Genetic diversity of serotype A foot-and-mouth disease viruses in Kenya from 1964-2013: implications for control strategies in eastern Africa

Sabenzia N. Wekesa[a, b,1], Abraham K. Sangula[a, 1], Graham J. Belsham[c], Vincent B. Muwanika[b], Rasmus Heller[d], Sheila N. Balinda[b], Charles Masembe[e], Hans R. Siegismund[d*]

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

[b]Department of Environmental Management, College of Agricultural and Environmental Sciences, Makerere University, P. O. Box 7062/7298 Kampala, Uganda

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

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

[e]Department of Biological Sciences, College of Natural Sciences, Makerere University, P. O. Box 7062 Kampala, Uganda

[1]Equal-joint authorship

***Correspondence**
Hans Redlef Siegismund
Department of Biology, University of Copenhagen
Ole Maaløes Vej 5, DK-2200, Copenhagen, Denmark
Tel.: +45 21 66 43 63
~~Fax: +45 35 32 13 39~~
E-mail: hsiegismund@bio.ku.dk

Abstract

Serotype A is the most genetically and antigenically diverse of the foot-and-mouth disease virus (FMDV) serotypes. Records of its occurrence records in Kenya date back to 1952 and the antigenic diversity of the outbreak viruses in this region is reflected by the current use of two different vaccine strains (K5/1980 and K35/1980) and previous use of two other strains (K18/66 and K179/71). This study aimed at enhancing the understanding of the patterns of genetic variation of serotype A FMDV in Kenya. The complete VP1 coding region sequences of 38 field isolates, identified as serotype A FMDV, collected between 1964 and 2013 were determined. Coalescent-based methods were used to infer times of divergence of the virus strains and the evolutionary rates alongside 27 other serotype A FMDV sequences from Genbank and the World Reference Laboratory (WRL). This study represents the first comprehensive genetic analysis of serotype A FMDVs from Kenya. The study detected four previously defined genotypes/clusters (termed G-I, G-III, G-VII and G-VIII), within the Africa topotype, together with a fifth lineage that has apparently emerged from G-I; these different lineages have each had a countrywide distribution. Genotypes G-III and G-VIII that were first isolated in 1964 are now apparently extinct, G-VII was last recorded in 2005, while G-I is currently still in widespread circulation. The high genetic diversity, widespread distribution and the transboundary spread of serotype A FMDVs across the region of eastern Africa was apparent. Continuous surveillance for the virus, coupled to genetic and antigenic characterization, ~~are~~is recommended for improved regional control strategies.

Key words; Africa, FMDV, genotype, Kenya, lineage, vaccination.

1. Introduction

Foot-and-mouth disease (FMD) is a highly contagious viral disease with well-known severe economic consequences that can affect both domestic and wild cloven-hoofed animals. FMD virus (FMDV), the causative agent, is an *Aphthovirus* within the family *Picornaviridae* whose genome is a single-stranded positive-sense RNA of about 8.4 kb. FMDV occurs in seven immunologically distinct serotypes namely O, A, C, Asia 1, SAT 1, SAT 2, and SAT 3. Serotype A is the most antigenically and genetically diverse of the FMDV serotypes making its control by vaccination difficult (Kitching, 2005). Three distinct topotypes of this serotype have been identified globally, which are named Africa, Asia and Europe-South America (Euro-SA) and are distinguished by >15% nucleotide divergence within the VP1 coding region sequences (Tosh et al., 2002a). These topotypes are subdivided into about 26 genotypes (Mohapatra et al., 2011), of which eight have been described within Africa (Knowles and Samuel, 2003; Mittal et al., 2005). Four of these African genotypes (G-I, G-III, G-VII and G-VIII) have been present in Kenya (WRLFMD, 2013).

FMD is endemic in Kenya and outbreaks date back to 1932~~,~~; however, proper documentation and characterization of viruses started in 1952 when serotypes O and A were first diagnosed (WRLFMD, 2013). Currently, serotypes O, A, SAT 1 and SAT 2 are in circulation in the country (Namatovu et al., 2013; WRLFMD, 2013), while serotype C was last reported in 2004, and Asia 1 has never occurred (Sangula et al., 2011). Records at the national FMD laboratory indicate that the highest frequency of serotype A FMDV outbreaks occurred in 1964 (74 out of a total of 162 (46%) positive samples), in 1965 (with 62/184 (36%)) and in 1985 (62/232 (27%)) which were diagnosed by the Complement Fixation Test. Since then, outbreaks caused by serotype A FMDV reduced to less than 5 positive samples per year between 2004 to 2009 and no serotype A virus positive cases were recorded in 2011 and in early 2012. However, there was an upsurge of cases in October 2012, when

12/17 (71%) of FMDV positive samples were diagnosed as serotype A by antigen ELISA. This upward trend persisted up to the close of the year and spilled over into 2013; some 16 out of 19 (84%) cases of FMD had been diagnosed as caused by serotype A virus by the end of February 2013; the highest percentage ever recorded. This trend necessitated characterization of the circulating strains of this serotype since control of the disease in Kenya involves vaccination using locally produced vaccines based on –two different FMDV strains (K5/1980 and K35/1980) in conjunction with animal movement restraint. For vaccination to be very effective, genetic and antigenic studies of circulating FMDVs is required to enable selection of the most appropriate vaccine strain (Bastos et al., 2003; Kitching, 2005), yet this has been seldom undertaken in Kenya.

