

2020-10

Multiple DNA viruses identified in multimammate mouse (*Mastomys natalensis*) populations from across regions of sub-Saharan Africa

Calvignac-Spencer, S

<http://hdl.handle.net/10026.1/16422>

10.1007/s00705-020-04738-9

Archives of Virology

Springer (part of Springer Nature)

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Title:

Multiple DNA viruses identified in multimammate mouse (*Mastomys natalensis*) populations from across regions of Sub-Saharan Africa

Authors: Sébastien Calvignac-Spencer¹, Léonce Kouadio^{2,3}, Emmanuel Couacy-Hymann², Nafomon Sogoba⁴, Kyle Rosenke⁵, Andrew J. Davison⁶, Fabian Leendertz³, Michael A. Jarvis⁷, Heinz Feldmann⁵, Bernhard Ehlers⁸

¹ P3 "Viral Evolution", Robert Koch-Institute, Berlin, Germany

² LANADA/Central Laboratory for Animal Diseases, Bingerville, Côte d'Ivoire

³ P3 "Epidemiology of highly pathogenic microorganisms", Robert Koch-Institute, Berlin, Germany

⁴ Malaria Research and Training Center, international Center of Excellence in Research, Faculty of Medicine and Odontostomatology, University of Sciences, Techniques and Technologies of Bamako, Bamako, Mali

⁵ Laboratory of Virology, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, MT, USA

⁶ MRC-University of Glasgow Centre for Virus Research, Glasgow, United Kingdom

⁷ School of Biomedical Sciences, University of Plymouth, Plymouth, United Kingdom; and The Vaccine Group Ltd, Plymouth, United Kingdom

⁸ Division 12 "Measles, Mumps, Rubella, and Viruses Affecting Immunocompromised Patients", Robert Koch Institut, Berlin, Germany

* Corresponding author:

Bernhard Ehlers,

Division 12 "Measles, Mumps, Rubella and Viruses Affecting Immunocompromised Patients"

Robert Koch-Institute, Berlin, Germany

ehlersb@rki.de

+49 30 18754 2347

Summary

The multimammate mouse (*Mastomys natalensis*; *M. natalensis*) serves as the main reservoir for the zoonotic arenavirus, Lassa virus (LASV), and this has led to considerable investigation into the distribution of LASV and other related arenaviruses within this host species. In contrast to the situation with arenaviruses, the presence of other viruses within *M. natalensis* remains largely unexplored. In this study, herpesviruses and polyomaviruses were identified and partially characterized by PCR methods, sequencing and phylogenetic analysis. In tissues sampled from *M. natalensis* populations in Côte d'Ivoire and Mali, 6 new DNA viruses (4 betaherpesviruses, 1 gammaherpesvirus and 1 polyomavirus) were identified. Phylogenetic analysis of the glycoprotein B amino acid sequence clustered the herpesviruses within the respective cytomegaloviruses and rhadinoviruses of multiple rodent species. PCR was able to amplify the complete circular genome of the newly identified polyomavirus. Amino acid sequence analysis of large T antigen or VP1 clustered this virus with a known polyomavirus of house mouse (species *Mus musculus polyomavirus 1*). These two polyomaviruses form a clade with other rodent polyomaviruses, and the newly identified virus represents the third known polyomavirus of *M. natalensis*. This study represents the first identification of herpesviruses and the discovery of a novel polyomavirus within *M. natalensis*. In contrast to arenaviruses, we anticipate that these newly identified viruses represent a low zoonotic risk due to the normally highly restricted specificity of members of these two DNA virus families to their individual mammalian host species.

Keywords: polyomavirus, herpesvirus, betaherpesvirus, gammaherpesvirus, multimammate mouse

INTRODUCTION

Herpesviruses (order *Herpesvirales*) and polyomaviruses (family *Polyomaviridae*) are double-stranded DNA viruses known to infect many different vertebrate species, including fish, birds and mammals [5, 8]. The multimammate mouse (*Mastomys natalensis*; *M. natalensis*) is a common rodent belonging to family Muridae, subfamily Murinae. It occupies a wide geographic range extending across the entirety of Sub-Saharan Africa [34]. The natural habitat of *M. natalensis* is equally diverse, with animals well-adapted to agro-ecosystems and cohabitation with humans [24]. *M. natalensis* is the major reservoir for the zoonotic arenavirus, Lassa virus (LASV) [31, 39]. Multiple related arenaviruses have also been identified in *M. natalensis*, including: Gairo, Luna, Mopeia and Morogoro viruses [19, 35].

Earlier studies identified herpesviruses and polyomaviruses in several rodent species [11, 14], including two polyomaviruses in *M. natalensis* corresponding to an alphapolyomavirus [14] and a betapolyomavirus [36]. However, *M. natalensis* has never been examined for the presence of herpesviruses. In the present study, we analysed tissues from *M. natalensis* collected in Côte d'Ivoire (CI) in 2014 and in Mali in 2017 for the presence of herpesviruses (CI, Mali) and polyomaviruses (Mali). Our results show that *M. natalensis* carries multiple herpesviruses, predominantly belonging to subfamily *Betaherpesvirinae* but also to subfamily *Gammaherpesvirinae*. We also identified a novel polyomavirus, representing the third polyomavirus identified in *M. natalensis*.

