Virology 581 (2023) 116–127

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology

Subspecific rodent taxa as the relevant host taxonomic level for mammarenavirus host specificity

Laura N. Cuypers [a,*], Sophie Gryseels [a,b], Natalie Van Houtte [a], Stuart J.E. Baird [c], Christopher A. Sabuni [d], Abdul S. Katakweba [d], Sebastiaan R.M. van den Burg [a], Josef Bryja [c,e], Herwig Leirs [a], Joëlle Goüy de Bellocq [c]

[a] Evolutionary Ecology Group, Department of Biology, University of Antwerp, Universiteitsplein 1, 2610, Antwerp, Belgium
[b] OD Taxonomy and Phylogeny, Royal Belgian Institute of Natural Sciences, Vautierstraat 29, 1000, Brussels, Belgium
[c] Institute of Vertebrate Biology of the Czech Academy of Sciences, Květná 8, 603 65, Brno, Czech Republic
[d] Institute of Pest Management, Sokoine University of Agriculture, P.O. Box 3110 Chuo Kikuu, Morogoro, Tanzania
[e] Department of Botany and Zoology, Masaryk University, Kotlářská 2, 611 37, Brno, Czech Republic

ARTICLE INFO

Keywords:
Host specificity
Host genetic structure
Comparative phylogeography
Songea virus
Berega virus
Ngerengere virus
Mus minutoides
Grammomys surdaster
Spill-over
Host jump

ABSTRACT

Mastomys natalensis-borne mammarenaviruses appear specific to subspecific *M. natalensis* taxa rather than to the whole species. Yet mammarenaviruses carried by *M. natalensis* are known to spill over and jump hosts in northern sub-Saharan Africa. Phylogeographic studies increasingly show that, like *M. natalensis*, small mammals in sub-Saharan Africa are often genetically structured into several subspecific taxa. Other mammarenaviruses may thus also form virus-subspecific host taxon associations. To investigate this, and if mammarenaviruses carried by *M. natalensis* in southern Africa are less prone to spill-over, we screened 1225 non-*M. natalensis* samples from Tanzania where many small mammal taxa meet. We found mammarenavirus RNA in 6 samples. Genetic/genomic characterisation confirmed they were not spill-over from *M. natalensis*. We detected host jumps among rodent tribe members and an association between mammarenaviruses and subspecific taxa of *Mus minutoides* and *Grammomys surdaster*, indicating host genetic structure may be crucial to understand virus distribution and host specificity.

1. Introduction

Mammarenaviruses are single-stranded, ambisense RNA viruses from the *Arenaviridae* family. Their genomes comprise two segments: a large segment (L, ±7200 nt) encoding an RNA-dependent RNA polymerase (L) and a zinc-binding matrix protein (Z), and a small segment (S, ±3500 nt) encoding a nucleoprotein (NP) and a glycoprotein precursor (GPC) (Radoshitzky et al., 2019). For several decades, members of the current *Mammarenavirus* genus were the only known arenaviruses, expressing a marked host specificity to a single rodent reservoir species. Only Tacaribe virus was detected in bats rather than rodents (Downs et al., 1963; and more recently Góes et al., 2022). With increased screening and metagenomic sequencing of non-rodent samples, the number of recognised arenaviruses and their hosts has increased considerably (Kuhn et al., 2022). The genus *Arenavirus* was renamed *Mammarenavirus* and the genera *Reptarenavirus*, *Hartmanivirus* and *Antennavirus* were established to accommodate arenaviruses discovered in reptiles and fish (Abudurexiti et al., 2019; Maes et al., 2018; Radoshitzky et al., 2015). A fifth genus, *Innmovirus*, has been approved by the International Committee on Taxonomy of Viruses (ICTV) to accommodate Hailar virus discovered in river sediment samples (Chen et al., 2021a,b) and 'Anole arenavirus' may represent an additional arenavirus genus (Harding et al., 2022). Most mammarenaviruses have still only been detected in a single rodent host species from the Muroidea superfamily, typically murines in Africa and sigmodontines and neotomines in the Americas (Radoshitzky et al., 2015). Moreover, sympatric rodent species were found to carry distinct mammarenaviruses (Fulhorst et al., 1999; Goüy de Bellocq et al., 2010; Ishii et al., 2012) and closely related viruses are often carried by closely related host species (Cuypers et al., 2020; Inizan et al., 2010; Těšíková et al., 2021).

More and more mammarenaviruses have been detected in multiple hosts. For some, it seems there is one principal host and other species

* Corresponding author.
E-mail address: lauracuypers@outlook.be (L.N. Cuypers).

https://doi.org/10.1016/j.virol.2023.02.014
Received 22 December 2022; Received in revised form 19 February 2023; Accepted 23 February 2023
Available online 8 March 2023

may rather have been infected through spill-over (Fulhorst et al., 1999; Milazzo et al., 2015). For others like Dhati Welel virus and Lassa virus, it is less clear if other host species were infected through spill-over or if they could play an important role in virus dynamics (Goüy de Bellocq et al., 2020; Lavergne et al., 2015; Olayemi et al., 2016a). Lassa virus and several other mammarenaviruses can also spill over to humans and can cause disease (Bausch and Mills, 2014; Blasdell et al., 2016). Some viruses appear to have been able to make a successful host jump (Cajimat et al., 2007; Kronmann et al., 2013; Milazzo et al., 2015) and We¯nzho¯u mammarenavirus appears to have an exceptionally broad host range and to circulate in multiple rodents and shrews in Asia (Wu et al., 2021).

When a mammarenavirus is detected in more than a single rodent host species, it is usually considered right away that the virus might be specific to a higher host taxonomic level than the species level. In contrast, a host taxonomic level lower than the species level is rarely considered: when a virus is only found in one host species it might not circulate in that species as a whole, but only in an subspecific taxon of that species. Meanwhile, an increasing number of phylogeographic studies on rodents reveal that many rodents show pronounced subspecific genetic structure, often linked to divergence due to climatic and geomorphological processes (Giorello et al., 2021; Puckett et al., 2021; Shimada et al., 2007). For example, in Africa humid and dry climatic phases alternated during the Plio- and Pleistocene (Trauth et al., 2007; Williamson, 1985). During more humid phases, savannahs contracted and savannah-dwelling rodents diverged in isolated fragments (Aghová et al., 2017; Hánová et al., 2021), while forests expanded connecting forest-dwelling rodents. Conversely, during drier phases the savannahs expanded bringing savannah-dwelling rodents into secondary contact, while forest-dwelling rodents diverged in forest fragments (Mizerovská et al., 2019; Nicolas et al., 2020). Divergence has resulted in genetically distinct groups, some of which are recognised as different species today, others as clades within species (Cuypers et al., 2022b). Regardless of the formal taxonomic level, the genetic structure appears to impact their mammarenaviruses.

The hypothesis that subspecific rodent taxa rather than the species as a whole might be the relevant host taxonomic level for some mammarenaviruses was first proposed for mammarenaviruses carried by *Mastomys natalensis* (Smith, 1834) individuals. Gryseels et al. (2017) found that *M. natalensis* B-IV and B–V (sensu Colangelo et al., 2013) meet in a narrow hybrid zone in Tanzania. Gairo virus occurs on the B-IV side, Morogoro virus on the B–V side, and they co-occur in the hybrid zone (Cuypers et al., 2020; Gryseels et al., 2017). There is no clear environmental barrier, so it appears that these viruses have the opportunity to spread into the range of the other taxon, but they have not been observed to do so. This may also be the case for other *M. natalensis*-mammarenavirus associations. The *M. natalensis cytochrome b* (*cytb*) gene reveals six mitochondrial (mt) clades in distinct geographic regions throughout its sub-Saharan African range (Colangelo et al., 2013). Not only Gairo and Morogoro virus have been found in the B-IV and B–V mt range, respectively, but Lassa virus has been detected in the approximate A-I mt range (Olayemi et al., 2016b), a mobala-like virus (Mayo-Ranewo) in the A-II mt range (Olayemi et al., 2016b), Dhati Welel virus in the A-III mt range (Goüy De Bellocq et al., 2020), and both Luna and Mopeia virus in the B-VI mt range (Cuypers et al., 2020; Ishii et al., 2011; Těšíková et al., 2021; Wulff et al., 1977, Fig. 1).

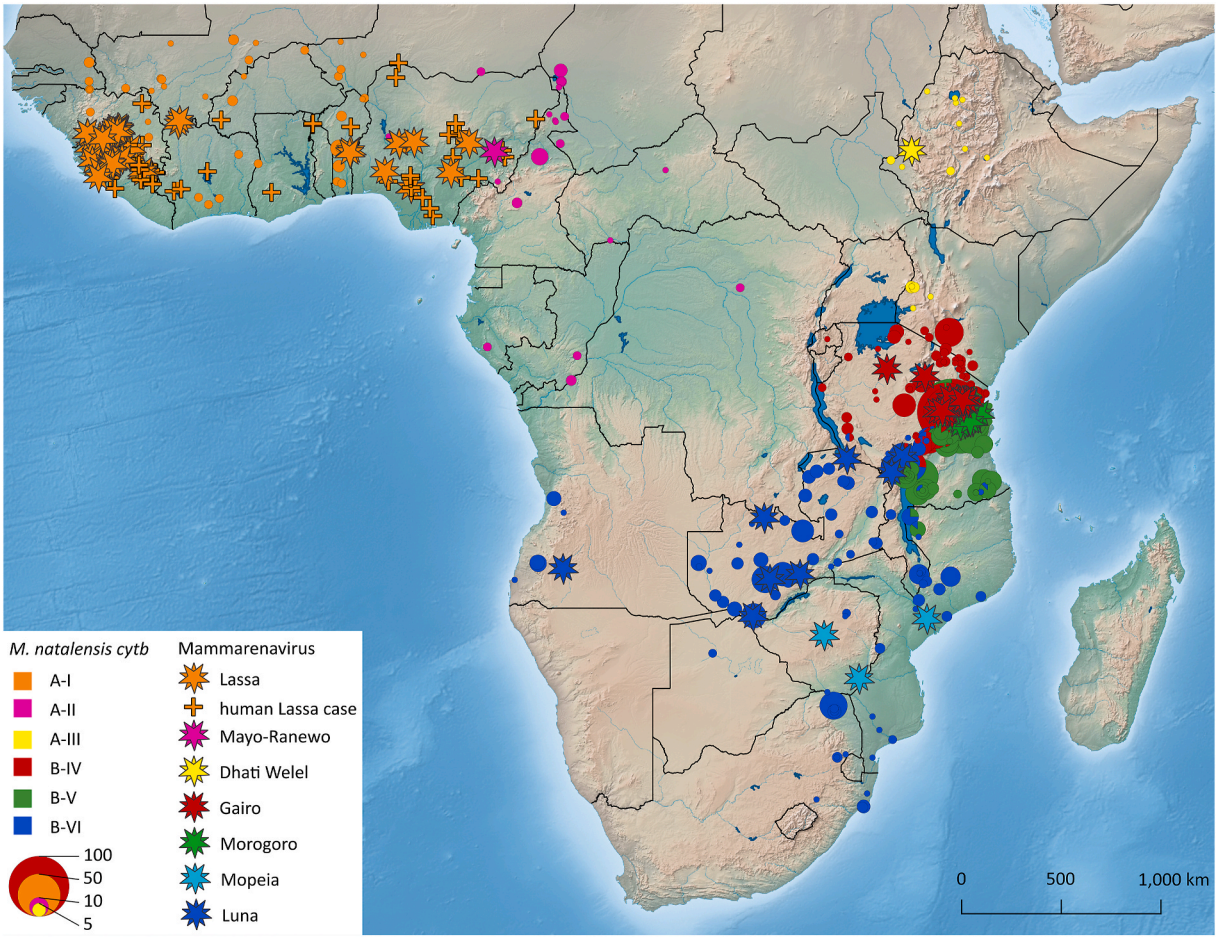

Fig. 1. Distribution of *Mastomys natalensis* and the mammarenaviruses they carry. Host pie charts are scaled to the number of individuals sampled. Based on Goüy de Bellocq et al. (2010) with additional data from Těšíková et al. (2021).