Within eastern Africa, although country reports on circulating FMD serotypes are scarce, recent occurrences of serotype A have been reported in Tanzania in 2008-2009 by Kasanga et al. (2012) and in 2012 and 2013 by the World Reference Laboratory for FMD (WRLFMD, 2013) as well as in Eritrea (2006-2009) and the Democratic Republic of Congo (DRC) in 2011 (WRLFMD, 2013). Elsewhere, large numbers of serotype A outbreaks have reported, in recent years, in Asia and the Middle-East (Jamal et al., 2011a; Nandi et al., 2013).

The FMDV genome encodes a polyprotein that is processed to the four structural proteins (VP1-VP4) which comprise the virus capsid (only VP1-VP3 are surface exposed) plus multiple non-structural proteins required for virus replication and protein processing (Belsham, 2005). The VP1 coding region sequence has been widely used to yield this information as well as for inferring the epidemiological dynamics of the disease such as tracing the origins of outbreaks (e.g.Bastos et al., 2003; Cottam et al., 2008; Knowles and Samuel, 2003; Sangula et al., 2010a; 2010b; 2011; Valdazo-González et al., 2011). Hence, a large number of such sequences are available in databases. Consequently, it has been demonstrated that certain serotype A viruses, which have circulated 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). The VP1 protein also contains important antigenic sites and an arg-gly-asp (RGD) motif that binds to certain integrin receptors on the host cell surface and is part of a larger sequence (RGDLXXL) that is conserved for effective interaction of the virus with the host cell (Sobrino et al., 2001; Jackson et al., 2003).

The genetic diversities of other previously circulating serotypes in Kenya including serotypes O (Balinda et al., 2010a; Wekesa et al., 2013), C (Sangula et al., 2011), SAT 1 (Sangula et al., 2010a) and SAT 2 (Sangula et al., 2010b) have been extensively studied recently but not serotype A. Therefore, this study was undertaken to enhance the understanding of the patterns of genetic variation and serotype A FMDV in Kenya using field viruses that were isolated and archived at the national FMD laboratory, Embakasi, Kenya (FMDL, Embakasi).

2. Materials and Methods

2.1. Virus isolates

Thirty eight serotype A FMD virus isolates (collected between 1966 and 2013) were obtained from the FMDL, Embakasi, for this study. This laboratory is a repository of all FMDV samples collected in Kenya. The details of the isolates are shown in Table 1 and the geographic origins of the Kenyan isolates are indicated in Figure 1. Twenty seven other published complete VP1 coding sequences of serotype A FMDV selected from representative topotypes, genotypes and lineages from some countries in the eastern Africa region including five more from Kenya were also included in the analysis (Table 1).

2.2. RNA extraction, cDNA synthesis, reverse transcriptase–PCR and sequencing

Total RNA extraction was performed with the QIAamp® Viral RNA Mini Kit (Qiagen, Hamburg, Germany) according to the manufacturer's instructions and cDNA was synthesized using the Ready-to-go-You-prime first strand beads (GE Healthcare Life Sciences, Uppsala, Sweden) with random hexamer primers (pdN6). DNA amplification for samples collected before 2008, was achieved using the primer set A-1C_{562} or A-1C_{612} (forward) and FMD-2B_{58} (reverse) (Knowles and Samuel, 1995). The amplicons were purified using the QIAquick® PCR purification kit (Qiagen, Hamburg, Germany) and both strands were 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). In addition, DNA amplification for samples collected in 2008 and thereafter was achieved with forward primer 10-PPN5 (5'TACCAAATTACACACGGGAA3') and reverse primer 10PPN7 (5'GAAGGGCCCAGGGTTGGACTC 3') as described by Jamal et al., (2011a) encompassing the complete VP1 coding sequence. The PCR products (ca. 840bp) were purified using SigmaSpin® Sequencing Reaction Clean-Up Columns (Sigma-Aldrich, St. Louis, MO, USA) in accordance with the manufacturer's instructions and quantified using a NanoDrop® Spectrophotometer (Thermo Scientific, Wilmington, DE, USA). Cycle sequencing, in both directions, was performed using the Big dye Terminator V 3.1 kit (Applied Biosystems, Foster City, CA, USA) using 10-15ng of amplicon for each reaction and run on an automated DNA Sequencer (ABI PRISM 3700; Applied Biosystems, Foster City, CA, USA) using the same forward and reverse primers as for the PCRs.

2.3. Sequence assembly, alignment and initial analysis

The generated sequences were initially assembled using the Seqman pro software (Lasergene package, DNAstar, Inc., WI, USA). Serotypes of these sequences were confirmed and compared with

previous data using the Basic Local Alignment Search Tool (BLAST) at (www.ncbi.nlm.nih.gov (Reference: Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410 to be put in the References). Multiple alignment of the complete VP1 coding region sequences for the whole data set was generated using MUSCLE (Edgar, 2004) incorporated in MEGA5.2 software (Tamura et al., 2011). The best fitting model of nucleotide substitution was determined using MEGA and the Hasegawa-Kishino-Yano model with gamma distribution and invariable rates (HKY+G+I) was chosen as the best that describes the substitution pattern (Tamura et al., 2011). Estimates of overall, as well as within group, nucleotide divergence and the amino acid sequence alignment of the VP1 coding sequences were obtained using MEGA5.2.

