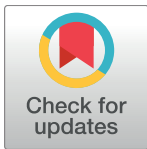


RESEARCH ARTICLE

A metagenomic viral discovery approach identifies potential zoonotic and novel mammalian viruses in *Neoromicia* bats within South Africa

Marike Geldenhuys¹, Marinda Mortlock¹, Jacqueline Weyer², Oliver Bezuidt³, Ernest C. J. Seamark^{4,5}, Teresa Kearney^{6,7}, Cheryl Gleasner⁸, Tracy H. Erkkila⁸, Helen Cui⁸, Wanda Markotter^{1*}



1 Centre for Viral Zoonoses, Department of Medical Virology, University of Pretoria, Pretoria, South Africa, **2** Centre for Emerging Zoonotic and Parasitic Diseases, National Institute for Communicable Diseases, National Health Laboratory Services, Johannesburg, South Africa, **3** Centre for Microbial Ecology and Genomics, University of Pretoria, Pretoria, South Africa, **4** AfricanBats NPC, Pretoria, South Africa, **5** Eugène Marais Chair of Wildlife Management, Mammal Research Institute, University of Pretoria, Pretoria, South Africa, **6** Ditsong National Museum of Natural History, Pretoria, South Africa, **7** School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg, South Africa, **8** Los Alamos National Laboratory, Los Alamos, New Mexico, United States of America

* wanda.markotter@up.ac.za

OPEN ACCESS

Citation: Geldenhuys M, Mortlock M, Weyer J, Bezuidt O, Seamark ECJ, Kearney T, et al. (2018) A metagenomic viral discovery approach identifies potential zoonotic and novel mammalian viruses in *Neoromicia* bats within South Africa. PLoS ONE 13 (3): e0194527. <https://doi.org/10.1371/journal.pone.0194527>

Editor: Stefan Pöhlmann, Deutsches Primatenzentrum GmbH - Leibniz-Institut für Primatenforschung, GERMANY

Received: December 14, 2017

Accepted: March 5, 2018

Published: March 26, 2018

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: Relevant data are found within the paper and its Supporting Information file. The raw metagenomics sequence data was submitted to the SRA archive of NCBI (<https://www.ncbi.nlm.nih.gov/>) under accession numbers SRR5889194 and SRR58891929. Viral sequences were submitted to NCBI (<https://www.ncbi.nlm.nih.gov/>) under accession numbers MF579865-MF579871 and MF593268-MF593281.

Abstract

Species within the *Neoromicia* bat genus are abundant and widely distributed in Africa. It is common for these insectivorous bats to roost in anthropogenic structures in urban regions. Additionally, *Neoromicia capensis* have previously been identified as potential hosts for Middle East respiratory syndrome (MERS)-related coronaviruses. This study aimed to ascertain the gastrointestinal virome of these bats, as viruses excreted in fecal material or which may be replicating in rectal or intestinal tissues have the greatest opportunities of coming into contact with other hosts. Samples were collected in five regions of South Africa over eight years. Initial virome composition was determined by viral metagenomic sequencing by pooling samples and enriching for viral particles. Libraries were sequenced on the Illumina MiSeq and NextSeq500 platforms, producing a combined 37 million reads. Bioinformatics analysis of the high throughput sequencing data detected the full genome of a novel species of the *Circoviridae* family, and also identified sequence data from the *Adenoviridae*, *Coronaviridae*, *Herpesviridae*, *Parvoviridae*, *Papillomaviridae*, *Phenuiviridae*, and *Picornaviridae* families. Metagenomic sequencing data was insufficient to determine the viral diversity of certain families due to the fragmented coverage of genomes and lack of suitable sequencing depth, as some viruses were detected from the analysis of reads-data only. Follow up conventional PCR assays targeting conserved gene regions for the *Adenoviridae*, *Coronaviridae*, and *Herpesviridae* families were used to confirm metagenomic data and generate additional sequences to determine genetic diversity. The complete coding genome of a MERS-related coronavirus was recovered with additional amplicon sequencing on the MiSeq platform. The new genome shared 97.2% overall nucleotide identity to a previous *Neoromicia*-associated MERS-related virus, also from South Africa. Conventional PCR

Funding: This work was financially supported in part by the National Research Foundation (NRF) of South Africa: the South African Research Chair held by WM grant no. 98339, as well as grant numbers 92524, 85756, and 91496. The opinions, findings and conclusions expressed are those of the authors alone, and the NRF accepts no liability in this regard for research supported. Additional support was obtained by WM from the Poliomyelitis Research Foundation (grant number: 12/14). The Research Trust of the National Health Laboratory Service and the Medical Research Council was awarded to JW. This research was partially supported by the Grant or Cooperative Agreement Number [5 NU2GGH001874-02-00], funded by the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Department of Health and Human Services. MG was supported by funding from the NRF's Innovation bursary award (grant UID: 79409), the Poliomyelitis Research Foundation (grant no. 13/48), and the postgraduate study abroad bursary program of the University of Pretoria, which funded a research visit to the Los Alamos National Laboratory in New Mexico.

Competing interests: The authors have declared that no competing interests exist.

analysis detected diverse adenovirus and herpesvirus sequences that were widespread throughout *Neoromicia* populations in South Africa. Furthermore, similar adenovirus sequences were detected within these populations throughout several years. With the exception of the coronaviruses, the study represents the first report of sequence data from several viral families within a Southern African insectivorous bat genus; highlighting the need for continued investigations in this regard.

Introduction

The role of bats as potential or confirmed reservoirs of various viral agents with public health importance has been increasingly appreciated in recent years. Several bat-borne viruses are considered emerging, with an increase in the number of human cases and outbreaks over the past two decades; such as the bat-borne viruses Marburg virus and Nipah virus [1,2]. These viruses are associated with direct zoonotic transmission and infection of exposed human populations. Other identified bat-associated viruses are only related to viruses of known public health importance, like the severe acute respiratory syndrome coronavirus (SARS-CoV) or Middle East respiratory syndrome coronavirus (MERS-CoV) [3–5]. Though no direct spillover have been identified, overall sequence similarities between human viral strains and related bat-borne viruses have implicated these viruses in the emergence of new human coronaviruses [6,7]. Detection of potentially zoonotic viruses in bat species has been the driving force for increased research and surveillance of the bat virome (as well as other infectious agents).

Viral discovery with conventional nucleic acid detection assays targeting conserved gene regions (such as for PCR-based surveillance) as well as viral metagenomic studies have been successful in analysing the viral richness harboured by bats [8–14]. The use of sequence-independent metagenomic approaches has enabled detection of highly divergent viruses, often sharing only low homologies to currently known virus species [11,14–17]. Viral metagenomic investigations provide opportunities for the detection of potentially zoonotic viruses, as well as other mammalian-infecting viruses not previously known to have been present [8,11,18]. An inventory of novel viruses can thus be constructed, and further investigated for their potential as zoonotic agents. Metagenomics and conventional PCR surveillance have identified bats as natural hosts for large genetic diversities of several viral families—such as the *Circoviridae*, *Coronaviridae*, *Paramyxoviridae*, *Parvoviridae*, and *Adenoviridae*, as well as the genus *Hepacivirus*, in the family *Flaviviridae* [6,10,19,20].

Insectivorous bats of the genus *Neoromicia* (family *Vespertilionidae*) referred to commonly as either Serotine or Pipistrelle bats, are distributed in the Afro-Malagasy regions. Thirteen of the seventeen currently recognized species occur on mainland Africa, with six species (*Neoromicia capensis*, *N. nana*, *N. rendalli*, *N. stanleyi*, *N. zuluensis* and *N. cf. helios*) recorded in South Africa [21,22]. *Neoromicia* bats habitually roost in small colony numbers (up to 10 individuals) in naturally occurring crevices, though some species may also be found in anthropogenic structures such as the roofs of houses [21].

The International Union for Conservation of Nature (IUCN) Red List conservation status varies per species. *Neoromicia nana* and *N. capensis* are both considered “Least Concern” (i.e. not threatened) in the global IUCN Red List assessment due to their wide geographic range and abundance, with no population declines currently known [23,24]. Investigations have identified that both species host various viruses. The novel hantavirus, Mouyassué virus, was detected from archival specimens of two *N. nana* individuals collected in the Ivory Coast in

West Africa [25]. In South Africa, novel paramyxovirus sequences were identified from the same host species [26]. Lastly, coronaviruses from both the genera *Alphacoronavirus* and *Beta-coronavirus* have been reported from *N. capensis* species [27,28]. A detected lineage C betacoronavirus, named NeoCoV/PML-PHE1, grouped taxonomically within the same species as MERS-CoV [4,28]. The virus shared an overall nucleotide identity of 85.5–85.6% to human and camel strains of MERS-CoV, with only 64.3–64.6% amino acid identity within the spike gene that is responsible for receptor recognition and attachment [4]. Considering the high diversity of genus-specific bat coronaviruses detected globally [6,27,29], other MERS-related viruses are likely circulating within this host genus.

This study focused on viral discovery within the genus *Neoromicia*, to identify viruses that may be present within this host, as well as detect additional sequences of MERS-related viruses for genetic comparisons. Viral metagenomic sequencing provided an initial assessment of the virome composition from gastrointestinal samples at a host genus-level, detecting sequences from several viral families. However, sequencing outputs produced fragmented coverage of several viruses, making sequence diversity difficult to discern, and lacked suitable sequencing depth as some viruses were only detected from sequencing reads-data. Three viral families (*Coronaviridae*, *Adenoviridae*, and *Herpesviridae*) were selected for further investigation by analysing samples pooled per host species. Conventional nucleic acid detection based on conserved gene regions confirmed diverse sequences of adenoviruses and herpesviruses, as well as the recovery of the complete coding genome of another MERS-related virus.

Materials and methods

Ethics statement and sample collection

This research was conducted with the approval of the University of Pretoria Animal Ethics committee (Project no. EC054-14 and EC059-14). Permits were obtained for bat sample collection from the South African provinces involved: the Department of Economic Development, Environment and Tourism Limpopo province directorate- wildlife permit no. CPM006806; Premier of the Province of Gauteng Nature conservation permit no. CPF6 no. 0027/ no.0109; Department of Agriculture, Conservation, Environment and Tourism no. 000039 NW-07; Mpumalanga Tourism and Parks Agency no. MBP5385. As part of a research programme focusing on zoonotic pathogens in small mammals, specimens were collected from four *Neoromicia* species, in five South African provinces between 2007 and 2015 (S1 Table). Bats were measured for morphological identification [21,30], and either released after sampling or taken as voucher specimens and deposited in the small mammal collection at the Ditsong National Museum of Natural History (Pretoria, SA). Further species confirmation was sought for individuals from which coronavirus RNA was detected with sequencing of the mitochondrial cytochrome c oxidase subunit I (COI) gene [31]. All specimen material collected including faecal pellets, as well as rectal and intestinal tissues from necropsies were immediately frozen in liquid nitrogen and stored at -80°C after returning to the laboratory.

Sample preparation and viral metagenomic enrichment steps

Samples from 58 individuals of four *Neoromicia* species were collected over an eight-year time span from the north eastern regions of South Africa (Fig 1). The study focused on samples from the lower gastrointestinal tract (faecal, rectal and intestinal samples), to identify viruses present in these tissues or excreted in faecal material (Fig 2A; S1 Table). Due to the size of these bat species and small quantities of faecal material at times available for collection, sample material for certain individuals were limited. All faecal or tissue samples were used sparingly to enable additional analyses. However, priority was given to sample preparation for high

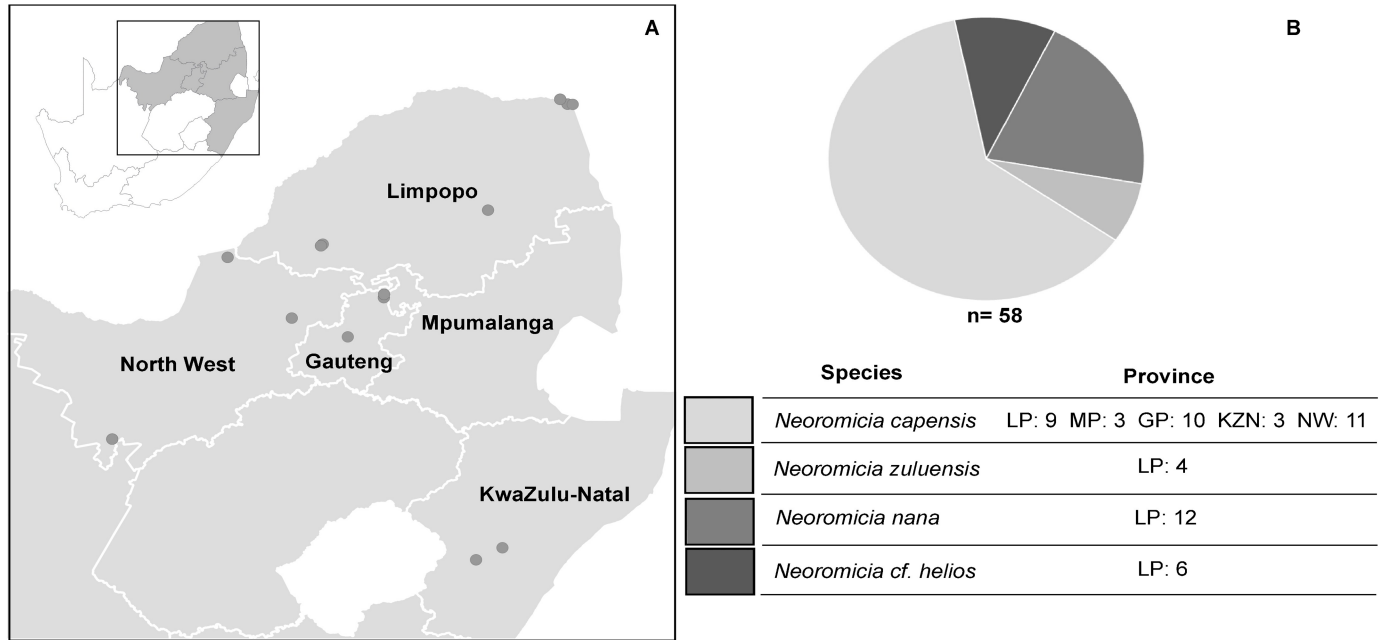


Fig 1. Map of *Neoromicia* sampling locations. A) Sampling locations of *Neoromicia* species from 2007 to 2015 in the north eastern regions in South Africa. The map was plotted according to GPS coordinates in QGIS 2.0.1. B) Proportion and number of *Neoromicia* species collected per province for the *Neoromicia* virome analysis. LP = Limpopo, MP = Mpumalanga, GP = Gauteng, KZN = KwaZulu-Natal, NW = North West.

