220KEN07 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU2223KEN07 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
SPU384001KEN97 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU45ZAMB85 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU52001RSA99 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU77NAMB04 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG SPU77RSA08 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
VRL1032ZIM78 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG VRL1290ZIM78 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG VRL1516ZIM78 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG VRL1887ZIM78 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG VRL2230ZIM78 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG VRL763ZIM70 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG VRL825ZIM79 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG ZH501EGY77 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG ZH548EGY77 (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG
Consensus (91) KLKEGNPSRDELTLSRVAAALAGWTCQALVVLSEWLPVTGTTMDG

Figure 18. Continued

## 136

180
214445 KEN83 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI 56KEN65 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
900085MAU88 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
An1000MAD91 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
An991MAD91 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
An999MAD91 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
AnK6087GUI84 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
AnTAMBULEGY94 (136) LSPAYPRHMMHPSFAGMVDPSLPGDYLRAILDAHSLYLLQFSRVI
Ar12568RSA71 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
Ar20368RSA81 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI Ar21229SA00 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI

Ar74RSA55 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
Ar811MAD79 (136) LSPAYPRHMMHPSFAGMVDPSLPGDYLRAILDAHSLYLLQFSRVI
ArD38661SEN81 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
B1143KEN77 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
B314KEN62 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
H1825RSA75 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
LunyoUGA55 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
R1662CAR85 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
SPU10301KEN07 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
SPU10302KENO7 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
SPU10307KEN07 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
SPU10315KEN07 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
SPU12002SOM98 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
SPU152RSA08 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
SPU204ANGL85 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI SPU2207KEN07 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI SPU2214KEN07 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI SPU2215KEN07 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI SPU2220KEN07 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI SPU2223KEN07 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
SPU384001KEN97 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI SPU45ZAMB85 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI SPU52001RSA99 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI SPU77NAMB04 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
SPU77RSA08 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI
VRL1032ZIM78 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI VRL1290ZIM78 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI VRL1516ZIM78 (136) LSPAYPRHMMHPSFAGMVDPSLPGDYLRAILDAHSLYLLQFSRVI VRL1887ZIM78 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI VRL2230ZIM78 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI VRL763ZIM70 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI VRL825ZIM79 (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI ZH501EGY77 (136) LSPAYPRHMMHPSFAGMVDPSLPGDYLRAILDAHSLYLLQFSRVI ZH548EGY77 (136) LSPAYPRHMMHPSFAGMVDPSLPGDYLRAILDAHSLYLLQFSRVI
Consensus (136) LSPAYPRHMMHPSFAGMVDPSLPEDYLRAILDAHSLYLLQFSRVI

Figure 18. Continued
$214445 \mathrm{KEN83}$ (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD 56KEN65 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD 900085MAU88 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD An1000MAD91 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD An991MAD91 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD An999MAD91 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD
AnK6087GUI84 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD
AnTAMBULEGY94 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD
Ar12568RSA71 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD
Ar20368RSA81 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD Ar21229SA00 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD

Ar74RSA55 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD Ar811MAD79 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD
ArD38661SEN81 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD B1143KEN77 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD B314KEN62 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD H1825RSA75 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD LunyoUGA55 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRRGFLKAFGLVD R1662CAR85 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU10301KEN07 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU10302KENO7 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU10307KEN07 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU10315KEN07 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU12002SOM98 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU152RSA08 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU204ANGL85 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU2207KEN07 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU2214KEN07 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU2215KEN07 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU2220KEN07 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU2223KEN07 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD
SPU384001KEN97 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU45ZAMB85 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU52001RSA99 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU77NAMB04 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD SPU77RSA08 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD VRL1032ZIM78 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD VRL1290ZIM78 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD VRL1516ZIM78 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD VRL1887ZIM78 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD VRL2230ZIM78 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD VRL763ZIM70 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD VRL825ZIM79 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD ZH501EGY77 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD ZH548EGY77 (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD Consensus (181) NPNLRGRTKEEVAATFTQPMNAAVNSNFISHEKRREFLKAFGLVD