MATERIALS and METHODS

Sample collection

M. natalensis were live-captured, deeply anesthetized with isoflurane, bled by cardiac puncture and euthanized by cervical dislocation. All procedures on live animals were conducted in compliance with the applicable institutional and national guidelines for use and handling of animals. Tissues were immediately flash frozen and stored at -80°C or below.

Samples were confirmed to originate from *M. natalensis* by *cytochrome b* PCR, sequencing and BLAST analysis of GenBank [25].

PCR methods

DNA was extracted and tested for the presence of herpesviruses by a generic nested PCR targeting a region of the herpesvirus DNA polymerase (DPOL) gene (Figure 1A), as described previously [11]. Glycoprotein B (gB) coding sequences of viruses in subfamilies *Betaherpesvirinae* and *Gammaherpesvirinae* were also amplified by a generic nested PCR [11]. For testing for the presence of polyomaviruses, a generic polyomavirus PCR (Figure 1B) was performed that targets the major capsid VP1 gene of mammalian polyomaviruses [26, 40]. Specific long-distance (LD)-PCR (Figure 1A and Figure 1B) was carried out using the TaKaRa Ex Taq PCR Kit (Clontech, California, USA) according to manufacturer's instructions.

Availability of data

The novel sequences were deposited in GenBank under accession numbers MN417224-MN417229.

Phylogenetic analysis

For phylogenetic analysis, reference viruses were selected representing all currently recognized species in the family or subfamily, as well as additional viruses that represented distinct viral lineages discussed in the literature but not yet integrated into the official taxonomy (sensu International Committee on Taxonomy of Viruses [ICTV]). This represented 111, 21 and 40 viruses for polyoma-, betaherpes- and gammaherpesviruses, respectively. The coding sequences (CDSs) of polyomavirus large T antigen (LTA_g) and VP1 or herpesvirus DPOL and gB were extracted using Geneious v11.1.5 [23]. The CDSs were translated into amino acid sequences, which were aligned using Muscle [9] as implemented

in Seaview v4 [16]. Conserved amino acid blocks were then selected using Gblocks, with selection criteria for a less stringent selection as implemented in Seaview [42].

For each dataset, a maximum-likelihood (ML) analysis was carried out using PhyML v3 with smart model selection (PhyML-SMS) and the Bayesian information criterion and a tree search using subtree pruning and regrafting [16, 18, 27]. Branch robustness was estimated using Shimodaira-Hasegawa-like approximate likelihood ratio tests (SH-like aLRT) [1]. The polyomavirus and gammaherpesvirus ML trees were rooted with TempEst v1.5 by minimizing the variance of root-to-tip distances [37]; the betaherpesvirus ML trees were rooted using roseolovirus outgroups. Bayesian Markov chain Monte Carlo (BMCMC) analysis was then carried out using BEAST v1.10.4 [41]. For each alignment, the amino acid substitution model identified by PhyML-SMS was used with an uncorrelated relaxed clock (lognormal) model and a speciation model (birth-death) as a tree prior. The output of multiple BMCMC runs was examined for convergence and appropriate sampling of the posterior using Tracer v1.7.1 [38], before being merged using LogCombiner v1.10.4 (distributed with BEAST). The maximum clade credibility tree (MCC tree) was identified from the posterior set of trees (PST) and annotated with TreeAnnotator v1.10.4 (also distributed with BEAST). Branch robustness was estimated based on posterior probability values in the PST. The final amino acid sequence alignments comprised 264 (VP1), 523 (LTA_g), 615 (betaherpesvirus DPOL), 384 (betaherpesvirus gB), 627 (gammaherpesvirus DPOL) and 284 (gammaherpesvirus gB) amino acid residues, respectively.

RESULTS

DNA was extracted from 72 archived *M. natalensis* spleen samples (Mali) and 103 *M. natalensis* lung samples (CI; these samples had already *been* tested for the presence of LASV [25] and polyomaviruses [14]). DNAs were tested for the presence of herpesviruses by generic nested DPOL PCR, and PCR products of the predicted size were sequenced. As indicated by BLAST analysis, 45 samples were herpesvirus-positive. The sequences from 39 samples matched members of subfamily *Betaherpesvirinae* and those from 6 samples

matched members of subfamily *Gammaherpesvirinae* (Table 1). Consistent with the lack of alphaherpesviruses in other murids, an alphaherpesvirus was not detected. The betaherpesvirus sequences originated from 4 distinct viruses that were tentatively named *M. natalensis* cytomegalovirus 1 to 4 (MnatCMV1-4) on the basis of phylogenetic clustering with members of the relevant genus. The 5 gammaherpesvirus sequences were identical and represented the same virus, which was tentatively named *M. natalensis* rhadinovirus 1 (MnatRHV1).