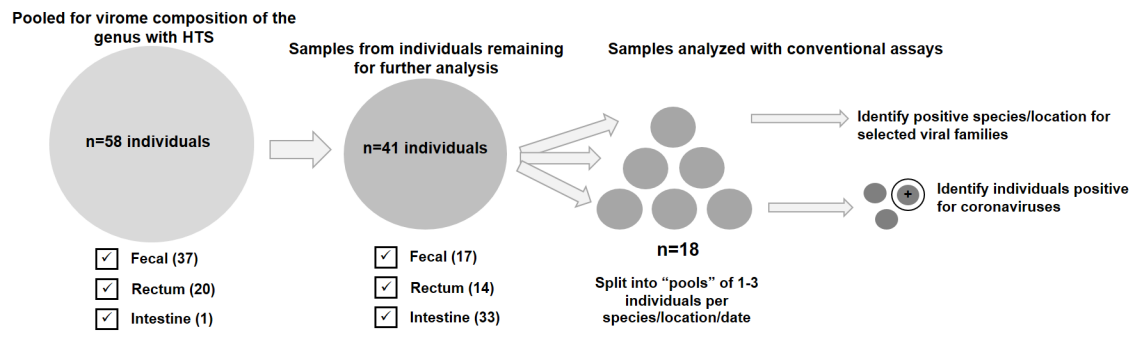
<https://doi.org/10.1371/journal.pone.0194527.g001>

throughput sequencing in order to determine the virome composition of the analysed host genus. Therefore, if sample material consisted of very small quantities of faecal or rectal material, these samples were completely used for this step. For high throughput sequencing sample preparation, faecal and rectal samples of all bats were pooled and processed by mechanical homogenization in 400 µl PBS (Lonza) with two stainless-steel beads (Qiagen). A Tissue Lyser II system (Qiagen) was utilized, shaking at 30Hz for 60 seconds. The lysates were cleared by low temperature (4°C) centrifugation for 10 minutes at 10,000x g. Cleared supernatants were pooled and filtered through a 0.45 µm cellulose acetone syringe filter (Corning Incorporated) to remove large particulate matter. The filtrate was ultracentrifuged at 130 000 x g for 2 hours at 4°C. The pellet was resuspended in 200 µl PBS and DNase treated with Turbo DNase (Ambion) before total nucleic acid extraction (ZR Viral DNA/RNA Kit, Zymo Research). Total nucleic acids were eluted in 35 µl nuclease free water (Ambion) and further depleted of ribosomal RNA with the RiboMinus™ Eukaryote System v2 (Ambion); nucleic acids were eluted in 12 µl nuclease free water (Ambion).

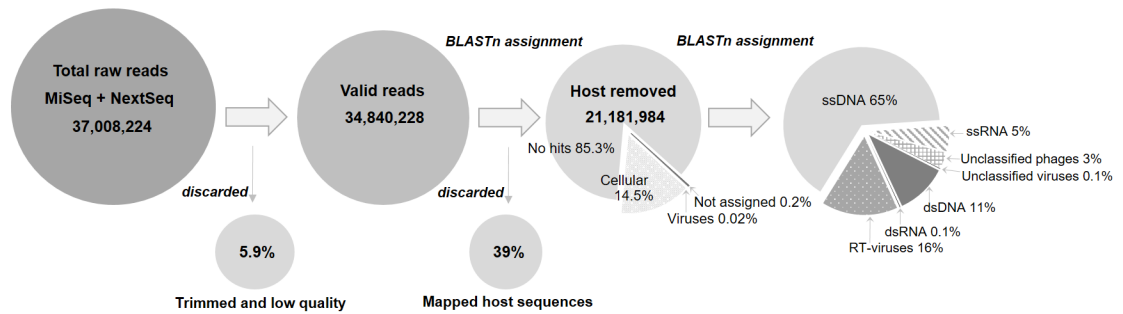
Double stranded cDNA preparation for Illumina sequencing

First strand cDNA synthesis and double stranded DNA was prepared as described in Kohl et al. [32] using the K-mer approach (5' GACCATCTAGCGACCTCCCANNNNNNNNN' 3 and 5' GACCATCTAGCGACCTCCCA' 3) described in Victoria et al. [33]. The double stranded cDNA was purified with AmPure XP beads (Agencourt, Beckman Coulter) and analysed for quality with both the Qubit fluorometer (Life technologies, Thermo Scientific) and Agilent 2100 Bioanalyzer (Agilent Technologies). Size selection of 500bp was performed with AmPure XP beads (Agencourt, Beckman Coulter) and the virome library was prepared with the NEB Next Ultra DNA library prep kit (New England Biolabs, Inc). Libraries were normalized to 10pM and 2pM, respectively, and sequenced on the MiSeq and the NextSeq 500 Illumina

A. Sample pooling- from initial viral metagenomic preparation to determination of positive species/ individuals



B. Bioinformatics overview of Illumina read analysis:



C. BLASTn viral assignment of reads:

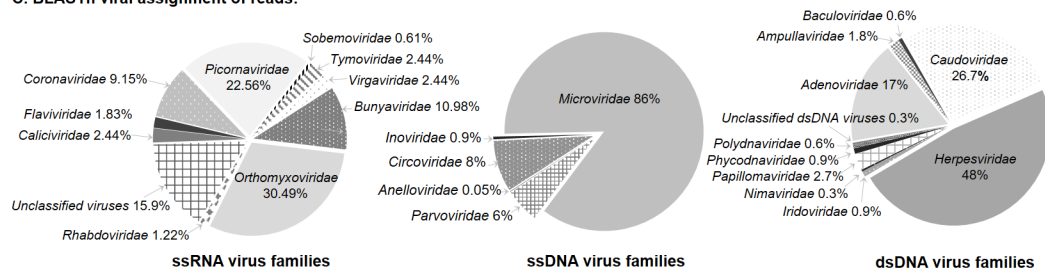


Fig 2. Strategies for sample pooling and data analysis workflows. A) The workflow depicts the sample pooling strategies utilized in the study. Faecal and rectal samples were primarily pooled for initial viral metagenomic sample processing. Subsequently, remaining sample material (faecal homogenates, rectum, or intestine) available for individual bats were pooled according to species and location to associate selected viral families with specific species. A further step was implemented to determine specific host individuals harbouring coronaviruses in order to clarify exact sampling locations. B) Bioinformatics data analysis of Illumina reads with a developed workflow in CLC genomic workbench that trims data for quality, removes adaptors and depletes host reads based on set mapping parameters. Remaining reads were taxonomically assigned with BLASTn and MEGAN v6. C) Shows the BLASTn assignments of reads according to virus genome types: ssRNA, ssDNA and dsDNA viruses.

<https://doi.org/10.1371/journal.pone.0194527.g002>

sequencers for 150bp paired reads at the Los Alamos National Laboratory (New Mexico, USA).

Bioinformatics and initial taxonomic assignment

Raw Illumina reads were processed in CLC Bio Genomics workbench (Qiagen) using an ‘in house’ developed workflow (S2 Table). After an initial data QC, reads were trimmed for quality and adapter sequences (removing indexes and primers). Host sequences were depleted by strict mapping to an available full genome of a related bat host using default parameters (*Myotis brandtii* assembly no. GCF_00412655.1). The remaining reads were searched for sequence

similarity by BLASTn against the NCBI's nucleotide (nt) database (downloaded July 2015) using an e-value of $<10^{-5}$. Due to limited bioinformatics computational power available, alternative strategies to reduce run time analysis of BLASTx were implemented. A smaller 'virus-only' database was obtained from NCBI (downloaded July 2015) and used for BLASTx analyses of reads and contigs. The BLAST xml output files were subsequently exported to the MetaGenome Analyzer (MEGAN v6) [34] for taxonomic assignments, and were further manually investigated. Combinations of assembly approaches were employed including CLC Bio Genomics workbench assembly, and assemblies in Velvet 1.2.10 [35] of all host depleted reads as well assemblies only after family-level identified reads were exported from MEGAN. Kmer sizes of 21, 33 and 55 were used and all contigs from different approaches were combined and curated.

Conventional nucleic acid detection for specific viral families

In a conserved approach to confirm the high throughput sequence data of specific viral families, primers targeting conserved gene regions were selected from the literature or designed based on obtained sequence data using CLC Bio Genomic Workbench and assessed with Annhyb v4.9. Confirmatory PCRs were performed using sample material remaining after high throughput sequencing. Due to limited sample quantity, original samples of certain individuals were depleted during metagenomic sample preparation. Therefore, remaining samples (faecal homogenates, rectum, or intestine) available from individual bats were pooled according to species, location and sampling date in order to associate selected viral families with specific species (S3 Table). A further step was implemented to analyse original samples of pools found to harbour coronavirus RNA to determine the exact sampling locations and dates of specific host individuals (Fig 2A). Total nucleic acids were extracted in parallel using the Duet RNA/DNA miniprep plus kit (ZymoResearch). A 20 μ l cDNA reaction was prepared with 100 ng random primers (IE HPLC Purified, Integrated DNA Technologies) and Superscript III (Invitrogen), followed by incubation for 20 minutes at 37°C in the presence of 2U RNase H (Thermo Fisher Scientific) and inactivated at 65°C for 10 minutes.

In-house *Alpha*- and *Betacoronavirus* specific hemi-nested RT-PCR assays targeting the RNA dependent RNA polymerase (RdRp) gene adapted from Geldenhuys et al. [27] were utilized for *Coronaviridae* confirmatory PCR. In a final reaction volume of 50 μ l, 2 μ l cDNA template was combined with 1.25U Dream Taq (Thermo Scientific), 1X Dream Taq buffer, 200 μ M deoxynucleoside triphosphate (dNTP), 10 pmol of each first round forward and reverse primer (S4 Table) and nuclease free water (Ambion). First round PCR cycling conditions were as follows: 94°C for 1 minute, with 40 cycles of 94°C for 30 s, 42°C for 30 s, 72°C for 1 min; and a final extension at 72°C for 10 minutes. The nested PCR was similarly set up with alterations to the cycling conditions: 94°C for 1 minute, with 35 cycles of 94°C for 30 s, 42°C for 30 s, 72°C for 45 sec; and a final extension at 72°C for 10 minutes. The RdRp gene of detected coronaviruses were extended with appropriate RdRp grouping unit (RGU) primers as described in Drexler et al. [36].

DNA virus confirmatory nucleic acid analysis was performed for the *Adenoviridae* and *Herpesviridae* families as described in Li et al. [37] and Van Devanter et al. [38], respectively. PCR products of appropriate size were excised, purified (Zymoclean™ Gel DNA Recovery Kit, Zymo Research) and sequenced on an ABI 3130 sequencer (Applied Biosystems) at Inqaba Biotech (Pretoria, SA).

MiSeq coronavirus amplicon sequencing

Amplicon sequencing was used to obtain the complete coding regions of the detected MERS-related betacoronavirus from the intestinal material of sample UP5038. Primers were designed

from obtained reads, as well as human, camel and *Neoromicia* MERS-related coronavirus reference genome sequences for 11 nested RT-PCR assays (primers available upon request). Amplicons were generated from randomly primed cDNA in PCR reactions using the Phusion high fidelity DNA polymerase (New England Biolabs, Inc). Amplicon sequencing at an estimated 1000x genome coverage was conducted on Illumina's MiSeq platform after preparation with a Nextera XT library prep kit at the National Institute for Communicable Diseases Sequencing Core Facility (Sandringham, SA). The genome assembly and annotation was performed in CLC Genomics workbench and open reading frames were identified with NCBI's ORF finder (<https://www.ncbi.nlm.nih.gov/orffinder/>) as well as the Craig Venter Institute's VIGOR software [39].

Phylogenetic analysis and pairwise estimations

Viral sequences for phylogenetic comparisons were downloaded from GenBank (NCBI). Sequence manipulations and alignments were performed with ClustalW in Bioedit [40]. Pairwise sequence similarity estimations were compared in MEGA v7 [41]. Phylogenetic analyses of selected viral contigs and confirmatory PCR sequences were performed with Bayesian phylogenetics using BEAST v. 1.8 [42]. Maximum clade credibility trees were constructed using suggested models selected from jModelTest.org [43]. Unless otherwise stated, Bayesian MCMC chains were set to 10,000,000 states, sampling every 1000 steps. Final trees were calculated from the 9000 generated trees after discarding the first 10% as burn-in. Computationally expensive analyses were run with the aid of CIPRES [44]. Trees were viewed and edited in Figtree v1.4.2.

Construction of an overlapping distribution map of known hosts of MERS virus and MERS-related coronaviruses

The African geographical distributions of dromedary camels (*Camelus dromedarius*) and the Cape serotine bats (*Neoromicia capensis*) were plotted in ArcMap v.10.4.1 to indicate possible regions of overlap that may be used as focus areas for coronavirus surveillance. The natural distribution range of dromedary camels was traced from the warm desert, semi-desert, woodland, scrub grassland vegetation subclass [45] based on the range described by Faye [46]. To identify introductions to areas outside the natural range, a Google search (terms included 'camel rides' and 'country') and Wilson [47] were employed. To create an estimated species distribution model for *N. capensis*, the museum records with point localities was extracted from the African Chiropteran Report [21] and modelled using climatic variables from present WorldClim [48], using MaxEnt version 3.3.3k [49]. MERS antibody seroprevalence data for dromedary camels was obtained from available publications as well as hosts from which MERS-CoV and MERS-related viral RNA have been detected [50–55].

Data submission and accession numbers

The raw sequence data was submitted to the Sequence Read Archive of the NCBI under accession numbers SRR5889194 and SRR58891929. Viral sequences were submitted under accession numbers MF579865-MF579871 and MF593268-MF593281.

Results and discussion

The study investigated the gastrointestinal virome of the bat genus *Neoromicia*, as viruses excreted in fecal material or present in these tissues (rectum and intestine) pose a greater risk of coming into contact with other animals. *Neoromicia* species do not roost in caves where

they can be readily caught and sampled [21]. Instead, most individuals were caught in flight away from their roosts, and therefore fewer individuals were available to investigate. The samples utilized in this study were processed with methods similar to those that have proven to be successful in bat virome analyses—with enrichment methods for conserving viral particles and reducing sequencing of host and non-viral nucleic acids [11,15,17,56]. Furthermore, an additional host ribosomal RNA depletion step was incorporated to simplify bioinformatics analysis. Prepared Illumina libraries were sequenced on the MiSeq platform as an initial examination, followed by sequencing on the NextSeq500. Combined data outputs produced 37 million reads of which 34.8 million remained after quality and adaptor trimming with an in-house developed workflow in CLC Genomics Workbench (Fig 2B). As no *Neoromicia* host genomes are available for sequence removal, the genome of *Myotis brandtii*, another species within the family Vespertilionidae, was utilized. Despite enrichment efforts, application of strict mapping parameters within the CLC workflow removed nearly 40% of reads as host (Fig 2B).

