
Article

Phylogenetic characterisation of Crimean-Congo haemorrhagic fever virus detected in African blue ticks feeding on cattle in a Ugandan abattoir.

Eddie M Wampande¹, Peter Waiswa¹, David J Allen², Roger Hewson^{2,3}, Simon DW Frost^{4,5} and Samuel CB Stubbs^{4,6,*}

¹ Makerere University, College of Veterinary Medicine, Animal Resources and Biosecurity, Kampala, Uganda

² London School of Hygiene and Tropical Medicine, Department of Infection Biology, London, UK

³ Virology and Pathogenesis, Public Health England, Porton Down, Salisbury, UK

⁴ London School of Hygiene and Tropical Medicine, Department of Infectious Disease Epidemiology, London, UK

⁵ Microsoft Research, Redmond, Washington, USA

⁶ University of Cambridge, Department of Veterinary Medicine, Cambridge, UK

* Correspondence: sam.stubbs@lshtm.ac.uk

Abstract: Crimean-Congo haemorrhagic fever virus (CCHFV) is the most geographically widespread tick-borne virus. However, African strains are poorly represented in sequence databases. In addition, almost all CCHFV sequence data have been obtained from cases of human disease, while information regarding circulation of the virus in tick and animal reservoirs is severely lacking. Here, we characterise the complete coding region of a novel CCHFV strain, detected in African blue ticks (*Rhipicephalus (Boophilus) decoloratus*) feeding on cattle in an abattoir in Kampala, Uganda. These cattle originated from a farm in Mbarara, a major cattle-trading hub for much of Uganda. Phylogenetic analysis indicates that the newly sequenced strain belongs to the African genotype II clade, which predominantly contains the sequences of strains isolated from West Africa in the 1950's and South Africa in the 1980's. Whilst, the viral S (nucleoprotein) and L (RNA polymerase) genome segments shared >90% nucleotide similarity with previously reported genotype II strains, the glycoprotein-coding M segment shared only 80% nucleotide similarity with the next most closely related strains, which were from India and China. This segment also displayed a large number of non-synonymous mutations previously unreported in genotype II strains. Characterisation of this novel strain adds to our limited understanding of the natural diversity of CCHFV circulating in both ticks and in Africa. Such data can be used to inform the design of vaccines and diagnostics, as well as studies exploring the epidemiology and evolution of the virus for the establishment of future CCHFV control strategies.

Keywords: Crimean-Congo hemorrhagic fever virus; hemorrhagic fever; viral genomics; Rhipicephalus (Boophilus) decoloratus; Rhipicephalus; ticks

1. Introduction

Crimean-Congo haemorrhagic fever virus (CCHFV) is a zoonotic pathogen primarily transmitted between vertebrate hosts through the bite of infected ticks, as well as through direct contact with infected livestock, ticks and bodily fluids [1]. CCHFV is the most geographically widespread of the tick-borne viruses; endemic throughout Africa, Asia, Eastern Europe and the Middle East, it circulates in an enzootic cycle between animals and ticks, sporadically emerging to cause severe outbreaks of human disease [2]. Outbreak fatality rates up to 30% have been reported, although CCHF symptoms can

range in severity from mild febrile illness, to haemorrhagic fever and multi-organ failure [2].

CCHFV is classified within the genus *Orthonairovirus* (family *Nairoviridae*) and carries a tripartite RNA genome of negative polarity. The three genome segments are labelled small (S), medium (M) and large (L), according to their respective lengths [3]. The ~1.7 Kb S segment encodes the viral nucleoprotein (NP) and, in the reverse orientation, an overlapping non-structural (NSs) protein. The M segment is approximately 5.3 Kb and encodes a single polyprotein precursor, which is post-translationally cleaved into two envelope glycoproteins (Gn and Gc), several secreted glycoproteins (GP38, GP85 and GP160) and the non-structural M protein, NSm. The pre-Gn precursor also contains a hyper-variable region known as the mucin-like domain (MLD). The ~12.1 Kb L segment encodes the RNA-dependent RNA polymerase.

Strains of CCHFV display a high degree of genetic variability for an arthropod-borne virus, with nucleotide diversity as high as 20% for the S segment and 31% for the M segment [4]. Genotypes are assigned to each genomic segment independently based on phylogenetic analysis, with each genotype largely defined by a distinct geographical region: three are African (genotypes I, II and III), two Asian (genotypes IVa and IVb), and two European (genotypes V and VI). However, limited sampling and uneven representation of CCHFV sequences has hindered phylogeographic and evolutionary analyses [2, 5]. This is particularly true for African strains, which are poorly represented in sequence databases in comparison to strains from Europe, Asia and the Middle East. Furthermore, the majority of CCHFV strains reported to date have been derived from cases of severe human disease, while insufficient sampling from tick vectors and reservoir hosts is almost certainly limiting our knowledge of the natural diversity of CCHFV and the potential for spillover of unanticipated viral variants.

Recent detection of an African genotype III strain following an outbreak in Spain [6], and the presence of African genotype II M segments in viruses obtained from pools of ticks in India and China [7, 8] has highlighted the need for a more complete genomic dataset, in order to better understand the epidemiology of the virus, including viral transmission routes and the extent of re-assortment among genotypes. The potential for such re-assortment is of importance, as many molecular tests for CCHFV are only capable of detecting a restricted number of closely related genotypes [9]. It also remains to be seen whether recently reported vaccine candidates, including one based on the S and M segments of a Hoti strain (genotype V) [10] and another based on the M segment of IbAr102000 strain (genotype III) [11], are capable of eliciting a protective immune response against viruses belonging to other more distantly related genotypes.

2. Materials and Methods

Ethical approval for this study (NS 673) was granted by the Uganda National Council for Science and Technology Review Board on 31/07/2019. Seventy ticks were collected from cattle, goats and a pig in an abattoir in Kampala, Uganda between September - December 2019 (Table S1) as part of a study into the virome of ticks in Uganda. Ticks were immediately frozen in liquid nitrogen following collection and stored at -80°C until ready for processing. Individual ticks were morphologically identified using taxonomic keys [12] before being crushed using a sterile micropestle and resuspended in 250 µL of phosphate buffered saline (PBS). Solid debris was pelleted by low speed centrifugation (400 × g) and total nucleic acid (TNA) was extracted from 200 µL of the supernatant using the High Pure viral nucleic acid extraction kit (Roche) following the manu-

facturer's instructions. The extracted nucleic acid was then lyophilized before shipping at ambient temperature to the University of Cambridge, UK for further analysis.

