bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Sarbecoviruses of British Horseshoe Bats; Sequence Variation and Epidemiology

Ternenge Apaa[1], Amy J. Withers[1], Ceri Staley[1], Adam Blanchard[1], Malcolm Bennett[1], Samantha Bremner-Harrison[2], Elizabeth A. Chadwick[3], Frank Hailer[3], Stephen W.R. Harrison[2], Mathew Loose[4], Fiona Mathews[5], Rachael Tarlinton[1*]

[i]1. School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington UK

2. School of Animal, Rural and Environmental Sciences , Nottingham Trent University, Southwell UK

3. Organisms and Environment, School of Biosciences, Cardiff University

4. School of Life Sciences, University of Nottingham, Nottingham, UK

5. School of Life sciences, University of Sussex

*Corresponding author: Rachael Tarlinton Rachael.tarlinton@nottingham.ac.uk

Keywords

Sarbecovirus, Horseshoe bat, *Rhinolophus hipposideros, Rhinolophus ferrumequinum,* Coronavirus, UK

Repositories

All GB lesser horseshoe Sarbecovirus genomes assembled from this study have been deposited in the NCBI GenBank database under accession numbers OP776338 – OP776340 and OP837780 – OP837781. In addition, Illumina read datasets generated have been submitted under Bioproject accession number PRJNA895830, SRA accession numbers SRR22103527 – SRR22103531 and Biosample accession numbers SAMN31525260 – SAMN31525264.

Abstract

Horseshoe bats are the natural hosts of the *Sarbecovirus* subgenus that includes SARS-CoV-1 and 2. Despite the devastating impacts of the COVID-19 pandemic, there is still little known about the underlying epidemiology and virology of sarbecoviruses in their natural hosts, leaving large gaps in

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

our pandemic preparedness. Here we describe the results of PCR testing for sarbecoviruses in the two horseshoe bat species (*Rhinolophus hipposideros* and *R. ferrumequinum*) present in Great Britain, collected in 2021-22 during the peak of COVID-19 pandemic. One hundred and ninety seven *R. hipposideros* samples from 33 roost sites and 277 *R. ferremequinum* samples from 20 roost sites were tested. No coronaviruses were detected in any samples from *R. ferrumequinum* whereas 44% and 56% of individual and pooled (respectively) faecal samples from *R. hipposideros* across multiple roost sites tested positive in a sarbecovirus-specific qPCR. Full genome sequences were generated from three of the positive samples (and partial genomes from two more) using Illumina RNAseq on unenriched samples. Phylogenetic analyses showed that the obtained sequences belong to the same monophyletic clade, with >95% similarity, as previously reported European isolates from *R. hipposideros*. The sequences differed in the presence or absence of accessory genes ORF 7b, 9b and 10. All lacked the furin cleavage site of SARS-CoV-2 spike gene and are therefore unlikely to be infective for humans. These results demonstrate a lack, or at least low incidence, of SARS-CoV-2 spill over from humans to susceptible GB bats, and confirm that sarbecovirus infection is widespread in *R. hipposideros*. Despite frequently sharing roost sites with *R. ferrumequinum*, no evidence of cross-species transmission was found.

Background

The most widely accepted explanation for the origin of the SARS-CoV-2 pandemic is that it arose from animals held in the Wuhan market in late 2019 (Holmes, Goldstein et al. 2021). It has been proposed that masked palm civets (*Paguma larvatai*) acted a bridging host for SARS-CoV transmission from an unknown ancestral reservoir in a species of horseshoe bats to humans (Lau, Woo et al. 2005, Wang, Yan et al. 2005). Widespread onward transmission of SARS-CoV-2 from humans to other mammals has also occurred during the human pandemic, including large outbreaks in farmed mink (*Neovison vison*) (Eckstrand, Baldwin et al. 2021, Larsen, Fonager et al. 2021, Lu, Sikkema et al. 2021, Rabalski, Kosinski et al. 2021), and the establishment of a new reservoir in wild white-tailed deer (*Odocoileus virginianus*) in the USA (Hale, Dennis et al. 2022, Willgert, Didelot et al. 2022), and repeated infection (though without the establishment of endemic transmission) in domestic cats (*Felis catus*) (Hosie, Hofmann-Lehmann et al. 2021, Bienzle, Rousseau et al. 2022, Goletic, Goletic et al. 2022, Jairak, Chamsai et al. 2022). Repeated reports of transmission back to humans from these species have also occurred (Lu, Sikkema et al. 2021, Pickering, Lung et al. 2022, Sila, Sunghan et al. 2022) highlighting the risk of new animal reservoirs developing, potentially as sources of new viral variants that might evade vaccine-induced immunity. Sporadic cases have also been reported in domestic dogs (*Canis familiaris*), hamsters (*Mesocricetus auratus*), cattle (*Bos Taurus)* and camels (*Camelus dromedarius)*, and a large range of primates, non-domestic felids, mustelids, hippos (*Hippopotamus amphibius*) and manatees (*Trichechus manatus*) (reviewed in (Pappas, Vokou et al. 2022).

The natural hosts of sarbecoviruses (SARS-like betacoronaviruses) are insectivorous horseshoe bats (represented by a single extant genus *Rhinolophus* within the family Rhinolophidae; superfamily Rhinolophoidea; sub-order Yinpterochiroptera) and the related old world roundleaf bats of the family Hipposideridae which share the same superfamily. There are around 180 known species of Rhinolophidae and Hipposideridae with ranges across Eurasia and Africa, and sarbecoviruses have been detected in at least 30 of them (Muylaert, Kingston et al. 2022). For species and sites where more extensive studies have been performed (generally SE Asia), it is apparent that species- and site-specific clustering of virus sequences occurs (Boni, Lemey et al. 2020, Lytras, Hughes et al. 2022). Geographical hotspots of both bat and viral species are apparent in SE Asia (Wong, Li et al. 2019, Muylaert, Kingston et al. 2022).

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

There are, in contrast, very few reports of sarbecoviruses in the sub-order Yangochiroptera: a small number of isolated reports in individual animals or pooled samples of Molossidae and Vespertilionidae families have been recorded in multispecies studies (Muylaert, Kingston et al. 2022, Seifert, Bai et al. 2022, Wang, Pan et al. 2022).

There is also evidence from multiple studies for considerable variation in both sarbecovirus carriage and sequences within host and viral lineages in bats, with implications for potential cross species transmission (Wang, Pan et al. 2022). While the exact determinants of what makes sarbecoviruses more likely to cross the species barrier into humans are not known, one clear limiting factor is the ability of the virus spike protein (the protein that binds to the host cell facilitating virus entry) to bind to the human ACE-2 protein (the receptor for SARS-CoV and SARS-CoV-2 in humans) (Conceicao, Thakur et al. 2020). Considerable mutation and antigenic variation is present in this spike protein in SARS-CoV-2 isolates and bat sarbecoviruses (Boni, Lemey et al. 2020, Li, Lai et al. 2021). In SARs-CoV-2 this protein also displays a distinctive cleavage site for the furin protease that is missing in many bat isolates and appears to be a critical factor in transmission and pathogenesis in humans (Chan and Zhan 2022). Complicating matters further, it is apparent that there is frequent viral recombination between different sarbecovirus lineages (Boni, Lemey et al. 2020, Lytras, Hughes et al. 2022).

Two species of horseshoe bats are present in Great Britain, the greater horseshoe bat (*Rhinolophus ferrumequinum)* and lesser horseshoe bat (*Rhinolophus hipposideros*), both with distributions largely restricted to southwest England and Wales. Neither are common at a national scale and both are regarded as rare in Europe being provided with considerable protection under the Habitats Directive and Eurobats Agreement (JNCC 2022, JNCC 2022). In Britain, populations comprise approximately 13,000 greater horseshoe bats and 50,000 lesser horseshoe bats, (Mathews, Kubasiewicz et al. 2018) but both are increasing over time, recovering from historically depressed levels. They are considered Least Concern on the British Red List, but the greater horseshoe bat is near threatened in Europe and there have been recent extinctions of populations in several countries (Temple and Terry 2007). Britain and Ireland form the extreme north-western end of their geographic ranges (only *R. hipposideros* is found in Ireland), which extends through central and southern Europe to Central Asia for lesser horseshoe bats and across Asia to Japan for greater horseshoe bats. The two species can share roost sites and both use a variety of roosts across the year, including maternity and hibernation sites (Russo 2022).

Sarbecoviruses have been reported previously from both bat species. Partial RDRP (Rnase dependent RNA polymerase) sequences have been described from Chinese, Italian and Bulgarian greater horseshoe bats (Drexler, Gloza-Rausch et al. 2010, Balboni, Palladini et al. 2011, Balboni, Gallina et al. 2012, Hu, Zeng et al. 2017, Muth, Corman et al. 2018) at a prevalence of 26-42% in European samples. Full sequences are also available for a number of Chinese, one Korean and one Russian greater horseshoe bat sarbecoviruses; the Asian sequences are phylogenetically very distinct from the European isolates (Lin, Wang et al. 2017, Alkhovsky, Lenshin et al. 2022, Sander, Moreira-Soto et al. 2022, Wu, Han et al. 2022).

