Faecal virome of the Australian grey-headed flying fox from urban/suburban environments contains novel coronaviruses, retroviruses and sapoviruses

Kate Van Brussel[a], Jackie E. Mahar[a], Ayda Susana Ortiz-Baez[a], Maura Carrai[b], Derek Spielman[c], Wayne S. J. Boardman[d], Michelle L. Baker[e], Julia A. Beatty[b], Jemma L. Geoghegan[f], Vanessa R. Barrs[b,g*] and Edward C. Holmes[a*]

[a]Sydney Institute for Infectious Diseases, School of Life & Environmental Sciences and School of Medical Sciences, The University of Sydney, NSW, 2006, Australia.
[b]Jockey Club College of Veterinary Medicine & Life Sciences, City University of Hong Kong, Kowloon Tong, Hong Kong, SAR China.
[c]School of Veterinary Science, Faculty of Science, University of Sydney, Sydney, NSW 2006, Australia.
[d]School of Animal and Veterinary Sciences, Faculty of Science, Engineering and Technology, University of Adelaide, Adelaide, SA 5371, Australia.
[e]CSIRO Australian Centre for Disease Preparedness, Health and Biosecurity Business Unit, Geelong, VIC 3220, Australia.
[f]Department of Microbiology and Immunology, University of Otago, Dunedin 9010, New Zealand; Institute of Environmental Science and Research, Wellington 5022, New Zealand.
[g] Centre for Animal Health and Welfare, City University of Hong Kong, Kowloon Tong, Hong Kong, China.

[*]Corresponding authors: edward.holmes@sydney.edu.au, vanessa.barrs@cityu.edu.hk

ABSTRACT

Bats are important reservoirs for viruses of public health and veterinary concern. Virus studies in Australian bats usually target the families *Paramyxoviridae*, *Coronaviridae* and *Rhabdoviridae*, with little known about their overall virome composition. We used metatranscriptomic sequencing to characterise the faecal virome of grey-headed flying foxes from three colonies in urban/suburban locations from two Australian states. We identified viruses from three mammalian-infecting (*Coronaviridae, Caliciviridae, Retroviridae*) and one possible mammalian-infecting (*Birnaviridae*) family. Of particular interest were a novel bat betacoronavirus (subgenus *Nobecovirus*) and a novel bat sapovirus (*Caliciviridae*), the first identified in Australian bats, as well as a potentially exogenous retrovirus. The novel betacoronavirus was detected in two sampling locations 1,375 km apart and falls in a viral lineage likely with a long association with bats. This study highlights the utility of unbiased sequencing of faecal samples for identifying novel viruses and revealing broad-scale patterns of virus ecology and evolution.

Keywords

Coronavirus, sapovirus, retrovirus, faecal, mammalian, grey-headed flying fox

1. Introduction

Bats (order Chiroptera) are one of the largest mammalian orders with a unique physiology adapted for flight. The number of bat colonies in urban habitats has increased in recent decades, leading to more frequent interactions with humans, companion animals and livestock that have in turn facilitated outbreaks of zoonotic disease (Plowright et al., 2011). This process has been dramatically highlighted by the emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the detection of SARS-like coronaviruses in Asian bat populations (Temmam et al., 2022, Zhou et al., 2021, Zhou et al., 2020, Wacharapluesadee et al., 2021, Murakami et al., 2020). In addition, bats have been associated with the emergence of Hendra virus (Halpin et al., 2000), Nipah virus (Yob et al., 2001), lyssaviruses (Botvinkin et al., 2003, Gould et al., 1998) and SARS-CoV (Li et al., 2005). In turn, these outbreaks have led to increased sampling of bat species, and the widespread use of metagenomic sequencing has enabled more detailed exploration of the bat virome (Wu et al., 2016, Hardmeier et al., 2021, Van Brussel and Holmes, 2022).

In Australia, bat species of the genus *Pteropus* are reservoir hosts for Hendra virus and Menangle virus, zoonotic pathogens of the family *Paramyxoviridae* (Halpin et al., 2000, Philbey et al., 1998), as well as Australian bat lyssavirus, a zoonotic virus of the *Rhabdoviridae* that causes rabies in mammals (Gould et al., 1998). Studies of viruses in bats in Australia have largely focused on these virus families and recently identified a new member of the *Paramyxoviridae* – Cedar virus – as well as a novel genotype of Hendra virus (Wang et al., 2021, Marsh et al., 2012). Although important, these studies lack information on overall virome composition, particularly those virus families not included in targeted PCR studies.

The grey-headed flying fox (*Pteropus poliocephalus*), a member of the megabat family Pteropodidae and native to Australia, is a species of importance in the context of zoonotic viruses. Grey-headed flying foxes are distributed throughout the eastern coastline of Australia (Queensland, New South Wales and Victoria) and more recently a colony was established in Adelaide (South Australia). Grey-headed flying foxes feed on fruit, pollen and nectar and roost in large colonies, sometimes sharing roosting locations with other species of *Pteropus,* allowing intraspecies and interspecies virus transmission (Timmiss et al., 2021). Roosting sites are commonly located alongside human communities including in densely populated

urban settings (Williams et al., 2006). As numerous viruses are transmitted by faeces and other excretions, the co-habitation between bats and humans likely increases the risk of zoonotic spill-over.

Herein, we used metatranscriptomic sequencing of faecal samples to describe the community of viruses present in the gastrointestinal tract of grey-headed flying foxes from three sampling locations in two Australian states – Centennial Park and Gordon in Sydney, New South Wales, and the Botanic Park, Adelaide in South Australia. Specifically, to reveal the composition and abundance of viruses in bats residing in metropolitan areas we sampled roosting sites either located in a residential setting or in parks that are frequented by humans.

2. Methods

Sample collection

Faecal samples were collected from grey-headed flying fox roosting sites in three regions of Australia: Centennial Parklands, Centennial Park New South Wales (NSW), Gordon NSW, and Botanic Park, Adelaide parklands, Adelaide, South Australia (Table 1, Fig. 1A). Sampling was conducted over two dates in 2019 for the Centennial Park and Gordon sites, while the roosting site in the Adelaide parklands was sampled over several months in 2019 (Table 1). A plastic sheet of approximately 3 x 5 metres was placed under densely populated trees the night before collection. The following morning samples captured by the plastic sheet were placed into 2 mL tubes and immediately stored at -80°C until processing. Any faecal sample touching or submerged in urine was discarded.

Table 1. Sampling overview, including number of samples allocated to sequencing pools and sequencing metadata.

Location	Sampling date	Pool no.	No. of samples	No. of reads	No. of contigs
Centennial Park, NSW 33. 89999°S, 151.23592°E	5 February 2019	01	12	24,732,494	159,527
		02	9	35,835,953	147,425
		03	9	31,960,624	107,431
	26 February 2019	04	9	19,833,973	111,196
		05	11	31,410,836	136,180
		06	9	29,318,213	105,118

		07	10	19,160,704	90,339
Gordon, NSW 33.75065°S, 151.16242°E	12 March 2019	01	12	52,605,108	89,247
		02	12	48,784,843	50,574
		03	9	27,396,450	118,509
	26 March 2019	04	11	36,591,148	181,524
		05	12	36,815,461	146,466
		06	12	52,934,611	97,013
		07	10	37,980,832	156,960
Adelaide, SA 34.91571°S, 138.6068°E	2019	01	8	25,977,712	135,969
	2019	02	9	21,113,731	113,546

RNA extraction, sequencing and read processing

Faecal samples were homogenised using the Omni Bead Ruptor 4 with 1.44 mm ceramic beads (Omni international). Total RNA was extracted from each sample individually using the RNeasy Plus Mini Kit (Qiagen) following the manufacturer's protocol. RNA was pooled in equimolar ratios and separated by sampling location, date and RNA concentration (Table 1). Ribosomal RNA was depleted, and libraries constructed using the Illumina Stranded Total RNA Prep with Ribo-Zero Plus (Illumina) preparation kit. Libraries were sequenced as 150 bp paired-end on the Illumina Novaseq 6000 platform at the Australian Genome Research Facility (AGRF). Read ends with a quality score of below 25 phred and adapter sequences were removed using cutadapt v1.8.3 (Kechin et al., 2017). Sortmerna v4.3.3 was used to remove 5S and 5.8S, eukaryotic 18S and 23S, bacterial 16S and 23S, and Archaea 16S and 23S ribosomal RNA (rRNA) reads (Kopylova et al., 2012). The filtered reads were then *de novo* assembled using Megahit v1.1.3 (Li et al., 2015) and contigs were compared to the non-redundant protein database using diamond v2.0.9. The Genemark heuristic approach (Besemer and Borodovsky, 1999, Zhu et al., 2010) and information from closely related viruses were used to predict genes and annotate genomes. Intact retrovirus genomes were detected using an in-house pipeline (Chang et al., manuscript in preparation). The Geneious assembler (Geneious Prime version 2022.1.1) was used to reassemble megahit contigs from multiple libraries for bat faecal associated retrovirus 2 (see Results). The final sequence for bat faecal associated retrovirus 2 (see Results) was determined by mapping reads from all

libraries to the reassembled genome on Geneious Prime using a 0% (majority) threshold for the final consensus sequence.

Abundance estimation

Virus and host abundance were estimated by mapping non-rRNA reads from each library to assembled contigs, and to the COX1 gene (accession no. KF726143) from the *P. alecto* (Black flying fox) genome using Bowtie2 v2.3.4.3 (Langmead and Salzberg, 2012). The impact of index-hopping was minimised by excluding the read abundance count for a contig in any library that was less than 0.1% of the highest read count for that assembled contig in any other library.

