

p

# **Identification and genetic characterization of MERS-related coronavirus isolated from Nathusius' pipistrelle (*Pipistrellus nathusii*) near Zvenigorod (Moscow region, Russia)**

**Speranskaya A.S.<sup>1,2,3\*</sup>, Artyushin I.V.<sup>3</sup>, Samoilov A.E.<sup>1,4</sup>, Korneenko E.V.<sup>1</sup>, Khabudaev K.<sup>1</sup>, Ilina E.N.<sup>1</sup>, Yusefovich A.P.<sup>3</sup>, Safonova M.V.<sup>5</sup>, Dolgova A.S.<sup>4</sup>, Gladkikh, A.S.<sup>4</sup>, Dedkov V.G.<sup>4,6</sup>, Daszak Peter<sup>7</sup>**

*<sup>1</sup>Scientific Research Institute for Systems Biology and Medicine, Federal Service on Consumers' Rights Protection and Human Well-Being Surveillance, Moscow, Russia;*

*<sup>2</sup>Department of Molecular Diagnostics and Epidemiology, Central Research Institute for Epidemiology, Federal Service on Consumers' Rights Protection and Human Well-Being Surveillance, Moscow, Russia;*

*<sup>3</sup>Lomonosov Moscow State University, Biological Department, Moscow, Russia;*

*<sup>4</sup>Saint-Petersburg Pasteur Institute, Federal Service on Consumers' Rights Protection and Human Well-Being Surveillance, Saint-Petersburg, Russia;*

*<sup>5</sup>Department of Particularly Dangerous Diseases, Anti-Plague Center, Federal Service on Consumers' Rights Protection and Human Well-Being Surveillance, Moscow, Russia;*

*<sup>6</sup>Martsinovskiy Institute of Medical Parasitology, Tropical and Vector Borne Diseases, Sechenov First Moscow State Medical University, Moscow, Russia;*

*<sup>7</sup>EcoHealth Alliance, New York, USA*

**Keywords:** *Bat-CoV, MERS-related coronaviruses, Pipistrellus nathusii, bats, hedgehogs, humans, camels, DPP4, Spike protein, molecular docking*

***\*hanna.s.939@gmail.com***

## Abstract

The majority of emerging infectious diseases are caused by pathogens with zoonotic origin, and most of these emerged from wildlife reservoirs. Bats are diverse, and widely distributed globally, and are the known or hypothesized reservoir of a series of emerging zoonotic viruses. Analyses of bat viromes have been used to identify novel viruses with potential to cause human infection. We characterized the fecal virome of 26 samples collected from six bat species captured during 2015 in Moscow Region. Of these 13/26 (50%) samples were found to be coronavirus positive. We sequenced and assembled the complete genome of a novel MERS-related Betacoronavirus from *Pipistrellus nathusii*, named MOW-BatCoV strain 15-22. Of *P. nathusii* 3/6 samples were found to carriers of MOW-BatCoVs. The genome organization of MOW-BatCoV/15-22 was identical to other known MERS-related coronaviruses. Phylogenetic analysis of whole genomes suggests that MOW-BatCoV/15-22 falls into a distinct subclade closely related to human and camel MERS-CoV, and MERS-related CoVs from the bat species *Hypsugo savii* and *Pipistrellus kuhlii* (from Italy) and *Neoromicia capensis* (from South Africa). Unexpectedly, phylogenetic analysis of the novel MOW-BatCoV 15-22 spike genes showed the closest similarity to a bat CoV *Neoromicia*/5038 and CoVs from *Erinaceus europaeus* (the European hedgehog), thus MOW-BatCoV could arise as result of recombination between ancestral viruses of bats and hedgehogs. Computer molecular docking analysis of MOW-BatCoV 15-22 Spike glycoprotein binding to DPP4 receptors of different mammal species predicted highest binding interaction with DPP4 of the bat *M. brandtii* (docking score -320.15) and the European hedgehog, *E. europaeus* (docking score -294.51). Hedgehogs are widely kept as pets, and are commonly found in areas of human habitation. Our finding of a novel bat-

p

CoV likely able to infect hedgehogs suggests the potential for hedgehogs to act as intermediate hosts for bat-CoVs between bats and humans.