The partial DPOL nucleic acid sequences of MnatCMV1-4 revealed pairwise identities of 69-87% (Table 1). On the basis of BLAST analysis, these viruses were most similar to known rodent cytomegaloviruses from gerbil (*Dipodillus spp.*) herpesvirus (MnatCMV1), Malayan field rat (*Rattus tiomanicus*) cytomegalovirus 1 (MnatCMV2), and wood mouse (*Apodemus sylvaticus*) cytomegalovirus 1 (MnatCMV3 and 4). The partial DPOL sequence of MnatRHV1 was most similar to a gammaherpesvirus of the house mouse (*Mus musculus* rhadinovirus 1; 55% identity). MnatCMV1, MnatCMV2, and MnatRHV1 were detected only in animals from CI, whereas MnatCMV3 and 4 were present in animals from both countries. In terms of the divergence of *M. natalensis* throughout these geographic regions, Mali and CI are believed to be represented by a single phylogenetic group (named A-I; [7]). The differences in virus distribution may therefore reflect real geographical differences or may be resolved by more extensive sampling. Overall, herpesvirus frequency in the *M. natalensis* lungs (CI) and spleens (Mali) was 34% and 13%, respectively. In the spleen samples from Mali, only MnatCMV3 and MnatCMV4 were detected, possibly suggesting distinct tissue tropism of individual viruses for lung compared to spleen.

Next, betaherpesvirus- and gammaherpesvirus-positive DNAs were tested by generic PCRs (Figure 1A) targeting the gB coding sequences of members of subfamilies *Betaherpesvirinae* and *Gammaherpesvirinae*, respectively, with sequencing of amplified products. Glycoprotein B sequences were identified for MnatCMV1 (2 samples), MnatCMV3 (5 samples) and MnatRHV1 (3 samples), but not for MnatCMV2 and 4. Based on the partial DPOL and gB sequences of MnatCMV1, MnatCMV3 and MnatRHV1, sequence-specific

primer pairs were selected for each virus from the respective gB and DPOL sequences and used in LD-PCR (Figure 1A). This resulted in amplification of a 3.2 kilobase pair (kbp) sequence. Assembly with the initial gB and DPOL sequences resulted in a 3.4 kbp contiguous sequence stretching from the 3' region of gB to the 5' region of DPOL. This approach was successful for two samples positive for MnatCMV3, and 3 samples containing MnatRHV1, but not for any of the MnatCMV1-containing samples.

Spleen DNA samples from Mali were also tested for the presence of polyomaviruses by generic PyV PCR (Figure 1B). One sample was positive, with the sequence being identified by BLAST analysis as originating from a new polyomavirus distinct from the two known polyomaviruses of *M. natalensis*. Specific nested 'back-to-back' primers were selected for the VP1 sequence, and a 5.2 kbp product was amplified by LD-PCR (Figure 1B), followed by sequencing using classical 'primer-walking'. The sequences of the initial generic PCR product and the LD-PCR product were then used to assemble a contiguous circular sequence, resulting in generation of a complete polyomavirus genome of 5338 bp. Open reading frame analysis using Geneious 11.1.5 software showed the genome to exhibit typical polyomavirus genome organization: i) an early region encoding large, middle and small T antigen CDSs, and ii) a late region on the opposite strand encoding VP1, VP2 and VP3 capsid CDSs. Early and late regions were separated by a non-coding control region (NCCR). The genome also contained a CDS encoding a putative ALTO protein [6] of 221 amino acid residues (Figure 1B).

Two *M. natalensis* polyomaviruses had been identified previously: i) a more distantly related one, named Mastomys polyomavirus (species *Mastomys natalensis polyomavirus 1*, genus *Betapolyomavirus*; accession number AB588640 [36]) from animals in Zambia, and ii) a more closely related alphapolyomavirus, named *M. natalensis polyomavirus 2* (accession number MG701350) [14] from animals in CI. We therefore tentatively named the new polyomavirus identified in the Mali animals as *M. natalensis polyomavirus 3* (MnatPyV3). The full genome of MnatPyV3 revealed a respective pairwise identity of 46% and 55% to those of

MnatPyV1 and MnatPyV2, and was most similar (66%) to the murine polyomavirus (species *Mus musculus polyomavirus 1*, genus *Alphapolyomavirus*; accession number AF442959).

As only a few polyomaviruses, all alphapolyomaviruses, have been shown to encode a middle T antigen or an ALTO protein, we compared the predicted middle T antigen CDS of MnatPyV3 with that of its closest relative, murine PyV. We found that the splice donor and acceptor sites are conserved in sequence and position, and that the encoded proteins have an amino acid sequence identity of 61%. The putative ALTO CDS of MnatPyV3 is similar in length and genomic position to the ALTO CDS of Merkel cell polyomavirus (MCPyV), with a hydrophobic motif at its C terminus similar to that of MCPyV [6]. These comparisons add strength to the prediction that MnatPyV3 encodes a middle T antigen and/or an ALTO protein.