Figure 18. Continued

|  |  | 246 |
| ---: | :--- | :--- |
| $214445 K E N 83$ | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| 56KEN65 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| $900085 M A U 88$ | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| An1000MAD91 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| An991MAD91 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| An999MAD91 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| AnK6087GUI84 | $(226)$ | SNGKPSAAVKAAAQAYKTAA- |
| AnTAMBULEGY94 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| Ar12568RSA71 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| Ar20368RSA81 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| Ar21229SA00 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| Ar74RSA55 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| Ar811MAD79 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| ArD38661SEN81 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| B1143KEN77 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| B314KEN62 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| H1825RSA75 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| LunyoUGA55 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| R1662CAR85 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU10301KEN07 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU10302KEN07 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU10307KEN07 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU10315KEN07 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU12002SOM98 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU152RSA08 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU204ANGL85 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU2207KEN07 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU2214KEN07 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU2215KEN07 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU2220KEN07 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU2223KEN07 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU384001KEN97 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU45ZAMB85 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU52001RSA99 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU77NAMB04 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| SPU77RSA08 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| VRL1032ZIM78 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| VRL1290ZIM78 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| VRL1516ZIM78 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| VRL1887ZIM78 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| VRL2230ZIM78 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| VRL763ZIM70 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| VRL825ZIM79 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |
| ZH501EGY77 | $(226)$ | SNGKPSAAVMAAAQAYKTAA- |

Figure 18. Continued

### 6.4.2. NSs protein amino acid alignment

## 1

45
214445 KEN83 (1) MDYFPVISVDLQSGRRVVSVEYIIGDGPPRIPYSMVGPCCVFLMH
56KEN65 (1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMR
$900085 M A U 88$ (1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
An1000MAD91 (1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
An991MAD91 (1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
An999MAD91 (1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
AnK6087GUI84
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH

AnTAMBULEGY94
(1) MDYFPVISVDLQSGRRVVSVEYFRGDGPPRIPYSMVGPCCVFLMH

Ar12568RSA71
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH

Ar20368RSA8
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYFRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVYLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIKGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYFRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYFRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYFRGDGPPRIPYSMVGPCCVFLMH
(1) MDYFPVISVDLQSGRRVVSVEYIRGDGPPRIPYSMVGPCCVFLMH

Figure 19. The deduced NSs protein amino acid alignment of 45 strains of Rift Valley fever generated in this study. Areas with black lettering and white background show identical nucleotides. Areas with black lettering and green background represent non-similar nucleotides.

214445 KEN83 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
56KEN65 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASSVAPPPAKPFQRLI
$900085 \operatorname{MAU88}(46)$ HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
An1000MAD91 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
An991MAD91 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
An999MAD91 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
AnK6087GUI84 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
AnTAMBULEGY94 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
Ar12568RSA71 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFVSNVAPPPAKPFQRLI
Ar20368RSA81 (46) HRPSYEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
Ar21229SA00 (46) HRPSHEVRLRFSDFYNAGEFPYRVGLGDFASNVAPPPAKPFQRLI
Ar74RSA55 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
Ar811MAD79 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
ArD38661SEN81 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
B1143KEN77 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
B314KEN62 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
H1825RSA75 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFVSNVAPPPAKPFQRLI
LunyoUGA55 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
R1662CAR85 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU10301KEN07 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU10302KEN07 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU10307KEN07 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU10315KEN07 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU12002SOM98 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU152RSA08 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU204ANGL85 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU2207KEN07 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU2214KEN07 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU2215KEN07 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU2220KEN07 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU2223KEN07 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU384001KEN97 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI SPU45ZAMB85 (46) HRPSHEVRLRFSDFYNVGEFPCRVGLGDFASNVAPPPAKPFQRLI
SPU52001RSA99 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU77NAMB04 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
SPU77RSA08 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
VRL1032ZIM78 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
VRL1290ZIM78 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
VRL1516ZIM78 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
VRL1887ZIM78 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
VRL2230ZIM78 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI VRL763ZIM70 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI VRL825ZIM79 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI ZH501EGY77 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI ZH548EGY77 (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI
Consensus (46) HRPSHEVRLRFSDFYNVGEFPYRVGLGDFASNVAPPPAKPFQRLI