Phylogenetic analysis

Virus amino acid sequences were aligned with related sequences (i.e., representing the same virus family and/or genus) retrieved from the NCBI/GenBank database using MAFTT v7.450 (Katoh and Standley, 2013) and the E-INS-I algorithm (Katoh et al., 2005). The partial RdRp sequence of P. alecto/Aus/SEQ/2009 was retrived from Smith et al. (2016). The gappyout method in TrimAL v1.4.1 was used to remove ambiguous regions in the alignment (Capella-Gutiérrez et al., 2009). Maximum likelihood trees of each data set were inferred using IQ-TREE v1.6.7 (Nguyen et al., 2014), employing the best-fit amino acid substitution model determined by the ModelFinder program (Kalyaanamoorthy et al., 2017) in IQ-TREE. Nodal support was accessed using 1000 ultrafast bootstrap replicates (Hoang et al., 2017). Any virus sequence with over 90% nucleotide similarity to another detected here was excluded from the phylogenetic analysis.

3. Results

Virome overview

In total, 164 faecal samples allocated to 16 libraries underwent metatranscriptomic sequencing. This generated 19,160,704 to 52,934,611 reads per library (average of 33,278,293 reads) after read filtering (Table 1). Reads were *de novo* assembled into 50,574 to 181,524 contigs (average of 121,689 contigs) per library (Table 1). A total of 5,933 contigs were assigned as of viral origin across all the libraries. The samples collected at Centennial Park, Sydney produced the most viral contigs, with 3,216 identified from 65 virus families (Supplementary Fig. 1). The Gordon, NSW sample site produced 2,399 virus contigs from 66

156 virus families, while the Adelaide site contained 318 virus contigs from 33 virus families,
157 although this site had only two sequencing libraries comprising 17 faecal samples, compared
158 to seven sequencing libraries in each of the other two locations (69 faecal samples from
159 Centennial Park, 78 from Gordon) (Table 1, Supplementary Fig. 1). Screening of the NCBI
160 protein database revealed assembled virus contigs were mostly associated with infection of
161 invertebrates (29.9% of total contigs), fungi (25.6%), plants (24.4%), and oomycetes (13%),
162 representing 79 virus families (Fig. 1B, Supplementary Fig. 1). These viruses were most
163 likely associated with host diet and differed in frequency depending on sampling site (Fig.
164 1B, Supplementary Fig. 1). The plant, fungal, and oomycete-associated viruses, as well as
165 those likely to be bacteriophage (including the picobirnaviruses) were not considered further.
166 Importantly, however, we also identified sequences from viruses likely associated with
167 mammalian infection (3% overall), including near complete genomes from members of the
168 *Coronaviridae, Caliciviridae* and *Retroviridae* (Fig. 1B).

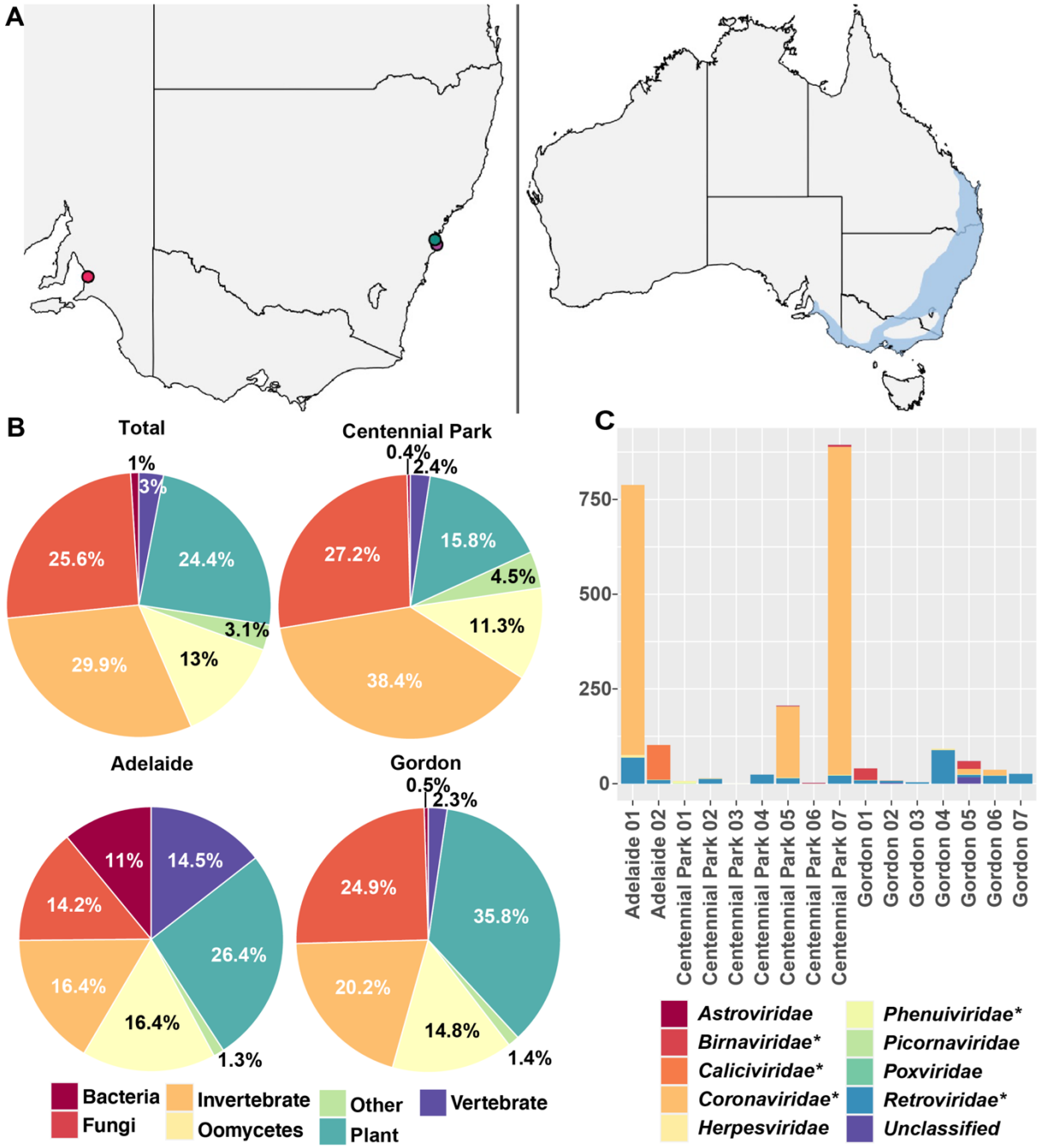


Fig. 1. Overview of sampling sites and bat faecal sample composition. (A) Sampling locations in Australia (left) and distribution map of the grey-headed flying fox (right) (IUCN, 2021). (B) Likely hosts of viral contigs based on host designation of the closest relatives in the NCBI non-redundant protein database. (C) Read abundance presented as reads per million (RPM) for the vertebrate-associated virus sequences for each library and separated by virus family. The virus families discussed in this study are highlighted with an asterisk.

Mammalian-associated viruses

We detected contigs from nine viral families likely to infect mammals (Fig. 1C). The *Coronaviridae* and *Retroviridae* were particularly abundant and present in five and 13 libraries, respectively (Fig. 1C). Members of the *Birnaviridae* and *Caliciviridae* were also abundant in specific libraries (Fig. 1C). The remaining mammalian-associated viral families were only detected at low abundance and the contigs were not of sufficient length for further characterisation.

Novel betacoronavirus (*Coronaviridae*)

A novel complete betacoronavirus genome (single-strand, positive-sense RNA virus; +ssRNA) – provisionally denoted bat faecal coronavirus CP07/aus/1 – was identified in a sequencing library sampled from Centennial Park (pool no. 07) and in a sequencing library from Adelaide (pool no. 01). These two sequences exhibited 99.8% identity over the complete viral genome indicating that they represent the same species. Additionally, three sequences with 99.2-100% sequence identity to CP07/aus/1 were identified in an additional Centennial Park library (pool no. 05).

CP07/aus/1 contains ten ORFs in the arrangement ORF1a, ORF1ab, spike, NS3, envelope, matrix, nucleocapsid, NS7a, NS7b and NS7c. Transcription Regulatory Sequences (TRS) preceeded all ORFs. Additional bat coronavirus contigs ranging from 318 to 1,309 bp were detected in sequencing libraries from two Gordon sampling locations. These short contigs shared 40-95% amino acid identity to CP07/aus/1. Three of these contigs contained RdRp or spike amino acid sequences of sufficient length for phylogenetic analysis, and these were provisionally denoted bat faecal coronavirus G05/aus/1, G05/aus/2 and G05/aus/3. Based on phylogenetic analysis of the RNA-dependent RNA polymerase (RdRp) and/or spike protein, the novel betacoronaviruses detected here fell within the *Betacoronavirus* subgenus *Nobecovirus* (Fig. 2) and were most closely related to P.alecto/Aus/SEQ/2009 (for which only a partial RdRp is available) sampled from a black flying fox in south east Queensland, Australia (Smith et al., 2016) and to Pteropus rufus nobecovirus sampled from a flying fox in Madagascar (accession no. OK067319; Fig. 2) (Kettenburg et al., 2022). CP07/aus/1 had 83% amino acid identity to Pteropus rufus nobecovirus over the complete ORF1ab replicase and 97% to P.alecto/Aus/SEQ/2009 over the partial RdRp. Amino acid identity to Pteropus rufus nobecovirus over the spike and non-structural proteins was 72% and 58%, respectively. The RdRp of G05/aus/1 shared 95% amino acid identity to CP07/aus/1, while the partial spike proteins of G05/aus/2 and G05/aus/3 shared 57% and 63% amino acid identity to

CP07/aus/1, respectively. It is possible that G05/aus/1 and G05/aus/2 represent transcripts from the same virus, while G05/aus/3 represents a different species to CP07/aus/1. However, this could not be confirmed as the G05/aus/3 genome was incomplete. Regardless, it is clear from the spike protein phylogeny that at least three different coronaviruses are circulating in the bats sampled here.