To investigate the evolutionary position of the novel herpesviruses, we carried out phylogenetic analysis on viruses for which the 3.4 kbp sequence was available, namely MnatCMV3 and MnatRHV1. ML and BMCMC analyses performed on MnatCMV3 and known representatives of the subfamily *Betaherpesvirinae* using an alignment of glycoprotein B amino acid sequences [11] suggested that MnatCMV3 is a member of a monophyletic group of rodent CMVs, which comprises 3 species assigned by ICTV to genus *Muromegalovirus* (Figure 2; Online Resource 1). Phylogenetic analyses of an alignment of gB sequences of viruses representing subfamily *Gammaherpesvirinae* showed that MnatRHV1 is in a sister taxon to a group of rodent gammaherpesviruses (Figure 3; Online Resource 2). This group was described previously as comprising rhadinoviruses and includes the rhadinovirus of house mouse (*Mus musculus rhadinovirus 1*) [11]. However, these viruses did not form a monophyletic group with recognized rhadinovirus species [30], at least using the tree rooting that we employed (minimization of root-to-tip distance variance with TempEst [37]). Phylogenetic analyses of DPOL alignments of the beta- and gammaherpesviruses yielded similar results (Online Resource 3 and Online Resource 4 [betaherpesviruses] and Online Resource 5 and Online Resource 6 [gammaherpesviruses]).

Phylogenetic analyses based on the amino acid sequences of LTA_g of MnatPyV3 and representatives of all polyomavirus species currently recognized by the ICTV showed that MnatPyV3 is a single member of a sister taxon to the murine polyomavirus (species *Mus musculus polyomavirus 1*). The clade formed by these two viruses is itself in sistership with a monophyletic group comprising only rodent alphapolyomaviruses, including rat (species *Rattus norvegicus polyomavirus 1*) and hamster (species *Mesocricetus auratus polyomavirus 1*) polyomaviruses, and another polyomavirus infecting *M. natalensis*, MnatPyV2 (Figure 4; Online Resource 7). Phylogenetic analyses of a VP1 alignment of the same polyomaviruses supported a similar topology (Online Resource 8 and Online Resource 9).

DISCUSSION

M. natalensis is highly adapted to close co-habitation with humans. This characteristic, combined with a high frequency and extensive geographic range, makes the rodent a high-value species in terms of potential for zoonotic microbial flow to humans. *M. natalensis* carries arenaviruses throughout Sub-Saharan Africa and is known to serve as a continuing source of human LASV infection in West African countries, mainly Sierra Leone, Guinea, Liberia and Nigeria. This has led to the intensive study of LASV and other arenaviruses within this rodent species [17, 34]. However, beyond arenaviruses little is known regarding the microbes that infect these animals. *M. natalensis* was reported to be a reservoir species of *Borrelia crocidurae* [3], the etiological agent of relapsing fever, and is also known to carry *Yersinia pestis*. The latter pathogen:host interaction is believed to play a key role in the plague cycle in southern Africa [22]. Similarly, a papillomavirus (*M. natalensis* papillomavirus) has been identified in *M. natalensis*, and is believed to be causally associated with a high incidence of cutaneous skin tumors in these animals [43]. Our results are compatible with the notion that the herpesviruses and polyomaviruses identified in *M. natalensis* arose from lineages of viruses closely associated with rodents. The occurrence of multiple herpesviruses and polyomaviruses in *M. natalensis* is not surprising, as infection with multiple species of a DNA virus family is common in all host species that have been examined in any detail (e.g.

[10-13, 40]). It is tempting to interpret this pattern as suggesting that these viruses co-diverged with their hosts and are *M. natalensis*-specific, since co-divergence is an important process in shaping herpesvirus and polyomavirus evolution ([2, 4, 14, 15, 28, 33]. Therefore, despite the commensality, abundance and extensive geographic range of this rodent species, we do not expect these DNA viruses to represent a major zoonotic threat.

CMVs are showing promise as the basis of a generation of host-specific vaccine vectors [20, 21, 29, 32, 44]. In this context, our study opens the way to developing MnatCMVs to target zoonotic pathogens in *M. natalensis*. As a component of an ongoing multi-institutional study, we have since isolated several infectious MnatCMVs and sequenced their genomes. We plan to clone these genomes as infectious bacterial artificial chromosomes, thereby enabling the development of a novel scalable vaccine platform for combating zoonotic pathogens in this important reservoir host.

CONCLUSIONS

In addition to a novel polyomavirus, this study represents the first identification of herpesviruses within *M. natalensis*. In contrast to the arenaviruses commonly found in this rodent species, we anticipate that these newly identified viruses represent a low zoonotic risk due to the normally highly restricted specificity of DNA viruses such as polyomaviruses and herpesviruses to their individual mammalian host species.