Preliminary analysis and tests for evidence of recombination in the data set were performed using the Genetic Algorithm for Recombination Detection (GARD) method (Kosakovsky Pond et al., 2006) available on the Datamonkey server (http://www.datamonkey.org) (Kosakovsky Pond and Frost, 2005a). Furthermore, selection forces were analyzed using the single-likelihood ancestor counting (SLAC), fixed effects likelihood (FEL) and random effects likelihood (REL) methods available in Datamonkey (Kosakovsky Pond and Frost, 2005b).

2.4. Phylogenetic analysis and determination of evolutionary rates and selection forces

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 1.7.2 (Drummond and Rambaut, 2007) (http://beast.bio.ed.ac.uk). The sample dates were used to provide information about the rate of molecular change. The Bayesian skyline plot (BSP) model, a flexible demographic model allowing changing population sizes over

Comment [h1]: years??

time, and a relaxed (uncorrelated exponential) clock were used. The MCMC chains were made of 50 million runs achieving Effective Sample Size (ESS) of values > 200 for all model parameters with a 10% burn-in as viewed in Tracer software version 1.5 (http://tree.bio.ed.ac.uk/software/tracer/). Mean evolutionary rates were measured as the number of nucleotide substitutions per site per year (s/s/y). Uncertainties of the estimates were summarized as highest posterior density (HPD) intervals. Similarly, the tree model height parameter was used to estimate the age of the tree in years. A maximum clade credibility tree was obtained using the Tree Annotator program in BEAST and visualized with FigTree version 1.3.1 (http://tree.bio.ed.ac.uk/software/tracer/) (Fig. 2). In addition, estimation of the mean divergence times and the ages of each of the lineages and sub-lineages were read from the phylogenetic tree using the FigTree software.

Comment [h2]: Is this the real reference to FigTree? It's the same as for Tracer!

Nucleotide diversities within and among different groups in the data set of 65 VP1 coding sequences were conducted in MEGA5.2 (Tamura et al., 2011). Analyses were conducted using the P-distance approach.

Selection forces acting on the serotype A FMDV sequences were estimated using the whole data set of 65 viruses by Maximum Likelihood computations of *dN* and *dS* conducted using the HyPhy software package (Kosakovsky Pond et al., 2005) incorporated in MEGA5.2 through analysis of natural selection codon-by-codon under a Muse-Gaut model of codon substitution (Muse and Gaut, 1994) and Felsenstein (1981) model of nucleotide substitution. The probability of rejecting the null hypothesis of neutral evolution (*P*-value) was calculated (Kosakovsky Pond and Frost, 2005c; Suzuki and Gojobori, 1999) and *P*-values ≤ 0.05 were considered significant.

Comment [h3]: Unclear to me. F1981 has nothing to do with dN and dS???

3. Results

3.1. Sequence characteristics

The complete VP1 coding sequences of 38 different serotype A FMDV isolates from Kenya, which were collected from different outbreaks that occurred in different locations (Fig. 1) between 1966 and 2013, were determined (Table 1). These sequences were analyzed together with 27 other serotype A FMDV VP1 sequences from eastern Africa and the rest of the world that were obtained from Genbank. Of these 65 sequences (in total), 43 were of Kenyan origin. The nucleotide sequences of these Kenyan viruses were aligned, translated and analyzed in comparison to the current vaccine strain K5/1980 (Fig. 2). A total of 639 nucleotides (nt), which encode 213 amino acids, were aligned. Out of these 639 nt, a total of 293 (46%) sites were variable, which resulted in amino acid substitutions at 77 out of the 213 (36%) residues (Fig. 2). Most substitutions in the protein sequences were observed in the known hyper-variable regions of the VP1 (the G-H loop and the C-terminus). In contrast, the RGD motif within the RGDLXXL (residues 144-150) receptor binding sequence was highly conserved, except for the Lum/KEN/64 sequence, which had E substituted for D at residue 146, and the K44/2005 virus that had Q at residue 144. Among the 43 Kenyan viruses, variation from L147 to M147 was observed in multiple viruses; however the residue L150 was completely conserved. As expected, the most recent viruses from 2012 to 2013, indicated as G-Id in Fig. 2, had, on average, the highest number of modified sites (up to 36/213) (compared to the vaccine strain from 1980 that groups within the G-I genotype) while viruses collected between 1991 and 1995 had, on average, the least number of modifications (ca. 9/213). Notably, within the G-I group, the 2012-2013 strains (comprising the G-Id sub-lineage) were characterized by the changes D31S, I35V, M54L, Q99A, A110E, A156T and A192S while, in contrast, the G-VII group (last seen in 2005) shared these

Comment [h4]: This is still a repetition from M&M. Please cut down.

residues with the vaccine strain K5/1980 (Fig. 2). Other changes within the G-Id sub-lineage compared to the rest of the viruses within the G-I genotype and to G-VII are marked in Fig. 2.