Before assembly into contigs, reads data were analysed for sequence similarity by BLASTn against the nucleotide database of the National Center for Biotechnology Information (NCBI). Of the remaining 21.2 million reads, 14% were attributed to cellular organisms (host and non-viral microbial reads), with less than 1% designated as viral in origin; which is comparable to other virome studies [15,18]. As with most metagenomic studies performed to date, the largest portion of reads (85%) were assigned as 'non-hits' by BLASTn analyses, as they did not match any available sequence data in the NCBI database [11,56].

Reads were assembled into contigs using CLC Genomics Workbench and Velvet assembler, and combined for the best results; often producing overlapping contigs. Depending on specific inputs (e.g. varying kmer lengths), the number of contigs ranged from 222,343 to 566,951. Superior contig lengths were more frequently found from CLC assemblies than those from multiple Velvet configurations with different kmer lengths. Again, less than 1% of all contigs were assigned as viral with BLASTn analyses. Contigs and unassembled reads assigned to specific virus families from BLAST outputs were inspected manually before being considered as valid (Fig 2C). Many of the assignments were deemed as false positives due to poor e-values, low complexity sequences (such as repeat runs of nucleotides) or having greater similarity matches to host/mammalian sequences [57]. Reads and contigs (e.g. rhabdoviruses, orthomyxoviruses and caliciviruses) identified as such were discarded from further analyses (Fig 2C). Despite a low percentage of viral reads, sequence data from eight mammalian-infecting virus families were detected, including the *Adenoviridae*, *Circoviridae*, *Coronaviridae*, *Herpesviridae*, *Parvoviridae*, *Papillomaviridae*, *Picornaviridae* and *Phenuiviridae*. In agreement with other virome investigations, many of detected sequences shared low genetic similarities to known viruses due to a lack of available sequence diversity for comparison [11,14]. This lack of available sequence data may only be overcome by further reports documenting the viral diversity of bats as well as other hosts. Though non-mammalian infecting viruses (phages and insect viruses) were identified, they were outside the scope of the study and not investigated any further.

New viral species in the family *Circoviridae*

Viruses from the family *Circoviridae* have small, single-stranded, DNA genomes (1.8–3.8 kb). Members of the genera *Circovirus* and *Cyclovirus* have proven to be readily detectable with metagenomic approaches [58,59]. A large diversity of these viruses has been reported from viromes of various mammalian hosts, including both domestic and wildlife species [58,59]. Infections of these viruses in domestic animals have been known to cause significant annual

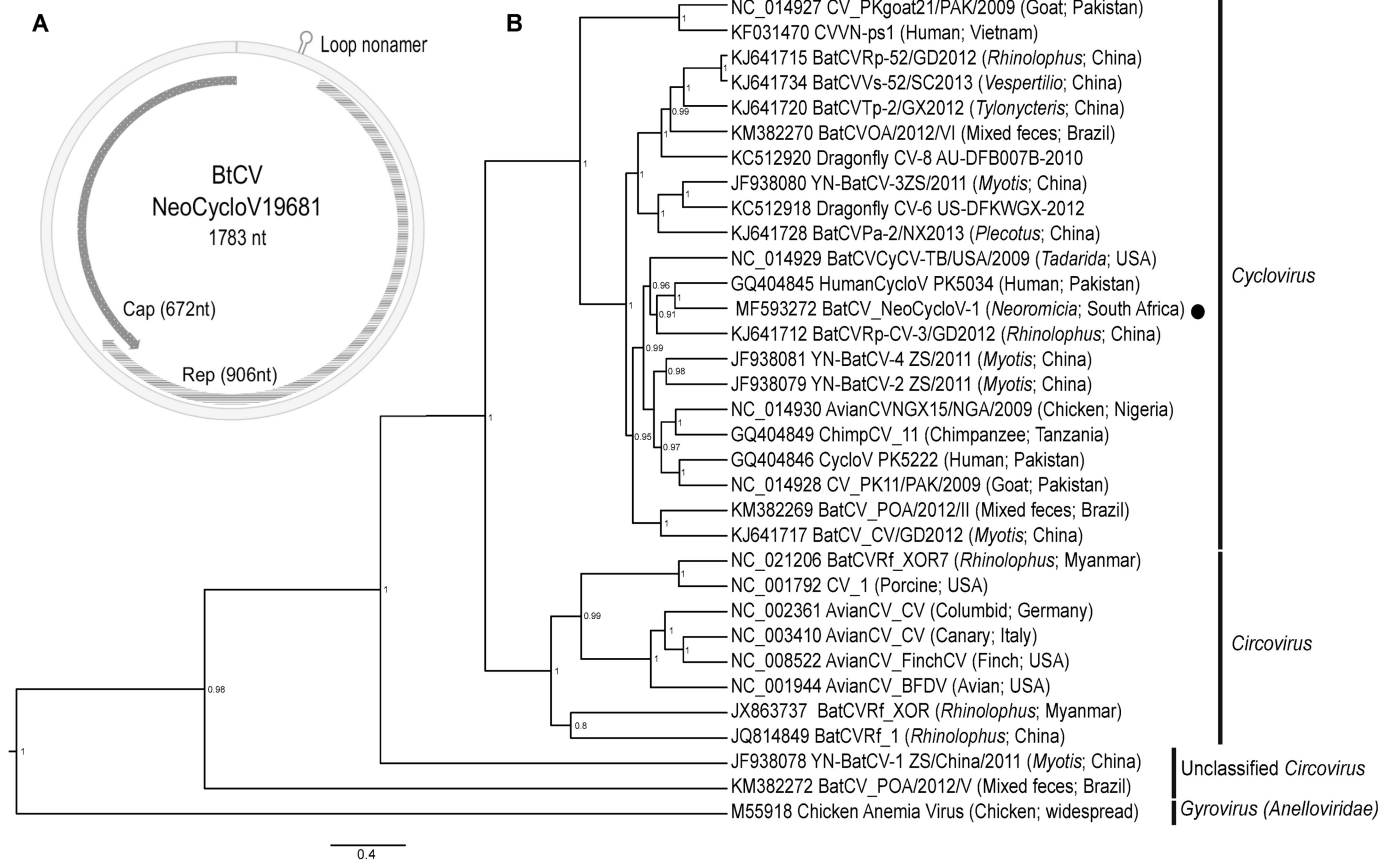


Fig 3. Novel *Neoromicia* cyclovirus. **A)** The circular genome organization of NeoCycloV-1 (19681/RSA) is shown with the rep and cap genes in opposing directions. The characteristic *Cyclovirus* nonamer is indicated at nucleotide position 99. **B)** Bayesian phylogeny of representative species of the *Circoviridae* using complete genomes. The phylogenetic tree was constructed in BEAST v1.8 using GTR plus invariant sites and gamma distribution substitution model. The genera are indicated on the right side of the sequence names (with indicated GenBank accession numbers), and the *Neoromicia* cyclovirus from this study is indicated with the black circle. Chicken anemia virus from the *Anelloviridae* family was used as an outgroup. Posterior probability values of less than 50% were omitted.

<https://doi.org/10.1371/journal.pone.0194527.g003>

economic losses to the poultry and pork industries [58]. The risk of zoonotic transmission for members of the *Circoviridae* is still largely unknown.

A putative novel *Cyclovirus* species was identified from the collective *Neoromicia* virome. The full-length genome of 1783bp was characterized to encode a 906bp Rep (replicase-associated protein) gene and 672bp Cap (capsid protein) gene (Fig 3A). The full genome phylogeny of the novel virus named *Neoromicia*-associated cyclovirus 1 strain 19681/RSA (or NeoCycloV-1), shows grouping within the genus *Cyclovirus*; including bat cycloviruses from the USA and China, as well as cycloviruses from human stool originating in Pakistan (Fig 3B) [11,14,58]. As with other cycloviruses, the genome of NeoCycloV-1 is smaller (1.7kb) than in relation to circoviruses (typically 2kb). Furthermore, characteristic to cycloviruses is the rolling circle replication initiation 5' intergenic loop nonamer (5' TAATACTAT' 3) identified at position 99-107bp of the genome (Fig 3A), as well as the absent 3' intergenic regions between the two major open reading frames [58].

The International Committee on Virus Taxonomy (ICTV) requires sequence similarities of less than 75% overall genome identity and less than 70% amino acid similarity of capsid proteins for designation as different species [58,59]. NeoCycloV-1 shares overall genome nucleotide identities of 44.2–64.9% between compared cycloviruses, with less than 40.8% identity to

members of the genus *Circovirus* (S5 Table). The capsid protein shared amino acid similarities of only 20.6–49.1% to all compared viruses in the family. In agreement with additional species demarcations suggested by Li et al. [58], requiring at least 85% amino acid similarity within the Rep protein, the NeoCycloV-1 shares between 48.4–73.3% similarity to compared cycloviruses. Similarities of the NeoCycloV-1 are well within accepted common thresholds for species in the genus *Cyclovirus*. Given it meets the ICTV criteria for species demarcation, we propose the new species, *Neoromicia*-associated cyclovirus 1 (strain 19681/RSA) within the genus *Cyclovirus*. Interestingly, the closest relative of the putative new *Neoromicia* cyclovirus originated from human stool; with an overall nucleotide identity of 64.9%, and amino acid similarities of 73.3% and 49.1% for the Rep protein and capsid protein, respectively. Additional cycloviruses are likely to be detected in bat species from Africa, which may be more closely related.

Contigs identified from the *Neoromicia* virome

Various contigs were identified from the data analysis of the *Neoromicia* virome. Shorter, fragmented contigs were obtained from the *Phenuiviridae* family as well as the *Parvoviridae* and *Papillomaviridae* (the latter two families are detailed in S1 File). The longest length contigs (1077–1806bp) originated from the *Picornaviridae* and *Adenoviridae* families.

Phenuiviridae. Recent taxonomic changes have reported the elevation of the family *Bunyaviridae* to the order *Bunyavirales*, with the known genera being reassigned as families (ICTV *Bunyavirales* taxon change, code 2016.030a-vM). Some families comprise arthropod-borne viruses, including a number of important public health pathogens such as severe fever with thrombocytopenia syndrome virus (SFTSV). Despite several contigs assigned as belonging to members of the *Bunyavirales*, only one was considered a valid viral contig after further inspection (as other contigs were shown to be false positives attributed to mammalian/host sequences).

The contig, Bt/PhenuiV/Neo11304/RSA, of 267bp aligned to a region of the L gene segment in the RNA dependent RNA polymerase (RdRp) region that may typically be used for phylogenetic analysis. Its closest matches were to novel Asian bunyaviruses identified in China from bats in the genus *Rhinolophus*, as well as mixed species bat guano from Vietnam (Fig 4) [8,60]. These viruses were classified as unassigned, though considered part of the *Peribunyaviridae* family, and shared between 63.9–68.7% pairwise nucleotide identities to the contig Bt/PhenuiV/Neo11304/RSA (S6 Table). A mosquito virus identified in China shared a basal position to the lineage of these bat viruses (60–61.8%). As in this study, these bat-associated virus sequences were identified from gastrointestinal specimens (faecal, rectum or gastrointestinal swabs), and therefore the possibility that these viruses may have originated from arthropods consumed by these insectivorous bat hosts cannot be excluded. Alternatively, part of the replication cycle of these bat-associated bunyaviruses may be shared in arthropod hosts.

Bt/PhenuiV/Neo11304/RSA shares low homologies (38.2–42.2% nucleotide identities) to viruses in the family *Peribunyaviridae* (such as Kaeng Khoi bat virus or Oropouche virus from the genus *Orthobunyavirus*). Similarities to members of the genus *Phlebovirus*, in the family *Phenuiviridae*, which includes SFTSV and Malsoor *Rousettus aegyptiacus* bat virus, range between 45–49% [61]. As a comparison, the *Neoromicia* hantavirus (*Hantaviridae*), Mouyassué virus from the Ivory Coast, shares only 32.9% nucleotide identity to the contig detected in this study [25].

The phylogeny indicates that Bt/PhenuiV/Neo11304/RSA and other bat viruses from China and Vietnam may be more related to the family *Phenuiviridae* (known to be frequently transmitted by arthropods) than to the family *Peribunyaviridae*. Though this is only tentatively

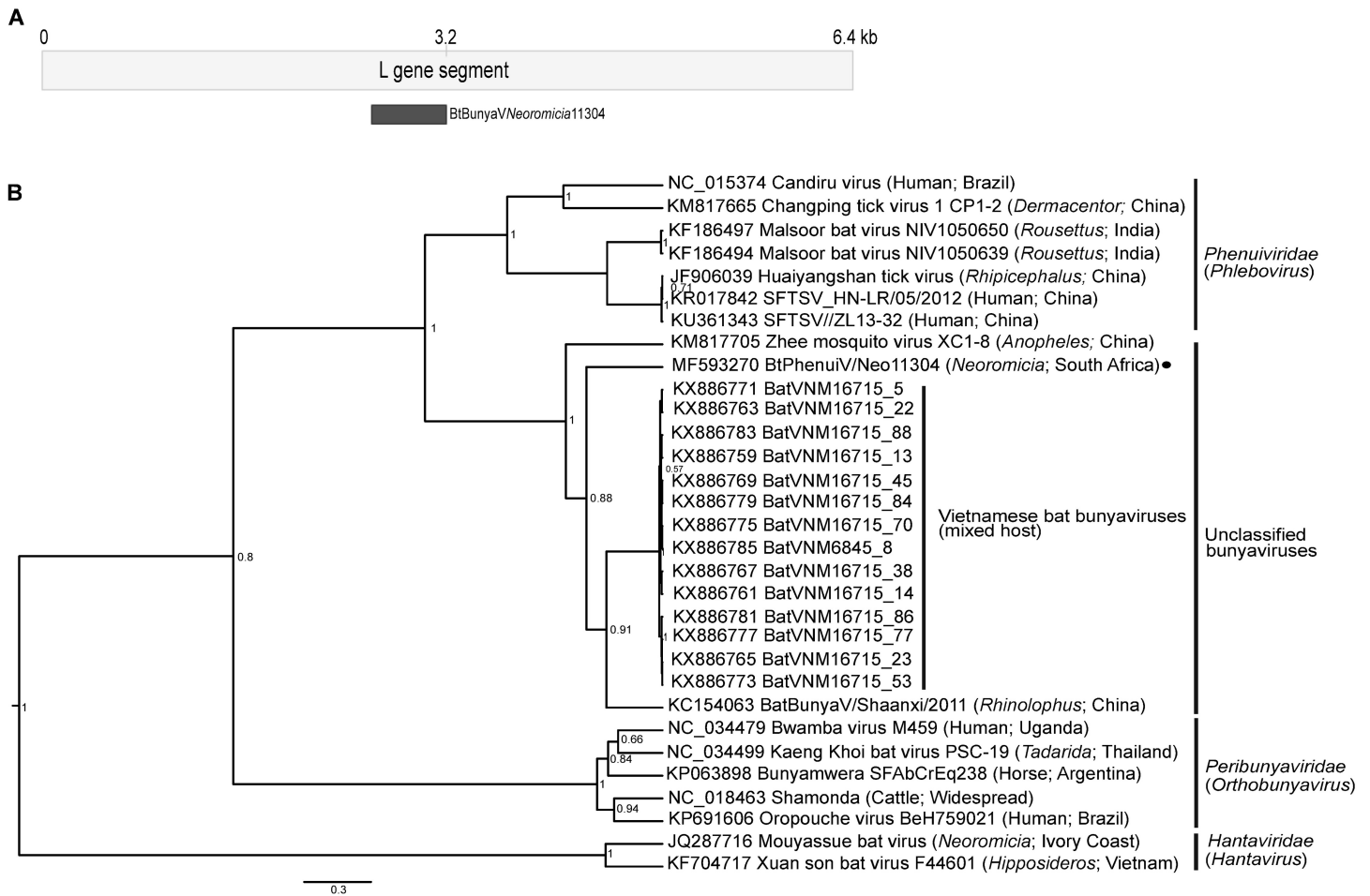


Fig 4. Phenuiviridae sequence identified from the *Neoromicia* virome. A) Alignment of the *Phenuiviridae* contig in reference to a typical L segment gene. B) The phylogenetic tree was constructed with a 267 bp region of the bunyavirus L genome segment using BEAST v1.8 with the GTR substitution model plus invariant sites. Relevant genera are shown on the right and GenBank accession numbers of each sequence are provided; the novel *Neoromicia* bunyavirus sequence from this study is indicated with a black circle. SFTSV = Severe fever with thrombocytopenia syndrome virus.