Figure 19. Continued

## 91 <br> 135

$214445 K E N 83$ (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
56KEN65 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRSDD 900085 MAU88 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD An1000MAD91 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD An991MAD91 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
An999MAD91 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
AnK6087GUI84 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
AnTAMBULEGY94 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
Ar12568RSA71 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRSDD
Ar20368RSA81 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
Ar21229SA00 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
Ar74RSA55 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
Ar811MAD79 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
ArD38661SEN81 (91) DLIGHMTLSDFTRFPNLREAISWPLGEPSLAFFDLSSTRVHRNDD
B1143KEN77 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
B314KEN62 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
H1825RSA75 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRSDD
LunyoUGA55 (91) DLIGHMTLSDFIRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
R1662CAR85 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU10301KEN07 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU10302KEN07 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU10307KEN07 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU10315KEN07 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU12002SOM98 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU152RSA08 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU204ANGL85 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU2207KEN07 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU2214KEN07 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU2215KEN07 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU2220KEN07 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU2223KEN07 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU384001KEN97 (91) DLMGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD SPU45ZAMB85 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD
SPU52001RSA99 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD SPU77NAMB04 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD SPU77RSA08 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD VRL1032ZIM78 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD VRL1290ZIM78 (91) DLIGHMTLSDFTRFPNLKEAMSWPLGEPSLAFFDLSSTRVHRNDD VRL1516ZIM78 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD VRL1887ZIM78 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD VRL2230ZIM78 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD VRL763ZIM70 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRSDD VRL825ZIM79 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD ZH501EGY77 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD ZH548EGY77 (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD Consensus (91) DLIGHMTLSDFTRFPNLKEAISWPLGEPSLAFFDLSSTRVHRNDD

Figure 19. Continued

```
136
180
    214445KEN83 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
        56KEN65 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    900085MAU88 (136) IRRDQIATLAMRSCKVTNDLEDSFVGLHRMIVTEAILRGIDLCLL
    An1000MAD91 (136) IRRDQIATLAMRSCKIANDLEDSFVGLHRMIVTEAILRGIDLCLL
        An991MAD91 (136) IRRDQIATLAMRSCKIANDLEDSFVGLHRMIVTEAILRGIDLCLL
        An999MAD91 (136) IRRDQIATLAMRSCKIANDLEDSFVGLHRMIVTEAILRGIDLCLL
        AnK6087GUI84 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
AnTAMBULEGY94 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIATEAILRGIDLCLI
    Ar12568RSA71 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    Ar20368RSA81 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRKIVTEAILRGIDLCLL
        Ar21229SA00 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
            Ar74RSA55 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLI
    Ar811MAD79 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIATEAILRGIDLCLL
ArD38661SEN81 (136) IRRDQIATLAMRSCKITNDLQDSFVGLHRMIVTEAILRGIDLCLL
            B1143KEN77 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
            B314KEN62 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
            H1825RSA75 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
            LunyoUGA55 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGVDLCLL
            R1662CAR85 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMVVTEAILRGIDLCLL
SPU10301KEN07 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
SPU10302KEN07 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
SPU10307KEN07 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
SPU10315KEN07 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
SPU12002SOM98 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLI
    SPU152RSA08 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    SPU204ANGL85 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    SPU2207KEN07 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    SPU2214KEN07 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    SPU2215KEN07 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    SPU2220KEN07 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    SPU2223KEN07 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
SPU384001KEN97 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
        SPU45ZAMB85 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
SPU52001RSA99 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    SPU77NAMB04 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
            SPU77RSA08 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    VRL1032ZIM78 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRIIVTEAILRGIDLCLL
    VRL1290ZIM78 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    VRL1516ZIM78 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIATEAILRGIDLCLL
    VRL1887ZIM78 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
    VRL2230ZIM78 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
        VRL763ZIM70 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
        VRL825ZIM79 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRKIVTEAILRGIDLCLL
            ZH501EGY77 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIATEAILRGIDLCLL
            ZH548EGY77 (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIATEAILRGIDLCLL
            Consensus (136) IRRDQIATLAMRSCKITNDLEDSFVGLHRMIVTEAILRGIDLCLL
```