3.2. Phylogenetic relationships and geographical distribution

The inferred maximum clade credibility tree for the entire data set~~, comprising 65 serotype A FMDV sequences,~~ is shown in Fig. 3 with the posterior probabilities indicated. The Kenyan serotype A FMD viruses all fell into one major distinct clade (within the Africa topotype) based on >15% nucleotide divergence cut-off between topotypes (Tosh et al., 2002a) and included the four distinct genotypes labelled as G-I, G-III, G-VII, G-VIII (Fig. 3) with several lineages and sub-lineages as described previously (Knowles and Samuel, 2003; Tosh et al., 2002a). The oldest of these Kenyan viruses (Ktl/KEN/64 and Lum/KEN/64), which were both isolated in 1964, belonged to G-VIII and G-III respectively, with no other viruses falling within these groups. The two current vaccine strains K5/1980 and K35/1980 belonged to G-I and G-VII respectively, while the most recently circulating viruses (K73/2008, K63/2009, K138/2012, K/143/2012, K148/2012, K154/2012 and K3/2013) all belonged to the G-Id sub-lineage within the G-I genotype (Fig. 2 and 3). The genotype G-I consisted of viruses exclusively of Kenyan origin while G-III and G-VIII had only one virus each from Kenya (Lum/KEN/64) and (Ktl/KEN/1964) respectively. In contrast, genotype G-VII included viruses from several countries in the eastern Africa region including Kenya, Ethiopia and Egypt. None of the Kenyan viruses isolated within the entire period belonged to genotype G-V, which includes the Nigerian strain (NGR/2/1973), or G-IV that includes viruses found in neighbouring Sudan (SUD3/1977, SUD 2/1984 and SUD3/2006). The G-I, G-III, G-VII and G-VIII genotypes were geographically distributed across 7 of the 8 administrative regions (provinces) within Kenya including Western, Rift Valley, Central, Nairobi, Eastern, North-Eastern and Coast (Table 1 and 2).

3.3. Mean evolutionary diversities and selection forces

The overall nucleotide diversity within the VP1 coding region for the entire data set of 65 viruses was 12.7%, while for the 43 Kenyan sequences it was 11.2% (data not shown). The nucleotide diversities within and distances between the different genotypes and sub-lineages of the Kenyan viruses are shown in Table 2. The distances between the 2003-2013 (G-Id) viruses and two groups G-Ib and G-VII to which the current vaccine strains K5/1980 (G-Ib) and K35/1980 (G-VII) belong were 14.4% and 13.3%, respectively (Table 2).

There was no evidence of recombination within any of the sequences in the data set (data not shown). The overall substitution rate for all the virus sequences analyzed here was 4.22×10^{-3} substitutions/site/year (s/s/yr) (95% HPD: $2.91 \times 10^{-3} – 5.51 \times 10^{-3}$). The estimated age of the tree was 130 years (Fig.4). The estimated divergence time of these 43 Kenyan viruses was about 90 years ago and the most recent common ancestor (TMRCA) is estimated to have emerged about 90 years before the present time, 2013 (ybp 2013) (95% HPD: 60-180). The oldest genotype, G-VIII, emerged ≈90 ybp followed by G-III (≈71 ybp), genotype G-I (≈66 ybp) and G-VII (≈52 ybp). The GId viruses, collected between 2003 and 2013, emerged ≈42 ybp, while the sub-lineages observed during 2008-2013 appeared ≈7 ybp and the most recent sub-lineages seen in 2012-2013 emerged ≈5ybp.

Comment [h5]: A little bit unclear. What is the dfference between divergence time and TMRCA?

Considering P values ≤ 0.05 as significant, test results for selection forces provided no evidence for positive selection ($dN > dS$ = -20.341; P = 1.00); but showed evidence of purifying (negative) selection ($dN < dS$ = 19.478; $P \approx 0.00$).

4. Discussion

This study presents the first comprehensive genetic analysis of serotype A FMD viruses in Kenya since the first recorded disease occurrence of this serotype in 1952. The phylogenetic analysis

identified the presence of four distinct genotypes that have circulated in Kenya during the period from 1964 until 2013 (Fig. 3); The presence of the G-I, G-III, G-VII and G-VIII genotypes in Kenya has been reported on the World Reference Laboratory for FMD website (WRL-FMD, 2013). A new genotype may be emerging from within the G-I genotype (labelled here as sub-lineage G-Id) that is between 11% - 14% nt different from other sub-lineages within this genotype.

According to the phylogenetic tree in this study (Fig. 3), genotype G-I viruses seem exclusive to Kenya suggesting a within country historical maintenance and evolution as shown by representative isolates from the 1960s (e.g. K18/66) to the present time (e.g. K3/2013). However, due to the limited number of recent sequences from the other countries of eastern Africa in this data set, it cannot be excluded that this genotype circulates outside of Kenya. Furthermore, recent unpublished data from (WRLFMD, 2013) indicates that viruses of the genotype G-I were first reported in DRC in 1953 and more recently in 2011, in Ethiopia in 1969 and in 1991, in Uganda in 2002, in Burundi in 1980, in Somalia in 1978, in Tanzania in 1967 and then again more recently in 2008, 2009, 2012 and 2013 (Kasanga et al., 2012; WRLFMD, 2013). Apart from eastern Africa, other occurrences of G-I, according to WRLFMD, have been detected in Malawi in 1981 and Zambia in 1980.

All eight Kenyan provinces except one (Nyanza) were reported in this study as having occurrences of serotype A FMDV outbreaks. This does not imply that Nyanza province has never had this serotype because records at the FMDL Embakasi have shown serotype A outbreaks in Nyanza in the past (K139/1971 (Migori district), K150/1971 (HomaBay district), K151/1971 (Siaya district), K155/1982 (Kisumu district) and K138/1991 (Nyamira district) but these viruses were not available for this study. This widespread distribution in the country is perhaps a reflection of the extensive livestock movements that occur across the country for trade and grazing.