Each mitochondrial clade might thus represent a distinct subspecific taxon that carries its own mammarenavirus (or its own mammarenaviruses for B-VI).

However, Dhati Welel virus has also been detected in an *M. awashensis* individual in the same locality as *M. natalensis* A-III positive individuals (Goüy de Bellocq et al., 2020). Similarly, Lassa virus has been detected in *M. natalensis* individuals with A-II mt (but in two localities where A-I mt predominated in the contact zone), as well as in *Mastomys erythroleucus* individuals (Olayemi et al., 2016a,b). Divergent Lassa or Lassa-like viruses have also been detected in *Hylomyscus pamfi*, *Mus baoulei* and *Mus setulosus* individuals (reviewed in Olayemi and Fichet-Calvet, 2020). Furthermore, Lassa repeatedly spills over to humans (Andersen et al., 2015), causing variable but sometimes high case fatality rates (Akpede et al., 2019). No spill-over has been reported for the mammarenaviruses carried by individuals of the *M. natalensis* B taxa in south-eastern Africa, but, to date, relatively few sympatric small mammals have been screened for mammarenaviruses there (Goüy de Bellocq et al., 2010; Gryseels et al., 2015; Ishii et al., 2012; Těšíková et al., 2021).

Gryseels et al. (2015) screened 370 *M. natalensis* and 39 other small mammal individuals from three localities. They detected two distinct mammarenaviruses: a virus in *Arvicanthis neumanni* in Chakwale, and Gairo virus in all three localities in *M. natalensis* individuals with B-IV mitochondria, but also in two individuals from Berega with B-IV mitochondria that are likely [B-IV, B–V] hybrids (Cuypers et al., 2020). Goüy de Bellocq et al. (2010) screened 489 *M. natalensis* and 34 other small mammal individuals from Morogoro city in Tanzania. They detected three distinct mammarenaviruses: Morogoro virus in B–V *M. natalensis* and two other viruses, one in a *Lemniscomys rosalia* clade A individual (sensu Hánová et al., 2021), and one in a *Mus (Nannomys) minutoides* (Smith, 1834) individual from the 'South East Africa clade' (SE, sensu Bryja et al., 2014). The virus detected in *L. rosalia*, Mafiga virus, has recently been fully characterised in Cuypers et al. (2022a) and is the only known mammarenavirus from *L. rosalia*. In *M. minutoides*, however, other mammarenaviruses have been detected as well. Seli virus was characterised from five 'West African clade' individuals (W, sensu Bryja et al., 2014) in Sierra Leone (Vučak et al., 2022). For both *L* and *NP*, Seli virus clusters within partial sequences of Kodoko virus detected in two mice from the same clade in Guinea (Lecompte et al., 2007), so we here consider them the same virus which we will refer to as Kodoko virus, the name given at initial discovery. Another virus, Lunk virus, was characterised from a 'Zambia and surrounding clade' individual (ZA, sensu Bryja et al., 2014) in Zambia (Ishii et al., 2012). Like *M. natalensis*, *M. minutoides* occurs throughout sub-Saharan Africa with distinct mt clades in different geographic regions (Bryja et al., 2014), which might host distinct mammarenaviruses.

In this study we screened 1225 non-*M. natalensis* rodents and shrews sampled in Tanzania over a period of more than ten years. The goal was to investigate if some of the small mammals living in sympatry with *M. natalensis* harbour mammarenaviruses and, if yes, are they spill-over from the three mammarenaviruses known from *M. natalensis* B-IV, B–V, B-VI in Tanzania (Cuypers et al., 2020) or are they distinct mammarenaviruses specific to other small mammals. Since many small mammal species are genetically structured across Africa and several are known to be reservoirs of mammarenaviruses, this study allows us to investigate if the host genetic structure shapes mammarenavirus evolutionary history. We focus on Tanzania as it is one of the most important biogeographic crossroads in Eastern Africa where many small mammal clades come together (Cuypers et al., 2022b), including the three *M. natalensis* B taxa. Moreover, Goüy de Bellocq et al. (2010) shows that there can be local hotspots in Tanzania where as many as three mammarenaviruses can be detected in sympatric rodents in a relatively small sampling area.

2. Material and methods

2.1. Trapping and sampling

Rodents and shrews were caught in Tanzania in 2008–2019. *Mastomys natalensis* captures of the 2008–2016 fieldwork expeditions have previously been published in Gryseels et al. (2017) and Cuypers et al. (2020). Sherman live traps (H. B. Sherman Traps Inc., Tallahassee, USA) and snap traps were baited with a mix of peanut butter, maize flour and in 2015–2016 canned fish. Rodents and shrews caught alive were killed by an overdose of isoflurane or by cervical dislocation. Kidney samples were stored in RNAlater and blood from the heart was preserved on pre-punched Serobuvard filter papers (Zoopole, Ploufragan, France). Field work was approved by the University of Antwerp Ethical Committee for Animal Experimentation (2011–52, 2014–98, 2017–75) and adhered to the Code of Conduct for Research Ethics of the Sokoine University of Agriculture.

2.2. Mammarenavirus screening

1225 rodent and shrew kidney, salivary gland or dried blood samples of at least 48 species belonging to 27 genera were screened for mammarenavirus RNA (see Supplementary Material 1). The bulk of the small mammals belonged to the Muridae family (N = 1114), mostly from the Murinae subfamily (N = 794). Except for the 2015–2019 *Mus minutoides* samples, samples were first pooled by two and then positive pools were resolved. RNA was extracted with the RTP DNA/RNA Virus Mini Kit (STRATEC Molecular, Berlin, Germany) and with the Nucleospin RNA Isolation kit (Macherey-Nagel, Düren, Germany), for dried blood samples and tissue samples, respectively.

For the 2008–2012 samples, 8 μL of RNA extract was reverse-transcribed to cDNA with Maxima Reverse Transcriptase and Random Hexamers (Thermo Scientific, Waltham, USA), followed by two mammarenavirus PCRs with 0.02 U Phusion DNA polymerase (Finnzymes, New England Biolabs, USA), 0.2 μM dNTPs (Thermo Scientific, Waltham, USA), 43% of ddH_2O and 13% of cDNA template in a total volume of 15 μL. The first PCR targeted a 340 nt fragment of the *L* gene with 0.4 μM LVL_3359D_Y+ and LVL_3359G_Y+ and 0.6 μM LVL_3754A_R- and LVL_3754D_R-primers (Vieth et al., 2007). The second PCR targeted a 558 nt fragment of the *NP* gene with 0.5 μM OWS2805-fwd, OWS2810-fwd, OWS3400-rev and OWS3400A-rev primers (Ehichioya et al., 2011). Both PCRs were run with 1 cycle at 98 °C for 30 s; 45 cycles of 98 °C for 5 s, 51 °C for 30 s, and 72 °C for 20 s; and final extension at 72 °C for 5 min. An *NP* fragment was also obtained this way for a *Mus minutoides*-virus from Goüy de Bellocq et al. (2010). Additionally, a PCR targeting 1000 nt of the S segment, including a part of the *GPC* gene, was attempted on all positive samples with primers OWS-1-fwd and OWS-1000-rev following Ölschläger et al. (2010).

The 2015–2019 samples were screened using the SuperScript II One-Step RT-PCR System kit (Invitrogen, Carlsbad, USA) with 4.5 μL of RNA extract, 10 μL 2X Reaction Mix, 0.3 μL $MgSO_4$, 0.4 μL SuperScript II RT/Platinum Taq Mix, 0.05 μL RNase free water and 0.8 μL of each of the four LVL primers from above and MoroL3359-forward and MoroL3753-reverse (Günther et al., 2009). *NP* and *GPC* fragments from positive samples were amplified separately with 0.8 μL of each of the OWS primers mentioned above. Cycling conditions followed those of Vieth et al. (2007).

All amplicons were Sanger sequenced in both directions at the Genetic Service Facility of the Vlaams Instituut voor Biotechnologie (Antwerp, Belgium). Amplicons of samples that were not whole-genome sequenced (see below) were deposited in GenBank under accession numbers OP775651– OP775653.

2.3. WGS and assembly

RNA from *Mus minutoides* kidney samples (TZ23131, TZ23187 and

TZ28088), a *Grammomys surdaster* (Thomas and Wroughton, 1908; sensu Musser and Carleton, 2005 and Bryja et al., 2017) salivary gland (TZ33709), and from a *Grammomys surdaster* dried blood sample (TZ31109) was extracted using the RNeasy Mini kit (Qiagen, Hilden, Germany), Nucleospin RNA Isolation kit (Macherey-Nagel, Düren, Germany) and the QIAmp Viral RNA Mini kit (Qiagen, Hilden, Germany), respectively. cDNA synthesis, amplification and library preparation were performed as in Goüy De Bellocq et al. (2020; TZ28088) or Těšíková et al. (2021; TZ23131, TZ23187 and TZ31109) and those libraries were sequenced at Novogene (Cambridge, UK). For TZ33709 the library was prepared with the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina and sequenced by Genewiz (Leipzig, Germany). Read quality was checked using FastQC (Andrews, 2010) and adapters and low-quality bases were trimmed from the raw reads using Skewer (Jiang et al., 2014). A de novo assembly was run in metaSPAdes (Nurk et al., 2017) and the contigs were classified using BlobTools 1.1.1 (Laetsch and Blaxter, 2017) as in Cuypers et al. (2022a). The trimmed reads were deduplicated in Geneious Prime (Biomatters, Auckland, New Zealand) using Dedupe Duplicate Read Remover 38.84. The metaSPAdes contigs were then polished by mapping the deduplicated reads to the contigs with low sensitivity and a consensus threshold of 75%. The raw data were deposited under ENA project no. PRJEB52502. The assembled segments were deposited in GenBank under accession numbers OP775654– OP775658.

2.4. Host characterisation and divergence

Host *cytb* sequences and identifications have already been published in other studies for 540 of the 1225 screened individuals (see Supplementary Material). We identified three additional mammarenavirus-positive hosts by BLASTing their *cytb* sequence. The *cytb* sequence of TZ22285, the mammarenavirus-positive *Mus minutoides* individual from Goüy de Bellocq et al. (2010), was amplified using L14723 and H15915 primers (Lecompte et al., 2002) and deposited in GenBank under accession number OP775650. For mammarenavirus-positive individuals whose samples were whole-genome sequenced, mitochondrial sequences were assembled in the same way as the mammarenavirus segments and deposited in GenBank under accession numbers OP764673-OP764677.

Cytb sequences of *M. minutoides* hosts from this study and the literature were aligned in Geneious Prime with the Geneious alignment algorithm. Pairwise Kimura 2-parameter (K2P) distances, and between- and within-group mean distances were calculated in MEGA11 (Tamura et al., 2021). Pairwise K2P-distances were also calculated for mammarenavirus-positive *G. surdaster*, but a non-mammarenavirus-hosting individual of the same species (RS0756) was added as well. This individual belongs to the same mt clade as the Solwezi virus host (13ZR68), allowing not only a *cytb* comparison, but a (near) full mitochondrial comparison.