Figure 19. Continued

```
181
225
    214445KEN83 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
        56KEN65 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
    900085MAU88 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
    An1000MAD91 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
        An991MAD91 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
        An999MAD91 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
        AnK6087GUI84 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSGLIALMEESLML
AnTAMBULEGY94 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIVLMEESLML
    Ar12568RSA71 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
    Ar20368RSA81 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVAPNSALIALMEESLML
        Ar21229SA00 (181) PGFDLMYEIAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
            Ar74RSA55 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
    Ar811MAD79 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIVLMEESLML
ArD38661SEN81 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
    B1143KEN77 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
        B314KEN62 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
        H1825RSA75 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
        LunyoUGA55 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
    R1662CAR85 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
SPU10301KEN07 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
SPU10302KEN07 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
SPU10307KEN07 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
SPU10315KEN07 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
SPU12002SOM98 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
    SPU152RSA08 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
    SPU204ANGL85 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
    SPU2207KEN07 (181) PGFDLVYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
    SPU2214KEN07 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
    SPU2215KEN07 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
    SPU2220KEN07 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
    SPU2223KEN07 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
SPU384001KEN97 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
        SPU45ZAMB85 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
SPU52001RSA99 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
        SPU77NAMB04 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
            SPU77RSA08 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
    VRL1032ZIM78 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
    VRL1290ZIM78 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
    VRL1516ZIM78 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIVLMEESLML
    VRL1887ZIM78 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
    VRL2230ZIM78 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
        VRL763ZIM70 (181) PGFDLMYEVAHVQCVRLLQAAREDISNAVVPNSALIALMEESLML
        VRL825ZIM79 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
            ZH501EGY77 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIVLMEESLML
            ZH548EGY77 (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIVLMEESLML
            Consensus (181) PGFDLMYEVAHVQCVRLLQAAKEDISNAVVPNSALIALMEESLML
```

Figure 19. Continued
226
266
214445 KEN83 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVD56KEN65 (226) RSSLPSMMGRNNWVPVVSPIPDVEMESEEESDDDGFVEVD900085MAU88 (226) RSSLPSMMGRNNWIPVVPPIPDVEMDSGEESDDDGFVEVDAn1000MAD91 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDAn991MAD91 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVD-
An999MAD91 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDAnK6087GUI84 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVD AnTAMBULEGY94 (226) RSSLPSMMGRNNWIPVIPPIPDVEMESEEESDDDGFVEVDAr12568RSA71 (226) RSSLPSMMGRNNWVPVVPPIPDVEIESEEESDDDGFVEVD Ar20368RSA81 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVDAr21229SA00 (226) RSSLPSMMGRNNWVPVVPPIPDVEIESGEESDDDGFVEVD-
Ar74RSA55 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVD-
Ar811MAD79 (226) RSSLPSMMGRNNWIPVIPPIPDVEMESEEESDDDGFVEVD-
ArD38661SEN81 (226) RSSLPSMMGRNNWIPVIPPIPDVEMDSEEESDDDGFVEVN-
B1143KEN77 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVD
B314KEN62 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVD
H1825RSA75 (226) RSSLPSMMGRNNWVPVVPPIPDVEIESEEESDDDGFVEVD-
LunyoUGA55 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVD-
R1662CAR85 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVD
SPU10301KEN07 (226) RLSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU10302KEN07 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU10307KEN07 (226) RLSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU10315KEN07 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU12002SOM98 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU152RSA08 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU204ANGL85 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVDSPU2207KEN07 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU2214KEN07 (226) RTSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU2215KEN07 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU2220KEN07 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU2223KEN07 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU384001KEN97 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU45ZAMB85 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVD-
SPU52001RSA99 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVDSPU77NAMB04 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVD-
SPU77RSA08 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESGEESDDDGFVEVD-
VRL1032ZIM78 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVD
VRL1290ZIM78 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVDVRL1516ZIM78 (226) RSSLPSMMGRNNWIPVIPPIPDVEMESEEESDDDGFVEVDVRL1887ZIM78 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVDVRL2230ZIM78 (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVDVRL763ZIM70 (226) RSSLPSMMGRNNWVPVVPPIPDVEIESEEESDDDGFVEVD VRL825ZIM79 (226) RSSLPSMMGRNNWIPVVPPIPDVEMGSEEESDDDGFVEVD-
ZH501EGY77 (226) RSSLPSMMGRNNWIPVIPPIPDVEMESEEESDDDGFVEVD-
ZH548EGY77 (226) RSSLPSMMGRNNWIPVIPPIPDVEMESEEESDDDGFVEVD-
Consensus (226) RSSLPSMMGRNNWIPVVPPIPDVEMESEEESDDDGFVEVD

Figure 19. Continued

### 6.5. Supplementary strain data

Table 5. Supplementary strain data, acquired from GenBank.