<https://doi.org/10.1371/journal.pone.0194527.g004>

based on a short segment of the RdRp gene and would require confirmation with complete genome segments. This short sequence provides an initial indication that members of this family are present in the genus *Neoromicia* and constitutes the only sequence data of a virus possibly associated with the *Phenuiviridae* family from insectivorous bats in Southern Africa. The low homology of Bt/PhenuiV/Neo11304/RSA to other available viruses from the *Bunyavirales* order is likely attributable to lacking comparable sequence data. Additionally, further viral surveillance activities directed toward insectivorous bats would greatly benefit from comparative surveillance initiatives of the arthropods frequently consumed by the bats.

Picornaviridae. Picornaviruses are small, single-stranded, positive sense RNA viruses with approximately 7.1–8.9 kb genomes [62]. These viruses are diverse and exploit several routes of transmission; resulting in respiratory disease caused by rhinoviruses, gastroenteric infections caused by entero- and kobuviruses, as well as hepatic infections caused by hepatitis A virus [62]. Due to rare reports of zoonotic spillover, picornaviruses have not generally been considered as zoonotic. The zoonotic potential of the few documented bat picornaviruses are thus unknown [11,62–64]. Bat-associated picornaviruses have been reported from Asia, Europe, Africa and the USA, and some have necessitated the creation of additional genera (e.g.

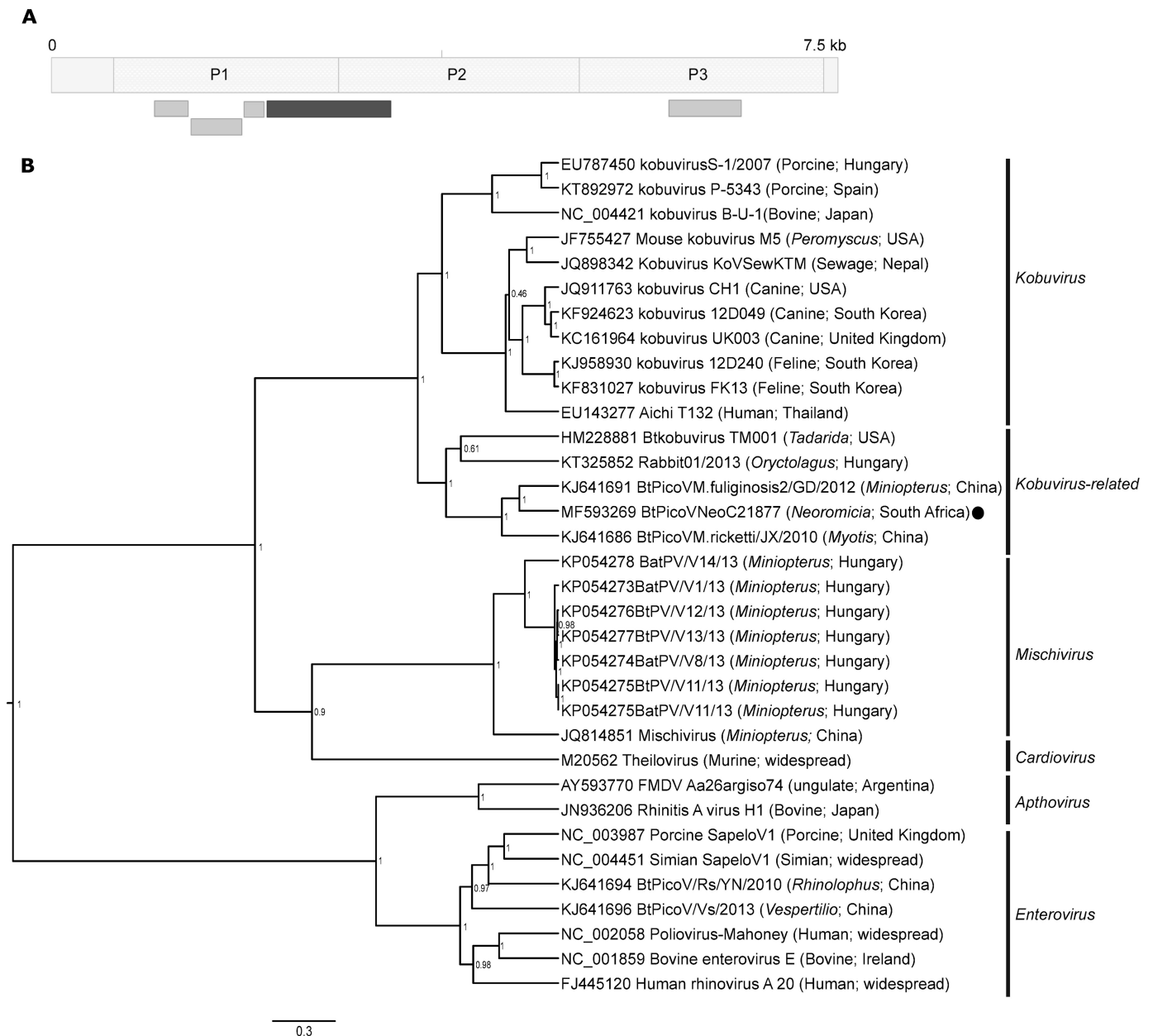


Fig 5. Novel *Neoromicia* picornavirus. **A)** Overview of the confirmed picornavirus contig alignment positions in reference to a typical picornavirus genome. The dark grey contig (1077bp) of the P1/P2 region was used to construct a Bayesian phylogeny. **B)** The maximum clade credibility tree constructed in BEAST v1.8 using the GTR plus invariant sites and gamma distribution substitution model. The genera are indicated on the right side of the sequence names (with GenBank accession numbers); the *Neoromicia* picornavirus from this study is indicated with a black circle. Posterior probability values of less than 50% were omitted.

<https://doi.org/10.1371/journal.pone.0194527.g005>

the genus *Mischivirus*) in a family that already contained 29 officially recognized genera [11,63]. Other bat-associated picornaviruses have also grouped with the genera *Sapelovirus*, *Kobuvirus* and *Enterovirus* [11,14,62,64].

Several contigs from the *Picornavirales* order were identified from the metagenomic data, some constituting insect infecting iflaviruses or were deemed as low complexity sequences upon further curation of the sequences (Fig 5A). A number of the contigs of mammalian-infecting picornavirus origin could be consolidated into a 1077bp contig, aligning to the P1

and P2 genome regions that encode for structural proteins (Fig 5B). Phylogenetic analysis groups the novel *Neoromicia* picornavirus sequence (BatPicV/NeoC21877/RSA) in a sister clade to those belonging to the genus *Kobuvirus* (Fig 5B). Pairwise alignments to other unassigned bat picornaviruses as well as a novel rabbit picornavirus within the clade indicate that they share between 55.1–73.6% nucleotide identity and 48–76.7% amino acid similarity. The *Miniopterus fuliginosus* bat picornavirus (BtMf-PicoV-2/GD2012) identified in China was the closest available relative [11]. The rabbit picornavirus had been tentatively placed into the *Kobuvirus* genus as it possessed certain qualifying characteristics [65]. However, this novel clade shares much lower similarities (46.3–54.9% nucleotide identities and 38.9–47.4% amino acid similarities) to other classified kobuviruses and decreased to below 39.7% nucleotide identity when compared to the genera *Mischivirus*, *Enterovirus*, *Sapelovirus*, *Aphovirus* and *Cardiovirus*.

To our knowledge, this is the first indication of a bat picornavirus from an insectivorous bat genus in Africa. The virome analysis of *Eidolon helvum*, a fruit bat from Ghana, also identified sequences of kobuvirus-related picornaviruses from urine; though it cannot be directly compared due to non-overlapping regions [17]. Kobuvirus species have been documented to infect both humans and domestic species such as cattle, dogs and cats, causing severe gastroenteritis [65]. Future surveillance for bat-associated kobuviruses will enable determination of zoonotic potential and discern whether these viruses may be classified as a new genus.

Viral families detected with insufficient sequencing depth

Low sensitivity and insufficient depth of sequencing was observed as fragmented viral genome coverage, and further noted from discrepancies between BLASTn assignments of reads and that of contigs; as reads data suggested the presence of more viral families than contigs. For example, no contigs were produced from the *Coronaviridae* family, though several pairs of reads were detected across distant genome regions. Several contigs of <430bp were designated as herpesvirus sequences, though were discarded after manual curation as being only mammalian-derived sequences. Some contigs were still identified as herpesvirus, though had poor e-value support. Severely fragmented viral genome coverage was observed for the *Adenoviridae*, with contigs aligning to multiple gene regions ranging in length from 200 to 1826bp (Fig 6A). Estimation of adenovirus genome coverage is further complicated by the detection of divergent sequences that align to diverse adenovirus reference genomes. These limitations may be attributed to overall insufficient depth of sequencing, the bioinformatics analyses methods implemented (BLASTn), or biases incorporated by certain nucleic acid preparation techniques for library construction—such as the use of randomly primed cDNA [66,67]. Similar events have been noted in the literature; where viruses were unsuccessfully or partially sequenced by high throughput sequencing though readily amplified with conventional PCR assays [10,68]. Due to limited computational resources, and surveillance emphasis focusing on identification of mammalian viral families with zoonotic potential, BLASTn analyses was deemed sufficient. However, more divergent viral sequences, particularly for non-frequently sequenced genome regions, may have been better detected with a BLASTx analyses using the complete NCBI nr database.

To confirm the presence of the herpes- and coronavirus sequences as well as determine the sequence diversity of adenoviruses, these three viral families were selected for further investigation with conventional nucleic acid detection assays based on conserved gene regions [37,38]. Analysis was conducted on sample material remaining from individuals following sample preparation for Illumina sequencing. Due to limited quantities of sample material, original samples of certain individuals were depleted during metagenomic sample processing

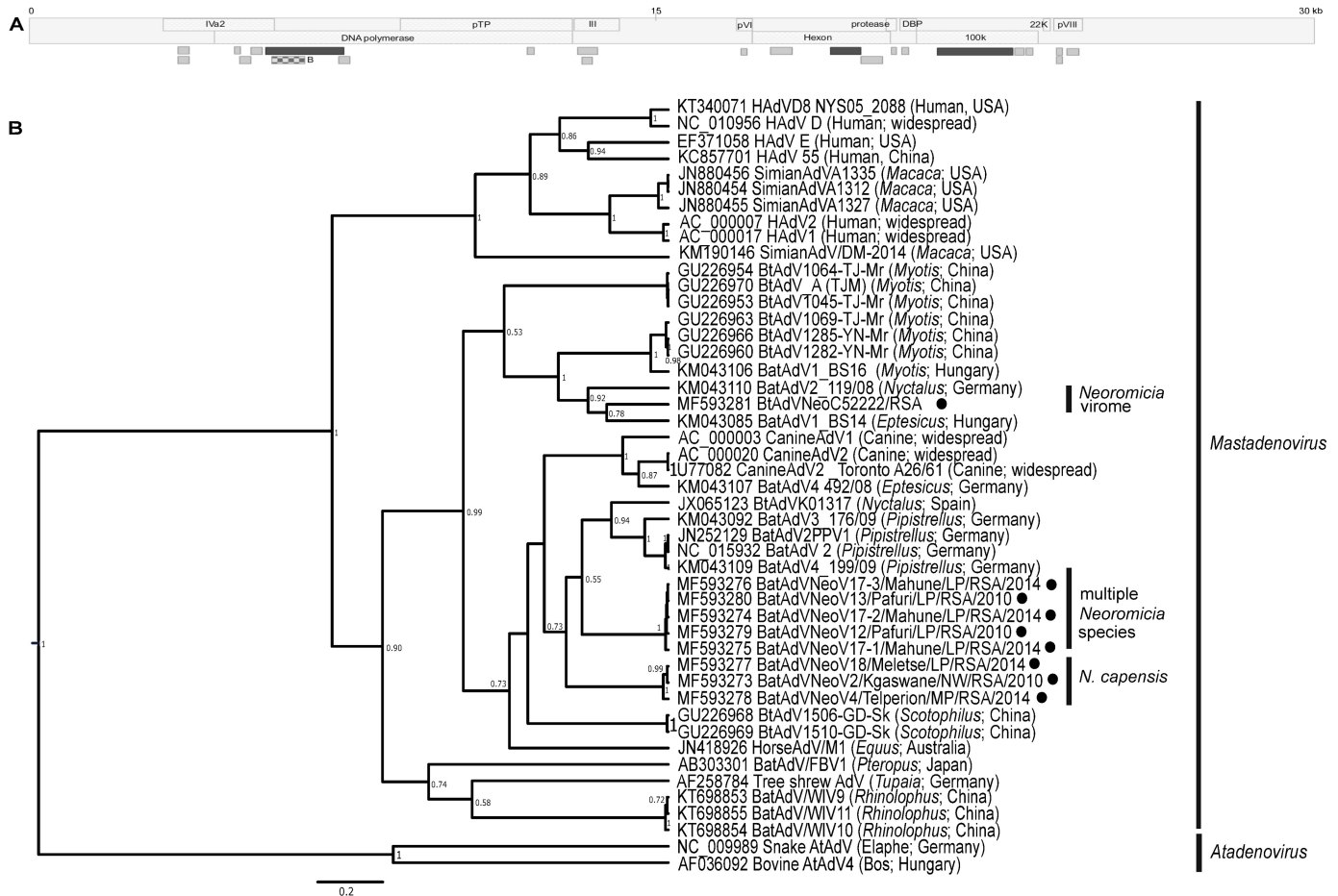


Fig 6. *Neoromicia* adenovirus sequences. A) Overview of confirmed mastadenovirus contigs from the *Neoromicia* virome created with Velvet and CLC assemblers as they align to a characteristic mastadenovirus genome. The dark grey contig was used in B, along with the amplicons produced by conventional PCR (depicted by the checkered block). B) Bayesian phylogenetic tree of a 237 bp region of the DNA polymerase gene. The phylogeny was constructed in BEAST v1.8 using the Hasegawa, Kishino and Yano (HKY) substitution model plus gamma distribution model suggested by J-model test. The MCMC chain was set to 20,000,000 generations sampled every 2000 steps, with a 10% burn-in of the first generated trees. Adenovirus sequences detected from this study are shown with black circles, and bat species from which adenoviruses originated are indicated on the right side of the sequence names. Posterior probability values of less than 50% were omitted. GenBank accession numbers are shown next to sequences.