2.5. Mammarenavirus phylogenetic analysis

Alignments per gene were produced using the translation alignment algorithm in Geneious Prime. New mammarenavirus sequences were aligned together with a sequence of each classified mammarenavirus from the Old World (OW) mammarenavirus complex and with additional sequences from mammarenaviruses carried by *M. natalensis*, *M. minutoides* and *G. surdaster* individuals. A suspected chimaeric region of the Mariental virus *L* gene (870 nt) was masked (Těšíková et al., 2021). Furthermore, one N was added at position 7074 in the Lunk L sequence to restore the reading frame and enable a longer sequence and the first four nucleotides of the Alxa virus *NP* gene were cut so it starts with a start codon (Cuypers et al., 2022a). Amino acid translations were aligned as well in order to calculate their pairwise identities in MEGA11. Nucleotide sequences were compared to known mammarenaviruses using the PAirwise Sequence Comparison (PASC) webtool (Bao et al., 2014).

For the *L*, *NP*, and *GPC* genes a phylogenetic tree was inferred with MrBayes v3.2.7 (Ronquist et al., 2012). Nucleotide sequences were partitioned by codon position and the nucleotide substitution model was set at GTR + G. Two independent runs with 4 chains ran for 2,500,000 generations with a burn-in of 25% and sampling set at every 500 generations. Average standard deviation of split frequencies dropped well below 0.01. Total effective sample sizes of all parameters exceeded 200 and their trace patterns were examined in Tracer (Rambaut et al., 2014). Trees were visualised in FigTree v1.4.4 (Rambaut, 2012).

3. Results

3.1. Mammarenavirus screening

No shrew samples (N = 85) tested positive, but mammarenavirus *L* gene RNA was detected in 6 out of 1140 rodent samples (see Supplementary Material 1 for more details). Notable rodent genera without any positive samples in our screening are *Acomys* (0/132), *Aethomys* (0/128) and *Gerbilliscus* (0/118). Six samples tested positive for mammarenavirus RNA, all individuals from family Muridae, subfamily Murinae. From the tribe Murini, none of the 226 *Mus Triton* samples tested positive, but 3 other samples from the *Mus* genus did: out of 50 *Mus minutoides* from 23 localities, 1/2 from Mkundi, 1/9 from Morogoro and 1/2 samples from Ngana tested positive. From the tribe Arvicanthini, two out of 26 *Grammomys surdaster* samples from 18 localities tested positive, a kidney sample from Lilondo (TZ31109) and a salivary gland sample from Berega (TZ33709); and one out of 27 *Lemniscomys rosalia* samples from 15 localities tested positive, one of the three kidney samples from Komtema (TZ29841).

Preliminary sequences obtained by Sanger sequencing showed that the viruses were not Morogoro, Gairo or Luna virus found in *Mastomys natalensis* in Tanzania, but related to the Tanzanian *Mus minutoides* virus or likely new mammarenaviruses. The mammarenavirus in the *L. rosalia* individual, Mafiga virus, was described in Cuypers et al. (2022a). That sample was co-infected with two L segments, a unique phenomenon for mammarenaviruses.

3.2. Mammarenavirus genomic characterisation

For the *Mus minutoides* sample from Ngana too few viral reads were produced during whole genome sequencing to assemble a complete genome. The four remaining mammarenaviruses were successfully sequenced and assembled. Three are fully assembled; TZ33709 misses a few non-coding nucleotides at all ends except the 3' end of the S segment. Table 1 summarises their main characteristics. All genomes show a typical mammarenavirus genomic organisation with two segments with each two ambisense open reading frames (ORFs) separated by noncoding intergenic regions with stem-loop structures. Trimmed, deduplicated read coverages varied from 334 ± 106 (SD) to 249,526 ± 88,664 for the L segments and from 524 ± 286 to 423,199 ± 164,018 for the S segments (Table 1). S segment coverage is higher than L segment coverage for TZ23131, TZ23187 and TZ31109, which was also the case for TZ29841 from the same sequencing batch (Cuypers et al., 2022a). However, for TZ33709, L segment coverage is more similar (and even slightly higher compared) to S segment coverage. The other mammarenaviruses from that sequencing batch have rather similar L and S segment coverages as well. As the batches differed in library protocol, we suggest this may have a major influence on the balance of L and S segment coverage.

The L segments are between 7182 and 7280+ (the incomplete segment of TZ33709) nt long and encode the *Z* and *L* genes. The Z proteins are 93–98 amino-acids (AAs) long and contain the regular motifs of OW mammarenaviruses (Fehling et al., 2012) such as a conserved N-terminal myristoylation site G_2 for myristic acid attachment; a conserved, centrally located, zinc-binding RING domain; and the

Table 1
A summary of host and genomic characteristics, best PAirwise Sequence Comparison (PASC) hits with PASC pairwise identities and amino acid (AA) identities, and accession numbers (AN) for new sequences. The pairwise identities with the Lunk L AA sequence are indicated with an asterisk because one N was added to the NC_018711 nucleotide sequence at position 7074 before translation to elongate the L protein sequence. See Supplementary Material 2 for p-distances between all sequences. The plus signs at the end of the TZ33709 segment lengths in nucleotides (nt) indicate that the segments are incomplete at (a) non-coding end(s).

	Sample	TZ23131	TZ23187	TZ31109	TZ33709
	Virus	Ngerengere	Ngerengere	Songea	Berega
	Locality	Morogoro	Mkundi	Lilondo	Berega
	AN raw read data	ERX10442960	ERX10443037	ERX10437543	ERX10442948
	Host	*Mus minutoides* SE	*Mus minutoides* SE	*Grammomys surdaster* su9	*Grammomys surdaster* su10
	Host mitochondrion AN	OP764674	OP764675	OP764676	OP764677
L segment	Segment AN	OP775654	OP775656	OP775658	OP775660
	Mean coverage ±SD	249,526 ± 88,664	4517 ± 1260	334 ± 106	875 ± 545
	Length (nt)	7193	7182	7217	7280+
	Length L (AA)	2213	2213	2217	2215
	Length Z (AA)	93	93	96	98
	Late domains in Z	YLCRDCL	YLCRDCL	YLCL	YLCL
		PSAP	PSAP	PSAP	PSAP
		PPPY	PPPY	PPPY	PPPY
	Best PASC hit (AN)	Lunk (NC_018711)	Lunk (NC_018711)	Solwezi (NC_038366)	Solwezi (NC_038366)
	PASC nt identity (%)	69.6	69.0	69.6	57.8
	L AA identity (%)	75.5*	75.6*	66.3	57.5
	Z AA identity (%)	61.3	58.1	62.6	65.6
S segment	Segment AN	OP775655	OP775657	OP775659	OP775661
	Mean coverage ± SD	423,199 ± 164,018	8306 ± 3994	524 ± 286	843 ± 525
	Length (nt)	3380	3382	3416	3328+
	Length GPC (AA)	495	495	496	492
	Length NP (AA)	558	558	569	568
	CTL epitope in NP	GVYLGNL	GVYLGNL	GVYMGNL	GVYMGNL
	Probable antigenic site in NP	RKEKRDDKD	RKEKRDDKD	RKDKRSDTD	RKEKRSEAD
	GP1/GP2 cleavage site	RRLS	RRLS	RRLL	RRIL
	N-glycosylation GP1	6	5	7	8
	N-glycosylation GP2	3	3	4	4
	Best PASC hit (AN)	Lunk (NC_018710)	Lunk (NC_018710)	Solwezi (NC_038367)	Mariental (KM272987)
	PASC nt identity (%)	77.4	77.3	62.8	69.8
	GPC AA identity (%)	88.5	88.9	78.6	81.3
	NP AA identity (%)	92.1	92.7	82.4	78.0

late domains YxxL, P[T/S]AP and PPxY which mediate interactions between viral proteins and components of the endosomal sorting complexes required for transport. The L proteins are 2213 to 2217 AAs long, and contain the conserved polymerase subdomains with the pre-A and A through E motifs (Ferron et al., 2017), and the key active site residues of the NL1 domain, a cap-snatching endonuclease for viral mRNA transcription (Morin et al., 2010). Due to our addition of the N nucleotide in the Lunk L sequence, the coding sequence is elongated so that it now also contains the key active site residues of the NL1 domain which were lacking before.

The three complete S segment lengths range between 3380 and 3416 nt. All S segments encode the *GPC* and *NP* genes. The glycoprotein precursors are 492–495 AAs long. Potential N-glycosylation sites were determined through the NetNGlyc 1.0 server (https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0 accessed on October 25, 2022) and revealed five to eight and three to four sites in GP1 and GP2, respectively, all in positions analogous to other mammarenaviruses (Bonhomme et al., 2011). The motif at the GP1/GP2 cleavage site is RRLS for both mammarenaviruses carried by *M. minutoides* individuals, RRLL for TZ31109 and RRIL for TZ33709. The nucleoproteins are 558–569 AAs long and contain a DEDDh exonuclease motif conserved among mammarenaviruses (Hastie et al., 2011). A cytotoxic T-lymphocyte epitope described for lymphocytic choriomeningitis virus (LCMV; Whitton et al., 1989) is also present in the mammarenaviruses carried by *G. surdaster* individuals. The epitope differs in one AA in the mammarenaviruses carried by the *M. minutoides* individuals, as in Lunk, Gairo, Dhati Welel and Alxa virus. The viruses carried by the *M. minutoides* individuals also contain the same probable antigenic site in the N-terminal portion of the nucleoprotein as LCMV: RKEKRDDKD. Homologous sequences occur in other mammarenaviruses (Gonzalez et al., 1995), including the viruses carried by the *G. surdaster* individuals.

3.3. Host identification and divergence

The *cytb* sequences of the three new *M. minutoides* hosts (TZ23131, TZ23187 and TZ28088) reveal that the three hosts' mitochondria belong to the 'South East Africa (SE) clade' (*sensu* Bryja et al., 2014). Mitochondrial sequences were assembled to (near) completion (93–100% complete), making the complete sequence the first full mitochondrial sequence for *M. minutoides* in GenBank. Pairwise K2P-distances are shown in Table 2. Within-group mean distances of *M. minutoides* mammarenavirus hosts were 1.7% for the SE clade and 0.4% for the 'West African (W) clade'. No within-group mean distance was calculated for the 'Zambia and surrounding (ZA) clade' as there is only one known mammarenavirus host. Between-group mean distances were 7.3% for Zambia and surrounding clade ZA-W, 6.8% for ZA-SE and 5.7% for W-SE. Nearly complete (99%) host mitochondrial sequences were also assembled for the *Grammomys surdaster* hosts. The individual from Lilondo has mt belonging to the su9 clade (sensu Bryja et al., 2017; Cuypers et al., 2022b), while the individual from Berega has mt belonging to the su10 clade. Between group pairwise K2P-distances for *cytb* and mitochondrial sequences varied between 4.2 and 6.8% (Table 2).

3.4. Mammarenavirus phylogeny and divergence

In Bayesian phylogenetic tree analyses, the new mammarenavirus sequences detected in *Mus minutoides* individuals cluster with the known *L* fragment and the newly sequenced *NP* fragment from the Tanzanian mammarenavirus carried by an *M. minutoides* individual (TZ22285) from Goüy de Bellocq et al. (2010) with posterior probabilities (PP) of 0.85 and 1, respectively (Fig. 2). These sequences then cluster with Lunk virus carried by an *M. minutoides* individual from Zambia (PP = 1), their

Table 2
Cytb pairwise K2P-distances (%) for *Mus minutoides* and *Grammomys surdaster*. For *G. surdaster* K2P-distances for (nearly complete) mitochondrial sequences are given within parentheses. Pairwise distances within the same mitochondrial clade are shaded. Incomplete *cytb* sequences are marked with a plus after their accession number. All individuals are known mammarenavirus-positive individuals except for the individual marked with an asterisk.

