EVOLUTIONARY GENETICS OF FOOT-AND-MOUTH DISEASE VIRUS IN KENYA

BY

ABRAHAM KIPROTICH SANGULA

BVM (MAK); MSc (KU)

A THESIS SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES IN FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY OF MAKERERE UNIVERSITY

SEPTEMBER 2010

DECLARATION

I, Abraham Kiprotich Sangula, hereby declare that this is my original work and that it has never been submitted as a thesis for a degree award to any other University. The similarities in work already existent have been expressly acknowledged and what is not indicated is a coincidence of ideas and or is my finding.

Candidate:

Signed…  … Date 22nd September, 2010…

Abraham Kiprotich Sangula

Registration Number: 2007/HD19/4875K

This thesis has been submitted for examination with the approval of the following supervisors.

Supervisor I:

Signed…  … Date…22nd September, 2010..

Vincent Muwanika, PhD

Senior Lecturer, Molecular Genetics

Supervisor II:

Signed…  Date 22nd September, 2010

Hans R. Siegismund, PhD

Associate Professor, Department of Biology, University of Copenhagen

DEDICATION

I dedicate this thesis to my family and all my friends who supported me throughout the study and to the memory of my sister Beatrice Rotich to whom my pursuit of a PhD at Makerere would have been a great inspiration.

ACKNOWLEDGEMENTS

I sincerely appreciate the assistance, constant guidance, constructive discussions and encouragement of my supervisors: Dr Vincent Muwanika, Associate Professor Hans R Siegismund and Professor Graham Belsham. To them I say many thanks for seeing me through the studies. To all members I met at Molecular Genetics lab of MUIENR including; Dr Sylvester Nyakaana, Dr John Okello, Allen Nalugwa and Mary Nakamya; your helpful discussions and interactions are most appreciated. I also thank all staff of MUIENR for their support. Gratitude goes to Dr Charles Masembe (post doc) and my course mates; Chris Ayebazibwe, Frank Mwiine and Sheila Nina Balinda for constant encouragement and assistance. Dr Kirsten Tjørnehøj of the Technical University of Denmark, Lindholm, Rasmus Heller of the Department of Biology, University of Copenhagen and other members of the two institutions whom I interacted with in many ways helped make the research a success. I sincerely acknowledge the patience of my wife Dorcas and sons Brian and Bradley for bearing with my absence during the study period. Mention should also be made to my friends and colleagues in Kenya including Drs Sabenzia Wekesa, Eunice Chepkwony, Arina Odek and Rufus Rumbria; Teresa Kenduiywo, William Birgen, Joseph Kabugi, Isaac Rotich, Jeremiah Ongaro and Rehema Munyasia for their valuable discussions and assistance throughout my study. I am grateful to the Director of Veterinary Services, Kenya, for providing the virus isolates. This work was fully financed by the Danish International Development Agency (DANIDA) under the Livestock-Wildlife Diseases in East Africa Project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

CHAPTER SIX

CHAPTER SEVEN

SUMMARY

Foot-and-mouth disease (FMD) is endemic in Kenya and the East African region in general. The epidemiology of the disease in the region is complex with six of the seven foot-and-mouth disease virus (FMDV) serotypes reported to cause outbreaks and a further complication from the presence of large populations of African buffalo that are known to play an important role in the epidemiology of the southern Africa Territories (SAT) serotypes. A sustained campaign since the 1950s in Kenya to control FMD mainly by vaccination, combined with quarantine and zoo-sanitary measures has been undertaken but with limited success partly due to inadequate knowledge on circulating viruses which diminishes chances of efficient vaccine choices and diagnosis. In order to get insights about processes shaping evolution of the five FMDV serotypes (O, A, C, SAT 1 and SAT 2) recorded to cause outbreaks in Kenya; complete VP1 coding sequences obtained from isolates preserved at the Embakasi FMD laboratory spanning a period of 50 years were analyzed. A total of 129 (O = 46, A = 32, C = 8, SAT 1 = 11 and SAT 2 = 32) sequences were generated in the study while 130 (O = 27, A = 19, C = 9, SAT 1 = 42 and SAT 2 = 33) others were sourced from published reports. Several relatively up-to-date sequence analysis algorithms such as genealogy-based coalescent methods were used to analyse the data. Patterns of evolution among serotype O foot-and-mouth disease viruses in East Africa, revealed the emergence and expansion of the topotype previously designated EA-2 within Kenya and Uganda with cross-border disease transmission within the region and incursions of topotypes EA-3 and EA-4 into Kenya and Uganda probably from neighboring Ethiopia and Sudan. The patterns of genetic variation and distribution of FMD serotype A virus in Kenya, identified two currently circulating virus lineages with a countrywide distribution. Very low nucleotide

diversity and remarkably little change were observed amongst serotype C isolates collected over a period of nearly 40 years. This result was interpreted as being suggestive of re-introductions of the vaccine strain into the field. Analysis of serotype SAT 1 isolates showed that two virus groups; one group exclusive to Uganda while the other is present within Kenya and Tanzania are in circulation. Our results also revealed remarkably low evolutionary rates for the SAT 1 viruses. Serotype SAT 2 analysis results presented evidence for co-circulation of two extremely divergent lineages in Kenya. In general, our observations highlight the importance of characterization of fairly representative virus isolates both on a temporal and spatial scale to be able to discern disease dynamics in a region. It is evident that FMD viruses circulate and are shared within and beyond the East African region necessitating a regional approach to the control of trans-boundary animal diseases.

CHAPTER ONE

Introduction and Literature Review

CHAPTER ONE

1.0 Introduction

1.1 General background

Foot-and-mouth disease (FMD) is a highly contagious disease that affects domestic and wild cloven-hoofed animals such as cattle, sheep, goats, pigs and buffalo (Brooksby, 1982; Thomson et al., 2003). The causative agent, the foot-and-mouth disease virus (FMDV), is a positive-sense, single-stranded RNA genome of 8.5 kb belonging to the *Aphthovirus* genus in the family *Picornaviridae.* Other pathogenic viruses in this family include hepatitis A virus, enteroviruses, rhinoviruses and parechoviruses (Racaniello, 2001). The FMDV genome is translated into a polyprotein which is post-translationally cleaved to yield 12 mature proteins (Belsham, 1993; Sáiz et al., 2002). These are; 4 structural (VP1, VP2, VP3, VP4 also known as 1D, 1B, 1C and 1A respectively) and 8 nonstructural proteins (L^{pro}, 2A, 2B, 2C, 3A, 3B, $3C^{pro}$, and $3D^{pol}$). The structural proteins form the capsid of the virion and with the exception of VP4, are surface exposed (Acharya et al., 1989). VP1 is involved in the host cell interaction via the RGD-dependent integrins and heparin sulphate proteoglycan receptors (Neff et al., 1998; Jackson et al., 2000; Jackson et al., 2007). The nonstructural proteins are involved in replicatory and other biological functions (Grubman and Baxt, 2004; Moffat et al., 2005). While the functions of some of the nonstructural proteins are less well understood, the 2B protein is known to be involved in membrane rearrangements required for viral RNA replication and capsid assembly while 3A and 3B are reported to play a role in virulence and host range (Nunez et al., 2001; Pacheco et al., 2003). The 3C protease is responsible for most of the proteolytic cleavage in the viral polyprotein and RNA replication while 3D forms the core subunit of the RNA-dependent RNA polymerase (Forss et al., 1984; Vakharia et al., 1987).

FMDV occurs as seven immunologically distinct serotypes namely: O, A C, the Southern Africa Territories (SAT) 1, 2, 3 and, Asia1, with a wide spectrum of antigenic and epidemiological subtypes distributed around the world (Kitching, 1998; Knowles and Samuel, 2003). The wide antigenic spectrum is a reflection of genetic variability due to the high mutation rates and population bottlenecks experienced during virus transmission (Drake and Holland, 1999; Escarmis et al., 2006).

In general, FMDV is transmitted by the oronasal route causing in most cases vesicular symptoms of varying severity in the oral mucosa of susceptible species (Brooksby, 1982; Alexandersen et al., 2003). Morbidity is usually high (100%) in FMD but mortality is low particularly in adults with full recovery within four weeks. While a majority of recovered animals clear the virus through the action of macrophages on antibody opsonized virus, a small number develop a persistent infection and become carriers by so far unclear mechanisms (McCullough et al., 1992; Salt, 1993; Woodbury, 1995; Alexandersen et al., 2002). Carrier state development, duration and the role of carrier animals in the epidemiology of FMD is varied (Condy et al., 1985; Dawe et al., 1994; Malirat et al., 1994; Vosloo et al., 1996; Bastos et al., 2000).

FMD is distinguished from other vesicular diseases by laboratory confirmation using techniques most of which are described in the Terrestrial Manual (OIE, 2009) of the World Organization for Animal Health (OIE), and include virus isolation in cell cultures (Hedger, 1968), serological tests to detect virus antigen or antibodies to virus such as: the complement fixation test (Forman, 1974), the serum neutralisation test (Rweyemamu et al., 1977), the enzyme linked immunosorbent assay (ELISA) (Ouldridge et al., 1982) and methods that detect the viral genome such as the PCR (Rodriguez et al., 1992; Rodriguez et al., 1994; Reid et al., 2001). The

choice of method is dependent on the available laboratory facilities and considerations to minimize on virus spread to the environment. The PCR is a highly sensitive and rapid method but requires the use of many serotype-specific primer sets to be able to detect all field strains due to the intratypic variation in FMD viruses (Reid et al., 2001).

Control of FMD is dependent on the disease status of the country, with FMD free countries applying a non-vaccination and stamping out policy in the event of an outbreak while countries with endemic disease adopt measures that depend on their financial and infrastructural capacity (Hunter, 1998; Sobrino et al., 2001; Kitching et al., 2007; Hegde et al., 2009). Vaccination is widely used in countries which try to control FMD, and the availability of antigenically relevant vaccines is always a challenge (Gonzalez et al., 1992; Bastos et al., 2003). The presence of wildlife reservoirs in some endemic settings such as Africa complicates further the epidemiology and control of FMD (Kitching, 1998; Thomson et al., 2003).

1.2 Geographical distribution of FMD

References to the description of the disease indicate that FMD was in Western Europe as far back as the 16th century, and was widely diffused by the 19th century affecting cattle, sheep, pigs and goats (Bulloch, 1927). It was the first animal disease to be recognized as being caused by a filterable agent (virus) in 1898 by Loeffler and Frosch (Brown, 2003), and the existence of antigenic types was demonstrated in the early 1920s (Vallée and Carré, 1922). Vallée and Carré showed the existence of two types designated by their areas of origin: O for Oise and A for Allemagne in France. Waldmann and Trautwein discovered a third serotype termed C (Waldmann and Trautwein, 1926). This was followed by the identification at Pirbright laboratory of three additional serotypes from Southern Africa in the 1940s which were designated

accordingly as Southern African Territories types SAT1, SAT2 and SAT3 and lastly Asia 1 found in the early 1950s in Asian countries of India and Pakistan (Brooksby, 1958).

Most countries of the world have experienced outbreaks of FMD except Greenland, New Zealand and the smaller islands of Oceania (Kitching, 1998). The United States of America, Canada, Mexico, Australia, and Scandinavia haven't had the disease for many years (Knowles and Samuel, 2003). At present the global distribution of serotypes is varied with types O and A occurring in many parts of Africa, Asia and South America (Vosloo et al., 2002; Kitching, 2005; Rweyemamu et al., 2008). While the SAT serotypes are restricted in their distribution mainly to sub-Saharan Africa, Asia1 occurs mostly in south Asia. Type C is at the present rare although it has been recorded in Europe, South America, Africa and Southern Asia with the last reported cases in 2004 in Brazil and Kenya (Roeder and Knowles, 2009).

FMD is endemic in sub-Saharan Africa and all the serotypes except Asia 1 have been recorded in Eastern Africa although the majority of the outbreaks are of serotypes O, A and SAT2 (Vosloo et al., 2002). SAT 3 is the most restricted in distribution in East Africa having been recorded in Uganda only. In the recent past, circulating serotypes in Kenya have included O, A, SAT 1 and SAT 2 with most of the reported outbreaks concentrated in the dairy farming districts of Western, Rift valley and Central provinces (e.g. see Figure 1-1 below for outbreaks reported in year 2006). However, non-reporting of disease outbreaks by pastoralist communities and the general fear of quarantine consequences contribute to the presumed low FMD prevalence in some districts of the country and the apparent concentration of outbreaks in the mostly dairy farming districts (Ngulo, 1980; Ngichabe, 1984).

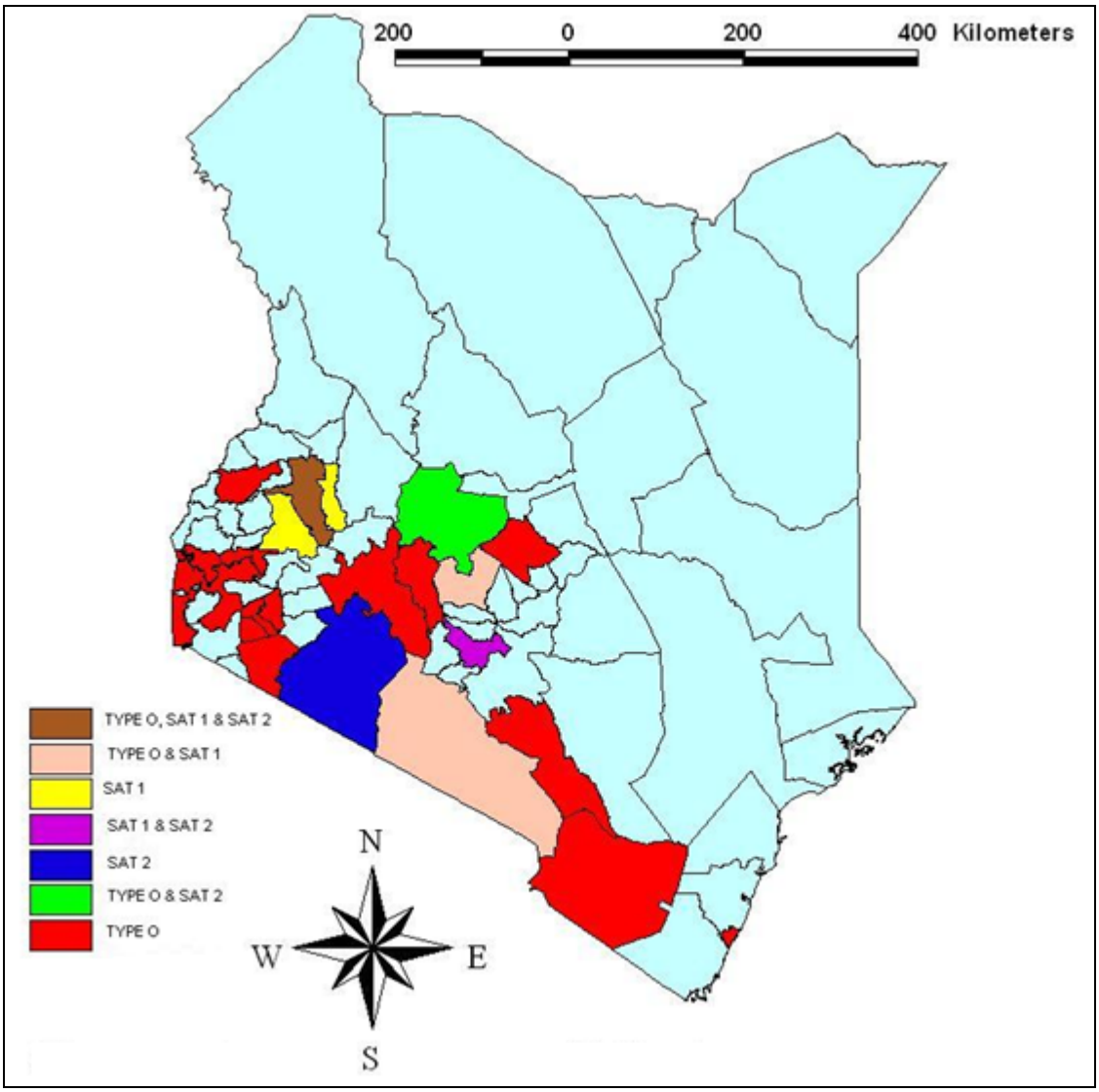


Figure 1-1 Map of Kenya showing the districts reporting FMD outbreaks in 2006

1.3 Economic importance of FMD

Foot-and-mouth disease is one of the most important animal diseases that undermine livestock production and marketing in countries where it is endemic (James and Rushton, 2002). Direct losses include reduced milk yields, loss of draught power, retardation of growth, abortion in pregnant animals, and death in calves and lambs. The severity of the disease depends on the virus strain and the type of animal affected, and the sequelae are found to be more important than the

clinical disease (Woodbury, 1995; Kitching and Hughes, 2002). Indirect losses are attributed to the severe restrictions on trade of animals and animal products. While restrictions on exports undermine poverty alleviation efforts in FMD endemic countries of sub-Saharan Africa, introduction or reemergence of the disease in FMD free countries has devastating economic consequences as exemplified by the 2001 outbreak in the United Kingdom (Samuel and Knowles, 2001b; Thompson et al., 2002). Economic benefits of FMD control in sub-Saharan Africa are highlighted by the southern Africa countries which have successfully employed the zoning system enabling them to export to the international market including the European Union (Brückner et al., 2002).

1.4 Evolutionary epidemiology of FMDV

Like other RNA viruses, the genetic diversity of FMDV is a consequence of its high mutation rate and quasi-species dynamics; the understanding of which is central to disease epidemiology and control (Moya et al., 2004; Domingo et al., 2005; Tully and Fares, 2009). The quasi-species nature of viruses necessitates that diagnostic as well as therapeutic and preventive measures are influenced by the understanding of virus complexity at the population level. For example, a polyvalent immune response is a requirement to prevent the selection of vaccine-escape mutants as partially immune animals may promote selection of variants with altered host cell tropism (Domingo et al., 2003). Various parts of the capsid proteins are recognized by the host immune system including the principal antigenic site located on the G-H loop of VP1 in all serotypes to which much of the antibody response is directed (Baxt et al., 1984; Mateu et al., 1994; Curry et al., 1996). FMDV capsid proteins are known to exhibit substantial genetic variability in the field (Martinez et al., 1992; Meyer et al., 1994; Haydon et al., 2001) and this requires that vaccine strains be updated periodically (Feigelstock et al., 1992).

Mutation produces new strains and phylogenetic lineages whose survival is dependent on the prevailing epidemiological and immunological forces (Grenfell et al., 2004). This can be equated to a continuous evolutionary arms race which is to a large part induced by a strong selection pressure from the immune system of the host to generate variation in the virus capsid surface proteins that the host recognizes in its defense against the virus. New variants of the virus that evade the immune system of the host are under a strong positive selection and may reach high frequencies in the virus population. The high rates of evolution pose great challenges in efforts to control viral infections in general, as an example one of the major challenges in the fight against human immunodeficiency virus (HIV) is the emergence of new resistant strains resulting from selection pressure when anti retroviral drugs are used (Lemey et al., 2005).

Phylogenetic analysis of the VP1coding region of FMDV genome has been widely used to characterize virus genealogy, transmission and molecular evolution (Beck and Strohmaier, 1987; Bastos et al., 2000; Konig et al., 2001; Gurumurthy et al., 2002; Tully and Fares, 2006). Percentage nucleotide differences have been used to guide virus relationships as follows: <1% same epizootic, although viruses from the same epizootic sampled over a 12 month period have been shown to differ by as much as 4% across the VP1 coding region; <7% common origin; 15% same genotype/topotype, and >20% different genetic lineages. Sequence differences of 30 to 50% occur between the seven serotypes with inter-serotypic differences being higher between SAT-types than between the European serotypes (Samuel and Knowles, 2001a). This information has been used to determine sources of outbreaks, follow disease spread and to select appropriate vaccine strains (Vosloo et al., 1996; Jangra et al., 2005; Knowles et al., 2007; Cottam et al., 2008; Klein et al., 2008). Phylogenetic analysis is now widely applied in human and animal health to trace the spread of disease e.g. HIV (Freeman and Herron, 2001), rift valley

fever (Shoemaker et al., 2002) and to predict strains to include in vaccines e.g. for influenza A virus.

1.5 Phylogenetic analysis

Phylogenetic inference is premised on obtaining correctly aligned sequence data sets. Many multiple alignment methods are commonly available and the choice of method will depend on among other considerations on the accuracy and time of execution (Edgar and Batzoglou, 2006). Evolutionary relationships are illustrated using phylogenetic trees which depict the common history of the group of organisms. There are many methods for constructing phylogenetic trees which are either character based e.g. the maximum likelihood and parsimony methods or distance based e.g. the unweighted-pair group method with arithmetic means (UPGMA) and neighbor-joining. Maximum likelihood methods and distance methods are based on an explicit model of substitution, and using an unrealistic evolutionary model can cause serious artifacts in the tree topology (Lockhart et al., 1994). The Bayesian analysis is a recent method that utilizes posterior likelihoods. Recent applications of coalescent theory (Kingman, 2000) together with Bayesian analysis methods have allowed novel theoretical analyses of genealogical processes in populations (Drummond and Rambaut, 2007) and have been used to infer details on the past demographic history of natural populations including FMDV (Drummond et al., 2005; Tully and Fares, 2008). However, it is important to keep in mind that not all tree-building algorithms are able to infer true trees of existing phylogenetic information in sequences and selecting the best method is still poorly known (Xia and Lemey, 2009). It is worth noting that evolutionary forces such as selection and recombination also bias estimates of phylogenies (Schierup and Hein, 2000; Moya et al., 2004).

1.6 Statement of the problem

Foot-and-mouth disease is one of the most important animal diseases affecting livestock production in Kenya with outbreaks of many serotypes occurring in a single year. The success of economic flagship projects for the country's vision 2030 such as the creation of disease free zones is premised on the control of key diseases such as FMD. The current efforts to control FMD in the country include: surveillance, diagnosis, vaccinations and enforcement of zoo-sanitary measures. These efforts are, however, hampered by inadequate knowledge about temporal and spatial dynamics of the virus that would guide the selection of vaccines to be used and also improve on diagnostic tests and thus enhance disease control measures. This study was therefore aimed at generating reliable information that can be utilized to secure an increased and stable supply of animal products in Kenya and thereby contribute to alleviation of poverty.

1.7 Aims and Objectives

The global aim of this study is to generate information on the dynamics of FMDV that is relevant for the enforcement of knowledge-based disease control strategies in the eastern Africa region.

1.7.1 General objective

To determine patterns of circulation of FMDV and study evolutionary processes shaping FMDV serotypes in Kenya.

1.7.2 Specific objectives

1. To determine the evolutionary relationships and the processes shaping FMDV serotype O isolates in East Africa.
2. To infer genetic divergence and diversity of FMDV serotype A viruses circulating in Kenya.

3. To estimate patterns of sequence divergence and evolution of FMDV serotype C isolates in Kenya.

4. To apply genealogy-based coalescent methods to infer evolutionary history of FMDV serotype SAT 1 in eastern Africa.

5. To establish patterns of evolution of FMDV serotype SAT 2 in Kenya.

1.8 Justification of the study

Eastern Africa which includes Kenya has probably the most complicated FMD status in the world (Rweyemamu et al., 2008) and yet little is known about the evolutionary forces which contribute to the complexity of the epidemiology of the disease in this region. Such knowledge can be obtained from analyzing the temporal and spatial genetic variation within and between different FMDV serotypes. Previous studies on some of the serotypes in East Africa have alluded to the complex epidemiology and genetic diversity of the viruses (Sahle, 2004). In southern Africa, establishing that the African buffalo harbors and transmits the SAT serotypes to livestock has contributed to the successful zoning and fencing strategy although the strategy has serious environmental implications (Brückner et al., 2002). In Kenya, there is widespread livestock-wildlife interaction. Some studies have shown that the African buffalo in Kenya harbor the SAT serotypes, which is similar to the situation in southern Africa (Anderson et al., 1979; Bronsvoort et al., 2008). Results from this study will contribute to our insight into the evolutionary processes that cause the complex situation of FMD in East Africa. This knowledge is critical for the desired control of the disease in the region. The study is of importance to regional transboundary disease control and is likely to transform our understanding of disease movements, virus threats, and provide essential information on which national and regional control strategy can be

developed. The information is directly relevant to global efforts in disease control which are being spearheaded by international organizations such as, the Food and Agriculture Organization (FAO) and the World Animal Health Organization (OIE). This includes the implementation of regional roadmaps on FMD control to the year 2020 along the FMD progressive control pathway. The first stage requires that viral threats and risk control points are identified in order to develop national strategies.

1.9 Organization of the thesis

This thesis is written and arranged in the form of published and submitted manuscripts in peer reviewed journals. Chapter one presents an introduction with a general background to the study, and outlines the objectives and the molecular genetics methods applied to achieve the study objectives. Chapter two presents an overview of the general methods employed in the study including virus isolates and genetic laboratory analyses. Chapter three addresses specific objective 1 and details the evolutionary relationships among serotype O foot-and-mouth disease viruses in East Africa. It is formatted for the *Infection, Genetics and Evolution* journal, where the manuscript has been submitted. Specific objective 2 is addressed by chapter four which presents patterns of genetic variation and distribution of FMD serotype A virus in Kenya and it is formatted for the *Preventive Veterinary Medicine* journal, where the manuscript has been submitted. Chapter five addresses specific objective 3 and presents evidence for probable FMDV serotype C vaccine strain introductions in the field in Kenya. It is formatted for the *Epidemiology and Infection* journal, where the manuscript has been accepted for publication. Chapter six addresses specific objective 4 and details the evolutionary analysis of foot-and-mouth disease virus serotype SAT 1 isolates from East Africa. It is formatted for the *BMC Evolutionary Biology* journal, where the manuscript has been submitted. And finally chapter seven addresses specific

objective 5 and presents evidence for co-circulation of two extremely divergent serotype SAT 2 lineages in Kenya. It is formatted for the *Archives of Virology* journal where the manuscript has been submitted. At the end of chapter seven, general conclusions about the findings and recommendations are presented.

References

Acharya, A., Fry, E., Stuart, D., Fox, G., Rowlands, D., Brown, F., 1989. The three-dimensional structure of foot and mouth disease virus at 2.9 Å resolution. Nature 337, 709-716.

Alexandersen, S., Zhang, Z., Donaldson, A.I., 2002. Aspects of the persistence of foot-and-mouth disease virus in animals—the carrier problem. *Microbes and Infection* **4**, 1099–1110.

Alexandersen, S., Zhang, Z., Donaldson, A.I., Garland, A.J.M., 2003. The pathogenesis and diagnosis of foot-and-mouth disease. *Journal of Comparative Pathology* **129**, 1–36.

Anderson, E.C., Doughty, W.J., Anderson, J., Paling, R., 1979. The pathogenesis of foot-and-mouth disease in the African buffalo (*Syncerus Caffer*) and the role of this species in the epidemiology of the disease in Kenya. *Journal of Comparative Pathology* **89**, 511 - 519.

Bastos, A.D.S., Boshoff, C.I., Keet, D., Bengis, R.G., Thomson, G.R., 2000. Natural transmission of foot-and-mouth disease virus between African buffalo (*Syncerus caffer*) and impala (*Aepyceros melampus*) in the Kruger National park, South Africa. *Epidemiology and Infection* **124**, 591-598.

Bastos, A.D.S., Haydon, D.T., Sangare´, O., Boshoff, C.I., Edrich, J.L., Thomson, G.R., 2003. The implications of virus diversity within the SAT 2 serotype for control of foot-and-mouth disease in sub-Saharan Africa. *Journal of General Virology* **84**, 1595–1606.

Baxt, B., Morgan, D.O., Robertson, B.H., Timpone, C.A., 1984. Epitopes on foot-and-mouth disease virus outer capsid protein VP1 involved in neutralization and cell attachment. *Journal of Virology* **51**, 298-305.

Beck, E., Strohmaier, K., 1987. Subtyping of European foot-and-mouth disease virus strains by nucleotide sequence determination. *Journal of Virology* **61**, 1621-1629.

Belsham, G.J., 1993. Distinctive features of foot-and-mouth disease virus, a member of the picorna virus family; aspects of virus protein synthesis, protein processing and structure. *Progress in Biophysics and Molecular Biology* **60**, 241-260.

Bronsvoort, B.M.D.C., Parida, S., Handel, I., McFarland, S., Fleming, L., Hamblin, P., Kock, R., 2008. Serological survey for foot-and-mouth disease virus in wildlife in eastern Africa and estimation of test parameters of a nonstructural protein enzyme-linked immunosorbent assay for buffalo. *Clinical and Vaccine Immunology* **15**, 1003–1011.

Brooksby, J.B., 1958. The virus of foot-and-mouth disease. *Advances in Virus Research* **5**, 1–37.

Brooksby, J.B., 1982. Portraits of viruses: foot-and-mouth disease virus. *Intervirology* **18**, 1-23.

Brown, F., 2003. The history of research in foot-and-mouth disease. *Virus Research* **91**, 3-7.

Brückner, G.K., Vosloo, W., Plessis, B.J.A.D., Kloeck, P.E.L.G., Connoway, L., Ekron, M.D., Weaver, D.B., Dickason, C.J., Schreuder, F.J., Marais, T., Mogajane, M.E., 2002. Foot and mouth disease: the experience of South Africa. *OIE Scientific and Technical Review* **21**, 751-764.

Bulloch, W., 1927. Foot-and-mouth disease in the 16th century. *Journal of Comparative Pathology* **40**, 75-76.

Condy, J.B., Hedger, R.S., Hamblin, C., Barnett, I.T.R., 1985. The duration of foot-and-mouth disease virus carrier state in African buffalo (i) in the individual animal and (ii) in a free-living herd. Comparative lmmunology, *Microbiology and Infectious diseases* **8**, 259- 265.

Cottam, E.M., Wadsworth, J., Shaw, A.E., Rowlands, R.J., Goatley, L., Maan, S., Maan, N.S., Mertens, P.P.C., Ebert, K., Li, Y., Ryan, E.D., Juleff, N., Ferris, N.P., Wilesmith, J.W., Haydon, D.T., King, D.P., Paton, D.J., Knowles, N.J., 2008. Transmission pathways of foot-and-mouth disease virus in the United Kingdom in 2007. *PLoS Pathogens* **4**, 1-8.

Curry, S., Fry, E., Blakemore, W., Abu-Ghazaleh, R., Jackson, T., King, A., Lea, S., Newman, J., Rowlands, D., Stuart, D., 1996. Perturbations in the surface structure of A22 Iraq foot-and-mouth disease virus accompanying coupled changes in host cell specificity and antigenicity. *Structure* **4**, 135-145.

Dawe, P.S., Flanagan, F.O., Madekurozwa, R.L., Sorensen, K.J., Anderson, E.C., Foggin, C.M., Ferris, N.P., Knowles, N.J., 1994. Natural transmission of foot-and-mouth disease from African buffalo (*Syncerus caffer*) to cattle in a wildlife area of Zimbabwe. *Veterinary Record* **134**, 230-232.

Domingo, E., Escarmı´s, C., Baranowski, E., Ruiz-Jarabo, C.M., Carrillo, E., Nu´n˜ez, J.I., Sobrino, F., 2003. Evolution of foot-and-mouth disease virus. *Virus Research* **91**, 47-63.

Domingo, E., Pariente, N., Airaksinen, A., Gonzalez-Lopez, C., Sierra, S., Herrera, M., Grande-Perez, A., Lowenstein, P., Manrubia, S., Lazaro, E., Escarmis, C., 2005. Foot-and-mouth disease virus evolution: exploring pathways towards virus extinction. *Current Topics in Microbiology and Immunology* **288**, 149-173.

Drake, J.W., Holland, J.J., 1999. Mutation rates among RNA viruses. *Proceedings of the National Academy of Sciences, USA* **96**, 13910–13913.

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

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

Edgar, R.C., Batzoglou, S., 2006. Multiple sequence alignment. *Current Opinion in Structural Biology* **16**, 368–373.

Escarmis, C., Lazaro, E., Manrubia, S.C., 2006. Population bottlenecks in quasispecies dynamics. *Current Topics in Microbiology and Immunology* **299**, 141–170.

Feigelstock, D., Mateu, M.G., Piccone, M.E., De Simone, F., Brocchi, E., Domingo, E., Palma, E.L., 1992. Extensive antigenic diversification of foot-and-mouth disease virus by amino acid substitutions outside the major antigenic site. *Journal of General Virology* **73**, 3307-3311.

Forman, A.J., 1974. A study of foot-and-mouth disease virus strains by complement fixation. I. A model for the fixation of complement by antigen/antibody mixtures. J*ournal of Comparative Pathology* **72**, 397-405.

Forss, S., Strebel, K., Beck, E., Schaller, H., 1984. Nucleotide sequence and genome organization of foot-and-mouth disease virus. *Nucleic Acids Research* **12**, 6587-6601.

Freeman, C., Herron, J.C., 2001. *Evolutionary Analysis*. Prentice-Hall New Jersey.

Gonzalez, M., Mateu, M.G., Martinez, M.A., Carrillo, C., F., S., 1992. Comparison of capsid protein VP1 of the viruses used for the production and challenge of foot and mouth disease vaccines in Spain. *Vaccine* **10**, 732-734.

Grenfell, B.T., Pybus, O.G., Gog, J.R., Wood, J.L.N., Daly, J.M., Mumford, J.A., Holmes, E.C., 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* **303**, 327-332.

Grubman, M.J., Baxt, B., 2004. Foot-and-mouth disease. *Clinical Microbiology Reviews* **17**, 465–493.

Gurumurthy, C.B., Sanyal, A., Venkataramanan, R., Tosh, C., George, M., Hemadri, D., 2002. Genetic diversity in the VP1 gene of foot-and-mouth disease virus serotype Asia 1. *Archives of Virology* **147**, 85–102.

Haydon, D.T., Bastos, A.D., Knowles, N.J., Samuel, A.R., 2001. Evidence for positive selection in foot-and-mouth disease virus capsid genes from field isolates. *Genetics* **157**, 7–15.

Hedger, R.S., 1968. The isolation and characterisation of foot and mouth disease virus from clinically normal herds of cattle in Botswana. *Journal of Hygiene* **66**, 27-37.

Hegde, N.R., Maddur, M.S., Rao, P.P., Kaveri, S.V., Bayry, J., 2009. Thermostable foot-and-mouth disease virus as a vaccine candidate for endemic countries: A perspective. *Vaccine* **27**, 2199–2201.