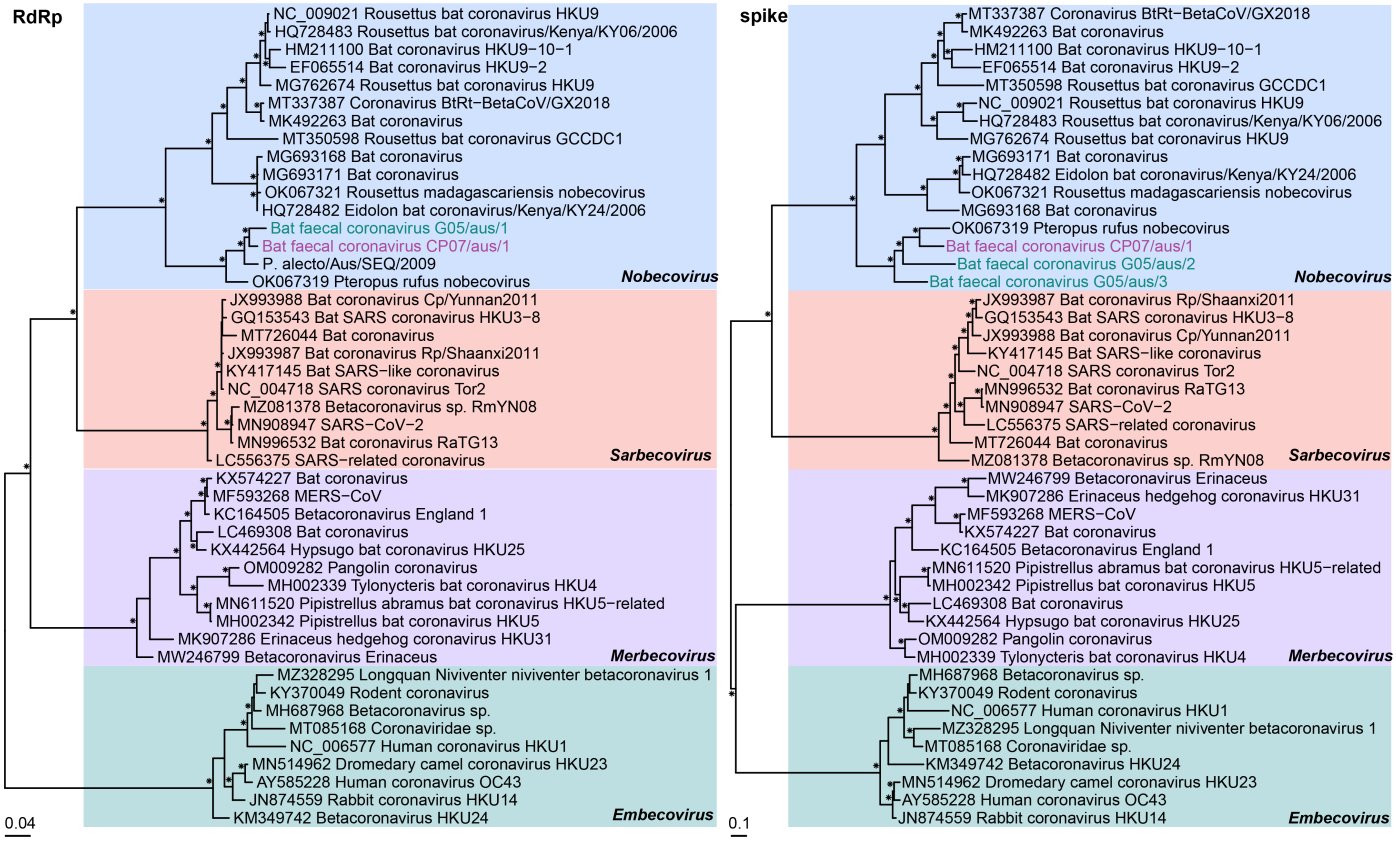


Fig. 2. Phylogenetic relationships of the novel bat betacoronaviruses based on the amino acid sequences of the RdRp and spike protein. Amino acid alignment lengths were 832 and 1,092 residues for the RdRp and spike protein, respectively. Representative betacoronavirus sequences from this study are coloured by sampling location (Centennial Park, Sydney – purple, and Gordon – green) and the subgenera are highlighted. Bootstrap values >70% are represented by the symbol shown at the branch node. The tree is rooted at midpoint for clarity and the scale bar represents the amino acid substitutions per site.

Novel sapovirus (*Caliciviridae*)

A near complete genome of a novel sapovirus (*Caliciviridae*, +ssRNA virus), tentatively named bat faecal sapovirus Ad02/aus/1, was detected in a sequencing library sampled from Adelaide (pool no. 2). Nine additional bat sapovirus sequences ranging from 340 to 783 bp

were detected in the same sequencing library. The nine sequences shared 66-74% nucleotide and 76-81% amino acid identity to Ad02/aus/1 over the polyprotein, suggesting the presence of additional diverse sapoviruses. The near complete Ad02/aus/1 genome is 7,254 bp and contains two ORFs encoding a polyprotein (near complete with likely 45 residues missing from the 5' end), and the VP2. Ad02/aus/1 exhibited 44.8% amino acid identity in the partial polyprotein to its closest relative – Bat sapovirus Bat-SaV/Limbe65/CAM/2014 (accession no. KX759620) – detected in the faeces of *Eidolon helvum* bats in Cameroon, Africa (Yinda et al., 2017). Phylogenetic analysis of the RdRp and VP1 revealed a clustering of bat sapoviruses in both trees that included the novel Australian bat sapoviruses found here (Fig. 3). Bat sapoviruses have been assigned to the putative genogroups GXIV, GXVI, GXVII, GXVIII and GXIX based on VP1 phylogeny and amino acid sequence identities. Using the same criteria, the novel sapovirus Ad02/aus/1 identified here should be assigned to its own genogroup, putatively named GXX, which would also include the partial VP1 Ad02/aus/4 sequence (Supplementary Table 1, Fig. 3).

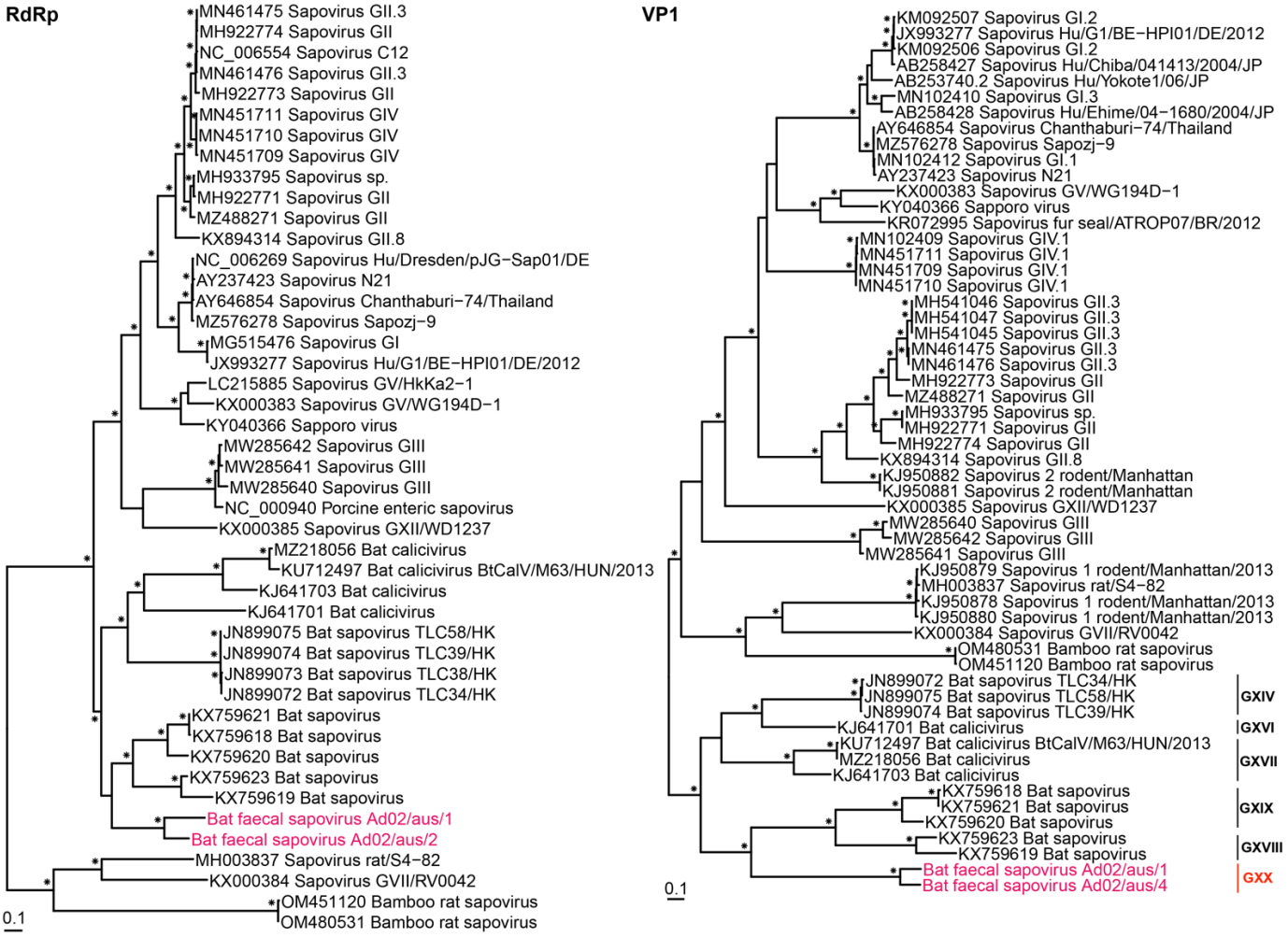


Fig. 3. Phylogenetic relationships of the novel bat sapoviruses using the amino acid sequences of the RdRp and VP1. Amino acid alignment lengths were 491 and 623 residues for the RdRp and VP1, respectively. Bat sapoviruses from this study are coloured by sampling location (Adelaide – pink) and bootstrap values >70% are represented by the symbol shown at the branch node. The putative bat sapovirus genogroups are displayed to the right of the VP1 tree and our proposed putative genogroup is coloured in red. The trees are rooted at midpoint for clarity and the scale bar represents the amino acid substitutions per site.