Mus minutoides	CZC10 LR34 _ZA	KD 42 _W	SLE 00094 _W	SLE 00127 _W	SLE 00207 _W	SLE 0036 _W	SLE 00414 _W	TZ 23131 _SE	TZ 23187 _SE
CZC10LR34_ZA_AB731581+									
KD42_W_EU603965+	7.7								
SLE00094_W_OL441084	7.3	0.7							
SLE00127_W_OL441085	7.3	0.7	0.0						
SLE00207_W_OL441086	7.0	0.0	0.7	0.7					
SLE00367_W_OL441087	7.3	0.7	0.2	0.2	0.7				
SLE00414_W_OL441088	7.2	0.5	0.1	0.1	0.6	0.1			
TZ23131_SE_ OP764674	6.7	5.2	5.7	5.7	5.0	5.7	5.6		
TZ23187_SE_ OP764675	6.9	5.5	5.9	5.9	5.2	5.9	5.8	0.7	
TZ28088_SE_ OP764673	6.9	6.1	6.2	6.2	5.5	6.2	6.1	2.1	2.4

Grammomys surdaster	13ZR68_su7	RS0756_su7	TZ31109_su9
13ZR68_su7_AB972437+			
RS0756_su7_MN807579*	0.7		
TZ31109_su9_ OP764676	5.7	4.9 (4.6)	
TZ33709_su10_ OP764677	6.8	6.6 (4.7)	4.7 (4.2)

closest PASC hit (Table 1). In the *NP* and *GPC* trees Ngerengere and Lunk cluster with Kodoko virus carried by *M. minutoides* individuals in Western Africa before clustering with Ryukyu virus from a *Mus caroli* individual and LCMV mostly carried by *Mus musculus* individuals. In the *L* gene tree Ngerengere and Lunk cluster with LCMV and Ryukyu virus before clustering with Kodoko virus (Fig. 2). The two fully sequenced mammarenavirus segments from Tanzanian *M. minutoides* individuals are 83–88% identical to each other and 69–77% identical to Lunk virus (Table 1). Gene pairwise identities with PASC hits are summarised in Table 1 and p-distances between all sequences of the phylogenetic tree alignment in Supplementary Material 2.

The mammarenaviruses from the Tanzanian *G. surdaster* individuals form a highly supported (PP: 0.96–1; Fig. 2) monophyletic clade with mammarenaviruses carried by other rodents of the Arvicanthini tribe: Solwezi virus from a *G. surdaster* su7 individual, Kitale virus from a *Grammomys macmillani* individual, Mafiga virus carried by *Lemniscomys rosalia* A individuals, Ippy virus from an *Arvicanthis* sp. individual, and Mariental from a *Micaelamys namaquensis* individual. In the *L* and *NP* trees, the mammarenavirus from the *G. surdaster* su9 individual clusters with Solwezi virus (PP = 1), its best PASC hit for both segments (Table 1), while in the *GPC* tree it clusters with Ippy and Kitale virus (PP = 0.99). *GPC* p-distances indicate that its *GPC* is more similar to that of Kitale, but less similar to that of the Ippy than to those of Solwezi (and Mafiga) virus (Supplementary Material 2). The mammarenavirus carried by the su10 individual always clusters with Mariental virus (PP = 1), its best PASC hit for the S segment.

4. Discussion

In our screening mammarenavirus RNA was detected in six murine non-*Mastomys natalensis* samples. As reported previously, Gairo, Morogoro and Luna viruses were detected in samples of the most abundant species of the same field sampling, *Mastomys natalensis* (Cuypers et al., 2020; Gryseels et al., 2015), but none of these mammarenaviruses were detected in samples from other species in this study. We thus, despite large sampling effort, did not detect current spill-over or additional hosts of mammarenaviruses carried by *M. natalensis* B individuals into other species, as have been reported for the mammarenavirus Lassa virus carried by *M. natalensis* A-I individuals in Western Africa (Andersen et al., 2015; Olayemi et al., 2016a) and Dhati Welel virus carried by *M. natalensis* A-III individuals in Ethiopia (Goüy de Bellocq et al., 2020). Rather, we detected mammarenaviruses that appear specific to other rodents. The mammarenaviruses carried by *Lemniscomys rosalia* A and *Mus minutoides* SE individuals had been detected before in individuals assigned to the same host taxa (Goüy de Bellocq et al., 2010), so we can now confirm their occurrence in those hosts. The two viruses detected in *Grammomys surdaster* individuals are reported for the first time.

4.1. Mammarenaviruses in Grammomys surdaster individuals

The viruses detected in *Grammomys surdaster* individuals carried mitochondria from the su9 and su10 clades (*sensu* Bryja et al., 2017, Fig. 3). Previously Solwezi virus was already discovered in a *Grammomys surdaster* su7 individual in Zambia (Ishii et al., 2016; Těšíková et al., 2021). (Nearly) complete mitochondrial and *cytb* sequences indicate a divergence between the host clades of about 4–5% and 5–7%, respectively, a genetic distance frequently observed among closely related sister species of murine rodents (e.g. Nicolas et al., 2012). Based on unpublished genome-scale ddRAD data (Bryja et al., in preparation), these three mitochondrial clades hosting mammarenaviruses correspond to distinct genomic clades. The su9 and su10 genomic clades are more closely related to each other than to the genomic clade that includes individuals with su7 mt (as well as individuals carrying several other mitochondrial clades, see Fig. 3). Depending on the method, the genomic clade that includes su7 individuals can be considered a separate Molecular Operational Taxonomic Unit (MOTU) from the MOTU that includes the su9 and su10 genomic clades.

By these measures, the three mammarenaviruses carried by *G. surdaster* individuals appear to be carried by distinct hosts, a criterion used by ICTV to demarcate mammarenavirus species. These host taxa and therefore likely their viruses, occur in distinct geographic areas, a second criterion. The host ranges are generally parapatric, though there

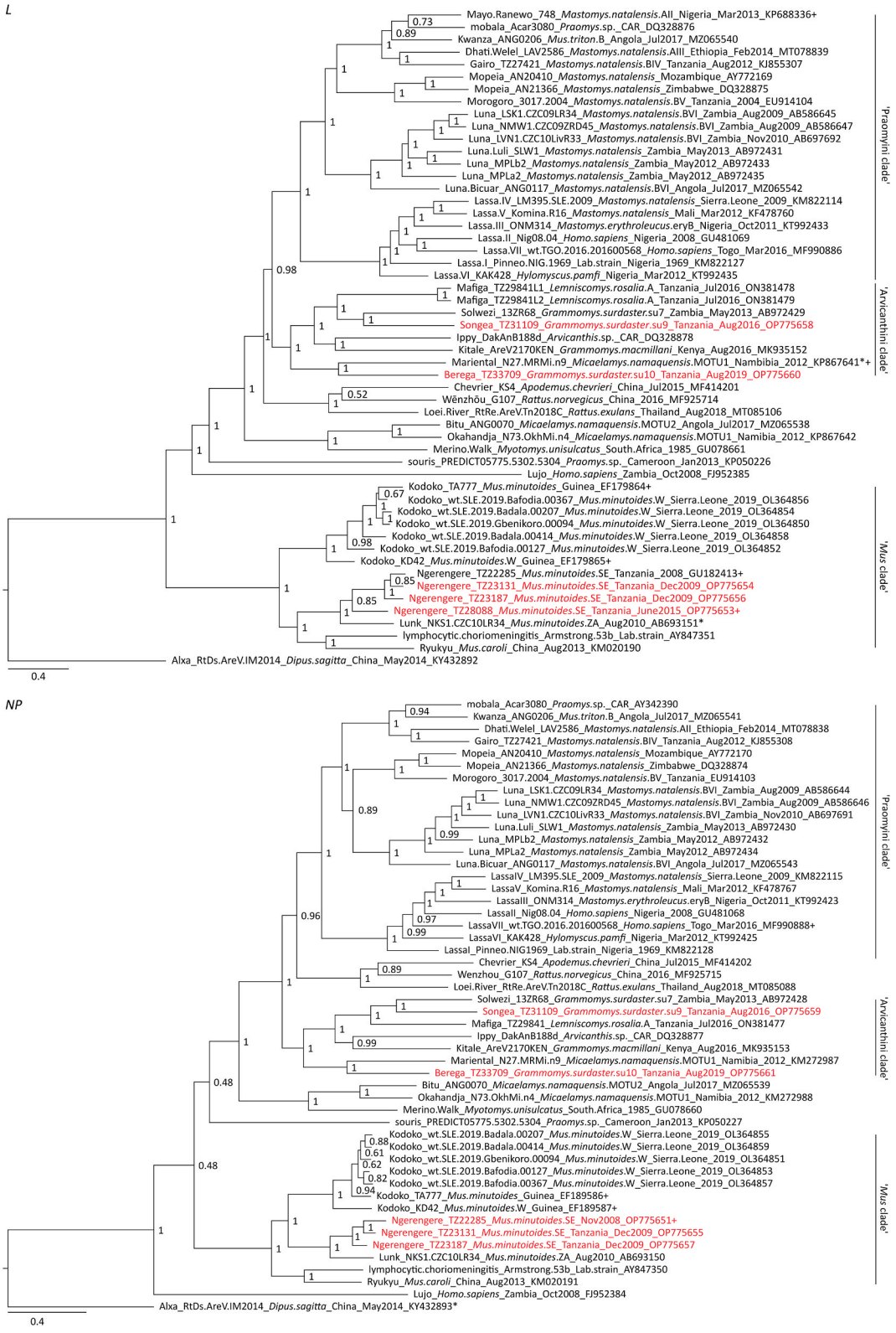

Fig. 2. *L*, *NP* and *GPC* Bayesian phylogenetic trees. Asterisks indicate sequences that were slightly modified from the corresponding GenBank sequences to improve translation alignment (see text). Plus signs at the end of sequence names indicate sequences which are not complete. New sequences are indicated in red. The scale bars indicate 0.4 nt substitutions per site. Numbers at the nodes represent posterior probabilities.

is likely some range overlap of the su9 taxon and the taxon that includes su7 individuals in the area around the Livingstone Mountains (Cuypers et al., 2022b; Fig. 3). The criteria for shared sequence identity are also met (<76%/<80%/<88% for L segment nucleotide, S segment nucleotide and NP amino acid sequences, respectively). We therefore name the mammarenavirus detected in the *G. surdaster* su9 individual Songea