List of abbreviations

BMCMC	Bayesian Markov chain Monte Carlo
CI	Côte d'Ivoire
DPOL	DNA polymerase
LASV	Lassa virus
gB	Glycoprotein B
LD-PCR	Long-distance PCR

LTA _g	Large T antigen
MCPyV	Merkel cell polyomavirus
ML	Maximum likelihood
<i>M. natalensis</i>	<i>Mastomys natalensis</i>
MnatCMV	Mastomys natalensis cytomegalovirus
MnatPyV	Mastomys natalensis polyomavirus
MnatRHV	Mastomys natalensis rhadinovirus
PhyML-SMS	Maximum-likelihood analysis using PhyML v3 with smart model selection
PST	Posterior set of trees
PyV	Polyomavirus
SH-like aLRT	Shimodaira-Hasegawa-like approximate likelihood ratio test
VP1	Virion protein 1

REFERENCES

1. Anisimova M, Gil M, Dufayard JF, Dessimoz C, Gascuel O (2011) Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst Biol* 60:685-699
2. Anoh AE, Murthy S, Akoua-Koffi C, Couacy-Hymann E, Leendertz FH, Calvignac-Spencer S, Ehlers B (2018) Cytomegaloviruses in a community of wild nonhuman primates in Taï National Park, Côte D'Ivoire. *Viruses* 10:11
3. Boardman K, Rosenke K, Safronetz D, Feldmann H, Schwan TG (2019) Host Competency of the Multimammate Rat *Mastomys natalensis* Demonstrated by Prolonged Spirochetemias with the African Relapsing Fever Spirochete *Borrelia crocidurae*. *Am J Trop Med Hyg* 101:1272-1275
4. Buck CB, Van Doorslaer K, Peretti A, Geoghegan EM, Tisza MJ, An P, Katz JP, Pipas JM, McBride AA, Camus AC, McDermott AJ, Dill JA, Delwart E, Ng TF, Farkas

- K, Austin C, Kraberger S, Davison W, Pastrana DV, Varsani A (2016) The ancient evolutionary history of polyomaviruses. *PLoS Path* 12:e1005574
5. Calvignac-Spencer S, Feltkamp MC, Daugherty MD, Moens U, Ramqvist T, Johne R, Ehlers B (2016) A taxonomy update for the family Polyomaviridae. *Arch Virol* 161:1739-1750
 6. Carter JJ, Daugherty MD, Qi X, Bheda-Malge A, Wipf GC, Robinson K, Roman A, Malik HS, Galloway DA (2013) Identification of an overprinting gene in Merkel cell polyomavirus provides evolutionary insight into the birth of viral genes. *Proc Nat Acad Sci* 110:12744-12749
 7. Colangelo P, Verheyen E, Leirs H, Tatar 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
 8. Davison AJ (2010) Herpesvirus systematics. *Vet Microbiol* 143:52-69
 9. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792-1797
 10. Ehlers B, Ulrich S, Goltz M (1999) Detection of two novel porcine herpesviruses with high similarity to gammaherpesviruses. *J Gen Virol* 80:971-978
 11. Ehlers B, K uchler J, Yasmum N, Dural G, Voigt S, Schmidt-Chanasit J, J akel T, Matuschka F-R, Richter D, Essbauer S, Hughes DJ, Stewart JP, Summers C, Bennett M, Ulrich R (2007) Identification of novel rodent herpesviruses, including the first gammaherpesvirus of *Mus musculus*. *J Virol* 81:8091-8100
 12. Ehlers B, Spie  K, Leendertz F, Peeters M, Boesch C, Gatherer D, McGeoch DJ (2009) Lymphocryptovirus phylogeny and the origins of Epstein-Barr virus. *J Gen Virol* 91:630-642
 13. Ehlers B, Wieland U (2013) The novel human polyomaviruses HPyV 6, 7, 9 and beyond. *Acta Pathol Microbiol Immunol Scand* 121:783-795

14. Ehlers B, Anoh AE, Ben Salem N, Broll S, Couacy-Hymann E, Fischer D, Gedvilaite A, Inghent N, Liebmann S, Martin M, Mossoune A, Mugisha L, Muyembe-Tamfum J-J, Pauly M, Pérez de Val B, Preugschas H, Richter D, Schubert G, Szentiks CA, Teichmann T, Walter C, Ulrich RG, Wiersma L, Leendertz FH, Calvignac-Spencer S (2019) Novel Polyomaviruses in Mammals from Multiple Orders and Reassessment of Polyomavirus Evolution and Taxonomy. *Viruses* 11:930
15. Gedvilaite A, Tryland M, Ulrich RG, Schneider J, Kurmauskaite V, Moens U, Preugschas H, Calvignac-Spencer S, Ehlers B (2017) Novel polyomaviruses in shrews (Soricidae) with close similarity to human polyomavirus 12. *J Gen Virol* 98:3060-3067
16. Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 27:221-224
17. Gryseels S, Rieger T, Oestereich L, Cuypers B, Borremans B, Makundi R, Leirs H, Günther S, Goüy de Bellocq J (2015) Gairovirus, a novel arena virus of the widespread *Mastomys natalensis*: genetically divergent, but ecologically similar to Lassa and Morogoro viruses. *Virology* 476:249-256
18. Guindon S, Lethiec F, Duroux P, Gascuel O (2005) PHYML Online--a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res* 33:W557-559
19. Günther S, Hoofd G, Charrel R, Röser C, Becker-Ziaja B, Lloyd G, Sabuni C, Verhagen R, Van der Groen G, Kennis J (2009) Mopeia virus--related arenavirus in natal multimammate mice, Morogoro, Tanzania. *Emerg Infect Dis* 15:2008
20. Hansen SG, Sacha JB, Hughes CM, Ford JC, Burwitz BJ, Scholz I, Gilbride RM, Lewis MS, Gilliam AN, Ventura AB (2013) Cytomegalovirus vectors violate CD8+ T cell epitope recognition paradigms. *Science* 340:1237874