Similarly, partial RDRP sequences have been reported in Slovenian and Polish lesser horseshoe bats (Rihtaric, Hostnik et al. 2010), at a prevalence of 31%, and Spanish lesser horseshoe bats at a prevalence of 7.1% (Muth, Corman et al. 2018). More recently full genome sequences from a Russian and British lesser horseshoe bat (Crook, Murphy et al. 2021, Alkhovsky, Lenshin et al. 2022) have also been reported. Resequencing of partial RDRP genes from the Spanish, Italian and Slovenian isolates (and the full genome sequences for the UK and Russian isolates) has demonstrated a lack of the furin cleavage site present in SARS-CoV-2 (Sander, Moreira-Soto et al. 2022). The European

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

greater and lesser horseshoe bat sequences are also phylogenetically distinct from each other (Sander, Moreira-Soto et al. 2022) with the two Russian isolates displaying 59-95% amino acid similarity to each other depending on the gene (Alkhovsky, Lenshin et al. 2022).

In addition to their being hosts of various sarbecoviruses (Wong, Li et al. 2019), horseshoe bats have been identified as potentially susceptible to infection with SARS-CoV-2 (Cook, Grant et al. 2021, Common, Shadbolt et al. 2022, Logeot, Mauroy et al. 2022). This led us to screen British horseshoe bat populations, at the height of the SARS-CoV-2 pandemic in western Europe, for the presence of sarbecoviruses with the aim to confirm or rule out the establishment of SARS-CoV-2 circulation in these animals.

Materials and methods

Sample collection

A total of 517 bat samples (oronasal swabs, external swabs of the rectal region, and faecal samples), including 474 from horseshoe bats, (Table 1) were collected from roost sites or animals in care in the UK between April 2021 and February 2022 during which time human cases in the UK underwent successive alpha, delta and omicron SARS-CoV-2 waves (UKHSA 2022) . Faecal samples included those collected from individuals in capture bags during routine population monitoring (during which sex and age were also recorded where possible) and pooled samples of fresh droppings collected from the floor of bat roosts. All samples were collected secondarily to regular licenced population monitoring efforts. Handling of animals was conducted under Natural England Licence 2022-61108-Sci-Sci (Mathews), and followed best practice guidelines for minimising the risk of human to bat SARS-CoV-2 transmission). Ethical approval was granted by the University of Nottingham School of Veterinary Medicine and Science Committee for Animal Research and Ethics (CARE), and the University of Sussex Animal Welfare and Ethical Review Board. Samples were preserved in RNAlater at room temperature and sent to the University of Nottingham where they were stored at -20 ℃ until RNA extraction. Samples were collected from five bat species from 54 sites: 197 from lesser horseshoe bats (*Rhinolophus hipposideros*) at 33 sites, 277 for greater horseshoe bats (*Rhinolophus ferrumequinum*) from 20 sites, 10 from common pipistrelles (*Pipistrellus pipistrellus*), 32 from Daubentons bats (*Myotis daubentonii*) and 1 from a Natterer's bat (*Myotis nattereri*) (Table 1 and Supplementary information).

RNA extraction, reverse transcriptase (RT) and RNA-dependent RNA polymerase (RDRP) gene coronaviruses generic conventional PCR and envelope gene sarbecovirus-specific real-time PCR

RNA extraction from bat faecal and oronasal swabs, and cell culture supernatant as positive control, was carried out using the Macherey-Nagel RNA tissue extraction kit as per manufacturer's instructions. The Wuhan SARS-CoV-2 strain positive control sample used throughout this study was kindly donated by Dr Christopher Coleman (Division of Infection, Immunity and Microbes, School of Life Sciences, University of Nottingham, UK). RT was performed in two steps, using M-MLV-RT and random hexamer primers (Promega) as per manufacturer's instructions. All cDNA products were stored at -20 °C for conventional PCR.

A generic pan-coronavirus PCR assay published by (Woo, Lau et al. 2005) was used to amplify a 440 bp fragment of the coronavirus RDRP gene using Q5® Hot Start High-Fidelity DNA Polymerase (New England Biolabs cat no: M0493S). PCR products were purified using the Nucleospin® extract II kit

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(Macherey-Nagel) according to manufacturer's instructions and were Sanger sequenced (Eurofins UK).

Real-time PCR was carried out using the Promega GoTaq® Probe 1-Step RT-qPCR System (Promega) with Sarbecovirus-specific envelope gene primers from bat RNA samples as published by (Corman, Landt et al. 2020).

RNA and cDNA quality control was assessed via partial amplification of 148 bp of the bat mitochondrial cytochrome b gene using a published conventional PCR protocol (Lopez-Oceja, Gamarra et al. 2016).

Next generation sequencing and genome analyses

RNA sequencing was performed by Novogene UK, using the Illumina NovaSeq 6000 platform. Quality filtering and trimming to remove adapters, duplicates and low quality reads was achieved using fastp v0.23.1 (Chen, Zhou et al. 2018). Kraken2 v2.1.2 was used for taxonomic classification reads against the Kraken2 viral Refseq database (Wood, Lu et al. 2019) (retrieved on 9th June 2022). Reads were assembled using the coronaSPAdes option in SPAdes genome assembler v3.15.4 (Meleshko, Hajirasouliha et al. 2021) using default parameters. While CheckV v1.0.1, a fully automated command-line pipeline, was used for identification and quality assessment of contigs, contigs were also queried against the NCBI custom BLASTN (v2.12.0) viral database (Altschul, Gish et al. 1990) (retrieved on 3rd July 2022) to confirm results from CheckV. Assembled contigs were indexed and contigs that were classified and assessed as complete bat Sarbecovirus genomes were extracted for downstream analysis using samtools v1.16.1 faidx option (Danecek, Bonfield et al. 2021). Individual reads for each sample were mapped back to identified contigs using Minimap2 (Li 2018), read coverage and depth were generated using samtools (Danecek, Bonfield et al. 2021). Assembled genomes were annotated in Geneious Prime® (v.2022.2.2) using NCBI coronavirus reference sequences and visualized for the presence of the structural, non-structural protein, accessory genes and to generate linear genome maps and data on individual gene location, composition, and nucleotide length. Reads were also mapped to the prototype European horseshoe bat (*Rhinolophus blasii,* Blasius horseshoe bat Bulgarian isolate) sarbecovirus reference genome (NC014470) using BBMap for variant calling, SNPs were viewed, and data exported in Geneious Prime® (v.2022.2.2).

Phylogenetic analysis

Complete coronavirus genomes, extracted RDRP, spike, envelope and nucleocapsid nucleotide sequences from sarbecovirus genomes assembled in this study, and a total of 198 reference sarbecovirus genomes (including non-human sarbecovirus isolates, SARS-CoV and SARS-CoV-2) downloaded from NCBI, were aligned using mafft v7.490 (Katoh and Standley 2013). Maximum likelihood phylogenetic trees were reconstructed based on complete coronavirus genomes, and four different genes using IQ-TREE v2.0.7 (Minh, Schmidt et al. 2020), using 1000 bootstrap approximations following implementation of UFBoot2 within IQ-TREE v2.0.7 to evaluate branch support (Hoang, Chernomor et al. 2018). The ModelFinder option was included in command-line to select the best fitting nucleotide substitution model for phylogenetic reconstruction (Kalyaanamoorthy, Minh et al. 2017). Phylogenetic trees were visualized and annotated in FigTree v1.4.4 (https://github.com/rambaut/figtree/).

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Spike glycoprotein comparison, identification of furin cleavage site (FCS) and transmembrane protease serine 2 (TMPRSS2), receptor binding domain (RBD) homology modelling and structural analysis

To identify and compare the presence of FCS and TMPRSS2 between UK Coronavirus genomes and related CoVs, RBD spike proteins from the six UK Coronaviruses (five from this study and one reported by Crook et al 2020), SARS-CoV, SARS-CoV-2 and related Beta coronaviruses were aligned, viewed and pictures generated in Jalview (Procter, Carstairs et al. 2021). The model for the receptor binding domain (RBD) and hACE2 protein complex were constructed using SWISS-MODEL to assess the amino acid residues and structural differences between UK bat CoV, SARS-CoV, and SARS-CoV-2 using an RBD-hACE2 complex. The SWISS-MODEL template library was search for evolutionary related structures matching the target UK Coronavirus RBD (residues 321 – 515) and hACE2 (PDB: Q9BYF1) protein sequences using BLAST HHblits database. The most suitable template (PDB: 6vw1) with the highest selected Global Model Quality Estimate (GMQE) of 0.75 (range: 0 – 1) and the lowest average Quaternary Structure Quality Mean Estimate (QMEAN) of – 0.64 was selected to build models for the five UK Coronavirus RBD proteins generated from this study. The structural analysis and verification server (SAVES) was used to validate 3D structures, create Ramachandran plots, and classify amino acid residue torsion angles of each structure via implementation of PROCHECK AND ERRAT2 (Laskowski, MacArthur et al. 1993), respectively. In addition, receptor binding residues detection and binding energies of 3D model complex structures generated were calculated using PRODIGY (Vangone and Bonvin 2015, Honorato, Koukos et al. 2021). Visualization, superimposition, alignment, and generation of figures for presentation was done using UCSF Chimera (Pettersen, Goddard et al. 2004).

Geospatial mapping and data analysis

Maps of roost sites positive and negative for coronavirus PCR were created in QGIS (version 3.28.2). to account for pooled and individual samples, map circle sizes were set to 5 proportional sizes as follows: <3 bats, 3-5 or one pooled sample, 5-10 (pools count as 3), 10-120 (pools count as 3) and >20 (pools count as 3). Chi-square test was used to evaluate potential differences in prevalence with sex and age-class (SPSS, Amos 28.0.0).