<https://doi.org/10.1371/journal.pone.0194527.g006>

(Fig 2A). Therefore, as selected confirmatory nucleic acid tests could not be performed for each individual bat, available samples (faecal homogenates, rectal or intestinal samples) were pooled according to species, sampling date and location to associate selected viral families with specific species (S3 Table). Original sample material from pooled samples found to harbour coronavirus RNA were separately assayed to trace back the exact sampling location and date (Fig 2A).

Adenoviridae. Adenoviruses are double-stranded DNA viruses of approximately 35kb, which may cause a variety of respiratory and gastrointestinal infections within their human, domestic pet animal, and wildlife hosts [37]. Surveillance initiatives in multiple countries have identified bats as important hosts for the genetic diversity and evolutionary history of adenoviruses [19,69,70]. Adenoviruses have been shown to be capable of interspecies transmission and establishment within a new host due to the close relatedness of certain bat and canine adenoviruses [19]. Even though bat adenoviruses could potentially be transmitted via the aerosolization of bat faecal material, their potential zoonotic risk to humans seems low [71].

The PCR assay selected to confirm *Neoromicia* adenovirus sequence diversity was based on a conserved region (260bp) of the DNA polymerase gene [37]. This region overlapped with a 860bp contig (BtAdV/NeoC52222/RSA) identified from the high throughput sequencing data, though few bat adenovirus sequences of greater than 400bp were available for comparison. PCR analysis identified adenovirus DNA in eight of the remaining pooled samples. These sequences represented three clades (Fig 6B). The *N. capensis* sequences originating from three different provinces, and sampled over five years, grouped together and shared between 97.7–99.6% nucleotide identities. This clade was 78.7–79.8% identical to sequences in the mixed species clade of *N. nana*, *N. zuluensis* and *N. cf. helios* hosts from Limpopo province, which shared 98.8–100% nucleotide identity. Two of these sequences (BtAdV/NeoV13/LP/RSA/2010 and BtAdV/NeoV17-3/LP/RSA/2014) were identical, even though they were identified from bats collected four years apart in different locations of the Limpopo province. The polymerase contig BtAdV/NeoC52222/RSA, identified directly from the *Neoromicia* virome, shared lower similarities (64.7–65.5%) to the other groups of *Neoromicia* adenovirus sequences.

When compared to non-*Neoromicia* adenovirus sequences, contig BtAdVNeoC52222 shared the greatest similarity (79%) to a vesperilionid adenovirus from *Eptesicus serotinus* from Hungary [70]. The other *Neoromicia* adenovirus sequences shared greater nucleotide similarities (72.2–79.9%) to bat adenovirus B (BtAdV2) from *Pipistrellus pipistrellus* in Germany and other vesperilionid species from Hungary [70,72]. The DNA polymerase assay did not amplify any sequences able to confirm the origin of contig BtAdVNeoC52222—detected directly from the *Neoromicia* virome. The contig may have originated from one of the individuals for which specimen material was depleted, and was thus unable to be further analysed with the nucleic acid detection assay. The metagenomic analysis approach was unable to identify the diversity of adenovirus sequences confirmed with a conventional PCR approach—highlighting a limitation of the methods utilized for sample preparation or a lack of sufficient depth of sequencing.

Herpesviridae. The *Herpesviridae* is a large family of viruses subdivided into three subfamilies, the *Alphaherpesvirinae*, *Betaherpesvirinae*, and *Gammaherpesvirinae*. Bat-associated herpesviruses have been detected from all three subfamilies, with several that have yet to be classified [73,74]. Herpesviruses are known to cause recurrent or latent infections that may later become reactivated within particular cell types. Though no bat herpesviruses have been linked to zoonotic infections, a gammaherpesvirus isolated from a *Myotis* cell line was shown capable of growth in human derived cell lines [73]. Evolutionary analysis of gammaherpesviruses originating from a number of host orders suggest that multiple host switching events and independent viral cospeciation of lineages may previously have occurred; with bats and primates acting as sources of genetic diversity [75].

The conventional nested PCR assay used for herpesvirus detection also targeted a conserved region of the DNA polymerase gene [38]. The assay confirmed the presence of three herpesvirus sequences from *N. cf. helios* and *N. capensis* pooled tissue from the Limpopo and Mpumalanga provinces, respectively. The two herpesvirus sequences from *N. cf. helios* shared 99.5% nucleotide identity, and only 60.9–61.4% nucleotide identity to the sequence from *N. capensis*. These DNA polymerase sequences were compared to Genbank sequences and grouped with other known bat gammaherpesviruses—particularly *Scotophilus kuhlii* bat herpesviruses from China (Fig 7). The *N. cf. helios* herpesvirus sequences were more divergent (44.7–67.4% nucleotide similarity) in comparison to other bat herpesviruses than the sequence from *N. capensis*, which shared 74.9% similarity to BtHVSK_13YF125 (KR261860) [74]. The lack of larger sequenced regions of African bat herpesviruses from the subfamily may have contributed to the unsuccessful recovery of *Neoromicia* herpesvirus sequence information from the metagenomic data analysis, which would have improved phylogenetic placement.

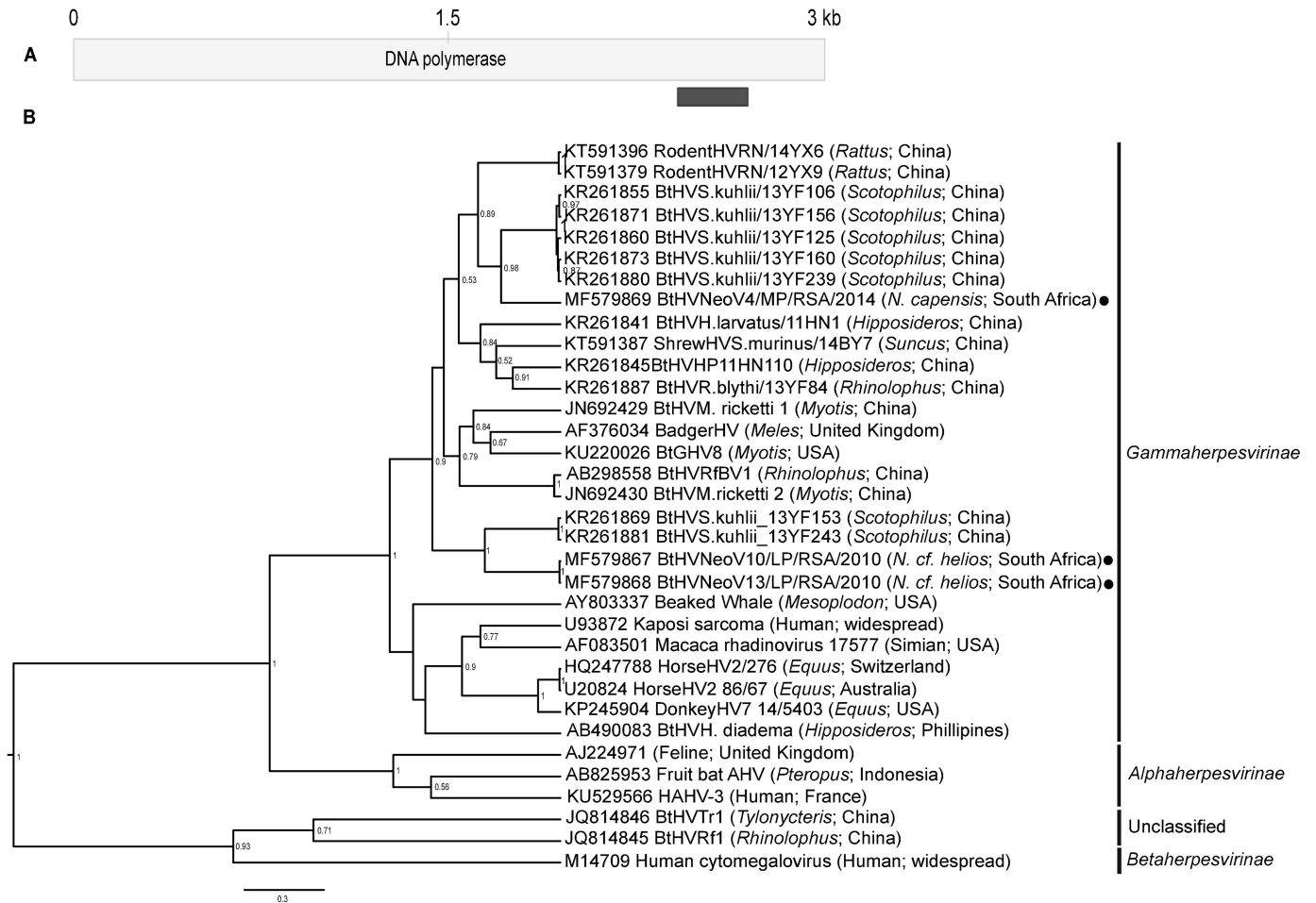


Fig 7. *Neoromicia herpesvirus* sequences. **A)** Region of the herpesvirus DNA polymerase targeted with the molecular detection assay [38]. **B)** The phylogenetic tree of a conserved 215 bp segment the herpesvirus DNA polymerase gene. The phylogeny was constructed in BEAST v1.8 using the GTR plus invariant sites and gamma distribution substitution model as suggested by J-model test. The *Neoromicia* herpesvirus sequences from this study are indicated with the black circles, and the species from which they originated are shown. The subfamily that compared sequences may possibly belong to, are indicated with GenBank accession numbers on the right side of the phylogenetic tree.

<https://doi.org/10.1371/journal.pone.0194527.g007>

Coronaviridae. Coronaviruses have large positive sense single-stranded RNA genomes of approximately 30kb. The *Coronavirinae* subfamily is split into four genera; species from the genera *Alphacoronavirus* and *Betacoronavirus* mainly infect mammalian hosts, while the species from the genera *Gammacoronavirus* and *Deltacoronavirus* largely infect avian hosts (with sporadic detections in mammals) [6]. The genus *Betacoronavirus* is divided into four lineages (A-D), with lineage C referring to MERS-CoV and related coronaviruses [76]. A number of paired reads suggested the presence of both alpha- and betacoronaviruses, with several pairs sharing high similarity to the betacoronavirus NeoCoV/PML-PHE1 [4]. Conventional nucleic acid detection with a genus-specific nested RT-PCR assay amplifying a 268-440bp conserved region of the RdRp gene confirmed coronavirus RNA in available specimen material from two *N. capensis* individuals. Sequences were extended to 648bp with combinations of RdRp grouping unit (RGU) primers [36].

The alphacoronavirus BtCoV*Neoromicia*1787/LP/RSA/2013 from intestinal tissue of an *N. capensis* bat, shared high similarity (94.6–95.3% nucleotide identities) to the previously mentioned South African alphacoronaviruses reported from the same host genus [27,28]. The

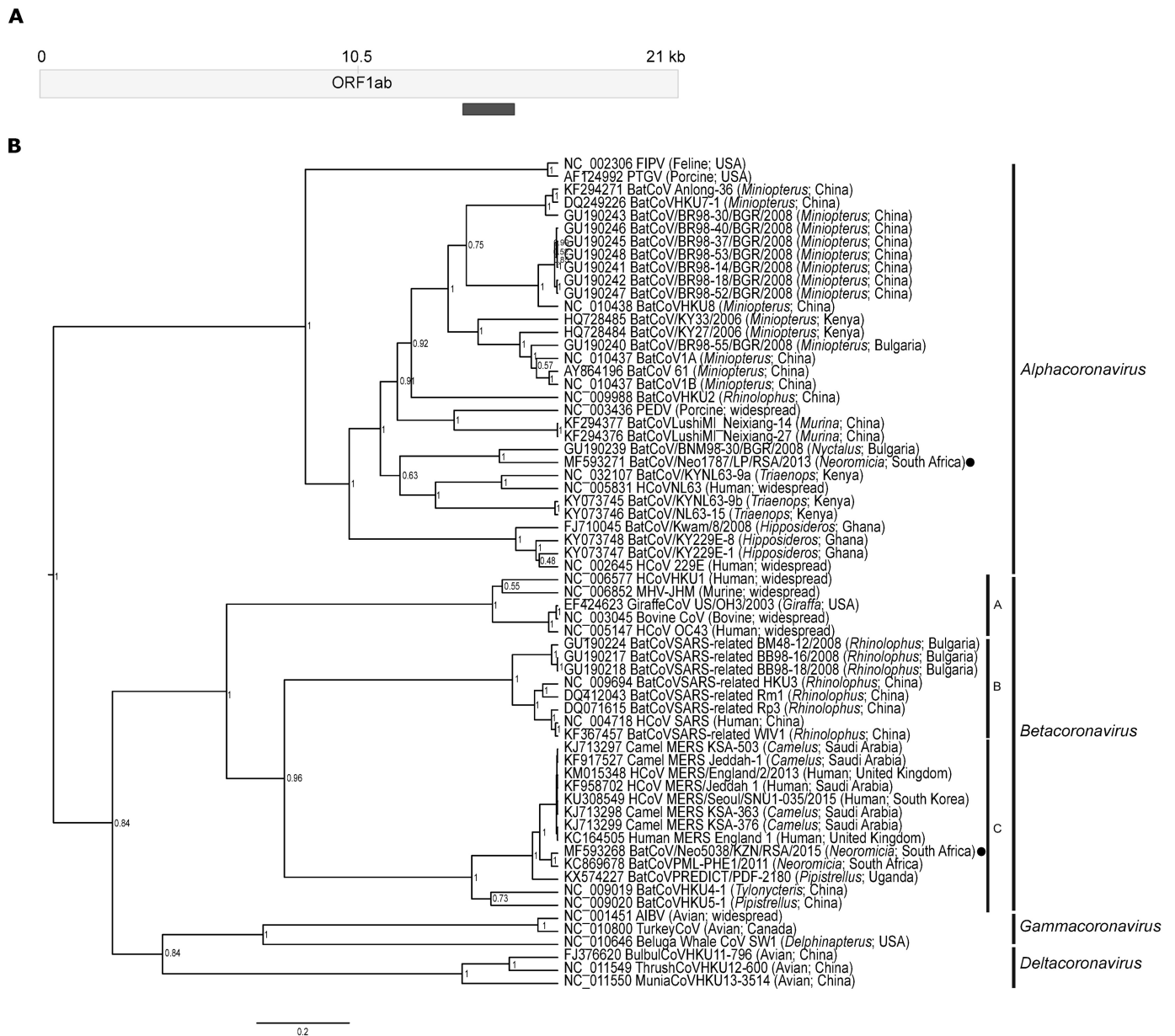


Fig 8. Conserved sequences of identified *Neoromicia* coronaviruses. A) Region of the coronavirus ORF1ab (within the RdRp gene) targeted with the molecular detection assay. B) The coronavirus phylogenetic tree generated using the 605bp region was constructed in BEAST v1.8 using the GTR plus invariant sites and gamma distribution substitution model. The MCMC chain was set to 20,000,000 generations, sampled every 2000 steps, with a 10% burn-in of the first generated trees. Coronavirus genera and *Betacoronavirus* lineages are indicated on the right side of the sequences. GenBank accession numbers are shown with each sequence and coronaviruses detected in this study are indicated with a black circle. Posterior probability values of less than 50% were omitted.