## Introduction

Coronaviruses (CoVs) have been responsible for three high impact outbreaks in the past 2 decades, including severe acute respiratory syndrome (SARS), the Middle East respiratory syndrome (MERS) and the ongoing coronavirus disease 2019 (COVID-19) pandemic. Each of these diseases affects the human respiratory system, causing a spectrum from asymptomatic and mild respiratory illness to severe pneumonia, acute respiratory failure and death.

The current COVID-19 pandemic is now in its third year and continues to be a global health emergency, with more than 500 million confirmed cases, including more than six million deaths, reported to WHO by the date of submitting this manuscript [1]. The earlier outbreak of MERS was caused by zoonotic virus, MERS-CoV, transmitted to humans from infected dromedary camels, with secondary human-to-human transmission mainly nosocomial [2]. MERS was first identified in Saudi Arabia in 2012 [3] and has now been reported in 27 countries, with the largest outbreaks in Saudi Arabia, United Arab Emirates, and the Republic of Korea leading to 858 known deaths due to the infection and related complications. The disease has a high fatality rate of up to 35% [4,5]. The origin of the virus is not fully understood yet but phylogenetic analysis of different virus genomes suggest it originated in bats and passed to humans in 2012 after circulating endemically in dromedary camels for around 30 years [6,7].

Bats are the known or putative reservoirs of several viruses that can cause severe human disease, including rabies, Hendra, Nipah, Marburg, SARS, MERS and Ebola viruses [8–10]. Bats also carry diverse coronaviruses, some of which are able

p

to bind to human cells *in vitro*, suggesting that bats are the likely reservoirs of potential future zoonotic CoVs [11]. In addition, bats are a diverse group of mammals (representing around 1/5th of all mammalian biodiversity), have wide geographic distribution, long life span and are known to feed and roost near to human communities, suggesting they are an important reservoir of other potential emerging diseases [12].

Two genera of CoVs infect mainly mammals: *Alphacoronavirus* ( $\alpha$ -CoV) and *Betacoronavirus* ( $\beta$ -CoV) [11,13,14]. Phylogenetic analysis shows that  $\beta$ -CoVs are grouped into several clades (from A to D). These four Betacoronavirus lineages have been reclassified into five subgenera: Embecovirus (A), Sarbecovirus (B), Merbecovirus (C), Nobecovirus (D), and an additional lineage Hibecovirus [15],[16]. MERS CoV, which causes fatal pneumonia in humans and has a dromedary camel reservoir, along with MERS-related (MERSr-) CoVs form clade C (or Merbecoviruses) within the *Betacoronaviruses* and includes CoVs discovered in bats and hedgehogs [14–17]. MERSr-CoVs have been reported from South Africa (NeoCoV from *Neoromicia capensis*) [18], Mexico (Mex\_CoV-9 from *Nyctinomops laticaudatus*) [19], Uganda (MERSr-CoV PREDICT/PDF-2180 from *Pipistrellus cf. hesperidus*) [20], Netherlands (NL-VM314 from *Pipistrellus pipistrellus*) [21], Italy (BatCoV-Ita1 strain 206645–40 from *Hypsugo savii* and BatCoV-Ita2 strain 206645–63 from *Pipistrellus kuhlii*) [17] and China (BatCoV/SC2013 from *Vespertilio superans* [22] and strains of HKU4- and HKU5-CoVs from *Tylonycteris* and *Pipistrellus* bats [23–26]. To date, few surveys for CoVs have been conducted in Russia, and no MERSr-CoVs have been reported.

Coronavirus cell tropism and ability to infect hosts is determined primarily by the spike protein which is part of the receptor binding domain (RBD) of the sCoV genome [27–29]. Unlike SARS-CoV and SARS-CoV-2, which bind to Angiotensin-

p

converting enzyme 2 (ACE2), MERS-CoV targets the cell surface receptor Dipeptidyl peptidase 4 (DPP4, also known as CD26). DPP4 is relatively conserved among mammalian species, so that MERS-CoV is capable of infecting a wide range of cell lines derived from humans, non-human primates, bats, swine, horse, rabbit, civet, and camel but not from mice, hamster, dog, ferret, and cat [10,27]. The MERS-CoV spike protein undergoes adaptive evolution when inoculated onto normally non-permissive hamster cells transfected to express DPP4 from different bat species [27,30]. HKU4 MERSr-CoVs from bats in China have RBDs that potentially can bind human DPP4 [31] with even low affinity for human cells suggesting potential to infect humans and adapt to more efficient cell entry [17].