Results

No samples tested positive in the pan-coronavirus assay, and no oronasal swabs tested positive in the sarbecovirus-specific PCR assay for either species, However, numerous lesser horseshoe bat faecal samples tested positive with the Sarbecovirus qPCR assay (Table 1). Within the five roosts from which >5 individual faecal samples and/or rectal swabs were collected, all yielded at least one positive sample, and the percentage positivity ranged from 5% to 25% (mean at individual level across those sites, 16/100 individuals, 16%, 95% CI (Wilson's) 10.0%-24.6%). Twenty-two sites yielded only single, pooled samples, of which 10 (45%; 95% CI (Wilson's) 27.0%-65.3%) were positive (Supplementary information). The locations and infection status of lesser horseshoe bat roost sites yielding either >3 individual faecal samples or at least 1 pooled faecal sample are shown in Figure 1; of these, 21/31 (68%; 95% CI (Wilson's) 50.0%-81.4%) roosts yielded at least one positive sample.

Age and sex were recorded for 41 and 37 animals, respectively, when individual faecal samples were collected and examined; among this subset of animals tested, there were no significant (chi-square) differences in the frequency of positive samples by age (19.5% adult, 8% juvenile) or sex (12.5% female, 21.9% male).

Table 1: Summary of bat species, sample type and positivity rates for Sarbecovirus E gene qPCR

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Bat Species	Type of sample	Number of samples	Sarbecovirus E gene qPCR
Lesser horseshoe	Oronasal swabs	71	0 (0%)
	Rectal swabs	56	3 (5.4%)
	Individual faeces samples	39	17 (43.5%)
	Pooled faeces samples	31	17 (54.8%)
Greater horseshoe	Oronasal swabs	90	0
	Rectal swabs	69	0
	Individual faeces samples	91	0
	Pooled faeces samples	28	0
Common pipistrelle	Oronasal swabs	3	0
	Rectal swabs	-	-
	Individual faeces samples	7	0
	Pooled faeces samples	-	-
Natterer's	Oronasal swabs	1	0
	Rectal swabs	-	-
	Individual faeces samples	-	-
	Pooled faeces samples	-	-
Daubenton's	Oronasal swabs	5	0
	Rectal swabs	8	0
	Individual faeces samples	16	0
	Pooled faeces samples	-	-

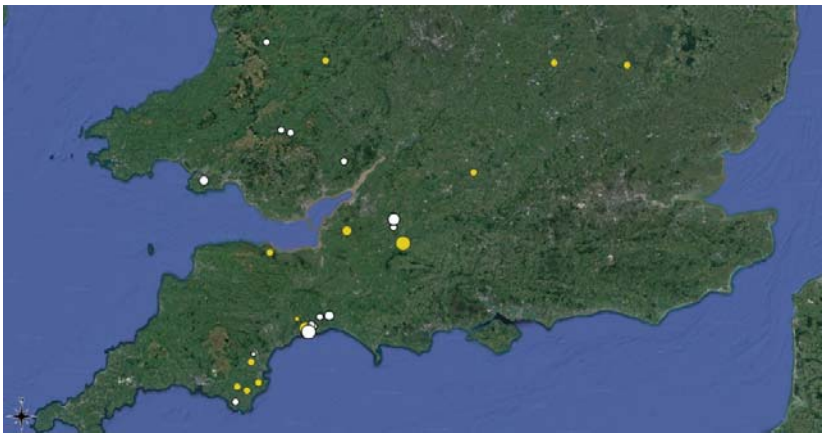


Figure 1: Roost sites of *Sarbecovirus* detected by PCR in faeces from lesser horseshoe bats (*Rhinolophus hipposideros)*, Yellow=negative, White= positive. Circle sizes are proportional to number of samples from the site.

Taxonomic classification, genome assembly

Taxonomic classification using Kraken2 identified reads assigned to other viral operational taxonomic units, however, only reads classified to the Coronaviridae viral family are reported in this study. Paired end reads were generated from RNA extracted from six lesser horseshoe bat faecal

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

samples. Between 98 to 99% of reads were unclassified. The percentage of classified viral reads from the datasets examined varied from 0.49 – 1.18%, while the total percentage of classified viral reads assigned to the family Coronaviridae varied from 0.01 – 63.58%. Three samples including RhGB04, RhGB05 and RhGB06 recorded approximately 50% of classified viral reads assigned to Coronaviridae (Supplementary information).

De novo assembly of datasets from lesser horseshoe bats yielded five genomes made up of single contigs. The length of the assembled genomes varied from 28.2 – 30.6 kb with one short contig of 12.6k. CheckV analysis demonstrated that all the assembled contigs were closely related to beta coronaviruses, specifically the subgenus Sarbecovirus. Sarbecovirus genomes identified were assessed to have 94 – 100% (4/5) and 42% (1/%) quality/completeness, all genomes had 0% contamination except for RhGB05 with 3.17% contamination. Overall, all the assembled genomes shared 94.5-97.7% average amino acid identity with the reference genome (RhGB01 MW1957) identified by CheckV (Supplementary information).

Assessment of reads mapped to the assembled genomes demonstrated over 7000, 4000, 302,000, 420,000 and 500,000 reads mapped to RhGB02, RhGB03, RhGB04, RhGB05 and RhGB06, respectively. The genome assembled from RhGB03 had the least average mean depth of 22X and was missing ORF10 and the 3' UTR, while RhGB04 (partial genome of 12.6 kb missing the 5' end of ORF 1ab) recorded the highest mean depth coverage of 3446X (Supplementary information).

Genome annotation and organization of British bat CoVs

Genome annotation in Geneious Prime using NCBI SARS-CoV-2 and SARS-related reference sequences, confirmed a similar genome organization to the UK lesser horseshoe bat reference genome available from the NCBI database (Figure 2). Sarbecovirus genomes consist of a leader sequence (5'UTR), followed by ORF1ab gene with sixteen non-structural proteins (nsp1 – 16) making up about 2/3 of the viral genome. The assembled sarbecoviruses were made up of four major structural proteins including the spike (S), membrane (M), envelope (E) and nucleocapsid (N) proteins. While accessory proteins including ORF3b, 6a, 6b, 7a and variable ORFs recognisable for 7b, 9b and 10 were reported within the 3' region interspaced between the major structural proteins (Figure 2).

Sequence ID	5'UTR	ORF1ab (consists of nsp1 – 16)																S	ORF 3a	ORF 3b	E	M	ORF 6a	ORF 6b	ORF 7a	ORF 7b	OR F9b	N	ORF 10	3'UTR
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16													
Bat CoV RhGB01 MW719567	279	-	-	5721	-	-	-	-	-	-	-	-	2796	-	-	-	-	3771	813	452	-	-	-	-	-	129	291	-	117	948
RhGB02	248	-	-	5721	-	-	-	-	-	-	-	-	2796	-	-	-	-	3771	813	452	-	-	-	-	-		291	-	117	948
RhGB03	239	-	-	5721	-	-	-	-	-	-	-	-	2796	-	-	-	-	3771	813	452	-	-	-	-	-	129		547		
RhGB04												-		1628	-	-	-	3771	813	452	-	-	-	-	-	129	291	-		589
RhGB05	247	-	-	5721	-	-	-	-	-	-	-	-	2796	-	-	-	-	3771	813	452	-	-	-	-	-	129	291	-	117	948
RhGB06	249	-	-	5721	-	-	-	-	-	-	-	-	2796	-	-	-	-	3771	813	452	-	-	-	-	-		291	-	117	948

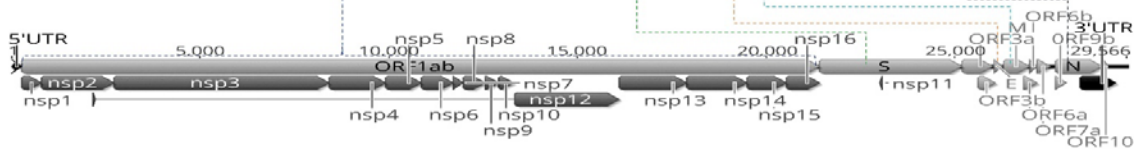


Figure 2: Genome organisation of lesser horseshoe bat Sarbecovirus sequences derived from this study and the reference UK genome RhGB001. Missing genes are shown in grey, lengths of nucleic acid segments are listed for each gene, LTR and isolate.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