Hunter, P., 1998. Vaccination as a means of control of foot-and-mouth disease in sub-saharan Africa. *Vaccine* **16**, 261-264.

Jackson, A.L., O'Neill, H., Maree, F., Blignaut, B., Carrillo, C., Rodriguez, L., Haydon, D.T., 2007. Mosaic structure of foot-and-mouth disease virus genomes. *Journal of General Virology* **88**, 487–492.

Jackson, T., Shepherd, D., Denyer, M., Blakemore, W., King, A.M.Q., 2000. The epithelial integrin alphavbeta6 is a receptor for foot-and-mouth disease virus. *Journal of Virology* **74**, 4949-4956.

James, A.D., Rushton, J., 2002. The economics of foot and mouth disease. *OIE Scientific and Technical Review* **21**, 637-644.

Jangra, R.K., Tosh, C., Sanyal, A., Hemadri, D., Bandyopadhyay, S.K., 2005. Antigenic and genetic analyses of foot-and-mouth disease virus type A isolates for selection of candidate vaccine strain reveals emergence of a variant virus that is responsible for most recent outbreaks in India. *Virus Research* **112**, 52–59.

Kingman, J.F.C., 2000. Origins of the coalescent: 1974–1982. Genetics 156, 1461–1463.

Kitching, P., Hammond, J.M., Jeggo, M., Charleston, B., Paton, D.J., Rodriguez, L., Heckert, R., 2007. Global FMD control—is it an option? . *Vaccine* **25**, 5660–5664.

Kitching, R.P., 1998. A Recent history of foot-and-mouth disease. *Journal of Comparative Pathology* **118**, 89 -108.

Kitching, R.P., 2005. Global epidemiology and prospects for control of foot-andmouth disease. *Current Topics in Microbiology and Immunology*. pp. 133–148.

Kitching, R.P., Hughes, G.J., 2002. Clinical variation in foot and mouth disease: sheep and goats. *OIE Scientific and Technical Review* **21**, 505-512.

Klein, J., Hussain, M., Ahmad, M., Afzal, M., Alexandersen, S., 2008. Epidemiology of foot-and-mouth disease in Landhi dairy colony, Pakistan, the world largest buffalo colony. *Virology Journal* **5**:53.

Knowles, N.J., Samuel, A.R., 1995. Polymerase chain reaction amplification and cycle sequencing of the 1D gene of foot-and-mouth disease viruses. *Session of the research group of the standing technical committee of the European commission for the control of foot-and-mouth disease.* FAO, Rome, 19-22 September 1994, Vienna, Austria.

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

Knowles, N.J., Wadsworth, J., Reid, S.M., Swabey, K.G., El-Kholy, A.A., El-Rahman, A.O.A., Soliman, H.M., Ebert, K., Ferris, N.P., Hutchings, G.H., Statham, R.J., King, D.P., Paton, D.J., 2007. Foot-and-mouth disease virus serotype A in Egypt. *Emerging Infectious Diseases* **13**.

Konig, G., Blanco, C., Knowles, N.J., Palma, E.L., Maradei, E., Piccone, M.E., 2001. Phylogenetic analysis of FMD viruses isolated in Argentina. *Virus Genes* **23**, 175-181.

Lemey, P., Dooren, S.V., Vandamme, A.-M., 2005. Evolutionary dynamics of human retroviruses investigated through full-genome scanning. *Molecular Biology and Evolution* **22**, 942–951.

Lockhart, P.J., Steel, M.A., Hendy, M.D., Penny, D., 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Molecular Biology and Evolution* **11**, 605-612.

Malirat, V., Mello, P.A.D., Tiraboschi, B., Beck, E., Gomes, I., Bergmann, I.E., 1994. Genetic variation of foot-and-mouth disease virus during persistent infection in cattle. *Virus Research* **34**, 31-48.

Martinez, M.A., Dopazo, J., Hernandez, J., Mateu, M.G., Sobrino, F., Domingo, E., Knowles, N.J., 1992. Evolution of the capsid protein genes of foot-and-mouth disease virus: antigenic variation without accumulation of amino acid substitutions over six decades. *Journal of Virology* **6**, 3557-3565.

Mateu, M.G., Hernandez, J., Martinez, M.A., Feigelstock, D., Lea, S., Perez, J.J., Giralt, E., Stuart, D., Palma, E.L., Domingo, E., 1994. Antigenic heterogeneity of a foot-and-mouth

disease virus serotype in the field is mediated by very limited sequence variation at several antigenic sites. *Journal of Virology* **68**, 1407-1417.

McCullough, K.C., Simone, F.D., Brocchi, E., Cappuci, L., Crowther, J.R., Kihm, U., 1992. Protective immune response against foot and mouth diseases. *Journal of Virology* **66**, 1835-1840.

Meyer, R.F., Pacciarini, M., Hilyard, E.J., Ferrari, S., Vakharia, V.N., Donini, G., Brocchi, E., Molitor, T.W., 1994. Genetic variation of foot-and-mouth disease virus from field outbreaks to laboratory isolation. *Virus Research* **32**, 299-312.

Moffat, K., Howell, G., Knox, C., Belsham, G.J., Monaghan, P., Ryan, M.D., Wileman, T., 2005. Effects of foot-and-mouth disease virus nonstructural proteins on the structure and function of the early secretory pathway: 2BC but not 3A blocks endoplasmic reticulum-to-golgi transport. *Journal of Virology* **79**, 4382-4395.

Moya, A., C.Holmes, E., González-Candelas, F., 2004. The population genetics and evolutionary epidemiology of RNA viruses. *Nature Reviews | Microbiology* **2**, 279-288.

Neff, S., Sa´-Carvalho, D., Rieder, E., Mason, P.W., Blystone, S.D., Brown, E.J., Baxt, B., 1998. Foot-and-mouth disease virus virulent for cattle utilizes the integrin avb3 as its receptor. *Journal of Virology* **72**, 3587–3594.

Ngichabe, C.K., 1984. Foot-and-mouth disease in Kenya: Surveillance, vaccination procedures and policy. *Kenya Veterinarian* **8**, 8-13.

Ngulo, W.K., 1980. Strategies and costs of animal disease control with indications for research on foot-and-mouth disease. *Kenya Veterinarian* **4**, 27-28.

Nunez, J.I., Baranowski, E., Molina, N., Ruiz-Jarabo, C.M., Sanchez, C., Domingo, E., Sobrino, F., 2001. A single amino acid substitution in nonstructural protein 3A can mediate

adaptation of foot-and-mouth disease virus to the guinea pig. *Journal of Virology* **75**, 3977–3983.

OIE, 2009. Chapter 2.1.5 . Foot and mouth disease. *OIE Terrestrial Manual*. OIE Paris.

Ouldridge, E.J., Barnett, P., Hingley, P., Head, M., Rweyemamu, M.M., 1982. The use of an indirect sandwich ELISA assay for the differentiation of FMD virus strains. *The ELISA* **22**, 152-160.

Pacheco, J.M., Henry, T.M., O'Donnell, V.K., Gregory, J.B., Mason, P.W., 2003. Role of nonstructural proteins 3A and 3B in host range and pathogenicity of foot-and-mouth disease virus. *Journal of Virology* **77**, 13017–13027.

Racaniello, V. (Ed.), 2001. *Picornaviridae: The viruses and their replication*. Lippincott Williams & Wilkins Philadelphia.

Reid, S.M., Ferris, N.P., Hutchings, G.H., Clercq, K.D., Newman, B.J., Knowles, N.J., Samuel, A.R., 2001. Diagnosis of foot-and-mouth disease by RT-PCR:use of phylogenetic data to evaluate primers for the typing of viral RNA in clinical samples. *Archives of Virology* **146**, 2421–2434.

Rodriguez, A., Martinez-Salas, E., Dopazo, J., Davila, M., Sáiz, J.C., Sobrino, F., 1992. Primer design for specific diagnosis by PCR of highly variable RNA viruses: typing of foot and mouth disease virus. *Virology* **189**, 363-367.

Rodriguez, A., Nunez, J.I., Nolasco, G., Ponz, F., Sobrino, F., De Blas, C., 1994. Direct PCR detection of foot and mouth disease virus. *Journal of Virological Methods* **47**, 345-349.

Roeder, P.L., Knowles, N.J., 2009. Foot-and-mouth disease virus type C situation: the first target for eradication? , *The Global control of FMD - Tools, ideas and ideals*. FAO, Rome, 14-17 October 2008, Erice, Italy.

Rweyemamu, M., Roeder, P., Mackay, D., Sumption, K., Brownlie, J., Leforban, Y., Valarcher, J.-F., Knowles, N.J., Saraiva, V., 2008. Epidemiological patterns of foot-and-mouth disease worldwide. *Transboundary and Emerging Diseases* **55**, 57-72.

Rweyemamu, M.M., Pay, T.W.F., Hingley, P.J., 1977. Serological differentiation of foot-and-mouth disease virus strains in relation to selection of suitable vaccine strains. *Developments in Biological Standards* **35**, 205-214.

Sahle, M., 2004. An epidemiological study on the genetic relationships of foot and mouth disease viruses in East Africa. *PhD thesis, University of Pretoria*, Pretoria, South Africa.

Sáiz, M., Núñez, J.I., Jimenez-Clavero, M.A., Baranowski, E., Sobrino, F., 2002. Foot-and-mouth disease virus: biology and prospects for disease control. *Microbes and Infection* **4**, 1183–1192.

Salt, J.S., 1993. The carrier state in foot and mouth disease - an immunological review. *British Veterinary Journal* **149**, 207-223.

Samuel, A.R., Knowles, N.J., 2001a. Foot-and-mouth disease type O viruses exhibit genetically and geographically distinct evolutionary lineages (Topotypes). *Journal of General Virology* **82**, 609-621.

Samuel, A.R., Knowles, N.J., 2001b. Foot-and-mouth disease virus: cause of the recent crisis for the UK livestock industry. *Trends in Genetics* **17**, 421–424.

Schierup, M.H., Hein, J., 2000. Consequences of recombination on traditional phylogenetic analysis. *Genetics* **156**, 879–891.

Shoemaker, T., Boulianne, C., Vincent, M.J., Pezzanite, L., Al-Qahtani, M.M., Al-Mazrou, Y., Khan, A.S., Rollin, P.E., Swanepoel, R., Ksiazek, T.G., Nichol, S.T., 2002. Genetic

analysis of viruses associated with emergence of rift valley fever in Saudi Arabia and Yemen, 2000-01. *Emerging Infectious Diseases* **8**, 1415-1420.

Sobrino, F., Saiz, M., Jimenez-Clavero, M.A., Nunez, J.I., Rosas, M.F., Baranowski, E., Ley, V., 2001. Foot-and-mouth disease virus: a long known virus, but a current threat. *Veterinary Research* **32**, 1–30.

Thompson, D., Muriel, P., Russell, D., Osborne, P., Bromley, A., Rowland, M., Creigh-Tyte, S., Brown, C., 2002. Economic costs of the foot and mouth disease outbreak in the United Kingdom in 2001. *OIE Scientific and Technical Review* **21**, 675-687.

Thomson, G.R., Vosloo, W., Bastos, A.D.S., 2003. Foot and mouth disease in wildlife. *Virus Research* **91**, 145-161.

Tully, D.C., Fares, M.A., 2006. Unravelling selection shifts among foot-and-mouth disease virus (FMDV) serotypes. *Evolutionary Bioinformatics* **2**, 211-225.

Tully, D.C., Fares, M.A., 2008. The tale of a modern animal plague: tracing the evolutionary history and determining the time-scale for foot and mouth disease virus. *Virology* **382**, 250–256.

Tully, D.C., Fares, M.A., 2009. Shifts in the selection-drift balance drive the evolution and epidemiology of foot-and-mouth disease virus. *Journal of Virology* **83**, 781–790.

Vakharia, V.N., Devaney, M.A., Moore, D.M., Dunn, J.J., Grubman, M.J., 1987. Proteolytic processing of foot-and-mouth disease virus polyproteins expressed in a cell-free system from clone-derived transcripts. *Journal of Virology* **61**, 3199–3207.

Vallée, H., Carré, H., 1922. *Sur la pluralité des virus aphteux*. CR Acad. Sci. Paris 174, 1498–1500.

Vosloo, W., Bastos, A.D.S., Kirkbride, E., Esterhuysen, J.J., Van Rensburg, D.J., Bengis, R.G., Keet, D.W., Thomson, G.R., 1996. Persistent infection of African buffalo (Syncerus caffer) with SAT-type foot-and-mouth disease viruses: rate of fixation of mutations, antigenic change and interspecies transmission. *Journal of General Virology* **77**, 1457-1467.

Vosloo, W., Bastos, A.D.S., Sangare, O., Hargreaves, S.K., Thomson, G.R., 2002. Review of the status and control of foot and mouth disease in sub-Saharan Africa. *OIE Scientific and Technical Review* **21**, 437-447.

Waldmann, O., Trautwein, K., 1926. Experimentelle Untersuchungen über die Pluralität des Maul-und Klauenseuchevirus. *Berl. Tierärztl. Wochenschrift*. **42**, 569-571.

Woodbury, E.L., 1995. A review of the possible mechanisms for the persistence of foot and mouth disease virus. *Epidemiology and Infection* **114**, 1-13.

Xia, X., Lemey, P., 2009. Assessing substitution saturation with DAMBE. In: Philippe, L., Salemi, M., Vandamme, A.-M. (Eds.), *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press, Cambridge, pp. 615-630.

CHAPTER TWO

Materials and Methods

CHAPTER TWO

Materials and Methods

2.1 Virus Isolates

FMDV isolates of the five serotypes were selected considering a broad temporal and geographic distribution representation from the virus collection at the Embakasi FMD laboratory in Nairobi. Isolates selected were collected between 1964 and 2008 and consisted of mainly epithelial tissue and cell culture suspensions preserved at -70^{o}C. Due to the long term storage at --70^{o}C, some of the virus isolates were passaged according to standard laboratory procedures at least once in baby hamster kidney (BHK) monolayer cells (Mowat and Chapman, 1962), before RNA extraction. The virus collection at Embakasi is as summarized in Table 2-1 and the details of the isolates selected for each serotype are found later in the respective chapters.

2.2 General laboratory methods

2.2.1 FMDV RNA extraction and cDNA synthesis

RNA was extracted from both cell culture and epithelial tissue samples at the Embakasi laboratory using standard procedures in the QIAamp® Viral RNA kit, according to the manufacturer's instructions (Qiagen, Germany). This was followed by cDNA synthesis using Ready-To-Go TM You-Prime First-Strand Beads (GE Healthcare Life Sciences, Sweden) with random hexamer (pdN6) primers, according to the manufacturer's instructions. cDNA was synthesized from a total of 430 FMDV isolates (Table 2-1).

2.2.2 PCR amplification and Cycle sequencing

Serotype specific primer sets targeting amplification of the entire VP1 coding region of the FMDV genome e.g. Knowles and Samuel (1995) and others designed specifically from East African FMDV sequences Balinda et al., (2009) were used. Details of primer sets used are described in the respective chapters. Generally, PCR reactions were performed in a final volume of 50 µl using 2-5 ng of cDNA, and 2.5 U of AmpliTaq DNA polymerase (Applied Biosystems), 200µM of each dNTP (dATP, dCTP, dGTP, dTTP) and 1.5 mM $MgCl_2$ with 0.2 pmol of each primer. The thermocycling conditions were generally as follows: initial activation of AmpliTaq DNA polymerase at 95˚C for 5 min was followed by denaturation at 95˚C for 15 s, followed by annealing at 60˚C for 2 min and elongation at 72˚C for 1 min 20 s. The denaturation, annealing and extension steps were repeated for 30 cycles followed by a final extension at 72˚C for 5 min. Modifications to the thermocycling conditions involved a touchdown procedure of 5-7 cycles, where in each next cycle, annealing temperature was decreased by one degree. The PCR products were analysed by electrophoresis on 2% agarose gels using ethidium bromide staining and ΦX 174-RF DNA (Amersham, Biosciences) as a molecular weight marker. Prior to cycle sequencing, the PCR products were cleaned using QIAquick® PCR purification kit (Qiagen).

The cleaned PCR products were then cycle sequenced in both directions following the dideoxy-chain termination method (Sanger et al., 1977) using Big dye Terminator V 3.1 kit (Applied Biosystems) and ran on an automated DNA Sequencer (ABI PRISM® 3700) by Macrogen in South Korea. The same primers were used for the cycle sequencing and PCR amplification. The sequence data were processed using the program Sequencher 4.8 (Gene Code Corporation, USA). A total of 253 complete VP1 coding sequences were obtained (Table 2-1).

2.2.3 Sequence analysis

The sequences were multiply aligned and analysed by different evolutionary genetics software to answer specific questions as detailed in the respective chapters that follow. Multiple alignments were obtained by log-expectation comparisons using MUSCLE (Edgar, 2004) incorporated within Geneious 4.7.6 (Drummond et al., 2009).

Preliminary sequence analysis using Popstr software (Siegismund, unpublished) was undertaken to determine genetically identical sequences. A total of 129 sequences were used for further analysis (Table 2-1).

Models of evolution were selected using hierarchical likelihood-ratio tests or Akaike information criteria implemented in PAUP*(v. 4.0 beta 10) (swofford, 2001) and MrModeltest (v 2.2) (Nylander et al., 2004).

Sequence characteristics including nucleotide diversity and amino acid variation in data sets were determined using MEGA 4 (Tamura et al., 2007). Selection forces were detected using genetic algorithms available on the Datamonkey server (Pond and Frost, 2005) while recombination was investigated using the RDP software (Martin and Rybicki, 2000) and the Datamonkey facility.

Phylogenetic relationships and or demographic histories were estimated using Bayesian Markov Chain Monte Carlo (MCMC) methods available in BEAST 1.4.7 (Drummond and Rambaut, 2007) and MrBayes (Huelsenbeck and Ronquist, 2001).

References

Balinda, S.N., Belsham, G.J., C.Masembe, Sangula, A.K., Siegismund, H.R., Muwanika, V.B., 2009. Molecular characterization of SAT 2 foot-and-mouth disease virus from post-

outbreak slaughtered animals: implications for disease control in Uganda. *Epidemiology and Infection* doi:10.1017/S0950268809991427.

Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., Wilson, A., 2009. *Geneious v4.6.* Available from http://www.geneious.com/.

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

Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**.

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

Knowles, N.J., Samuel, A.R., 1995. Polymerase chain reaction amplification and cycle sequencing of the 1D gene of foot-and-mouth disease viruses. *Session of the research group of the standing technical committee of the European commission for the control of foot-and-mouth disease*. FAO, Rome, 19-22 September 1994, Vienna, Austria.

Martin, D., Rybicki, E., 2000. RDP: detection of recombination amongst aligned sequences. *Bioinformatics* **16**, 562-563.

Mowat, G.N., Chapman, W.G., 1962. Growth of foot-and-mouth disease virus in a fibroblastic cell derived from hamster kidneys. *Nature*, **194**, 253-255.

Nylander, J.A.A., 2004. *MrModeltest v2.* Program distributed by the author Evolutionary Biology Centre, Uppsala University.

Pond, S.L.K., Frost, S.D.W., 2005. A genetic algorithm approach to detecting lineage-specific variation in selection pressure. *Molecular Biology and Evolution* **22**, 478–485.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, USA* **74**, 5463-5467.

Swofford, D.L., 2003. *PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version4*. Sinauer Associates, Sunderland, Massachusetts.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596–1599.

Table 2-1 Summary of the number of FMD virus isolates selected and used in the study

Serotype	**Total Collection in Bank**	**RNA/cDNA obtained (No.)**	**VP1 sequences obtained (No.)**	**Sequences in final analysis (No.)**
O	473	142	81	46
A	196	106	59	32
C	13	13	12	8
SAT 1	91	62	23	11
SAT 2	388	117	78	32
Total	1161	430	253	129

CHAPTER THREE

Diversity and transboundary mobility of serotype O foot-and-mouth disease virus in East Africa: implications for vaccination policies

Diversity and transboundary mobility of serotype O foot-and-mouth disease virus in East Africa: implications for vaccination policies

Sheila N. Balinda[1§], Abraham K. Sangula[1, 4§], Rasmus Heller[2], Vincent B. Muwanika[1], Graham J. Belsham[3], Charles Masembe[1], Hans R. Siegismund[2]

[§] Equal-Joint first authorship

[1]Makerere University, Institute of Environment and Natural Resources P.O.Box 7298, Kampala, Uganda

[2]Department of Biology, Ole Maaløes Vej 5, DK-2200, Copenhagen N, Denmark

[3]National Veterinary Institute, Technical University of Denmark, Lindholm, DK-4771 Kalvehave, Denmark

[4] Foot-and-Mouth Disease Laboratory, Embakasi, P. O. Box 18021, 00500, Nairobi, Kenya

Corresponding author:

Abraham K. Sangula

Makerere University Institute of Environment and Natural Resources,

Molecular Biology Laboratory,

P.O. Box 7298, Kampala, Uganda

Fax: +256 41 530134 Phone + 256775633133

E-mail: aksangula@muienr.mak.ac.ug / aksangula@gmail.com

Abstract

Foot-and-mouth disease (FMD) virus serotype O has been responsible for most reported outbreaks of the disease in East Africa. A sustained campaign for the past 40 years to control FMD mainly by vaccination, combined with quarantine and zoosanitary measures has been undertaken with limited success. We investigated the genetic relationships among serotype O strains in eastern Africa using complete VP1 coding region sequences obtained from 46 FMD virus isolates collected in Kenya in the years 1964 to 2008 and 8 Ugandan isolates collected between 1999 and 2006. In addition, 21 selected FMDV sequences from Genbank representing reference strains from eastern Africa and elsewhere were included in the Bayesian inference analyses and the detection of selection forces. The results confirmed previous observations that eastern Africa harbours four distinct topotypes (clades with >15% sequence divergence). All but one strain isolated post-2000 belonged to either topotypes EA-2, EA-3 or EA-4. EA-1 had only one strain isolated post-2000 and all three vaccines have been based on strains in this topotype. The estimated dN/dS ratios across the individual codons of the entire VP1 coding region revealed that purifying (negative) selection constituted the dominant evolutionary force. Cross-border disease transmission within the region has been suggested with probable incursions of topotypes EA-3 and EA-4 into Kenya and Uganda from neighboring Ethiopia and Sudan. We conclude that the vaccines have probably been effective in controlling EA-1, but less so for the other topotypes and propose a more comprehensive representation of topotypes in the development of new vaccines in recognition of the considerable diversity and transboundary nature of serotype O.

Keywords: FMDV, Topotype, Selection, East Africa

Introduction

Foot-and-mouth disease (FMD) is a vesicular disease affecting cloven hoofed domestic and wild animals (Coetzer et al., 1994). It is mainly transmitted through direct contact involving infected aerosol deposition in the respiratory tract and mechanical transfer of the virus to susceptible animals. Disease dissemination may occur indirectly through contaminated surfaces and products (Alexandersen and Mowat, 2005). The causative agent, foot-and-mouth disease virus (FMDV), comprises a positive sense RNA within a protein capsid. The RNA genome has a single large open reading frame (ORF) encoding a polyprotein that is eventually processed into 12 mature proteins. The virus capsid contains sixty copies of each of the four structural proteins with VP1-3 exposed on the outside while VP4 is located internally (Acharya, 1989; Belsham, 2005). FMD viruses, like other members of *Picornaviridae*, are very diverse and exist in seven serotypes: O, A, C, Asia 1 and the Southern African Territories (SAT) 1, SAT 2 and SAT 3 as determined using virus neutralization assays. The VP1 coding region has been used extensively in molecular characterization and to determine evolutionary relationships (Bastos et al., 2003; Knowles and Samuel, 2003; Sangare et al., 2003), resulting in a comprehensive nucleotide data library of this part of the virus coding region.

In East Africa, serotype O has been responsible for most of the reported outbreaks (Vosloo et al., 2002). A recent study has found that type O viruses in this region belong to four topotypes, EA-1 to EA-4 (Ayelet et al., 2009), based on the concept of topotype classification (Samuel & Knowles, 2001; Knowles & Samuel, 2003). Over the last 40 years, a sustained campaign to control FMD has been undertaken in East Africa with limited apparent success. This has mainly been through the use of vaccination to control outbreaks and for routine prophylaxis applied together with quarantine measures adopted during FMD outbreaks which are usually coupled to zoo-sanitary measures to prevent indirect disease dissemination. A single vaccine

strain for serotype O FMDV (K77/78) has been in use in Kenya and Uganda for over 25 years but other vaccine strains used in the past include K120/64 and K83/79. The limited efficiency of vaccination in controlling FMD can be attributed, in part, to the existence of six of the seven known serotypes in this region (Vosloo et al., 2002) since immunity to one serotype does not confer protection against another. Different strains within a single serotype also exhibit significant diversity, as was shown initially by Bedson et al. (1927) through the distinction between two type A strains using cross neutralization tests (Brown, 2003). Consequently, antigenic diversity may limit the efficiency of vaccines to combat different strains of the same serotype (Mumford, 2007).

For effective disease control, an understanding of the mechanisms responsible for the maintenance of FMDV is required. Key to this process is improved insight into the evolutionary forces such as selection and genetic drift operating on populations of the virus. In addition, as previously suggested by Heath et al. (1996), the evolution of aphthoviruses and enteroviruses (other members of the family *Picornaviridae*) includes recombination and thus analyses towards its identification in FMD viral sequence data sets may be desirable as it is known to bias estimates of phylogenies (Schierup & Hein, 2000).

In this study, the complete VP1 coding region sequences were used to determine the evolutionary relationships and processes shaping this most prevalent FMD virus serotype in eastern Africa.

Materials and Methods

Virus Isolates

A total of 54 type O virus isolates obtained between the years 1964 and 2008 from Kenya and Uganda were included in this study (Table 1). Figure 1 indicates their respective districts of

origin. The Kenyan isolates were part of the collection of outbreak samples submitted to the Embakasi FMD laboratory, Nairobi, and were mainly generated from epithelium samples from which viruses were isolated and passaged in BHK cells. All the Ugandan samples were from oropharyngeal fluid and swab samples (from buccal and foot lesions) collected during field outbreak investigations.

RNA extraction, cDNA synthesis, PCR and Cycle sequencing

RNA was extracted from cell culture supernatants and directly from clarified samples of oropharyngeal fluid or swab samples using the QIAamp® Viral RNA kit (Qiagen, Germany) according to the manufacturer's instructions. The cDNA synthesis was carried out using Ready-To-Go ^{TM} You-Prime First – Strand Beads (GE Healthcare Life Sciences, Sweden) with random hexamer (pdN_6) primers.

PCR reactions of two overlapping fragments (to cover the VP1 region) using 0.2 pmol primers 8-A PN 64 with 8-A PN 98 and 8-A PN 84 with 8-A PN 85; SNB-84 with SNB-85 (designed specifically for Eastern African viruses) as well as primers previously used to target the entire VP1 region O-1C_{564} and FMD-2A_{34}(NK-72) (Table 2) were performed in a final volume of 50 µl using 2-5 ng of cDNA, and 2.5 U of AmpliTaq Gold® DNA polymerase (Applied Biosystems), 200µM of each dNTP (dATP, dCTP, dGTP, dTTP) and 1.5 mM $MgCl_2$. Following the activation of AmpliTaq Gold® DNA polymerase at 95˚C for 5 min, reaction mixtures were denatured at 95˚C for 15 s followed by 60˚C for 2 min to allow for primer annealing. For each cycle, a chain elongation step at 72˚C for 1 min 20 s was allowed. This process was repeated 30 times and final extension continued at 72˚C for 5 min. The resultant PCR products were analysed by electrophoresis on 2% agarose gels using as a molecular weight marker ΦX 174-RF DNA (Amersham, Biosciences). Purification of the PCR products to remove

oligonucleotide primers, dNTPs and enzyme was achieved with a QIAquick® PCR purification kit (Qiagen). Cycle sequencing was performed in both directions using Big dye Terminator V 3.1 kit (Applied Biosystems) and ran on an automated DNA Sequencer (ABI PRISM® 3700) by Macrogen in Korea.

Sequence Analysis

Sequencher software 4.8 (Gene Code Corporation) was used to assemble the sequences and multiple alignments by log-expectation comparison were carried out using MUSCLE incorporated within Geneious 4.7.6 software (Drummond *et al.,* 2009). Determination of the most appropriate model of evolution was performed using a hierarchical likelihood-ratio test of 24 models implemented in PAUP*(v. 4.0 beta 10) (Swofford, 2001) and MrModeltest (v 2.2) (Nylander *et al.*, 2004). The GTR+I+G model was used with rate variation following a gamma distribution across sites and a proportion of invariable sites was identified as the best model of evolution and was used in subsequent analyses. Most previous studies concerned with topotype designation within FMDV serotypes have used distance-based phylogenetic approaches such as neighbour-joining or UPGMA trees without taking the temporal scale of sampling into account. Instead, we applied the Markov Chain Monte Carlo (MCMC) method available in BEAST 1.4.7 (Drummond et al., 2005), incorporating the sampling dates in a Bayesian framework that co-estimates divergence times as well as tree topology. This approach also allows several different demographic scenarios to be tested together with the phylogenetic relationship between the samples. First, we tested the fit of the data to two different demographic models, the constant population size and the Bayesian skyline plot (BSP) and to two types of molecular clock, a strict and a relaxed (uncorrelated exponential) molecular clock (Drummond & Rambaut, 2007). This was done by calculating approximate Bayes factors comparing the marginal likelihoods of each

model in Tracer v1.4 (http://beast.bio.ed.ac.uk/) following recommendations in Suchard et al (2001). The 10,000 trees saved during the MCMC chain were summarized using TreeAnnotator v1.4.7 (part of the BEAST package) to produce a maximum clade credibility (MCC) tree, discarding the first 1000 trees as burn-in. We used DnaSP v. 4.10.9 (Rozas et al., 2003) to estimate the average number of nucleotide substitution per site (Nei 1987, eq. 10.20) between putative topotypes.

The translation, alignment and identification of variable sites within amino acids were carried out using MEGA 4 (Tamura et al., 2007). Overall, site-specific selection pressures acting within this region were determined by estimating the *d*N/*d*S using two likelihood procedures available in the HyPhy package and accessed through the Datamonkey web-server. These methods are the single-likelihood ancestor counting (SLAC) method and the random-effects (REL) method (Kosakovsky Pond & Frost, 2005).

Recombination was investigated using RDP (Martin & Rybicki, 2000), GENECONV (Padidam et al., 1999), Chimaera (Posada & Crendall, 2001) and Maxichi (Posada and Crendall, 2001) within the RDP3 beta 22 software (Martin et al., 2005). The analyses using each of the mentioned methods were conducted twice for repeatability.

Results

Phylogenetic relationships

The complete VP1 coding region of 54 FMDV samples have been amplified by RT-PCR and then the sequences determined. The 46 samples with origins from Kenya were obtained between the years 1964-2008 while the 8 Ugandan samples were obtained during the years 1999-2006 within the districts of origin summarized in Figure 1. The sequences used here were selected to include representation of serotype O strains from West Africa and elsewhere in addition to the

East African strains. Analyses involving the detection of recombination revealed no evidence for this evolutionary process within the dataset (of VP1 coding sequences), hence phylogenetic analyses were performed under the assumption that the sequences in this study were not generated by recombination. The Bayes factor test showed no significant support for the more complex BSP demographic model, but strongly supported a relaxed (uncorrelated exponential) rather than a strict clock. Branch rate heterogeneity was severe with a coefficient of variation of 0.90 (95% HPD: 0.79–1.01) and inferred rates on individual branches ranging from $9.81 * 10^{-5}$ to $1.02 * 10^{-2}$. Hence, results from the constant population size model evolving under a relaxed clock are reported throughout.

The overall mean rate of substitutions per site per year was estimated to be $2.76 * 10^{-3}$ (95% HPD: $1.84 * 10^{-3} – 3.63 * 10^{-3}$). This resulted in an estimated root of the tree at 164 years (95% HPD: 80–299 years). Phylogenetic relationships inferred from the maximum clade credibility tree confirmed that the type O samples from eastern Africa belonged to four distinct viral topotypes (with a minimum sequence divergence of 15% among them) termed EA-1–EA-4 (Figure 2). Four samples did not cluster into the topotype designations (K/120/64, K/51/92, K/52/92 and K/131/85) but should probably be attributed to topotypes EA-1 (first three samples) and EA-2 (latter sample), respectively. They were not sufficiently divergent from these topotypes to be designated as separate topotypes. Interestingly, the West African sample included in this analysis grouped as a sister taxon to EA-1, but with sufficient sequence divergence to be designated in a separate topotype, WA (following the nomenclature in Ayelet et al., 2009). Hence, WA and EA-1 collectively formed a sister clade to EA-2. The posterior probability of some of the nodes separating topotypes were not very high: EA-1/WA: 0.29; EA-1,WA /EA-2: 0.46; EA-3/EA-4: 0.47.

The earliest topotype identified (EA-1) was found in Kenya and Uganda. The majority of the viruses in this topotype are from the period 1964–1996. However, K/31/07 is an exception being the only recent strain from this topotype. The topotype EA-2 was represented by Kenyan, Ugandan, Tanzanian and Malawian isolates. It was the predominant topotype within the East African samples, and all but one (K/31/07) of the recently collected (2000-2008) samples from Kenya and Uganda belonged to this topotype. EA-3 and EA-4 are mainly restricted to Ethiopia, Eritrea and Sudan in recent times, although they did previously circulate in Kenya and Uganda. The most recent representative of topotype EA-4 in Uganda occurred about 10 years ago (strain U/97/99 in Figure 2), and EA-3 was last observed in Kenya about twenty years ago (K/114/87).

Amino acids

The alignment of the deduced amino acid sequences for the entire VP1 proteins of 54 Kenya and Uganda viruses belonging to topotypes EA-1 and EA-2 is shown in Figure 3. Two hundred and eighty three (44%) variable sites have been determined across the 642 nucleotides of these viruses which encode substitutions to 80 (37.4%) of the 214 amino acids (Figure 3). The majority of the amino acid variations were clustered within antigenic site 1(A) corresponding to the G-H loop (residues 135-161) and the carboxy terminus region (residues 191-212).