Upon arrival in the UK, TNA was reconstituted in 20 μ L of RNase-free water (RFW) supplemented with 2 U of RNaseOUT (Thermo Fisher Scientific). Two μ L of reconstituted TNA was subjected to reverse transcription in 20 μ L reactions using Superscript III reverse transcriptase (Thermo Fisher Scientific) and priming with 50 ng random hexamers (Thermo Fisher Scientific), as per the enzyme manufacturer's instructions. Two μ L from each reverse transcription reaction was used as direct input to screen for the presence of Nairovirus RNA by (RT-)PCR. PCR was performed in 20 μ L reactions, using Phusion DNA polymerase (New England Biolabs) and 0.5 μ M pan-nairovirus primers targeting the S segment [13], set up according to the enzyme manufacturer's instructions. PCR cycling conditions were as follows: incubation at 98 $^{\circ}$ C for 2 minutes, followed by 45 cycles of 98 $^{\circ}$ C for 5 secs, 60 $^{\circ}$ C for 20 secs and 72 $^{\circ}$ C for 20 secs. A final extension was performed at 72 $^{\circ}$ C for a further 5 mins, followed by incubation at 4 $^{\circ}$ C. The resulting PCR products were visualised by gel electrophoresis and amplicons of the expected size (~400 bp) were individually purified and confirmed to be CCHFV by Sanger dideoxy sequencing.

Five RT-PCR positive ticks from Mbarara were selected at random for further characterisation by Illumina metagenomic sequencing. Five μ L of TNA from each tick was pooled for a total volume of 25 μ L before being treated with 2 U of TURBO DNase (Ambion), incubating for 20 mins at 37 $^{\circ}$ C. DNase-treated RNA was purified using RNA clean & concentrator-5 spin columns (Zymo Research), eluted in 15 μ L of RFW and quantified using the Qubit High Sensitivity RNA assay (Invitrogen). Illumina sequencing libraries were prepared from 200 ng of purified RNA using the Zymo-Seq RiboFree Total RNA Library kit (Zymo Research), which incorporates a ribosomal RNA depletion step. The resulting library was sequenced on the Illumina NextSeq 500 platform, using the 300-cycle, mid-output kit (v2.5). Resulting paired-end reads were imported into CLC Genomics Workbench v7.5.1 (Qiagen) and sequences sharing 95% sequence identity over 35% of their length were assembled *de novo*. Contiguous sequences corresponding to the CCHFV S, M and L segments were identified using BLASTn against the redundant nucleotide sequence database with default settings [14].

The CCHFV genome segments (S, M, and L) were independently characterised through phylogenetic analysis. MAFFT v7.427 [15] was used to align the open reading frame of each genome segment against a representative set of CCHFV sequences from GenBank, including members of each genotype (Africa, Asia & European etc.) and all available sequences belonging to the genotype to which Mbarara strain was found to belong (genotype II). For the S and L segments, complete ORF sequences were analysed. For the M segment, an 857 bp region at the 5' end, corresponding to the MLD, was excluded from phylogenetic analysis due to the extreme variability of the region, which produced sub-optimal alignments. Inclusion of this region did not significantly alter phylogenetic clustering, however bootstrap values were adversely affected. Following trimming of the M segment alignment, the final datasets were 1,449, 4,323 and 11,838 nt long and included 50, 56 and 46 sequences for S, M and L segments, respectively.

Maximum likelihood phylogenies were calculated using IQTree v1.6.11 [16], employing 1,000 iterations of the ultrafast bootstrap estimation [17]. Optimal substitution models for each alignment were selected using the ModelFinder algorithm [18] according to Bayesian information criteria (GTR+F+I+G4 for S and L segments and GTR+F+R3 for M segment). Nucleotide identity matrices were generated for each segment using CLUSTAL Omega [19] and intra-genotype nucleotide sequence similarity plots were generated using the EMBOSS v6.6.0 [20] Plotcon function across a window size $l = 4$. Mid-point

rooted, maximum likelihood phylogenies and sequence similarity plots were visualised in R v3.6.3 [21] using the ggplot2 [22] and ggtree [23] packages. Genotype II M segment sequences were also analysed for evidence of recombination using the RDP5 tool [24] and for evidence of episodic diversifying selection using the adaptive branch-site random effects likelihood (aBRSSEL) method [25]. Nucleotide alignments of genotype II sequences were also translated into predicted amino acid alignments in Seaview v4.7 [26] and polymorphic and informative sites were identified using the ape v5.0 package [27] in R v3.6.3.

3. Results

A total of 70 ticks were collected from cattle (n = 48), goats (n = 21) and a pig (n = 1) and morphologically identified as *Rhipicephalus appendiculatus* (n = 16), *Rhipicephalus (Boophilus) decoloratus* (n = 40) and *Ambylomma variegatum* (n = 14) (Table S1). Nairovirus RNA was detected by RT-PCR in 8 of the 70 ticks (11.4 %). These 8 ticks were all identified as the African blue tick *Rhipicephalus (Boophilus) decoloratus* and all were picked from cattle. Seven of the CCHFV positive ticks were sampled from cattle originating from a farm in Mbarara, western Uganda, and 1 from a farm in Nakasongola in central Uganda, as determined based on cattle movement permits. Sanger dideoxy sequencing of the RT-PCR products confirmed the presence of CCHFV and revealed all 8 400 bp fragments of the S segment to be identical.