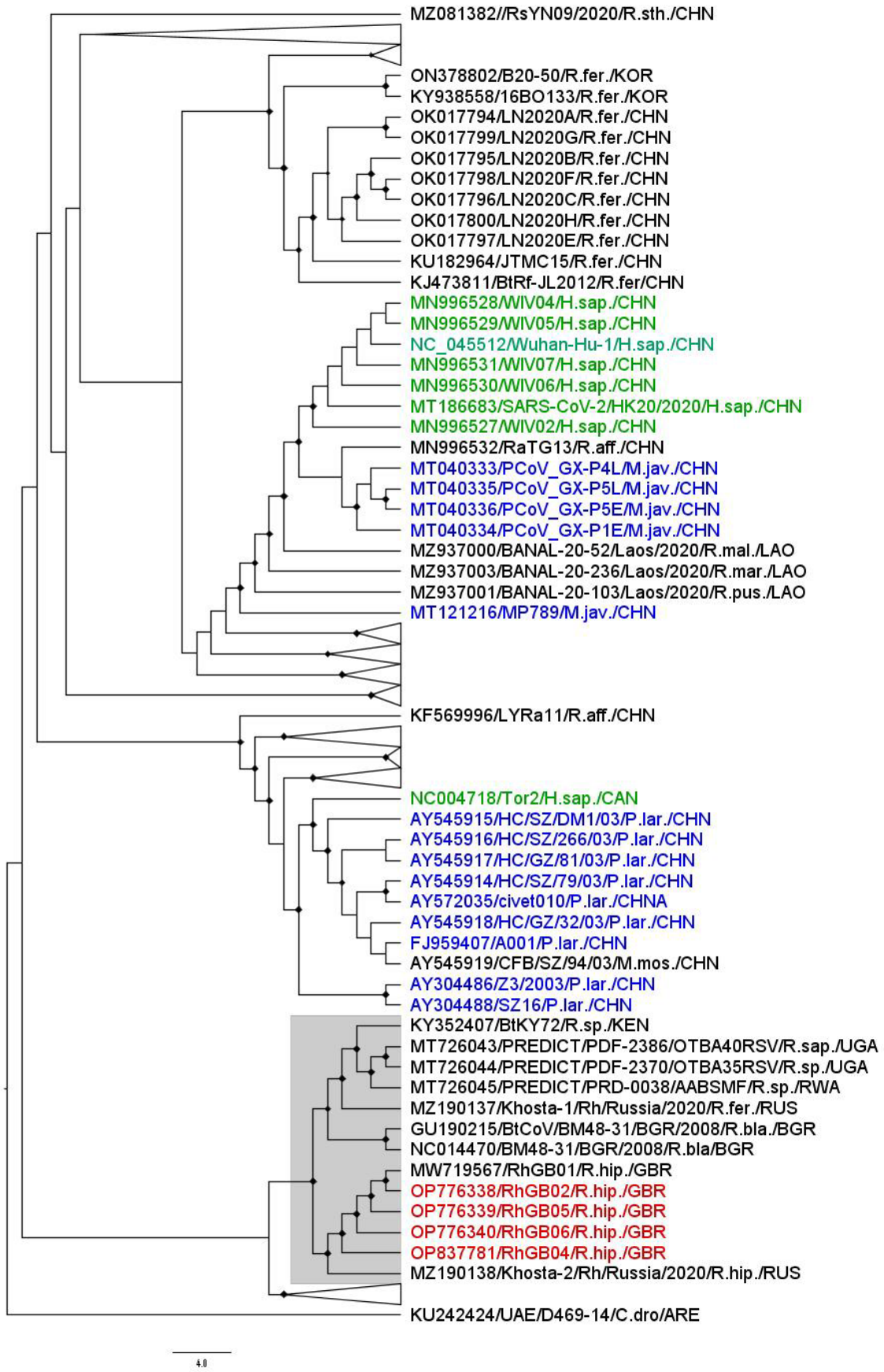
Variant calling

Variant calling following mapping of raw reads to the bat Bulgarian Sarbecovirus reference genome (Bat CoV BM48-31) available in Europe, demonstrated the presence of a series of single nucleotide polymorphism (SNPs) characterized by amino acid substitution distributed throughout the entire length of the genome. However, only a total of 22 unique SNPs were identified within the spike glycoprotein (Supplementary information) with minimum read coverage ≥ 10X, mapped quality ≥ 10% and variant frequency > 95% from the four full genomes assembled (Supplementary information for details). Overall, 15/22 SNPs resulting in amino acid substitution were found in RhGB02, 12/22 SNPs identified in RhGB03(12/22) and 8/22 SNPs were reported from both RhGB05 and RhGB06 with the highest read-depth coverage. One unique SNP at coding sequence position 3641 (V to A) was found in all four assembled genomes, and 5/22 unique SNPs, including changes at position 2399 (SR to KR), 2766 (D to E), 2781 (TT to TA), 3208 (E to K), and 3268 (I to V), in at least 3/4 of the assembled genomes. The position of these SNPs coincided with location of Sarbecovirus SD-1 and SD-2 subdomains, S1/S2 cleavage region and the S2 fusion subunit (Supplementary information).

Phylogenetic analysis

Results from the maximum likelihood phylogenetic trees drawn using complete coronavirus genomes (Supplementary information), spike glycoprotein (Figure 3) RDRP envelope and nucleocapsid (Supplementary information) nucleotide sequences showed that the Sarbecovirus genomes assembled from UK lesser horseshoe bats belong to the same monophyletic clade as the published European horseshoe bat Sarbecovirus genomes (RhGB01, Khosta1 and 2 and BM48-31), and are more similar to the small number of African horseshoe bat sarbecoviruses than to any Asian isolates, even those from the same host species.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

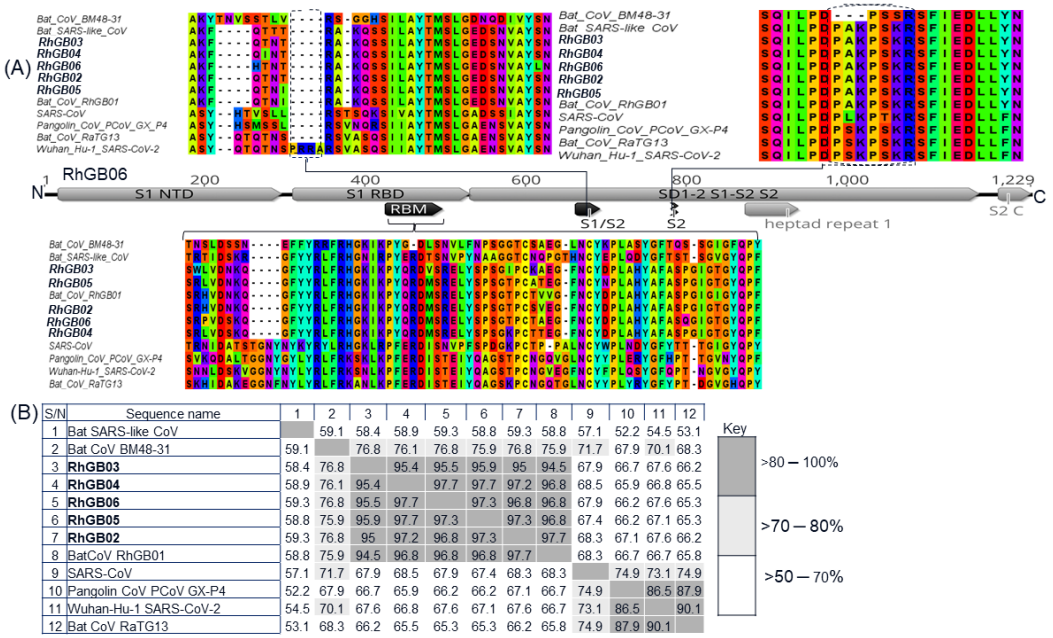
Figure 3: Maximum likelihood phylogenetic tree of S gene nucleic acid constructed with 1000 bootstrap approximation, rooted on the MERs coronavirus reference sequence. One hundred and ninety eight non-human Sarbecovirus genomes and reference sequences for major variants of SARS-CoV and SARS-CoV-2 were included (non-human SARS-CoV-2 isolates were not included). Clades of Asian bat coronavirus sequences (apart from greater horseshoe bat sequences) have been collapsed for clarity (represented as broad triangles). Red=isolates from this study, Green=human isolates, blue=isolates from other mammals. Sequences are named with Genbank ID, name from original study species of origin (eg R.hip= *Rhinolophus hipposideros*) and country of origin (eg GBR= Great Britain)

Spike glycoprotein comparison, identification of furin cleavage site (FCS) and Transmembrane protease serine 2 (TMPRSS2), receptor binding domain (RBD) homology modelling and structural analysis

All the UK bat coronavruses reported in this study shared the same spike glycoprotein amino acid length of 1256 aa with the only available complete bat SARS-like coronavirus genome reported from UK lesser horseshoe bats (Crook et al., 2021). The S1 RBD of UK bat coronaviruses reported from this study and the previously reported RhGB01 (Crook, Murphy et al. 2021) consisted of 220 amino acids, the RBM (receptor binding motif) of UK lesser horseshoe bat coronaviruses reported from this study have the same amino acid length (71 aa) as RhGB01.

The RBD (receptor binding domain) pairwise alignment percentage identity estimation (figure 4B) between the reported UK bat coronaviruses and selected Beta coronavirus reference sequences revealed an estimated percentage identity of 68% with SARS-CoV and 65 – 67% with SARS-CoV-2 related viruses. The RBD of all five UK bat coronaviruses reported from this study shared ≥ 95% amino acid homology with the RhGB01 reported by Crook, Murphy et al. (2021), and 76% amino acid homology with the Bulgarian horseshoe bat sarbecovirus (BM48-31) reported by Drexler, Gloza-Rausch et al. (2010). Further assessment of the S1 and S2 regions of UK bat coronaviruses RBDs showed the absence of a furin or S1/S2 cleavage site (figure 4A), while demonstrating the presence of host transmembrane serine protease 2 (TMPRSS2) or S2 cleavage site from reported UK bat CoVs (figure 4A).