Genotype G-VII consists of both older and fairly recent viruses that are also distributed both across the country and beyond the national borders; examples of this include K35/1980 (Eastern province) which is also a current vaccine strain; K60/2001 (Nairobi); K131/1979 (Central province); K4/1981 (Western province); K83/1985 (Coast province); K50/1981 (Rift Valley province) and some more recent isolates (K1973/2001 and K29/2005 both from Eastern province); together with viruses from the wider eastern Africa region, e.g. ETH/3/2005 (Ethiopia) and EGY/1/2006 (Egypt). The presence of this genotype could be indicative of introductions of serotype A FMDVs into Kenya from countries across the region at various times with resultant spread countrywide and other, subsequent, transboundary exchanges of viruses among countries in the region. Strain K35/1980, 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.

The two genotypes G-III and G-VIII are represented in the phylogenetic tree by two of the oldest viruses in the data set that were both isolated in 1964. No other viruses isolated after that period have been observed to belong to these genotypes suggesting that they may be extinct in the country. This observation, however, does not rule out their existence elsewhere in the region (as is also apparent from the lack of reports on G-I referred to above) and due to the extensive transboundary movement of animals and transmission of viruses as recently reported for serotype O FMDV by Balinda et al., (2010a) and Wekesa et al., (2013).

Genotype G-IV was not observed among the Kenyan isolates within the entire study period though it has been found elsewhere in the region such as in Eritrea in 2006-2009 (WRLFMD, 2013) and Sudan 2006-2011 (Habiela et al.,2010; WRLFMD,2013). A similar situation apparently occurs with serotype O FMDVs; the EA-3 topotype has been identified in Eritrea, Ethiopia and Sudan but has not been found in Kenya since 1999 (Wekesa et al., 2013). However, as observed by the latter

study, considering the livestock movement trends from these countries, including from Somalia into Kenya, as described by Di Nardo et al. (2011), it may be that genotype G-IV can easily move into Kenya if transboundary livestock movement control measures are inefficient.

The transboundary spread of serotype A FMDV strains across the region to different countries, as shown here for genotypes G-VII and G-I, reaffirms the need for regional efforts to improve FMD control. Livestock movement seems to be the major factor in the epidemiology of serotype A FMDV 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). Thus, improved regional approach to FMD control should include strategies on livestock movement control across the borders.

The Department of Veterinary Services (DVS) in Kenya has been controlling FMDV outbreaks by animal movement restraint and ring vaccination with monovalent vaccines. For serotype A outbreaks, vaccination has used either of the two vaccine strains depending on the geographical area of the outbreak and/or the antigenic characterization (vaccine matching) of field isolates using serological methods. The last time vaccine matching was performed on serotype A FMDV was in 2001 when AK5/1980 was shown to be protective against virus strains from all provinces except Eastern, while the A/K35/1980 vaccine only failed against viruses from the Rift valley and Coast provinces, thus the two vaccine strains apparently complemented each other well during that period. These have been appropriately incorporated and used as part of multivalent vaccines in selected districts of Kenya as part of the FMD control policy. Occasionally, they have been used for ring vaccinations befitting the prevailing virus prevalence status. Generally, vaccines based on AK5/1980 have been used in Central, Rift valley, Western and Nyanza provinces, while the AK35/1980 vaccines were employed in Eastern, Coast and some parts of Central province. However, this control

policy has not been routinely undertaken since the 1990s due to limited government resources occasioned by structural adjustment programs that had led to privatization of some government services. This may have resulted in the spread of strains to other regions traditionally not known to have them as is evident from the distribution of the serotype A lineages in Kenya reported in this study (Fig.1). Hence, the strategy of using particular vaccine strains for specific regions of the country may no longer be entirely appropriate since the different virus lineages are spread across the country.

Genetic diversities have been defined and measured by estimating genetic distances from a common ancestor (Korbinian and Haeseler, 2009; Tamura, 1992). In this study, nucleotide diversities and genetic distances between the recent 2003-2013 (G-Id) viruses within the G-I genotype and those that cluster with the current vaccine strain K5/1980 (G-Ib) were found to be approximately 14%. The sub-lineages labelled (G-Ia to G-Id are also evident from the alignment of the deduced VP1 amino acid sequences (Fig. 2). The amino acid substitutions D31S, I35V, M54L, Q99A, A110E, A156T and A192S are common to each of the 2012-2013 strains within the G-Id lineage, while on the contrary, the G-VII group had these residues generally conserved compared to the vaccine strain (K5/1980). This finding could probably explain the recent upsurge of this serotype despite the considerable vaccination efforts that are in place. Similarly, diversity of these recent viruses from the G-VII group that includes the vaccine strain K35/1980 of ≈13% in the same period indicates a move towards the emergence of a different lineage; hence the two vaccine strains may have become less effective with time. Two other type A vaccine strains (K18/1966 and K179/1971) have been previously used but in 1982, their use was discontinued after vaccine matching tests and the use of AK5/1980 and AK35/1980 commenced. These tests have not been consistently undertaken within the country due to limited resources, thus decisions on vaccine usage have mostly relied on intermittent results from the

WRLFMD. Further studies to ascertain the efficacy of these vaccines should be undertaken considering that it has now been shown that although the most recent viruses cluster within genotype G-1 with the vaccine strain K5/1980, the level of diversity is fairly high and may render the vaccine ineffective.

In Asia, including India, Pakistan, Afghanistan and Bangladesh, where type A FMDV is also endemic, this serotype (albeit from the Asia topotype) co-exists with serotypes O and Asia 1 but serotype A is the most divergent and is responsible for most of the recent FMD outbreaks (Jamal et al., 2011a; Jamal et al., 2011b; Mittal et al., 2005; Nandi et al., 2013). In contrast, in Kenya serotypes O, SAT 1 and SAT 2 have caused most of the outbreaks, until the recent 2012 to 2013 upsurge of serotype A outbreaks. From this study, it is possible that this change in prevalence of serotype A FMDV may be due, at least in part, to the high level of diversity from the vaccine strain among the current circulating virus lineages.