Metagenomic sequencing was performed on a pool of 5 randomly selected (RT-)PCR positive *Rhipicephalus (Boophilus) decoloratus* ticks picked from Mbarara cattle, producing 27,745,476 150 bp paired end reads. A contiguous sequence corresponding to the S segment was 1624 nt in length, with an open reading frame (ORF) of 1,449 nt (482 AA), the M segment was 5,341 nt (ORF length: 5,102 nt; 1700 AA) and L segment was 11,838 nt (ORF length: 11,838 nt; 3,945 AA). The average coverage was 6189 x, 3882 x and 1833 x for S, M and L respectively. The full contiguous sequences have been deposited in NCBI GenBank under accession numbers MW452933 - MW452935.

Phylogenetic analyses revealed that all 3 segments of the Mbarara CCHFV strain clustered within the African genotype II (Africa 2) clade, together with strains from the Democratic Republic of Congo, Uganda and South Africa (Figures 1-3). The S and L genome segments were highly similar to genotype II strains previously reported, sharing between 91.79 – 94.34 % nucleotide identity (98.76 – 99.38 % amino acid identity) and 91.57 – 94.12 % nucleotide identity (97.49 – 98.56 % amino acid identity), respectively (Figure S1).

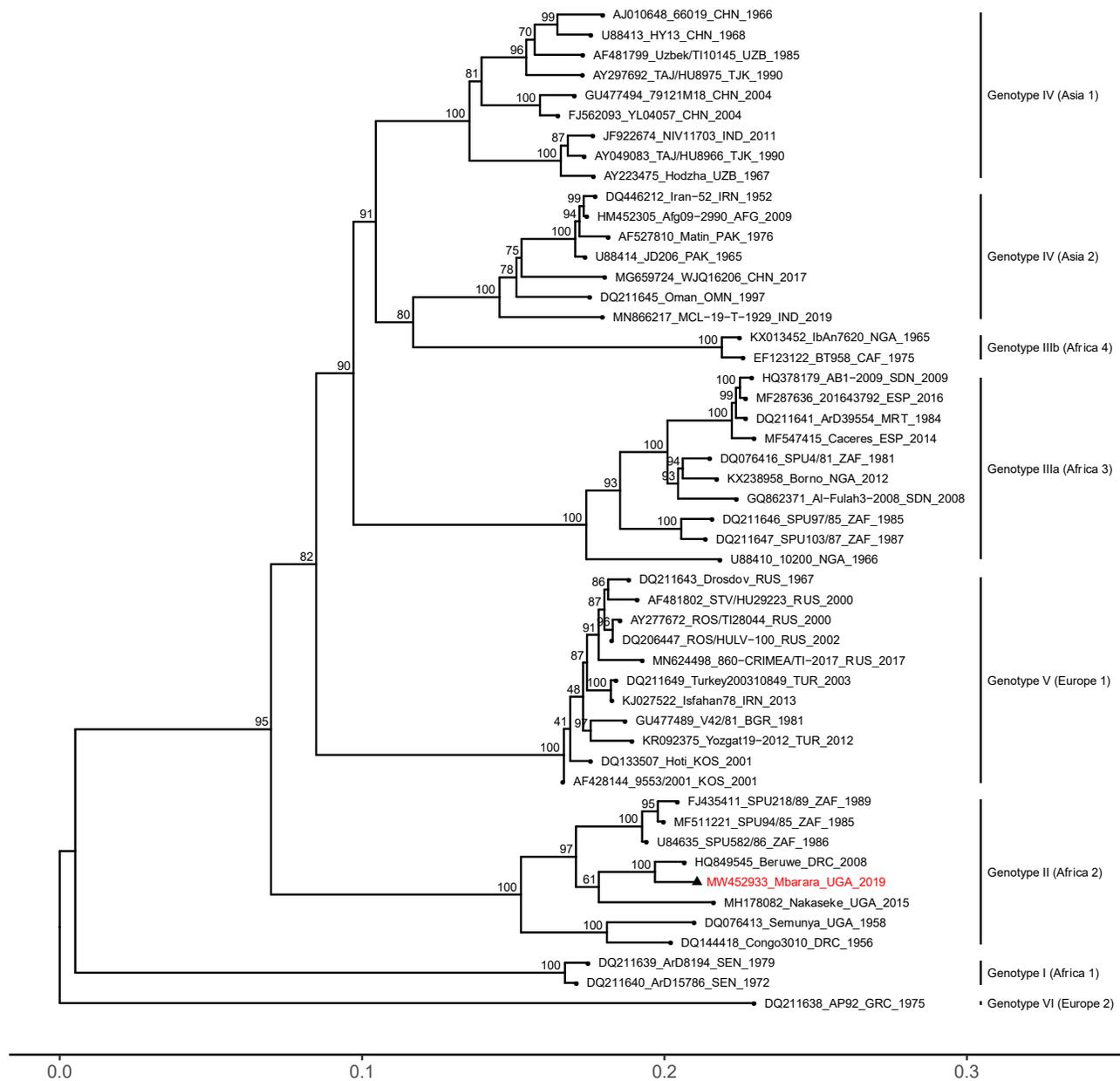


Figure 1. Phylogenetic analysis of the CCHFV S genome segment. Maximum likelihood phylogenies were estimated from complete ORF nucleotide sequences where available. Genbank accession numbers, strain, country of origin and year of sampling are listed in taxon names. Numbers at the tree nodes represent bootstrap support values estimated using the ultrafast bootstrap approximation. Mbarara strain is highlighted in red with a triangular node tip.