21. Hansen SG, Wu HL, Burwitz BJ, Hughes CM, Hammond KB, Ventura AB, Reed JS, Gilbride RM, Ainslie E, Morrow DW (2016) Broadly targeted CD8+ T cell responses restricted by major histocompatibility complex E. *Science* 351:714-720
22. Isaäcson M (1975) The ecology of *Praomys* (*Mastomys*) *natalensis* in southern Africa. *Bull World Health Organ* 52:629
23. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647-1649
24. Kingdon J, Happold D, Butynski T, Hoffmann M, Happold M, Kalina J (2013) *Mammals of Africa*. A&C Black
25. Kouadio L, Nowak K, Akoua-Koffi C, Weiss S, Allali BK, Witkowski PT, Krüger DH, Couacy-Hymann E, Calvignac-Spencer S, Leendertz FH (2015) Lassa virus in multimammate rats, Côte d'Ivoire, 2013. *Emerg Infect Diss* 21:1481
26. Leendertz FH, Scuda N, Cameron KN, Kidega T, Zuberbühler K, Leendertz SAJ, Couacy-Hymann E, Boesch C, Calvignac S, Ehlers B (2011) African great apes are naturally infected with polyomaviruses closely related to Merkel cell polyomavirus. *J Virol* 85:916-924
27. Lefort V, Longueville JE, Gascuel O (2017) SMS: Smart Model Selection in PhyML. *Mol Biol Evol* 34:2422-2424
28. Madinda NF, Ehlers B, Wertheim JO, Akoua-Koffi C, Bergl RA, Boesch C, Akonkwa DBM, Eckardt W, Fruth B, Gillespie TR (2016) Assessing host-virus codivergence for close relatives of Merkel cell polyomavirus infecting African great apes. *J Virol* 90:8531-8541
29. Marzi A, Murphy AA, Feldmann F, Parkins CJ, Haddock E, Hanley PW, Emery MJ, Engelmann F, Messaoudi I, Feldmann H (2016) Cytomegalovirus-based vaccine expressing Ebola virus glycoprotein protects nonhuman primates from Ebola virus infection. *Sci Rep* 6:21674

30. McGeoch DJ, Gatherer D, Dolan A (2005) On phylogenetic relationships among major lineages of the Gammaherpesvirinae. *J Gen Virol* 86:307-316
31. Monath TP, Newhouse VF, Kemp GE, Setzer HW, Cacciapuoti A (1974) Lassa virus isolation from *Mastomys natalensis* rodents during an epidemic in Sierra Leone. *Science* 185:263-265
32. Murphy AA, Redwood AJ, Jarvis MA (2016) Self-disseminating vaccines for emerging infectious diseases. *Expert Rev Vaccines* 15:31-39
33. Murthy S, O'Brien K, Agbor A, Angedakin S, Arandjelovic M, Ayimisin EA, Bailey E, Bergl RA, Brazzola G, Dieguez P (2019) Cytomegalovirus distribution and evolution in hominines. *Virus Evol* 5:vez015
34. Olayemi A, Obadare A, Oyeyiola A, Igbokwe J, Fasogbon A, Igbahenah F, Ortsega D, Asogun D, Umeh P, Vakkai I (2016) Arenavirus diversity and phylogeography of *Mastomys natalensis* rodents, Nigeria. *Emerg Infect Dis* 22:694
35. Olayemi A, Obadare A, Oyeyiola A, Igbokwe J, Fasogbon A, Igbahenah F, Ortsega D, Asogun D, Umeh P, Vakkai I (2016) Arenavirus diversity and phylogeography of *Mastomys natalensis* rodents, Nigeria. *Emerg Infect Dis* 22:694
36. Orba Y, Kobayashi S, Nakamura I, Ishii A, Hang'ombe BM, Mweene AS, Thomas Y, Kimura T, Sawa H (2011) Detection and characterization of a novel polyomavirus in wild rodents. *J Gen Virol* 92:789-795
37. Rambaut A, Lam TT, Max Carvalho L, Pybus OG (2016) Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* 2:vew007
38. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst Biol* 67:901-904
39. Richmond JK, Baglolle DJ (2003) Lassa fever: epidemiology, clinical features, and social consequences. *Brit Med J* 327:1271-1275
40. Scuda N, Madinda NF, Akoua-Koffi C, Adjogoua EV, Wevers D, Hofmann J, Cameron KN, Leendertz SAJ, Couacy-Hymann E, Robbins M, Boesch C, Jarvis MA, Moens U,