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



(B)

S/N	Sequence name	1	2	3	4	5	6	7	8	9	10	11	12
1	Bat SARS-like CoV		59.1	58.4	58.9	59.3	58.8	59.3	58.8	57.1	52.2	54.5	53.1
2	Bat CoV BM48-31	59.1		76.8	76.1	76.8	75.9	76.8	75.9	71.7	67.9	70.1	68.3
3	**RhGB03**	58.4	76.8		95.4	95.5	95.9	95	94.5	67.9	66.7	67.6	66.2
4	**RhGB04**	58.9	76.1	95.4		97.7	97.7	97.2	96.8	68.5	65.9	66.8	65.5
5	**RhGB06**	59.3	76.8	95.5	97.7		97.3	96.8	96.8	67.9	66.2	67.6	65.3
6	**RhGB05**	58.8	75.9	95.9	97.7	97.3		97.3	96.8	67.4	66.2	67.1	65.3
7	**RhGB02**	59.3	76.8	95	97.2	96.8	97.3		97.7	68.3	67.1	67.6	66.2
8	BatCoV RhGB01	58.8	75.9	94.5	96.8	96.8	96.8	97.7		68.3	66.7	66.7	65.8
9	SARS-CoV	57.1	71.7	67.9	68.5	67.9	67.4	68.3	68.3		74.9	73.1	74.9
10	Pangolin CoV PCoV GX-P4	52.2	67.9	66.7	65.9	66.2	66.2	67.1	66.7	74.9		86.5	87.9
11	Wuhan-Hu-1 SARS-CoV-2	54.5	70.1	67.6	66.8	67.6	67.1	67.6	66.7	73.1	86.5		90.1
12	Bat CoV RaTG13	53.1	68.3	66.2	65.5	65.3	65.3	66.2	65.8	74.9	87.9	90.1	

Key
>80 — 100%
>70 — 80%
>50 — 70%

Figure 4 British bat Sarbecovirus (in bold text) spike glycoprotein organization and sites of interest. (A) The presence of conserved S2 or TMPRSS2 cleavage site (top right) and the absence of furin cleavage site (S1/S2 top left). The FCS consisted of four amino acids (PRRA) identified to be present in only SARS-CoV-2. Below these two is the British bat CoVs spike glycoprotein linear map. (B) Estimated percentage pairwise alignment identity heatmap following alignment of RBD proteins from sarbecovirus reference genomes and those reported in this study.

Comparative homology modelling of this study's sarbecoviruses to examine the interaction between RBD and the human ACE2, yielded five UK bat CoV 3D models sharing ≥ 92% identity and 3.0Å root-mean-square deviation (RMSD) of C-atoms with the SARS-CoV-2 (PDB: 6vw1) X-ray crystal template structure (figure 5A and B). The overall quality factor of 3D models was calculated through validation with ERRAT2 and Ramachandran plots (Supplementary information) as ≥ 96 for all RBD and hACE2 3D structures constructed. Amino acid residues rejected at 95% and 99% confidence intervals and those known to be characteristic for the interaction between SARS-CoV-2 and hACE2 are highlighted along the RBD proteins in Figure 5. Ramachandran plots generated following ERRAT2 validation demonstrated that 91% of residues were in the most favoured regions, 8.7% in additional allowed regions and 0.3% were found to be in the generously allowed regions (Supplementary information).

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

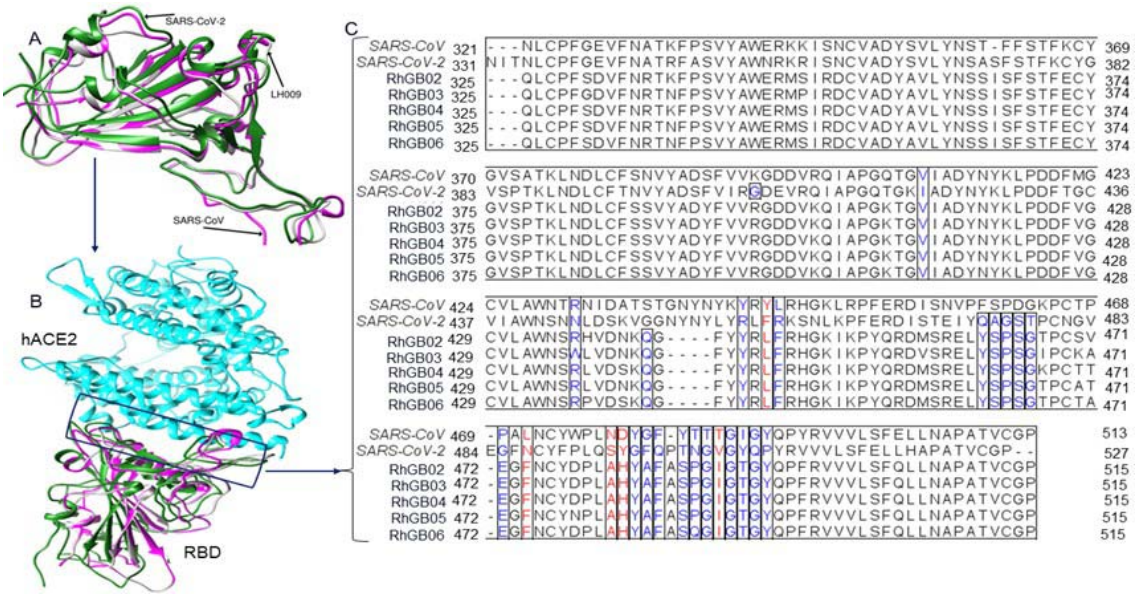


Figure 5 Protein-protein homology complex model of British bat and related sarbecoviruses (A) Superimposed RBD 3D complexes from UK bat CoVs (gray LH009), SARS-CoV (magenta), and SARS-CoV-2 (green). (B) Superimposed model between RBDs of UK bat CoVs (gray), SARS-CoV (magenta) and SARS-CoV-2 (green) and hACE2 (cyan). (C) Sequence alignment and comparison of S1 RBD and RBM of UK bat CoVs, SARS-CoV, and SARS-CoV-2. RBM for UK bat CoVs is positioned between 426 aa to 496 aa. Amino acid residues predicted to interact with hACE2 are shown in blue text, amino acid residues in red text are the critical residues within the RBM previously reported to play key role in cross-species transmission.

Prediction of amino acid residues using Prodigy showed that amino acid residues at interface with hACE2 receptor in British bat Sarbecovirus spike glycoproteins can be located at positions V409, R431, Q437, Y441, L443, F444, Y461, S462, P463, S464, G465, E471, F473, Y482, A483, F484, S486, P487, G488, I489, G490, T491, and Y493; as reported previously (Crook, Murphy et al. 2021). RBM of these sarbecoviruses is located between amino acid position 426 aa to 496 aa of the spike protein, coinciding with the location of hACE2 amino acid residues. In addition, RBD comparative analysis revealed the presence of critical amino acid residues in British bat sarbecoviruses (L443, L443, F473, A480, H481 and I489) previously reported to play a key role in cross-species transmission of SARS-like CoVs (Figure 5).

Discussion

This study found no detectable SARS-CoV-2 in British bats, in particular none in horseshoe bats which might be expected to be most at risk of infection. This is similar to the findings of other studies of European bats during the COVID-19 pandemic, none of which demonstrated SARS-CoV-2 infection including from, three common pipistrelles in the UK (Jones, Bell et al. 2022), 503 samples from 20 bat species including 58 lesser horseshoe bats in Poland (Orłowska, Smreczak et al. 2022), 197 samples from five bat species including 82 samples from lesser horseshoe bats, 104 from greater horseshoe bats and five from Mediterranean horseshoe bats (*Rhinolophus euyale)* from Sochi in Russia (Alkhovsky, Lenshin et al. 2022) and 76 lesser horseshoe bats in the UK (Crook, Murphy et al. 2021).

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The pan-coronavirus screening assay used in this study (Woo, Lau et al. 2005) is known to be insensitive and may miss some coronavirus strains due to sequence mismatch for example, it is not a good match for the known UK isolate MW719567, RhGB01. The Sarbecovirus E gene qPCR (Corman, Landt et al. 2020) was designed to detect all known sarbecoviruses and is in common use in human SARS-CoV-2 diagnostics due to the relative stability of the E gene in SARS isolates and sarbecoviruses in general. Serial dilution of the SARS-CoV-2 positive control in this study (data not shown) found that the E gene qPCR was approximately a 10^2 times more sensitive than the pan coronavirus screening assay, though SARS-CoV-2 is routinely detected by both assays. This difference in sensitivity of the two assays may explain the discrepancies in results between the two screens (CT values for the E gene qPCR ranged between 22 and 40). It is therefore possible that other coronaviruses, in particular non-sarbecoviruses are present in these sample but were not detected.

British lesser horseshoe bats were, however, found to be frequently infected with a Sarbecovirus similar to that described previously in this species (Rihtaric, Hostnik et al. 2010, Crook, Murphy et al. 2021, Alkhovsky, Lenshin et al. 2022, Orłowska, Smreczak et al. 2022, Sander, Moreira-Soto et al. 2022) and distinct from sarbecoviruses previously described in greater horseshoe bats in Bulgaria (Muth, Corman et al. 2018) and Russia (Alkhovsky, Lenshin et al. 2022). Although the sampling strategy in this study, based on opportunistic sampling linked to bat survey and conservation studies, did not allow a prevalence to be calculated, the frequency at which the virus was detected within and between populations suggests a prevalence not dissimilar to that found in Poland of around 30% (Orłowska, Smreczak et al. 2022). That study, unlike this, detected virus in oronasal swabs as well as faeces, and this may reflect the smaller amount of material collected on swabs in this study (rectal swabs were also less frequently positive in the qPCR assay than faecal samples). The sites sampled reflected the distribution of horseshoe bats in Great Britain, and infection was clearly common within and between populations of lesser horseshoe bats in this study, with no clear sex or age differences in likelihood of shedding. This high frequency of shedding and widespread distribution geographically and demographically, likely indicates either persistent infection and excretion, or frequent reinfection, of the gastrointestinal tract as has been reported for many other coronaviruses, such as those of cats, chickens and pigs (Quinteros, Noormohammadi et al. 2022, Thayer, Gogolski et al. 2022, Zhang, Chen et al. 2022).