Novel birna-like virus (*Birnaviridae*)

Sequences related to the *Birnaviridae* (double-stranded RNA viruses; dsRNA) were detected in one Centennial Park and two Gordon libraries. All the birna-like virus sequences identified in the Centennial Park and Gordon libraries shared >99% nucleotide identity, and the complete coding region of segment B, which encodes the RdRp, was obtained from one library (Gordon 05). The *Birnaviridae* segment A that encodes the polyprotein and a small overlapping ORF was not identified in our data. Phylogenetic analysis revealed that the birna-like virus RdRp sequence, denoted G05/aus/1, was most closely related (50% amino acid identity) to the disease-causing virus Chicken proventricular necrosis virus (Fig. 4) (Guy et al., 2011), forming a distinct clade that is distantly related to the birnaviruses that infect a wide range of hosts.

Bat retrovirus (*Retroviridae)*

A near complete genome of a retrovirus was identified in Gordon library 04 and provisionally named bat faecal associated retrovirus 1 G04/aus/1. Four ORFs were observed over the 7,455 bp genome and assigned as the gag, pro, pol and env genes based on the presence of conserved domains. In the pro gene we were able to identify an active site motif DTGAD predominately observed in functional retroviruses, and a helix motif GRDVL (Turnbull and Douville, 2018). We were unable to identify complete long terminal repeat (LTR) regions in the 7,455 bp genome, although this may be due to incomplete assembly at the 5' and/or 3' end, rather than a true absence of LTRs. Importantly, as the four ORFs contained the appropriate retrovirus conserved domains and were uninterrupted by stop codons, it is possible that G04/aus/1 is potentially exogenous and functional. A BLASTn analysis of the complete G04/aus/1 genome revealed no match to any bat reference genome on NCBI/GenBank. G04/aus/1 exhibited 56% amino acid identity in the pol protein to its closest

281 relative, Simian retrovirus 2 (accession M16605), a presumably exogenous retrovirus (Thayer
282 et al., 1987). The abundance for this novel retrovirus in the Gordon 04 library was 67 RPM
283 (2,457 reads) (Fig. 1C).
284
285 A further near complete retroviral genome was identified by reassembling 31 partial contig
286 sequences from 12 libraries from all three sample locations. This Bat faecal associated
287 retrovirus 2 AdCPG/aus/1 is 6,630 bp and contains four open reading frames encoding the
288 gag, pro, pol and env genes. It also contains the conserved domains expected in functional
289 retroviruses, although the terminal end of the env gene is missing (either from true truncation
290 or incomplete assembly). The virus is most closely related to AdCPG/aus/1 sampled from the
291 lung tissue of Malayan pangolins (Ning et al., 2022). BLASTn analysis of the complete
292 genome of AdCPG/aus/1 showed the absence of this genome in any bat reference genome on
293 NCBI/GenBank. AdCPG/aus/1 reads were detected in 13 libraries (two Adelaide, four
294 Centennial Park and seven Gordon) and the abundance in each library ranged from 3.7 – 68.8
295 RPM (127 – 1786 reads) (Fig. 1C). Phylogenetic analysis of the pol protein that contains the
296 reverse transcriptase (RT) domain revealed that G04/aus/1 and AdCPG/aus/1 fell within the
297 genus *Betaretrovirus,* clustering with both exogenous and endogenous retroviruses associated
298 with various mammalian species (Fig. 4).

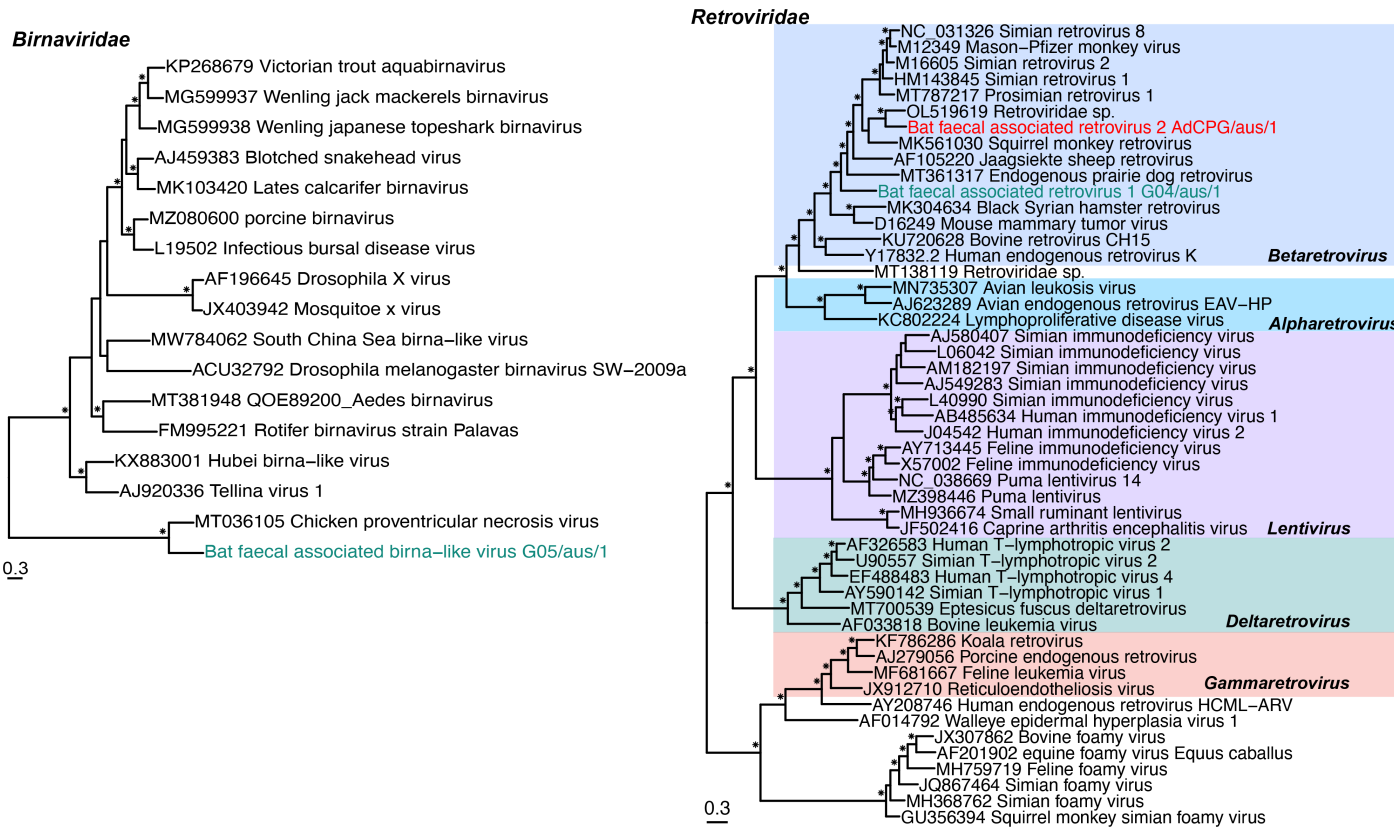


Fig. 4. Phylogenetic analysis of the birna-like virus and bat retroviruses based on the RdRp and pol amino acid sequences, respectively. The *Birnaviridae* RdRp sequence alignment was 767 amino acid resides in length while the *Retroviridae* pol alignment comprised 1,356 residues. The viruses from this study are coloured by sampling location (Gordon – green) and the reassembled retrovirus sequence is in red (to indicate multiple locations). The *Retroviridae* genera are highlighted and bootstrap values >70% are represented by the symbol shown at the branch node. The tree is midpoint rooted for clarity, with the scale bar representing the amino acid substitutions per site.

Invertebrate-associated viruses

We detected likely invertebrate-associated virus sequences from seven single-strand negative-sense RNA viruses (-ssRNA), three +ssRNA virus and one dsRNA virus families, in addition to the order *Bunyavirales* (-ssRNA). The virus sequences from the *Chuviridae*, *Lispiviridae*, *Artoviridae*, *Nyamiviridae*, *Xinmoviridae*, *Qinviridae*, *Disctroviridae* and *Iflaviridae* are not discussed further, although information on positive libraries is provided (Supplementary Fig. 1) and phylogenetic analysis was performed (Supplementary Fig. 2). Virus sequences from the *Orthomyxoviridae*, *Nodaviridae*, *Reoviridae* and *Bunyavirales* are considered further as these viral groups include mammalian-infecting viruses, are important

318 vector-borne viruses, or are able to infect mammals experimentally (*Nodaviridae*, genus
319 *Alphanodavirus*).
320
321 Orthomyxovirus (-ssRNA virus) segments were identified in five libraries from Centennial
322 Park. Full coding regions for two polymerase segments – PB2 and PA – and the
323 hemagglutinin segment 2 and nucleocapsid segment 5 were present in all libraries, although a
324 full coding region for polymerase segment PB1 was only present in a single Centennial Park
325 library. The three polymerase proteins of Centennial Park library 06 were used for
326 phylogenetic analysis, which revealed that this sequence was most closely related to an
327 orthomyxovirus sampled from jumping plant lice in Australia (Fig. 5) (Käfer et al., 2019).
328 Nodaviruses (+ssRNA virus) were detected in five Centennial Park libraries and three
329 Gordon libraries. Both the RNA1 (RdRp) and RNA2 segments were identified, including two
330 sequences with the complete RdRp. Nodavirus CP01/aus/1 and CP02/aus/1 were related to a
331 nodavirus sampled from birds in China (Zhu et al., 2022) and most likely belong to the same
332 viral species, although these fragments were only 476 and 232 amino acids, respectively. The
333 nodavirus CP07/aus/1 RdRp segment was related to a nodavirus from arthropod hosts from
334 China (Fig. 5) (Shi et al., 2016). Gene segments related to the *Reoviridae* (dsRNA) were
335 present in all Centennial Park, three Gordon and one Adelaide library. The reovirus VP1 Pol
336 segments detected here were related, albeit distantly (~40% amino acid identity) to reoviruses
337 associated with ticks (Harvey et al., 2019, Vanmechelen et al., 2021), moths (Graham et al.,
338 2006), bat flies (Xu et al., 2022) and the Asian citrus psyllid (Nouri et al., 2015) (Fig. 5).