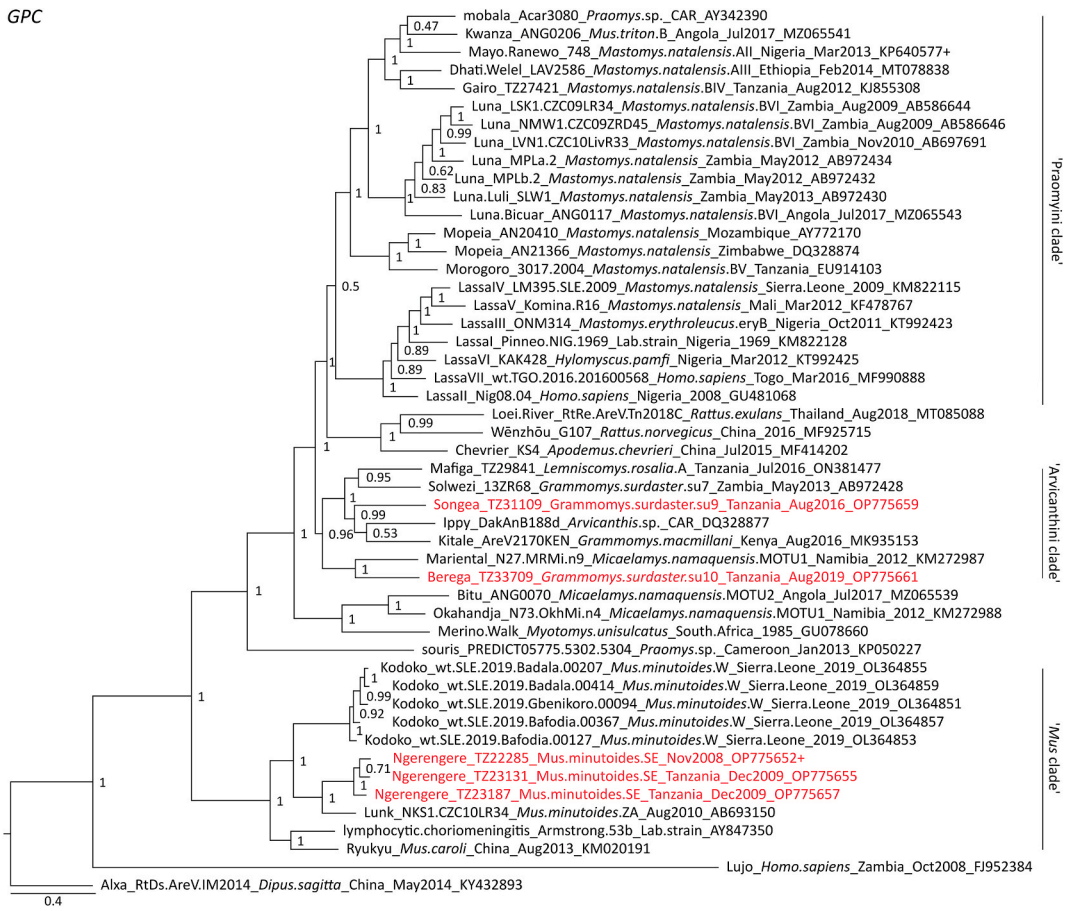

Fig. 2. (*continued*).

mammarenavirus (SOGV), after the Songea Rural District where it was detected. It is 69.6%/62.8%/82.4% identical to its closest relative Solwezi virus, for the L segment, S segment and NP amino acid sequence, respectively (Table 1). We name the mammarenavirus detected in a *G. surdaster* su10 individual Berega mammarenavirus (BEGV), after the locality where it was discovered. It is 69.8%/78.0% identical to Mariental virus, found in a *Micaelamys namaquensis* individual in Namibia for the S segment nucleotide and NP amino acid sequence, respectively (Table 1). However, Mariental virus does not appear in the top 15 hit results for the L segment in the PASC webtool. It does consistently cluster with Mariental virus in our phylogenetic trees (Fig. 2), including for the *L* gene which makes up most of the L segment. This discrepancy is likely because we masked a suspected chimaeric region of 870 nt for the Mariental *L* gene, and such a masking feature is absent in the PASC webtool.

4.2. Mammarenaviruses in Mus minutoides individuals

The mammarenaviruses detected in the *Mus minutoides* individuals resemble the mammarenavirus previously detected in a Tanzanian *M. minutoides* individual of which a 340 nt *L* gene fragment was described in Goüy de Bellocq et al. (2010). We name these viruses Ngerengere mammarenavirus (NGEV) after the river near its first sampling locality in Morogoro. Based on the TZ23131 and TZ23187 virus genomes, it meets the criteria to be recognised as a new mammarenavirus species, except for the NP AA sequence identity which is above 88% (92.1–92.7% identical to their closest relative Lunk virus). L segment and S segment nucleotide identities are under the required 76% (69–70%) and 80% (77%). We were not able to sequence the mammarenavirus genome of TZ28088. The very short *L* fragment is more intermediate between the Zambian Lunk virus and the other Tanzanian Ngerengere viruses and clusters with the Ngerengere viruses with a lower posterior probability of 0.85 in a basal position in the *L* gene phylogenetic tree (Fig. 2). However, the sequence only contains 291 non-ambiguous nucleotides in a very conserved region of the genome and its amino acid sequence (which is less prone to saturation) is identical to the mammarenavirus sequence of TZ23187. More research is needed to determine to what extent Ngerengere and Lunk represent distinct evolutionary clades, ideally in the contact zone of the SE and ZA host clades, as Ngerengere has been found in hosts with SE mt and Lunk in an individual with ZA mt. Mean K2P distances ranged from 0 to 2% within-mitochondrial clades and from 5 to 8% between-mitochondrial clades for *Mus minutoides* hosts (Table 2). The latter are larger than the K2P distances reported in Bryja et al. (2014), ranging from 1 to 4% over all mitochondrial clade comparisons, perhaps due to a lower number of samples included in our analysis.

4.3. Host specificity, taxonomy and spill-over/host jumps

The strongest evidence for host specificity from nature is when taxa are shown to meet but not exchange distinct viruses they carry (the case for Morogoro and Gairo virus in the *Mastomys natalensis* B-IV and B–V taxa). While our current data for mammarenavirus-positive *Grammomys surdaster* or *Mus minutoides* individuals does not have this extreme level of geographic resolution, it appears that, like *M. natalensis*, *Grammomys surdaster* and *Mus minutoides* carry distinct mammarenaviruses in distinct subspecific host taxa in different geographic regions. They therefore join a growing list of mammarenaviruses consistent with this pattern. Most recently it has been shown four distinct LCMV clades can be distinguished: clade I comprises infections that can be linked to

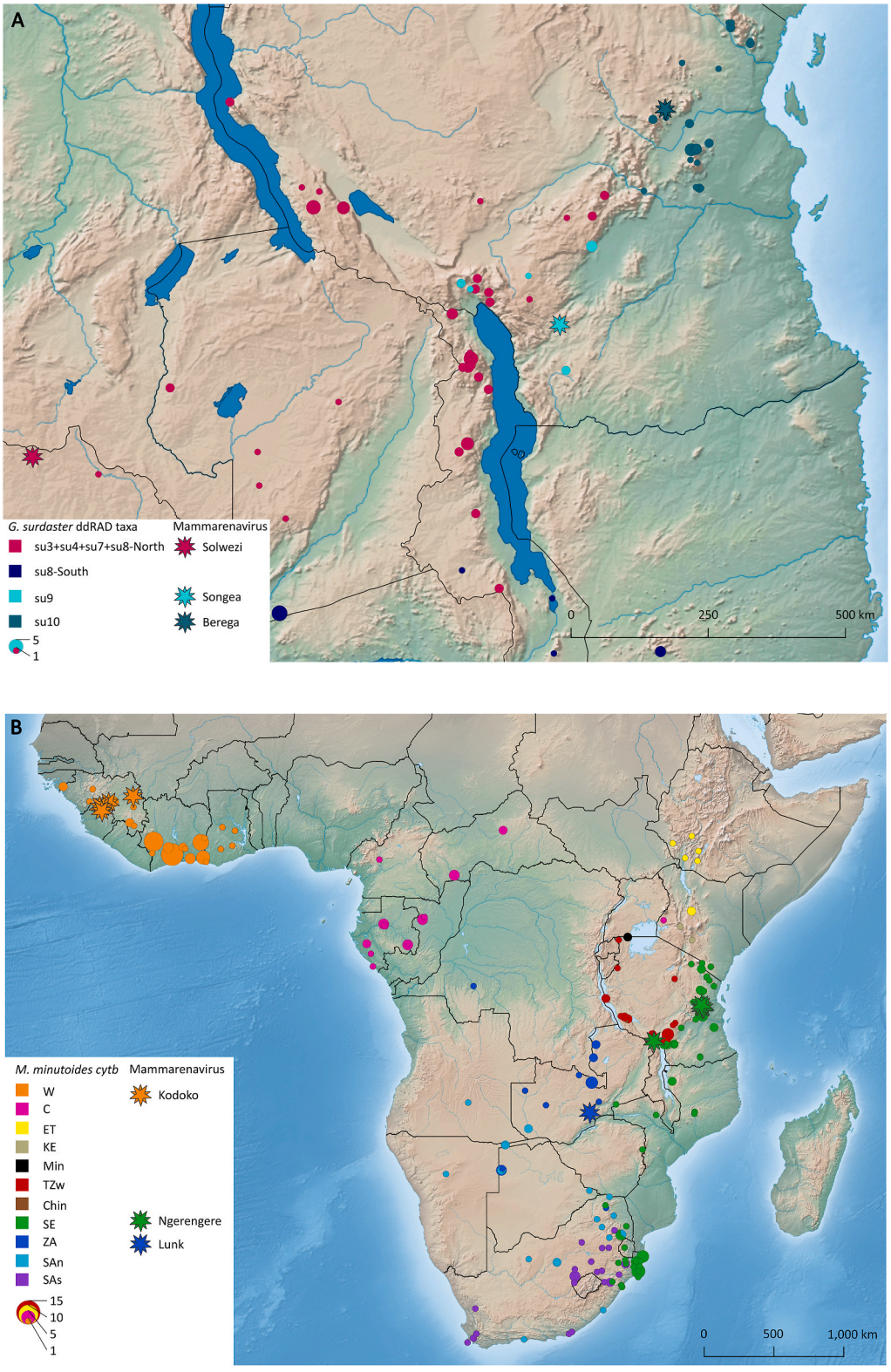

Fig. 3. *Grammomys surdaster* (A), *Mus minutoides* (B) and the mammarenaviruses they carry. *G. surdaster* rather occurs in higher-elevation miombo woodlands, montane forests and their margins, as opposed to *M. natalensis* and *M. minutoides*, which prefer lowland savannah-like habitats. Host pie charts are scaled to the number of individuals sampled. As in Cuypers et al. (2022b), *G. surdaster* mitochondrial clades are grouped based on unpublished ddRAD data (Bryja et al., in preparation): *G. surdaster* su7 individuals form a single genomic clade with individuals carrying su3 and su4 and with the northern individuals carrying su8 mitochondria, and individuals from the northernmost locality with su9 mitochondria belong to the genomic clade with su10 individuals rather than to the genomic clade with the other su9 individuals.

M. musculus domesticus, clade II infections that can be linked to *M. musculus musculus*, clade III a single virus from a human infection, and clade IV from infected *Apodemus sylvaticus* individuals (Fornůsková et al., 2021). Subspecies of *M. musculus* therefore harbour distinct LCMV despite forming an extensive contact zone in Europe (Baird and Macholán, 2012).

Taxonomic uncertainties impose a limitation on interpretation of the current study. Except for Ngerengere virus, the viruses concerned all meet the ICTV criteria to be considered different viral species. However, despite meeting all these criteria, the decision to recognise Morogoro virus as a new species was postponed when the current sequence identity cut-offs were introduced (Radoshitzky et al., 2015). The cut-offs were chosen to cause minimal changes and disruption to the arenavirus classification at that time (Radoshitzky et al., 2015) and differ from

species delimitation criteria used in other virus families, which also vary considerably. The notion of a species is less arbitrary for their sexually reproducing hosts, but still attracts a wide variety of opinions. Species and subspecific taxa may therefore not be comparable among small mammal hosts as species can be more 'clumped' in some groups and more 'split' in others. In that case the words used for the host taxonomic rank relevant to virus structure will differ. As taxonomic reviews are outside of our scope, we have used current small mammal taxonomical designations for simplicity and do not taxonomically rank subspecific levels. A molecular taxonomic revision of the small mammals surveyed here could potentially upgrade some of our subspecific taxa to species status in the future as the repeated patterns of subspecific genome-wide small mammal divergence maintained in the face of geneflow that are now coming to light are consistent with operational definitions at the species level following (Mallet, 1995). Likewise, mammarenavirus species delimitation criteria may evolve in the future, certainly given that several mammarenaviruses such as Lassa prove difficult to classify with the current criteria.