Comment [h6]: Do you mean "genetic distance"

Evolutionary forces such as selection and recombination are known to affect phylogenetic inferences (Schierup and Hein, 2000; Moya et al., 2004; Lewis-Rogers et al., 2008) and could bias evolutionary estimates. Recombination within some regions of the FMDV genome contributing to genetic heterogeneity has been reported in previous studies involving both the NSP coding regions (Balinda et al., 2010b; Lee et al., 2009 and Li et al., 2007; Mohapatra et al., 2008) and the capsid coding regions (Haydon et al., 2004; Tosh et al., 2002b). However, in the current study, there was no evidence for recombination between any of the VP1 coding sequences. This observation concurs with a previous report by Jackson et al., (2007) that recombination is known to be rare within the VP1 coding region and more frequent in the non-structural coding regions. Thus the genetic diversities and the selection pressures (purifying or negative selection) observed in this study may have been due to other factors including those related to viral replication and are not due to recombination.

The nucleotide substitution rate inferred here (4.22×10^{-3} s/s/yr) is similar to the average rate of nucleotide substitution in RNA viruses of $\approx 10^{-3}$ s/s/yr (Jenkins et al., 2002). However, it was much lower than that recently reported by Jamal et al. (2011a) for the A Iran05 lineage within the Asia topotype ($\approx 1.25 \times 10^{-2}$ s/s/yr) that has spread rapidly through the Middle East. Conversely, this rate is slightly higher than for SAT 2 recently reported in Kenya ($\approx 2.42 \times 10^{-3}$) by Sangula et al. (2010b) and serotype O in East Africa ($\approx 2.7 \times 10^{-3}$) by Balinda et al. (2010a). The high evolutionary rate resulting in the genetic heterogeneity of type A in Kenya is consistent with the use of several different vaccine strains by the DVS. However, in Kenya, this scenario is not unique to serotype A because, similarly, several vaccine virus strains have been used for other serotypes, e.g. six strains for SAT 2 and three for serotype O as reported by Sangula et al., (2011).

5. Conclusion

Multiple lineages and sub-lineages of the genotype G-I within the Africa topotype of serotype A FMDV are currently circulating in Kenya. The high nucleotide substitution rates and genetic diversity of this serotype in Kenya reported in this study indicate that the viruses are rapidly evolving and underscores the limited success of vaccination programs. Therefore continuous surveillance, genetic and antigenic characterization including vaccine matching of field isolates both on a local and regional scale are important and are recommended for discerning virus circulation patterns and the relevance of the vaccines that are currently in use. The importance of transboundary livestock mobility in the distribution of FMD viruses in eastern Africa is apparent from the circulation pattern of serotype A FMDV observed, thus emphasizing the need for regional control strategies for FMD control.

Acknowledgements

We thank the Director of Veterinary Services, Kenya, Dr. Peter Ithondeka for granting access to the virus isolates from FMDL Embakasi; Drs. Kenneth Ketter and Eunice Chepkwony for preservation, packaging and information on the Kenyan isolates; Teresa Kenduiywo, William Birgen, Eugene Arinaitwe, Preben Norman, Inge Nielsen, Jane Borch, Jane Christiansen for their technical assistance; and Dr Kirsten Tjørnehøj for her professional assistance. We are grateful to Margaret Mwangi and Teddy Ocheing, USAID Kenya, for assistance with the map. 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 (LWDEA) and the Trans-boundary Animal Diseases in East Africa (TADEA) project, DFC no. 10-006KU.

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. J. Comp. Pathol. 89, 511-519.

Balinda, S.N., Sangula, A.K., Heller, R., Muwanika, V.B., Belsham, G.J., Masembe, C., Siegismund, H.R., 2010a. Diversity and transboundary mobility of serotype O foot-and-mouth disease virus in East Africa: implications for vaccination policies. Infect. Genet. Evol. 10, 1058-1065.

Balinda, S.N., Siegismund, H.R., Muwanika, V.B., Sangula, A.K., Masembe, C., Ayebazibwe, C., Normann, P., Belsham, G.J., 2010b. Phylogenetic analyses of the polyprotein coding sequences of serotype O foot-and-mouth disease viruses in East Africa: evidence for interserotypic recombination. Virol. J. 7, 199.

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. J. Gen. Virol. 84, 1595-1606.

Belsham G.J., 2005. Translation and replication of FMDV RNA, in: Foot and Mouth Disease Virus. Mahy, B (Ed.), Springer-Verlag, Berlin, Heidelberg. Curr. Top. Microbiol. Immunol. 288, 43-70.

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. Clin. Vaccine. Immunol. 15, 1003-1011.

Cottam, E. M., Thebaud, G., Wadsworth, J., Gloster, J., Mansley, L., Paton, D. J., King, D. P., Haydon, D. T., 2008. Integrating genetic and epidemiological data to determine transmission pathways of foot-and-mouth disease virus. Proc. Biol. Sci. 275, 887-895.

Di Nardo A., Knowles N.J., Paton D.J., 2011. Combining livestock trade patterns with phylogenetics to help understand the spread of foot and mouth disease in sub-Saharan Africa, the Middle East and Southeast Asia. Rev. Sci. Tech. Off. Int. Epiz. 30, 63-85.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5, 113.