<https://doi.org/10.1371/journal.pone.0194527.g008>

host species was confirmed by morphological identification and cytochrome C barcode analysis [21,31]. The *Neoromicia* alphacoronaviruses form an independent lineage with genetically similar members detected throughout the geographic distribution of the *N. capensis* host species [27,28]. BtCoV*Neoromicia*1787/LP/RSA/2013 was utilized as a lineage representative for these *Neoromicia* alphacoronaviruses in Fig 8 (as the other sequences were of an insufficient length for comparison). The closest relatives to this lineage was a *Nyctalus lysleri* bat alphacoronavirus (BNM98-30/BGR/2008) from Bulgaria, as well as a clade containing the human

coronavirus NL63 (HCoVNL63) and related *Triaenops* bat viruses from Kenya [36,77,78]. The *Neoromicia* and *Nyctalus* alphacoronaviruses shared between 75.5–79.2% nucleotide identities to HCoVNL63, which is somewhat lower than the similarities of the NL63-related *Triaenops* bat viruses within this gene region (76–84% nucleotide identity) (S7 Table). Comparatively, *Hipposideros* bat coronaviruses identified in Ghana were suggested to have been possible early progenitors of the human alphacoronavirus HCoV229E [79], and shared greater similarities of 87–92% nucleotide identity to HCoV229E. Further surveillance and sequence analysis of the alphacoronaviruses from *N. capensis* may identify whether they can be considered as potential recombinant progenitors that may have participated in the emergence of HCoVNL63 [78].

The BtCoVNeo5038/KZN/RSA/2015 betacoronavirus (referred to as BtCoVNeo5038), was detected from the intestine (and the remaining faecal/rectal homogenate) of a *N. capensis* collected in the KwaZulu-Natal province. This virus shared 97.7% nucleotide identity to NeoCoV/PML-PHE1 within the compared RdRp gene region. This second variant was detected four years later, confirming the circulation of MERS-related viruses within the host genus. The complete coding genome of BtCoVNeo5038 (30009bp) was sequenced by further amplicon sequencing from 11 amplicons with 1000x coverage on Illumina's MiSeq platform. The 11 coding regions were in the characteristic coronavirus order, and compared to other lineage C betacoronaviruses (S8 Table). The overall genome similarities between BtCoVNeo5038 and its closest relative, NeoCoV/PML-PHE1, were 97.2% nucleotide identity and 98.5% amino acid similarity. The shared nucleotide identities of all genes were between 96–98.3% (98–99.5% amino acid similarities); with the gene sharing the lowest nucleotide similarity being the spike protein gene (96% nucleotide and 98.7% amino acid similarity). Due to similarities within the seven conserved ORFs required for species demarcation (97.7–100% amino acid similarity) [76], BtCoVNeo5038 may also be assigned to the species, *Middle East respiratory syndrome-related coronavirus*. The species includes the human and camel MERS-CoV strains, as well as the bat viruses NeoCoV/PML-PHE1 and PREDICT/PDF-2108 [4,5]. The latter virus was identified from *Pipistrellus cf. hesperidus* collected in Uganda.

BtCoVNeo5038 shares overall nucleotide identities of 85.4–85.5% and concatenated genome amino acid similarities of 88.1–88.2% to selected human and camel strains of MERS-CoV; comparable to similarities of these strains to NeoCoV/PML-PHE1 (S8 Table). Overall genome similarities to MERS-related *Pipistrellus* CoV PREDICT/PDF-2108 [5] were 86.2% nucleotide identity and 91% amino acid similarity. Overall nucleic and amino acid similarities show that the two *Neoromicia* MERS-related coronaviruses share higher genetic similarity to human and camel strains of MERS-CoV (S9 Table).

All four betacoronavirus lineages (A–D) were included in a full genome phylogenetic comparison (Fig 9A). Similarities of lineage C betacoronaviruses to other lineages range from 43–49% nucleotide identities, and within lineage C similarities are greater than 66%. Comparison of the SARS-CoV lineage (lineage B) to the MERS-CoV lineage, shows shorter phylogenetic distances between the human/civet SARS-CoV and bat associated SARS-related viruses, than between human/camel MERS-CoV and the presently detected bat associated MERS-related viruses (Fig 9B & 9C). At a genetic level, the SARS-related CoV strains circulating in *Rhinolophus* bat hosts share 87.8–95.5% nucleotide identities to the human/civet SARS-CoV strains. Comparatively, the bat MERS-related viruses from the *Neoromicia* and *Pipistrellus* genera share only 83–85.5% nucleotide identities to human/camel MERS-CoV strains.

Dromedary camels have been implicated as the reservoir hosts responsible for MERS-CoV transmission to humans in the Middle East [50]. However, it has also been suggested that the original progenitor viruses of MERS-CoV may have originated as a MERS-related coronavirus harboured by a vespertilionid bat [4,5,80]. The currently identified diversity of bat-borne MERS-related viruses, however, are not sufficiently similar (particularly within the spike

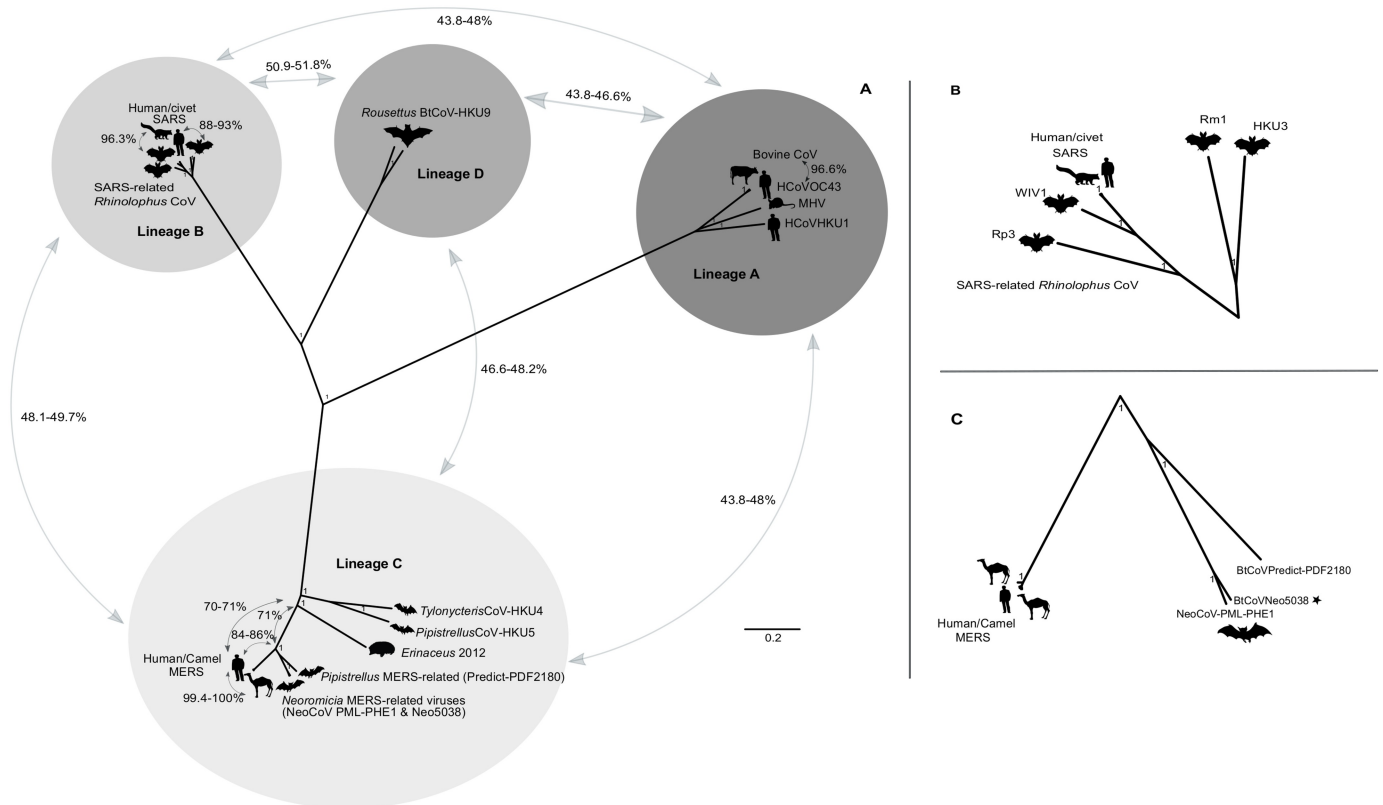


Fig 9. Betacoronavirus full genome phylogeny. A) The full genome phylogeny of 4 lineages (A-D) of the genus *Betacoronavirus* constructed using BEAST software with the GTR substitution model using invariant sites and gamma distribution. The MCMC chain was set to 15,000,000 generations sampled every 1500 steps, with a 10% burn-in of the first generated trees and displayed as a radial tree in Figtree. The lineages are indicated with clipart images of host species. Also displayed are the averaged pairwise similarities between lineages as well as highlighted similarities between human coronaviruses and related viruses identified in bats (and other animals). B) Close-up of the external nodes of the lineage B phylogeny to show relative distances of human and civet SARS-CoV strains and SARS-related *Rhinolophus* strains (WIV1, Rp3, Rm1 and HKU3). C) Close-up of the lineage C external nodes depicting the human and camel MERS strains with the SARS-related viruses (BtCoVNeo5038 from this study is indicated with a star). Sequence abbreviations and GenBank accession numbers are listed in S10 Table.

<https://doi.org/10.1371/journal.pone.0194527.g009>

protein) to human/camel strains of MERS-CoV to be considered as the possible progenitors. This would mirror the identification of similar SARS-related coronaviruses from *Rhinolophus* hosts. After 11 years of continued surveillance within *Rhinolophus* species, the WIV1 strain capable of utilizing the same binding receptor as human SARS-CoV was finally detected [7]. The finding would lead to the suggestion that through recombination of multiple SARS-related strains, certain SARS-related bat coronaviruses may have been capable of direct human infection [7].

Since there is no evidence supporting an equivalent assumption for MERS-related coronaviruses, coronavirus surveillance to identify additional strains or recombinant MERS-related viruses is crucial. Investigations focused in regions where the distributions of the known reservoir, dromedary camels, overlap with known vespertilionid bat hosts (such as *N. capensis*), would aid in ascertaining what viruses are circulating. Investigation of camels in these regions would also identify coronaviruses with the potential of recombining with other coronavirus species native to these hosts [81]. Overlapping distributions between dromedary camels and *N. capensis*, particularly where MERS antibody prevalence has been reported from camel populations in Africa and the Arabian Peninsula [50–55], indicate the best sampling localities as Sudan, Ethiopia, Somalia and Kenya (Fig 10). Coronavirus surveillance in these regions can be

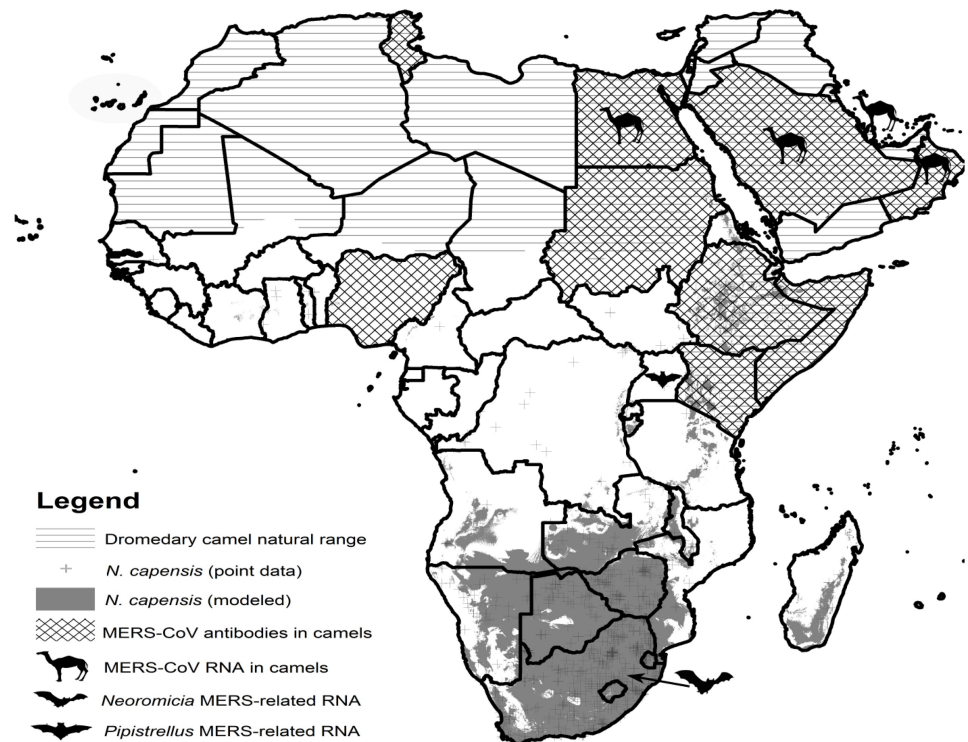


Fig 10. Map of Africa depicting the overlapping distributions of dromedary camels and cape serotine bats as hosts of MERS and related coronaviruses. The map was constructed in ArcMap v.10.4.1. The geographic distribution of dromedary camels are depicted with horizontal lines, with seroprevalence data of MERS antibodies detected from surveillance activities in camels shown with crossed lines. The distributions of *N. capensis* were taken from museum collections (point data) and thus extrapolated as modelled data. Clipart images of camels or bats show where viral RNA of MERS and MERS-related strains and have been reported.