| Strain | Location of Origin | Year | Source | Accession Number |
| :--- | :--- | :--- | :--- | :---: |
| 10911SA00 | Saudi Arabia | 2000 | Human | DQ380170 |
| 73HB1449CAR73 | CAR | 1973 | Human | DQ380162 |
| 74HB59CAR74 | CAR | 1974 | Human | DQ380163 |
| 9800523KEN98 | Kenya | 1998 | Human | DQ380169 |
| AnK3837GUI81 | Guinea | 1981 | Bat | DQ380165 |
| AnK6087GUI84 | Guinea | 1984 | Bat | DQ380166 |
| ArD38388BF83 | Burkina Faso | 1983 | Mosquito | DQ380181 |
| Clone 13 | CAR | 1974 | Plaque pick of 74HB59 | DQ380182 |
| EntebbeUGA44 | Uganda | 1944 | Mosquito | DQ380156 |
| HvB375CAR85 | CAR | 1985 | Human | DQ380161 |
| MgH824MAD79 | Madagascar | 1979 | Human | DQ380144 |
| MP12 | Egypt | 1977 | Vaccine strain | derived |
|  |  | from Zh548 |  |  |
| OS1MAU87 | Mauritania | 1987 | Human |  |
| OS3MAU87 | Mauritania | 1987 | Human | DQ380180 |
| OS8MAU87 | Mauritania | 1987 | Human | DQ380178 |
| OS9MAU87 | Mauritania | 1987 | Human | DQ380177 |
| R1622CAR85 | CAR | 1985 | Human | DQ380179 |
| SA51RSA51 | South Africa | 1951 | Ovine | DQ380160 |
| SA75RSA75 | South Africa | 1975 | Human | DQ380158 |
| SNS | 1944 |  | DQ380175 |  |
| VRL1260ZIM74 | Zimbabawe | 1978 | Bovine | DQ380157 |
| VRL1853ZIM74 | Zimbabwe | 1978 | Bovine | DQ380164 |
| VRL2250ZIM74 | Zimbabwe | 1974 | Bovine | DQ380168 |
| VRL2269ZIM74 | Zimbabwe | 1974 | Bovine | DQ380143 |
| VRL2373ZIM74 | Zimbabwe | 1974 | Bovine | DQ380173 |
| ZC3349EGY78 | Egypt | 1978 | Bovine | DQ380159 |
| ZH1776EGY78 | Egypt | 1978 | Human | DQ380152 |
| ZingaCAR69 | CAR | 1969 | Human | DQ380153 |
| ZM657EGY78 | Egypt | 1978 | Mosquito | DQ380167 |
| ZS6365EGY79 | Egypt | 1979 | Ovine | DQ380146 |
|  |  |  |  |  |

### 6.6. Evolutionary relatedness of two passaged strains of Rift Valley fever virus



Figure 20. Evolutionary relationships of 2 passaged strains of Rift Valley fever virus. The evolutionary history was inferred using the Neighbour-Joining method (Saitou and Nei 1987). The bootstrap consensus tree inferred from 1000 replicates (Felsenstein 1985) is taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than $50 \%$ bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test ( 1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method (Kimura 1980) and are in the units of the number of base substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). There were a total of 1691 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007).

### 6.7. Ethics clearance certificate

| Human Research Ethics Committee (Medical) of the Witwa(formerly Committee for Research on Human Subjects (Medical) |  |  |
| :---: | :---: | :---: |
| Secretariat: Research Office, Room SH10005, 10th floor, Senate House • Telephone: +27 11 717-1234 - Fax: +27 11 339-5708 Private Bag 3, Wits 2050, South Africa |  |  |
| Ref: W-CJ-07 | 234 | 26/01/2007 |
| TO WHOM IT MAY CONCERN: |  |  |
| Waiver: | This certifies that the the Human Research | does not require clearance from Medical). |
| Investigator: | Susan Aitken |  |
| Project title: | Variations in the $S$ reference to the non | ey fever virus with special ling |
| Reason: | Reference strains in participants. | be used. There are no human |


[^0]:    (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC
    (841) GCTGCCCCACCCCCC--ACCCCCAATCCCGACCGTAACCC
    (841) GCTGCCCCACCCCCCCACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC--ACCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCC--ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GСTGCCCCACCCCCC-ACCCCCCAATCCCAACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCTAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCTAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC--ACCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTTACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCC-AATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCC--ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCTACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAGTCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCC--ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACTGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCTCACCCCCC-ACCCCCCAATCCTGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GСTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC--ACCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC (841) GCTGCCCCACCCCCC-ACCCCCCAATCCCGACCGTAACCC
    (841) GCTGCCCCACCCCCC ACCCCCCAATCCCGACCGTAACCC