Felsenstein J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368-376.

Habiela M., Ferris N.P., Hutchings G.H., Wadsworth J., Reid S.M., Madi M., Ebert K., Sumption K.J., Knowles N.J., King D.P., Paton D.J., 2010. Molecular characterization of foot-and-mouth disease viruses collected from Sudan. Transbound. Emerg. Dis. 57, 305-314.

Haydon, D.T., Bastos, A. D., Awadalla, P., 2004. Low linkage disequilibrium indicative of recombination in foot-and-mouth disease virus gene sequence alignments. J. Gen. Virol. 85, 1095-1100.

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

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. J. Gen. Virol. 88, 487-492.

Jamal, S.M., Ferrari, G., Ahmed, S., Normann, P., Curry, S., Belsham, G.J., 2011a. Evolutionary analysis of serotype A foot-and-mouth disease viruses circulating in Pakistan and Afghanistan during 2002-2009. J. Gen. Virol. 92, 2849-2864.

Jamal, S. M., Ferrari, G., Ahmed, S., Normann, P., Belsham, G. J., 2011b. Molecular characterization of serotype Asia-1 foot-and-mouth disease viruses in Pakistan and Afghanistan; emergence of a new genetic group and evidence for a novel recombinant virus. Infect. Genet. Evol. 11, 2049-2062.

Jenkins, G. M., Rambaut A., Pybus, O.G., Holmes. E.C., 2002. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J. Mol. Evol. 54, 152-161.

Kasanga, C., Sallu, R., Mpelumbe-Ngeleja, C., Wadsworth, J., Ferris, N., Hutchings, G., Cleaveland, S., Lembo, T., Wambura, P., Yongolo, M., Knowles, N., King, D., Rweyemamu, M., 2012. Molecular epidemiology of foot-and-mouth disease virus in Tanzania GFRA scientific workshop, Hazyview, RSA: 17th April 2012. Available at; http://www.ars.usda.gov/GFRA/presentations/Session3/3.2%20Kasanga.GFRA%202012.Molecular%20epidemiology%20FMD%20Tanzania.pdf. Accessed 11th August 2013.

Kitching, R.P., 2005. Global epidemiology and prospects for control of foot-and-mouth disease, in: Foot and Mouth Disease Virus. Mahy, B (Ed.), Springer-Verlag, Berlin, Heidelberg. Curr. Top. Microbiol. Immunol. 288, 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. In: Open Session of Research Group of the Standing Technical Committee of the EU-FMD, Vienna, Austria, 1994, 19-22 September.

Knowles, N.J., Samuel, A.R., 2003. Molecular epidemiology of foot-and-mouth disease virus. Virus Res. 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. Emerg. Infect. Dis. 13, 1593-1596.

Korbinian, S., Haeseler, A., 2009. Genetic distances and nucleotide substitution models, in: Lemey, P., Salemi, M., Vandamme, A. (Eds.), The Phylogenetic Handbook: A practical Approach to Phylogenetic Analysis and Hypothesis Testing: Cambridge University Press, New York. pp. 111-113.

Kosakovsky Pond, S.L., Frost, S.D.W., 2005a. A genetic algorithm approach to detecting lineage-specific variation in selection pressure. Mol. Biol. Evol. 22, 478-485.

Kosakovsky Pond, S.L., Frost, S.D.W., 2005b. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21 (10), 2531-2533.

Kosakovsky Pond, S.L., Frost, S.D.W., 2005c. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol. Biol. Evol. 22, 1208-1222.

Kosakovsky Pond, S.L., Frost, S.D.W., Muse, S.V., 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21, 676-679.

Kosakovsky Pond, S.L., Posada, D., Gravenor, M.B., Woelk, C.H., Frost, S.D.W., 2006. Automated phylogenetic detection of recombination using a genetic algorithm. Mol. Biol. Evol. 23, 1891-1901.

Lee, K.N., Oem, J.K., Park, J.H., Kim, S.M., Lee, S.U., Tserendorj, S., Sodnomdarjaa, R., Joo, Y.S., Kim, H., 2009. Evidence of recombination in a new isolate of foot-and-mouth disease virus serotype Asia 1. Virus Res. 139, 117-121.

Li, D., Shang, Y.J., Liu, Z.X., Liu, X.T., Cai, X.P., 2007. Comparisons of the complete

genomes of two Chinese isolates of a recent foot-and-mouth disease type Asia1 virus. Arch. Virol. 152, 1699-1708.

Lewis-Rogers, N., McClellan, D.A., Crandall, K.A., 2008. The evolution of foot-and-mouthdisease virus: impacts of recombination and selection. Infect. Genet. Evol. 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. Arch. Virol. 150, 911-928.

Mohapatra, J.K., Sanyal, A., Hemadri, D., Tosh, C., Biswas, S., Knowles, N.J., Rasool, T.J., Bandyopashyay, S.K., Pattnaik, B., 2008. Comparative genomics of serotype Asia 1 foot-and-mouth disease virus isolates from India sampled over the last two decades. Virus Res. 136, 16-29.

Mohapatra, J. K., Subramaniam, S., Pandey, L. K., Pawar, S. S., De, A., Das, B., Sanyal, A., Pattnaik, B., 2011. Phylogenetic structure of serotype A foot-and-mouth disease virus: global diversity and the Indian perspective. J. Gen. Virol. 92, 873-879.