In both topotypes, some level of amino acid conservation was observed within the G-H loop, with the cysteine amino acid residue (135) at the base of the loop conserved within all the strains. Similarly, the RGD motif within this loop of VP1, which binds to the cellular integrin receptor located at residues 146-148 was completely conserved. Residues flanking the RGD motif were also conserved as far as the -2 position. Conservation was also observed as far as +4 (RGDLQVL), in most of the viruses with the exception of U/13B/04 (EA-2) at +2 with P (Proline) instead of Q (glutamine) and UGA/5/96 (EA-1) with V (Valine) instead of L (Leucine)

at position +4. A variation between topotypes EA-1 and EA-2 at position 69-70 was observed, with the former comprising T (Threonine), A (Alanine) and the latter A (Alanine), S (Serine) respectively.

Selection

To identify the forces shaping evolution, the ratio of non-synonymous to synonymous substitutions (dN/dS) was determined. An overall mean of dN/dS of 0.12 for the entire VP1 coding region was estimated. However, the per site dN/dS analyses revealed 130 of the sites to be under negative selection and no positively selected sites as determined using the SLAC method at $P = 0.1$. On the other hand, REL, at a specified significance of Bayes Factor = 50 identified three positively selected sites at positions 45, 47 and 109, and 156 negatively selected sites.

Discussion

The East African dataset generated in this study is the largest and the most geographically representative of FMDV serotype O assembled so far for Kenya and Uganda in a single analysis. Comparing our findings with those of Tully and Fares (2008), who used the same Bayesian phylogenetic approach to infer divergence dates within the whole FMDV tree, some interesting differences are apparent. The mean rate of evolution inferred in the present study was slightly lower than that obtained by Tully and Fares (2008), but 95% HPD intervals overlapped considerably between the studies. However, the inferred age of the root of the tree presented in this study was much older (164 years) than that derived for all serotype O viruses by Tully and Fares (2008) (92 years), even though a broader diversity of topotypes was included there. This is probably attributable to the fact that they applied a demographic model of exponential growth, which was not supported by our data. The appropriateness of inferring any dynamic demographic model from a broad variety of FMDV topotypes is questionable, given that they cannot be

considered to constitute a single population, hence violating the assumptions of demographic inferences based on the coalescent (Halloran & Holmes, 2009). It is also possible that the sample composition plays an important role in dating divergences in this manner, as we have shown that rates of evolution differ substantially among branches in the type O phylogeny. Given the variability and inconsistencies in the estimated evolutionary rates and the divergence times between studies, the possibility of a radically changed time scale of historical divergence within FMDV should be addressed in future studies.

The phylogenetic relationships found here are consistent in some respects with those previously determined (Samuel & Knowles, 2001; Knowles & Samuel, 2003; Ayelet et al., 2009) in that four topotypes within serotype O exist within the Eastern African region. However, our phylogeny differed from those previous published by grouping the only West African sample within the clade comprising the four East African topotypes, making the East African topotypes paraphyletic. Of the known topotypes within East Africa, only EA-1 and EA-2 were observed in Kenya and Uganda within the last 10 years. Of these, EA-2 was by far the most dominant. On the other hand, topotypes EA-3 and EA-4 were mainly found in Ethiopia, Eritrea and Sudan, although both topotypes appeared to have previously circulated in Kenya and Uganda; EA-3 was found in Kenya in the 1980s while an incursion of EA-4 into Uganda occurred in 1999.

Vaccines against type O in East Africa have all been based on virus isolates from the topotype EA-1. Only one virus sample from all the post-2000 samples included in this study was attributable to this topotype, whereas the remaining 31 were from the other three topotypes. This could suggest that vaccine programs have been successful, but that type O vaccines generally do not cross-protect well against other topotypes, confirming what has been observed for type A (Araujo et al. 2002). Based on the dated phylogeny inferred in this study, we conclude that all

four East African topotypes probably already existed in 1964, the time of the collection of the first strain on which a vaccine was based. Furthermore, there was no indication of positive selection in the amino acid positions that were variable among the vaccine strains and the other topotypes. This leads us to conclude that the diversification of type O FMDV was most likely not an evolutionary response to the selection pressure afforded by vaccine programs. The current dominance of topotypes EA-2 and EA-3 in East Africa probably just reflects the lack of protection induced by the previous vaccine programs against these topotypes. This instills some confidence that vaccine programs can be used to control the disease as long as the full diversity of the serotype is taken into account and vaccines are not exclusively based on single topotypes. However, as EA-1 has continued to circulate, in Kenya at least, in the recent past (2007), the vaccines have not been able to eliminate this topotype completely. This could be attributable to lack of comprehensive vaccine programs and to the survival of virus strains in untreated livestock reservoirs or potentially type O is being maintained in wildlife populations. Anti-O antibodies have been detected among buffalo populations in Uganda indicating previous exposure to serotype O (Ayebazibwe et al. 2010).

Despite the limited representation of virus sequences from some countries in the region, the phylogenetic relationships provided indications of transboundary events within the East African region. The Sudanese (1999) isolate, within topotype EA-3 appeared closely related to Ethiopian isolates obtained in 2004 and 2007. Similarly, the close evolutionary relationship between the Malawian (MAL/1/98), Tanzanian (TAN/7/98), Ugandan (UG/kumi/2002) and the Kenyan K/141/00 strains suggests that FMD virus strains can transgress borders relatively quickly. This is perhaps not very surprising, as uncontrolled animal movement across the border points is known to occur within the East African region but our evolutionary analysis underlines

the transboundary nature of type O FMDV. More surprising was the observation that the only West African sample included in this study grouped as a sister taxon to EA-1 and hence within the East African topotype diversity. If true, this further stresses the lack of geographical sorting in type O genetic diversity but it should be examined in more detail by including more West African samples along with dense East African sampling.

Taken together, the considerable natural diversity and the transboundary nature of FMDV serotype O leads us to propose comprehensive control programs that use frequently administered vaccines based on strains from all four topotypes in East Africa, and possibly even also on WA strains. This is necessary to prevent "exotic" varieties of type O originating from introduced livestock from starting new epidemics. Of course it cannot be excluded that FMDV has the capability to escape vaccine programs by acquiring adaptive resistance to the selection pressure, given its rapid rate of evolution. However, on the basis of our evolutionary analyses we conclude that vaccines have been reasonably successful in controlling the topotype they were based upon.

An overall $dN/dS < 1$ suggests that negative selection (purifying selective pressure) within the VP1 coding region is dominant resulting in amino acid conservation as observed previously (Tully and Fares, 2006). Nevertheless, positive selection on a few codon sites has been revealed in this study. In other studies, several sites were identified as undergoing positive selection within this widely geographically distributed serotype. Interestingly, even the entire RGD motif, a region known for extensive conservation in all the serotypes was suggested as undergoing positive selection in this serotype (Tully and Fares, 2006). The differences in observed sites undergoing positive selection in these two studies may be attributed to the nature of sampling with the current study mainly focusing on the East African strains while the analysis of Tully and Fares (2006) included samples with a wide geographic distribution coupled to

various sampling times for an attempted random representation of the world wide situation (Tully and Fares, 2006).

This work has analyzed the evolutionary relationships within serotype O of FMDV in Eastern Africa. The occurrence of transboundary disease transmission was evident, and effective control strategies to combat this should be adopted by applying vaccines that more broadly reflect the diversity of East African topotypes. This should be implemented by the respective governments in a joint regional effort.

Acknowledgements

The authors thank the Director of Veterinary Services, Kenya, for providing the Kenyan virus isolates used in the study and the Managing Director of KEVEVAPI. Dr Sabenzia Wekesa, Teresa Kenduiywo, William Birgen and Eugene Arinaitwe are particularly appreciated for excellent technical assistance, and Sheila Naka- wombe for her support. This work was funded by the Danish International Development Agency (DANIDA) under the Livestock- Wildlife Diseases in East Africa Project.

References

Acharya, R., Fry, E., Stuart, D., Fox, G., Rowlands, D., Brown, F., 1989. The three-dimensional structure of foot-and-mouth disease virus. *Nature* **337**, 709-716.

Araujo Jr J.P., Montassier H.J. & Pinto A.A. 2002. Extensive antigenic and genetic variation among foot and-mouth disease type A viruses isolated from the 1994 and 1995 foci in Sao Paulo, Brazil. *Veterinary Microbiology* **84**, 15-27.

Ayebazibwe, C, Mwiine F. N., Balinda S. N., Tjørnehøj, K., Masembe C., Muwanika, V. B., Okurut, A. R. A., Siegismund, H. R. and Alexandersen S_2010. Antibodies against foot-and-mouth disease (FMD) virus in African buffalos (*Syncerus caffer*) in selected national

parks in Uganda (2001–2003). *Transboundary and Emerging Diseases*, Online Early, DOI 10.1111/j.1865-1682.2010.01147.x

Ayelet, G., Mahapatra M., Gelaye E., Egziabher B. G., Rufeal T., Sahle M., Ferris N. P., Wadsworth J., Hutchings G. H., Knowles N. J., 2009. Genetic characterization of foot-and-mouth disease viruses, Ethiopia, 1981–2007. *Emerging Infectious Diseases* **15**, 1409-1417.

Belsham, G. J. 2005. Translation and replication of FMDV RNA. In 'Foot and mouth disease virus'. *Currents Topics Microbiology and Immunology* **288**, 43-70.

Brown, F. 2003. The history of research in FMDV. *Virus Research* **91**, 3-7.

Coetzer, J. A. W., Thomsen, G. R., Tustin, R. C., Kriek, N. P. J. 1994. Foot-and-mouth disease. In: *Infectious Diseases of Livestock with Special Reference to Southern Africa.* Cape Town: Oxford University Press.

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

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

Drummond A. J., Ashton, B., Cheung, M., Heled J., Kearse, M., Moir, R., Stones-Havas,S., Thierer T., Wilson, A., 2009. *Geneious v. 4.7*. http://www.geneious.com

Fares, M. A., Moya, A., Escarmis, C., Baranowski, E., Domingo, E., Barrio, E., 2001. Evidence of positive selection in the capsid protein-coding region of foot-and-mouth disease virus (FMDV) subjected to experimental passage regimens. *Molecular Biology and Evolution* **18**, 10-21.

Halloran, M. E., Holmes, E. C., 2009. Evaluating vaccination programs using genetic sequence data. *American Journal of Epidemiology* **170**, 1464-1466.

Haydon, D. T., Bastos, A. D., Knowles, N. J., Samuel, A. R., 2001. Evidence for positive selection in foot-and-mouth disease virus capsid genes from field isolates. *Genetics* **157**, 7-15.

Heath, L., van der Walt, E., Varsani, A., Martin, D. P., 2006. Recombination patterns in Aphthoviruses mirror those found in other picornaviruses. *Journal of Virology*, **80**(23), 11827-11832.

Kosakovsky Pond, S. L., Frost, S. D. W., 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* **21**, 2531–2533.

King, A. M. Q., McCahon, D., Saunders, K., Newman, J. W. I., Slade, W. R., 1985. Mutiple sites of recombination within the RNA genome of foot-and-mouth disease virus. *Virus Research* **3**, 373-384.

Knowles, N. J., Samuel, A. R., 1994. Polymerase chain reaction amplification and cycle sequencing of the 1D (VP1) gene of foot-and-mouth disease viruses. Paper presented at the *session of the Research group of the standing Technical committee of European commission for the control of FMD*, Vienna Austria, 19-22, September, 1994.

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

Martin, D., Rybicki, E., 2000. RDP: detection of recombination amongst aligned sequences. *Bioinformatics* **16**, 562-563.

Martin, D. P., William, C., Posada, D., 2005. RDP2: Recombination detection and analysis from sequence alignments. *Bioinformatics* **21**, 260-262.

Mumford J. A., 2007. Vaccines and viral antigenic diversity. *Revue Scientifique et Technique (International Office of Epizootics)* **26**, 69-90.

Nei, M., 1987. Molecular evolutionary genetics. Columbia University Press, New York, NY.

Nylander, J. A., Ronquist, F., Huelsenbeck, J. P., Nieves-Aldrey, J. L., 2004. Bayesian phylogenetic analysis of combined data. *Systematic Biology* **53**, 47-67.

Padidam, M., Sawyer, S., Fauquet, C. M., 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology* **265**, 218-225.

Posada, D., Crendall, K. A., 2001. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proceedings of the National Academy Sciences USA* **98,** 13757-13762.

Rozas, J., Sanchez-Del Barrio, J. C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by coalescent and other methods. *Bioinformatics* **19**, 2496-2497.

Samuel, A. R., Knowles, N. J., 2001. Foot-and-mouth disease type O viruses exhibit genetically and geographically distinct evolutionary lineages (topotypes). *Journal of General Virology* **82**, 609-621.

Schierup, M. H., Hein, J., 2000 Consequences of recombination on traditional phylogenetic analysis. *Genetics* **156**(2): 879-891.

Swofford, D. L. (2001). *PAUP* Phylogenetic Analysis Using Parsimony (*and other methods). Version 4.0*. Sunderland: Sinauer Associates.

Tamura, K., Dudley, J., Nei M., Kumar S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution* **24**, 1596-1599.

Tully, D. C., Fares, M. A., 2006. Unraveling selection shifts among foot-and-mouth disease virus (FMDV) serotypes. *Evolutionary Bioinformatics* **2**, 211-225.

Tully, D. C., Fares, M. A., 2008. The tale of a modern animal plague: Tracing the evolutionary history and determining the scale for foot-and-mouth disease virus. *Virology* **382**(2):250-256

Vosloo, W., Bastos, A. D. S., Sangare, O., Hargreaves, S. K., Thomson, G. R., 2002. Review of the status and control of foot and mouth disease in sub-Saharan Africa. *Revue Scientifique et Technique de L' Office International des Epizooties* **21**, 437-449.

Wilson, V., Taylor. P., Desselberger, U., 1988. Crossover regions in foot-and-mouth disease virus (FMDV) recombinants correspond to high local secondary structure. *Archives of Virology* **102**, 131-139.

Table 3-1 Foot-and-mouth disease virus strains and sequences used in the study

Isolate Designation	District/Country	Accession No.	Topotype
K120/64	Laikipia, Kenya	This study	EA-1
BFS/UK/67	United Kingdom	AY593815	EURO-SA
K77/78	Nakuru, Kenya	This study	EA-1
K83/79	Nyeri, Kenya	AJ303511	EA-1
K101/80	Laikipia, Kenya	This study	EA-3
K103/82	Thika, Kenya	This study	EA-1
HKN/6/83	Hong Kong	AJ294919	CATHAY
ISA/8/83	Indonesia	AJ303503	ISA-1
K11/84	Kiambu, Kenya	This study	EA-3
K131/85	Kiambu, Kenya	This study	EA-2
K114/87	Kiambu, Kenya	This study	EA-3
K121/91	Kiambu, Kenya	This study	EA-1
K51/92	Nakuru, Kenya	This study	EA-1
K52/92	Kiambu, Kenya	This study	EA-1
K11/93	Kiambu, Kenya	This study	EA-2
K34/93	Laikipia, Kenya	This study	EA-1
GHA/5/93	Ghana	AJ303488	WA
K29/95	Kiambu, Kenya	This study	EA-1
K56/95	Kiambu, Kenya	This study	EA-1

UGA/5/96	Uganda	AJ296327	EA-1
K82/98	Kiambu, Kenya	This study	EA-2
MAL/1/98	Malawi	DQ165074	EA-2
UGA/17/98	Uganda	WRL	EA-4
K117/99	Nakuru, Kenya	This sudy	EA-2
SUD/1/99	Sudan	WRL	EA-2
U/97/99	Uganda	This study	EA-4
K63/00	Trans Nzoia, Kenya	This study	EA-2
K109/00	Uasin Gishu, Kenya	This study	EA-2
K117/00	Nyeri, Kenya	This study	EA-2
K130/00	Trans Nzoia, Kenya	This study	EA-2
K131/00	Nairobi, Kenya	This study	EA-2
K141/00	West Pokot, Kenya	This study	EA-2
K145/00	Laikipia, Kenya	This study	EA-2
K147/00	Trans Nzoia, Kenya	This study	EA-2
K150/00	Uasin Gishu, Kenya	This study	EA-2
K45/01	Nakuru, Kenya	This study	EA-2
K61/01	Mombasa, Kenya	This study	EA-2
UKG/35/01	United Kingdom	AJ294910	ME-SA
KEN/5/02	Nakuru, Kenya	DQ165073	EA-2
K79/02	Nakuru, Kenya	This study	EA-2
UGA/Kumi/02	Kumi, Uganda	FJ461344	EA-2
UGA/Kap/02	Kapchorwa, Uganda	FJ461345	EA-2

UGA/3/02	Uganda	DQ165077	EA-2
K55/03	Nakuru, Kenya	This study	EA-2
ETH/3/04	Ethiopia	FJ798109	EA-3
U/13B/04	Hoima, Uganda	This study	EA2
U/14B/04	Hoima, Uganda	This study	EA2
U/17B/04	Hoima, Uganda	This study	EA2
U/20B/04	Hoima, Uganda	This study	EA2
TAN/2/04	Tanzania	WRL	EA-2
K5/05	Laikipia, Kenya	This study	EA2
K31/05	Laikipia, Kenya	This study	EA2
K48/05	Kiambu, Kenya	This study	EA2
ETH/58/05	Ethiopia	FJ798141	EA-4
U/12/05	Wakiso, Uganda	This study	EA-2
K50/06	Uasin Gishu, Kenya	This study	EA-2
ETH/2/06	Ethiopia	FJ798127	EA-3
U/18/06	Mpigi, Uganda	This study	EA-2
U/25/06	Mpigi, Uganda	This study	EA-2
UGA/KSE/06	Kasese, Uganda	EF611987	EA-2
K2/07	Kiambu, Kenya	This study	EA2
K6/07	Koibatek, Kenya	This study	EA2
K28/07	Laikipia, Kenya	This study	EA2
K30/07	Laikipia, Kenya	This study	EA2
K31/07	Kiambu, Kenya	This study	EA-1

K82/07	Muranga, Kenya	This sudy	EA-2
ETH/1/07	Ethiopia	FJ798137	EA-3
K1/08	Nairobi, Kenya	This study	EA-2
K4/08	Thika, Kenya	This study	EA-2
K11/08	Kiambu, Kenya	This study	EA-2
K14/08	Baringo, Kenya	This study	EA-2
K31/08	Kajiado, Kenya	This study	EA-2
K32/08	Thika, Kenya	This study	EA-2

Table 3-2 Primers used in the study

Primer ID	Primers
8-A PN 64	GGTTGGACTCCACATCTCC
8-A PN 84	TACTACACCCAGTACAGCG
8-A PN 85	TGCAGCTTCGGGTGCTCC
8-A PN 98	GCATCCACTTACTACTTTGC
O – 1C $_{564*}$	AATTACACATGGCAAGGCCGACGG
FMD-2A $_{34}$(NK72)*	GAAGGGCCCAGGGTTGGACTC
SNB-85	TGC(AG)(GC)CTTC(AG)GGTGC(CT)CCGTTCGGG
SNB-84	GCCCAATACTACACCCAGTACAGCGGC

*(Knowles and Samuel, 1995)

Figure legends

Figure 3-1 Kenya – Uganda Map

The districts of origin of the FMD virus samples used in the study from both Kenya and Uganda are shown.

Figure 3-2 Phylogenetic relationships between East African strains of serotype O FMD Viruses .

Phylogenetic relationships between East African strains of serotype O FMD Viruses. The tree is based on the nucleotide sequence of the VP1 (1D) coding region with selected reference strains from East Africa and elsewhere. Other sequences deposited in Genbank from East African strains were also included as indicated. Trees were estimated using Bayesian inference analysis

(BEAST) with a relaxed uncorrelated exponential molecular clock (see Section 2). The scale bar shows the estimated number of amino acid substitutions per site per year.

Figure 3-3 Amino acid alignments of 54 serotype O FMD viruses representative of topotypes EA-1 and EA-2.
A '.' indicates an amino acid site identical to that of the sequence KEN/77/78 and '?' denotes an ambiguous (undetermined) site; UGA/3/02 with an "N" at position 631 probably translates into either lysine (K), glutamic acid (E) or glutamine (Q).

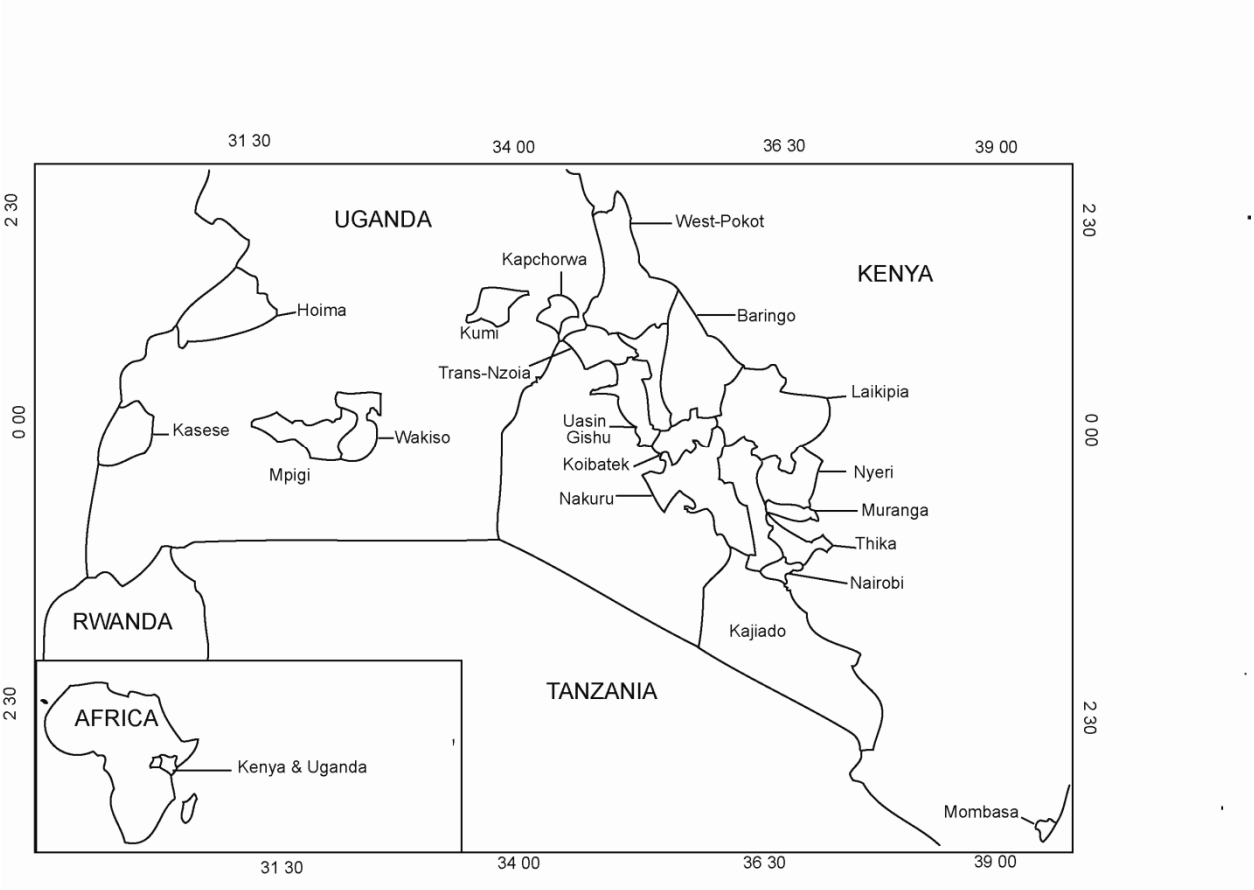


Figure 3-1

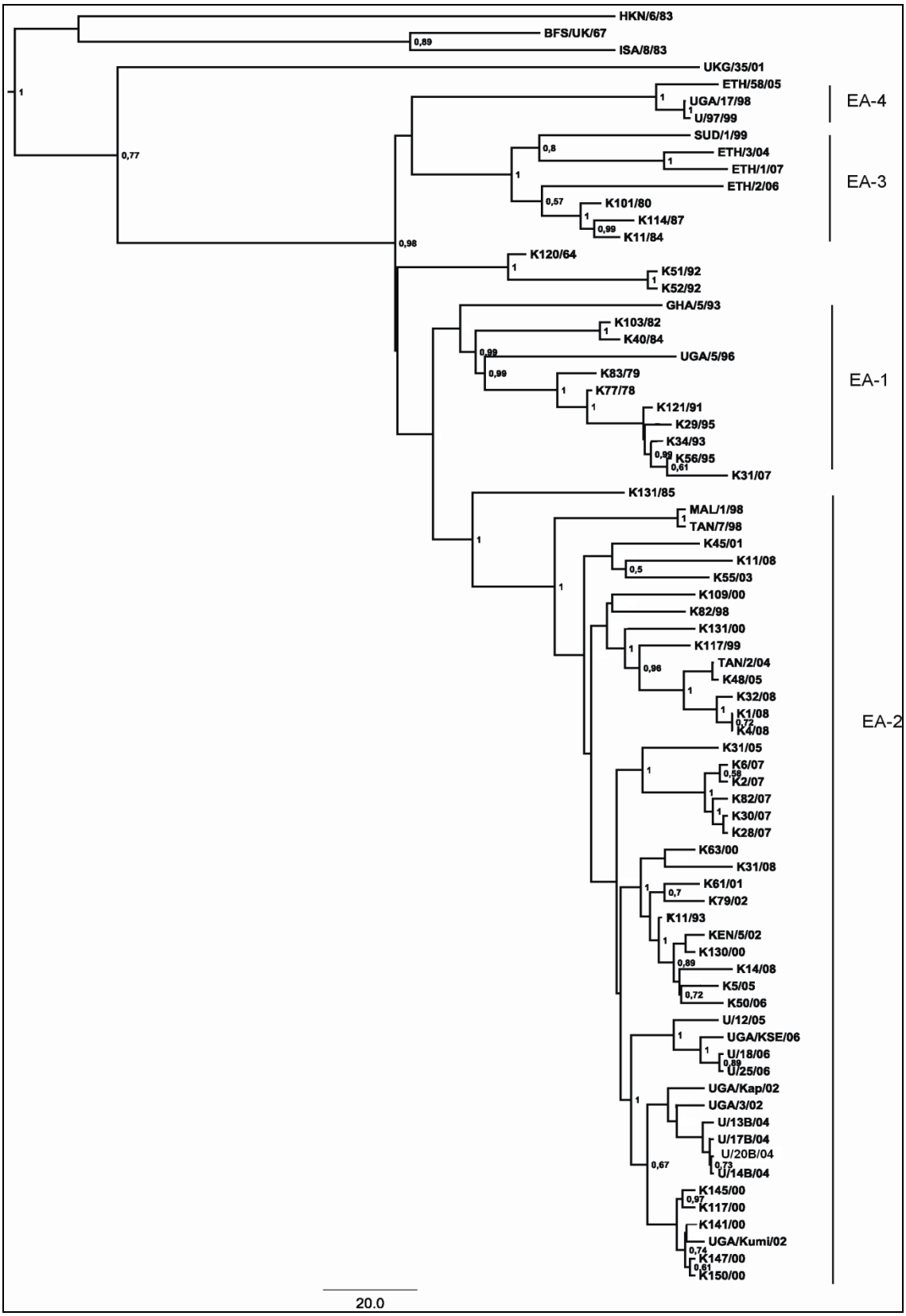

HKN/6/83
BFS/UK/67
ISA/8/83
UKG/35/01
ETH/58/05
UGA/17/98
U/97/99
SUD/1/99
ETH/3/04
ETH/1/07
ETH/2/06
K101/80
K114/87
K11/84
K120/64
K51/92
K52/92
GHA/5/93
K103/82
K40/84
UGA/5/96
K83/79
K77/78
K121/91
K29/95
K34/93
K56/95
K31/07
K131/85
MAL/1/98
TAN/7/98
K45/01
K11/08
K55/03
K109/00
K82/98
K131/00
K117/99
TAN/2/04
K48/05
K32/08
K1/08
K4/08
K31/05
K6/07
K2/07
K82/07
K30/07
K28/07
K63/00
K31/08
K61/01
K79/02
K11/93
KEN/5/02
K130/00
K14/08
K5/05
K50/06
U/12/05
UGA/KSE/06
U/18/06
U/25/06
UGA/Kap/02
UGA/3/02
U/13B/04
U/17B/04
U/20B/04
U/14B/04
K145/00
K117/00
K141/00
UGA/Kumi/02
K147/00
K150/00
EA-4
EA-3
EA-1
EA-2
20.0

Figure 3-2

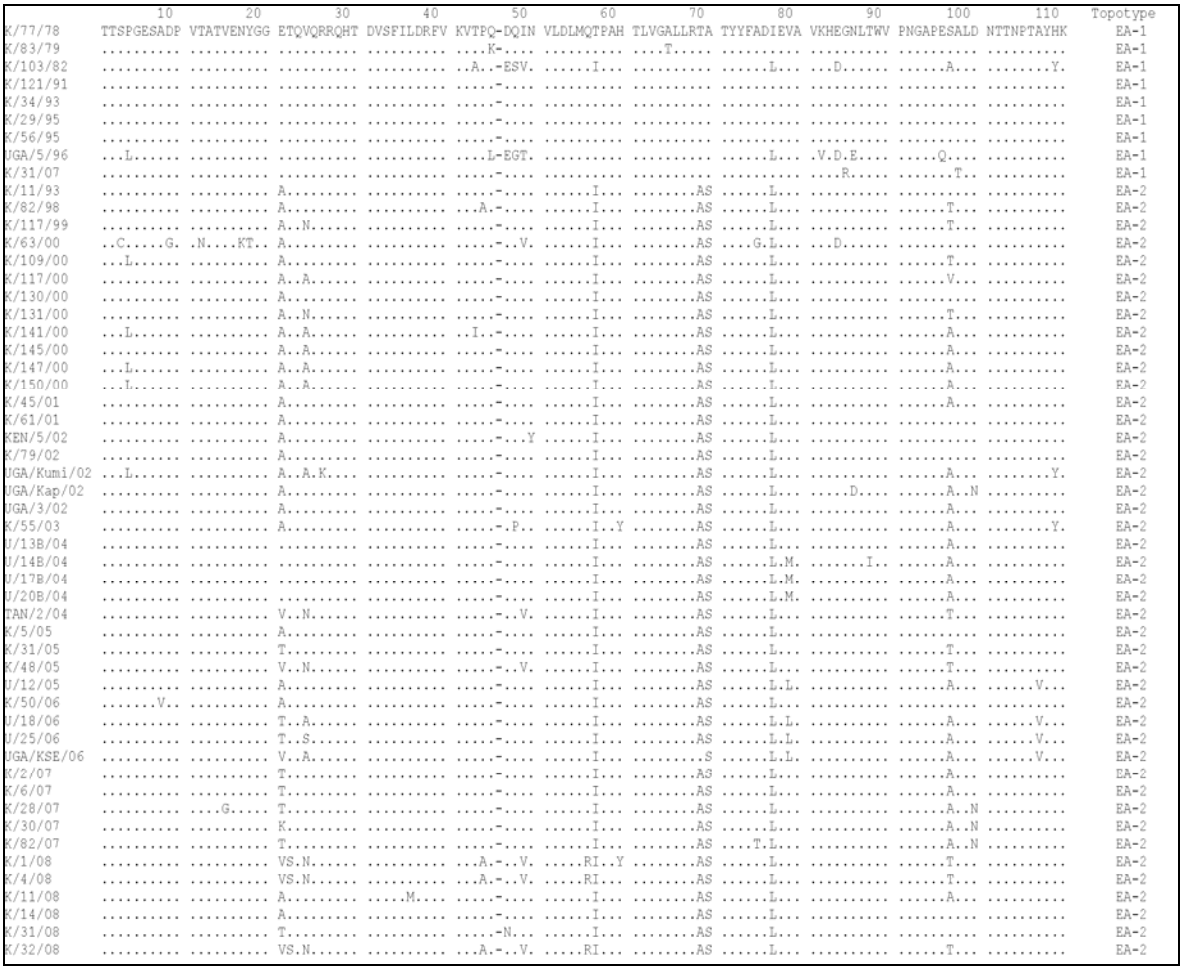
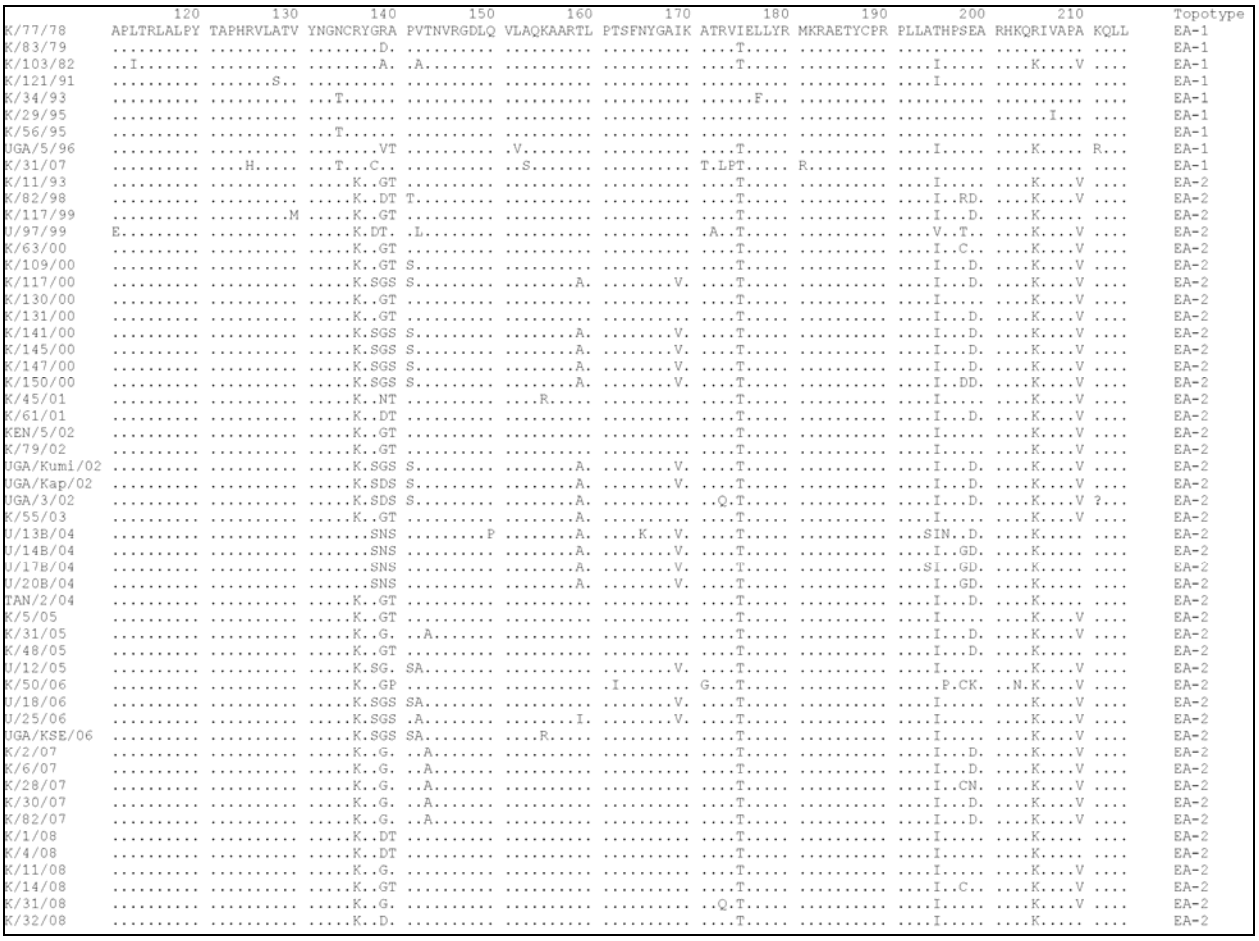