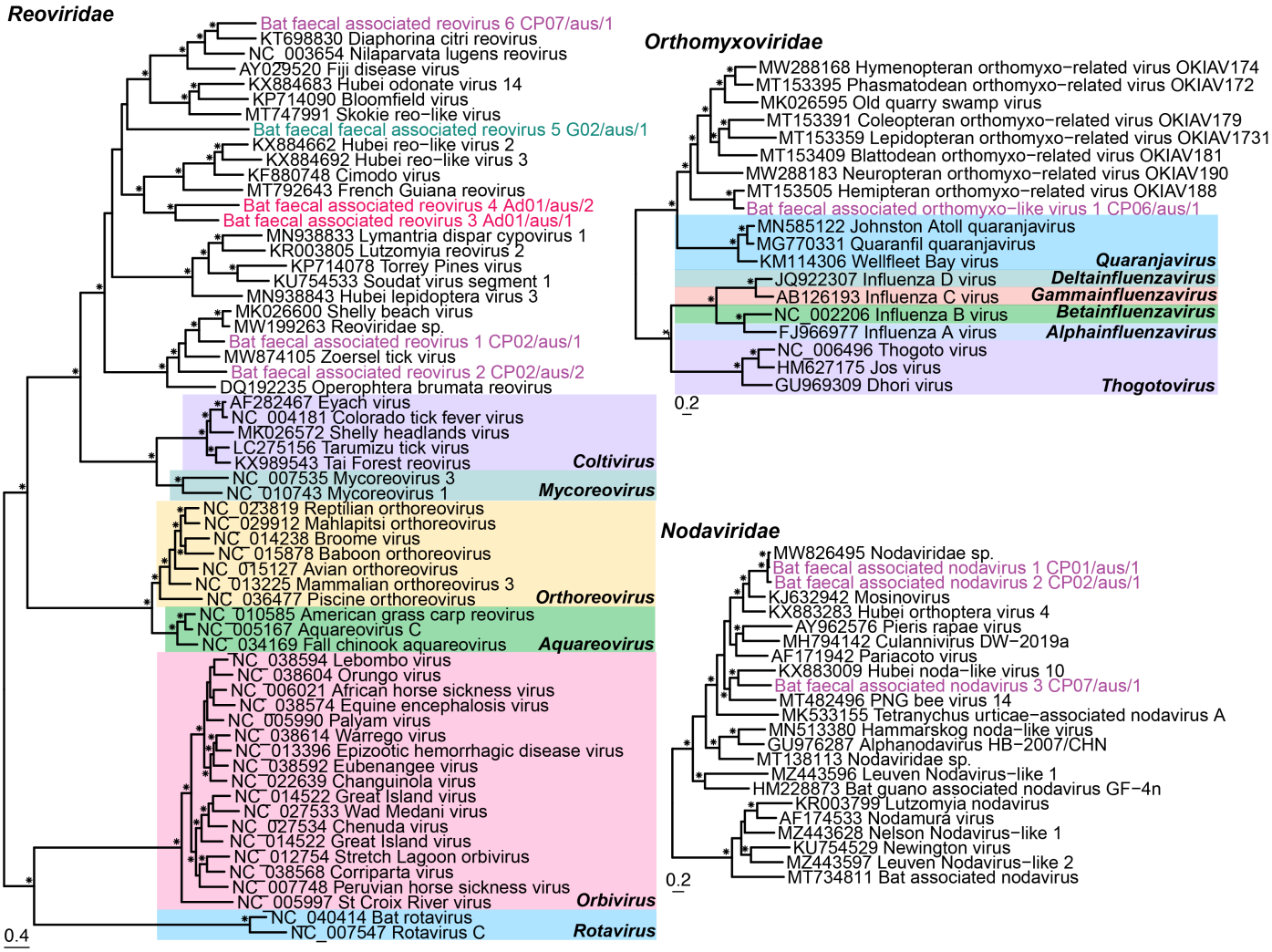
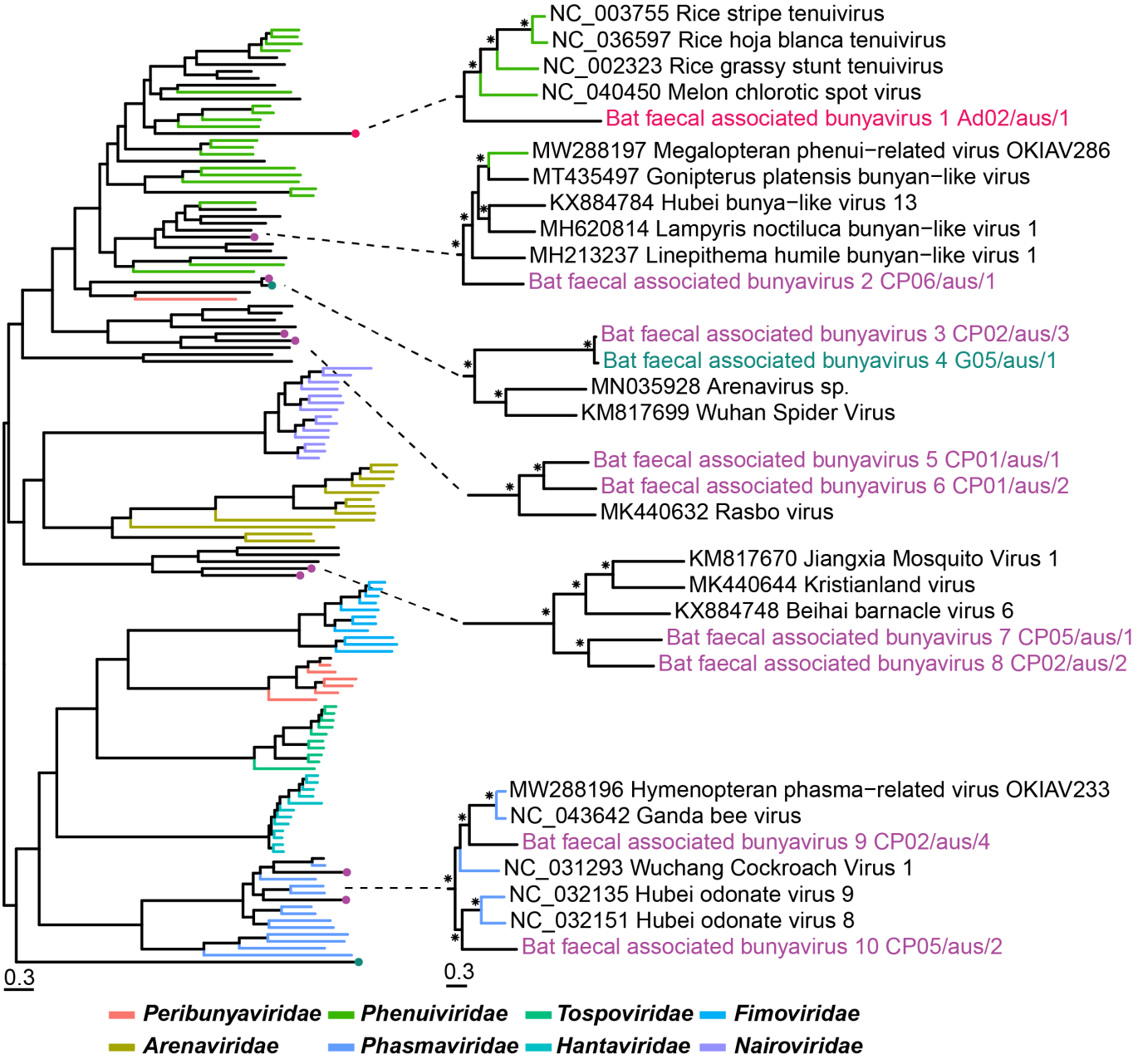


Fig. 5. Phylogenetic analysis of the invertebrate-associated reoviruses, orthomyxoviruses and nodaviruses based on the VP1 Pol, concatenated PB2-PB1-PA and RdRp amino acid sequences, respectively. Amino acid alignment length were 1,020 residues for *Reoviridae*, 2,233 residues for the *Orthomyxoviridae* and 774 residues for the *Nodaviridae*. Viruses from this study are coloured by sampling location (Adelaide – pink, Centennial Park – purple and Gordon – green) and genera are highlighted in the *Reoviridae* and *Orthomyxoviridae* tress. Bootstrap values >70% are represented by the symbol shown at the branch node. The tree is rooted at midpoint for clarity and the scale bar represents the amino acid substitutions per site.

Finally, bunyavirus fragments were detected in all the Adelaide and Centennial Park libraries and six Gordon libraries. Eleven RdRp coding regions were used for phylogenetic analysis which revealed that two bunyavirus sequences fell into the *Phenuiviridae* and four were basal to that family, while two sequences fell into the *Phasmaviridae*, two were basal to the

354 *Arenaviridae*, and one was basal to a grouping of five families (Fig. 6). The Adelaide
355 bunyavirus Ad02/aus/1 was related to the plant associated genus *Tenuivirus* and the
356 remaining 10 were related to invertebrate hosts (Fig. 6).



357

358 **Fig. 6.** Phylogenetic analysis of viruses from the order *Bunyavirales*. The RdRp amino acid
359 sequence was used to estimate phylogenetic trees and the alignment length was 1,434 amino
360 acid residues. Viruses from this study are coloured by sampling location and bootstrap values
361 >70% are represented by the symbol shown at the branch node. The tree is midpoint rooted
362 for clarity and the scale bar represents the amino acid substitutions per site.