Many of the roosts of lesser horseshoes bats in this study were shared with greater horseshoe bats, yet none of the latter were shedding detectable virus. This suggests lack of cross-species transmission and that the virus is relatively host-species restricted. This contrasts with several studies of SE Asian horseshoe bat species and their sarbecoviruses (Latinne, Hu et al. 2020, Cappelle, Furey et al. 2021, Zhou, Ji et al. 2021), which found strong evidence of cross-species transmission of coronaviruses. Lesser horseshoe bats were primarily sampled in the summer due to the greater sensitivity of this species to disturbance, with greater horseshoe bats sampled across both summer and winter (53/277). The numbers of greater horseshoe bats in this study should have been adequate to detect, if it were present, the virus previously detected in Italian, Bulgarian and Russian greater horseshoe bats at the prevalence rates of 26-42% recorded in those studies (Drexler, Gloza-Rausch et al. 2010, Balboni, Palladini et al. 2011, Balboni, Gallina et al. 2012, Muth, Corman et al. 2018, Alkhovsky, Lenshin et al. 2022) and the assay should be able to detect the Russian isolate (the only one for which an E gene is available). It is not clear why no sarbecoviruses were detected in these animals, though annual variation in viral loads is possible as is a “founder effect” in an island population on the edge of the host’s geographic range.

The apparent host specificity of the European lesser horseshoe bat Sarbecovirus, combined with its lack of the furin cleavage site thought to be critical for human spread of SARS-Cov-2, indicate that

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

these viruses are likely of low potential for zoonotic transmission, although modelling studies (Sander, Moreira-Soto et al. 2022) have indicated that these viruses could acquire such features with minimal mutation and that these viruses could potentially bind to the human ACE-2 receptor (Crook, Murphy et al. 2021).

The sarbecoviruses found in European lesser horseshoe bats cluster monophyletically and probably represent a distinct genus of Sarbecovirus. Recombination analysis with RDP5 (data not shown) indicates that these viruses are not recombinants.

Some variation in the presence of the accessory genes 7b, 9b and 10 was found amongst the isolates studied. These genes are not essential for viral replication, but are modulators of the host's innate immune system response and as such can affect strain pathogenicity (Fang, Fang et al. 2021, Forni, Cagliani et al. 2022, McGrath, Xue et al. 2022) . Plasticity in accessory gene complement has also been evident in different SARs-CoV-2 isolates and appears to be a feature of Sarbecovirus isolate variability (Forni, Cagliani et al. 2022, McGrath, Xue et al. 2022).

We do not have any indication if infection with these viruses have any adverse affects on their hosts or whether excretion patterns vary with age or reproductive status. This type of work is mature in relation to other bat-borne viruses, for example Hendra virus (a member of the Paramyxoviridae family) in bats of the genus Pteropus (flying foxes), for which it is clear that virus excretion is associated with maternity roosts with large numbers of birthing and juvenile animals and that virus excretion peaks in times of nutritional stress (Eby, Peel et al. 2022). Such work is however in its infancy with horseshoe bat sarbecoviruses, although some longitudinal studies in SE Asia hint at a summer/maternity roost excretion pattern (Cappelle, Furey et al. 2021).

A better understanding of the ecology of horseshoe bat viruses requires further and longer term studies of sarbecoviruses in their natural hosts. The samples with the highest detection rates in this study were pooled faecal samples from bat roosts so this sample type may provide the most reliable method of detection of virus in a roost site as well as being the most convenient, causing least disruption to bat colonies and presenting least threat of cross-species transmission between bats or between bats and humans.

Overall, this study provides several critical pieces of data in the overall picture of sarbecoviruses in their natural hosts. The growing picture is one of relative stability in terms of viral diversity and cross-species transmission potential in European horseshoe bat species, in contrast to higher diversity and cross species sharing of isolates in SE Asia, potentially related to the greater diversity and number of horseshoe bat species in the latter geographic region. This likely partly explains the repeated cross-species spill over of sarbecoviruses with an ancestral origin in bats, into civets, pangolins and humans in SE Asia, though human behaviour and wildlife/human interactions, habitat disruption and farming of wildlife for meat are also likely to be contributing factors (Holmes, Goldstein et al. 2021). Ongoing work with these European sarbecoviruses remains to determine the epidemiology, temporal changes in viral prevalence and sequence and effects of the virus in its natural host with a real need for longitudinal studies of roost sites across the host range and spatial comparisons of sequences and prevalence.

Conflicts of interest

The authors declare that there are no conflicts of interest

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Funding information

This work was funded by the BBSRC Grant number: BB/W009501/1

Ethical approval

Ethical approval was granted by the University of Nottingham School of Veterinary Medicine and Science Committee for Animal Research and Ethics (CARE) and the Animal Welfare and Ethical Review Board of the University of Sussex.

Acknowledgements

We are grateful for the help of many licensed bat workers who assisted in the collection of samples for this project, particularly Devon Bat Research and Conservation Group, Wiltshire Bat Group, and the Vincent Wildlife Trust. We also appreciate constructive discussions with the Bat Conservation Trust that informed this project.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

Alkhovsky, S., S. Lenshin, A. Romashin, T. Vishnevskaya, O. Vyshemirsky, Y. Bulycheva, D. Lvov and A. Gitelman (2022). "SARS-like Coronaviruses in Horseshoe Bats (Rhinolophus spp.) in Russia, 2020." Viruses **14**(1).

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." Journal of molecular biology **215**(3): 403-410.

Balboni, A., L. Gallina, A. Palladini, S. Prosperi and M. Battilani (2012). "A real-time PCR assay for bat SARS-like coronavirus detection and its application to Italian greater horseshoe bat faecal sample surveys." ScientificWorldJournal **2012**: 989514.

Balboni, A., A. Palladini, G. Bogliani and M. Battilani (2011). "Detection of a virus related to betacoronaviruses in Italian greater horseshoe bats." Epidemiol Infect **139**(2): 216-219.

Bienzle, D., J. Rousseau, D. Marom, J. MacNicol, L. Jacobson, S. Sparling, N. Prystajecky, E. Fraser and J. S. Weese (2022). "Risk Factors for SARS-CoV-2 Infection and Illness in Cats and Dogs(1)." Emerg Infect Dis **28**(6): 1154-1162.

Boni, M. F., P. Lemey, X. Jiang, T. T. Lam, B. W. Perry, T. A. Castoe, A. Rambaut and D. L. Robertson (2020). "Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic." Nat Microbiol **5**(11): 1408-1417.

Cappelle, J., N. Furey, T. Hoem, T. P. Ou, T. Lim, V. Hul, O. Heng, V. Chevalier, P. Dussart and V. Duong (2021). "Longitudinal monitoring in Cambodia suggests higher circulation of alpha and betacoronaviruses in juvenile and immature bats of three species." Sci Rep **11**(1): 24145.

Chan, Y. A. and S. H. Zhan (2022). "The Emergence of the Spike Furin Cleavage Site in SARS-CoV-2." Mol Biol Evol **39**(1).

Chen, S., Y. Zhou, Y. Chen and J. Gu (2018). "fastp: an ultra-fast all-in-one FASTQ preprocessor." Bioinformatics **34**(17): i884-i890.

Common, S. M., T. Shadbolt, K. Walsh and A. W. Sainsbury (2022). "The risk from SARS-CoV-2 to bat species in england and mitigation options for conservation field workers." Transbound Emerg Dis **69**(2): 694-705.

Conceicao, C., N. Thakur, S. Human, J. T. Kelly, L. Logan, D. Bialy, S. Bhat, P. Stevenson-Leggett, A. K. Zagrajek, P. Hollinghurst, M. Varga, C. Tsirigoti, M. Tully, C. Chiu, K. Moffat, A. P. Silesian, J. A. Hammond, H. J. Maier, E. Bickerton, H. Shelton, I. Dietrich, S. C. Graham and D. Bailey (2020). "The SARS-CoV-2 Spike protein has a broad tropism for mammalian ACE2 proteins." PLoS Biol **18**(12): e3001016.

Cook, J. D., E. H. C. Grant, J. T. H. Coleman, J. M. Sleeman and M. C. Runge (2021). "Risks posed by SARS-CoV-2 to North American bats during winter fieldwork." Conserv Sci Pract **3**(6): e410.

Corman, V. M., O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D. K. Chu, T. Bleicker, S. Brunink, J. Schneider, M. L. Schmidt, D. G. Mulders, B. L. Haagmans, B. van der Veer, S. van den Brink, L. Wijsman, G. Goderski, J. L. Romette, J. Ellis, M. Zambon, M. Peiris, H. Goossens, C. Reusken, M. P. Koopmans and C. Drosten (2020). "Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR." Euro Surveill **25**(3).

Crook, J. M., I. Murphy, D. P. Carter, S. T. Pullan, M. Carroll, R. Vipond, A. A. Cunningham and D. Bell (2021). "Metagenomic identification of a new sarbecovirus from horseshoe bats in Europe." Sci Rep **11**(1): 14723.

Danecek, P., J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies and H. Li (2021). "Twelve years of SAMtools and BCFtools." Gigascience **10**(2).