10 20 30 40 50 60 70 80 90 100 110 Topotype
K/77/78 TTSPGESADP VTATVENYGG ETQVQRRQHT DVSFILDRFV KVTPQ-DQIN VLDLMQTPAH TLVGALLRTA TYYFADIEVA VKHEGNLTWV PNGAPESALD NTTNPTAYHK EA-1
K/83/79 EA-1
K/103/82 EA-1
K/121/91 EA-1
K/34/93 EA-1
K/29/95 EA-1
K/56/95 EA-1
UGA/5/96 EA-1
K/31/07 EA-1
K/11/93 EA-2
K/82/98 EA-2
K/117/99 EA-2
K/63/00 EA-2
K/109/00 EA-2
K/117/00 EA-2
K/130/00 EA-2
K/131/00 EA-2
K/141/00 EA-2
K/145/00 EA-2
K/147/00 EA-2
K/150/00 EA-2
K/45/01 EA-2
K/61/01 EA-2
KEN/5/02 EA-2
K/79/02 EA-2
UGA/Kumi/02 EA-2
UGA/Kap/02 EA-2
UGA/3/02 EA-2
K/55/03 EA-2
U/13B/04 EA-2
U/14B/04 EA-2
U/17B/04 EA-2
U/20B/04 EA-2
TAN/2/04 EA-2
K/5/05 EA-2
K/31/05 EA-2
K/48/05 EA-2
U/12/05 EA-2
K/50/06 EA-2
U/18/06 EA-2
U/25/06 EA-2
UGA/KSE/06 EA-2
K/2/07 EA-2
K/6/07 EA-2
K/28/07 EA-2
K/30/07 EA-2
K/82/07 EA-2
K/1/08 EA-2
K/4/08 EA-2
K/11/08 EA-2
K/14/08 EA-2
K/31/08 EA-2
K/32/08 EA-2

Figure 3-3



	120	130	140	150	160	170	180	190	200	210	Topotype
K/77/78	APLTRLALPY	TAPHRVLATV	YNGNCRYGRA	PVTNVRGDLQ	VLAQKAARTL	PTSFNYGAIK	ATRVIELLYR	MKRAETYCPR	PLLATHPSEA	RHKQRIVAPA	KQLL	EA-1
K/83/79	....	....	........D.	....	....	....	....T.....	....	....	....	....	EA-1
K/103/82	..I.......	....	.......A.	.A........	....	....	....T.....	....	....I.....	....K....V	....	EA-1
K/121/91	....	.......S..	....	....	....	....	....	....	....I.....	....	....	EA-1
K/34/93	....	....	...T......	....	....	....	......F...	....	....	....	....	EA-1
K/29/95	....	....	....	....	....	....	....	....	....	......I...	....	EA-1
K/56/95	....	....	...T......	....	....	....	....	....	....	....	....	EA-1
UGA/5/96	....	....	.......VT	....	.V........	....	....T.....	....	....I.....	....K.....	R...	EA-1
K/31/07	....	....H.....	...T...C..	....	..S.......	....	T.LPT.....	R.........	....	....	....	EA-1
K/11/93	....	....	.....K..GT	....	....	....	....T.....	....	....I.....	....K....V	....	EA-2
K/82/98	....	....	.....K..DT	T.........	....	....	....T.....	....	....I..RD.	....K....V	....	EA-2
K/117/99	....	.........M	.....K..GT	....	....	....	....T.....	....	....I...D.	....K.....	....	EA-2
U/97/99	E.........	....	.....K.DT.	.L........	....	....	.A..T.....	....	....V..T..	....K....V	....	EA-2
K/63/00	....	....	.....K..GT	....	....	....	....T.....	....	....I..C..	....K....V	....	EA-2
K/109/00	....	....	.....K..GT	S.........	....	....	....T.....	....	....I...D.	....K....V	....	EA-2
K/117/00	....	....	.....K.SGS	S.........	........A.	.........V.	....T.....	....	....I...D.	....K....V	....	EA-2
K/130/00	....	....	.....K..GT	....	....	....	....T.....	....	....I.....	....K....V	....	EA-2
K/131/00	....	....	.....K..GT	....	....	....	....T.....	....	....I...D.	....K....V	....	EA-2
K/141/00	....	....	.....K.SGS	S.........	........A.	.........V.	....T.....	....	....I...D.	....K....V	....	EA-2
K/145/00	....	....	.....K.SGS	S.........	........A.	.........V.	....T.....	....	....I...D.	....K....V	....	EA-2
K/147/00	....	....	.....K.SGS	S.........	........A.	.........V.	....T.....	....	....I...D.	....K....V	....	EA-2
K/150/00	....	....	.....K.SGS	S.........	........A.	.........V.	....T.....	....	....I..DD.	....K....V	....	EA-2
K/45/01	....	....	.....K..NT	....	....R.....	....	....T.....	....	....I.....	....K....V	....	EA-2
K/61/01	....	....	.....K..DT	....	....	....	....T.....	....	....I...D.	....K....V	....	EA-2
KEN/5/02	....	....	.....K..GT	....	....	....	....T.....	....	....I.....	....K....V	....	EA-2
K/79/02	....	....	.....K..GT	....	....	....	....T.....	....	....I.....	....K....V	....	EA-2
UGA/Kumi/02	....	....	.....K.SGS	S.........	........A.	.........V.	....T.....	....	....I...D.	....K....V	....	EA-2
UGA/Kap/02	....	....	.....K.SDS	S.........	........A.	.........V.	....T.....	....	....I...D.	....K....V	....	EA-2
UGA/3/02	....	....	.....K.SDS	S.........	........A.	....	..Q.T.....	....	....I...D.	....K....V	?...	EA-2
K/55/03	....	....	.....K..GT	....	........A.	....	....T.....	....	....I.....	....K....V	....	EA-2
U/13B/04	....	....	.......SNS	.........P	........A.	....K...V.	....T.....	....	...SIN..D.	....K.....	....	EA-2
U/14B/04	....	....	.......SNS	....	........A.	.........V.	....T.....	....	....I..GD.	....K.....	....	EA-2
U/17B/04	....	....	.......SNS	....	........A.	.........V.	....T.....	....	...SI..GD.	....K.....	....	EA-2
U/20B/04	....	....	.......SNS	....	........A.	.........V.	....T.....	....	....I..GD.	....K.....	....	EA-2
TAN/2/04	....	....	.....K..GT	....	....	....	....T.....	....	....I...D.	....K.....	....	EA-2
K/5/05	....	....	.....K..GT	....	....	....	....T.....	....	....I.....	....K....V	....	EA-2
K/31/05	....	....	.....K..G.	..A.......	....	....	....T.....	....	....I...D.	....K....V	....	EA-2
K/48/05	....	....	.....K..GT	....	....	....	....T.....	....	....I...D.	....K.....	....	EA-2
U/12/05	....	....	.....K.SG.	SA........	....	........V.	....T.....	....	....I.....	....K....V	....	EA-2
K/50/06	....	....	.....K..GP	....	....	.I........	G...T.....	....	.....P.CK.	..N.K....V	....	EA-2
U/18/06	....	....	.....K.SGS	SA........	....	........V.	....T.....	....	....I.....	....K....V	....	EA-2
U/25/06	....	....	.....K.SGS	.A........	........I.	........V.	....T.....	....	....I.....	....K....V	....	EA-2
UGA/KSE/06	....	....	.....K.SGS	SA........	....R.....	....	....T.....	....	....I.....	....K....V	....	EA-2
K/2/07	....	....	.....K..G.	..A.......	....	....	....T.....	....	....I...D.	....K....V	....	EA-2
K/6/07	....	....	.....K..G.	..A.......	....	....	....T.....	....	....I...D.	....K....V	....	EA-2
K/28/07	....	....	.....K..G.	..A.......	....	....	....T.....	....	....I..CN.	....K....V	....	EA-2
K/30/07	....	....	.....K..G.	..A.......	....	....	....T.....	....	....I...D.	....K....V	....	EA-2
K/82/07	....	....	.....K..G.	..A.......	....	....	....T.....	....	....I...D.	....K....V	....	EA-2
K/1/08	....	....	.....K..DT	....	....	....	....T.....	....	....I.....	....K.....	....	EA-2
K/4/08	....	....	.....K..DT	....	....	....	....T.....	....	....I.....	....K.....	....	EA-2
K/11/08	....	....	.....K..G.	....	....	....	....T.....	....	....I.....	....K....V	....	EA-2
K/14/08	....	....	.....K..GT	....	....	....	....T.....	....	....I..C..	....K....V	....	EA-2
K/31/08	....	....	.....K..G.	....	....	....	..Q.T.....	....	....I.....	....K....V	....	EA-2
K/32/08	....	....	.....K..D.	....	....	....	....T.....	....	....I.....	....K.....	....	EA-2

Figure 3-3 Continued

CHAPTER FOUR

High genetic diversity of foot-and-mouth disease virus serotype A in Kenya: a challenge to vaccination strategies

High genetic diversity of foot-and-mouth disease virus serotype A in Kenya: a challenge to vaccination strategies

A.K. Sangula[a, d*], V.B. Muwanika[a], G.J. Belsham[b], R. Heller[c], S.N. Balinda[a], C. Masembe[a], H.R. Siegismund[c]

[a] Makerere University, Institute of Environment and Natural Resources, Molecular Biology Laboratory, P. O. Box 7298, Kampala, Uganda

[b] National Veterinary Institute, Technical University of Denmark, Lindholm, DK-4771 Kalvehave, Denmark

[c] Department of Biology, Ole Maaløes Vej 5, DK-2200, Copenhagen ØN, Denmark

[d] Foot-and-Mouth Disease Laboratory, Embakasi, P. O. Box 18021, 00500, Nairobi, Kenya

[*] Corresponding author:

Abraham K. Sangula

Makerere University Institute of Environment and Natural Resources,

Molecular Biology Laboratory,

P.O. Box 7298, Kampala, Uganda

Fax: +256 414 530134 Phone + 256 712629425

E-mail: aksangula@muienr.mak.ac.ug / aksangula@gmail.com

Abstract

Serotype A is considered to be the most antigenically and genetically diverse of the foot-and-mouth disease virus (FMDV) serotypes. Records of its occurrence in Kenya date back to 1932 and the antigenic diversity of the outbreak viruses has been reflected in the use of two current vaccine strains (K5/80 and K35/80) as well as two other strains used in the past (K18/66 and K179/71). To enhance our understanding of the patterns of genetic variation and distribution of FMD serotype A virus in Kenya, we analyzed complete VP1 region coding sequences of 51 field isolates collected between 1966 and 2008. Thirty two of these sequences were generated in this study while 19 were from published reports. We used genealogy-based coalescent methods to infer times of divergence of the current circulating strains as well as the evolutionary rates. Two currently circulating virus lineages with a countrywide distribution and an extinct lineage were identified. The trans-boundary spread of this serotype across the eastern Africa region was inferred. The implications on the vaccination strategy are discussed.

Key words; FMDV serotype A, Kenya, genetic diversity, vaccination

4.1 Introduction

Foot-and-mouth disease (FMD) is a highly contagious viral disease with well known economic consequences affecting cloven-hoofed animals both domestic and wild. FMD virus (FMDV), the causative agent, is an *Aphthovirus* within the family *Picornaviridae* whose genome is a single-stranded positive-sense RNA of about 8.5 kb. The genome encodes a polyprotein that is processed to the four structural proteins (VP1-VP4) of the virus capsid (only VP1-VP3 are surface exposed) plus several non-structural proteins required for virus replication and protein processing (Belsham, 1993). FMDV occurs in seven immunologically distinct serotypes namely O, A, C, Southern African Territories (SAT) 1-3 and Asia 1. The seven serotypes have different global distributions with types O and A occurring in many parts of Africa, Asia and South America while the SATs are restricted to sub-Saharan Africa. Asia1 is restricted to south Asia while type C has occurred in Asia, South America and Africa but is currently infrequent (or perhaps extinct). Records of type A outbreaks in Kenya date back to 1932 when A and O types were first characterized. However, proper documentation of outbreaks started in 1952 and indicates that before the mid 1970s when SAT2 recorded an upsurge, serotype A was second only to type O in prevalence (Department of Veterinary Services, Kenya reports). Serotype A is considered to be the most antigenically and genetically diverse of the FMDV serotypes with at least 10 major genotypes (I to X) (distinguished by >15% nucleotide divergence of the VP1 coding region sequences) reported of which 8 have been found in Africa (Knowles and Samuel, 2003; Mittal et al., 2005). Three of these reported genotypes (I, III and VIII) are or have been present in Kenya.

For effective control strategies, genetic and antigenic studies of circulating FMDVs are required to be able to select appropriate vaccine strains (Bastos et al., 2003; Kitching, 2005). The

VP1 coding region sequence has been widely used to yield valuable information on the epidemiological dynamics of the disease such as tracing the origins of outbreaks (Beck and Strohmaier, 1987; Bastos et al., 2003; Knowles and Samuel, 2003). The VP1 protein contains important antigenic sites and the RGD motif required for cell receptor recognition (Sobrino et al., 2001; Jackson et al., 2003). Genetic characterization of FMDV strains, largely based on the VP1 coding sequence, has, in the past, showed that type A viruses circulating in the horn of Africa have been able to spread as far as Egypt in North Africa and Kenya in East Africa through animal trade (Knowles et al., 2007).

This study was undertaken to enhance our understanding of the patterns of genetic variation and distribution of FMD type A virus in Kenya by analyzing the partial nucleotide sequences of field isolates stored at Embakasi using genealogy-based coalescent methods to infer times of divergence of the current circulating strains as well as the evolutionary rates. Evolutionary forces known to affect phylogenetic inferences such as selection and recombination (Schierup and Hein, 2000; Moya et al., 2004; Lewis-Rogers et al., 2008), which could bias our estimates, were also tested for in the data set.

4.2 Materials and Methods

4.2.1 Virus isolates

Thirty two serotype A virus isolates (collected between 1966 and 2008) were obtained from the Embakasi FMD laboratory, Nairobi for this study. This laboratory is a repository of all FMD sample materials collected in Kenya. The details of the isolates are shown in Table 1 and the geographic origins of the Kenyan isolates are shown in Figure 1. Nineteen other published serotype A FMDV complete VP1 coding sequences selected from representative genotypes and from the eastern African region were also included in the study (see Table 1).

4.2.2 RNA extraction, reverse transcriptase – PCR, cycle sequencing and sequence assembly

Total RNA extraction was performed with QIAamp® Viral RNA Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions and cDNA synthesized using Ready-to-go-You-prime first strand beads (GE Healthcare Life Sciences, Sweden) with random hexamer primers (pdN6). DNA amplification was achieved using the primer sets : A-1C_{562} or A-1C_{612} (forward) and FMD-2B_{58} (reverse) (Knowles and Samuel, 1995) in a 50µl reaction volume consisting of 5µl of template cDNA and 45µl of PCR reaction mixture (Sangula et al., 2010). The PCR products were examined by electrophoresis on a 2% agarose gel using ethidium bromide staining and a molecular weight marker ΦX174-RF DNA (Amersham, Biosciences). The amplicons were purified using the QIAquick PCR purification kit (QIAGEN) and both strands cycle-sequenced by Macrogen, South Korea utilizing the same forward primers as for the PCRs and the reverse primer FMD-2A_{34} (Knowles and Samuel, 1995) .

The generated sequences were initially assembled using the software program Sequencer version 4.8 (Gene Codes Corporation, USA). Multiple alignment of the complete VP1 coding region sequences for the whole data set were obtained using MUSCLE (Edgar, 2004) incorporated in Geneious version 4.6 (Drummond et al., 2009).

4.2.3 Sequence characteristics and phylogenetic relationships

The best fitting model of nucleotide substitution was determined by means of a hierarchical likelihood ratio test using PAUP* version 4b10 (Swofford, 2003) and MrModeltest version 2.2 (Nylander, 2004). The GTR (Rodriguez et al., 1990) with gamma-distributed rates among sites and a proportion of invariable sites was the preferred model.

Estimates of overall, as well the within group, nucleotide diversity and the amino acid sequence alignment of the VP1 coding sequences was obtained using MEGA version 4 (Tamura et al., 2007).

As a preliminary analysis, tests for evidence of recombination in the data set were done using the GARD method (Pond et al., 2006) available on the datamonkey server (Pond and Frost, 2005). Furthermore, selection forces were tested for using the single-likelihood ancestor counting (SLAC), fixed effects likelihood (FEL) and random effects likelihood (REL methods) available on the datamonkey server.

The Phylogenetic relationships, evolutionary rates and population size changes were co-estimated by applying a Bayesian Markov Chain Monte Carlo (MCMC) method implemented in the BEAST package version 1.4.8 (Drummond and Rambaut, 2007) (http://beast.bio.ed.ac.uk). The Bayesian skyline plot (BSP) model with an uncorrelated exponential clock was used following a selection process involving the preliminary testing of different demographic models and molecular clocks using $\log_{10}$ BF comparisons (Suchard et al., 2001) available in BEAST. The MCMC chains were made of 40 million runs allowing ESSs above 250 with a 10% burn-in as viewed in Tracer software version 1.4 (http://tree.bio.ed.ac.uk/software/tracer/) to be achieved. Uncertainties of the estimates were summarized as highest posterior density (HPD) intervals. Overall evolutionary rates were measured as the number of nucleotide substitutions per site per year (s/s/y). A maximum clade credibility tree was obtained using the Tree Annotator program in BEAST and visualized with FigTree version 1.1.2 (http://tree.bio.ed.ac.uk/software/tracer/).

4.3 Results

4.3.1 Sequence characteristics

The complete VP1 coding sequences of 51 FMDV serotype A isolates comprising of 32 sequences generated in this study and 19 obtained from published reports were analyzed. Of the 51 sequences, 36 were of Kenyan origin. Out of 639 nucleotide sites in the sequence alignment, a total of 363 (57%) were invariant which translated to 142 amino acid residues being completely conserved (Figure 2). Variation was observed in the known hypervariable region of the VP1 (G-H loop and the C-terminus) although the RGD (144-146) motif was conserved, save for Lum/KEN/64 which had a glutamic acid (E146) substitution in place of aspartic acid (D). The overall nucleotide diversity for the entire data set was 16% and the within group diversity for the Kenyan sequences was 14%.

We also estimated the selection forces acting on the FMDV serotype A using the whole data set. The majority of codons in the VP1 coding region of FMDV type A were under purifying (negative) selection. Of the 213 codons analyzed, 165 were under strong negative selection using three methods (SLAC at $p = 0.1$, FEL at $p = 0.1$ and REL with $BF > 50$) while no site was under positive selection at these confidence levels, although 2 sites (codons 139 and 171) appeared to be under positive selection under the REL method. Increasing the p value to 0.2 for the SLAC and FEL methods also identified site 171 to be under positive selection.

No evidence of recombination in the data set was established as determined by the GARD method.

The overall substitution rate was 3.57×10^{-3} substitutions/site/year (95% HPD: 2.31×10^{-3} - 4.87×10^{-3}).

4.3.2 Phylogenetic relationships

The inferred maximum clade credibility tree is shown in Fig. 3 with the posterior probabilities for significant branches (≥ 0.5) indicated. Two distinct clades (lineages) of Kenyan type A viruses were observed with one lineage (labeled G-I) consisting of viruses exclusively of Kenyan origin while the lineage labeled G-III includes viruses from both the eastern Africa region and the rest of Africa. No local structuring was observed, as the Kenyan viruses in the two lineages both have a countrywide distribution. A third lineage (G-VIII) with no currently circulating viruses among the dataset is represented on the phylogenetic tree by a single virus (Ktl/KEN/64). These lineages correspond to the 3 genotypes reported to be present in Kenya on the World Reference Laboratory for FMD website (http://www.wrlfmd.org/fmd_genotyping/africa/ken.htm).

The estimated divergence time between the Kenyan clades was about 75 years ago with the most recent common ancestor (TMRCA) estimated to have emerged 110 years before present (ybp) (95% HPD: 60-180). The two vaccine strains (K5/80 and K35/80) belong to the different clades i.e. G-I and G-III respectively.

4.4 Discussion

The phylogenetic analysis of serotype A FMDV (Figure 3) identified the presence of two lineages of circulating viruses within Kenya. Lineage G-I viruses seem exclusive to Kenya suggesting a within country maintenance and evolution as shown by representative isolates from the 1960s (e.g. K18/66) to the present time (e.g. K73/08). However, the limited number of sequences from the other countries in the data set may not exclude the possibility of the existence of this lineage outside Kenya. This lineage has a countrywide distribution as shown by representative isolates from districts in different provinces of the country e.g. K5/80, K288/83,

K49/84, and K44/05 (Rift Valley province); K67/85 and K15/92 (Western province); K87/91 and K11/92 (Eastern province); K64/95 (North Eastern province); and K48/81 and K293/83 (Coast province). This countrywide distribution is perhaps a reflection of the extensive livestock movements across the country for trade and grazing. Strain K5/80, originally isolated from Kajiado district in the Rift valley province and belonging to this lineage, has been used as a vaccine to control outbreaks mainly in the Rift Valley and western parts of the country. Lineage G-III consists of viruses also distributed across the country e.g. K35/80 (Eastern); K60/01 (Nairobi); K131/79 (Central); K4/81 (Western); K83/85 (Coast); and K50/81 (Rift Valley) together with viruses from the wider East African region e.g. ETH/3/05 (Ethiopia); SUD/3/77 (Sudan) and UGA/13/66 (Uganda). This lineage could be indicative of type A introductions into the country from countries across the region at various times with resultant spread countrywide as a result of livestock movements. Strain K35/80, originally isolated from Embu in Eastern province, has been used as a vaccine to control outbreaks mostly in the Central, Eastern and North Eastern provinces. A third lineage which seems to have become extinct is represented by a single virus isolate (Ktl/KEN/64) in the phylogeny. Outbreaks of type A have been controlled by ring vaccination with mono-valent vaccines using either of the vaccine strains depending on the geographical area. Two vaccine strains (K5/80 and K35/80) are currently in use based on the antigenic characterization of field isolates using serological methods (unpublished reports Embakasi FMD laboratory). Two other type A vaccine strains (K18/66 and K179/71) have been previously used. Routine bi-annual vaccination using multi-valent vaccines incorporating the appropriate type A strain was undertaken in select districts as part of FMD control policy in the 1970s and 1980s. This period also coincides with the highest record of outbreaks which could be attributed to the high awareness and reporting levels. However, this policy has not been

effectively implemented since the 1990s due to budgetary constraints resulting in vaccination programs being limited to ring vaccination during reported outbreaks. It is apparent from the distribution of the type A lineages in Kenya reported in this study that the strategy of using particular vaccine strains for specific regions of the country may not be rational since the viruses spread across the country.

The transboundary spread of the type A virus across the region, as represented by lineage G-III here, reaffirms the need for regional efforts towards FMD control. Type A has been reported in all countries of the eastern Africa region with reports indicating first records in Uganda (1953), Tanzania (1980), Ethiopia (1969) and Somalia (1978) (Vosloo et al., 2002). The results of this study confirm reports that livestock movement across the countries in the region may be responsible for FMD spread to countries as far distant as Egypt in North Africa (Knowles et al., 2007). Livestock movement seems to be the major factor in the epidemiology of type A in the eastern Africa region since wildlife has not been suggested to play a significant role in contrast to the SAT serotype viruses (Anderson et al., 1979; Bronsvoort et al., 2008). In India, where type A is also endemic, the serotype co-exists with types O and Asia 1 but is the most divergent and is responsible for most of the recent FMD outbreaks (Mittal et al., 2005) in contrast to Kenya were type O and SAT 2 have caused most of the recent outbreaks (reports of the Department of Veterinary Services, Kenya). The substitution rates observed here were similar to those reported for serotype A previously (Tully and Fares, 2008) and are higher than those observed for the SAT serotypes in eastern Africa (unpublished results by the authors). The high evolutionary rates are consistent with the existence of many genetic and antigenic variants reported for the serotype. In Kenya, the genetic heterogeneity of type A is consistent with the

limited success of using several different vaccine strains as compared to, for example, serotypes C and SAT 1 where one vaccine strain has been in use over extended periods of time.

4.5 Conclusion

The high genetic diversity of foot-and-mouth disease virus serotype A in Kenya reported in this study highlights the limited success of vaccination programs. The importance of livestock mobility in the distribution of FMD viruses in eastern Africa is apparent from the circulation pattern of FMDV serotype A observed. Continuous characterization of field isolates on a regional scale is important for discerning virus circulation patterns emphasizing the need for regional approaches to FMD control.

Acknowledgements

We thank; the Director of Veterinary Services, Kenya, for providing the virus isolates from the Embakasi laboratory used in the study, Dr. Sabenzia Wekesa for providing information on the Kenyan isolates, Teresa Kenduiywo, William Birgen and Eugene Arinaitwe for technical assistance. Part of this work was carried out by using the resources of the Computational Biology Service Unit from Cornell University which is partially funded by Microsoft Corporation. This work was supported by the Danish International Development Agency (DANIDA) under the Livestock-Wildlife Diseases in East Africa Project.

References

Anderson, E.C., Doughty, W.J., Anderson, J., Paling, R., 1979. The pathogenesis of foot-and-mouth disease in the African buffalo (*Syncerus Caffer*) and the role of this species in the epidemiology of the disease in Kenya. *Journal of Comparative Pathology* **89**, 511 - 519.

Bastos, A.D.S., Haydon, D.T., Sangare´, O., Boshoff, C.I., Edrich, J.L., Thomson, G.R., 2003. The implications of virus diversity within the SAT 2 serotype for control of foot-and-mouth disease in sub-Saharan Africa. *Journal of General Virology* **84**, 1595–1606.

Beck, E., Strohmaier, K., 1987. Subtyping of European foot-and-mouth disease virus strains by nucleotide sequence determination. *Journal of Virology* **61**, 1621-1629.

Belsham, G.J., 1993. Distinctive features of foot-and-mouth disease virus, a member of the picorna virus family; aspects of virus protein synthesis, protein processing and structure. *Progress in Biophysics and Molecular Biology* **60**, 241-260.

Bronsvoort, B.M.D.C., Parida, S., Handel, I., McFarland, S., Fleming, L., Hamblin, P., Kock, R., 2008. Serological survey for foot-and-mouth disease virus in wildlife in eastern Africa and estimation of test parameters of a nonstructural protein enzyme-linked immunosorbent assay for buffalo. *Clinical and Vaccine Immunology* **15**, 1003–1011.

Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., Wilson, A., 2009. *Geneious v4.6.* Available from http://www.geneious.com/.

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

Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**.

Jackson, T., King, A.M.Q., Stuart, D.I., Fry, E., 2003. Structure and receptor binding. *Virus Research* **91**, 33-46.

Kitching, R.P., 2005. Global epidemiology and prospects for control of foot-andmouth disease. *Current Topics in Microbiology and Immunology*. pp. 133–148.

Knowles, N.J., Samuel, A.R., 1995. Polymerase chain reaction amplification and cycle sequencing of the 1D gene of foot-and-mouth disease viruses. *Session of the research group of the standing technical committee of the European commission for the control of foot-and-mouth disease*. FAO, Rome, 19-22 September 1994, Vienna, Austria.

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

Knowles, N.J., Wadsworth, J., Reid, S.M., Swabey, K.G., El-Kholy, A.A., El-Rahman, A.O.A., Soliman, H.M., Ebert, K., Ferris, N.P., Hutchings, G.H., Statham, R.J., King, D.P., Paton, D.J., 2007. Foot-and-mouth disease virus serotype A in Egypt. *Emerging Infectious Diseases* **13**.

Lewis-Rogers, N., McClellan, D.A., Crandall, K.A., 2008. The evolution of foot-and-mouth disease virus: impacts of recombination and selection. *Infection, Genetics and Evolution* **8**, 786–798.

Mittal, M., Tosh, C., Hemadri, D., Sanyal, A., Bandyopadhyay, S.K., 2005. Phylogeny, genome evolution, and antigenic variability among endemic foot-and-mouth disease virus type A isolates from India. *Archives of Virology* **150**, 911–928.

Moya, A., C.Holmes, E., González-Candelas, F., 2004. The population genetics and evolutionary epidemiology of RNA viruses. *Nature Reviews | Microbiology* **2**, 279-288.

Nylander, J.A.A., 2004. *MrModeltest v2*. Program distributed by the author Evolutionary Biology Centre, Uppsala University.

Pond, S.L.K., Frost, S.D.W., 2005. A genetic algorithm approach to detecting lineage-specific variation in selection pressure. *Molecular Biology and Evolution* **22**, 478–485.

Pond, S.L.K., Posada, D., Gravenor, M.B., Woelk, C.H., Frost, S.D.W., 2006. Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology and Evolution* **23**, 1891–1901.

Rodriguez, F., Oliver, J.L., Marfn, A., Medina, J.R., 1990. The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology* **142**, 485-501.

Sangula, A.K., Siegismund, H.R., Belsham, G.J., Balinda, S.N., Masembe, C., Muwanika, V.B., 2010. Low diversity of foot-and-mouth disease serotype C virus in Kenya: evidence for probable vaccine strain re-introductions in the field. *Epidemiology and Infection* (In press).

Schierup, M.H., Hein, J., 2000. Consequences of recombination on traditional phylogenetic analysis. *Genetics* **156**, 879–891.

Sobrino, F., Saiz, M., Jimenez-Clavero, M.A., Nunez, J.I., Rosas, M.F., Baranowski, E., Ley, V., 2001. Foot-and-mouth disease virus: a long known virus, but a current threat. *Veterinary Research* **32**, 1–30.

Suchard, M.A., Weiss, R.E., Sinsheimer, J.S., 2001. Bayesian selection of continuous-time markov chain evolutionary models. *Molecular Biology and Evolution* **18**, 1001–1013.

Swofford, D.L., 2003. *PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version4*. Sinauer Associates, Sunderland, Massachusetts.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596–1599.

Tully, D.C., Fares, M.A., 2008. The tale of a modern animal plague: tracing the evolutionary history and determining the time-scale for foot and mouth disease virus. *Virology* **382**, 250–256.

Vosloo, W., Bastos, A.D.S., Sangare, O., Hargreaves, S.K., Thomson, G.R., 2002. Review of the status and control of foot and mouth disease in sub-Saharan Africa. *OIE Scientific and Technical Review* **21**, 437-447.

Table 4-1 List of the serotype A viruses included in this study of which 32 sequences were generated in this work while 19 were sourced from published reports

Lab Ref No.	Year of Isolation	District/ Country	Accession No.
Cruz/BRA/55	1955	Brazil	AY593768
Bagk/TAI/60	1960	Thailand	AY593755
Lum/KEN/64	1964	Kericho, Kenya	AY593761
Ktl/KEN/64	1964	Trans Nzoia, Kenya	AY593766
K18/66*	1966	Kericho, Kenya	This study
UGA/13/66	1966	Uganda	WRL
K179/71*	1971	Kajiado	This study
EGY/1/72	1972	Egypt	EF208756
NGR/2/73	1973	Nigeria	WRL
MOR/7/77	1977	Morocco	EU553870
SUD/3/77	1977	Sudan	WRL
K28/78	1978	Meru Central, Kenya	This study
K61/78	1978	Kajiado, Kenya	This study
K131/79	1979	Kiambu, Kenya	This study
K151/79	1979	Thika, Kenya	This study
K158/80	1980	Laikipia, Kenya	This study
K35/80**	1980	Embu, Kenya	This study
K16/80	1980	Kajiado, Kenya	This study
K5/80**	1980	Kajiado, Kenya	This study
K50/81	1981	Trans Nzoia, Kenya	This study

K4/81	1981	Bungoma, Kenya	This study
K7/81	1981	Teso, Kenya	This study
K48/81	1981	Kilifi, Kenya	This study
K51/81	1981	Kwale, Kenya	This study
Alem/ARG/81	1981	Argentina	AJ306219
K288/83	1983	Kericho, Kenya	This study
K293/83	1983	Taita Taveta, Kenya	This study
K36/84	1984	Meru Central, Kenya	This study
K49/84	1984	Narok, Kenya	This study
K67/85	1985	Isiolo, Kenya	This study
K83/85	1985	Mombasa, Kenya	This study
K87/91	1991	Meru North, Kenya	This study
K11/92	1992	Meru North, Kenya	This study
K15/92	1992	Kakamega, Kenya	This study
K9/94	1994	Narok, Kenya	This study
ETH/23/94	1994	Ethiopia	EF208767
K64/95	1995	Garissa, Kenya	This study
IRN/1/96	1996	Iran	EF208771
K1/97	1997	Meru North, Kenya	This study
KEN/15/98	1998	Meru, Kenya	EF208774
CAR/15/00	2000	Cameroon	EF208755
K73/01	2001	Meru Central, Kenya	This study
K60/01	2001	Nairobi, Kenya	This study

IND/68/01	2001	India	AF390659
K129/03	2003	Kajiado, Kenya	This study
K44/05	2005	Nakuru, Kenya	This study
KEN/29/05	2005	Embu, Kenya	EF208773
IRN/1/05	2005	Iran	EF208769
ETH/3/05	2005	Ethiopia	EF208762
EGY/1/06	2006	Egypt	EF208757
K73/08	2008	Kajiado, Kenya	This study

* Previous vaccine strain

** Current vaccine strain

WRL = accessed from the World reference laboratory sequence data website (http://www.wrlfmd.org/fmd_genotyping/prototypes.htm).