Regardless of the host and virus taxonomic labels, host genetic structure appears to impact mammarenavirus diversity and distribution. For example, all Ngerengere and Lunk viruses were detected in localities where different mammarenaviruses have been detected in *M. natalensis* (Figs. 1 and 3). Further, numerous mammarenavirus pairs have now been shown to geographically co-occur. In the locality Lusaka East, Lunk virus co-occurs with Luna virus (Ishii et al., 2012); in the localities Morogoro and Mkundi, Ngerengere virus co-occurs with Morogoro virus (Goüy de Bellocq et al., 2010; Gryseels et al., 2017; Supplementary Material 1); and in the locality Ngana, Ngerengere virus co-occurs with Luna virus (Cuypers et al., 2020; Supplementary Material 1). While *M. natalensis* and *M. minutoides* are both savannah-dwelling species with similar phylogeographic patterns around Tanzania, the pattern is somewhat different in southwestern Tanzania (Cuypers et al., 2022b). Genetic groups of both species are separated by the same landscape features (the Livingstone Mountains, the Eastern Arc Mountains, the Ufipa Plateau and Lake Malawi), but in Ngana on the Kyela Plain at the base of the Livingstone Escarpment, *M. natalensis* B-VI are found as in Lusaka rather than *M. natalensis* B–V as in Morogoro and Mkundi, while *M. minutoides* SE are found as in Morogoro and Mkundi rather than *M. minutoides* ZA as in Lusaka (Figs. 1 and 3; see Supplementary Material 3 for more detail). Based on the mt clade of the hosts, it is thus possible to predict the presence of Luna/Ngerengere virus in Ngana rather than other mammarenaviruses carried by *M. minutoides/M. natalensis* individuals.

Mammarenavirus-subspecific host taxon associations do not necessarily imply that these mammarenaviruses co-diverged with their subspecific host taxa. In the *L* gene tree, the mammarenaviruses carried by *Mus minutoides* individuals do not form a monophyletic clade, nor do the mammarenaviruses carried by *M. natalensis* and *G. surdaster* individuals form monophyletic clades in any of the trees (Fig. 2). It is clear that several host jumps have occurred during the evolutionary history of mammarenaviruses (Cajimat et al., 2007; Kronmann et al., 2013; Milazzo et al., 2015) and more will likely become apparent as more mammarenaviruses are being discovered.

Host jumps can be constrained by phylogenetic and geographic distance of the hosts. In New World mammarenaviruses host switching appeared to be driven by host relatedness in one clade, but by geographic proximity in three other clades (Irwin et al., 2012). Host geographic range overlap may have facilitated a host jump in a mammarenavirus carried by *G. surdaster* individuals mammarenavirus: Songea virus is more closely related to Solwezi than to Berega virus, and the *G. surdaster* su9 range overlaps more with that of the taxon that includes su7 individuals than with that of su10, even though su9 is more closely related to su10. However, host relatedness appears a major constraining factor as well, as most of the host jumps appear to have occurred among hosts of the same rodent tribe (Fig. 2): the 'Arvicanthini clade' consists exclusively of viruses carried by Arvicanthini hosts (but does not contain Bitu and Okahandja); most of the viruses carried by Praomyini (except Souris, but there is no information available on how the *Praomys* sp. host was identified) cluster together; and most viruses carried by Murini (more specifically *Mus*) cluster together, including Ryukyu virus from a different continent. In that regard it is remarkable that several Praomyini to *Mus* host jumps appear to have taken place in Western Africa and Angola, resulting in a mobala-like virus in *Mus triton*, Kwanza (Těšíková et al., 2021, Fig. 2), and Lassa-like viruses in *Mus baoulei*, Jirandogo (Kronmann et al., 2013), and *Mus setulosus*, Gbagroube (Coulibaly-N'Golo et al., 2011). (For the latter two only partial sequences are available, therefore they are not included in the phylogenetic trees in our study.)

Despite evidence of past host jumps in the phylogenetic tree, and large sample size, we did not detect any current spill-over of mammarenaviruses carried by *M. natalensis* individuals into our non-*Mastomys natalensis* individuals, nor was spill-over of the mammarenaviruses characterised in this study detected in *M. natalensis* individuals sampled at the same sites (Cuypers et al., 2020; Gryseels et al., 2017). As they were sampled at the same sites, generally inside or at the edges of agricultural fields, and as several sampled species share similar habitats (e.g. *M. minutoides* and *M. natalensis* are both grassland species that also thrive in agricultural landscapes), the absence of detected spill-over does not appear to be due to a lack of ecological opportunity.

5. Conclusion

Mammarenaviruses are known for their strong host specificity. Spill-over and host jumps can thus be expected to be relatively rare. Indeed, we were not able to detect any current spill-over in 1225 Tanzanian samples, but from phylogenetic reconstruction we can see that host jumps have occurred in the past, though mostly within the same rodent tribe. Furthermore, subspecific host genetic structure appears to shape mammarenavirus diversity and distribution ranges as different subspecific taxa of *Mastomys natalensis*, *Mus minutoides* and *Grammomys surdaster* that occur in different geographic regions appear to carry distinct mammarenaviruses.

CRediT authorship contribution statement

Laura N. Cuypers: Formal analysis, Investigation, Writing – original draft, Visualization. **Sophie Gryseels:** Resources, Writing – review & editing, Supervision, Funding acquisition. **Natalie Van Houtte:** Investigation. **Stuart J.E. Baird:** Conceptualization, Writing – review & editing. **Christopher A. Sabuni:** Resources, Writing – review & editing. **Abdul S. Katakweba:** Resources, Writing – review & editing. **Sebastiaan R.M. van den Burg:** Investigation. **Josef Bryja:** Resources, Writing – review & editing. **Herwig Leirs:** Resources, Writing – review & editing, Funding acquisition. **Joëlle Goüy de Bellocq:** Conceptualization, Supervision, Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank S. Olschläger for his help in characterising parts of the TZ22285 Ngerengere virus and T. Aghová, E. Bendová, A. Hánová, Z. Hiadlovská, R. Idd, K. Kibwana, J. Komárková, A. Konečný, J. Krásová, J. Kreisinger, S. Lutea, R. Makundi, V. A. Massawe, Mazoch, G. Mphamphi, O. Mikula, C. Pangapanga, H. Patzenhauerová-Konvičková, G. Sabuni, F. Sedláček, R. Smolinksy and R. Šumbera for their help in the field. This study was supported by the Czech Science Foundation (18-19629S) and the Research Foundation—Flanders (FWO; G0A4815N).

This collaborative work was also part of the Czech Republic—Flanders, Belgium Mobility projects 2020 (FWO no. VS07521 N and CAS no. FWO-21-02). Computational resources were supplied by the project 'e-Infrastruktura CZ' (e-INFRA CZ LM2018140) supported by the Ministry of Education, Youth and Sports of the Czech Republic. LC was a Ph.D. fellow with the FWO (1171921N).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virol.2023.02.014.

References

Abudurexiti, A., Adkins, S., Alioto, D., Alkhovsky, S.V., Avšič-Županc, T., Ballinger, M.J., Bente, D.A., Beer, M., Bergeron, É., Blair, C.D., Briese, T., Buchmeier, M.J., Burt, F.J., Calisher, C.H., Cháng, C., Charrel, R.N., Choi, I.R., Clegg, J.C.S., de la Torre, J.C., et al., 2019. Taxonomy of the order *Bunyavirales*: update 2019. Arch. Virol. 164, 1949–1965. https://doi.org/10.1007/s00705-019-04253-6.

Aghová, T., Šumbera, R., Piálek, L., Mikula, O., McDonough, M.M., Lavrenchenko, L.A., Meheretu, Y., Mbau, J.S., Bryja, J., 2017. Multilocus phylogeny of East African gerbils (Rodentia, *Gerbilliscus*) illuminates the history of the Somali-Masai savanna. J. Biogeogr. 44 (10), 2295–2307. https://doi.org/10.1111/jbi.13017.

Akpede, G.O., Asogun, D.A., Okogbenin, S.A., Dawodu, S.O., Momoh, M.O., Dongo, A.E., Ike, C., Tobin, E., Akpede, N., Ogbaini-Emovon, E., Adewale, A.E., Ochei, O., Onyeke, F., Okonofua, M.O., Atafo, R.O., Odia, I., Adomeh, D.I., Odigie, G., Ogbeifun, C., et al., 2019. Caseload and case fatality of Lassa fever in Nigeria, 2001-2018: a specialist center's experience and its implications. Front. Public Health 7, 2001–2018. https://doi.org/10.3389/fpubh.2019.00170.

Andersen, K.G., Shapiro, B.J., Matranga, C.B., Sealfon, R., Lin, A.E., Moses, L.M., Folarin, O.A., Goba, A., Odia, I., Ehiane, P.E., Momoh, M., England, E.M., Winnicki, S., Branco, L.M., Gire, S.K., Phelan, E., Tariyal, R., Tewhey, R., Omoniwa, O., et al., 2015. Clinical sequencing uncovers origins and evolution of Lassa virus. Cell 162, 738–750. https://doi.org/10.1016/j.cell.2015.07.020.

Andrews, S., 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Baird, S.J.E., Macholán, M., 2012. What can the *Mus musculus musculus/M. m. domesticus* hybrid zone tell us about speciation? In: Macholán, M., Baird, S.J.E., Munclinger, P., Piálek, J. (Eds.), Evolution of the House Mouse. Cambridge University Press, Cambridge, pp. 334–372.

Bao, Y., Chetvernin, V., Tatusova, T., 2014. Improvements to pairwise sequence comparison (PASC): a genome-based tool for virus classification. Arch. Virol. 159, 3293–3304. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4221606/.

Bausch, D.G., Mills, J.N., 2014. Arenaviruses : Lassa fever, lujo hemorrhagic fever, lymphocytic choriomeningitis, and the south American hemorrhagic fevers. In: Viral Infections of Humans. Springer, pp. 147–171. https://doi.org/10.1007/978-1-4899-7448-8.

Blasdell, K.R., Duong, V., Eloit, M., Chretien, F., Ly, S., Hul, V., Deubel, V., Morand, S., Buchy, P., 2016. Evidence of human infection by a new mammarenavirus endemic to Southeastern Asia. Elife 5, e13135. https://doi.org/10.7554/eLife.13135.

Bonhomme, C.J., Capul, A.A., Lauron, E.J., Bederka, L.H., Knopp, K.A., Buchmeier, M.J., 2011. Glycosylation modulates arenavirus glycoprotein expression and function. Virology 409 (2), 223–233. https://doi.org/10.1016/j.virol.2010.10.011.

Bryja, J., Mikula, O., Šumbera, R., Meheretu, Y., Aghová, T., Lavrenchenko, L.A., Mazoch, V., Oguge, N., Mbau, J.S., Welegerima, K., Amundala, N., Colyn, M., Leirs, H., Verheyen, E., 2014. Pan-African phylogeny of Mus (subgenus Nannomys) reveals one of the most successful mammal radiations in Africa. BMC Evol. Biol. 14, 256. https://doi.org/10.1186/s12862-014-0256-2.

Bryja, J., Šumbera, R., Kerbis Peterhans, J.C., Aghová, T., Bryjová, A., Mikula, O., Nicolas, V., Denys, C., Verheyen, E., 2017. Evolutionary history of the thicket rats (genus Grammomys) mirrors the evolution of African forests since late Miocene. J. Biogeogr. 44, 182–194. https://doi.org/10.1111/jbi.12890.

Cajimat, M.N.B., Milazzo, M.L., Hess, B.D., Rood, M.P., Fulhorst, C.F., 2007. Principal Host Relationships and Evolutionary History of the North American Arenaviruses, vol. 367, pp. 235–243. https://doi.org/10.1016/j.virol.2007.05.031.

Chen, Y.-M., Sadiq, S., Holmes, E.C., Zhang, Y.-Z., 2021a. ICTV Taxonomic Proposal 2022.003M.A.v1.Arenaviridae_1ngen_1nsp. Create one new genus including one new species in the family *Arenaviridae*. https://ictv.global/filebrowser/download/10871. (Accessed 13 March 2023).