Moya, A., Holmes, E., González-Candelas, F., 2004. The population genetics and evolutionary epidemiology of RNA viruses. Nat. Rev. Microbiol. 2, 279-288.

Muse, S.V., Gaut, B.S., 1994. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Mol. Biol. Evol. 11, 715-724.

Nandi, S.P., Rahman, M.Z., Momtaz, S., Sultana, M., Hossain, M.A., 2013. Emergence and distribution of foot-and-mouth disease virus serotype A and O in Bangladesh. Transbound. Emerg. Dis. doi: 10.1111/tbed.12113. In Press.

Namatovu A., Wekesa S.N., Tjørnehøj K., Dhikusooka M.T., Muwanika V.B., Siegsmund H.R., Ayebazibwe C., 2013. Laboratory capacity for diagnosis of foot-and-mouth disease in Eastern Africa: implications for the progressive control pathway. BMC Vet. Res. 9, 19.

Sangula, A. K., Belsham, G. J., Muwanika, V. B., Heller, R., Balinda, S. N., Masembe, C., Siegismund, H. R., 2010a. Evolutionary analysis of foot-and-mouth disease virus serotype SAT 1 isolates from East Africa suggests two independent introductions from southern Africa. BMC Evol. Biol. 10, 371.

Sangula, A. K., Belsham, G. J., Muwanika, V. B., Heller, R., Balinda, S. N., Siegismund, H. R., 2010b. Co-circulation of two extremely divergent serotype SAT 2 lineages in Kenya highlights challenges to foot-and-mouth disease control. Arch. Virol. 155, 1625-1630.

Sangula, A.K., Siegismund, H.R., Belsham, G.J., Balinda, S.N., Masembe, C., Muwanika, V.B., 2011. Low diversity of foot-and-mouth disease serotype C virus in Kenya: evidence for probable vaccine strain re-introductions in the field. Epidemiol. Infect. 139, 189-96.

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. Vet. Res. 32, 1-30.

Suzuki Y., Gojobori T., 1999. A method for detecting positive selection at single amino acid sites. Mol. Biol. Evol. 16,1315-28.

Tamura K., 1992. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C-content biases. Mol. Biol. Evol. 9, 678-687.

Tamura K. P.D., Peterson N., Stecher G., Nei M., Kumar S., (2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731-2739.

Tosh, C., Sanyal, A., Hemadri, D., Venkataramanan, R., 2002a. Phylogenetic analysis of serotype A foot-and-mouth disease virus isolated in India between 1977 and 2000. Arch. Virol. 147, 493-513.

Tosh, C., Hemadri, D., Sanyal, A., 2002b. Evidence of recombination in the capsid-coding region of type A foot-and-mouth disease virus. J. Gen. Virol. 83, 2455-2460.

Valdazo-González, B., Knowles, N. J., Wadsworth, J., King, D. P., Hammond, J. M., Ozyoruk, F., Firat-Sarac, M., Parlak, U., Polyhronova, L., Georgiev, G. K., 2011. Foot-and-mouth disease in Bulgaria. Vet. Rec. 168, 247.

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. Rev. Sci. Tech. Off. Int. Epiz. 21, 437-447.

Wekesa, S.N., Muwanika, V.B., Siegismund, H.R., Sangula A. K., Namatovu, A., Dhikusooka, M.T., Tjørnehøj, K., Balinda S. N., Wadsworth, J., Knowles, N. J., Belsham, G.J., 2013. Analysis of recent serotype O foot-and-mouth disease viruses from livestock in Kenya: evidence of four independently evolving lineages. Transbound. Emerg. Dis. doi: 10.1111/tbed.12152. In Press.

WRLFMD., 2013. ME reports: 2013. World Reference Laboratory for Foot-and-Mouth Disease. Genotyping report for Africa. Available at; http://www.wrlfmd.org/fmd_genotyping/africa.html. Accessed 10th August 2013.

Figure legends

Figure 1

Map of Kenya illustrating the geographical origins of serotype A FMDV isolates included in this study demonstrating the distribution of the four different genotypes. The encircled numbers correspond to the serial numbers in Table 1.

Figure 2

Alignment of the deduced VP1 amino acid sequences of 43 serotype A FMDVs from Kenya isolated between 1964 and 2013. Genotypes G-I, G-III, G-VII, G-VIII and sub-lineages G-Ia, G-Ib, G-Ic, and G-Id are labelled. Single asterisks (*) indicate previous vaccine strains and double asterisks (**) depict current vaccine strains. Dots represent an amino acid residue identical to that of the top sequence K5/1980 (one of the current vaccine strains). The large rectangle marks the integrin receptor-binding motif (RGDLXXL), while the smaller rectangles indicate amino acid substitutions observed specifically within the G-Id lineage.

Figure 3

Maximum clade credibility tree based on nucleotide sequences of the complete VP1 coding region of the 65 serotype A viruses in this study, inferred using BEAST, showing lineage divergence since the most recent common ancestor (TMRCA) and the estimated age of each virus strain. Time axis is shown in years before 2013. Branches are labelled with posterior probabilities. The three global topotypes Africa, Asia and Euro-SA are labelled. The four genotypes (G-I, G-III, G-VII and G-VIII) that have been observed within Kenya are encircled on their respective clades, while two other African genotypes (G-IV) and (G-V) from eastern Africa and western Africa are also indicated. Four G-I sub-lineages (G-Ia to G-Id) are shown. The older and current vaccine strains are marked with black dots (●) and crosses (+) respectively.