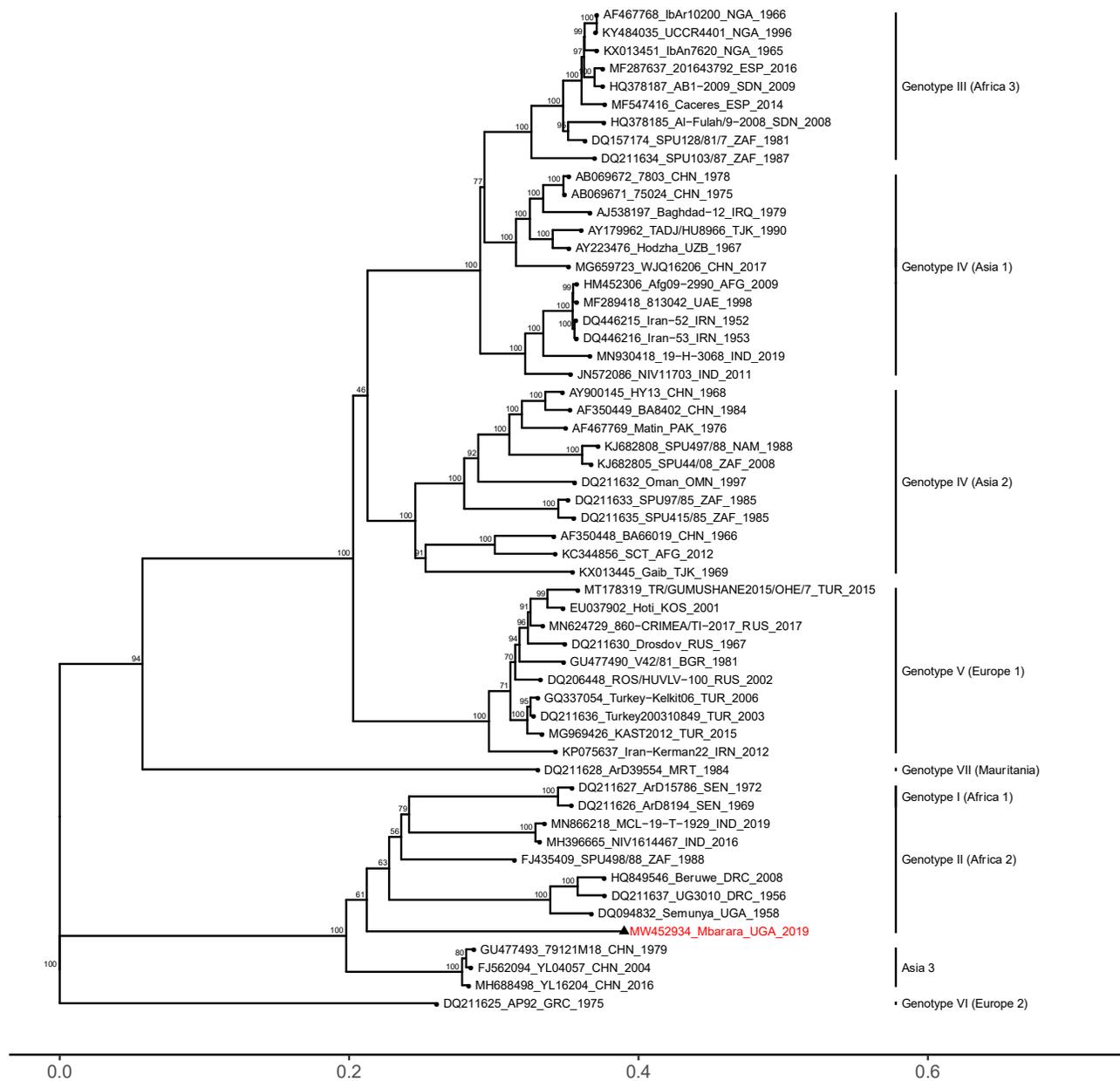


Figure 2. Phylogenetic analysis of the CCHFV M genome segment. Maximum likelihood phylogenies were calculated from nucleotide sequences. The hyper-variable (857 bp) mucin-like domain was trimmed from the 5' end prior to analysis, to facilitate accurate alignment, particularly between genotypes. Genbank accession numbers, strain, country of origin and year of sampling are listed in taxon names. Numbers at the tree nodes represent bootstrap support values estimated using the ultrafast bootstrap approximation. Mbarara strain is highlighted in red with a triangular node tip.