Figure legends

Figure 4-1 Map of Kenya showing the origins of the serotype A isolates the study. The location of each isolate is marked by a colour code of the lineage marked in the phylogenetic tree in Figure 4-3.

Figure 4- 2 Amino acid sequence alignment of the VP1 for 23 serotype A FMD viruses. A `.´ indicates an amino acid residue identical to that of the sequence K5/80 (vaccine strain). The last 2 amino acid residues are from the 2A

Figure 4-3 Maximum clade credibility tree of serotype A viruses based on complete VP1 coding sequences inferred using BEAST showing lineage divergence since the most recent common ancestor. The posterior clade probabilities are shown for selected tree nodes. Kenyan lineages are marked G-I, G-III and G-VIII. Vaccine strains are circled.

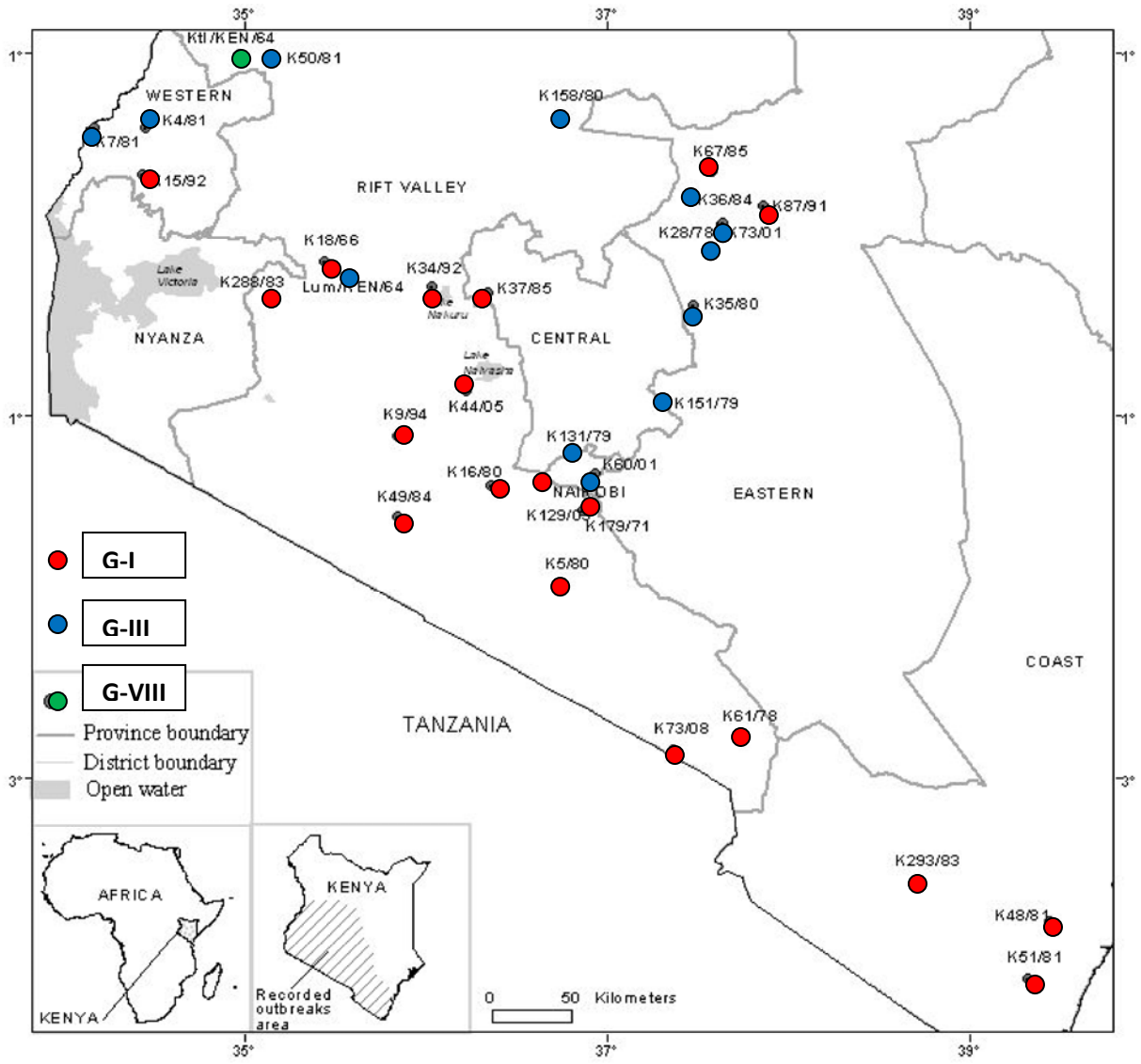
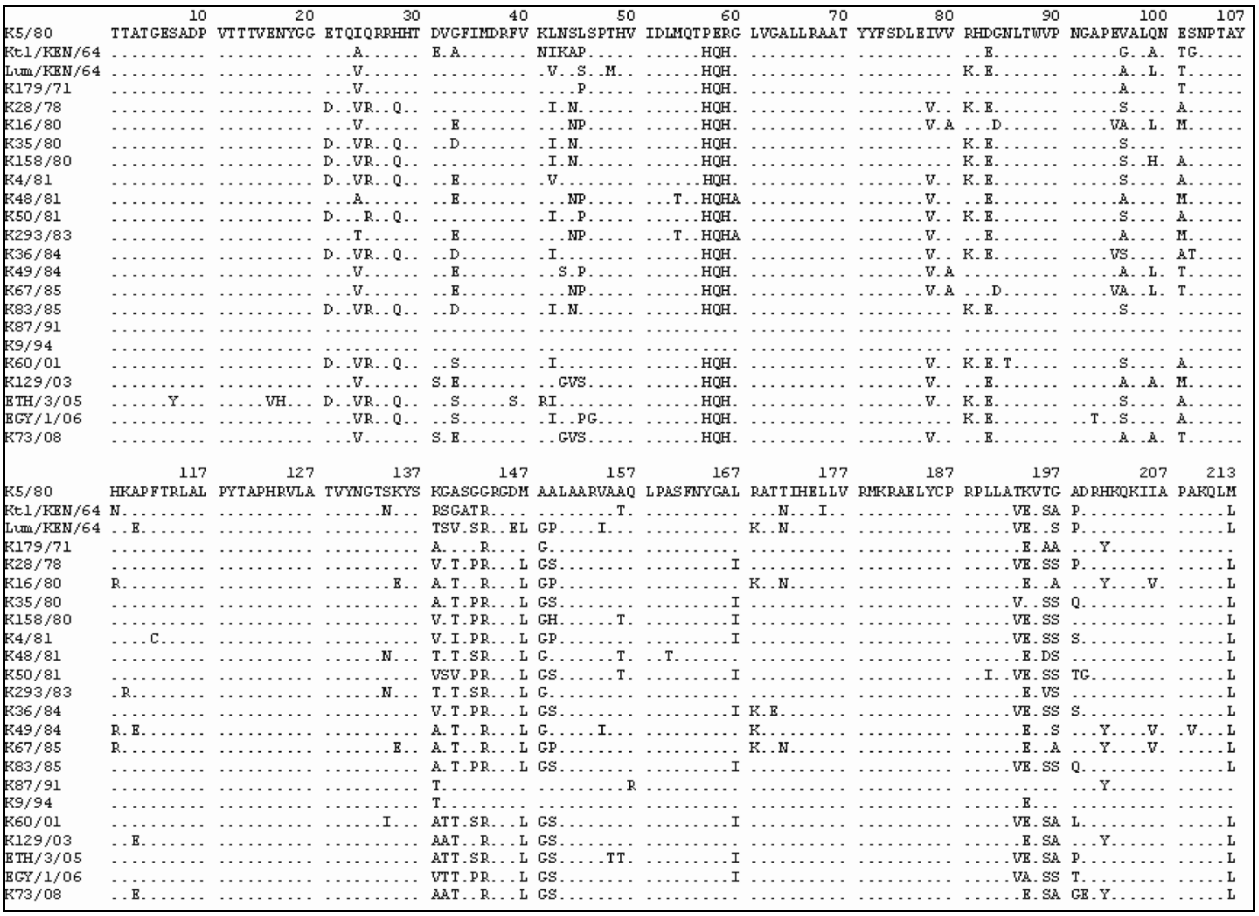

35°
37°
39°
1°
3°
Ktl/KEN/64
K50/81
WESTERN
K4/81
K7/81
K15/92
K158/80
K67/85
RIFT VALLEY
K36/84
K87/91
K28/78
K73/01
K18/66
K34/92
K288/83
Lum/KEN/64
K37/85
K35/80
Lake Victoria
Nakuru
NYANZA
CENTRAL
Lake Naivasha
K44/05
K151/79
K9/94
K131/79
K16/80
K60/01
K49/84
NAIROBI
EASTERN
K129/03
K179/71
K5/80
G-I
G-III
G-VIII
Province boundary
District boundary
Open water
TANZANIA
K73/08
K61/78
COAST
K293/83
K48/81
K51/81
AFRICA
KENYA
Recorded outbreaks area
0
50
Kilometers

Figure 4-1



	10	20	30	40	50	60	70	80	90	100	107
K5/80	TTATGESADP	VTTTVENYGG	ETQIQRRHHT	DVGFIMDRFV	KLNSLSPTHV	IDLMQTPERG	LVGALLRAAT	YYFSDLEIVV	RHDGNLTWVP	NGAPEVALQN	ESNPTAY
Ktl/KEN/64	....	....	...A......	E.A.......	NIKAP.....	......HQH.	....	....	..E.......	.....G..A.	TG.....
Lum/KEN/64	....	....	...V......	....	.V..S..M..	......HQH.	....	....	K.E.......	.....A..L.	T......
K179/71	....	....	...V......	....	....P.....	......HQH.	....	....	....	.....A....	T......
K28/78	....	....	D..VR..Q..	....	.I.N......	......HQH.	....	.......V..	K.E.......	.....S....	A......
K16/80	....	....	...V......	..E.......	...NP.....	......HQH.	....	.......V.A	...D......	....VA..L.	M......
K35/80	....	....	D..VR..Q..	..D.......	.I.N......	......HQH.	....	....	K.E.......	.....S....	....
K158/80	....	....	D..VR..Q..	....	.I.N......	......HQH.	....	....	K.E.......	.....S..H.	A......
K4/81	....	....	D..VR..Q..	..E.......	.V........	......HQH.	....	.......V..	K.E.......	.....S....	A......
K48/81	....	....	...A......	..E.......	...NP.....	...T..HQHA	....	.......V..	..E.......	.....A....	M......
K50/81	....	....	D...R..Q..	....	.I..P.....	......HQH.	....	.......V..	K.E.......	.....S....	A......
K293/83	....	....	...T......	..E.......	...NP.....	...T..HQHA	....	.......V..	..E.......	.....A....	M......
K36/84	....	....	D..VR..Q..	..D.......	.I........	......HQH.	....	.......V..	K.E.......	....VS....	AT.....
K49/84	....	....	...V......	..E.......	..S.P.....	......HQH.	....	.......V.A	....	.....A..L.	T......
K67/85	....	....	...V......	..E.......	...NP.....	......HQH.	....	.......V.A	...D......	....VA..L.	T......
K83/85	....	....	D..VR..Q..	..D.......	.I.N......	......HQH.	....	....	K.E.......	.....S....	....
K87/91	....	....	....	....	....	....	....	....	....	....	....
K9/94	....	....	....	....	....	....	....	....	....	....	....
K60/01	....	....	D..VR..Q..	..S.......	.I........	......HQH.	....	.......V..	K.E.T.....	.....S....	A......
K129/03	....	....	...V......	S.E.......	..GVS.....	......HQH.	....	.......V..	..E.......	.....A..A.	M......
ETH/3/05	......Y...	.....VH...	D..VR..Q..	..S.....S.	RI........	......HQH.	....	.......V..	K.E.......	.....S....	A......
EGY/1/06	....	....	...VR..Q..	..S.......	.I..PG....	......HQH.	....	....	K.E.......	..T..S....	A......
K73/08	....	....	...V......	S.E.......	..GVS.....	......HQH.	....	.......V..	..E.......	.....A..A.	T......

	117	127	137	147	157	167	177	187	197	207	213
K5/80	HKAPFTRLAL	PYTAPHRVLA	TVYNGTSKYS	KGASGGRGDM	AALAARVAAQ	LPASFNYGAL	RATTIHELLV	RMKRAELYCP	RPLLATKVTG	ADRHKQKIIA	PAKQLM
Ktl/KEN/64	N.........	....	......N...	RSGATR....	........T.	....	...N...I..	....	.....VE.SA	P.........	L
Lum/KEN/64	..E.......	....	....	TSV.SR..EL	GP....I...	....	K..N......	....	.....VE..S	P.........	L
K179/71	....	....	....	A....R....	G.........	....	....	....	.....E.AA	...Y......	......
K28/78	....	....	....	V.T.PR...L	GS........	I	....	....	.....VE.SS	P.........	L
K16/80	R.........	....	.......E..	A.T..R...L	GP........	....	K..N......	....	.....E..A	...Y....V.	L
K35/80	....	....	....	A.T.PR...L	GS........	I	....	....	.....V..SS	Q.........	L
K158/80	....	....	....	V.T.PR...L	GH......T.	I	....	....	.....VE.SS	....	L
K4/81	....C.....	....	....	V.I.PR...L	GP........	I	....	....	.....VE.SS	S.........	L
K48/81	....	....	......N...	T.T.SR...L	G.......T.	..T.......	....	....	.....E.DS	....	L
K50/81	....	....	....	VSV.PR...L	GS......T.	I	....	....	..I..VE.SS	TG........	L
K293/83	.R........	....	......N...	T.T.SR...L	G.........	....	....	....	.....E.VS	....	L
K36/84	....	....	....	V.T.PR...L	GS........	I	K.E.......	....	.....VE.SS	S.........	L
K49/84	R.E.......	....	....	A.T..R...L	G.....I...	....	K..........	....	.....E..S	...Y....V.	.V...L
K67/85	R.........	....	.......E..	A.T..R...L	GP........	....	K..N......	....	.....E..A	...Y....V.	L
K83/85	....	....	....	A.T.PR...L	GS........	I	....	....	.....VE.SS	Q.........	L
K87/91	....	....	....	T..........	.........R	....	....	....	....	...Y......	......
K9/94	....	....	....	T..........	....	....	....	....	.....E...	....	......
K60/01	....	....	......I...	ATT.SR...L	GS........	I	....	....	.....VE.SA	L.........	L
K129/03	..E.......	....	....	AAT..R...L	GS........	....	....	....	.....E.SA	...Y......	L
ETH/3/05	....	....	....	ATT.SR...L	GS.....TT.	I	....	....	.....VE.SA	P.........	L
EGY/1/06	....	....	....	VTT.PR...L	GS........	I	....	....	.....VA.SS	T.........	L
K73/08	..E.......	....	....	AAT..R...L	GS........	....	....	....	.....E.SA	GE.Y......	L

Figure 4-2

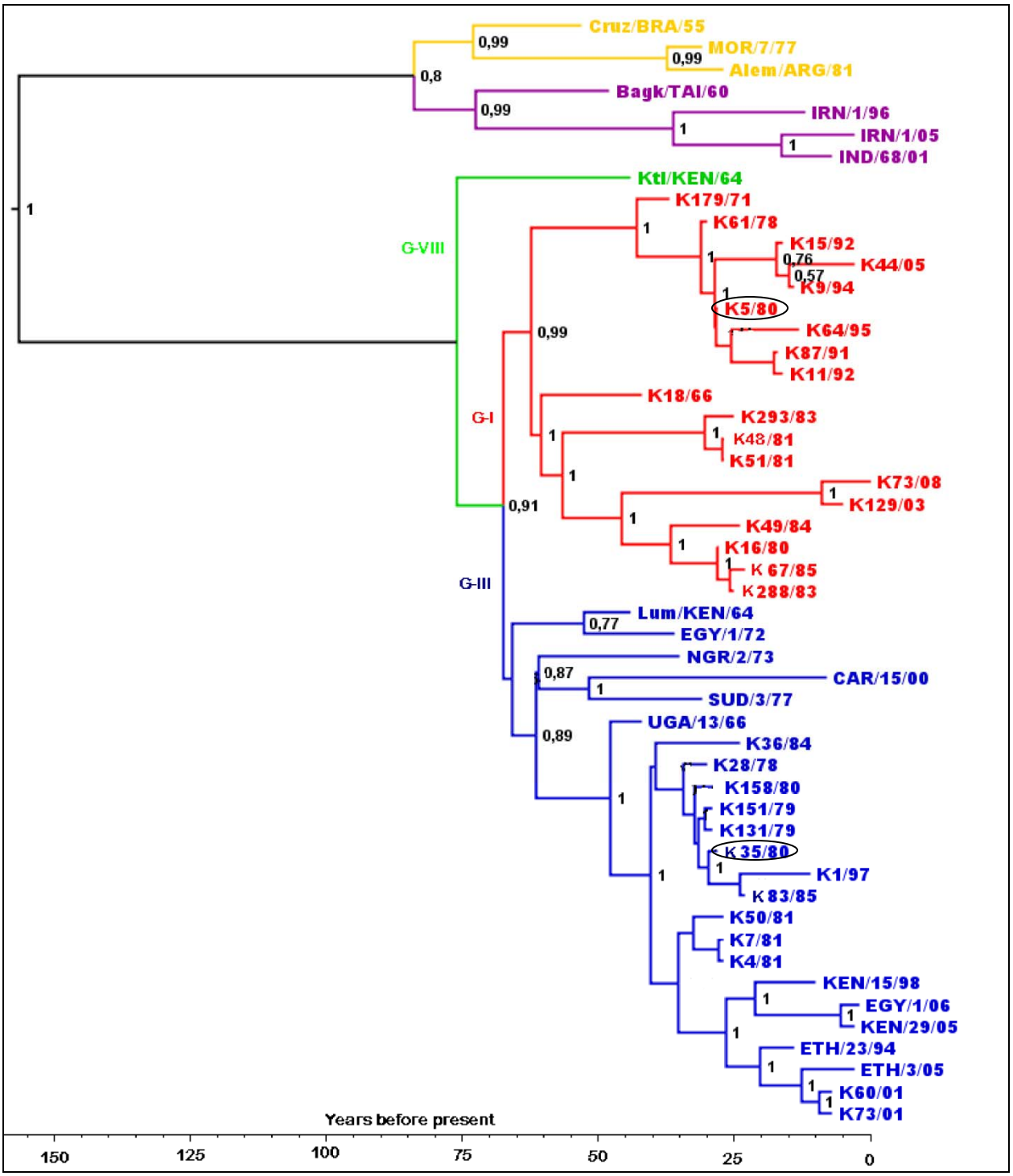

Cruz/BRA/55
MOR/7/77
Alem/ARG/81
Bagk/TAI/60
IRN/1/96
IRN/1/05
IND/68/01
Ktl/KEN/64
K179/71
K61/78
K15/92
K44/05
K9/94
K5/80
K64/95
K87/91
K11/92
K18/66
K293/83
K48/81
K51/81
K73/08
K129/03
K49/84
K16/80
K 67/85
K 288/83
Lum/KEN/64
EGY/1/72
NGR/2/73
CAR/15/00
SUD/3/77
UGA/13/66
K36/84
K28/78
K158/80
K151/79
K131/79
K 35/80
K1/97
K 83/85
K50/81
K7/81
K4/81
KEN/15/98
EGY/1/06
KEN/29/05
ETH/23/94
ETH/3/05
K60/01
K73/01
G-VIII
G-I
G-III
0,99
0,99
0,8
0,99
1
1
1
1
1
1
0,76
0,57
1
0,99
1
1
1
0,91
1
1
1
0,77
0,87
1
0,89
1
1
1
1
1
1
1
1
1
1
Years before present
150
125
100
75
50
25
0

Figure 4-3

CHAPTER FIVE

Low diversity of foot-and-mouth disease serotype C virus in Kenya: evidence for probable vaccine strain re-introductions in the field

Low diversity of foot-and-mouth disease serotype C virus in Kenya: evidence for probable vaccine strain re-introductions in the field

A.K. Sangula[14*], H.R. Siegismund[3], G.J. Belsham[2], S.N. Balinda[1], C. Masembe[1], V.B. Muwanika[1]

[1]Makerere University, Institute of Environment and Natural Resources, Molecular Biology Laboratory, P. O. Box 7298, Kampala, Uganda

[2]National Veterinary Institute, Technical University of Denmark, Lindholm, DK-4771 Kalvehave, Denmark

[3]Department of Biology, Ole Maaløes Vej 5, DK-2200, Copenhagen ØN, Denmark

[4]Foot-and-Mouth Disease Laboratory, Embakasi, P. O. Box 18021, 00500, Nairobi, Kenya

[*] Corresponding author:

Abraham K. Sangula

Makerere University Institute of Environment and Natural Resources,

Molecular Biology Laboratory,

P.O. Box 7298, Kampala, Uganda

Fax: +256 414 530134 Phone + 256 712629425

E-mail: aksangula@muienr.mak.ac.ug / aksangula@gmail.com

Running title: Low diversity of FMDV serotype C in Kenya

Summary

Most viruses are maintained by complex processes of evolution that enable them to survive but also complicate efforts to achieve their control. In this paper, we study patterns of evolution in foot-and-mouth disease serotype C virus isolates from Kenya, one of the few places in the world where serotype C has been endemic and is suspected to remain. The nucleotide sequences encoding the capsid protein VP1 from eight isolates collected between 1967 and 2004 were analysed for patterns of sequence divergence and evolution. Very low nucleotide diversity (π = 0.0025) and remarkably little change (only 5 segregating sites and 3 amino acid changes) were observed amongst these isolates collected over a period of nearly 40 years. We interpret these results as being suggestive of re-introductions of the vaccine strain into the field. The implications of these results for the maintenance of serotype C FMDV and the use of vaccination as a control measure in Kenya are discussed.

Key words; FMDV serotype C, Kenya, sequence divergence, vaccine strain

5.1 Introduction

Foot-and-mouth disease virus (FMDV) is an *Aphthovirus* within the family *Picornaviridae*. It causes a highly contagious disease (FMD) in cloven-hoofed animals (domestic and wild) which is the most significant constraint to international trade in livestock and livestock products today (Grubman and Baxt, 2004). The genome of FMDV is a positive-sense RNA of about 8.5 kb which encodes a polyprotein that is processed to the four structural proteins (VP1-VP4) of the virus capsid (only VP1-VP3 are surface exposed) plus several non-structural proteins required for virus replication and protein processing (Belsham, 1993). FMDV occurs in seven immunologically distinct serotypes namely O, A, C, Southern African Territories (SAT) 1-3 and Asia 1. The seven serotypes have different global distributions with type C belonging to the Euro-Asiatic serotypes which include O, A and Asia1. Type C FMDV was first described by Waldmann and Trautwein (Waldmann and Trautwein, 1926) and has since had a limited distribution. It has been recorded in Europe, South America, East Africa, North Africa, Angola and Southern Asia (Knowles and Samuel, 2003). At present, the last reported cases of type C were in 2004 in Brazil and Kenya (Roeder and Knowles, 2009). In the wider East African region, type C was last reported in Uganda in the 1970s and Ethiopia in 1983 (Knowles and Samuel, 2003; Roeder and Knowles, 2009).

Records of serotype C outbreaks of FMD in Kenya date back to 1957 with almost yearly occurrences in some districts across the country which peaked in the mid 1970s (Department of Veterinary Services, Kenya). However, from the mid 1980s, the distribution of these outbreaks has been geographically limited to two districts in the central Rift Valley (Baringo and Koibatek). The numbers of recorded outbreaks of FMD, by serotype, for each year in Kenya from 1957 to 2006 are illustrated in Figure 5-1. Most of the reported outbreaks have been in

cattle with serotypes C and SAT 1 being the least prevalent while serotypes O and SAT 2 were the most common. In Kenya, only one type C vaccine strain (termed K267/67 and originally isolated from Laikipia district) produced by the Kenya Veterinary Vaccines Production Institute (KEVEVAPI) has been used to contain outbreaks of type C since the 1970s (Department of Veterinary Services, Kenya records). KEVEVAPI also produces vaccines for serotypes O (K77/78), A (K5/80, K35/80), SAT 1 (T155/71) and SAT 2 (K52/84).

Molecular epidemiology studies are a helpful guide for understanding disease dynamics and for vaccine development. Most viruses are maintained by complex processes of evolution that enable them to survive, and which complicates efforts aimed at achieving disease control. High levels of genetic sequence diversity, such as that reported for SAT 2 in Africa, complicates diagnosis and vaccination since the development of suitable serotype specific primers and vaccine strains is difficult (Bastos et al., 2003). Molecular characterization of FMDV field strains is useful for tracing the origin of outbreaks, and can aid in identifying possible vaccine related outbreaks, which are genetically very close to vaccine strains (Beck and Strohmaier, 1987).

Kenya is one of the few remaining suspected habitats of type C FMDV globally and efforts towards progressive control of FMD will benefit from establishing the current status of this serotype in this country. Understanding the evolutionary forces shaping the epidemiology of this intermittently occurring serotype is thus desirable. This information will be useful for decisions on control strategies, and in particular the use of vaccination.

This study was undertaken to establish patterns of evolution of type C from FMD outbreaks that occurred during the period 1967-2004 using isolates that had been stored at the Embakasi FMD laboratory. The VP1 coding region was used for sequence analysis as it codes for one of the structural proteins forming the virus capsid. It contains important antigenic sites

plus the integrin receptor binding motif and is likely to reflect the evolutionary dynamics of the virus population. It has also been widely analysed previously and thus much information is available for most serotypes.

5.2 Methods

5.2.1 Virus isolates

Eight type C virus isolates were obtained from the Embakasi FMD laboratory, Nairobi for this study. This laboratory is a repository of all FMD sample materials collected in Kenya. The type C samples had been collected over a period of almost 40 years (1967-2004). Due to the long term storage at -70^{o}C, some of the viruses were passaged at least once in BHK monolayer cells, one at a time, before RNA extraction. Attempts at passaging more of the type C samples from the bank in BHK cells were unsuccessful possibly due to deterioration as a result of unreliable power supplies. All of the eight virus isolates were collected from outbreaks that occurred in the districts of the central part of the Rift Valley province of Kenya as shown in Figure 5-2. The details of the isolates are shown in Table 5-1.

5.2.2 Viral RNA extraction, cDNA synthesis and amplification

Total RNA was extracted using the QIAamp® Viral RNA Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. cDNA synthesis was performed using Ready-to-go-You-prime first strand beads (GE Healthcare Life Sciences, Sweden) with random hexamer primers (pdN_6). PCR amplification of the VP1 coding region was achieved using forward primers; C-1C_{536} (5′-TACAGGGATGGGTCTGTGTGTACC-3′) or C-1C_{616} (5′-AAAGACTTTGAGCTCCGGCTACC-3′) with the reverse primer FMD-2B_{58} (5′-GACATGTCCTCCTGCATCTG-3′) (Knowles and Samuel, 1995) yielding products of ~880 bp

and ~800 bp respectively. Five µl of template cDNA were added to 45µl of PCR reaction mixture containing 0.2µM primers, 200 µM of each dNTP, 1.5mM $MgCl_2$ and 1U of amplitaq gold DNA polymerase (Applied Biosystems, UK). Amplifications were performed using an Eppendorf Mastercycler (Eppendorf) incorporating a touchdown profile as follows: enzyme activation at 95ºC for 5 min, 7 cycles of 95ºC for 15 sec, touchdown of 57^{o}C to 51^{o}C (with a decrease of 1ºC in the subsequent cycle) for 1 min 30 sec, 72ºC for 1 min 20 sec. This was followed by a standard PCR of 30 cycles at 95ºC for 15 sec, 50ºC 1 min 30 sec, 72ºC for 1 min 20 sec and a final extension step at 72ºC for 20 min. PCR products were examined by electrophoresis on 2% agarose gel using ethidium bromide staining and a molecular weight marker ΦX174-RF DNA (Amersham, Biosciences). The expected products were purified using the QIAquick PCR purification kit (QIAGEN). Both strands of the amplicons were cycle-sequenced using BigDye technology on an ABI 3700 automated DNA sequencer (Applied Biosystems) employing the forward primers used in the PCRs and FMD-$2A_{34}$ (5′-GAAGGGCCCAGGGTTGGACTC-3′) (Knowles and Samuel, 1995) within the 2A region as the reverse primer.

5.2.3 Sequence analysis

The entire VP1 coding regions (630 nt) from the eight Kenyan FMDV type C sequences generated in this study and nine other reference sequences (Table 5-1) were aligned using the software programs Sequencer4.8 (Gene Codes Corporation, USA) and Geneious v 4.6 (Drummond et al., 2009). The model of evolution that best fitted the data was selected using Akaike information criteria and hierarchical likelihood ratio tests in MrModeltest2.2 (Nylander, 2004) as implemented in PAUP*4b10 (Swofford, 2003) resulting in the selection of the HKY

evolutionary model (Hasegawa et al., 1985) with gamma-distributed rate variation across sites and a proportion of invariable sites (HKY+I+G) as the preferred model.

5.2.4 Sequence characteristics

The level of nucleotide sequence divergence was inferred using Dnasp v5 (Librado and Rozas, 2009) to determine values of nucleotide diversity (π) (Nei, 1987) and the number of segregating sites. The amino acid substitutions predicted from the sequences were identified using MEGA4 (Tamura et al., 2007).

5.2.5 Phylogenetic relationships

Phylogenetic relationships between the type C virus isolates were determined using MrBayes (Huelsenbeck and Ronquist, 2001) assuming an HKY + I + G model. The Markov Chain Monte Carlo search was run with 3 chains for 500000 generations; with trees being sampled every 100 generations (the first 500 trees were discarded as "burn-in").

5.3 Results

5.3.1 Sequence characteristics

A very low level of nucleotide diversity (values of $\pi = 0.0025$) was observed among the eight Kenyan FMDV type C VP1 coding sequences. There were only five segregating sites at nucleotide positions 309, 327, 331, 335 and 435. These substitutions encoded only three amino acid changes within the VP1 protein (at residues 109, 111 and 112) resulting from the non-synonymous nucleotide changes at positions 327, 331 and 335 amongst the Kenyan virus isolates (Figure 5-3). While the other Kenyan type C virus isolates were invariant at amino acid residues 109 and 112, the isolate K15/97 had two substitutions (glutamine (Q) and arginine (R) in place of histidine (H109) and proline (P112) respectively). However, the K15/97 isolate, together with

isolates K267/67 and K14/96, had a glycine (G) residue at position 111, like the reference topotypes, while the other Kenyan virus isolates had an arginine (R). It should be noted that the 'RGD' motif (residues 144-146), required for binding to the integrin receptor (Neff et al., 1998), (Jackson et al., 2000), was completely conserved amongst the Kenyan serotype C virus isolates.

5.3.2 Phylogenetic relationships

The Kenyan FMDV type C virus isolates are very closely related to each other within the African topotype (Figure 5-4). They belong to a single clade on the phylogenetic tree. The vaccine strain K267/67 is identical to K14/96 in this region while the others (K11/96, K41/97, K41/98 and K6/04*)* are identical to K60/92. This grouping reflects the identity of residue 111; which is a glycine (G) in the former and arginine (R) in the latter strains respectively. The other nucleotide changes (positions 309, 327, 335 and 435) observed were all in isolate K15/97.

5.4 Discussion

There has been very little change in the sequence of the VP1 coding region over a period of almost 40 years in the Kenyan type C FMDV isolates. Moreover, the vaccine strain (K267/67) that has been used to contain type C outbreaks is genetically closely related to each of the field strains. This lack of significant variation among viruses isolated over a long period of time is inconsistent with the expected levels of divergence given the rapid evolution of FMD viruses previously reported, e.g. (Sobrino et al., 1986; Martinez et al., 1992; Haydon et al., 2001; Tully and Fares, 2008). FMDV serotypes are known to evolve rapidly resulting in a high lineage turnover (Mittal et al., 2005). Differences exist in the selection forces shaping the evolution of the serotypes with, for example, type C and SAT 3 major antigenic regions reported to be under no positive Darwinian selection when VP1 coding sequences were analysed (Tully and Fares, 2006). However, the low level of sequence divergence observed in this study is comparable to

that observed among C_1 subtype strains from Europe, which showed very little change from 1953-1989 (Knowles and Samuel, 2003). The absence of significant sequence divergence among type C strains has also been reported in India (Nagendrakumar et al., 2005). The apparently static evolution in the European subtype C_1 strains was attributed to multiple re-introductions of vaccine strains into the field through laboratory escapes or improperly inactivated vaccines. Many FMD outbreaks in Europe in the past have been attributed to improperly inactivated vaccines when formaldehyde was used (Beck and Strohmaier, 1987). For Kenya, available records at KEVEVAPI indicate that by the early 1980s, formaldehyde had been abandoned and binary ethyleneimine (BEI) was in use for virus inactivation during vaccine production as recommended in Doel, (2003). Type C vaccine has been distributed to the districts reporting type C outbreaks solely through the Directorate of Veterinary Services. The practice has been to apply ring vaccination to contain the outbreaks once reported and this applies to all the isolates in this study i.e. vaccine was applied following the outbreak. Type C vaccine use peaked in the 1970s and early 1980s when the number of reported outbreaks was high (up to ca. 50 outbreaks in a year affecting many districts) but then reduced in the 1990s when only a few districts were affected. Kajiado district in the southern Rift Valley received the highest number of type C FMDV vaccine doses in 1973 (totaling 105,800 doses). From the 1990s up to the last reported outbreak of type C in 2004, Baringo and Koibatek were the only districts that were supplied with type C vaccine to ring vaccinate and control respective outbreaks.

In a preliminary study, Roeder and Knowles (2009) reported a close relationship between some Kenyan type C strains and the vaccine strain K267/67, and suggested that this could be due to re-introductions of the vaccine strain into the field. They recommended the

suspension of the use of the vaccine. Additionally, Kenyan field isolates of type C were observed to be antigenically closely related to the vaccine strain (Embakasi FMD lab records).