<https://doi.org/10.1371/journal.pone.0194527.g010>

readily conducted with non-invasive sampling of bats and camels (collection of faecal material, oral swabs, and nasal secretions). Improved surveillance may yield additional stains of MERS-related viruses that could be used for isolation attempts in cell lines derived from bats, camels and human tissue. This may in turn allow for functional studies involving isolates that determine host susceptibility and permissibility.

Conclusions

Multi-pathogen surveillance approaches are integral to pathogen discovery programs that aim to identify potential public health risks, and also increase our knowledge of virus diversity and evolution [82]. The sequence-independent manner utilized by metagenomic high throughput sequencing methodologies enables detection of both known and unknown viral species that may not have been detected with conventional nucleic acid methods. Marked limitations of the metagenomic approach implemented here were highlighted by inadequate sequencing of several viral families, which could subsequently be detected with conventional PCR assays; such as the adenoviruses, herpesviruses and coronaviruses. In spite of the lack of coronavirus contigs produced from the metagenomic data, the complete coding genome of a MERS-related betacoronavirus was still recovered with additional amplicon sequencing. Since the experimental portion of the study was conducted, alternative methods have been suggested that may reduce these biases and improve sequencing results from nonclinical samples with low viral

nucleic acid concentrations—such as utilization of the ScriptSeq library kit (Illumina) directly after extraction of RNA [66].

Despite limitations with the metagenomic sequencing output, the study identified a novel *Cyclovirus* species, confirmed MERS-related virus circulation within this host genus, detected diverse adenoviruses and herpesviruses which are widespread among *Neoromicia* populations in South Africa, and determined that adenoviruses seemingly persist within these populations throughout several years. Follow up longitudinal studies can be implemented to confirm this finding and establish the total duration of the viral persistence. The *Neoromicia* adenovirus sequences shared high similarity to those identified in European bats, whereas the *Neoromicia* herpesviruses were much more diverse than previously identified bat-associated viruses. This observation may reflect differences in sampling efforts applied to each viral family. Further investigation of the identified viral families are required to sequence complete genes involved in receptor recognition and attachment to host cells. In the absence of viral isolates, this would allow functional assessment of the receptors utilized for cell entry, enable estimations of their potential to spread to new species, and assess the risks they pose to public or veterinary health. Lastly, the novel sequence data generated from the *Neoromicia* virome can be utilized in assay development for additional nucleic acid detection surveillance activities, to determine the prevalence rates of selected novel viruses.

At present, metagenomic high throughput sequencing may be unsuited for routine viral surveillance practices, as it may be restrictive in terms of sensitivity, incapable of detecting of the complete viral diversity, slow in turn-over time due to extensive bioinformatics data analysis or limited as a result of the high cost of large sequencing volumes. Future improvements to sample preparation and data analysis techniques would be invaluable, and enable these methodologies to be used routinely in strategies for pathogen discovery programs, with the ultimate goal of being aware of high-risk viral species that may be present in wildlife populations.

Supporting information

S1 File. Virome sequence data information from the *Parvoviridae* and *Papillomaviridae* families (with references and figures).

(DOCX)

S1 Table. *Neoromicia* samples collected to investigate the South African *Neoromicia* virome.

(PDF)

S2 Table. CLC Genomics workbench bioinformatics workflow parameters.

(PDF)

S3 Table. *Neoromicia* samples pooled for molecular detection of selected viruses.

(PDF)

S4 Table. *Alpha*- and *Betacoronavirus* genera hemi-nested RT-PCR primers.

(PDF)

S5 Table. Pairwise similarities inferred from distance estimations of full length genomes of the *Circoviridae* family. Sequence similarities of viruses in the *Circoviridae* family inferred from estimated evolutionary divergence calculated from pairwise distances. Full genomes were aligned and trimmed to 1075 overlapping positions. All ambiguous positions were removed for each sequence pair. Analyses were conducted in MEGA7 [41].

(PDF)

S6 Table. Pairwise similarities inferred from distance estimations of an L gene region between selected *Bunyavirales*. The table shows pairwise sequence similarities inferred from evolutionary divergence estimates of 249 positions of compared *Bunyavirales*. The number of base differences per site from between sequences were converted to percentage of similarities. Standard errors for distance estimates are shown above the diagonal. Codon positions included were 1–3 as well as noncoding. Ambiguous positions were removed for each sequence pair as per pairwise deletion. Estimates were analysed in MEGA7 [41].
(PDF)

S7 Table. Pairwise similarities inferred from distance estimations of a 605bp conserved segment of the coronavirus RNA dependent RNA polymerase gene. The table shows pairwise sequence similarities inferred from evolutionary divergence estimates of 605 positions of compared coronaviruses. The number of base differences per site from between sequences were converted to percentage of similarities. The sequences from this study are highlighted in grey and closest similarities to sequences from other studies are indicated in bold. Standard errors for distance estimates are shown above the diagonal in grey text. Codon positions included were 1–3 as well as noncoding and ambiguous positions were removed for each sequence pair as per pairwise deletion. Estimates were analysed in MEGA 7 [41].
(PDF)

S8 Table. Genome annotation of BtCoVNeo5038 with similarities to compared lineage C betacoronaviruses. The percentage similarities were inferred from pairwise distance estimates of the base pair and amino acid differences per site for each gene and of the overall genome. The estimates were calculated in MEGA7 [41] using pairwise deletion to treat gaps. All percentage similarities are given in comparison to BtCoVNeoV5038. Accession number of compared betacoronaviruses are listed in order: KC869678.4, KX574227, EF065505.1, EF065509.1, JX869059.2, KF958702.1, KF917527.1, and KJ477102.1.
(PDF)

S9 Table. Pairwise similarities inferred from distance estimations between betacoronavirus full genomes. The table shows pairwise similarities inferred from evolutionary divergence estimates of betacoronavirus full genomes. The number of base differences per site between sequences were converted to percentage similarities. Within lineage similarities are indicated in shaded blocks. Standard errors for distance estimates are shown above the diagonal in grey text. Codon positions included were 1–3 as well as noncoding and ambiguous positions were removed for each sequence pair as per pairwise deletion. Estimates were analysed in MEGA 7 [41].
(PDF)

S10 Table. Coronavirus strains and Genbank accession numbers of sequences used in the full genome phylogeny.
(PDF)

Acknowledgments

We would like to acknowledge the following people for allowing access to properties and logistical support: Aquila Steel Thabazimbi (PTY) Ltd. (Mike Halliday and Johann van Brede), E. Oppenheimer & Son (Duncan McFadden), Jonker Family Trust (Attie, Kathy, Riaan and Theamari), Leon Labuschagne and Yolanda Roux, Wilderness Safaris (Chris Roche and Walter Jubber).

Author Contributions

Conceptualization: Marike Geldenhuys, Marinda Mortlock.

Data curation: Marike Geldenhuys, Teresa Kearney.

Formal analysis: Marike Geldenhuys, Oliver Bezuidt, Tracy H. Erkkila.

Funding acquisition: Jacqueline Weyer, Wanda Markotter.

Investigation: Marike Geldenhuys, Ernest C. J. Seamark, Teresa Kearney.

Methodology: Marike Geldenhuys, Marinda Mortlock, Jacqueline Weyer, Cheryl Gleasner.

Project administration: Helen Cui.

Resources: Ernest C. J. Seamark, Teresa Kearney, Cheryl Gleasner, Helen Cui, Wanda Markotter.

Software: Oliver Bezuidt, Tracy H. Erkkila.

Supervision: Cheryl Gleasner, Tracy H. Erkkila, Helen Cui, Wanda Markotter.

Writing – original draft: Marike Geldenhuys.

Writing – review & editing: Marinda Mortlock, Jacqueline Weyer, Oliver Bezuidt, Ernest C. J. Seamark, Teresa Kearney, Cheryl Gleasner, Tracy H. Erkkila, Helen Cui, Wanda Markotter.

References

1. Rahman MA, Khan SU, Rahman M, Gurley ES, Rollin PE, Lo MK, et al. Date Palm Sap Linked to Nipah Virus Outbreak in Bangladesh, 2008. *Vector borne zoonotic Dis.* 2012; 12(1):65–72. <https://doi.org/10.1089/vbz.2011.0656> PMID: 21923274
2. Amman BR, Carroll SA, Reed ZD, Sealy TK, Balinandi S, Swanepoel R, et al. Seasonal pulses of Marburg virus circulation in juvenile *Rousettus aegyptiacus* bats coincide with periods of increased risk of human infection. *PLoS Pathog.* 2012; 8(10).
3. Lau SKP, Woo PCY, Li KSM, Huang Y, Tsoi H-W, Wong BHL, et al. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *PNAS.* 2005; 102(39):14040–5. <https://doi.org/10.1073/pnas.0506735102> PMID: 16169905
4. Corman VM, Ithete NL, Richards R, Schoeman MC, Preiser W, Drosten C, et al. Rooting the phylogenetic tree of Middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat. *J Virol.* 2014; 88(19):11297–303. <https://doi.org/10.1128/JVI.01498-14> PMID: 25031349
5. Anthony SJ, Gilardi K, Menachery VD, Goldstein T, Ssebide B, Mbabazi R, et al. Further evidence for bats as the evolutionary source of Middle East respiratory syndrome coronavirus. *MBio.* 2017; 8(2):1–13.
6. Drexler JF, Corman VM, Drosten C. Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. *Antiviral Res.* 2014; 101(1):45–56.
7. Ge X, Li J, Yang X, Chmura AA, Zhu G, Epstein JH, et al. Isolation and Characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature.* 2013; 503(7477):535–8. <https://doi.org/10.1038/nature12711> PMID: 24172901
8. Wu Z, Yang L, Ren X, He G, Zhang J, Yang J, et al. Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *ISME J.* 2015; 10:1–12.
9. Wang J, Moore NE, Murray ZL, McInnes K, White J, Tompkins DM, et al. Discovery of novel virus sequences in an isolated and threatened bat species, the New Zealand lesser short-tailed bat (*Mystacina tuberculata*). *J Virol.* 2015; 96:2442–52.
10. Drexler JF, Corman VM, Müller MA, Maganga GD, Vallo P, Binger T, et al. Bats host major mammalian paramyxoviruses. *Nat Commun.* 2012; 3(796):1–12.
11. Wu Z, Ren X, Yang L, Hu Y, Yang J, He G, et al. Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces. *J Virol.* 2012; 86(20):10999–1012. <https://doi.org/10.1128/JVI.01394-12> PMID: 22855479

12. Drexler JF, Geipel A, König A, Corman VM, van Riel D, Leijten LM, et al. Bats carry pathogenic hepadnaviruses antigenically related to hepatitis B virus and capable of infecting human hepatocytes. *PNAS*. 2013; 110(40):16151–6. <https://doi.org/10.1073/pnas.1308049110> PMID: 24043818
13. Tang XC, Zhang JX, Zhang SY, Wang P, Fan X, H, Li LF, et al. Prevalence and genetic diversity of coronaviruses in bats from China. *J Virol*. 2006; 80(15):7481–90. <https://doi.org/10.1128/JVI.00697-06> PMID: 16840328
14. Li L, Victoria JG, Wang C, Jones M, Fellers GM, Kunz TH, et al. Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses. *J Virol*. 2010; 84(14):6955–65. <https://doi.org/10.1128/JVI.00501-10> PMID: 20463061
15. Ge X, Li Y, Yang X, Zhang H, Zhou P, Zhang Y, et al. Metagenomic analysis of viruses from bat fecal samples reveals many novel viruses in insectivorous bats in China. *J Virol*. 2012; 86(8):4620–30. <https://doi.org/10.1128/JVI.06671-11> PMID: 22345464
16. Donaldson EF, Haskew AN, Gates JE, Huynh J, Moore CJ, Frieman MB. Metagenomic analysis of the viromes of three North American bat species: viral diversity among different bat species that share a common habitat. *J Virol*. 2010; 84(24):13004–18. <https://doi.org/10.1128/JVI.01255-10> PMID: 20926577
17. Baker KS, Leggett RM, Bexfield NH, Alston M, Daly G, Todd S, et al. Metagenomic study of the viruses of African straw-coloured fruit bats: Detection of a chiropteran poxvirus and isolation of a novel adenovirus. *Virology*. 2013; 441(2):95–106. <https://doi.org/10.1016/j.virol.2013.03.014> PMID: 23562481
18. Dacheux L, Cervantes-Gonzalez M, Guigon G, Thiberge JM, Vandenbogaert M, Maufrais C, et al. A preliminary study of viral metagenomics of french bat species in contact with humans: Identification of new mammalian viruses. *PLoS One*. 2014; 9(1):e87194. <https://doi.org/10.1371/journal.pone.0087194> PMID: 24489870
19. Kohl C, Vidovszky MZ, Mühldorfer K, Dabrowski PW, Radonić A, Nitsche A, et al. Genome analysis of bat adenovirus 2: indications of interspecies transmission. *J Virol*. 2012; 86(3):1888–92. <https://doi.org/10.1128/JVI.05974-11> PMID: 22130531
20. Quan P, Firth C, Conte JM, Williams SH, Zambrana-Torrel CM, Anthony SJ, et al. Bats are a major natural reservoir for hepaciviruses and pegiviruses. *PNAS*. 2013; 110(20):8194–9. <https://doi.org/10.1073/pnas.1303037110> PMID: 23610427
21. ACR. African Chiroptera report. Van Cakenberghe V, Seamark ECJ, editors. Pretoria: AfricanBats, Pretoria, Republic of South Africa ISSN: 1990-6471; 2016.
22. Goodman SM, Kearney T, Ratsimbazafy MM, Hassanin A. Description of a new species of *Neoromicia* (Chiroptera: Vespertilionidae) from southern Africa: A name for *N. cf. melckorum*. *Zootaxa*. 2017; 4236(2):zootaxa.4236.2.10.
23. Jacobs D, Cotterill FPD, Taylor PJ. *Neoromicia capensis*. The IUCN Red List of Threatened Species 2014: e.T44918A67358046. 2014.
24. Hutson AM, Racey PA, Goodman S, Jacobs D. *Neoromicia nana*. The IUCN Red List of Threatened Species 2014: e.T44923A67357605. 2014.
25. Sumibcay L, Kadjo B, Gu SH, Kang HJ, Lim BK, Cook JA, et al. Divergent lineage of a novel hantavirus in the banana pipistrelle (*Neoromicia nanus*) in Cote d'Ivoire. *Virology*. 2012; 9(34).
26. Mortlock M, Kuzmin I V, Weyer J, Gilbert AT, Agwanda B, Rupprecht CE, et al. Novel paramyxoviruses in bats from sub-Saharan Africa, 2007–2012. *Emerg Infect Dis*. 2015; 21(10):1840–3. <https://doi.org/10.3201/eid2110.140368> PMID: 26402433
27. Geldenhuys M, Weyer J, Nel LH, Markotter W. Coronaviruses in South African bats. *Vector borne zoonotic Dis*. 2013; 13(7):516–9. <https://doi.org/10.1089/vbz.2012.1101> PMID: 23473214
28. Ithete NL, Stoffberg S, Corman VM, Cottontail VM, Richards LR, Schoeman MC, et al. Close relative of human Middle East respiratory syndrome coronavirus in bat, South Africa. *Emerg Infect Dis*. 2013; 19(10):1697–9. <https://doi.org/10.3201/eid1910.130946> PMID: 24050621
29. Gloza-Rausch F, Ipsen A, Seebens A, Götsche M, Panning M, Felix Drexler J, et al. Detection and prevalence patterns of group I coronaviruses in bats, northern Germany. *Emerg Infect Dis*. 2008; 14(4):626–31. <https://doi.org/10.3201/eid1404.071439> PMID: 18400147
30. Monadjem A, Taylor PJ, Cotterill W, Schoeman MC. Bats of southern and central Africa: a biogeographic and taxonomic synthesis. Johannesburg: Wits University Press; 2010.
31. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol*. 1994; 3(5):294–9. PMID: 7881515
32. Kohl C, Brinkmann A, Dabrowski PW, Radoni A, Nitsche A, Kurth A. Protocol for metagenomic virus detection in clinical specimens. *Emerg Infect Dis*. 2015; 21(1):49–57.