- Mugisha L, Calvignac S, Ehlers B (2013) Novel polyomaviruses of nonhuman primates: genetic and serological predictors for the existence of multiple unknown polyomaviruses within the human population. *PLoS Path* 9:e1003429
41. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A (2018) Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol* 4:vey016
 42. Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* 56:564-577
 43. Tan C-H, Tachezy R, Van Ranst M, Chan S-Y, Bernard H-U, Burk RD (1994) The *Mastomys natalensis* papillomavirus: nucleotide sequence, genome organization, and phylogenetic relationship of a rodent papillomavirus involved in tumorigenesis of cutaneous epithelia. *Virology* 198:534-541
 44. Tsuda Y, Parkins CJ, Caposio P, Feldmann F, Botto S, Ball S, Messaoudi I, Cicin-Sain L, Feldmann H, Jarvis MA (2015) A cytomegalovirus-based vaccine provides long-lasting protection against lethal Ebola virus challenge after a single dose. *Vaccine* 33:2261-2266

Acknowledgements

We thank Cornelia Walter and Melanie Fechtner, Robert Koch-Institute, Berlin, for excellent technical assistance. Rodent work in Mali was supported by the International Centers for Excellence in Research (ICER) program of the Division of Intramural Research, NIAID, NIH.

Funding

This work was partially funded by the Deutsche Forschungsgemeinschaft (grants LE1813/7–1 and KR1293/15–1), the International Centers for Excellence in Research (ICER) program, the Intramural Research Program of the Division of Intramural Research, NIAID, NIH, and the Medical Research Council (grant MC_UU_12014/3). This work was also partially funded through awards to the Vaccine Group Company, Ltd, and the University of Plymouth.

ETHICS DECLARATIONS

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

Live-trapping and euthanization: In CI, ethical approval to sample and analyze small mammal organs was obtained from the National Commission of Ethics and Research (CNER) under protocol number 033 / MSLS / CNER-dkm. Animal studies with *M. natalensis* from Mali were approved by the Institutional Animal Care and Use Committee of the Rocky Mountain Laboratories, Division of Intramural Research, NIAID, NIH under protocol number #2015-011. All procedures on live animals were conducted in compliance with the applicable institutional and national guidelines for use and handling of animals.

FIGURE LEGENDS

Fig. 1 Amplification of herpesvirus and polyomavirus sequences

(A) Sequences of the herpesvirus DNA polymerase (DPOL) and glycoprotein B (gB) CDSs (depicted schematically by blue bars) were amplified by generic DPOL and gB PCRs. The respective products are depicted by magenta bars. The sequence between these regions of DPOL and gB was then amplified by long-distance PCR (grey bar). **(B)** The genome of *M. natalensis* polyomavirus 3 was amplified by generic PCR (magenta bar) targeting the VP1 CDS. The remaining part of the genome was amplified by inverse long-distance PCR using 'back-to-back' primers (grey bar). The CDSs in the early region are depicted by red arrows, those in the late region by blue arrows, and the NCCR by a green bar.

Fig. 2 Maximum-likelihood tree of betaherpesviruses based on conserved amino acid blocks of glycoprotein B

Betaherpesviruses are denoted by Latin taxonomic name or common name, followed by GenBank accession number and host common name. Denotation of the novel *M. natalensis* cytomegalovirus 3 identified herein is in bold font and includes also the sample ID. For ICTV-recognized virus species, genera are indicated by the colored background of the virus name. Grey branches are relatively weakly supported with posterior probability values <0.95.

Fig. 3 Maximum-likelihood tree of gammaherpesviruses based on conserved amino acid blocks of glycoprotein B

For explanation see legend of Figure 2.

Fig. 4 Maximum-likelihood tree of polyomaviruses based on conserved amino acid blocks of large T antigen

Polyomavirus nomenclature follows recommendations of the ICTV *Polyomaviridae* Study Group using the Latin binomials of their hosts followed by a serial number; GenBank accession numbers and vernacular names of the host are also given. Virus genera are indicated by colored background. *Mastomys natalensis* cytomegalovirus 3 identified in this study is identified in bold font. Grey branches are relatively weakly supported with posterior probability values <0.95.

ONLINE RESOURCES

ESM1. Maximum clade credibility tree analysis of betaherpesviruses based on glycoprotein B. Phylogenetic relationships of betaherpesviruses, including the novel *Mastomys natalensis* cytomegalovirus 3, based on conserved amino acid blocks of glycoprotein B sequence. For further details see legend of Figure 2.

ESM2. Maximum clade credibility tree analysis of gammaherpesviruses based on glycoprotein B. Phylogenetic relationships of gammaherpesviruses, including the novel *M. natalensis* rhadinovirus 1, based on conserved amino acid blocks of glycoprotein B sequence. For further details see legend of Figure 2.

ESM3. Maximum likelihood tree analysis of betaherpesviruses based on DNA polymerase. Phylogenetic relationships of betaherpesviruses, including the novel *M. natalensis* cytomegaloviruses, based on conserved amino acid blocks of DNA polymerase sequence. For further details see legend of Figure 2.