363

364 **4. Discussion**

Virological surveillance of bats in Australia has largely focused on screening for known zoonotic viruses such as Hendra virus and Australian bat lyssavirus, although the paramyxovirus Tioman virus, for which flying foxes are the natural host, and coronaviruses are also targeted (Boardman et al., 2020, Prada et al., 2019a, Smith et al., 2016). The primary aim of these studies is to identify specific viruses using either PCR or serological data. Although such surveillance has been successful in determining the active circulation of these specific viruses, these approaches necessarily have restricted capacity to detect novel or unexpected viruses, thus providing a limited understanding of viruses circulating in Australian bats. As bats are frequently found near human populations, they are of particular concern regarding potential zoonoses (Plowright et al., 2011, Williams et al., 2006, Halpin et al., 2000). Herein, we used metatranscriptomics to reveal the natural faecal virome of the grey-headed flying fox. Although most of the viruses identified were likely associated with bat diet, as expected from faecal sampling, we also identified viruses from three mammalian-associated families (*Coronaviridae, Caliciviridae, Retroviridae*) and one virus from the *Birnaviridae* family that may also have a mammalian association.

Both alpha- and betacoronaviruses have been identified in a variety of bat species (Smith et al., 2016, Prada et al., 2019b). Here, we characterised the complete genome of a betacoronavirus in grey-headed flying foxes that was closely related to two other betacoronaviruses sampled in flying foxes in Australia and Madagascar (Smith et al., 2016, Kettenburg et al., 2022). The current ICTV classification for coronavirus species states that less than 90% amino acid identity in the ORF1ab conserved replicase domains constitutes a new species. Although bat faecal coronavirus CP07/aus/1 shares high sequence similarity to another reported bat betacoronavirus, the P.alecto/Aus/SEQ/2009 sequence is only 146 amino acids in length, does not span the complete RdRp and is therefore difficult to classify. Accordingly, we suggest that betacoronavirus bat faecal coronavirus CP07/aus/1 represents a novel species, to which P.alecto/Aus/SEQ/2009 may also belong. The complete genome of this virus was found in both Adelaide and New South Wales (99.8% nucleotide similarity between the two genomes) and abundance counts were high in both locations (Fig. 1C), indicative of virus exchange between bat populations. Flying foxes are known to travel long distances to feed, roosting sites change depending on season, and in Australia several flying fox species share roosting sites (Timmiss et al., 2021), all of which provide opportunities for viruses to infect new individuals. Importantly, while we were only able to assemble the complete genome of one novel coronavirus, we identified partial genome fragments of at

399 least two more diverse coronaviruses (Fig. 2), indicating that Australian bats carry a high
400 diversity of coronaviruses as has been seen in other bat species.
401
402 This is the first report of a sapovirus in Australian bats. Previously, bat sapoviruses have been
403 sampled from *Eidolon helvum* (Straw-coloured fruit bat) in Cameroon (Yinda et al., 2017)
404 and Saudi Arabia (Mishra et al., 2019) and *Hipposideros Pomona* (Pomona leaf-nosed bat)
405 from Hong Kong (Tse et al., 2012). Currently, the bat sapoviruses characterised have been
406 from bats with no apparent disease (Tse et al., 2012, Yinda et al., 2017, Mishra et al., 2019).
407 Whether this is the case here is unknown because the reliance on faecal sampling meant that
408 there was no direct interaction with individual animals. The disease potential of bat
409 sapoviruses should be investigated further as sapoviruses have been linked to acute
410 gastroenteritis outbreaks in humans (Oka et al., 2015) and some animal sapoviruses are
411 closely related to those found in humans (Mombo et al., 2014, Firth et al., 2014, Martella et
412 al., 2008).
413
414 Until the metagenomic detection of porcine birnavirus (Yang et al., 2021) and porcupine
415 birnavirus (He et al., 2022) it was believed the *Birnaviridae* infected fish, insects and birds
416 exclusively (Crane et al., 2000, Da Costa et al., 2003, Chung et al., 1996, Brown and Skinner,
417 1996, Guy et al., 2011). We identified the segment B sequence of a novel bat faecal
418 associated birna-like virus that was most closely related to a divergent pathogenic avian
419 birnavirus (50% amino acid identity). Given its divergent phylogenetic position it is currently
420 unclear whether this virus actively infects grey-headed flying foxes or is associated with a
421 component of their diet or microbiome. While grey-headed flying foxes are not insectivores,
422 the ingestion of insects through the consumption of fruit and nectar seems likely given the
423 high number of invertebrate, plant and fungi viruses sequenced here (Fig. 1B, Supplementary
424 Fig. 1). The moderate abundance values (30.5 and 19.9 RPM) cannot exclude either scenario
425 as using a host reference gene such as COX1 for sequencing depth comparison may not be as
426 reliable for faecal samples as it would be when analysing tissue. Further investigation is
427 needed to determine the natural host of bat faecal associated birna-like virus and to determine
428 what tissue types are affected.
429
430 Two intact, possibly exogenous retrovirus near complete genomes were also identified in this
431 study and were most closely related to mammalian infecting retroviruses from the genus
432 *Betaretrovirus*. Six retroviruses have been previously characterised from Australian bat brain

tissue and excretions (including faeces), all from the genus *Gammaretrovirus* (Hayward et al., 2020, Cui et al., 2012) and hence highly divergent from the viruses identified here. Although the exogenous status needs to be confirmed, it is possible that bat faecal associated retrovirus 1 G04/aus/1 and bat faecal associated retrovirus 2 AdCPG/aus/1 constitute the first exogenous and intact betaretroviruses sampled from the faeces of bats in Australia. Unfortunately, virus identification through metatranscriptomics does not provide reliable information on whether a virus is endogenous and defective, or still functional and exogenous (Hayward et al., 2013, Hayward and Tachedjian, 2021). That the retroviruses detected here have all the necessary genes to comprise a functional virus, with undisrupted ORFs, were not detected in every library, and are not present in the bat genome, at the very least suggests that they are only recently endogenized and currently unfixed in the bat population. Further work confirming the nature of the retroviruses detected here is warranted since bats are known to be major hosts for retroviruses (Cui et al., 2015) and their cross-species transmission across mammalian orders is commonplace (Hayward et al., 2013).

In addition to mammalian viruses, we detected virus sequences that are likely invertebrate-associated. Of particular interest were those from the *Orthomyxoviridae* and *Reoviridae* that span a wide variety of hosts including mammals and were at high abundance in some of the Centennial Park libraries. Notably, bat faecal associated reovirus 1 CP02/aus/1 groups with members of the *Reoviridae* associated with ticks that are vectors for numerous pathogenic microorganisms, particularly from the genus *Coltivirus* – Colorado tick fever virus and Eyach virus (Goodpasture et al., 1978, Rehse-Küpper et al., 1976). Additionally, a novel coltivirus – Tai Forest reovirus – was sampled from bats in Cote d'Ivoire and shown to infect human cells (Weiss et al., 2017). The current evidence for tick-borne reovirus infection in humans highlights the importance of assessing the pathogenic potential of new tick associated reoviruses, especially those viruses discovered in urban wildlife.

Our study highlights the diversity of viruses in wildlife species from metropolitan areas. In this this context it is notable that the bat coronaviruses identified fall within the subgenus *Nobecovirus* of betacoronaviruses. Currently, this subgenus is strongly associated with bats sampled on multiple continents, with the phylogenetic depth of the *Nobecovirus* lineage further suggesting that bats have harbored these viruses for millennia with no apparent infection of humans. Hence, although the bats studied were resident in urban/suburban locations, this does not necessarily translate into a clear risk of human emergence.

Data statement

The raw data generated for this study are available in the NCBI SRA database under the BioProject accession number PRJNA851532 and SRA accession numbers SRR19790899-SRR19790914. All genome sequences presented in phylogenetic trees are available in NCBI GenBank under the accession numbers ON872523-ON872588.

Declaration of interests

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

Acknowledgements

Ethics statement

Ethics approval was granted by the University of Sydney Animal Ethics Committee (AEC 2018/1460).

Funding source

The work was funded by an Australian Research Council Australian Laureate Fellowship to ECH (FL170100022) and the Sydney Institute for Infectious Diseases.

CRediT authorship contribution statement

Kate Van Brussel: Investigation, Data curation, Formal analysis, Writing – original draft, Writing – preparation, review and editing. **Jackie E. Mahar**: Formal analysis, Writing – preparation, review and editing. **Ayda Susana Ortiz-Baez**: Formal analysis, Writing – preparation, review and editing. **Maura Carrai**: Investigation, Writing – review. **Derek Spielman**: Investigation. **Wayne S. J Boardman**: Investigation, Writing – review. **Michelle L. Baker**: Resources, Writing – review. **Julia A. Beatty**: Investigation, Supervision. **Jemma L. Geoghegan**: Resources, Writing – review. **Vanessa R. Barrs**: Conceptualisation, Methodology, Funding acquisition, Supervision, Writing – review and editing. **Edward C. Holmes**: Conceptualisation, Funding acquisition, Supervision, Writing – preparation, review and editing.

References

Besemer, J., Borodovsky, M. 1999. Heuristic approach to deriving models for gene finding. Nucleic Acids Res. 27, 3911-3920.

Boardman, W.S.J., Baker, M.L., Boyd, V., Crameri, G., Peck, G.R., Reardon, T., Smith, I.G., Caraguel, C.G.B., Prowse, T.A.A. 2020. Seroprevalence of three paramyxoviruses; Hendra virus, Tioman virus, Cedar virus and a rhabdovirus, Australian bat lyssavirus, in a range expanding fruit bat, the Grey-headed flying fox (*Pteropus poliocephalus*). PLoS One 15, e0232339.

Botvinkin, A.D., Poleschuk, E.M., Kuzmin, I.V., Borisova, T.I., Gazaryan, S.V., Yager, P., Rupprecht, C.E. 2003. Novel lyssaviruses isolated from bats in Russia. Emerg. Infect. Dis. 9, 1623-1625.