33. Victoria JG, Kapoor A, Dupuis K, Schnurr DP, Delwart EL. Rapid identification of known and new RNA viruses from animal tissues. *PLoS Pathog.* 2008; 4(9):e1000163. <https://doi.org/10.1371/journal.ppat.1000163> PMID: 18818738
34. Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic data. *2007*;377–86.
35. Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 2008; 18(5):821–9. <https://doi.org/10.1101/gr.074492.107> PMID: 18349386
36. Drexler JF, Gloza-Rausch F, Glende J, Corman VM, Muth D, Goettsche M, et al. Genomic characterization of severe acute respiratory syndrome-related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. *J Virol.* 2010; 84(21):11336–49. <https://doi.org/10.1128/JVI.00650-10> PMID: 20686038
37. Li Y, Ge X, Zhang H, Zhou P, Zhu Y, Zhang Y, et al. Host range, prevalence, and genetic diversity of adenoviruses in bats. *J Virol.* 2010; 84(8):3889–97. <https://doi.org/10.1128/JVI.02497-09> PMID: 20089640
38. Van Devanter DR, Warrener P, Bennett L, Schultz ER, Coulter S, Garber RL, et al. Detection and analysis of diverse herpesviral species by consensus primer PCR. *J Clin Microbiol.* 1996; 34(7):1666–71. PMID: 8784566
39. Wang S, Sundaram JP, Spiro D. VIGOR, an annotation program for small viral genomes. *BMC Bioinformatics.* 2010; 11(451):1–10.
40. Hall T. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* 1999; 41:95–8.
41. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 2016; 33(7):1870–4. <https://doi.org/10.1093/molbev/msw054> PMID: 27004904
42. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol.* 2012; 29(8):1969–73. <https://doi.org/10.1093/molbev/mss075> PMID: 22367748
43. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and high-performance computing. *Nat Methods.* 2015; 9(8):6–9.
44. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE)*. New Orleans, LA; 2010. p. 1–8.
45. Sayre R, Comer P, Hak J, Josse C, Bow J, Warner H, et al. A new map of standardized terrestrial ecosystems of Africa. *African Geogr Rev.* 2013;1–24.
46. Faye B. Role, distribution and perspective of camel breeding in the third millennium economies. *Emirates J Food Agric.* 2015; 27(4):318–27.
47. Wilson RT. The one-humped camel in Southern Africa: Unusual and new records for seven countries in the Southern African development community. *African J Agric Res.* 2013; 8(28):3716–23.
48. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol.* 2005; 25:1965–78.
49. Phillips SB, Aneja VP, Kang D, Arya SP. Modelling and analysis of the atmospheric nitrogen deposition in North Carolina. *Int J Glob Environ Issues.* 2006; 6(2–3):231–52.
50. Reusken CBEM, Haagmans BL, Müller MA, Gutierrez C, Godeke G-J, Meyer B, et al. Middle East respiratory syndrome coronavirus in dromedary camels: an outbreak investigation. *Lancet.* 2013; 3099(13):1–6.
51. Reusken CBEM, Messadi L, Feyisa A, Ullaramu H, Godeke GJ, Danmarwa A, et al. Geographic distribution of MERS coronavirus among dromedary camels, Africa. *Emerg Infect Dis.* 2014; 20(8):1370–4. <https://doi.org/10.3201/eid2008.140590> PMID: 25062254
52. Corman VM, Jores J, Meyer B, Younan M, Lijander A, Said MY, et al. Antibodies against MERS coronavirus in Dromedary camels, Kenya, 1992–2013. *Emerg Infect Dis.* 2014; 20(8):1319–22. <https://doi.org/10.3201/eid2008.140596> PMID: 25075637
53. Müller MA, Corman VM, Jores J, Meyer B, Younan M, Lijander A, et al. MERS coronavirus neutralizing antibodies in camels, eastern Africa, 1983–1997. *Emerg Infect Dis.* 2014; 20(12):2093–5. <https://doi.org/10.3201/eid2012.141026> PMID: 25425139
54. Chu DKW, Poon LLM, Goma MM, Shehata MM, Perera RAPM, Zeid DA, et al. MERS coronaviruses in dromedary camels, Egypt. *Emerg Infect Dis.* 2014; 20(6):1049–53. <https://doi.org/10.3201/eid2006.140299> PMID: 24856660
55. Hemida MG, Elmoslemany A, Al-Hizab F, Alnaeem A, Almuthen F, Faye B, et al. Dromedary camels and the transmission of Middle East respiratory syndrome coronavirus (MERS-CoV). *Transbound Emerg Dis.* 2015; 64(2):344–53. <https://doi.org/10.1111/tbed.12401> PMID: 26256102

56. He B, Li Z, Yang F, Zheng J, Feng Y, Guo H, et al. Virome profiling of bats from Myanmar by metagenomic analysis of tissue samples reveals more novel Mammalian viruses. *PLoS One*. 2013; 8(4): e61950. <https://doi.org/10.1371/journal.pone.0061950> PMID: 23630620
57. Zhao G, Wu G, Lim ES, Droit L, Krishnamurthy S, Barouch DH, et al. VirusSeeker, a computational pipeline for virus discovery and virome composition analysis. *Virology*. 2017; 503:21–30. <https://doi.org/10.1016/j.virol.2017.01.005> PMID: 28110145
58. Li L, Kapoor A, Slikas B, Bamidele OS, Wang C, Shaikat S, et al. Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. *J Virol*. 2010; 84(4):1674–82. <https://doi.org/10.1128/JVI.02109-09> PMID: 20007276
59. de S Lima FE, Cibulski SP, dos Santos HF, Teixeira TF, Varela APM, Roehle PM, et al. Genomic characterization of novel circular ssDNA viruses from insectivorous bats in Southern Brazil. *PLoS One*. 2015; 10(2):e0118070. <https://doi.org/10.1371/journal.pone.0118070> PMID: 25688970
60. Oude Munnink BB, Phan MVT, van der Hoek L, Kellam P, Cotten M. Genome Sequences of a Novel Vietnamese Bat Bunyavirus. *Genome Announc*. 2016; 4(6):e01366–16. <https://doi.org/10.1128/genomeA.01366-16> PMID: 28007850
61. Mourya DT, Yadav PD, Basu A, Shete A, Patil DY, Zawar D, et al. Malsoor Virus, a novel bat phlebovirus, is closely related to Severe Fever with Thrombocytopenia Syndrome virus and Heartland virus. *J Virol*. 2014; 88(6):3605–9. <https://doi.org/10.1128/JVI.02617-13> PMID: 24390329
62. Lau SKP, Woo PCY, Lai KKY, Huang Y, Yip CCY, Shek C-T, et al. Complete genome analysis of three novel picornaviruses from diverse bat species. *J Virol*. 2011; 85(17):8819–28. <https://doi.org/10.1128/JVI.02364-10> PMID: 21697464
63. Kemenesi G, Zhang D, Marton S, Dallos B, Gorfol T, Estok P, et al. Genetic characterization of a novel picornavirus detected in *Miniopterus schreibersii* bats. *J Gen Virol*. 2014; 96:815–21. <https://doi.org/10.1099/jgv.0.000028> PMID: 25516541
64. Lukashev AN, Corman VM, Schacht D, Gloza-rausch F, Seebens-hoyer A, Gmyl AP, et al. Close genetic relatedness of picornaviruses from European and Asian bats. *J Gen Virol*. 2017; 98:955–61. <https://doi.org/10.1099/jgv.0.000760> PMID: 28555547
65. Pankovics P, Boros Á, Bíró H, Barbara K, Gia T, Delwart E, et al. Novel picornavirus in domestic rabbits (*Oryctolagus cuniculus* var. *domestica*). *Infect Genet Evol*. 2016; 37:117–22. <https://doi.org/10.1016/j.meegid.2015.11.012> PMID: 26588888
66. Li L, Deng X, Mee ET, Collot-teixeira S, Anderson R, Schepelmann S, et al. Comparing viral metagenomics methods using a highly multiplexed human viral pathogens reagent. *J Virol Methods*. 2015; 213:139–46. <https://doi.org/10.1016/j.jviromet.2014.12.002> PMID: 25497414
67. Grard G, Fair JN, Lee D, Slikas E, Steffen I, Muyembe JJ, et al. A novel rhabdovirus associated with scute hemorrhagic fever in Central Africa. *PLoS Pathog*. 2012; 8(9):e1002924. <https://doi.org/10.1371/journal.ppat.1002924> PMID: 23028323
68. Yuan L, Li M, Li L, Monagin C, Chmura AA, Schneider BS, et al. Evidence for retrovirus and paramyxovirus infection of multiple bat species in China. *Viruses*. 2014; 6:2138–54. <https://doi.org/10.3390/v6052138> PMID: 24841387
69. Jánoska M, Vidovszky M, Molnár V, Liptovszky M, Harrach B, Benko M. Novel adenoviruses and herpesviruses detected in bats. *Vet J*. 2011; 189(1):118–21. <https://doi.org/10.1016/j.tvjl.2010.06.020> PMID: 20813566
70. Vidovszky MZ, Kohl C, Bologh S, Gorfol T, Wibbelt G, Kurth A, et al. Random sampling of the Central European bat fauna reveals the existence of numerous hitherto unknown adenoviruses. *Acta Vet Hung*. 2015; 63(4):508–25. <https://doi.org/10.1556/004.2015.047> PMID: 26599097
71. Benkő M, Harrach B, Kremer EJ. Do nonhuman primate or bat adenoviruses pose a risk for human health? *Future Microbiol*. 2014; 9:269–72. <https://doi.org/10.2217/fmb.13.170> PMID: 24762301
72. Sonntag M, Mühldorfer K, Speck S, Wibbelt G, Kurth A. New adenovirus in bats, Germany. *Emerg Infect Dis*. 2009; 15(12):2052–5. <https://doi.org/10.3201/eid1512.090646> PMID: 19961700
73. Shabman RS, Shrivastava S, Tsibane T, Attie O, Jayaprakash A, Mire CE, et al. Isolation and characterization of a novel Gammaherpesvirus from a microbat cell line. *mSphere*. 2016; 1(1):1–15.
74. Zheng X, Qiu M, Chen S, Xiao J, Ma L, Shan L, et al. High prevalence and diversity of viruses of the subfamily Gammaherpesvirinae, family Herpesviridae, in fecal specimens from bats of different species in southern China. *Arch Virol*. 2016; 161:135–40. <https://doi.org/10.1007/s00705-015-2614-9> PMID: 26446885
75. Escalera-Zamudio M, Rojas-Anaya E, Kolokotronis S-O, Taboada B, Loza-Rubio E, Méndez-Ojeda ML, et al. Bats, primates, and the evolutionary origins and diversification of mammalian gammaherpesviruses. *MBio*. 2016; 7(6):e01425–16. <https://doi.org/10.1128/mBio.01425-16> PMID: 27834200

76. de Groot RJ, Baker SC, Baric R, Enjuanes L, Gorbalenya AE, Holmes K V., et al. Family Coronaviridae. In: King A, Adams MJ, Carstens EB, Lefkowitz E, editors. *Virus Taxonomy: Classification and nomenclature of viruses Ninth Report of the International Committee on Taxonomy of Viruses*. London, United Kingdom: Academic Press; 2012. p. 806–20.
77. van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJM, Wolthers KC, et al. Identification of a new human coronavirus. *Nat Med*. 2004; 10(4):368–73. <https://doi.org/10.1038/nm1024> PMID: 15034574
78. Tao Y, Shi M, Chommanard C, Queen K, Zhang J, Markotter W, et al. Surveillance of bat coronaviruses in Kenya identifies relatives of human coronaviruses NL63 and 229E and their recombination history. *J Virol*. 2017; 91:e01953–16. <https://doi.org/10.1128/JVI.01953-16> PMID: 28077633
79. Pfefferle S, Oppong S, Drexler JF, Gloza-Rausch F, Ipsen A, Seebens A, et al. Distant relatives of severe acute respiratory syndrome coronavirus and close relatives of human coronavirus 229E in bats, Ghana. *Emerg Infect Dis*. 2009; 15(9):1377–84. <https://doi.org/10.3201/eid1509.090224> PMID: 19788804
80. Annan A, Baldwin HJ, Corman VM, Klose SM, Owusu M, Nkrumah EE, et al. Human Betacoronavirus 2c EMC/2012-related viruses in bats, Ghana and Europe. *Emerg Infect Dis*. 2013; 19(3):3–6.
81. Sabir JSM, Lam TT-Y, Ahmed MMM, Li L, Shen Y, E. M. Abo-Aba S, et al. Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. *Science*. 2015; 351(6268):81–4. <https://doi.org/10.1126/science.aac8608> PMID: 26678874
82. Temmam S, Davoust B, Berenger J, Raoult D, Desnues C. Viral metagenomics on animals as a tool for the detection of zoonoses prior to human infection? *Int J Mol Sci*. 2014; 15:10377–97. <https://doi.org/10.3390/ijms150610377> PMID: 24918293