ESM4. Maximum clade credibility tree analysis of betaherpesviruses based on DNA polymerase. Phylogenetic relationships of betaherpesviruses, including the novel *M. natalensis* cytomegaloviruses, based on conserved amino acid blocks of DNA polymerase sequence. For further details see legend of Figure 2.

glycoprotein B CDS

DNA polymerase CDS

1

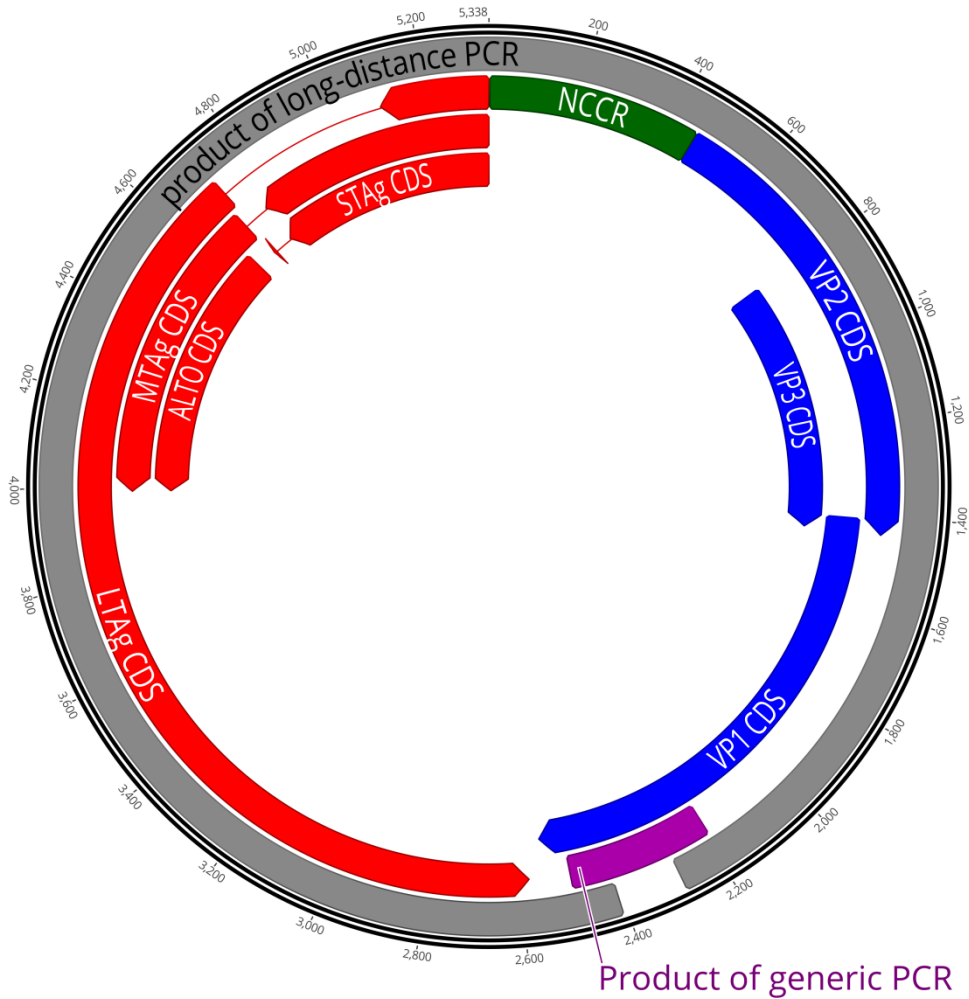
Appr. 5500 (bp)

product of generic gB PCR

product of generic DPOL PCR

product of specific long-distance PCR

A



B

ESM5. Maximum likelihood tree analysis of gammaherpesviruses based on DNA polymerase. Phylogenetic relationships of gammaherpesviruses, including the novel M. natalensis rhadinovirus 1, based on conserved amino acid blocks of DNA polymerase sequence. For further details see legend of Figure 2.

ESM6. Maximum clade credibility tree analysis of gammaherpesviruses based on DNA polymerase. Phylogenetic relationships of gammaherpesviruses, including the novel M. natalensis rhadinovirus 1, based on conserved amino acid blocks of DNA polymerase sequence. For further details see legend of Figure 2.

ESM7. Maximum clade credibility tree analysis of polyomaviruses based on large T antigen. Phylogenetic relationships of polyomaviruses, including the novel M. natalensis polyomavirus 3, based on conserved amino acid blocks of large T sequence. For further details see legend of Figure 4.

ESM8. Maximum likelihood tree analysis of polyomaviruses based on VP1. Phylogenetic relationships of polyomaviruses, including the novel M. natalensis polyomavirus 3, based on conserved amino acid blocks of VP1 sequence. For further details see legend of Figure 4.

ESM7. Maximum clade credibility tree analysis of polyomaviruses based on VP1. Phylogenetic relationships of polyomaviruses, including the novel M. natalensis polyomavirus 3, based on conserved amino acid blocks of VP1 sequence. For further details see legend of Figure 4.