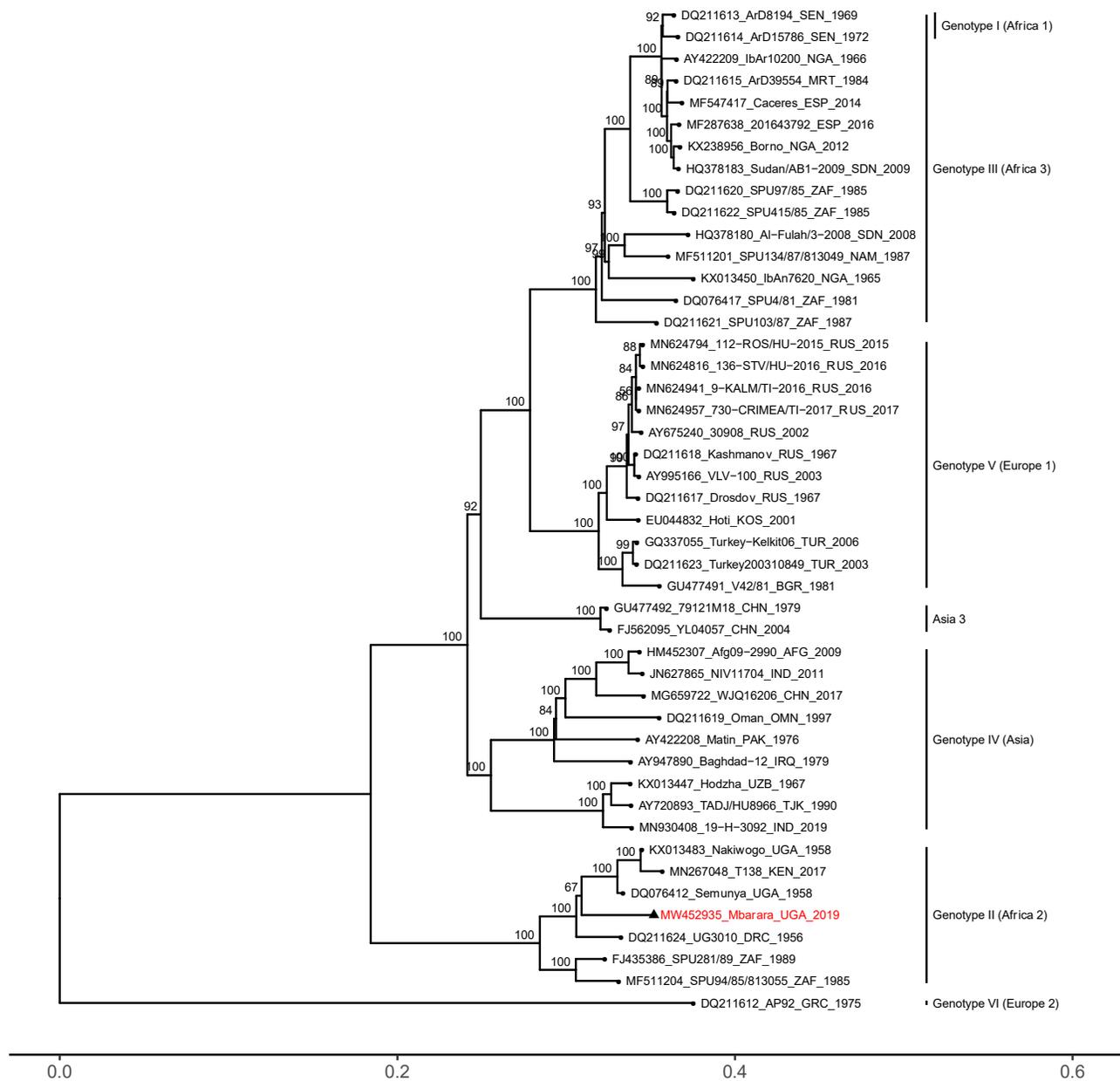


Figure 3. Phylogenetic analysis of the CCHFV L genome segment. Maximum likelihood phylogenies were estimated from complete ORF nucleotide sequences where available. Genbank accession numbers, strain, country of origin and year of sampling are listed in taxon names. Numbers at the nodes represent bootstrap support values estimated using the ultrafast bootstrap approximation. Mbarara strain is highlighted in red with a triangular node tip.

Compared to the S and L segments, the M segment of Mbarara strain was more distantly related, sharing between 77.54 – 79.62 % nucleotide identity (80.45 – 83.40 % amino acid identity) with previously reported genotype II sequences. Those with the greatest similarity to Mbarara strain included several sequences from India (sampled in 2016 and 2019) and China (sampled in 1979, 2004 and 2016) (Figure S1). Average nucleo-

tide sequence similarity within the genotype II clade was lowest (~40 %) in the 5' region that encompasses the signal peptide and mucin-like domains (Figure 4). Regions coding for the secreted glycoprotein GP38 and the surface glycoproteins Gn and Gc also displayed some degree of intra-genotype nucleotide variability (80 – 85 % similarity), whilst nucleotide conservation was highest in the non-structural coding region. Analysis of the M segment for recombination and diversifying episodic selection revealed no evidence for either ($p > 0.05$).

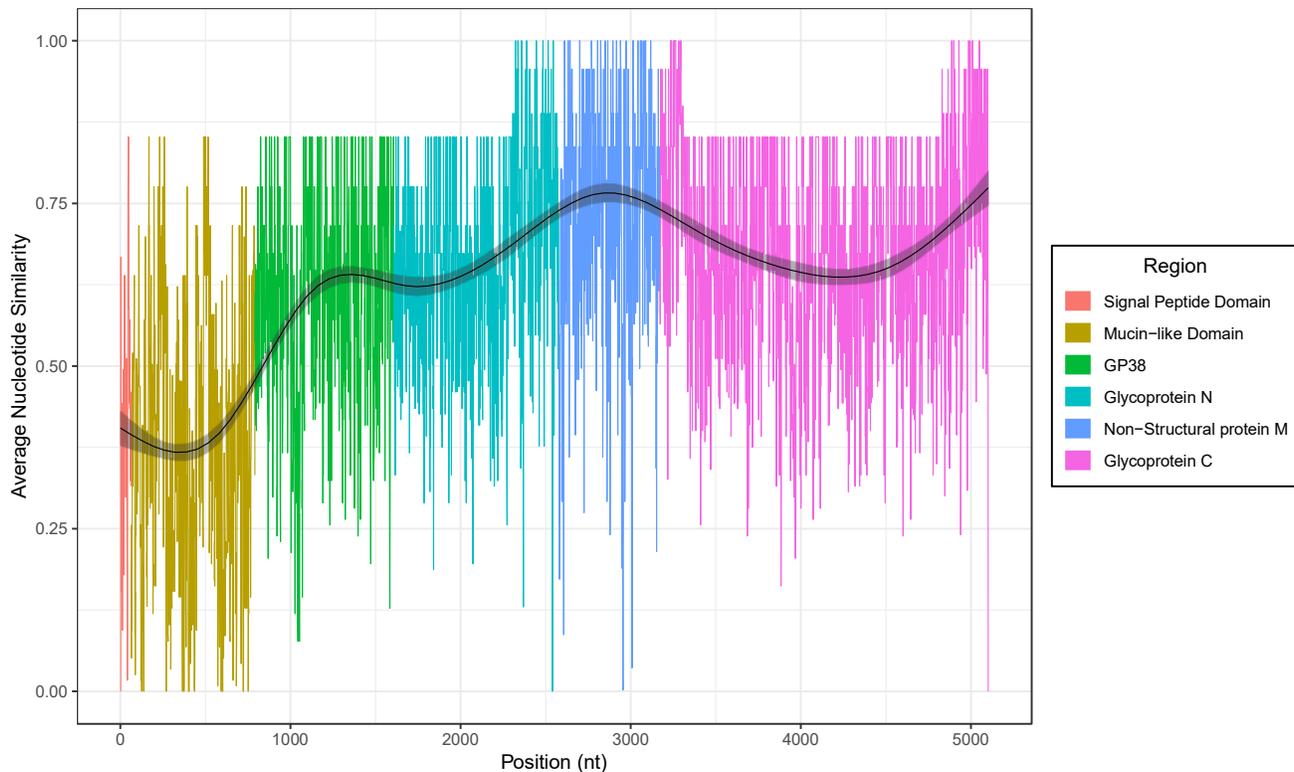


Figure 4. Nucleotide conservation between the M segments of genotype II strains. Nucleotide sequences of genotype II M segments ($n = 12$) were aligned and the average nucleotide similarity was calculated using a sliding window size of 4.