In this paper, we describe a survey of bats in Russia for CoVs, report a novel MERSr-CoV, describe its genome organization and relationship with known coronaviruses.

## **Materials and methods**

### **Sample collection**

In summer 2015 fecal samples were collected from 26 bats of the following species: *Myotis dasycneme* (n=5), *Myotis daubentonii* (n=5), *Myotis brandtii* (n=3), *Nyctalus noctula* (n=4), *Pipistrellus nathusii* (n=6), *Plecotus auritus* (n=2), *Vespertilio murinus* (n=1), inhabiting the Zvenigorodsky district of the Moscow region (Sharapovskoe forestry, coordinates N55.69, E36.70). No bats were killed for this study and all bats were captured in mist nets and released at the site of capture. Bat capture and sampling was conducted by professionally trained staff of the biological department of Lomonosov Moscow State University. Fecal samples, rectal swabs and ectoparasites were collected after capture, and species, sex, reproductive and health status visually determined by trained field biologists. Swab samples were kept in a transport media for transportation and storage of swabs with

p

mucoytic agent (AmpliSens, Russia) at 4°C during transport to the laboratory, and were then stored at −80 °C before processing.

### **RNA extraction and reverse transcription**

RNA was extracted from bat faecal samples using QIAamp Viral RNA Mini Kit (Qiagen, Germany). Carrier RNA was dissolved in buffer AVE and added to buffer AVL according to manufacturer's recommendations before extraction. 140 ul of fecal sample was added to the prepared AVL buffer with carrier RNA–Buffer AVE. Further steps were performed according to the original protocol. RNA was eluted with 60 µL of the AVE Buffer and stored at −70 °C until evaluation. 10 ul of RNA was taken to reverse transcription using Reverta-L (AmpliSens, Russia). Second strand of cDNA was obtained using NEBNext Ultra II Non-Directional RNA Second Strand Module (E6111L, New England Biolabs). In order to increase input concentration 24 ul of first strand product were added to 10 ul of the H2O (milliQ) for the further steps.

### **PCR and NGS screening of faecal samples for CoV RNA**

PCR-screening for CoV RNA was performed using primers targeting Alpha- and Betacoronavirus species: 5'-CTTATGGGTTGGGATTATCC (CoV2A-F) and 5'-TTATAACAGACAACGCCATCATC (CoV2A-R), described in [32–34]. This generated ~400-500-bp amplicons from the RNA-dependent RNA polymerase (RdRp) gene. The following thermal cycling parameters were used: 94°C for 3 min, followed by 10 cycles of 94°C for 20 s, 55°C to 45°C (−1°C per cycle) for 20 s, and 72°C for 30 s; and then 42 cycles of 94°C for 20 s, 45°C for 20 s, and 72°C for 30 s; and finally, 72°C for 3 min [35]. PCR amplification products were analyzed by agarose gel electrophoresis. Positive PCR products were purified using AMPure beads and prepared for high throughput sequencing using the TruSeq protocol for

p

Illumina. Sequencing was performed using Illumina MiSeq system to generate 250-bp paired-end reads. Reads were subjected to analysis via the following pipeline: Reads were filtered using Trimmomatic [36]. Then sequences of PCR primers, as well as simple repeats, were masked; filtered reads with an unmasked region of greater than 30 bp were collected and used for taxonomic analysis by comparing assembled contigs or individual reads to the NCBI non-redundant nucleotide and protein sequence databases using the blastn, blastx or tblastn algorithms as described in [35].

### **Library preparation and high throughput sequencing of viral genome**

Double stranded cDNA was taken to library preparation using NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, England). End prep was performed according to the manufacturer's protocol. For the adaptor ligation step we chose Y-shaped adaptors compatible with Nextera XT Index Kit in the amount of 4 pM per reaction. PCR amplification with index adaptors in the amount of 7.5 pM per reaction was performed with Nextera XT Index Kit (Illumina) in 25 µL total volume according to NEBNext Ultra II DNA Library Prep Kit for Illumina protocol with 15 cycles.