In the Kenyan situation, it is highly improbable that direct laboratory escape of the virus could be a source of these particular outbreaks as the concerned districts are geographically distant (more than 200 km) from Nairobi where the laboratory is located. Previous laboratory escapes of FMD viruses have been reported to occur near laboratories as was the case in the United Kingdom in 2007 (Cottam et al., 2008). Epidemiological information on the Kenyan outbreaks obtained by the Veterinary department on the likely source and history of each outbreak also excluded possible escapes from the laboratory based on the fact that they occurred in districts far away from Nairobi and no obvious transmission method, such as through laboratory personnel or visitors, could be found.

A possible explanation for these results is improperly inactivated vaccines as a probable source of field re-introductions, which is consistent with the conclusions from Roeder and Knowles (2009) and similar reports in Europe (Beck and Strohmaier, 1987), although the use of BEI early in the history of vaccine production at KEVEVAPI and the fact that vaccine was only used after the outbreak was reported makes this inference uncertain. High sequence similarity could also result from cross contamination of the samples, for example during cell culture passage. Such a possibility is also unlikely in this study since the isolates were passaged at different times and incorporated negative controls of uninfected monolayer cells incubated in parallel and checked for lack of any CPE in strict adherence to good laboratory practice. The suggestion of improperly inactivated vaccines, although not unequivocally proven, highlights the importance of strengthening the availability of laboratory diagnostic capability at national FMD laboratories. However, to our knowledge, neither a definitive confirmation of vaccine related

outbreaks nor its possible epidemiological impact in Kenya has been reported. Although the conservative magnitude of change as well as the close similarity to the vaccine strain observed in this study for type C suggests that type C outbreaks are likely of vaccine origin, we found little evidence of vaccine strain re-introductions for the other serotypes in Kenya (unpublished results by the authors). Instead we observe a high rate of virus turn over for the other Kenya FMD virus serotypes as exemplified by as many as 6 (SAT 2), 4 (type A) and 3 (type O) changes in vaccine strains over a similar period. This is despite the vaccine production procedure being the same for all the serotypes and the other serotype vaccines being used in much larger quantities and over a similar time period as the type C vaccine. Similarly, no outbreaks have been associated with vaccine inoculation during potency experiments. However, on the World Reference Laboratory for FMD website (http://www.wrlfmd.org/fmd_genotyping/africa/ken.htm), there appears to be evidence of Kenyan field viruses being isolated sporadically which have very close genetic relationships to the other Kenyan vaccine strains.

The suggestion of probable vaccine strain re-introductions of this serotype in Kenya has implications for the control of the disease through vaccination. It may therefore be necessary to obtain definitive evidence for the possible linkage of the use of vaccines in the maintenance of serotype C FMD in Kenya. The association of vaccine usage to outbreaks poses a dilemma for the veterinary authorities in the country who have to decide on whether to continue maintaining expensive vaccine stocks for a seemingly disappearing serotype.

Acknowledgements

We sincerely thank the Director of Veterinary Services, Kenya, for providing the virus isolates used in the study and the Managing Director of KEVEVAPI, Dr Rufus Rumberia, for helpful comments and information on FMD vaccine production and use in Kenya. Dr Sabenzia Wekesa

is thanked for providing information on the isolates. Teresa Kenduiywo, William Birgen and Eugene Arinaitwe are particularly appreciated for excellent technical assistance. This work was supported by the Danish International Development Agency (DANIDA) under the Livestock-Wildlife Diseases in East Africa Project.

References

Bastos, A.D.S., Haydon, D.T., Sangare´, O., Boshoff, C.I., Edrich, J.L., Thomson, G.R., 2003. The implications of virus diversity within the SAT 2 serotype for control of foot-and-mouth disease in sub-Saharan Africa. *Journal of General Virology* **84**, 1595–1606.

Beck, E., Strohmaier, K., 1987. Subtyping of European foot-and-mouth disease virus strains by nucleotide sequence determination. *Journal of Virology* **61**, 1621-1629.

Belsham, G.J., 1993. Distinctive features of foot-and-mouth disease virus, a member of the picornavirus family; aspects of virus protein synthesis, protein processing and structure. *Progress in Biophysics and Molecular Biology* **60**, 241-260.

Cottam, E.M., Wadsworth, J., Shaw, A.E., Rowlands, R.J., Goatley, L., Maan, S., Maan, N.S., Mertens, P.P.C., Ebert, K., Li, Y., Ryan, E.D., Juleff, N., Ferris, N.P., Wilesmith, J.W., Haydon, D.T., King, D.P., Paton, D.J., Knowles, N.J., 2008. Transmission pathways of foot-and-mouth disease virus in the United Kingdom in 2007. *PLoS Pathogens* **4**, 1-8.

Doel, T.R., 2003. FMD vaccines. *Virus Research* **91**, 81-99.

Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., Wilson, A., 2009. *Geneious v4.6.* Available from http://www.geneious.com/.

Grubman, M.J., Baxt, B., 2004. Foot-and-Mouth Disease. *Clinical microbiology reviews* **17**, 465–493.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **21**, 160-174.

Haydon, D.T., Bastos, A.D., Knowles, N.J., Samuel, A.R., 2001. Evidence for positive selection in foot-and-mouth disease virus capsid genes from field isolates. *Genetics* **157**, 7–15.

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

Jackson, T., Shepherd, D., Denyer, M., Blakemore, W., King, A.M.Q., 2000. The epithelial integrin alphavbeta6 is a receptor for foot-and-mouth disease virus. *Journal of virology* **74**, 4949-4956.

Knowles, N.J., Samuel, A.R., 1995. Polymerase chain reaction amplification and cycle sequencing of the 1D gene of foot-and-mouth disease viruses. *Session of the research group of the standing technical committee of the European commission for the control of foot-and-mouth disease*. FAO, Rome, 19-22 September 1994, Vienna, Austria.

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

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

Martinez, M.A., Dopazo, J., Hernandez, J., Mateu, M.G., Sobrino, F., Domingo, E., Knowles, N.J., 1992. Evolution of the capsid protein genes of foot-and-mouth disease virus: antigenic variation without accumulation of amino acid substitutions over six decades. *Journal of Virology* **6**, 3557-3565.

Mittal, M., Tosh, C., Hemadri, D., Sanyal, A., Bandyopadhyay, S.K., 2005. Phylogeny, genome evolution, and antigenic variability among endemic foot-and-mouth disease virus type A isolates from India. *Archives of Virology* **150**, 911-928.

Nagendrakumar, S.B., Reddy, G.S., Chandran, D., Thiagarajan, D., Rangarajan, P.N., Srinivasan, V.A., 2005. Molecular characterization of foot-and-mouth disease virus type C of Indian origin. *Journal of Clinical Microbiology* **43**, 966–969.

Neff, S., Sa-Carvalho, D., Rieder, E., Mason, P.W., Blystone, S.D., Brown, E.J., Baxt, B., 1998. Foot-and-mouth disease virus virulent for cattle utilizes the integrin alphavbeta3 as its receptor. *Journal of Virology* **72**, 3587-3594.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press New York.

Nylander, J.A.A., 2004. *MrModeltest v2.* Program distributed by the author Evolutionary Biology Centre, Uppsala University.

Roeder, P.L., Knowles, N.J., 2009. Foot-and-mouth disease virus type C situation: The first target for eradication? , *The Global control of FMD - Tools, ideas and ideals*. FAO, Rome, 14-17 October 2008, Erice, Italy.

Sobrino, F., Palma, E.L., Beck, E., Davila, M., Torre, J.C.d.l., Negro, P., Villanueva, N., Ortin, J., Domingo, E., 1986. Fixation of mutations in the viral genome during an outbreak of foot-and-mouth disease: heterogeneity and rate variations. *Gene* **50**, 149-159.

Swofford, D.L., 2003. *PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version4.* Sinauer Associates, Sunderland, Massachusetts.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution* **24**, 1596–1599.

Tully, D.C., Fares, M.A., 2006. Unravelling selection shifts among foot-and-mouth disease virus (FMDV) serotypes. *Evolutionary Bioinformatics* **2**, 211-225.

Tully, D.C., Fares, M.A., 2008. The tale of a modern animal plague: tracing the evolutionary history and determining the time-scale for foot and mouth disease virus. *Virology* **382**, 250–256.

Waldmann, O., Trautwein, K., 1926. Experimentelle untersuchungen über die pluralität des maul-und klauenseuchevirus. *Berl. Tierärztl. Wochenschrift.* **42**, 569-571.

Table 5-1 List of the type C viruses included in this study of which 8 sequences derived from Kenyan viruses were generated in this work while 9 were sourced from published reports

Lab Ref No.	Year of Isolation	District/Country	Accession No.	Topotype
ETH/1/71	1971	Ethiopia	FJ798151	Africa
USSR/Tadjn/67	1967	Tadjikstan	WRL	Asia
IND/42/77	1977	India	WRL	Asia
BRA/Resnd/55	1955	Brazil	M90381	Euro-SA
SPA/StPa/70	1970	Spain	AJ133357	Euro-SA
BRA/Indl/71	1971	Brazil	M90376	Euro-SA
PHI/7/84	1984	Philipines	WRL	Euro-SA
GER/CGC/26	1926	Germany	EU553893	?
UK/149/34	1934	United Kingdom	AY593810	?
K60/92	1992	Baringo, Kenya	GU451110*	Africa
K15/97	1997	Koibatek, Kenya	GU451111*	Africa
K267/67	1967	Laikipia, Kenya	GU451109*	Africa
K14/96	1996	Koibatek, Kenya	GU451116*	Africa
K41/97	1997	Baringo, Kenya	GU451112*	Africa
K6/04	2004	Koibatek, Kenya	GU451113*	Africa
K41/98	1998	Koibatek, Kenya	GU451114*	Africa
K11/96	1996	Koibatek, Kenya	GU451115*	Africa

? = not assigned WRL = From World reference laboratory sequence data website (http://www.wrlfmd.org/fmd_genotyping/prototypes.htm). * = This study

Figure legends

Figure 5-1 Frequency of recorded outbreaks of FMD by serotype in Kenya (1954-2006)

Figure 5-2 Map of Kenya indicating the geographic origin of the type C FMDV isolates included in this study**.**

Figure 5-3 VP1 coding region and amino acid sequence comparison between the serotype C sequences analysed in this study. Only variable sites to indicate changes in the Kenyan sequences are shown with nucleotide and amino acid positions marked above and below the sequences respectively. Variable amino acids (aa) are indicated in bold. A `.´ indicates a nucleotide site identical to that of the sequence GER/CGC/26 (Germany, 1926) and `-´denotes a missing nucleotide.

Figure 5-4 Bayesian inference tree indicating the phylogenetic relationships between serotype C FMDV isolates. The three topotypes (Africa, Asia and Europe-South America) are indicated. Posterior probabilities >50% for nodes are shown.

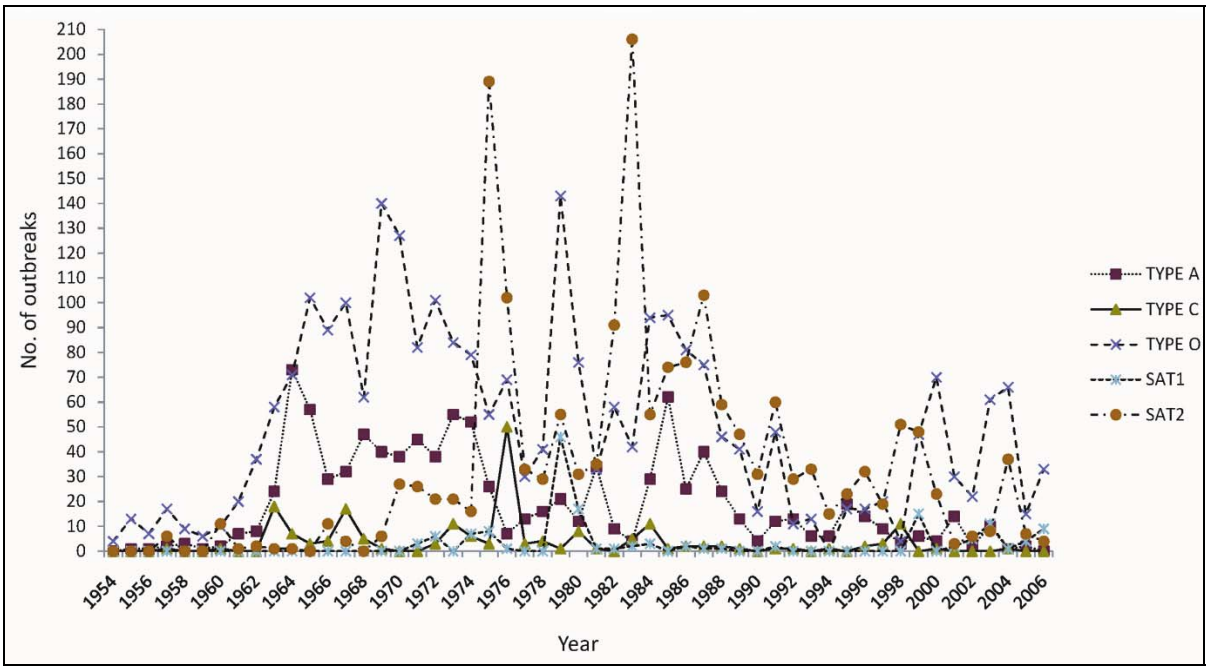


Figure 5-1

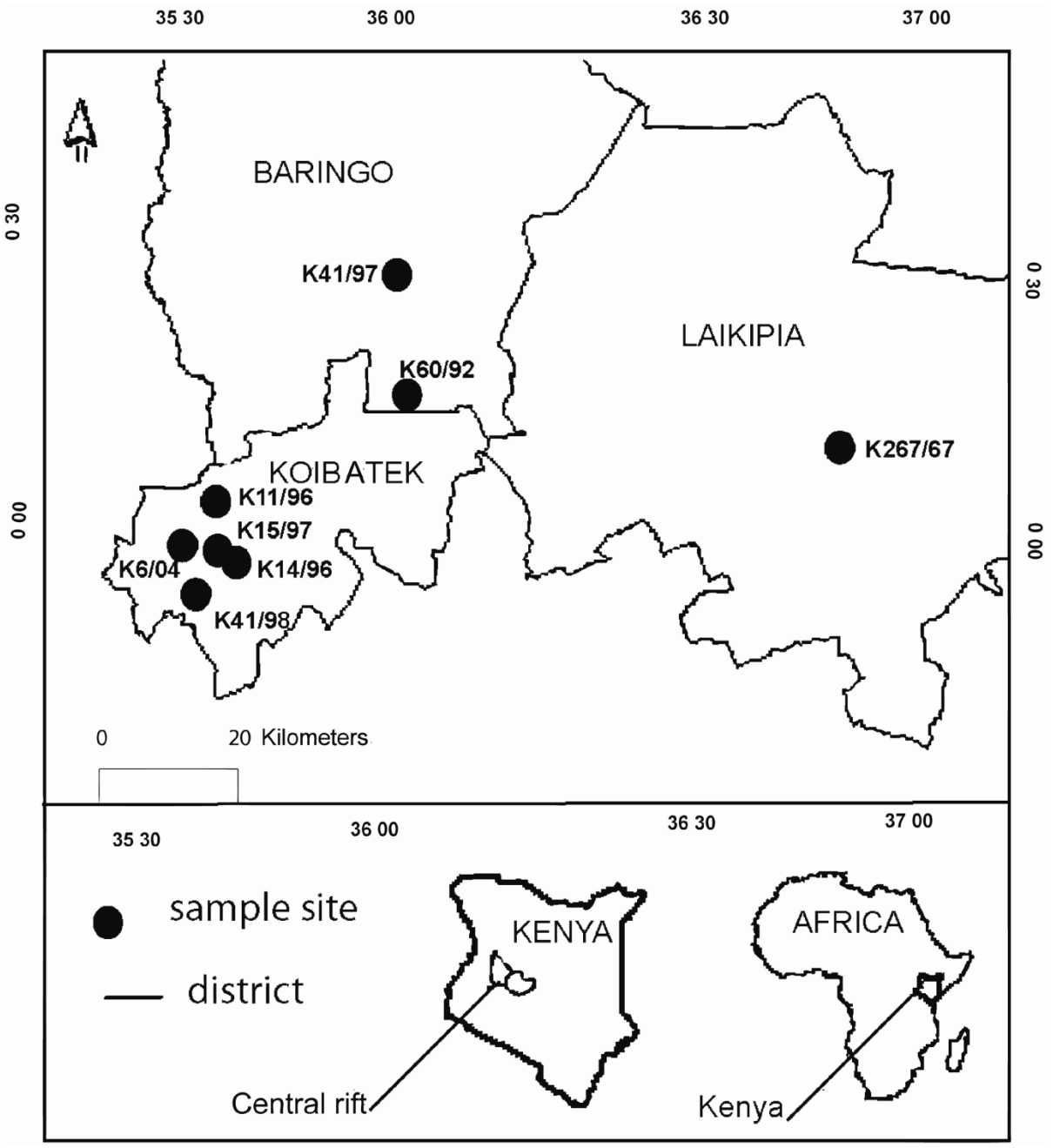

35 30
36 00
36 30
37 00
0 30
0 00
BARINGO
K41/97
LAIKIPIA
K60/92
K267/67
KOIBATEK
K11/96
K15/97
K6/04
K14/96
K41/98
0
20 Kilometers
sample site
district
KENYA
AFRICA
Central rift
Kenya

Figure 5-2

	T/G/A	Q	R	R	G/S	L/M/P			T/G	Q	
	S	H	G	P	R	G	D	S	A	H	L
	333	333	333	333	444	444	444	444	444	444	444
	000	222	333	333	333	333	333	344	444	444	445
	789	567	123	456	012	345	678	901	234	567	890
GER/CGC/26	AGC	CAC	GGG	CCG	AGA	GGG	GAT	TCG	GCT	CAC	CTG
UK/149/34	...	...	...	...	G..	...	...	CT.	A..	...	...
BRA/Res/55	.C.	...	..A	...	...	..A	...	CT.	..C	..T	T..
USSR/Ta/67	G..	...	..A	...	..T	...	..C	AT.	...	...	T..
SPA/StP/70	.C.	...	..C	...	---	...	...	.T.	...	...	..A
BRA/Ind/71	GC.	...	..A	...	..G	..A	...	CTA	..C	...	T..
IND/42/77	...	...	..A	...	...	...	..C	AT.	...	..T	T..
PHI/7/84	.C.	...	..A	...	G..	..A	...	.T.	..C	..G	...
ETH/1/71	...	...	..A	...	...	..A	...	C..	.G.	..A	...
K267/67	...	...	..A	...	...	..A	..C	CT.	.G.	..A	...
K60/92	...	...	A.A	...	...	..A	..C	CT.	.G.	..A	...
K11/96	...	...	A.A	...	...	..A	..C	CT.	.G.	..A	...
K14/96	...	...	..A	...	...	..A	..C	CT.	.G.	..A	...
K15/97	..T	..A	..A	.G.	...	...	..C	CT.	.G.	..A	...
K41/97	...	...	A.A	...	...	..A	..C	CT.	.G.	..A	...
K41/98	...	...	A.A	...	...	..A	..C	CT.	.G.	..A	...
K6/04	...	...	A.A	...	...	..A	..C	CT.	.G.	..A	...
aa position	**103**	**109**	**111**	**112**	**144**	**145**	**146**	**147**	**148**	**149**	**150**

Figure 5-3

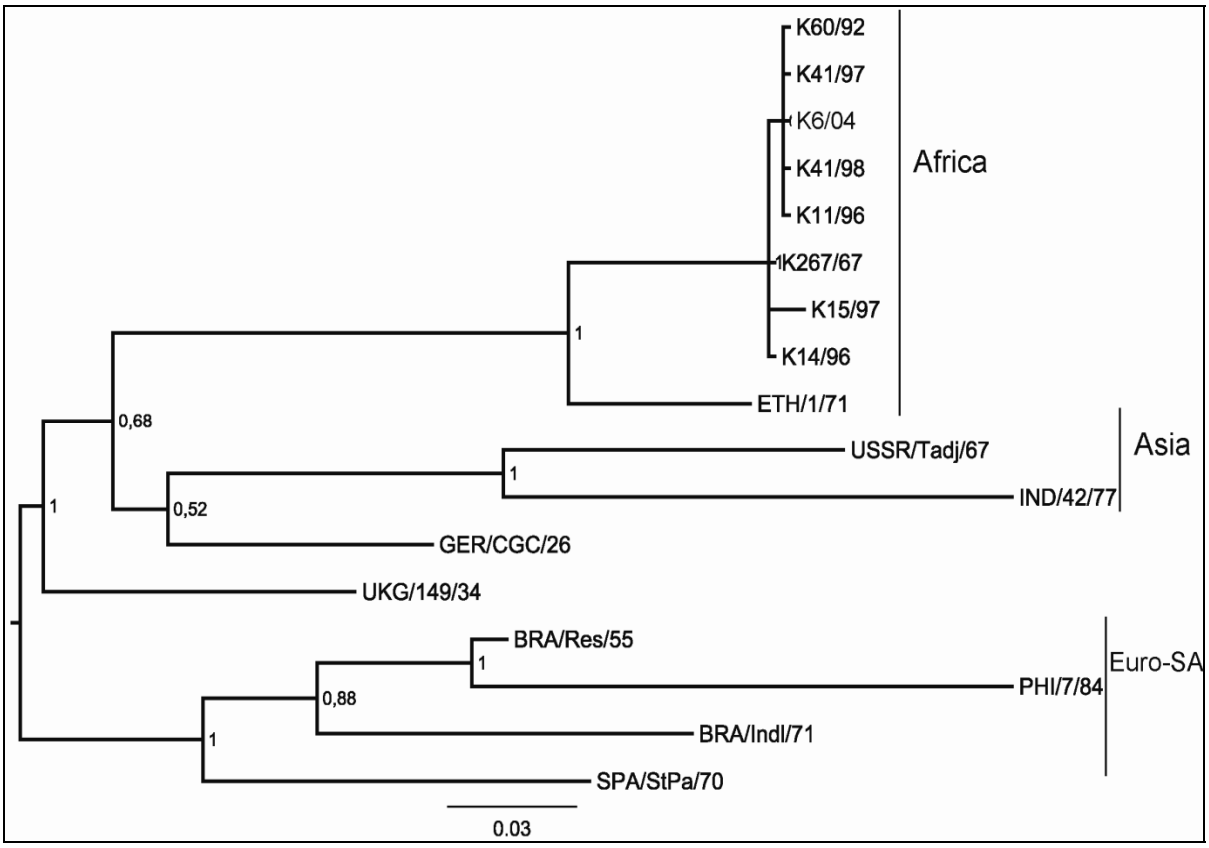


Figure 5-4

CHAPTER SIX

Evolutionary analysis of foot-and-mouth disease virus serotype SAT 1 isolates from East Africa suggests two independent introductions from southern Africa

Evolutionary analysis of foot-and-mouth disease virus serotype SAT 1 isolates from East Africa suggests two independent introductions from southern Africa

A.K. Sangula[1,4§], G.J. Belsham[2], V.B. Muwanika[1], R. Heller[3], S.N. Balinda[1], C. Masembe[1], H.R. Siegismund[3]

[1] Makerere University, Institute of Environment and Natural Resources, Molecular Biology Laboratory, P. O. Box 7298, Kampala, Uganda

[2] National Veterinary Institute, Technical University of Denmark, Lindholm, DK-4771 Kalvehave, Denmark

[3] Department of Biology, Ole Maaløes Vej 5, DK-2200, Copenhagen N, Denmark

[4] Foot-and-Mouth Disease Laboratory, Embakasi, P. O. Box 18021, 00500, Nairobi, Kenya

[§]Corresponding author

Abraham K. Sangula

Makerere University Institute of Environment and Natural Resources,

Molecular Biology Laboratory,

P.O. Box 7298, Kampala, Uganda

Fax: +256 414 530134 Phone + 256 712629425

Email addresses:

aksangula@muienr.mak.ac.ug / aksangula@gmail.com

Abstract

Background

In East Africa, foot-and-mouth disease virus serotype SAT 1 is responsible for occasional severe outbreaks in livestock and is known to be maintained within the buffalo populations. Little is known about the evolutionary forces underlying its epidemiology in the region. To enhance our appreciation of the epidemiological status of serotype SAT 1 virus in the region, we inferred its evolutionary and phylogeographic history by means of genealogy-based coalescent methods using 53 VP1 coding sequences covering a sampling period from 1948-2007.

Results

The VP1 coding sequence of 11 serotype SAT 1 FMD viruses from East Africa has been determined and compared with known sequences derived from other SAT 1 viruses from sub-Saharan Africa. Purifying (negative) selection and low substitution rates characterized the SAT 1 virus isolates in East Africa. Two virus groups with probable independent introductions from southern Africa were identified. One group was exclusive to Uganda while the other was present within Kenya and Tanzania.

Conclusions

Our results provide a baseline characterization of the inter-regional spread of SAT 1 in sub-Saharan Africa and highlight the importance of a regional approach to trans-boundary animal disease control in order to monitor circulating strains and apply appropriate vaccines.

Background

Foot-and-mouth disease (FMD) is an acute, highly communicable and economically important disease of livestock and it also affects wild ruminants (Grubman and Baxt, 2004). The causative agent, foot-and-mouth disease virus (FMDV) belongs to the *Aphthovirus* genus in the family *Picornaviridae.* Its positive-sense, single-stranded RNA genome of 8.5 kb is translated into a polyprotein which is post-translationally cleaved to 4 structural (VP1, VP2, VP3, VP4) and 8 nonstructural proteins (Belsham, 199). The structural proteins form the capsid of the virion and, with the exception of VP4, are surface exposed. VP1 is involved in the interaction with the host cell via the RGD-dependent integrins (Jackson et al., 2003).

The coding sequence for VP1 has been widely used in studies of evolutionary dynamics of FMDV needed for the understanding of the epidemiological patterns of these viruses and for determining possible sources of outbreaks (Haydon et al., 2001a; Knowles and Samuel, 2003; Tully and Fares, 2006). The genetic diversity of FMDV is a consequence of the high mutation rate due to the error-prone RNA polymerase lacking proofreading activity (Domingo et al., 2003).

There are seven immunologically distinct serotypes of FMDV (O, A, C, SAT 1, SAT 2, SAT 3 and Asia 1), each with a wide spectrum of antigenic and epidemiological subtypes distributed around the world (Knowles and Samuel, 2003). The Southern Africa Territories (SAT) serotypes are restricted in their distribution mainly to sub-Saharan Africa and they co-exist with the Euro-Asiatic (O, A, C) serotypes in the East African region although serotype C has not been reported since 2004. In southern Africa, the epidemiology of the SAT serotypes is mainly associated with African buffalos (*Syncerus caffer*) which act as reservoirs and sources of outbreaks (Condy et al., 1985; Thomson et al, 2003). In eastern Africa, FMD is prevalent in wildlife and within the African buffalo in particular although their role in the epidemiology of

the disease has not been as widely studied as in southern Africa. Most outbreaks of FMD in the region are reported among livestock populations. The African buffalo has been reported to be a carrier of serotypes SAT 1 and SAT 2 but not the Euro-Asiatic serotypes in Kenya (Anderson et al., 1979; Bronsvoort et al., 2008). This is similar to the situation in southern Africa.

Widespread animal movements in the eastern Africa region are possibly responsible for long-term circulation and reintroductions of FMDV strains, including SAT 1 (Sahle et al., 2007). However, little quantitative information exists about the extent of such livestock- and wildlife mediated dispersal of FMDV as well as the origin and evolutionary history of the SAT 1 viruses circulating in eastern Africa (Sahle et al., 2007; Bastos et al., 2001). Furthermore, the connectivity between the individual countries and the main routes of dispersal remain unknown, although such information would be of great value in containing the spread of the disease and avoiding introduction of novel strains against which existing vaccine programs offer little protection.

We have investigated the emergence of FMDV SAT 1 diversity in the region by inferring the phylogeographic history and selection forces shaping them by means of genealogy-based coalescent methods. Furthermore, we have tested for evidence of recombination in the data set which is known to bias phylogenetic inferences as observed in (Schierup and Hein, 2000; Moya et al., 2004; Heath et al., 2006)

Results

Phylogenetic relationships, substitution rates and divergence times

The VP1 coding sequences of 11 additional serotype SAT 1 FMD viruses from East Africa have been determined. Using this information the complete VP1 coding sequences of 8 southern Africa, 14 western Africa, 4 horn of Africa and 27 eastern Africa FMD serotype SAT 1

viruses from the period 1948 to 2007 were analysed to determine phylogenetic relationships, divergence times and substitution rates. Estimates of divergence times from a common root showed considerable uncertainty in determination with an estimated TMRCA at 360 years before present (ybp) (95% HPD: 130-710). The inferred maximum clade credibility tree is shown in Fig. 1 with the posterior probabilities for the significant branches shown. The East African SAT 1 viruses formed two main clades (lineages) labelled A and B supported by high posterior probabilities. The Ugandan viruses differed from those of Tanzania and Kenya and are of mainly one lineage (A) while one isolate (UGA/13/74) grouped with viruses of the Sudan and West Africa. Kenyan and Tanzanian viruses grouped together in lineage B and were related to a Zimbabwean isolate. No further geographic structure was observed within lineage B isolates from Kenya and Tanzania, suggesting high migration rates between these countries.

The mean nt substitution rate was 2.05×10^{-3} substitutions/site/year (s/s/yr) (95% HPD: 7.51×10^{-4} - 3.42×10^{-3}) with distinct variation in rates among the clades. We analysed the eastern African viruses (comprising 27 samples) separately in BEAST and found relatively lower rates at 2.75×10^{-4} s/s/yr (4.69×10^{-5} - 7.39×10^{-4}), while the West African viruses (comprising 14 samples) had higher rates at 6.91×10^{-3}s/s/yr (3.32×10^{-3} - 1.04×10^{-2}).

While the location of the root of the SAT1 tree could not be identified with particular confidence (Bayes factor = 1.4 comparing the posterior probability of the root being in southern Africa against the Horn of Africa), there was relatively strong support for the location of several of the remaining nodes in the MCC tree (Figure 1). From the location-annotated MCC tree, two separate introductions from southern Africa to eastern Africa were supported by the data, namely one leading to lineage A and one leading to lineage B. In addition, there was strong support for two separate introductions of SAT 1 from the Horn of Africa to western Africa. Bayes factor

tests revealed that the most significant routes of inter-regional dispersal were the horn of Africa-western Africa (BF = 77) and southern Africa-Kenya/Tanzania (BF = 5). No link between the Ugandan and the Kenyan/Tanzanian samples (between lineage A and B) could be identified, and this was in fact found to be the link with the second-lowest posterior support. The western Africa-Kenya/Tanzania link had the lowest support (results not shown).

Predominant purifying selection in the VP1 coding region of FMDV SAT 1

The majority of codons in the VP1 coding region of FMDV SAT 1 appeared to be under purifying (negative) selection. Of the 221 codons analysed, 153 were found to be under negative selection using three methods (SLAC at $P = 0.1$, FEL at $P = 0.1$ and REL with BF > 50) as summarized in Table 2. Five sites (codons 47, 61, 99, 143 and 147) were identified to be under positive selection by at least one method but no site was identified by all the three methods together at values of $P = 0.1$ (SLAC and FEL) or BF > 50 (REL). However, at $P = 0.2$, codon 147 (H/N/E/T 2 codon positions before the receptor binding motif RGD) was identified by all the methods and was mostly likely to be under true positive selection.

The GA Branch analysis showed that 5 rate classes were supported with a large number of models (over 1500 in the 95% confidence set). No branches had significant support for $dN > dS$ although differences existed in the branch selection pattern indicating that some branches may have been under weak positive selection.

Recombination

The GARD method detected a putative recombination breakpoint at nucleotide position 168 with a change in Akaike's Information Criterion (AIC_c) >100 which suggested support for recombination model while the Kishino-Hasegewa test showed support for significant topological incongruence at $P = 0.01$. Indeed, the exploratory analysis using RDP2 had at least

one method detect some recombinant sequences (TAN/60/99 and K66/80). However, further analysis did not support the view that these sequences were recombinant and the exclusion of these sequences from the analysis did not affect the phylogenetic results, indicating that they are not likely to be true recombinants (results not shown).

Discussion

FMDV serotype SAT 1 virus strains from East Africa analysed in this study grouped into 2 distinct clades (lineages with >20% nucleotide divergence) designated here as lineages A and B. While one of these lineages was found exclusively in Uganda, the other had virus strains from Tanzania and Kenya. Over the whole sampling period, Kenyan and Tanzanian isolates were interspersed in the phylogenetic tree, suggesting that these countries form a coherent ecosystem for SAT 1. The separate introduction of lineages A and B to eastern from southern Africa was supported by the high posterior probabilities of the location states in the phylogeographic analysis. A close association of Kenyan/Tanzanian and southern African lineages has been observed earlier (Sahle et al., 2007), but the link between Ugandan and SA lineages reported here reveals a previously undiscovered aspect of the ancestry of the East African SAT 1 lineages. This new finding was not due to a different data set, but rather to our Bayesian phylogeographic analysis framework. When we constructed a neighbour-joining tree using similar methods to (Sahle et al., 2007) using our data set, we were not able to infer a SA origin to the Ugandan lineage.

Several interesting aspects about the history of sub-Saharan SAT 1 viruses emerge from our continental phylogeographic approach. First, we found that the most likely root location of SAT 1 is in southern Africa. Because of the relatively deep root of the tree (~350 ybp), we could not achieve unequivocal posterior support for this root location (see also Lemey et al., 2009). We

found a strong link between western African and Horn of Africa SAT 1s (in agreement with Sahle et al., 2007) and our results suggest that the route of entry of SAT 1 into western Africa has been along the Sahel rather than through the rain forest belt surrounding equatorial Africa. A Ugandan isolate from 1974 was found to belong to to a lineage otherwise consisting of Sudanese, Ethiopian and western African strains, and the phylogeographic analysis suggested this was an incursion from the Horn of Africa. Hence, Ugandan SAT 1 strains appeared to be derived from two different sources, southern Africa and the Horn of Africa, respectively.