Alignment of genotype II predicted amino acid sequences revealed 164 residues unique to the Mbarara strain (singletons) (Table 1). These novel residues were located throughout all 3 genomic segments, at a frequency ranging from average to above average for the genotype. The majority of singletons (148/164) were observed in the M segment, in which Mbarara strain displayed the greatest number of unique residues across every gene, including the MLD, the non-structural NSm and the glycoproteins GP38, Gn and Gc.

Table 1. Predicted amino acid differences in the coding region of Mbarara strain CCHFV in comparison to previously sequenced African genotype II strains.

Genome Segment	Gene	Number of sequences (full length)	Gene Length (AA)	Polymorphic sites* (Informative sites [†])	Median singleton [‡] count (range) [§]	Mbarara strain singletons	% Sites unique to Mbarara strain
S	NP and Ns	8 (5)	482	11 (3)	1 (0-2)	1	0.2
M	Signal peptide	10 (10)	20	16 (13)	0 (0-1)	5	25.0
	MLD	10 (10)	242	206 (167)	4 (1-29)	76	31.4
	GP38	10 (10)	271	80 (55)	7 (4-14)	22	8.1
	Gn	11 (10)	322	58 (41)	1 (0-4)	12	3.7
	N5m	11 (11)	197	41 (27)	1.5 (0-3)	10	5.1
	Gc	12 (10)	643	103 (53)	2 (0-12)	23	3.6
L	RdRp	7 (5)	3945	153 (32)	7 (0-60)	32	0.8
S, M and L	All	-	6144	668 (391)		164	2.7

* Sites with at least two alleles.

[†] Sites with at least two alleles, where each of those alleles is found in at least two sequences.

[‡] Sites displaying amino acid residues unique within the genotype II clade.

[§] Genotype II strains, excluding Mbarara strain.

4. Discussion

Here we describe the detection and characterisation of a genotype II strain of CCHFV, present in African blue ticks feeding on cattle from a farm in Mbarara, Uganda. To our knowledge, CCHFV has only rarely been described in this tick species (*Rhipicephalus (Boophilus) decoloratus*): twice in Senegal (1972 and 1975) and twice in Uganda (1978 and 2015) [1, 28]. It was not possible to determine whether the virus was derived from within the tick or the host bloodmeal. However, the tick appears to be the more likely source, as a limited number of experimental studies have found CCHFV viremia to be transient and short lived in cattle [29]. As a single-host tick species, *Rhipicephalus (Boophilus) decoloratus* is unlikely to play a major role in zoonotic transmission of the virus. However, it is possible that these ticks may be acting as a reservoir for the virus, maintaining infection of the animal host and facilitating the infection of major vector species (e.g. *Hyalomma spp.*) through co-feeding [30]. Additionally, a number of two- and three-host *Rhipicephalus* species have also been shown to support the presence of CCHFV [30]. Therefore, given that *Rhipicephalus* ticks are common throughout Africa, further investigation into their role with regards to CCHFV epizootology and epidemiology is clearly warranted.

Previous reports of high seroprevalence in livestock and animal handlers throughout Africa (e.g. Bukbuk et al 2016 [31]), suggest that CCHFV infected ticks and/or animals can often be found in abattoirs, highlighting the potential for focusing surveillance measures around these facilities. Notably, the majority of CCHFV-infected ticks sampled from the abattoir were picked from cattle originating from Mbarara. This district, situated in the cattle corridor of western Uganda, is a major source of beef and dairy products for the country. Indeed, cattle herds in Nakasongola have been largely sourced from Mbarara, following their depletion during rebel occupation. This may account for the surprisingly high degree of similarity between the Mbarara CCHFV strain and the strain detected in a tick picked from a cow from Nakasongola. Further sampling of CCHFV strains from other regions of Uganda is required in order to examine transmission dynamics within the country. Possible alternative explanations for this similarity between the CCHFV strain detected in Mbarara and Nakasongola ticks include host-switching of the tick upon arrival at the abattoir, or that the CCHFV-infected tick population originated from the abattoir itself.

Two of the Mbarara strain genomic segments: the nucleocapsid (S) and RNA polymerase (L) showed a high degree of nucleotide conservation, sharing >90 % similarity with other African genotype II sequences, despite the fact that these were derived from strains isolated several decades ago. A similar observation has previously been made for viruses from Pakistan and Nigeria [31], suggesting a tendency for these genomic segments to remain genetically stable over time. In contrast, sequence diversity was greatest in the surface glycoprotein-coding (M) segment, which shared only 78 % nucleotide similarity with strains previously reported from Africa and elsewhere. Interestingly, the closest related M segment sequences belonged to several strains from Asia (e.g. MCL-19-T-1929 (India, 2019) and YL04057 (China, 2004)), which had also clustered within the African genotype II clade, although nucleotide homology was still low (80 %). The S and L segments of these Asian strains belonged to the Asian genotype IV, indicating that the M segment had likely been obtained through re-assortment with an African virus, a process made possible through intercontinental movement of CCHFV-infected ticks on livestock and migratory birds [32]. It is unclear whether the relatively low homology of the Mbarara strain M segment to previously reported genotype II strains, is due to it belonging to a novel lineage or simply the result of several decades of evolution. Testing this theory is made difficult by the paucity of closely related sequences. However, the stark dissimilarity to previously reported genotype II M segment sequences, in contrast to the S and L segments, implies recombination or re-assortment with an as of yet unidentified viral lineage. Unfortunately, many methods designed to identify instances of recombination and re-assortment are reliant upon inclusion of the parent lineages in the analysis, which is not possible with the current, small sample of genotype II strains.