High throughput sequencing was performed using Illumina HiSeq 1500 with HiSeq PE Rapid Cluster Kit v2 and HiSeq Rapid SBS Kit v2 (500 cycles). Paired reads were filtered with Trimmomatic using parameters SLIDINGWINDOW:4:25 MINLEN:40. After read trimming, genome assembly and selection of Coronaviridae sequences, we obtained two contigs with lengths of 20098 and 10135 bp. Genome assembly was completed by SPAdes 3.15.0 [37]. Coronaviridae sequences were selected by BLASTn [38] of assembled contigs using all of the available Coronaviridae genomes as a reference. Read mapping was performed using bowtie2 [39].



p

Alignment of contigs to the closest MERSr-CoV genomes (MG596802.1) revealed an uncovered 62 bp fragment between these contigs. Gaps within the assembled genome were closed and confirmed using Sanger sequencing. We performed Sanger sequencing of this area to connect two contigs and obtain full-genome sequencing with following primers: 1-forward ACATACGTGACAATGGTTCATTAG, and 1-reverse CTGTTGACTCTCTATAAATATAGAAC.

Genome annotation was performed by Geneious 7.1.9 and edited manually.

TRS-L and TRS-B alignment was made by Vector NTI software

### **Phylogenetic analysis**

We downloaded beta-CoV sequences from GenBank using the keyword “Merbecovirus” as the primary filter to identify beta-CoVs from bats, humans, camels and other mammals. We included 9 full genomes of CoVs from bats, 1 from a lama, 232 from camels and 254 from humans. For phylogenetic tree construction we used only complete genomes of viruses from highly represented hosts: individual trees were previously constructed separately for full genomes of camel and human Merbecoviruses. Clusters with a p-distance >0,001 were collapsed and one sequence per cluster was selected randomly. The complete genome of the newly discovered CoV was aligned with all sequences using MAFFT v7 via the online service of RIMD (Research Institute for Microbial Disease of Osaka University, Japan): <https://mafft.cbrc.jp/alignment/software> then bp on 5’-end and 3’-end were trimmed.

For phylogenetic analysis of genes encoding Nucleocapsid protein the partial sequences designated in metadata as “merbecovirus nucleocapsid gene” were downloaded from GenBank and combined with sequences extracted from complete genomes (sampling described above). The obtained 128 sequences were aligned



p

using MAFFT v7, trimmed the ~110 bp on 5'-end and ~200 bp on 3'-end. At last, the phylogenetic tree was constructed on alignment of ~1308 bp, the Best-fit model according to BIC: TIM2+F+I+G4.

For phylogenetic analysis of RdRp-encoding region of ORF/b as well as Spike-genes the similar method of sampling was used, namely: the partial sequences designated in metadata as “merbecovirus RdRp” or “merbecovirus Spike gene” were downloaded from GenBank, (2) combined with relevant sequences extracted from complete genomes, aligned using MAFFT v7, trimmed in the 5'-end and 3'-end. For the RdRp-encoding region, the ~2801 bp alignment of 118 sequences was used for phylogenetic tree construction, the Best-fit model according to BIC: GTR+F+I+G4. For the Spike, the ~4026 bp alignment of 123 sequences was used for phylogenetic tree construction, the Best-fit model according to BIC: GTR+F+I+G4.

Phylogenetic analyses were performed using W-IQ-TREE (<http://iqtree.cibiv.univie.ac.at/>) with ModelFinder [40], tree reconstruction [41], and ultrafast bootstrap (1000 replicates) [42]. Phylogenetic trees and coevolutionary events were visualized using the online website (<https://itol.embl.de/>) with iTOL software [43].