The sampling scheme employed in this study may have affected the outcome of phylogeographic analyses to some extent. For example, we cannot exclude that the inclusion of more samples from Uganda would alter the posterior state probability of some nodes in the tree to reflect an earlier introduction of SAT 1 to Uganda. Given that Uganda is represented in both of the two major clades, it may have played a more prominent role connecting southern African SAT 1s with those of the horn of Africa and western Africa. Such a scenario seems plausible given the central location of Uganda according to our definition of location states. Furthermore, we cannot exclude that additional samples from Uganda will show phylogenetic affinity with the surrounding countries. This could be easily tested by acquiring more Ugandan samples. In fact, more recent SAT 1 virus isolates from Uganda have been shown to group within the Ugandan lineage A (Ayebazibwe and others, unpublished results), in agreement with the phylogeographic conclusions reported here. In general, however, we stress that our findings should be viewed as a null hypothesis about continental SAT 1 dispersal against which studies based on more comprehensive sampling can be tested. Denser sampling (both temporally and spatially) will surely reveal novel dispersal patterns not observed here and further address the fine-scale historical movement of the serotype.

The substitution rate inferred in our study differs considerably from Tully and Fares (2008). This leads to a significantly deeper tree, and hence it is difficult for us to put our results in to a historical context that includes all FMDV serotypes. Our mean estimate of 575 ybp for the TMRCA of SAT 1 actually predates that of the whole FMDV found in Tully and fares (2008), although it is within the 95% HPDs reported in that study (218–1250 ybp). We caution that the time line of our phylogeographic tree should not be regarded as conclusive and that further studies are needed to establish the rate of evolution in FMDV. Our inferred rate is, however, closer to the reported mean rate of evolution across all serotypes (2.48×10^{-3}) for the VP1 gene. In Tully and Fares (2008), SAT 1 was found to have a roughly 3-fold faster rate than the species average. We speculate that this exceptionally fast rate could be derived from the sampling scheme in Tully and Fares (2008), where many of the included SAT 1 isolates are from the same epidemic outbreaks. This tends to yield faster rates of evolution, since what is recovered is actually the mutation rate rather than the long-term substitution rate subject to selection and other forces Ho et al., 2005), leading to a bias towards higher rates and a shallower tree. In accordance with this, we did find much faster rates of evolution in the western African samples, all collected during two epidemic outbreaks each spanning just two years. However, regionally variable evolutionary rates may in fact reflect real differences in the epidemiological dynamics and host-interaction of FMDV. For example, buffalos and other wildlife may play a more prominent role in the epidemiology of SAT 1 in eastern than in western Africa, and this may give rise to changed patterns of evolution of virus lineages in the two regions. Considerable localized differentiation in evolutionary rates has not previously been observed in FMDV, and although potentially informative concerning epidemiology and evolution, it also complicates evolutionary estimates based on global or widespread sample collections. Given these two (not necessarily

mutually exclusive) causes of the observed rate heterogeneity, it is vital that future studies address the caveats in using the VP1 gene sequence to infer evolutionary rates and history.

Purifying (negative) selection was the most predominant evolutionary force at play among the SAT 1 viruses. At least 153 codon positions including the RGD motif (amino acid residue positions 149-151) of VP1 were estimated to be under purifying selection signifying amino acid conservation as reflected in the low evolutionary rates. There was less evidence for positive selection although a few sites may have been under adaptive selection. Amino acid sites that are distinct between the regional virus groups as well as conservation of the RGD motif were observed when inferred using MEGA version 4 (Tamura et al., 2007) and is in agreement with previous reports in Sahle et al.,(2007). These evolutionary patterns may reflect the observed apparent long term circulation of some virus strains in the region previously reported in Sahle et al., (2007). It has also been observed that genetic heterogeneity may be limited by evolutionary constraints (Haydon et al., 2001b). There was no evidence for the presence of recombination within the VP1 coding sequences (in agreement with observations that recombination is largely restricted to non-structural coding regions with very few phylogenetic incongruities in the capsid proteins (Carrillo et al., 2005; Tully and Fares, 2009) adding confidence to our results.

Conclusions

We have inferred the most likely phylogeographic history of SAT 1 in sub-Saharan Africa. We found evidence that the SAT 1 viruses circulating in Uganda and Kenya/Tanzania represent independent phylogeographic lineages. Kenya and Tanzania appear to experience a much greater exchange of viruses at their respective southern and northern borders through the transboundary livestock and wildlife movements (a common feature in this area) than with Uganda. This highlights the importance of a regional approach to transboundary animal disease

control. It is apparent from the SAT 1 analysis presented here that monitoring of the emerging strains in the region is required for the success of vaccination strategies.

Methods

Virus Isolates

Eleven (10 Kenyan and 1 Tanzanian) SAT 1 virus isolates for this study (collected between 1977 and 2006) were obtained from the Embakasi FMD laboratory in Nairobi which is a repository of all FMD sample materials collected in Kenya. Virus was isolated from clinical material according to standard procedures on baby hamster kidney (BHK) cells. The details of the isolates are shown in Table 1.

Viral RNA extraction, cDNA synthesis and amplification

Total RNA was extracted and cDNA synthesized as previously described in Sangula et al., (2009). The complete VP1 coding region was amplified using the primer pair, FMD AKS (5′-ATGGGACACAGGTCTGAACTCGA-3′) and FMD-2B58 (Knowles and Samuel, 1995) applying PCR reagent volumes and conditions as previously described in Sangula et al., (2009). PCR products were visualized, purified and cycle-sequenced using the same primers as for PCR above.

Phylogeographic analysis

In addition to the eleven sequences generated in the study, 42 (17 from Eastern Africa and 25 the rest of Africa) other complete VP1 coding sequences available in the GenBank covering a sampling period from 1948-2007 were included to be able to put the results from eastern Africa into a continental SAT 1 context.

The sequences were assembled and aligned using the software program Geneious version 4.6 (Drummond et al., 2009) The best fitting nucleotide substitution model was tested by means

of a hierarchical likelihood ratio test (LRT) and the Akaike information criteria (AIC) as implemented in MrModeltest version 2.2 software (Nylander, 2004) and executed in PAUP* version 4b10 software (Swofford, 2003). The selected model was GTR (Rodriguez, 1990) with gamma-distributed rates among sites and a proportion of invariable sites.

Phylogenetic relationships, evolutionary rates and population size changes were co-estimated for the whole data set and geographic subsets using a Bayesian Markov Chain Monte Carlo (MCMC) method implemented in the BEAST software version 1.5.3 package (Lemey et al., 2009) (http://beast.bio.ed.ac.uk) using the selected model of nucleotide substitution. The method utilizes the sampling time of the sequences to infer rates of evolution along lineages, time of the most recent common ancestor (TMRCA) and demographic history. A recent extension of the software allows tracking of the geographic location state along the phylogenetic tree, yielding posterior estimates of the location of each branch/node in the tree given the phylogenetic uncertainties (Lemey et al., 2009). Given our limited data set with low representation of many countries, we defined the geographical states as five coherent regions roughly corresponding to areas separated by known topotype boundaries. The regions include: western Africa, the Horn of Africa, Uganda, Kenya/Tanzania and southern Africa (Table1). We used a Bayesian stochastic search variable selection (BSSVS) without distance informed priors on diffusion rates, as this has been shown not to improve confidence in the phylogeographical state assignment when dispersal patterns are complex such as in many viruses (e.g. Lemey et al., 2009). Rate indicator log files were inspected in Tracer software version 1.4 (http://tree.bio.ed.ac.uk/software/tracer/), and Bayes factor tests were carried out to test the most significant routes of dispersal.

In a preliminary analysis, we tested four different demographic models/coalescent priors as suggested in Suchard et al., (2001). In addition, we tested the appropriateness of a strict clock versus various versions of relaxed clocks available in BEAST. This process of model selection suggested the constant population size model with an uncorrelated exponential clock to be the best fit to the data. The MCMC chains were run long enough (40 million steps) to allow ESSs above 250 with a 10% burn-in as viewed in Tracer. Statistical uncertainties of the substitution rates and the MRCA were summarized as the lower 95%, mean, and upper 95% values of the highest posterior density (HPD) interval. Mean evolutionary rates (averaged over branches) were measured as the number of nucleotide substitution per site per year (s/s/y). Maximum clade credibility trees were obtained using Tree Annotator program in BEAST and visualized with FigTree version 1.1.2 software (http://tree.bio.ed.ac.uk/software/figtree/).

Selection and recombination detection

Tests for selection were performed using four methods which estimate selection in a phylogenetic context, available in the Datamonkey web interface (Pond and Frost, 2005). The best-fitting nucleotide substitution model was selected using the automated link. To identify codon sites under positive (adaptive) or negative (purifying) selection, we used the single-likelihood ancestor counting (SLAC), the fixed effects likelihood (FEL) and the random effects likelihood (REL) methods. The SLAC and FEL methods estimate selection on a site-by-site basis with the former method comparing observed to expected synonymous and non-synonymous rates while the latter uses two models which assume independent and equal rates, and a likelihood ratio test to determine significance. The REL method determines independent general discrete distributions for the global synonymous and non-synonymous rates using a codon based model which are then used as priors for Empirical Bayes analysis of site selection (Pond and Frost,

2005). The integrative selection analysis option in Datamonkey was then used to increase confidence on the estimation of selection at a site if all three methods support it. To test the hypothesis that different selective environments were acting on the branches of the phylogeny, we used the GA Branch method to estimate dN / dS.

To add confidence to our coalescent inferences, the presence of recombination in the data was tested using the GARD method (Pond et al., 2006) on the Datamonkey server with topological incongruence significance estimated by the Kishino-Hasegawa test (Kishino and Hasegawa, 1989) and also by the exploratory methods implemented in RDP version 2 beta 0.8 software (Martin et al., 2005) which included: RDP, (Martin and Rybicki, 2000) GENECONV, (Padidam et al., 1999) Bootscan, (Salminen et al., 1995 MaxChi, (Maynard Smith, 1992) and Chimaera (Posada and Crandall, 2001).

Authors' contributions

AKS, GJB, VM and HRS designed and conceived the study. AKS, SNB and CM generated, collected and aligned the sequences. AKS, RH, SNB and HRS carried out the analysis of the data. AKS, GJB, RH, VM and HRS wrote the paper. All authors read and approved the final manuscript.

Acknowledgements

We sincerely thank the Director of Veterinary Services, Kenya, for providing the virus isolates used in the study and Dr Sabenzia Wekesa of the FMD laboratory, Embakasi in particular for the information on the isolates and comments. Teresa Kenduiywo, William Birgen and Eugene Arinaitwe are particularly appreciated for excellent technical assistance. Part of this work was carried out by using the resources of the Computational Biology Service Unit from Cornell University which is partially funded by Microsoft Corporation. This work was supported by the

Danish International Development Agency (DANIDA) under the Livestock-Wildlife Diseases in East Africa Project.

References

Anderson EC, Doughty WJ, Anderson J, Paling R: The pathogenesis of foot-and-mouth disease in the African buffalo (*Syncerus caffer*) and the role of this species in the epidemiology of the disease in Kenya. *Journal of Comparative Pathology* 1979, **89**:511 - 519.

Bastos ADS, Haydon DT, Forsberg R, Knowles NJ, Anderson EC, Bengis RG, Nel LH, Thomson GR: Genetic heterogeneity of SAT-1 type foot-and-mouth disease viruses in southern Africa. *Archives of Virology* 2001, **146**:1537–1551.

Belsham GJ: Distinctive features of foot-and-mouth disease virus, a member of the picornavirus family; aspects of virus protein synthesis, protein processing and structure. *Progress in Biophysics and Molecular Biology* 1993, **60**:241-260.

Bronsvoort BMDC, Parida S, Handel I, McFarland S, Fleming L, Hamblin P, Kock R: Serological survey for foot-and-mouth disease virus in wildlife in eastern Africa and estimation of test parameters of a nonstructural protein enzyme-linked immunosorbent assay for buffalo. *Clinical and Vaccine Immunology* 2008, **15**(6):1003–1011.

Carrillo C, Tulman ER, Delhon G, Lu Z, Carreno A, Vagnozzi A, Kutish GF, Rock DL: Comparative genomics of foot-and-mouth disease virus. *Journal of Virology* 2005, **79**(10):6487–6504.

Condy JB, Hedger RS, Hamblin C, Barnett ITR: The duration of foot-and-mouth disease virus carrier state in African buffalo (i) in the individual animal and (ii) in a free-living herd. *Comparative lmmunology, Microbiology and Infectious diseases* 1985, **8**(3/4):259- 265.

Domingo E, Escarmı´s C, Baranowski E, Ruiz-Jarabo CM, Carrillo E, Nunez JI, Sobrino F: Evolution of foot-and-mouth disease virus. *Virus Research* 2003, **91**:47-63.

Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A: Geneious v4.6. In.: Available from http://www.geneious.com/; 2009.

Drummond AJ, Rambaut A: BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 2007, **7**. Doi. 10.1186/1471-2148-7-214.

Grubman MJ, Baxt B: Foot-and-mouth disease. *Clinical Microbiology Reviews* 2004, **17**(2):465–493.

Haydon DT, Samuel AR, Knowles NJ: The generation and persistence of genetic variation in foot-and-mouth disease virus. *Preventive Veterinary Medicine* 2001, **51**:111-124.

Haydon DT, Bastos AD, Knowles NJ, Samuel AR: Evidence for positive selection in foot-and-mouth disease virus capsid genes from field isolates. *Genetics* 2001, **157**:7–15.

Heath L, Walt Evd, Varsani A, Martin DP: Recombination patterns in aphthoviruses mirror those found in other picornaviruses. *Journal of Virology* 2006, **80**(23):11827–11832.

Ho SYW, Phillips MJ, Cooper A, Drummond AJ: Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution* 2005, **22**(7)_ 1561-156820.

Jackson T, King AMQ, Stuart DI, Fry E: Structure and receptor binding. *Virus Research* 2003, **91**:33-46.

Jackson AL, O'Neill H, Maree F, Blignaut B, Carrillo C, Rodriguez L, Haydon DT: Mosaic structure of foot-and-mouth disease virus genomes. *Journal of General Virology* 2007, **88**:487–492.

Kishino H, Hasegawa M: Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominodea. *Journal of Molecular Evolution* 1989, **29**:170-179.

Knowles NJ, Samuel AR: Molecular epidemiology of foot-and-mouth disease virus. *Virus Res* 2003, **91**:65-80.

Knowles NJ, Samuel AR: Polymerase chain reaction amplification and cycle sequencing of the 1D gene of foot-and-mouth disease viruses In: *Session of the research group of the standing technical committee of the European commission for the control of foot-and-mouth disease.* 19-22 September 1994, Vienna, Austria: FAO, Rome; 1995.

Lemey P, Rambaut A, Drummond AJ, Suchard MA: Bayesian phylogeny finds its roots. *PLoS Computational Biology* 2009, **5**(9): 1-16.

Martin DP, Williamson C, Posada D: RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics* 2005, **21**(2):260–262.

Martin D, Rybicki E: RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 2000, **16**(6):562-563.

Maynard Smith J: Analyzing the mosaic structure of genes. *Journal of Molecular Evolution* 1992, **34**:126-129.

Moya A, C.Holmes E, González-Candelas F: The population genetics and evolutionary epidemiology of RNA viruses. *Nature Reviews | Microbiology* 2004, **2**:279-288.

Nylander JAA: MrModeltest v2. *Program distributed by the author* Evolutionary Biology Centre, Uppsala University; 2004.

Padidam M, Sawyer S, Fauquet CM: Possible emergence of new geminiviruses by frequent recombination. *Virology* 1999, **265**:218-225.

Pond SLK, Frost SDW: A genetic algorithm approach to detecting lineage-specific variation in selection pressure. *Molecular Biology and Evolution* 2005, **22**(3):478–485.

Pond SLK, Posada D, Gravenor MB, Woelk CH, Frost SDW: Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology and Evolution* 2006, **23**(10):1891–1901.

Posada D, Crandall KA: Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proceedings of the National Academy of Sciences, USA* 2001, **98**:13757-13762.

Rodriguez F, Oliver JL, Marfn A, Medina JR: The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology* 1990, **142**:485-501.

Sahle M, Dwarka RM, Venter EH, Vosloo W: Comparison of SAT-1 foot-and-mouth disease virus isolates obtained from East Africa between 1971 and 2000 with viruses from the rest of sub-Saharan Africa. *Archives of Virology* 2007, **152**:797–804.

Salminen MO, Carr JK, Burke DS, McCutchan FE: Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Research and Human Retroviruses* 1995, **11**:1423-1425.

Sangula AK, Siegismund HR, Belsham GJ, Balinda SN, Masembe C, Muwanika VB: Low diversity of foot-and-mouth disease serotype C virus in Kenya: evidence for probable vaccine strain re-introductions in the field. *Epidemiology and Infection* 2010, doi:10.1017/S0950268810000580.

Schierup MH, Hein J: Consequences of recombination on traditional phylogenetic analysis. *Genetics* 2000, **156**:879–891.

Suchard MA, Weiss RE, Sinsheimer JS: Bayesian selection of continuous-time markov chain evolutionary models. *Molecular Biology and Evolution* 2001, **18**(16):1001–1013.

Swofford DL: PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version4. In., 4 edn. Sunderland, Massachusetts: Sinauer Associates; 2003.

Tamura K, Dudley J, Nei M, Kumar S: MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 2007, **24**(8):1596–1599.

Thomson GR, Vosloo W, Bastos ADS: Foot and mouth disease in wildlife. *Virus Research* 2003, **91**:145-161.

Tully DC, Fares MA: Unravelling selection shifts among foot-and-mouth disease virus (FMDV) serotypes. *Evolutionary Bioinformatics* 2006, **2**:211-225.

Tully DC, Fares MA: The tale of a modern animal plague: tracing the evolutionary history and determining the time-scale for foot and mouth disease virus. *Virology* 2008, **382**:250–256.

Tully DC, Fares MA: Shifts in the selection-drift balance drive the evolution and epidemiology of foot-and-mouth disease virus. *Journal of Virology* 2009, **83**(2):781–790.

Table 6-1 List of the SAT 1 virus isolates included in this study.

Isolate Code[a]	District/Country	Geographical group[b]	Accession No.
BEC/1/48	Bechuanaland (Botswana)	SA	AY593838
SWA/1/49	South West Africa (Namibia)	SA	AY593840
SA/13/61	South Africa	SA	AY593842

SWA/40/61	South West Africa (Namibia)	SA	AY593843
RHO/5/66	Rhodesia (Zimbabwe)	SA	AY593846
BOT/1/68	Botswana	SA	AY593845
UGA/21/70	Uganda	UG	WRL[b]
T155/71	Tanzania	KT	This study
SUD/8/74	Sudan	HA	AY441998
TAN/13/74	Tanzania	KT	AY442001
UGA/13/74	Uganda	UG	AY442010
NIG/11/75	Nigeria	WA	AF431711
NIG/14/75	Nigeria	WA	AF431709
NIG/17/75	Nigeria	WA	AF431712
NIG/24/75	Nigeria	WA	AF431714
NIG/1/76	Nigeria	WA	AF431721
NIG/5/76	Nigeria	WA	AF431723
NIG/14/76	Nigeria	WA	AF431725
NIG/20/76	Nigeria	WA	AF431727
NGR/2/76	Niger	WA	AF431718
NGR/5/76	Niger	WA	AF431720
SUD/3/76	Sudan	HA	AY441966
SUD/4/76	Sudan	HA	AY441997
TAN/2/77	Tanzania	KT	AY442008
NIG/2/79	Nigeria	WA	AF431728
K66/80	Narok, Kenya	KT	This study

NIG/3/80	Nigeria	WA	AF431729
TAN/3/80	Tanzania	KT	AY442006
K110/81	Kiambu, Kenya	KT	This study
NIG/5/81	Nigeria	WA	AF431730
NIG/10/81	Nigeria	WA	AF431731
KEN/9/91	Kenya	KT	AY441995
KEN/11/91	Kenya	KT	AY441994
KNP/196/91[c]	South Africa	SA	AF283429
TAN/5/96	Tanzania	KT	AY442007
TAN/19/96	Tanzania	KT	AY442013
UGA/1/97[c]	Uganda	UG	AF283439
KEN/4/98	Kenya	KT	AY441993
K58/99	Thika, Kenya	KT	This study
K96/99	Kajiado, Kenya	KT	This study
K114/99	Nairobi, Kenya	KT	This study
TAN/37/99	Tanzania	KT	AY442005
TAN/51/99	Tanzania	KT	AY442004
TAN/60/99	Tanzania	KT	AY442002
UGA/3/99	Uganda	UG	AY442009
UGA/7/99	Uganda	UG	AY442011
ZIM/23/03	Zimbabwe	SA	WRL[b]
K57/05	Thika, Kenya	KT	This study
K67/05	Trans Nzoia, Kenya	KT	This study

K10/06	Thika, Kenya	KT	This study
K21/06	Nyeri, Kenya	KT	This study
K28/06	Keiyo, Kenya	KT	This study
ETH/3/07	Ethiopia	HA	EJ798154

[a] The letters indicate the country of origin of the isolate while the last numerical field refers to the year of isolation. [b] Geographical group assignment according to the five defined regions of sub-Saharan Africa: SA (southern Africa), UG (Uganda), KT (Kenya/Tanzania), HA (Horn of Africa), WA (western Africa). [c] WRL = From World reference laboratory sequence data website (http://www.wrlfmd.org/fmd_genotyping/prototypes.htm). [c] African buffalo origin.

Table 2 Evidence for negative and positive selection using SLAC, FEL, and REL methods

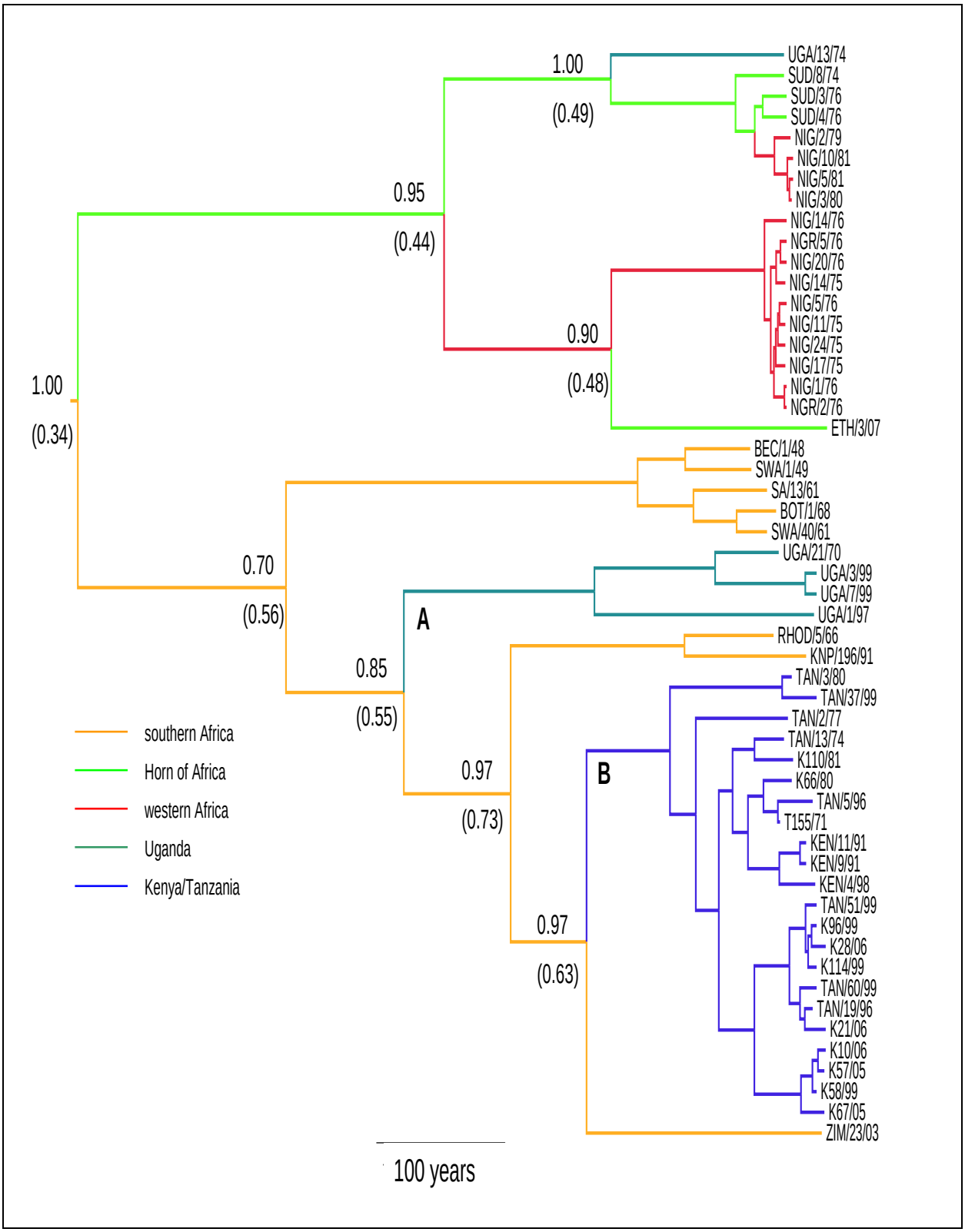
Number of sites	SLAC ($P < 0.1$)	FEL ($P < 0.1$)	REL (BF > 50)	Integrated
Positively selected	0	2	4	5
Negatively selected	143	152	112	153

A summary of the number of codon sites identified by the various methods to have been under selection at the default significance values for the three methods.

Figure legends

Figure 6-1 Bayesian time-scaled phylogeny of FMDV serotype SAT 1 with inferred geographical location states.

Maximum clade credibility tree of SAT 1 viruses based on complete VP1 coding sequences inferred using BEAST assuming a constant size coalescent prior showing lineage divergence since the most recent common ancestor. The inferred geographical location of each tree node is marked by a colour code defined in the insert legend. The posterior clade probability as well as the posterior geographical location state (in parenthesis) are shown for selected tree nodes. East African lineages A and B are marked.


1.00
(0.49)
UGA/13/74
SUD/8/74
SUD/3/76
SUD/4/76
NIG/2/79
NIG/10/81
NIG/5/81
NIG/3/80
0.95
(0.44)
NIG/14/76
NGR/5/76
NIG/20/76
NIG/14/75
NIG/5/76
NIG/11/75
NIG/24/75
NIG/17/75
NIG/1/76
NGR/2/76
0.90
(0.48)
ETH/3/07
1.00
(0.34)
BEC/1/48
SWA/1/49
SA/13/61
BOT/1/68
SWA/40/61
UGA/21/70
UGA/3/99
UGA/7/99
UGA/1/97
0.70
(0.56)
A
RHOD/5/66
KNP/196/91
0.85
(0.55)
TAN/3/80
TAN/37/99
TAN/2/77
TAN/13/74
K110/81
K66/80
TAN/5/96
T155/71
KEN/11/91
KEN/9/91
KEN/4/98
0.97
(0.73)
B
TAN/51/99
K96/99
K28/06
K114/99
TAN/60/99
TAN/19/96
K21/06
K10/06
K57/05
K58/99
K67/05
0.97
(0.63)
ZIM/23/03
southern Africa
Horn of Africa
western Africa
Uganda
Kenya/Tanzania
100 years

CHAPTER SEVEN

Co-circulation of two extremely divergent serotype SAT 2 lineages in Kenya highlights challenges to foot-and-mouth disease control

Co-circulation of two extremely divergent serotype SAT 2 lineages in Kenya highlights challenges to foot-and-mouth disease control

A.K. Sangula[a, d*], G.J. Belsham[b], V.B. Muwanika[a], R. Heller[c], S.N. Balinda[a], C. Masembe[a], H.R. Siegismund[c]

[a] Makerere University, Institute of Environment and Natural Resources, Molecular Biology Laboratory, P. O. Box 7298, Kampala, Uganda

[b] National Veterinary Institute, Technical University of Denmark, Lindholm, DK-4771 Kalvehave, Denmark

[c] Department of Biology, Ole Maaløes Vej 5, DK-2200, Copenhagen N, Denmark

[d] Foot-and-Mouth Disease Laboratory, Embakasi, P. O. Box 18021, 00500, Nairobi, Kenya

* Corresponding author:

Abraham K. Sangula

Makerere University Institute of Environment and Natural Resources,

Molecular Biology Laboratory,

P.O. Box 7298, Kampala, Uganda

Fax: +256 414 530134 Phone + 256 712629425

E-mail: aksangula@muienr.mak.ac.ug / aksangula@gmail.com

Abstract

Amongst the SAT serotypes of foot-and-mouth disease virus (FMDV) the SAT 2 serotype is the most widely distributed throughout sub-Saharan Africa. Kenyan serotype SAT 2 viruses have been reported to display the highest genetic diversity for the serotype globally. This complicates diagnosis and control and it is essential that patterns of virus circulation are known in order to overcome these difficulties. This study was undertaken to establish patterns of evolution of FMDV serotype SAT 2 in Kenya using complete VP1 coding sequences in a data set of 65 sequences from Africa collected over a period of 50 years. Two highly divergent lineages were observed to co-circulate and occasional trans-boundary spread was inferred emphasizing the value of constant monitoring and characterization of field strains for improved diagnosis and appropriate vaccine application as well as the need for regional approaches to control.

7.1 Introduction

Foot-and-mouth disease (FMD) is one of the most economically important infectious diseases of livestock. It is a vesicular disease of domesticated and wild cloven-hoofed animals. The agent of the disease, FMD Virus (FMDV), is a single stranded, positive-sense RNA virus in the genus *Aphthovirus,* family *Picornaviridae.* The genome of approximately 8400 nt encodes a single polyprotein, which is cleaved into four structural (VP1, VP2, VP3, VP4) and several non-structural proteins (Grubman and Baxt, 2004). FMDV exists as seven immunologically distinct serotypes known as O, A, C, Asia 1 and the Southern African Territories (SAT) 1, SAT 2 and SAT 3. The SAT serotypes are endemic to sub-Saharan Africa with the African buffalo (*Syncerus caffer*) reported to play an important role in their epidemiology in Southern Africa (Condy et al., 1985; Thomson et al., 2003). SAT 2 is the most widely distributed of these serotypes throughout sub-Saharan Africa with viruses in Kenya reported to be the most heterogeneous and having the highest genetic diversity (Bastos et al., 2003).

The history of confirmed SAT2 FMDV in Kenya dates back to 1956 with outbreaks due to this serotype becoming prominent from the late 1960s. It is currently, together with serotype O, the most prevalent serotype in Kenya (Ndiritu et al., 1983; Vosloo et al., 2002); records of the Embakasi FMD laboratory). Like type O, SAT 2 outbreaks have been distributed throughout Kenya with many antigenic variants observed, which has been reflected in the greatest change of vaccine strains used to control FMD outbreaks since vaccination was introduced in the 1950s (Ndiritu et al., 1983; records of the Embakasi FMD laboratory). The current FMDV SAT 2 vaccine strain is K52/84 and previous vaccine strains have included Ken 3/57, Tan 5/68, K183/74, R1215 (Anderson et al., 1982) and K65/82. The introduction of new vaccine strains has been undertaken whenever significant antigenic variants were observed among field isolates.

Serotype SAT 2 outbreaks in Kenya have mostly been reported in cattle and little is known about its circulation patterns in the country, although the African buffaloes have been reported to be carriers (Anderson et al., 1979) and serological evidence indicates high prevalence rates among buffalo populations (Bronsvoort et al., 2008). The high levels of genetic sequence diversity reported for SAT 2 in Kenya complicates diagnosis and vaccination since the development of suitable serotype specific primers to identify all field isolates and the selection of appropriate vaccine strains may be quite difficult (Bastos et al., 2003).

The VP1 coding region of FMDV has been extensively analysed to yield valuable molecular epidemiological information (Sobrino et al., 2001). This study was undertaken to establish patterns of evolution of serotype SAT 2 in Kenya using complete VP1 coding sequences from 36 Kenyan isolates in a data set of 65 sequences from Africa during the period 1948-2007. We specifically infer divergence times and evolutionary rates using genealogy-based coalescent methods.

7.2 Materials and Methods

7.2.1 Virus isolates

Thirty two (31 Kenyan and 1 Ugandan) SAT 2 virus isolates for this study, (collected between 1981 and 2007 from catttle) were obtained from the Embakasi FMD Laboratory Nairobi, Kenya. Virus was isolated from clinical material according to standard procedures using a single passage in baby hamster kidney (BHK) cells. The details of the isolates are shown in Table 7-1.

7.2.2 Viral RNA extraction, cDNA synthesis and amplification

Total RNA was extracted and cDNA synthesized as previously described in (Balinda et al., 2009). The complete VP1 coding region was amplified with a forward primer, in the VP3 coding sequence (5′- TTAACTACCACTTCATGTACAC(CG)G-3′) and the reverse primer FMD-2A$_{34}$

(Knowles and Samuel, 1995) using PCR reagent volumes and conditions as previously described in (Balinda et al., 2009) yielding a product of ~1050 bp. PCR products were visualized, purified and cycle-sequenced employing the same primers as for the PCR amplification and the other procedures were performed as previously described in (Balinda et al., 2009).

7.2.3 Sequence analysis

The generated sequences were initially assembled using the sequencher software 4.8 (Gene Code Corporation, USA). Multiple alignments of the whole data set were carried out using MUSCLE by log-expectation comparison incorporated within the software program Geneious version 4.6 (Drummond et al., 2009). Furthermore, to be able to put the phylogeny into a continental SAT 2 context, 33 other complete VP1 coding sequences representing topotypes defined in other studies (Bastos et al., 2003; Knowles and Samuel, 2003; Sangare et al., 2004) and available in the GenBank covering a sampling period from 1948-2007 were included in the study.

Models of evolution for the phylogenetic analysis were determined by the hierarchical likelihood ratio tests (LRT) using PAUP* version 4b10 (Swofford, 2003) and MrModeltest version 2.2 software (Nylander, 2004). The GTR (Rodriguez et al., 1990) model with gamma-distributed rates among sites and a proportion of invariable sites was used to co-estimate the phylogenetic relationships, evolutionary rates and demographic histories applying a Bayesian Markov Chain Monte Carlo (MCMC) method implemented in the BEAST software version 1.4.8 package (Drummond and Rambaut, 2007) (http://beast.bio.ed.ac.uk). Preliminary analysis involved testing different demographic models/coalescent priors and the appropriateness of a strict clock versus various versions of relaxed clocks available in BEAST (Suchard et al., 2001). The constant population size model with an uncorrelated exponential clock was used and the MCMC chain lengths established as sufficient by viewing the runs in Tracer software version 1.4

(http://tree.bio.ed.ac.uk/software/tracer/). Uncertainty in the data was reflected in the 95% high-probability density (HPD) intervals. The maximum clade credibility trees were obtained using Tree Annotator program in BEAST and visualized with FigTree version 1.1.2 software (http://tree.bio.ed.ac.uk/software/figtree/).