Brown, M.D., Skinner, M.A. 1996. Coding sequences of both genome segments of a European 'very virulent' infectious bursal disease virus. Virus Res. 40, 1-15.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972-1973.

Chung, H.K., Kordyban, S., Cameron, L., Dobos, P. 1996. Sequence analysis of the bicistronic Drosophila X virus genome segment A and its encoded polypeptides. Virology 225, 359-368.

Crane, M.S., Hardy-Smith, P., Williams, L.M., Hyatt, A.D., Eaton, L.M., Gould, A., Handlinger, J., Kattenbelt, J., Gudkovs, N. 2000. First isolation of an aquatic birnavirus from farmed and wild fish species in Australia. Dis. Aquat. Organ. 43, 1-14.

Cui, J., Tachedjian, G., Wang, L.F. 2015. Bats and rodents shape mammalian retroviral phylogeny. Sci. Rep. 5, 16561.

Cui, J., Tachedjian, M., Wang, L., Tachedjian, G., Wang, L.F., Zhang, S. 2012. Discovery of retroviral homologs in bats: implications for the origin of mammalian gammaretroviruses. J. Virol. 86, 4288-4293.

Da Costa, B., Soignier, S., Chevalier, C., Henry, C., Thory, C., Huet, J.C., Delmas, B. 2003. Blotched snakehead virus is a new aquatic birnavirus that is slightly more related to avibirnavirus than to aquabirnavirus. J. Virol. 77, 719-725.

Firth, C., Bhat, M., Firth, M.A., Williams, S.H., Frye, M.J., Simmonds, P., Conte, J.M., Ng, J., Garcia, J., Bhuva, N.P., Lee, B., Che, X., Quan, P.L., Lipkin, W.I. 2014. Detection of zoonotic pathogens and characterization of novel viruses carried by commensal *Rattus norvegicus* in New York City. mBio 5, e01933-14.

533 Goodpasture, H.C., Poland, J.D., Francy, D.B., Bowen, G.S., Horn, K.A. 1978. Colorado tick
534 fever: clinical, epidemiologic, and laboratory aspects of 228 cases in Colorado in 1973-
535 1974. Ann. Intern. Med. 88, 303-310.
536 Gould, A.R., Hyatt, A.D., Lunt, R., Kattenbelt, J.A., Hengstberger, S., Blacksell, S.D. 1998.
537 Characterisation of a novel lyssavirus isolated from Pteropid bats in Australia. Virus Res.
538 54, 165-187.
539 Graham, R.I., Rao, S., Possee, R.D., Sait, S.M., Mertens, P.P., Hails, R.S. 2006. Detection
540 and characterisation of three novel species of reovirus (*Reoviridae)*, isolated from
541 geographically separate populations of the winter moth *Operophtera brumata*
542 (Lepidoptera: Geometridae) on Orkney. J. Invertebr. Pathol. 91, 79-87.
543 Guy, J.S., West, A.M., Fuller, F.J. 2011. Physical and genomic characteristics identify
544 chicken proventricular necrosis virus (R11/3 virus) as a novel birnavirus. Avian Dis. 55,
545 2-7.
546 Halpin, K., Young, P.L., Field, H.E., Mackenzie, J.S. 2000. Isolation of Hendra virus from
547 pteropid bats: a natural reservoir of Hendra virus. J. Gen .Virol. 81, 1927-1932.
548 Hardmeier, I., Aeberhard, N., Qi, W., Schoenbaechler, K., Kraettli, H., Hatt, J.M., Fraefel, C.,
549 Kubacki, J. 2021. Metagenomic analysis of fecal and tissue samples from 18 endemic bat
550 species in Switzerland revealed a diverse virus composition including potentially
551 zoonotic viruses. PLoS One16, e0252534.
552 Harvey, E., Rose, K., Eden, J.-S., Lo, N., Abeyasuriya, T., Shi, M., Doggett, S.L., Holmes,
553 E.C. 2019. Extensive diversity of RNA viruses in Australian ticks. J. Virol. 93, e01358-
554 18.
555 Hayward, A., Grabherr, M., Jern, P. 2013. Broad-scale phylogenomics provides insights into
556 retrovirus-host evolution. Proc Natl Acad Sci USA110, 20146-20151.
557 Hayward, J.A., Tachedjian, G. 2021. Retroviruses of bats: a threat waiting in the wings?
558 mBio 12, e0194121.
559 Hayward, J.A., Tachedjian, M., Kohl, C., Johnson, A., Dearnley, M., Jesaveluk, B., Langer,
560 C., Solymosi, P.D., Hille, G., Nitsche, A., Sánchez, C.A., Werner, A., Kontos, D.,
561 Crameri, G., Marsh, G.A., Baker, M.L., Poumbourios, P., Drummer, H.E., Holmes, E.C.,
562 Wang, L.F., Smith, I., Tachedjian, G. 2020. Infectious KoRV-related retroviruses
563 circulating in Australian bats. Proc Natl Acad Sci USA 117, 9529-9536.
564 He, W.T., Hou, X., Zhao, J., Sun, J., He, H., Si, W., Wang, J., Jiang, Z., Yan, Z., Xing, G.,
565 Lu, M., Suchard, M.A., Ji, X., Gong, W., He, B., Li, J., Lemey, P., Guo, D., Tu, C.,

566 Holmes, E.C., Shi, M., Su, S. 2022. Virome characterization of game animals in China
567 reveals a spectrum of emerging pathogens. Cell 185, 1117-1129.e8.
568 Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S. 2017. UFBoot2:
569 improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518-522.
570 International Union for Conservation of Nature. 2021. Data from "*Pteropus poliocephalus*".
571 The IUCN Red List of Threatened Species, 2021-3.
572 https://www.iucnredlist.org/species/18751/22085511
573 Käfer, S., Paraskevopoulou, S., Zirkel, F., Wieseke, N., Donath, A., Petersen, M., Jones,
574 T.C., Liu, S., Zhou, X., Middendorf, M., Junglen, S., Misof, B., Drosten, C. 2019.
575 Reassessing the diversity of negative strand RNA viruses in insects. PLoS Pathog. 15,
576 e1008224.
577 Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S. 2017.
578 ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Meth. 14,
579 587-589.
580 Katoh, K., Kuma, K., Toh, H., Miyata, T. 2005. MAFFT version 5: improvement in accuracy
581 of multiple sequence alignment. Nucleic Acids Res. 33, 511-518.
582 Katoh, K., Standley, D.M. 2013. MAFFT multiple sequence alignment software version 7:
583 improvements in performance and usability. Mol. Biol. Evol. 30, 772-780.
584 Kechin, A., Boyarskikh, U., Kel, A., Filipenko, M. 2017. cutPrimers: a new tool for accurate
585 cutting of primers from reads of targeted next generation sequencing. J. Comput. Biol.
586 24, 1138-1143.
587 Kettenburg, G., Kistler, A., Ranaivoson, H.C., Ahyong, V., Andrianiaina, A., Andry, S.,
588 DeRisi, J.L., Gentles, A., Raharinosy, V., Randriambolamanantsoa, T.H.,
589 Ravelomanantsoa, N.A.F., Tato, C.M., Dussart, P., Heraud, J.M., Brook, C.E. 2022. Full
590 genome *Nobecovirus* sequences from Malagasy fruit bats define a unique evolutionary
591 history for this coronavirus clade. Front Public Health 10, 786060.
592 Kopylova, E., Noé, L., Touzet, H. 2012. SortMeRNA: fast and accurate filtering of ribosomal
593 RNAs in metatranscriptomic data. Bioinformatics 28, 3211-3217.
594 Langmead, B., Salzberg, S.L. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Meth. 9,
595 357-359.
596 Li, D., Liu, C.M., Luo, R., Sadakane, K., Lam, T.W. 2015. MEGAHIT: an ultra-fast single-
597 node solution for large and complex metagenomics assembly via succinct de Bruijn
598 graph. Bioinformatics 31, 1674-1646.

599 Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J.H., Wang, H., Crameri, G., Hu, Z.,
600 Zhang, H., Zhang, J., McEachern, J., Field, H., Daszak, P., Eaton, B.T., Zhang, S.,
601 Wang, L.F. 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science 310,
602 676-679.
603 Marsh, G.A., de Jong, C., Barr, J.A., Tachedjian, M., Smith, C., Middleton, D., Yu, M.,
604 Todd, S., Foord, A.J., Haring, V., Payne, J., Robinson, R., Broz, I., Crameri, G., Field,
605 H.E., Wang, L.F. 2012. Cedar virus: a novel Henipavirus isolated from Australian bats.
606 PLoS Pathog. 8, e1002836.
607 Martella, V., Lorusso, E., Banyai, K., Decaro, N., Corrente, M., Elia, G., Cavalli, A.,
608 Radogna, A., Costantini, V., Saif, L.J., Lavazza, A., Di Trani, L., Buonavoglia, C. 2008.
609 Identification of a porcine calicivirus related genetically to human sapoviruses. J. Clin.
610 Microbiol. 46, 1907-1913.
611 Mishra, N., Fagbo, S.F., Alagaili, A.N., Nitido, A., Williams, S.H., Ng, J., Lee, B.,
612 Durosinlorun, A., Garcia, J.A., Jain, K., Kapoor, V., Epstein, J.H., Briese, T., Memish,
613 Z.A., Olival, K.J., Lipkin, W.I. 2019. A viral metagenomic survey identifies known and
614 novel mammalian viruses in bats from Saudi Arabia. PLoS One 14, e0214227.
615 Mombo, I.M., Berthet, N., Bouchier, C., Fair, J.N., Schneider, B.S., Renaud, F., Leroy, E.M.,
616 Rougeron, V. 2014. Characterization of a genogroup I sapovirus isolated from
617 chimpanzees in the Republic of Congo. Genome Announc. 2, e00680-14..
618 Murakami, S., Kitamura, T., Suzuki, J., Sato, R., Aoi, T., Fujii, M., Matsugo, H., Kamiki, H.,
619 Ishida, H., Takenaka-Uema, A., Shimojima, M., Horimoto, T. 2020. Detection and
620 characterization of bat sarbecovirus phylogenetically related to SARS-CoV-2, Japan.
621 Emerg. Infect. Dis. 26, 3025-3029.
622 Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q. 2014. IQ-TREE: A fast and
623 effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol.
624 Biol. Evol. 32, 268-274.
625 Ning, S., Dai, Z., Zhao, C., Feng, Z., Jin, K., Yang, S., Shen, Q., Wang, X., Sun, R., Zhang,
626 W. 2022. Novel putative pathogenic viruses identified in pangolins by mining
627 metagenomic data. J. Med. Virol. 94, 2500-2509.
628 Nouri, S., Salem, N., Nigg, J.C., Falk, B.W. 2015. Diverse array of new viral sequences
629 identified in worldwide populations of the Asian citrus psyllid (*Diaphorina citri*) using
630 viral metagenomics. J. Virol. 90, 2434-4245.
631 Oka, T., Wang, Q., Katayama, K., Saif, L.J. 2015. Comprehensive review of human
632 sapoviruses. Clin. Microbiol. Rev. 28, 32-53.