Drexler, J. F., F. Gloza-Rausch, J. Glende, V. M. Corman, D. Muth, M. Goettsche, A. Seebens, M. Niedrig, S. Pfefferle, S. Yordanov, L. Zhelyazkov, U. Hermanns, P. Vallo, A. Lukashev, M. A. Müller, H. Deng, G. Herrler and C. Drosten (2010). "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 **84**(21): 11336-11349.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Eby, P., A. J. Peel, A. Hoegh, W. Madden, J. R. Giles, P. J. Hudson and R. K. Plowright (2022). "Pathogen spillover driven by rapid changes in bat ecology." Nature.
Eckstrand, C. D., T. J. Baldwin, K. A. Rood, M. J. Clayton, J. K. Lott, R. M. Wolking, D. S. Bradway and T. Baszler (2021). "An outbreak of SARS-CoV-2 with high mortality in mink (Neovison vison) on multiple Utah farms." PLoS Pathog **17**(11): e1009952.
Fang, P., L. Fang, H. Zhang, S. Xia and S. Xiao (2021). "Functions of Coronavirus Accessory Proteins: Overview of the State of the Art." Viruses **13**(6).
Forni, D., R. Cagliani, C. Molteni, F. Arrigoni, A. Mozzi, M. Clerici, L. De Gioia and M. Sironi (2022). "Homology-based classification of accessory proteins in coronavirus genomes uncovers extremely dynamic evolution of gene content." Mol Ecol **31**(13): 3672-3692.
Goletic, S., T. Goletic, A. Softic, A. Zahirovic, D. Rukavina, A. Kavazovic, J. Omeragic, S. Umihanic and M. Hukic (2022). "The Evidence of SARS-CoV-2 Human-to-Pets Transmission in Household Settings in Bosnia and Herzegovina." Front Genet **13**: 839205.
Hale, V. L., P. M. Dennis, D. S. McBride, J. M. Nolting, C. Madden, D. Huey, M. Ehrlich, J. Grieser, J. Winston, D. Lombardi, S. Gibson, L. Saif, M. L. Killian, K. Lantz, R. M. Tell, M. Torchetti, S. Robbe-Austerman, M. I. Nelson, S. A. Faith and A. S. Bowman (2022). "SARS-CoV-2 infection in free-ranging white-tailed deer." Nature **602**(7897): 481-486.
Hoang, D. T., O. Chernomor, A. Von Haeseler, B. Q. Minh and L. S. Vinh (2018). "UFBoot2: improving the ultrafast bootstrap approximation." Molecular biology and evolution **35**(2): 518-522.
Holmes, E. C., S. A. Goldstein, A. L. Rasmussen, D. L. Robertson, A. Crits-Christoph, J. O. Wertheim, S. J. Anthony, W. S. Barclay, M. F. Boni, P. C. Doherty, J. Farrar, J. L. Geoghegan, X. Jiang, J. L. Leibowitz, S. J. D. Neil, T. Skern, S. R. Weiss, M. Worobey, K. G. Andersen, R. F. Garry and A. Rambaut (2021). "The origins of SARS-CoV-2: A critical review." Cell **184**(19): 4848-4856.
Honorato, R. V., P. I. Koukos, B. Jimenez-Garcia, A. Tsaregorodtsev, M. Verlato, A. Giachetti, A. Rosato and A. Bonvin (2021). "Structural Biology in the Clouds: The WeNMR-EOSC Ecosystem." Front Mol Biosci **8**: 729513.
Hosie, M. J., R. Hofmann-Lehmann, K. Hartmann, H. Egberink, U. Truyen, D. D. Addie, S. Belák, C. Boucraut-Baralon, T. Frymus, A. Lloret, H. Lutz, F. Marsilio, M. G. Pennisi, S. Tasker, E. Thiry and K. Möstl (2021). "Anthropogenic Infection of Cats during the 2020 COVID-19 Pandemic." Viruses **13**(2).
Hu, B., L. P. Zeng, X. L. Yang, X. Y. Ge, W. Zhang, B. Li, J. Z. Xie, X. R. Shen, Y. Z. Zhang, N. Wang, D. S. Luo, X. S. Zheng, M. N. Wang, P. Daszak, L. F. Wang, J. Cui and Z. L. Shi (2017). "Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus." PLoS Pathog **13**(11): e1006698.
Jairak, W., E. Chamsai, K. Udom, K. Charoenkul, S. Chaiyawong, N. Techakriengkrai, R. Tangwangvivat, K. Suwannakarn and A. Amonsin (2022). "SARS-CoV-2 delta variant infection in domestic dogs and cats, Thailand." Sci Rep **12**(1): 8403.
JNCC. (2022). "1303 Lesser horseshoe bat Rhinolophus hipposideros." Species list, 2022, from https://sac.jncc.gov.uk/species/S1303/.
JNCC. (2022). "1304 Greater horseshoe bat Rhinolophus ferrumequinum." Species list, 2022, from https://sac.jncc.gov.uk/species/S1304/.
Jones, S., T. Bell, C. M. Coleman, D. Harris, G. Woodward, L. Worledge, H. Roberts, L. McElhinney, J. Aegerter, E. Ransome and V. Savolainen (2022). "Testing bats in rehabilitation for SARS-CoV-2 before release into the wild." Conserv Sci Pract **4**(7): e12707.
Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler and L. S. Jermiin (2017). "ModelFinder: fast model selection for accurate phylogenetic estimates." Nat Methods **14**(6): 587-589.
Katoh, K. and D. M. Standley (2013). "MAFFT multiple sequence alignment software version 7: improvements in performance and usability." Mol Biol Evol **30**(4): 772-780.
Larsen, H. D., J. Fonager, F. K. Lomholt, T. Dalby, G. Benedetti, B. Kristensen, T. R. Urth, M. Rasmussen, R. Lassaunière, T. B. Rasmussen, B. Strandbygaard, L. Lohse, M. Chaine, K. L. Møller, A. N. Berthelsen, S. K. Nørgaard, U. W. Sönksen, A. E. Boklund, A. S. Hammer, G. J. Belsham, T. G.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Krause, S. Mortensen, A. Bøtner, A. Fomsgaard and K. Mølbak (2021). "Preliminary report of an outbreak of SARS-CoV-2 in mink and mink farmers associated with community spread, Denmark, June to November 2020." Euro Surveill **26**(5).
Laskowski, R. A., M. W. MacArthur, D. S. Moss and J. M. Thornton (1993). "PROCHECK: a program to check the stereochemical quality of protein structures." Journal of Applied Crystallography **26**(2): 283-291.
Latinne, A., B. Hu, K. J. Olival, G. Zhu, L. Zhang, H. Li, A. A. Chmura, H. E. Field, C. Zambrana-Torrelio, J. H. Epstein, B. Li, W. Zhang, L. F. Wang, Z. L. Shi and P. Daszak (2020). "Origin and cross-species transmission of bat coronaviruses in China." Nat Commun **11**(1): 4235.
Lau, S. K., P. C. Woo, K. S. Li, Y. Huang, H. W. Tsoi, B. H. Wong, S. S. Wong, S. Y. Leung, K. H. Chan and K. Y. Yuen (2005). "Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats." Proc Natl Acad Sci U S A **102**(39): 14040-14045.
Li, H. (2018). "Minimap2: pairwise alignment for nucleotide sequences." Bioinformatics **34**(18): 3094-3100.
Li, J., S. Lai, G. F. Gao and W. Shi (2021). "The emergence, genomic diversity and global spread of SARS-CoV-2." Nature **600**(7889): 408-418.
Lin, X. D., W. Wang, Z. Y. Hao, Z. X. Wang, W. P. Guo, X. Q. Guan, M. R. Wang, H. W. Wang, R. H. Zhou, M. H. Li, G. P. Tang, J. Wu, E. C. Holmes and Y. Z. Zhang (2017). "Extensive diversity of coronaviruses in bats from China." Virology **507**: 1-10.
Logeot, M., A. Mauroy, E. Thiry, N. De Regge, M. Vervaeke, O. Beck, V. De Waele and T. Van den Berg (2022). "Risk assessment of SARS-CoV-2 infection in free-ranging wild animals in Belgium." Transbound Emerg Dis **69**(3): 986-996.
Lopez-Oceja, A., D. Gamarra, S. Borragan, S. Jimenez-Moreno and M. M. de Pancorbo (2016). "New cyt b gene universal primer set for forensic analysis." Forensic Sci Int Genet **23**: 159-165.
Lu, L., R. S. Sikkema, F. C. Velkers, D. F. Nieuwenhuijse, E. A. J. Fischer, P. A. Meijer, N. Bouwmeester-Vincken, A. Rietveld, M. C. A. Wegdam-Blans, P. Tolsma, M. Koppelman, L. A. M. Smit, R. W. Hakze-van der Honing, W. H. M. van der Poel, A. N. van der Spek, M. A. H. Spierenburg, R. J. Molenaar, J. Rond, M. Augustijn, M. Woolhouse, J. A. Stegeman, S. Lycett, B. B. Oude Munnink and M. P. G. Koopmans (2021). "Adaptation, spread and transmission of SARS-CoV-2 in farmed minks and associated humans in the Netherlands." Nat Commun **12**(1): 6802.
Lytras, S., J. Hughes, D. Martin, P. Swanepoel, A. de Klerk, R. Lourens, S. L. Kosakovsky Pond, W. Xia, X. Jiang and D. L. Robertson (2022). "Exploring the Natural Origins of SARS-CoV-2 in the Light of Recombination." Genome Biol Evol **14**(2).
Mathews, F., L. M. Kubasiewicz, J. Gurnell, C. A. Harrower, R. A. McDonald and R. F. Shore (2018). A review of the population and conservation status of British mammals. A report by the Mammal Society under contract to Natural England, Natural Resources Wales and Scottish Natural Heritage. Natural England Joint Report. **JP025:** pp701.
McGrath, M. E., Y. Xue, C. Dillen, L. Oldfield, N. Assad-Garcia, J. Zaveri, N. Singh, L. Baracco, L. J. Taylor, S. Vashee and M. B. Frieman (2022). "SARS-CoV-2 variant spike and accessory gene mutations alter pathogenesis." Proc Natl Acad Sci U S A **119**(37): e2204717119.
Meleshko, D., I. Hajirasouliha and A. Korobeynikov (2021). "coronaSPAdes: from biosynthetic gene clusters to RNA viral assemblies." Bioinformatics.
Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler and R. Lanfear (2020). "IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era." Mol Biol Evol **37**(5): 1530-1534.
Muth, D., V. M. Corman, H. Roth, T. Binger, R. Dijkman, L. T. Gottula, F. Gloza-Rausch, A. Balboni, M. Battilani, D. Rihtarič, I. Toplak, R. S. Ameneiros, A. Pfeifer, V. Thiel, J. F. Drexler, M. A. Müller and C. Drosten (2018). "Attenuation of replication by a 29 nucleotide deletion in SARS-coronavirus acquired during the early stages of human-to-human transmission." Sci Rep **8**(1): 15177.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Muylaert, R. L., T. Kingston, J. Luo, M. H. Vancine, N. Galli, C. J. Carlson, R. S. John, M. C. Rulli and D. T. S. Hayman (2022). "Present and future distribution of bat hosts of sarbecoviruses: implications for conservation and public health." Proc Biol Sci **289**(1975): 20220397.
Orłowska, A., M. Smreczak, K. Thor, M. Niedbalska, D. Pawelec, P. Trębas and J. Rola (2022). "The Genetic Characterization of the First Detected Bat Coronaviruses in Poland Revealed SARS-Related Types and Alphacoronaviruses." Viruses **14**(9).
Pappas, G., D. Vokou, I. Sainis and J. M. Halley (2022). "SARS-CoV-2 as a Zooanthroponotic Infection: Spillbacks, Secondary Spillovers, and Their Importance." Microorganisms **10**(11).
Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin (2004). "UCSF Chimera--a visualization system for exploratory research and analysis." J Comput Chem **25**(13): 1605-1612.
Pickering, B., O. Lung, F. Maguire, P. Kruczkiewicz, J. D. Kotwa, T. Buchanan, M. Gagnier, J. L. Guthrie, C. M. Jardine, A. Marchand-Austin, A. Massé, H. McClinchey, K. Nirmalarajah, P. Aftanas, J. Blais-Savoie, H. Y. Chee, E. Chien, W. Yim, A. Banete, B. D. Griffin, L. Yip, M. Goolia, M. Suderman, M. Pinette, G. Smith, D. Sullivan, J. Rudar, O. Vernygora, E. Adey, M. Nebroski, G. Goyette, A. Finzi, G. Laroche, A. Ariana, B. Vahkal, M. Côté, A. J. McGeer, L. Nituch, S. Mubareka and J. Bowman (2022). "Divergent SARS-CoV-2 variant emerges in white-tailed deer with deer-to-human transmission." Nat Microbiol.
Procter, J. B., G. M. Carstairs, B. Soares, K. Mourao, T. C. Ofoegbu, D. Barton, L. Lui, A. Menard, N. Sherstnev, D. Roldan-Martinez, S. Duce, D. M. A. Martin and G. J. Barton (2021). "Alignment of Biological Sequences with Jalview." Methods Mol Biol **2231**: 203-224.
Quinteros, J. A., A. H. Noormohammadi, S. W. Lee, G. F. Browning and A. Diaz-Méndez (2022). "Genomics and pathogenesis of the avian coronavirus infectious bronchitis virus." Aust Vet J **100**(10): 496-512.
Rabalski, L., M. Kosinski, T. Smura, K. Aaltonen, R. Kant, T. Sironen, B. Szewczyk and M. Grzybek (2021). "Severe Acute Respiratory Syndrome Coronavirus 2 in Farmed Mink (Neovison vison), Poland." Emerg Infect Dis **27**(9): 2333-2339.
Rihtaric, D., P. Hostnik, A. Steyer, J. Grom and I. Toplak (2010). "Identification of SARS-like coronaviruses in horseshoe bats (Rhinolophus hipposideros) in Slovenia." Arch Virol **155**(4): 507-514.
Russo, D. (2022). Chiroptera, Springer.
Sander, A. L., A. Moreira-Soto, S. Yordanov, I. Toplak, A. Balboni, R. S. Ameneiros, V. Corman, C. Drosten and J. F. Drexler (2022). "Genomic determinants of Furin cleavage in diverse European SARS-related bat coronaviruses." Commun Biol **5**(1): 491.
Seifert, S. N., S. Bai, S. Fawcett, E. B. Norton, K. J. Zwezdaryk, J. Robinson, B. Gunn and M. Letko (2022). "An ACE2-dependent Sarbecovirus in Russian bats is resistant to SARS-CoV-2 vaccines." PLoS Pathog **18**(9): e1010828.
Sila, T., J. Sunghan, W. Laochareonsuk, S. Surasombatpattana, C. Kongkamol, T. Ingviya, P. Siripaitoon, N. Kositpantawong, S. Kanchanasuwan, T. Hortiwakul, B. Charernmak, O. F. Nwabor, K. Silpapojakul and S. Chusri (2022). "Suspected Cat-to-Human Transmission of SARS-CoV-2, Thailand, July-September 2021." Emerg Infect Dis **28**(7): 1485-1488.
Temple, H. J. and A. Terry (2007). The Status and Distribution of European Mammals. Luxembourg: , Office for Official Publications of the European Communities.
Thayer, V., S. Gogolski, S. Felten, K. Hartmann, M. Kennedy and G. A. Olah (2022). "2022 AAFP/EveryCat Feline Infectious Peritonitis Diagnosis Guidelines." J Feline Med Surg **24**(9): 905-933.
UKHSA (2022). SARSCoV2 variants of concern and variants under investigation in England Technical briefing 8 April 2022 report 40
Vangone, A. and A. M. Bonvin (2015). "Contacts-based prediction of binding affinity in protein-protein complexes." Elife **4**: e07454.
Wang, J., Y. Pan, L. Yang, W. Yang, C. Luo, J. Wang, G. Kuang, W. Wu, Q. Gou, G. Xin, B. Li, H. Luo, Y. Chen, Y. Shu, D. Guo, Z. Gao, G. Liang, J. Li, E. C. Holmes, Y. Feng and M. Shi (2022). "Individual bat