To infer amino acid variation in the VP1 coding region in the data set, the amino acid sequence alignment was obtained using MEGA version 4 software (Tamura et al., 2007). Evolutionary divergence and mean diversity were also estimated using the Maximum Composite Likelihood method in MEGA4.

7.3 Results

7.3.1 Phylogenetic relationships

Complete VP1 coding sequences of 32 (31 Kenyan and one Ugandan) FMDV serotype SAT 2 isolates were generated in this study. In total, 65 (36 Kenyan and 29 from the rest of Africa) VP1 sequences including those from published reports were used to infer phylogenetic relationships, divergence times and evolutionary rates. The maximum clade credibility tree is shown in Figure 7-1 with the posterior probabilities for significant branches indicated. Two main clades of Kenyan serotype SAT 2 viruses with more than 20% sequence divergence were identified with one clade (Lineage V in Bastos et al., (2003)) consisting of viruses from 1956 up to the 1990s and the other clade (Lineage I in Bastos et al., (2003)) consisting of viruses from the 1980s and onwards. The Kenyan viruses of the two clades both have a countrywide distribution. The Kenyan Lineage V viruses were related to viruses circulating in the horn of Africa, the Sudan, East Africa (Uganda) and central Africa (Bastos et al., 2003; Sangare et al., 2004). The viruses in lineage I (Figure 7-1) grouped into 2 distinct sub-clades labeled here as IA and IB. Sub-clade IA had exclusively Kenyan viruses from the mid 1980s to the present day, while sub-clade IB had

Kenyan viruses from the early 1980s to 2005 together with viruses circulating in the Horn of Africa, East and Central Africa (Bastos et al., 2003).

The estimated divergence time from the MRCA was estimated to have been about 340 years before present (ybp) (95% HPD: 170-615). The overall substitution rate was 2.42 x 10^{-3} substitutions/site/year (95% HPD: 1.75 x 10^{-3} - 3.12 x 10^{-3}) as showed by the coefficient of variation >0.9 obtained with the relaxed molecular clock. Since it has been suggested that recombination may be responsible for non molecular clock like evolution in data sets (Schierup and Hein, 2000a, b), we tested for evidence of recombination within the VP1 coding region and found none as determined by the GARD (Pond et al., 2006) method in the Datamonkey server as well the exploratory methods implemented in RDP 2 beta 0.8 software (Martin et al., 2005).

7.3.2 Distribution of mutations

Of the 648 nucleotide sites characterized for the whole data set, a total of 239 (37%) were invariant. At the amino acid level, 92 amino acids (43%) were completely conserved (Figure 7-2). Variation was observed in the known hypervariable (GH loop and the C-terminus) regions of VP1 although the RGD motif at amino acid residues 144-146 as well as the cysteine residue at the base of the GH loop (position 134) were completely conserved in agreement with previous reports (Bastos et al., 2003; Sangare et al., 2004). At the VP1/2A cleavage site, the amino acid sequence KQ/LC predominated although other sequences including KQ/LL were observed among the Kenyan isolates.

The net evolutionary divergence among the Kenyan isolates was 23% with the mean diversity highest for the viruses in clade K2 at 15% and lower in clade K1 at 6%.

7.4 Discussion

The phylogenetic analysis of serotype SAT 2 in this study identified two very divergent clades of viruses circulating within Kenya that appear to have diverged an estimated 350 years ago. Lineage V is comprised of the earliest reported SAT 2 viruses e.g. KEN/3/57 and KEN/11/60 and seems to have become extinct in the mid 1990s with the last representative isolate being K77/96. This lineage had a wide distribution as shown by representative isolates from districts in different provinces of the country, e.g. KEN/3/57, K32/92 and K5/94 (Rift Valley province); K65/82 and K14/91 (Central province) and K37/95 (Nairobi province). This countrywide distribution is perhaps a reflection of the extensive livestock movements across the country for trade and grazing (Ndiritu, 1984). These viruses were related to viruses circulating in the horn of Africa, e.g. ETH/2/07, the Sudan e.g. SUD/6/77, East Africa (Uganda e.g. UGA/19/98) and central Africa e.g. ZAI/1/82, indicative of cross border introductions and spread. Two previous vaccine strains, KEN/3/57 and K65/82, belonged to this clade and the apparent disappearance of this lineage may perhaps be a reflection of the success of vaccination in controlling these outbreaks. The inclusion of additional Kenyan sequences covering a wider temporal space has helped to better resolve the genetic relationships of viruses in this grouping identified previously as lineage V; topotypes H,I and J which were only represented in earlier studies by a limited number of isolates (Bastos et al., 2003).

Lineage I consists of two sub-clades i.e. IA and IBB (Figure 7-1). While sub-clade IA contains exclusively Kenyan viruses from the mid 1980s to the present day, IB comprises Kenyan viruses from the early 1980s to 2005 together with viruses circulating in the Horn of Africa (ETH/1/90), East Africa (U267/83) and southern Africa (MAL/3/75). Clade K2 appears to

be related to viruses in southern Africa, in agreement with previous findings suggesting introductions northwards into eastern Africa.

The widespread distribution of very divergent virus strains during the same period was demonstrated by the co-existence of the two lineages in one district, e.g. K5/94 (lineage V) and K37/94 (lineage I) found in Nakuru district in 1994. The identification of two very divergent clades circulating sympatrically for hundreds of years without going extinct or outcompeting each other is interesting and highlights the extreme genetic diversity of FMD virus lineages persisting today (Bastos et al., 2003). This diversity may be a reflection of the complex epidemiological activity (history) of FMD and its persistence across the livestock and wildlife interface. Serotype SAT 2 of FMDV has been reported to be prevalent among the African buffalo populations in the Kenyan wildlife sanctuaries where interaction with livestock is frequent (Anderson et al., 1979; Bronsvoort et al., 2008). A recent study in Uganda found serological evidence for a high prevalence of SAT 2 FMDV infection amongst buffalo in Ugandan National Parks (Ayebazibwe C and others, unpublished results). However, little is known so far about the role of this interaction in the pattern of viruses circulating in Kenya. Different rates of nucleotide substitution have been observed in SAT 2 virus groups recovered from cattle and wildlife which perhaps is reflected in the observed different epidemiological roles of these hosts (Vosloo et al., 1996; Bastos et al., 2000) The nucleotide substitution rates within the SAT2 VP1 coding sequences observed in this study were similar to those reported elsewhere (Tully and Fares, 2008) and are higher than those observed for the SAT 1 serotype in eastern Africa (unpublished results by the authors).

The observations in this study emphasize the need for constant characterization of field strains which is essential for efficient diagnosis and the selection of appropriate vaccine strains.

It is apparent from this study that the inclusion of additional genetically closely related virus isolates from within a country e.g. Kenya, in this study permits a higher level of resolution of the relationships and circulation patterns. Hence, because of the small number of isolates from the other East Africa countries used in this study, more comprehensive studies of trans-boundary spread may require collection of additional samples. Nevertheless, the complex circulation patterns and the trans-boundary spread across the region observed in this study reaffirm the need for regional efforts towards FMD control.

Acknowledgements

We sincerely thank the Director of Veterinary Services, Kenya, for providing the virus isolates used in the study. Dr Sabenzia Wekesa is thanked for the information on the isolates. Teresa Kenduiywo, William Birgen and Eugene Arinaitwe are appreciated for technical assistance. Part of this work was carried out by using the resources of the Computational Biology Service Unit from Cornell University which is partially funded by Microsoft Corporation. This work was supported by the Danish International Development Agency (DANIDA) under the Livestock-Wildlife Diseases in East Africa Project.

References

Anderson, E.C., Doughty, W.J., Anderson, J., Paling, R., 1979. The pathogenesis of foot-and-mouth disease in the African buffalo (*Syncerus Caffer*) and the role of this species in the epidemiology of the disease in Kenya. *Journal of Comparative Pathology* **89**, 511 - 519.

Anderson, E.C., Doughty, W.J., Spooner, P.R., 1982. Variation in the thermal stability of isolates of foot and mouth disease type SAT 2 and its significance in the selection of vaccine strains. *Journal of Comparative Pathology* **92**, 495-507.

Balinda, S.N., Belsham, G.J., C.Masembe, Sangula, A.K., Siegismund, H.R., Muwanika, V.B., 2009. Molecular characterization of SAT 2 foot-and-mouth disease virus from post-outbreak slaughtered animals: implications for disease control in Uganda. *Epidemiology and Infection* doi:10.1017/S0950268809991427.

Bastos, A.D.S., Boshoff, C.I., Keet, D., Bengis, R.G., Thomson, G.R., 2000. Natural transmission of foot-and-mouth disease virus between African buffalo (*Syncerus caffer*) and impala (*Aepyceros melampus*) in the Kruger National park, South Africa. *Epidemiology and Infection* **124**, 591-598.

Bastos, A.D.S., Haydon, D.T., Sangare´, O., Boshoff, C.I., Edrich, J.L., Thomson, G.R., 2003. The implications of virus diversity within the SAT 2 serotype for control of foot-and-mouth disease in sub-Saharan Africa. *Journal of General Virology* **84**, 1595–1606.

Bronsvoort, B.M.D.C., Parida, S., Handel, I., McFarland, S., Fleming, L., Hamblin, P., Kock, R., 2008. Serological survey for foot-and-mouth disease virus in wildlife in eastern Africa and estimation of test parameters of a nonstructural protein enzyme-linked immunosorbent assay for buffalo. *Clinical and Vaccine Immunology* **15**, 1003–1011.

Condy, J.B., Hedger, R.S., Hamblin, C., Barnett, I.T.R., 1985. The duration of foot-and-mouth disease virus carrier state in African buffalo (i) in the individual animal and (ii) in a free-living herd. *Comparative lmmunology, Microbiology and Infectious diseases* **8**, 259- 265.

Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., Wilson, A., 2009. *Geneious v4.6.* Available from http://www.geneious.com/.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**. Doi. 10.1186/1471-2148-7-214.

Grubman, M.J., Baxt, B., 2004. Foot-and-mouth disease. *Clinical Microbiology Reviews* **17**, 465–493.

Knowles, N.J., Samuel, A.R., 1995. Polymerase chain reaction amplification and cycle sequencing of the 1D gene of foot-and-mouth disease viruses. *Session of the research group of the standing technical committee of the European commission for the control of foot-and-mouth disease*. FAO, Rome, 19-22 September 1994, Vienna, Austria.

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

Martin, D.P., Williamson, C., Posada, D., 2005. RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics* **21**, 260–262.

Ndiritu, C.G., 1984. Foot and mouth disease virus antigenic variation and its implications on vaccines. *Kenya Veterinarian* **8**, 14-19.

Ndiritu, C.G., Ouldridge, E.J., Head, M., Rweyemamu, M.M., 1983. A serological evaluation of 1979-1982 Kenyan foot-and-mouth disease type SAT 2 viruses. *The Journal of Hygiene* **91**, 335-341.

Nylander, J.A.A., 2004. *MrModeltest v2*. Program distributed by the author Evolutionary Biology Centre, Uppsala University.

Pond, S.L.K., Posada, D., Gravenor, M.B., Woelk, C.H., Frost, S.D.W., 2006. Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology and Evolution* **23**, 1891–1901.

Rodriguez, F., Oliver, J.L., Marfn, A., Medina, J.R., 1990. The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology* **142**, 485-501.

Sangare, O., Bastos, A.D.S., Venter, E.H., Vosloo, W., 2004. A first molecular epidemiological study of SAT-2 type foot-and-mouth disease viruses in West Africa. *Epidemiology and Infection* **132**, 525-532.

Schierup, M.H., Hein, J., 2000a. Consequences of recombination on traditional phylogenetic analysis. *Genetics* **156**, 879–891.

Schierup, M.H., Hein, J., 2000b. Recombination and the molecular clock. *Molecular Biology and Evolution* **17**, 1578–1579.

Sobrino, F., Saiz, M., Jimenez-Clavero, M.A., Nunez, J.I., Rosas, M.F., Baranowski, E., Ley, V., 2001. Foot-and-mouth disease virus: a long known virus, but a current threat. *Veterinary Research* **32**, 1–30.

Suchard, M.A., Weiss, R.E., Sinsheimer, J.S., 2001. Bayesian selection of continuous-time markov chain evolutionary models. *Molecular Biology and Evolution* **18**, 1001–1013.

Swofford, D.L., 2003. *PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version4.* Sinauer Associates, Sunderland, Massachusetts.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596–1599.

Thomson, G.R., Vosloo, W., Bastos, A.D.S., 2003. Foot and mouth disease in wildlife. *Virus Research* **91**, 145-161.

Tully, D.C., Fares, M.A., 2008. The tale of a modern animal plague: tracing the evolutionary history and determining the time-scale for foot and mouth disease virus. *Virology* **382**, 250–256.

Vosloo, W., Bastos, A.D.S., Kirkbride, E., Esterhuysen, J.J., Van Rensburg, D.J., Bengis, R.G., Keet, D.W., Thomson, G.R., 1996. Persistent infection of African buffalo (Syncerus

caffer) with SAT-type foot-and-mouth disease viruses: rate of fixation of mutations, antigenic change and interspecies transmission. *Journal of General Virology* **77**, 1457-1467.

Vosloo, W., Bastos, A.D.S., Sangare, O., Hargreaves, S.K., Thomson, G.R., 2002. Review of the status and control of foot and mouth disease in sub-Saharan Africa. *OIE Scientific and Technical Review* **21**, 437-447.

Table 7-1 List of the SAT 2 virus isolates included in this study

Isolate Code	Date of Isolation	District/Country	Accession No.
RHO/1/48	1948	Rhodesia	AY593847
KEN/3/57[a]	1957	Samburu, Kenya	AJ251473
SA/106/59	1959	South Africa	AY593848
KEN/11/60	1960	Kenya	AY593849
MAL/3/75	1975	Malawi	AF367099
ANG/4/74	1974	Angola	AF479417
SEN/5/75	1975	Senegal	AF367140
UGA/51/75	1975	Uganda	AY343963
SUD/6/77	1977	Sudan	AY343939
MOZ/1/79	1979	Mozambique	AF367137
K81/81	1981	Laikipia, Kenya	This study
K46/82	1982	Trans Mara, Kenya	This study
K65/82[a]	1982	Kiambu, Kenya	This study
ZAI/1/82	1982	Zaire	AF367100
K70/83	1983	Kajiado, Kenya	This study
U267/83	1983	Uganda	This study
K151/83	1983	Nakuru, Kenya	This study
PAL/5/83	1983	South Africa	AF367102
ZIM/7/83	1983	Zimbabwe	AF136607
K34/84	1984	Nakuru, Kenya	This study
K52/84[b]	1984	Nakuru, Kenya	This study

KEN/1/84	1984	Uasin Gishu, Kenya	AY344505
KEN/2/84	1984	Nakuru, Kenya	AY343941
K37/86	1986	Kiambu, Kenya	This study
K13/87	1987	Kajiado, Kenya	This study
KNP/19/89	1989	South Africa	AF367110
ETH/1/90	1990	Ethiopia	AY343935
GHA/2/90	1990	Ghana	AF479415
K40/90	1990	Kiambu, Kenya	This study
K14/91	1991	Kirinyaga, Kenya	This study
BUN/1/91	1991	Burundi	AF367111
K32/92	1992	Kajiado, Kenya	This study
KNP/1/92	1992	South Africa	AF367114
KNP/132/92	1992	South Africa	AF367115
K3/93	1993	Kajiado, Kenya	This study
ZAM/10/93	1993	Zambia	AF367117
K5/94	1994	Nakuru, Kenya	This study
K25/94	1994	Kajiado, Kenya	This study
K37/94	1994	Nakuru, Kenya	This study
K37/95	1995	Nairobi, Kenya	This study
K39/95	1995	Kajiado, Kenya	This study
K77/96	1996	Nakuru, Kenya	This study
ZAM/10/96	1996	Zambia	AF367121
BOT/18/98	1998	Botswana	AF367123

ERI/12/98	1998	Eritrea	AF367126
NAM/286/98	1998	Namibia	AF367127
NAM/304/98	1998	Namibia	AF367129
UGA/19/98	1998	Uganda	AY343969
ZIM/267/98	1998	Zimbabwe	AF367130
K49/99	1999	Trans Mara, Kenya	This study
KEN/7/99	1999	Kiambu, Kenya	AF367132
RWA/1/00	2000	Rwanda	AF367134
ZIM/1/00	2000	Zimbabwe	AF367136
K13/02	2002	Kericho, Kenya	This study
K120/04	2004	Kiambu, Kenya	This study
K70/05	2005	Nairobi, Kenya	This study
K6/06	2006	Laikipia, Kenya	This study
K12/07	2007	Nyandarua, Kenya	This study
K15/07	2007	Narok, Kenya	This study
K17/07	2007	Kiambu, Kenya	This study
K20/07	2007	Kajiado, Kenya	This study
K42/07	2007	Kericho, Kenya	This study
K59/07	2007	Trans Mara, Kenya	This study
K67/07	2007	Nakuru, Kenya	This study
ETH/2/07	2007	Ethiopia	FJ798161

[a] Previous vaccine strain

[b] Current vaccine strain

Figure legends

Figure 7-1 Maximum clade credibility tree of SAT 2 viruses based on complete VP1 coding sequences inferred using BEAST assuming a constant size coalescent prior showing lineage divergence since the most recent common ancestor (MRCA). Time axis is shown in years and ranges from the MRCA to the present year. Branches with posterior probabilities ≥ 50% are labeled.

Figure 7-2 Sequence alignment of 214 amino acids of the VP1 coding region (plus the first 2 amino acids of 2A) of 23 SAT 2 FMD viruses. A `.´ indicates an amino acid residue identical to that of the sequence K139/81.

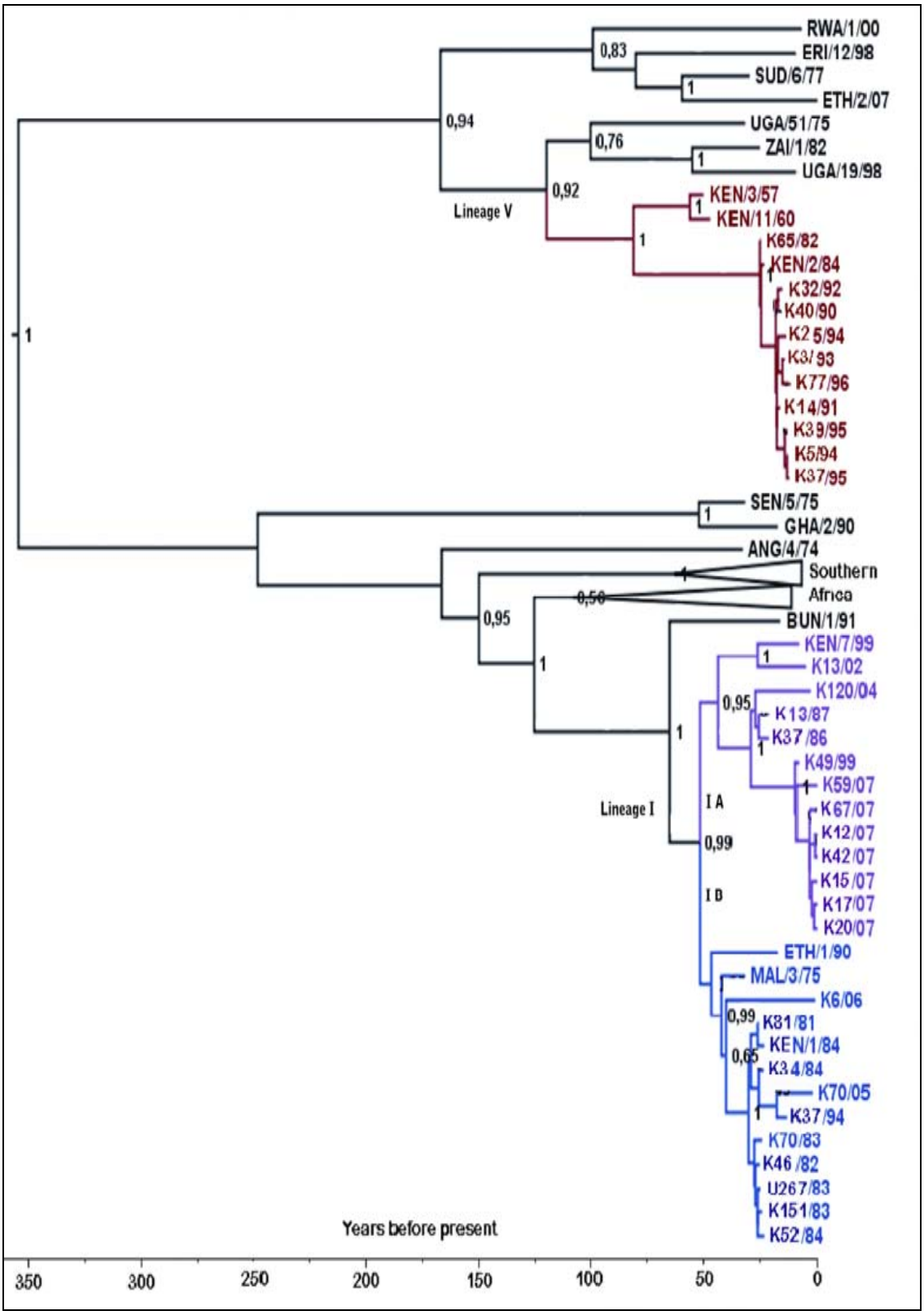

RWA/1/00
ERI/12/98
SUD/6/77
ETH/2/07
UGA/51/75
ZAI/1/82
UGA/19/98
KEN/3/57
KEN/11/60
K65/82
KEN/2/84
K32/92
K40/90
K25/94
K37/93
K77/96
K14/91
K39/95
K5/94
K37/95
SEN/5/75
GHA/2/90
ANG/4/74
Southern
Africa
BUN/1/91
KEN/7/99
K13/02
K120/04
K13/87
K37/86
K49/99
K59/07
K67/07
K12/07
K42/07
K15/07
K17/07
K20/07
ETH/1/90
MAL/3/75
K6/06
K81/81
KEN/1/84
K34/84
K70/05
K37/94
K70/83
K46/82
U267/83
K151/83
K52/84
Lineage V
Lineage I
I A
I D
0,94
0,83
0,76
0,92
0,95
0,50
0,99
0,65
Years before present
350
300
250
200
150
100
50
0

Figure 7-1

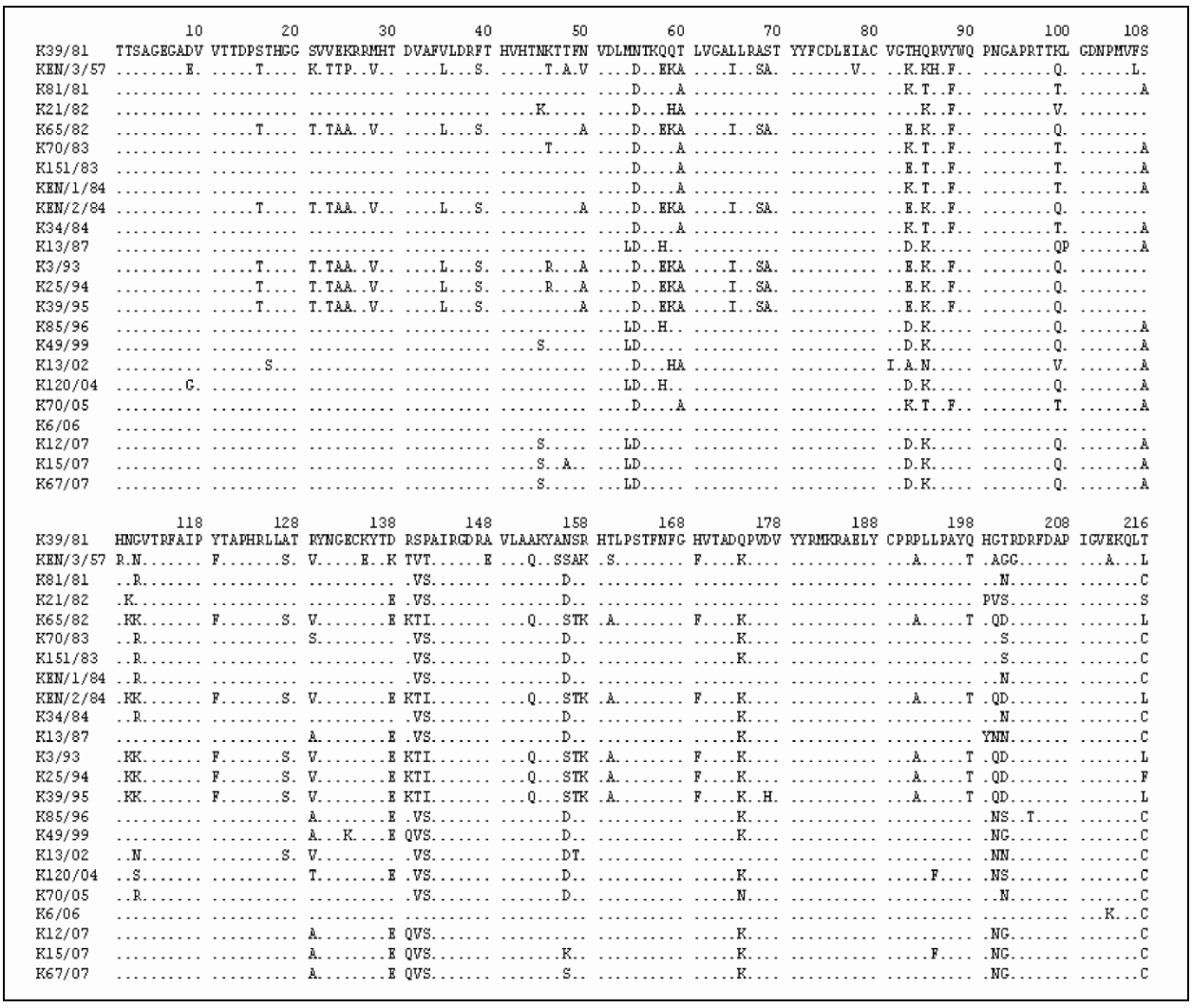


Figure 7-2

CHAPTER EIGHT

GENERAL DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

Results from this study show that, of the known FMDV serotype O topotypes within East Africa, only EA-1 and EA-2 were observed in Kenya and Uganda within the last 10 years. Of these, EA-2 was by far the most dominant. On the other hand, topotypes EA-3 and EA-4 were mainly found in Ethiopia, Eritrea and Sudan, although both topotypes appeared to have previously circulated in Kenya and Uganda; EA-3 was found in Kenya in the 1980s while an incursion of EA-4 into Uganda occurred in 1999. Only one virus sample from all the post-2000 samples included in this study was attributed to topotype EA-1. In contrast, the vaccine currently in use belongs to EA-1. Tests will be required to ascertain whether the vaccine based on topotype EA-1 protects against the other topotypes that are currently in circulation. Additionally, negative selection (purifying selective pressure) resulting in amino acid conservation was observed as being the dorminant evolutionary process. Nevertheless, positive selection on a few codon sites has been revealed in this study.

The phylogenetic analysis of serotype A showed that the isolates in the study belonged to 3 lineages (G-I, GIII, and GVIII). Only GI and GIII were observed in the recent past. GVIII was last observed in the 1960s. Lineage G-III was found in the other East African countries of Uganda, Horn of Africa, Sudan and West Africa. No positive selection on codon sites was observed. These results highlight the extent to which viruses circulate within the region through possibly animal movements and or trade. The existence of many circulating strains calls for constant search for relevant vaccine strains.

Results from FMDV serotype C analysis showed that Kenyan viruses were genetically very closely related. Very low nucleotide diversity (pi=0.0025) and remarkably little change (only five segregating sites and three amino-acid changes) over nearly forty years were observed. This is in contrast with what is known that FMD viruses evolve quickly. These results have been interpreted as being suggestive of re-introductions of the vaccine strain into the field. This has implications for maintenance of serotype C FMD virus in the field and the use of vaccination as a control measure in Kenya. Constant improvements in vaccine production processes is desirable to eliminate any possibilities of vaccine strain re-introductions

Phylogeographical analysis of FMDV serotype SAT 1 strains in this study showed that East Africa viruses grouped into 2 distinct lineages with one lineage found exclusively in Uganda, while the other in Tanzania and Kenya. High posterior probabilities of the location states supported the observed separate introductions of these lineages into eastern from southern Africa. . A Ugandan isolate from 1974 was found to belong to a lineage otherwise consisting of Sudanese, Ethiopian and western African strains. A mean nt substitution rate of 2.05×10^{-3} substitutions/site/year (s/s/yr) was observed. Distinct variation in rates among the eastern and West African viruses was observed with West Africa viruses having higher rates. This could imply that there are faster changes and greater diversity among the West African viruses.

Finally, results from analysis of serotype SAT 2 viruses showed that 2 very divergent lineages co-circulate in Kenya from the past to the present. The two lineages circulate in the wider East, Central, southern and horn of Africa. SAT 2 viruses like the type A demonstrated the widespread dispersal across the region and possible ineffectiveness of the vaccinations to eliminate or add to the diversity through evolutionary pressures.

The data sets generated in this study are the largest and the most geographically representative of FMDV serotypes assembled so far for Kenya and Uganda in single analyses. These results have all been accepted and / or submitted to international peer review journals of good standards. The phylogenetic relationships found here are consistent in some respects with those previously determined for serotypes O, A, C, SAT1 and SAT 2. However, the appropriateness of inferring any dynamic demographic model from a broad variety of FMDV topotypes need to be taken with caution given that they cannot be considered to constitute a single population, hence violating the assumptions of demographic inferences based on the coalescent theory. It is also possible that the sample composition plays an important role in dating divergences in this manner, as we have shown that rates of evolution differ substantially among branches in the phylogenies. Given the variability and inconsistencies in the estimated evolutionary rates and the divergence times between studies, the possibility of a radically changed time scale of historical divergence within FMDV should be addressed in future studies.

This work has analyzed the evolutionary relationships within serotype O, A, C, SAT 1 and SAT 2 of FMDV in Kenya and the wider eastern Africa. Despite the limited representation of virus sequences from some countries in the region, the phylogenetic relationships provided indications of transboundary events within the East African region. The occurrence of transboundary disease transmission was evident, and effective control strategies to combat this should be adopted by applying vaccines that are more broadly reflect the diversity of East African topotypes and lineages. This study revealed that the diversification of type O FMDV was most likely not an evolutionary response to the selection pressure afforded by vaccine programs and the dominance of some topotypes probably just reflects the lack of protection induced by the previous vaccine programs against these topotypes. This instills some confidence that vaccine

programs can be used to control the disease as long as the full diversity of the serotype is taken into account and vaccines are not exclusively based on single topotypes. Persistence of vaccine lineages and topotypes could be attributable to lack of comprehensive vaccine programs and to the survival of virus strains in untreated livestock reservoirs or maintenance in wildlife populations.

The study has also revealed an interesting observation on the continental SAT 1 dispersal in which Uganda may have played a more prominent role connecting southern African SAT 1s with those of the horn of Africa and western Africa. The sampling scheme employed in this study may have affected the outcome of phylogeographic analyses to some extent and we cannot exclude that additional samples from Uganda will show phylogenetic affinity with the surrounding countries. This could be easily tested by acquiring more Ugandan samples. More intense sampling (both temporally and spatially) could reveal novel dispersal patterns not observed here and further address the fine-scale historical movement of the serotype.

The substitution rates inferred in this study differ considerably from that reported in previous studies probably because a larger number of samples were included in this study. However despite the larger data set included in this study, sampling was far from representative in time and space. Further studies are therefore needed to establish the accurate rate of evolution in FMDV. Although the sampling scheme is likely to affect the rates, regionally variable evolutionary rates may in fact reflect real differences in the epidemiological dynamics and host-interaction of FMDV. For example, buffalos and other wildlife may play a more prominent role in the epidemiology of the SATs and this may give rise to changed patterns of evolution of virus lineages although considerable localized differentiation in evolutionary rates has not previously been observed in FMDV. Given these two (not necessarily mutually exclusive) causes of the

observed rate heterogeneity, it is vital that future studies address the caveats in using the VP1 gene sequence to infer evolutionary rates and history.

Overall Conclusions

1. The circulation patterns of FMDV serotypes O and A are diffuse within the east African region suggesting that countries in this region are actively exchanging viruses probably through livestock mobility across borders.
2. There was high genetic diversity observed in the FMDV serotypes A and SAT 2 in Kenya which may be the reason for the limited success of vaccination programs and emphasizes the importance of representative temporal and spatial characterization of virus isolates to be able to select appropriate vaccine strains.
3. There was unexpectedly low genetic variation and little sequence change over a long period of time within serotype C in Kenya suggesting possibly vaccine strain re-introductions in the field.
4. There is a high rate of virus strain turnover among serotypes A, O and SAT 2 FMD viruses in the East African region.
5. Serotype SAT 1 viruses in Uganda are not shared with the other East African countries of Kenya and Tanzania with implications on regional vaccine strains use.

Overall Recommendations

1. There is need for constant characterization of field strains so as to guide the choice of appropriate vaccines given the high rate of strain turnover among viruses observed in this study.

2. A regional strategy to control of FMD must be adopted by the East African countries given the observation from this study that there is active exchange of viruses among the East African countries.
3. Results of analysis of serotype C suggest a possible re-introduction of vaccine strain in the field and further studies to obtain definitive evidence for the possible linkage of the use of vaccines in the maintenance of serotype C FMD in Kenya is needed.
4. The use of two vaccines (K5/80 and K35/80) to control outbreaks due to serotype A is recommended since in this study two highly divergent lineages with cosmopolitan distribution exist in Kenya.
5. Antigenic comparisons of Ugandan SAT 1 field isolates with vaccine strain (T155/71) is necessary given the genetic distinctiveness of the Ugandan from the Kenyan and Tanzanian viruses.
6. There is need for antigenic comparison of current type O field strains with the vaccine strain (K77/78) given the observation in this study that they are genetically different.