633 Philbey, A.W., Kirkland, P.D., Ross, A.D., Davis, R.J., Gleeson, A.B., Love, R.J., Daniels,
634 P.W., Gould, A.R., Hyatt, A.D. 1998. An apparently new virus (family
635 *Paramyxoviridae*) infectious for pigs, humans, and fruit bats. Emerg. Infect. Dis. 4, 269-
636 271.
637 Plowright, R.K., Foley, P., Field, H.E., Dobson, A.P., Foley, J.E., Eby, P., Daszak, P. 2011.
638 Urban habituation, ecological connectivity and epidemic dampening: the emergence of
639 Hendra virus from flying foxes (*Pteropus* spp.). Proc Biol Sci. 278, 3703-3712.
640 Prada, D., Boyd, V., Baker, M., Jackson, B., O'Dea, M. 2019a. Insights into Australian bat
641 lyssavirus in insectivorous bats of Western Australia. Trop. Med. Infect. Dis. 4, 46.
642 Prada, D., Boyd, V., Baker, M.L., O'Dea, M., Jackson, B. 2019b. Viral diversity of microbats
643 within the south west botanical province of Western Australia. Viruses 11, 1157.
644 Rehse-Küpper, B., Casals, J., Rehse, E., Ackermann, R. 1976. Eyach-an arthropod-borne
645 virus related to Colorado tick fever virus in the Federal Republic of Germany. Acta
646 Virol. 20, 339-342.
647 Shi, M., Lin, X.D., Tian, J.H., Chen, L.J., Chen, X., Li, C.X., Qin, X.C., Li, J., Cao, J.P.,
648 Eden, J.S., Buchmann, J., Wang, W., Xu, J., Holmes, E.C., Zhang, Y.Z. 2016.
649 Redefining the invertebrate RNA virosphere. Nature 540, 539-543.
650 Smith, C.S., de Jong, C.E., Meers, J., Henning, J., Wang, L., Field, H.E. 2016. Coronavirus
651 infection and diversity in bats in the Australasian region. Ecohealth 13, 72-82.
652 Temmam, S., Vongphayloth, K., Baquero, E., Munier, S., Bonomi, M., Regnault, B.,
653 Douangboubpha, B., Karami, Y., Chrétien, D., Sanamxay, D., Xayaphet, V.,
654 Paphaphanh, P., Lacoste, V., Somlor, S., Lakeomany, K., Phommavanh, N., Pérot, P.,
655 Dehan, O., Amara, F., Donati, F., Bigot, T., Nilges, M., Rey, F.A., van der Werf, S.,
656 Brey, P.T., Eloit, M. 2022. Bat coronaviruses related to SARS-CoV-2 and infectious for
657 human cells. Nature 604, 330-336.
658 Thayer, R.M., Power, M.D., Bryant, M.L., Gardner, M.B., Barr, P.J., Luciw, P.A. 1987.
659 Sequence relationships of type D retroviruses which cause simian acquired
660 immunodeficiency syndrome. Virology,157, 317-329.
661 Timmiss, L.A., Martin, J.M., Murray, N.J., Welbergen, J.A., Westcott, D., McKeown, A.,
662 Kingsford, R.T. 2021. Threatened but not conserved: flying-fox roosting and foraging
663 habitat in Australia. Aust. J. Zool. 68, 226-233.
664 Tse, H., Chan, W.M., Li, K.S., Lau, S.K., Woo, P.C., Yuen, K.Y. 2012. Discovery and
665 genomic characterization of a novel bat sapovirus with unusual genomic features and
666 phylogenetic position. PLoS One 7, e34987.

667 Turnbull, M.G., Douville, R.N. 2018. Related endogenous retrovirus-K elements harbor
668 distinct protease active site motifs. Front. Microbiol. 9, 1577.
669 Van Brussel, K., Holmes, E.C. 2022. Zoonotic disease and virome diversity in bats. Curr.
670 Opin. Virol. 52, 192-202.
671 Vanmechelen, B., Merino, M., Vergote, V., Laenen, L., Thijssen, M., Martí-Carreras, J.,
672 Claerebout, E., Maes, P. 2021. Exploration of the *Ixodes ricinus* virosphere unveils an
673 extensive virus diversity including novel coltiviruses and other reoviruses. Virus Evol.7,
674 veab066.
675 Wacharapluesadee, S., Tan, C.W., Maneeorn, P., Duengkae, P., Zhu, F., Joyjinda, Y.,
676 Kaewpom, T., Chia, W.N., Ampoot, W., Lim, B.L., Worachotsueptrakun, K., Chen,
677 V.C., Sirichan, N., Ruchisrisarod, C., Rodpan, A., Noradechanon, K., Phaichana, T.,
678 Jantarat, N., Thongnumchaima, B., Tu, C., Crameri, G., Stokes, M.M., Hemachudha, T.,
679 Wang, L.F. 2021. Evidence for SARS-CoV-2 related coronaviruses circulating in bats
680 and pangolins in Southeast Asia. Nat. Commun. 12, 972.
681 Wang, J., Anderson, D.E., Halpin, K., Hong, X., Chen, H., Walker, S., Valdeter, S., van der
682 Heide, B., Neave, M.J., Bingham, J., O'Brien, D., Eagles, D., Wang, L.F., Williams, D.T.
683 2021. A new Hendra virus genotype found in Australian flying foxes. Virol. J. 18, 197.
684 Weiss, S., Dabrowski, P.W., Kurth, A., Leendertz, S.A.J., Leendertz, F.H. 2017. A novel
685 Coltivirus-related virus isolated from free-tailed bats from Côte d'Ivoire is able to infect
686 human cells *in vitro.* Virol. J. 14, 181.
687 Williams, N.S.G., McDonnell, M.J., Phelan, G.K., Keim, L.D., Van Der Ree, R. 2006. Range
688 expansion due to urbanization: Increased food resources attract Grey-headed Flying-
689 foxes (*Pteropus poliocephalus*) to Melbourne. Austral. Ecol. 31, 190-198.
690 Wu, Z., Yang, L., Ren, X., He, G., Zhang, J., Yang, J., Qian, Z., Dong, J., Sun, L., Zhu, Y.,
691 Du, J., Yang, F., Zhang, S., Jin, Q. 2016. Deciphering the bat virome catalog to better
692 understand the ecological diversity of bat viruses and the bat origin of emerging
693 infectious diseases. ISME J. 10, 609-620.
694 Xu, Z., Feng, Y., Chen, X., Shi, M., Fu, S., Yang, W., Liu, W.J., Gao, G.F., Liang, G. 2022.
695 Virome of bat-infesting arthropods: highly divergent viruses in different vectors. J. Virol.
696 96, e0146421.
697 Yang, Z., He, B., Lu, Z., Mi, S., Jiang, J., Liu, Z., Tu, C., Gong, W. 2021. Mammalian
698 birnaviruses identified in pigs infected by classical swine fever virus. Virus Evol. 7,
699 veab084.

Yinda, C.K., Conceição-Neto, N., Zeller, M., Heylen, E., Maes, P., Ghogomu, S.M., Van Ranst, M., Matthijnssens, J. 2017. Novel highly divergent sapoviruses detected by metagenomics analysis in straw-colored fruit bats in Cameroon. Emerg. Microbes Infect. 6, e38.

Yob, J.M., Field, H., Rashdi, A.M., Morrissy, C., van der Heide, B., Rota, P., bin Adzhar, A., White, J., Daniels, P., Jamaluddin, A., Ksiazek, T. 2001. Nipah virus infection in bats (order Chiroptera) in peninsular Malaysia. Emerg. Infect. Dis. 7, 439-441.

Zhou, H., Ji, J., Chen, X., Bi, Y., Li, J., Wang, Q., Hu, T., Song, H., Zhao, R., Chen, Y., Cui, M., Zhang, Y., Hughes, A.C., Holmes, E.C., Shi, W. 2021. Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses. Cell 184, 4380-4391.e14.

Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., Chen, H.D., Chen, J., Luo, Y., Guo, H., Jiang, R.D., Liu, M.Q., Chen, Y., Shen, X.R., Wang, X., Zheng, X.S., Zhao, K., Chen, Q.J., Deng, F., Liu, L.L., Yan, B., Zhan, F.X., Wang, Y.Y., Xiao, G.F., Shi, Z.L. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579, 270-273.

Zhu, W., Lomsadze, A., Borodovsky, M. 2010. *Ab initio* gene identification in metagenomic sequences. Nucleic Acids Res. 38, e132.

Zhu, W., Yang, J., Lu, S., Jin, D., Pu, J., Wu, S., Luo, X.L., Liu, L., Li, Z., Xu, J. 2022. RNA virus diversity in birds and small mammals from Qinghai-Tibet plateau of China. Front. Microbiol. 13, 780651.