Chen, Y.-M., Sadiq, S., Tian, J.-H., Chen, X., Lin, X.-D., Shen, J.-J., Chen, H., Hao, Z.-Y., Yang, W.-D., Zhou, Z.-C., Wu, J., Li, F., Wang, H.-W., Xu, Q.-Y., Wang, W., Gao, W.-H., Holmes, E.C., Zhang, Y.-Z., 2021b. RNA virome composition is shaped by sampling ecotype. SSRN Electron. J. https://doi.org/10.2139/ssrn.3934022.

Colangelo, P., Verheyen, E., Leirs, H., Tatard, C., Denys, C., Dobigny, G., Duplantier, J.-M., Brouat, C., Granjon, L., Lecompte, E., 2013. A mitochondrial phylogeographic scenario for the most widespread African rodent, Mastomys natalensis. Biol. J. Linn. Soc. 108, 901–916. https://doi.org/10.1111/bij.12013.

Coulibaly-N'Golo, D., Allali, B., Kouassi, S.K., Fichet-Calvet, E., Becker-Ziaja, B., Rieger, T., Ölschläger, S., Dosso, H., Denys, C., ter Meulen, J., Akoua-Koffi, C., Günther, S., 2011. Novel arenavirus sequences in Hylomyscus sp. and Mus (nannomys) setulosus from côte d'Ivoire: implications for evolution of arenaviruses in Africa. PLoS One 6 (6), e20893. https://doi.org/10.1371/journal.pone.0020893.

Cuypers, L.N., Baird, S.J.E., Hánová, A., Locus, T., Katakweba, A.S., Gryseels, S., Bryja, J., Leirs, H., Goüy de Bellocq, J., 2020. Three arenaviruses in three subspecific natal multimammate mouse taxa in Tanzania: same host specificity, but different spatial genetic structure? Virus Evolution 6 (2), veaa039. https://doi.org/10.1093/ve/veaa039.

Cuypers, L.N., Čížková, D., de Bellocq, J.G., 2022a. Co-infection of mammarenaviruses in a wild mouse, Tanzania. Virus Evolution 8 (2), veac065. https://doi.org/10.1093/ve/veac065.

Cuypers, L.N., Sabuni, C., Šumbera, R., Aghová, T., Lišková, E., Leirs, H., Baird, S.J.E., Goüy de Bellocq, J., Bryja, J., 2022b. Biogeographical importance of the Livingstone Mountains in southern Tanzania: comparative genetic structure of small non-volant mammals. Frontiers in Ecology and Evolution 9, 742851. https://doi.org/10.3389/fevo.2021.742851.

Downs, W., Anderson, C., Spence, L., Aitken, T., Greenhall, A., 1963. Tacaribe virus, a new agent isolated from *Artibeus* bats and mosquitoes in Trinidad, West Indies. Am. J. Trop. Med. Hyg. 12 (4), 640–646.

Ehichioya, D.U., Hass, M., Becker-Ziaja, B., Ehimuan, J., Asogun, D.A., Fichet-Calvet, E., Kleinsteuber, K., Lelke, M., ter Meulen, J., Akpede, G.O., Omilabu, S.A., Günther, S., Ölschläger, S., 2011. Current molecular epidemiology of Lassa virus in Nigeria. J. Clin. Microbiol. 49 (3), 1157–1161. https://doi.org/10.1128/JCM.01891-10.

Fehling, S.K., Lennartz, F., Strecker, T., 2012. Multifunctional nature of the arenavirus RING finger protein Z. Viruses 4 (11), 2973–3011. https://doi.org/10.3390/v4112973.

Ferron, F., Weber, F., de la Torre, J.C., Reguera, J., 2017. Transcription and replication mechanisms of Bunyaviridae and Arenaviridae L proteins. Virus Res. 234, 118–134. https://doi.org/10.1016/j.virusres.2017.01.018.

Fornůsková, A., Hiadlovská, Z., Macholán, M., Piálek, J., Goüy de Bellocq, J., 2021. New perspective on the geographic distribution and evolution of lymphocytic Choriomeningitis virus, Central Europe. Emerg. Infect. Dis. 27 (10), 2638.

Fulhorst, C.F., Bowen, M.D., Salas, R.A., Duno, G., Utrera, A., Ksiazek, T.G., De Manzione, N.M.C., De Miller, E., Vasquez, C., Peters, C.J., Tesh, R.B., 1999. Natural rodent host associations of Guanarito and Pirital viruses (family *Arenaviridae*) in central Venezuela. Am. J. Trop. Med. Hyg. 61 (2), 325–330.

Giorello, F.M., D'Elía, G., Lessa, E.P., 2021. Genomic footprints of Quaternary colonization and population expansion in the Patagonian-Fuegian region rules out a separate southern refugium in Tierra del Fuego. J. Biogeogr. 48 (10), 2656–2670. https://doi.org/10.1111/jbi.14231.

Góes, L.G.B., Fischer, C., Campos, A.C.A., de Carvalho, C., Moreira-Soto, A., Ambar, G., da Rosa, A.R., de Oliveira, D.C., Jo, W.K., Cruz-Neto, A.P., Pedro, W.A., Queiroz, L. H., Minoprio, P., Durigon, E.L., Drexler, J.F., 2022. Highly diverse arenaviruses in neotropical bats, Brazil. Emerg. Infect. Dis. 28 (12), 2528–2533. https://doi.org/10.3201/eid2812.220980.

Gonzalez, J.-P., Sanchez, A., Rico-Hesse, R., 1995. Molecular phylogeny of Guanarito virus, an emerging arenavirus affecting humans. Am. J. Trop. Med. Hyg. 53 (1), 1–6. https://doi.org/10.4269/ajtmh.1995.53.1.

Goüy de Bellocq, J., Borremans, B., Katakweba, A., Makundi, R., Baird, S.J.E., Becker-Ziaja, B., Günther, S., Leirs, H., 2010. Sympatric occurrence of 3 arenaviruses, Tanzania. Emerg. Infect. Dis. 16 (4), 692–695. https://doi.org/10.3201/eid1604.091721.

Goüy de Bellocq, J., Bryjová, A., Martynov, A.A., Lavrenchenko, L.A., 2020. Dhati Welel virus, the missing mammarenavirus of the widespread Mastomys natalensis. Journal of Vertebrate Biology 69 (2), 20018. https://doi.org/10.25225/jvb.20018.

Gryseels, S., Baird, S.J.E., Borremans, B., Makundi, R., Leirs, H., Goüy de Bellocq, J., 2017. When viruses don't go viral: the importance of host phylogeographic structure in the spatial spread of arenaviruses. PLoS Pathog. 13 (1), e1006073 https://doi.org/10.1371/journal.ppat.1006073.

Gryseels, S., Rieger, T., Oestereich, L., Cuypers, B., Borremans, B., Makundi, R., Leirs, H., Günther, S., Goüy de Bellocq, J., 2015. Gairo virus, a novel arenavirus of the widespread Mastomys natalensis: genetically divergent, but ecologically similar to Lassa and Morogoro viruses. Virology 476, 249–256. https://doi.org/10.1016/j.virol.2014.12.011.

Günther, S., Hoofd, G., Charrel, R., Röser, C., Becker-Ziaja, B., Lloyd, G., Sabuni, C., Verhagen, R., van der Groen, G., Kennis, J., Katakweba, A., Machang'u, R., Makundi, R., Leirs, H., 2009. Mopeia virus-related arenavirus in natal multimammate mice, Morogoro, Tanzania. Emerg. Infect. Dis. 15 (12), 2008–2012. https://doi.org/10.3201/eid1512.090864.

Hánová, A., Konečný, A., Nicolas, V., Denys, C., Granjon, L., Lavrenchenko, L.A., Šumbera, R., Mikula, O., Bryja, J., 2021. Multilocus phylogeny of African striped grass mice (*Lemniscomys*): stripe pattern only partly reflects evolutionary relationships. Mol. Phylogenet. Evol. 155, 107007 https://doi.org/10.1016/J.YMPEV.2020.107007.

Harding, E.F., Russo, A.G., Yan, G.J.H., Mercer, L.K., White, P.A., 2022. Revealing the uncharacterised diversity of amphibian and reptile viruses. ISME Communications 2, 95. https://doi.org/10.1038/s43705-022-00180-x.

Hastie, K.M., Kimberlin, C.R., Zandonatti, M.A., MacRae, I.J., Saphire, E.O., 2011. Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3′ to 5′ exonuclease activity essential for immune suppression. Proc. Natl. Acad. Sci. U.S.A. 108 (6), 2396–2401. https://doi.org/10.1073/pnas.1016404108.

Inizan, C.C., Cajimat, M.N.B., Milazzo, M.L., Barragán-Gomez, A., Bradley, R.D., Fulhorst, C.F., 2010. Genetic evidence for a tacaribe serocomplex virus, Mexico. Emerg. Infect. Dis. 16 (6), 1007–1010. https://doi.org/10.3201/eid1606.091648.

Irwin, N.R., Bayerlová, M., Missa, O., Martínková, N., 2012. Complex patterns of host switching in New World arenaviruses. Mol. Ecol. 21 (16), 4137–4150. https://doi.org/10.1111/j.1365-294X.2012.05663.x.

Ishii, A., Orba, Y., Sasaki, M., Kobayashi, S., Moonga, L., Hang'ombe, B.M., Mweene, A. S., Omori, R., Ito, K., Hall, W.W., Kuhn, J.H., Bao, Y., Sawa, H., 2016. ICTV

taxonomic proposal 2016.019aM.A.v2.Mammarenavirus_sp. create species Solwezi mammarenavirus in the genus mammarenavirus, family Arenaviridae. http://www.ictv.global/proposals-16/2016.019aM.A.v2%0A.Mammarenavirus_sp.pdf.

Ishii, A., Thomas, Y., Moonga, L., Nakamura, I., Ohnuma, A., Hang'ombe, B.M., Takada, A., Mweene, A.S., Sawa, H., 2012. Molecular surveillance and phylogenetic analysis of Old World arenaviruses in Zambia. J. Gen. Virol. 93, 2247–2251. https://doi.org/10.1099/vir.0.044099-0.

Ishii, A., Thomas, Y., Moonga, L., Nakamura, I., Ohnuma, A., Hang'ombe, B., Takada, A., Mweene, A., Sawa, H., 2011. Novel arenavirus, Zambia. Emerg. Infect. Dis. 17 (10), 1921–1924.

Jiang, H., Lei, R., Ding, S.-W., Zhu, S., 2014. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinf. 15, 182. https://doi.org/10.1186/1471-2105-15-182.

Kronmann, K.C., Nimo-Paintsil, S., Guirguis, F., Kronmann, L.C., Bonney, K., Obiri-Danso, K., Ampofo, W., Fichet-Calvet, E., 2013. Two novel arenaviruses detected in pygmy mice, Ghana. Emerg. Infect. Dis. 19 (11), 1832–1835. https://doi.org/10.3201/eid1911.121491.

Kuhn, J.H., Adkins, S., Alkhovsky, S.V., Avšič-Županc, T., Ayllón, M.A., Bahl, J., Balkema-Buschmann, A., Ballinger, M.J., Bandte, M., Beer, M., Bejerman, N., Bergeron, É., Biedenkopf, N., Bigarré, L., Blair, C.D., Blasdell, K.R., Bradfute, S.B., Briese, T., Brown, P.A., et al., 2022. 2022 taxonomic update of phylum *Negarnaviricota (Riboviria: orthornavirae)*, including the large orders *Bunyavirales* and *Mononegavirales*. Arch. Virol. 167 (12), 2857–2906. https://doi.org/10.1007/s00705-022-05546-z.