Unique amino acid residues were detected throughout the Mbarara strain genome. However, much of the observed diversity, at both nucleotide and amino acid levels, was located in the genomic M segment. A high proportion of this variability could be attributed to the region encoding the MLD, a highly glycosylated protein of unproven function but hypothesized to be involved in shielding viral epitopes against the host immune response, based on the role of a similar domain in filoviruses [33]. This hypothesis is supported by a recent epitope-mapping study, which reported a high degree of immune reactivity by CCHFV survivors against the MLD [34]. Amino acid variation was also high in the genes encoding the viral glycoproteins, the secreted glycoprotein GP38 in particular. The function of GP38 is similarly unknown, however it is believed to act as a chaperone, assisting in PreGn folding [3]. Interestingly, GP38 has recently shown promise as an antigen following the use of anti-GP38 monoclonal antibodies to protect mice

against lethal challenge by CCHFV [35]. Knowledge of the spectrum of GP38 variants in animal and vector reservoirs may therefore prove useful for the design of future vaccines.

5. Conclusions

A severely neglected pathogen, the lack of genomic data from African CCHFV makes exploration of the evolutionary history of this novel strain difficult. In addition, a lack of viral data from animal and tick reservoirs has highlighted the need to better characterise the natural variation of CCHFV strains in circulation outside of severe, human cases of disease. Indeed, to our knowledge only one other CCHFV strain has been reported from African ticks (IbAr102000). This genotype III virus was isolated in 1966 from a *Hyalomma* tick taken from a Camel in Nigeria. Notably, the virus was passaged multiple times over several decades prior to sequencing, likely introducing a number of mutations throughout this period. It is therefore uncertain how representative this sequence is of the original virus, particularly for a tick virus adapting to mammalian cell culture. In conclusion, generation of additional CCHFV genome sequence directly from ticks, such as that presented here, is vital in order to grant a more complete understanding of CCHFV ecology, transmission and evolution, all of which are necessary for the effective design and deployment of future vaccines and molecular diagnostic tests.

Supplementary Materials: Figure S1: Pairwise nucleotide identity comparisons for CCHFV genome segments S, M and L, Table S1: List of collected tick specimens.

Author Contributions: Conceptualization, Eddie M Wampande and Samuel CB Stubbs; Data curation, Samuel CB Stubbs; Formal analysis, Samuel CB Stubbs; Funding acquisition, Eddie M Wampande, David J Allen, Roger Hewson, Simon DW Frost and Samuel CB Stubbs; Investigation, Eddie M Wampande, Peter Waiswa and Samuel CB Stubbs; Methodology, Eddie M Wampande, Peter Waiswa and Samuel CB Stubbs; Project administration, Samuel CB Stubbs; Resources, David J Allen and Simon DW Frost; Visualization, Samuel CB Stubbs; Writing – original draft, Samuel CB Stubbs; Writing – review & editing, Eddie M Wampande, David J Allen, Roger Hewson and Simon DW Frost.

Funding: This research was funded by the ALBORADA Trust, grant number RG G101892, awarded to SCBS and EMW and the Government of the Republic of Uganda through Makerere University Research and Innovations Fund, which support EMW. SCBS is supported in part by an unrestricted gift from Microsoft Research. DJA and RH are supported by the Medical Research Council, grant number MR/T029196/1.

Data Availability Statement: All consensus sequences generated and analysed for this study were uploaded to NCBI Genbank under accession numbers MW452933 - MW452935. Raw sequence data have been deposited in ENA under study PRJEB42607. Nucleotide and amino acid multiple sequence alignments and maximum likelihood phylogenies for each segment are available at: DOI: 10.5281/zenodo.4455021 [all data will be made available in R1 of the manuscript].

Acknowledgments: We thank Dr. Robert Aruho the wildlife Veterinary officer and Ms Dorothy Kirumira in charge monitoring and research LMNP Uganda wildlife Authority, the contribution of Dr. Rose Ademun Okurut the commissioner Animal Health Ministry of Agriculture Animal Industry and Fisheries (MAAIF), the area Veterinarian (Dr Assimwe) of Kiruhura District (part of original greater Mbarara), the cattle traders and the staff of Uganda Meat packers' abattoir, especially Mr. Joseph Amagoro in the collection of ticks.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Hoogstraal H. The epidemiology of tick-borne Crimean-Congo hemorrhagic fever in Asia, Europe, and Africa. *J Med Entomol* **1979**, 15:307-417.
2. Bente DA, Forrester NL, Watts DM, McAuley AJ, Whitehouse CA, Bray M. Crimean-Congo hemorrhagic fever: History,