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.14.528476; this version posted February 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

viromes reveal the co-infection, spillover and emergence risk of potential zoonotic viruses." bioRxiv: 2022.2011.2023.517609.

Wang, M., M. Yan, H. Xu, W. Liang, B. Kan, B. Zheng, H. Chen, H. Zheng, Y. Xu, E. Zhang, H. Wang, J. Ye, G. Li, M. Li, Z. Cui, Y. F. Liu, R. T. Guo, X. N. Liu, L. H. Zhan, D. H. Zhou, A. Zhao, R. Hai, D. Yu, Y. Guan and J. Xu (2005). "SARS-CoV infection in a restaurant from palm civet." Emerg Infect Dis **11**(12): 1860-1865.

Willgert, K., X. Didelot, M. Surendran-Nair, S. V. Kuchipudi, R. M. Ruden, M. Yon, R. H. Nissly, K. J. Vandegrift, R. K. Nelli, L. Li, B. M. Jayarao, N. Levine, R. J. Olsen, J. J. Davis, J. M. Musser, P. J. Hudson, V. Kapur and A. J. K. Conlan (2022). "Transmission history of SARS-CoV-2 in humans and white-tailed deer." Sci Rep **12**(1): 12094.

Wong, A. C. P., X. Li, S. K. P. Lau and P. C. Y. Woo (2019). "Global Epidemiology of Bat Coronaviruses." Viruses **11**(2).

Woo, P. C., S. K. Lau, C. M. Chu, K. H. Chan, H. W. Tsoi, Y. Huang, B. H. Wong, R. W. Poon, J. J. Cai, W. K. Luk, L. L. Poon, S. S. Wong, Y. Guan, J. S. Peiris and K. Y. Yuen (2005). "Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia." J Virol **79**(2): 884-895.

Wood, D. E., J. Lu and B. Langmead (2019). "Improved metagenomic analysis with Kraken 2." Genome Biol **20**(1): 257.

Wu, Z., Y. Han, Y. Wang, B. Liu, L. Zhao, J. Zhang, H. Su, W. Zhao, L. Liu, S. Bai, J. Dong, L. Sun, Y. Zhu, S. Zhou, Y. Song, H. Sui, J. Yang, J. Wang, S. Zhang, Z. Qian and Q. Jin (2022). "A comprehensive survey of bat sarbecoviruses across China in relation to the origins of SARS-CoV and SARS-CoV-2." National Science Review.

Zhang, Y., Y. Chen, J. Zhou, X. Wang, L. Ma, J. Li, L. Yang, H. Yuan, D. Pang and H. Ouyang (2022). "Porcine Epidemic Diarrhea Virus: An Updated Overview of Virus Epidemiology, Virulence Variation Patterns and Virus-Host Interactions." Viruses **14**(11).

Zhou, H., J. Ji, X. Chen, Y. Bi, J. Li, Q. Wang, T. Hu, H. Song, R. Zhao, Y. Chen, M. Cui, Y. Zhang, A. C. Hughes, E. C. Holmes and W. Shi (2021). "Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses." Cell **184**(17): 4380-4391.e4314.

i