### **Structural modeling and molecular docking**

The SWISS-MODEL server [44] was used to determined the three-dimensional structure of the MOW-BatCoV spike protein, and the structures of DPP4 for *Myotis brandtii* [EPQ03439.1], *Pipistrellus kuhlii* (KAF6353216.1), *Erinaceus europaeus* [XP\_016043930.1], *Felis catus* [NP\_001009838.1] and *Mus musculus* [NP\_034204.1]. The DPP4 sequence of *E. europaeus* contains X in the sequence, which we replaced with the W sequence. The 5T4E crystal structure of DPP4 of Homo sapiens was obtained from the RCSB.

p

The constructed models and the crystal structure of DPP4s of the studied organisms were docked to the modeled structure of the MOW-BatCoV Spike protein using the HDOCK server [45]. The RBD (360-610 a.r.) of MOW-BatCoV and all DPP4 were used as docking sites. The docking results were analyzed in the PyMOL Molecular Graphics System, Version 2.0 [46].

## Results

### General description

We collected and analyzed 26 fecal samples from six species of bats: *Myotis dasycneme*, *Myotis daubentonii*, *Myotis brandtii*, *Nyctalus noctula*, *Pipistrellus nathusii*, *Plecotus auritus*. All bats were visually healthy. Ectoparasite analysis yielded mites in 21/26 samples with 2 samples containing both mites and fleas. PCR fragments of RdRp yielded results in 13 of 26 fecal samples, giving an overall detection rate of 50%. Using Illumina MiSeq high throughput sequencing analysis and data analysis as described in [35], we confirmed presence of CoVs in 8/13 samples (the five remaining samples were not sequenced). Five of the six species of bats sampled were infected by different CoV strains (*M. dasycneme*, *M. brandtii*, *M. daubentonii*, *N. noctula* and *P. nathusii*), and only one sample was negative for CoVs (from *P. auritus*). Of investigated species the only *Pipistrellus nathusii* were carriers of betacoronaviruses: sample numbers 16, 22 and 33, see table 1. The RdRp fragments were sequenced, all they shared >98% sequence identity to KC243390.1 (BtCoV/8-724/Pip\_pyg/ROU/2009) which was previously found in *Pipistrellus pygmaeus* from Romania and published in 2013 [47]. We propose that these represent a potentially novel Betacoronavirus in the *P. nathusii* and sequenced the complete genome of virus from sample №22.

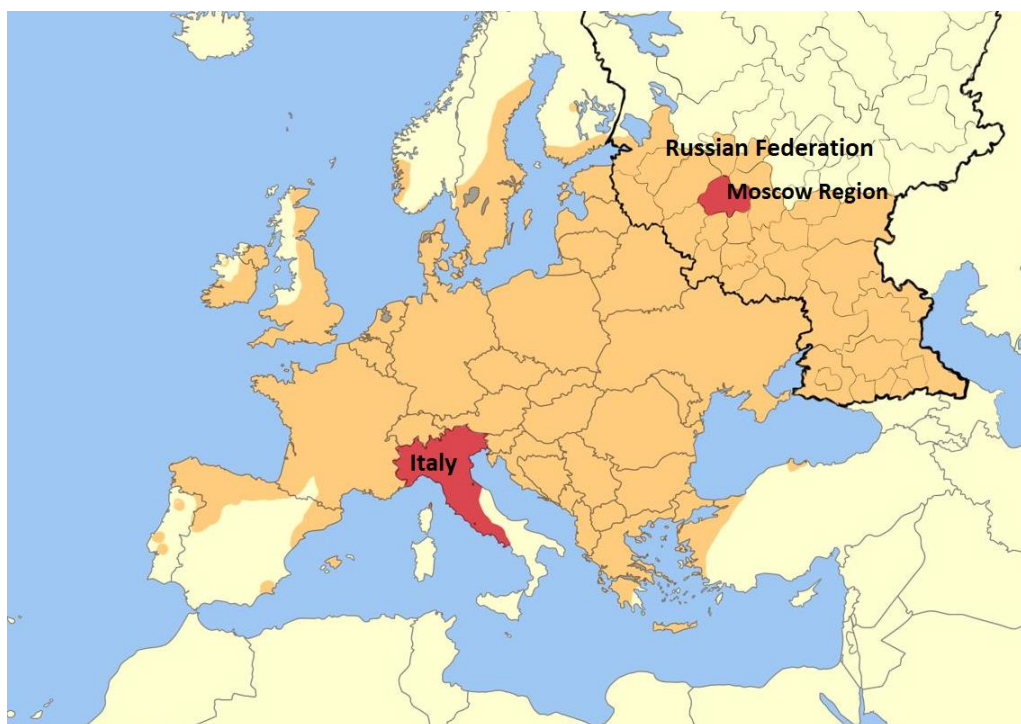
p

**Table 1. *P. nathusii* in which natural coronavirus infection has been found. Samples where Betacoronaviruses were detected are marked in bold.**