- epidemiology, pathogenesis, clinical syndrome and genetic diversity. *Antiviral Res* **2013**, 100:159-89.
3. Zivcec M, Scholte FEM, Spiropoulou CF, Spengler JR, Bergeron É. Molecular insights into Crimean-Congo hemorrhagic fever virus. *Viruses* **2016**, 8:106.
 4. Deyde VM, Khristova ML, Rollin PE, Ksiazek TG, Nichol ST. Crimean-Congo Hemorrhagic Fever Virus Genomics and Global Diversity. *J Virol* **2006**, 80:8834–8842.
 5. Lukashev AN, Klimentov AS, Smirnova SE, Dzagurova TK, Drexler JF, Gmyl AP. Phylogeography of Crimean Congo Hemorrhagic Fever Virus. *PLoS One* **2016**, 11:e0166744.
 6. Ramírez de Arellano E, Hernández L, Goyanes MJ, Arsuaga M, Fernández Cruz A, Negredo A, Paz Sánchez-Seco M. Phylogenetic characterization of Crimean-Congo hemorrhagic fever virus, Spain. *Emerg Infect Dis* **2017**, 23:2078–2080.
 7. Sahay RR, Dhandore S, Yadav PD, Chauhan A, Bhatt L, Garg V, Gupta N, Nyayanit DA, Shete AM, Singh R, Patil S, Sharma H, Meena D, Gupta J, Gawande P, Kalele K, Jain R, Mali D, Malhotra B, Nag V, Khatri PK, Kaur H, Vijay N, Singh S, Gopalakrishnan M. Detection of African genotype in Hyalomma tick pools during Crimean Congo hemorrhagic fever outbreak, Rajasthan, India, 2019. *Virus Res* **2020**, 286:198046.
 8. Zhang Y, Shen S, Fang Y, Liu J, Su Z, Liang J, Zhang Z, Wu Q, Wang C, Abudurexiti A, Hu Z, Zhang Y, Deng F. Isolation, Characterization, and Phylogenetic Analysis of Two New Crimean-Congo Hemorrhagic Fever Virus Strains from the Northern Region of Xinjiang Province, China. *Virol Sin* **2018**, 33:74–86.
 9. Gruber CEM, Bartolini B, Castilletti C, Mirazimi A, Hewson R, Christova I, Avšič T, Grunow R, Papa A, Sánchez-Seco MP, Koopmans M, Ippolito G, Capobianchi MR, Reusken CBEM, Di Caro A. Geographical variability affects CCHFV detection by RT-PCR: A tool for in-silico evaluation of molecular assays. *Viruses* **2019**, 11:953.
 10. Ahlén G, Sofia Appelberg K, Feldmann H, Sällberg M, Mirazimi A. A DNA-based vaccine protects against Crimean-Congo haemorrhagic fever virus disease in a Cynomolgus macaque model. *Nat Microbiol* **2021**. doi: 10.1038/s41564-020-00815-6.
 11. Buttigieg KR, Dowall SD, Findlay-Wilson S, Miloszewska A, Rayner E, Hewson R, Carroll MW. A novel vaccine against Crimean-Congo haemorrhagic fever protects 100% of animals against lethal challenge in a mouse model. *PLoS One* **2014**, 9:e91516.
 12. Walker AR, Bouattour A, Camicas J-L, Estrada-Peña A, Horak IG, Latif AA, Pegram RG, Preston PM. *Ticks of Domestic Animals in Africa: a Guide to Identification of Species*. 1st Ed. Bioscience Reports: Edinburgh, UK, 2003.
 13. Lambert AJ, Lanciotti RS. Consensus amplification and novel multiplex sequencing method for S segment species identification of 47 viruses of the Orthobunyavirus, Phlebovirus, and Nairovirus genera of the family Bunyaviridae. *J Clin Microbiol* **2009**, 47:2398–2404.
 14. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990, 215:403–410.
 15. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **2002**, 30:3059–3066.
 16. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol* **2015**, 32:268–274.
 17. Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast Approximation for Phylogenetic Bootstrap. *Mol Biol Evol* **2013**, 30:1188–1195.
 18. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **2017**, 14:587–589.
 19. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **2019**, 47:W636–W641.
 20. Rice P, Longden L, Bleasby A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet* **2000**, 16:276-277.

21. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria, 2020.
22. Wickham H. Ggplot2 : elegant graphics for data analysis. Springer-Verlag: New York, USA, 2009.
23. Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. ggtree : an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol* **2017**, 8:28–36.
24. Martin DP, Varsani A, Roumagnac P, Botha G, Maslamoney S, Schwab T, Kelz Z, Kumar V, Murrell B. RDP5: A computer program for analysing recombination in, and removing signals of recombination from, nucleotide sequence datasets. *Virus Evol* **2020**, doi: 10.1093/ve/veaa087.
25. Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. Less is more: An adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol Biol Evol* **2015**, 32:1342–1353.
26. Gouy M, Guindon S, Gascuel O. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Mol Biol Evol* **2010**, 27:221–224.
27. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **2019**, 35:526–528.
28. Balinandi S, Patel K, Ojwang J, Kyondo J, Mulei S, Tumusiime A, Lubwama B, Nyakarahuka L, Klena JD, Lutwama J, Ströher U, Nichol ST, Shoemaker TR. Investigation of an isolated case of human Crimean-Congo hemorrhagic fever in Central Uganda, 2015. *Int J Infect Dis* **2018**, 68:88–93.
29. Causey OR, Kemp GE, Madbouly MH, David-West TS. Congo virus from domestic livestock, African hedgehog, and arthropods in Nigeria. *Am J Trop Med Hyg* **1970**, 19:846–850.
30. Gargili A, Estrada-Peña A, Spengler JR, Lukashev A, Nuttall PA, Bente DA. The role of ticks in the maintenance and transmission of Crimean-Congo hemorrhagic fever virus: A review of published field and laboratory studies. *Antiviral Res* **2017**, 144:93-119.
31. Bukbuk DN, Dowall SD, Lewandowski K, Bosworth A, Baba SS, Varghese A, Watson RJ, Bell A, Atkinson B, Hewson R. Serological and Virological Evidence of Crimean-Congo Haemorrhagic Fever Virus Circulation in the Human Population of Borno State, Northeastern Nigeria. *PLoS Negl Trop Dis* **2016**, 10:e0005126.
32. Lindeborg M, Barboutis C, Ehrenborg C, Fransson T, Jaenson TGT, Lindgren PE, Lundkvist Å, Nyström F, Salaneck E, Waldenström J, Olsen B. Migratory birds, ticks, and crimean- congo hemorrhagic fever virus. *Emerg Infect Dis* **2012**, 18:2095-2097.
33. Fusco ML, Hashiguchi T, Cassan R, Biggins JE, Murin CD, Warfield KL, Li S, Holtsberg FW, Shulenin S, Vu H, Olinger GG, Kim DH, Whaley KJ, Zeitlin L, Ward AB, Nykiforuk C, Aman MJ, Berry J, Saphire EO. Protective mAbs and Cross-Reactive mAbs Raised by Immunization with Engineered Marburg Virus GPs. *PLoS Pathog* **2015**, 11:e1005016.
34. Fritzen A, Risinger C, Korukluoglu G, Christova I, Corli Hitzeroth A, Viljoen N, Burt FJ, Mirazimi A, Blixt O. Epitope-mapping of the glycoprotein from Crimean-Congo hemorrhagic fever virus using a microarray approach. *PLoS Negl Trop Dis* **2018**, 12:e0006598.
35. Golden JW, Shoemaker CJ, Lindquist ME, Zeng X, Daye SP, Williams JA, Liu J, Coffin KM, Olschner S, Flusin O, Altamura LA, Kuehl KA, Fitzpatrick CJ, Schmaljohn CS, Garrison AR. GP38-targeting monoclonal antibodies protect adult mice against lethal Crimean-Congo hemorrhagic fever virus infection. *Sci Adv* **2019**, 5:eaaw9535.