Laetsch, D.R., Blaxter, M.L., 2017. BlobTools: interrogation of genome assemblies. F1000Research 6, 1287. https://doi.org/10.12688/f1000research.12232.1.

Lavergne, A., de Thoisy, B., Donato, D., Guidez, A., Matheus, S., Catzeflis, F., Lacoste, V., 2015. Patawa virus, a new arenavirus hosted by forest rodents in French guiana. EcoHealth 12, 339–346. https://doi.org/10.1007/s10393-014-0971-6.

Lecompte, É., Granjon, L., Kerbis Peterhans, J., Denys, C., 2002. Cytochrome b-based phylogeny of the Praomys group (Rodentia, Murinae): a new African radiation? C. R. Biologies 325, 827–840.

Lecompte, E., ter Meulen, J., Emonet, S., Daffis, S., Charrel, R.N., 2007. Genetic identification of Kodoko virus, a novel arenavirus of the African pigmy mouse (*Mus Nannomys minutoides*) in West Africa. Virology 364 (1), 178–183. https://doi.org/10.1016/j.virol.2007.02.008.

Maes, P., Alkhovsky, S.V., Bào, Y., Beer, M., Birkhead, M., Briese, T., Buchmeier, M.J., Calisher, C.H., Charrel, R.N., Choi, I.R., Clegg, C.S., de la Torre, J.C., Delwart, E., DeRisi, J.L., Di Bello, P.L., Di Serio, F., Digiaro, M., Dolja, V.V., Drosten, C., et al., 2018. Taxonomy of the family *Arenaviridae* and the order *Bunyavirales*: update 2018. Arch. Virol. 163, 2295–2310. https://doi.org/10.1007/s00705-018-3843-5.

Mallet, J., 1995. A species definition for the modern synthesis. Trends in Ecol. Evol. 10 (7), 294–299. https://doi.org/10.1016/0169-5347(95)90031-4.

Milazzo, M.L., Cajimat, M.N.B., Mauldin, M.R., Bennett, S.G., Hess, B.D., Rood, M.P., Conlan, C.A., Nguyen, K., Wekesa, J.W., Ramos, R.D., Bradley, R.D., Fulhorst, C.F., 2015. Epizootiology of tacaribe serocomplex viruses (*Arenaviridae*) associated with neotomine rodents (cricetidae, neotominae) in southern California. Vector Borne Zoonotic Dis. 15 (2), 156–166. https://doi.org/10.1089/vbz.2014.1625.

Mizerovská, D., Nicolas, V., Demos, T.C., Akaibe, D., Colyn, M., Denys, C., Kaleme, P.K., Katuala, P., Kennis, J., Kerbis Peterhans, J.C., Laudisoit, A., Missoup, A.D., Šumbera, R., Verheyen, E., Bryja, J., 2019. Genetic variation of the most abundant forest-dwelling rodents in Central Africa (*Praomys jacksoni* complex): evidence for Pleistocene refugia in both montane and lowland forests. J. Biogeogr. 46 (7), 1466–1478. https://doi.org/10.1111/JBI.13604.

Morin, B., Coutard, B., Lelke, M., Ferron, F., Kerber, R., Jamal, S., Frangeul, A., Baronti, C., Charrel, R., de Lamballerie, X., Vonrhein, C., Lescar, J., Bricogne, G., Günther, S., Canard, B., 2010. The N-terminal domain of the arenavirus L protein is an RNA endonuclease essential in mRNA transcription. PLoS Pathog. 6 (9), e1001038. https://doi.org/10.1371/journal.ppat.1001038.

Musser, G.G., Carleton, M.D., 2005. Superfamily Muroidea. In: Wilson, D.E., Reeder, D. M. (Eds.), Mammal Species of the World: a Taxonomic and Geographic Reference. Johns Hopkins University Press, Baltimore, pp. 894–1531.

Nicolas, V., Fabre, P.H., Bryja, J., Denys, C., Verheyen, E., Missoup, A.D., Olayemi, A., Katuala, P., Dudu, A., Colyn, M., Kerbis Peterhans, J., Demos, T., 2020. The phylogeny of the African wood mice (Muridae, *Hylomyscus*) based on complete mitochondrial genomes and five nuclear genes reveals their evolutionary history and undescribed diversity. Mol. Phylogenet. Evol. 144, 106703 https://doi.org/10.1016/J.YMPEV.2019.106703.

Nicolas, V., Schaeffer, B., Missoup, A.D., Kennis, J., Colyn, M., Denys, C., Tatard, C., Cruaud, C., Laredo, C., 2012. Assessment of three mitochondrial genes (16S, Cytb, CO1) for identifying species in the Praomyini tribe (Rodentia: Muridae). PLoS One 7 (5), e36586. https://doi.org/10.1371/journal.pone.0036586.

Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.A., 2017. MetaSPAdes: a new versatile metagenomic assembler. Genome Res. 27 (5), 824–834. https://doi.org/10.1101/gr.213959.116.

Olayemi, A., Cadar, D., Magassouba, N., Obadare, A., Kourouma, F., Oyeyiola, A., Fasogbon, S., Igbokwe, J., Rieger, T., Bockholt, S., Jérôme, H., Schmidt-Chanasit, J., Garigliany, M., Lorenzen, S., Igbahenah, F., Fichet, J.-N., Ortsega, D., Omilabu, S., Günther, S., Fichet-Calvet, E., 2016a. New hosts of the Lassa virus. Sci. Rep. 6, 25280 https://doi.org/10.1038/srep25280.

Olayemi, A., Obadare, A., Oyeyiola, A., Igbokwe, J., Fasogbon, A., Igbahenah, F., Ortsega, D., Asogun, D., Umeh, P., Vakkai, I., Abejegah, C., Pahlman, M., Becker-Ziaja, B., Günther, S., Fichet-Calvet, E., 2016b. Arenavirus diversity and phylogeography of Mastomys natalensis rodents, Nigeria. Emerg. Infect. Dis. 22 (4), 694–697. https://doi.org/10.3201/eid2204.150155.

Olayemi, A., Fichet-Calvet, E., 2020. Systematics, ecology, and host switching: attributes affecting emergence of the Lassa virus in rodents across western Africa. Viruses 12, 312. https://doi.org/10.3390/v12030312.

Ölschläger, S., Lelke, M., Emmerich, P., Panning, M., Drosten, C., Hass, M., Asogun, D., Ehichioya, D., Omilabu, S., Günther, S., 2010. Improved detection of Lassa virus by reverse transcription-PCR targeting the 5′ region of S RNA. J. Clin. Microbiol. 48 (6), 2009–2013. https://doi.org/10.1128/JCM.02351-09.

Puckett, E.E., Murphy, S.M., Bradburd, G., 2021. Phylogeographic analysis delimits three evolutionary significant units of least chipmunks in North America and identifies unique genetic diversity within the imperiled Peñasco population. Ecol. Evol. 11 (17), 12114–12128. https://doi.org/10.1002/ece3.7975.

Radoshitzky, S.R., Bào, Y., Buchmeier, M.J., Charrel, R.N., Clawson, A.N., Clegg, C.S., DeRisi, J.L., Emonet, S., Gonzalez, J.P., Kuhn, J.H., Lukashevich, I.S., Peters, C.J., Romanowski, V., Salvato, M.S., Stenglein, M.D., de la Torre, J.C., 2015. Past, present, and future of arenavirus taxonomy. Arch. Virol. 160 (7), 1851–1874. https://doi.org/10.1007/s00705-015-2418-y.

Radoshitzky, S.R., Buchmeier, M.J., Charrel, R.N., Clegg, J.C.S., Gonzalez, J.P.J., Günther, S., Hepojoki, J., Kuhn, J.H., Lukashevich, I.S., Romanowski, V., Salvato, M. S., Sironi, M., Stenglein, M.D., de La Torre, J.C., Consortium, I.R., 2019. ICTV virus taxonomy profile: *Arenaviridae*. J. Gen. Virol. 100, 1200–1201. https://doi.org/10.1099/jgv.0.001280.

Rambaut, A., 2012. FigTree v1.4. University of Edinburgh, Institute of Evolutionary Biology.

Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. Tracer v1.6. http://beast.bio.ed.ac.uk/Tracer.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61 (3), 539–542. https://doi.org/10.1093/sysbio/sys029.

Shimada, T., Aplin, K.P., Jogahara, T., Lin, L.-K., Herbreteau, V., Gonzalez, J.-P., Suzuki, H., 2007. Complex phylogeographic structuring in a continental small mammal from East Asia, the rice field mouse, *Mus caroli* (Rodentia, Muridae). Mamm. Stud. 32 (2), 49–62. https://doi.org/10.3106/1348-6160(2007)32[49:cpsiac]2.0.co;2.

Tamura, K., Stecher, G., Kumar, S., 2021. MEGA11: molecular evolutionary genetics analysis version 11. Mol. Biol. Evol. 38 (7), 3022–3027. https://doi.org/10.1093/molbev/msab120.

Těšíková, J., Krásová, J., Goüy de Bellocq, J., 2021. Multiple mammarenaviruses circulating in Angolan rodents. Viruses 13, 982. https://doi.org/10.3390/v13060982.

Trauth, M.H., Maslin, M.A., Deino, A.L., Strecker, M.R., Bergner, A.G.N., Dühnforth, M., 2007. High- and low-latitude forcing of Plio-Pleistocene East African climate and human evolution. J. Hum. Evol. 53 (5), 475–486. https://doi.org/10.1016/j.jhevol.2006.12.009.

Vieth, S., Drosten, C., Lenz, O., Vincent, M., Omilabu, S., Hass, M., Becker-Ziaja, B., ter Meulen, J., Nichol, S.T., Schmitz, H., Günther, S., 2007. RT-PCR assay for detection of Lassa virus and related Old World arenaviruses targeting the L gene. Trans. R. Soc. Trop. Med. Hyg. 101, 1253–1264. https://doi.org/10.1016/j.trstmh.2005.03.018.

Vučak, M., Bangura, J., Ghersi, B.M., Nichols, J., Hughes, J., da Silva Filipe, A., Tremeau-Bravard, A., Wolking, D.J., Amara, E., Bangura, A., Kanu, M.C., Kanu, O.T., Kargbo, D., Lavalie, E.G., Lungay, V., Robert, W., Turay, M., Fornie, S., Samba, T.T., et al., 2022. Genome sequences of five arenaviruses from pygmy mice (*Mus minutoides*) in Sierra Leone. Microbiology Resource Announcements, e0009522. https://doi.org/10.1128/mra.00095-22.

Whitton, J.L., Tishon, A., Lewicki, H., Gebhard, J., Cook, T., Salvato, M., Joly, E., Oldstone, M.B.A., 1989. Molecular analyses of a five-amino-acid cytotoxic T-lymphocyte (CTL) epitope: an immunodominant region which induces nonreciprocal CTL cross-reactivity. J. Virol. 63 (10), 4303–4310. https://doi.org/10.1128/jvi.63.10.4303-4310.1989.

Williamson, P.G., 1985. Evidence for an early plio-pleistocene rainforest expansion in East Africa. Nature 315 (6019), 487–489. https://doi.org/10.1038/315487a0.

Wu, J.-Y., Guo, C., Xia, Y., Bao, H.-M., Zhu, Y.-S., Guo, Z.-M., Wei, Y.-H., Lu, J.-H., 2021. Genomic characterization of Wenzhou mammarenavirus detected in wild rodents in Guangzhou City, China. One Health 13, 100273. https://doi.org/10.1016/j.onehlt.2021.100273.

Wulff, H., McIntosh, B.M., Hamner, D.B., Johnson, K.M., 1977. Isolation of an arenavirus closely related to Lassa virus from Mastomys natalensis in south-east Africa. Bull. World Health Organ. 55 (4), 441–444.