Sample ID	Gender	Age	Ectoparasites	PCR product	Characterisation of PCR product by NGS			
					fragment length, bp	Genbank ID, nearest	Identity, %	Genus
Bat№16	F	Semi-adult	Mites	+	418	KC243390.1	98.35	Beta
					420	EU375869.1	99.51	Alpha
Bat№21	F	Semi-adult	Mites	+	420	EU375864.1	98.77	Alpha
Bat№22	M	Semi-adult	Mites	+	418	KC243390.1	98.51	Beta
Bat№23	M	Adult	Mites	+	420	EU375869.1	99.51	Alpha
Bat№25	F	Adult	A lot of Mites	-	-	-	-	-
Bat№33	M	Semi-adult	A lot of Mites	+	420	EU375864.1	98.77	Alpha
					418	KC243390.1	98.35	Beta

*Pipistrellus nathusii* is widely distributed across Europe, see Figure 1 (b). In Figure 1 (a) shows the bat caught in the process of work on the fecal samples collection. The metagenome sequencing of total RNA extracted from Sample №22 resulted in 248.8 million paired reads (SRR15508267), 0.01% of them were mapped to initially obtained Coronaviridae contigs resulting in the complete genome. It was named as MOW-BatCoV strain 15-22 and has been deposited in the GenBank under the accession numbers ON325306. RdRp fragments of viral genomes from two other samples (№16 and №33) have been deposited in the GenBank as Bat-CoV/P.nathusii/Russia/MOW15-16/1/15 (accession number ON676527) and Bat-CoV/P.nathusii/Russia/MOW15-33/1/15 (accession number ON676528).

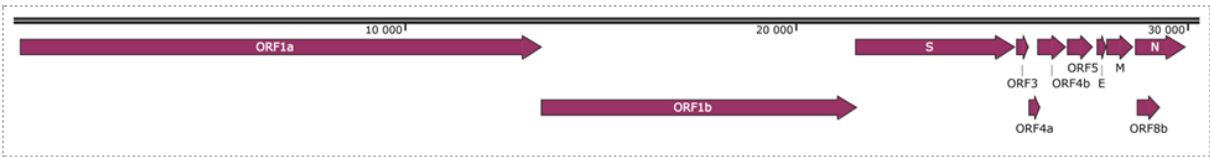
p



**Figure 1 (a) *Pipistrellus nathusii*, also known as Nathusius' pipistrelle, studied here; (b) Geographical distribution of *Pipistrellus nathusii* in Europe. The dark-yellow labeled region represents their habitats according to [48] and [49]. The red labeled area represents the location where MERS-related Bat-CoV were obtained.**

p

Bat-CoV/P. nathusii/Russia/MOW-BatCoV 15-22/2015 contains 30257 bases, with G+C content 43,72%, and a genome organization similar to other members of *Merbecovirus*, namely: ORF1ab encoding putative mature nonstructural proteins, including RdRp (RNA-dependent RNA polymerase) — S (spike protein) — the genes encoding nonstructural proteins NSP3, NSP4a, NSP4b and NSP5 — E (envelope protein) — M (membrane glycoprotein) — N (nucleocapsid phosphoprotein) — the gene encoding nonstructural protein NSP8 (Figure 2). Predicted functional domains in the ORFs are summarized in Table 2.



**Figure 2. Genome organization of MW-BatCoV, strain 15-22. Nonstructural genes (ORF1ab, ORF3, ORF4a, ORF4b, ORF5, orf8b) and structural genes (S, E, M, N) are illustrated. The major open reading frames (ORFs) had the order: ORF1ab (putative mature nonstructural proteins, including RdRp (RNA-dependent RNA polymerase)) — S (Spike) — the genes encoding nonstructural proteins (NS3, NS4a, NS4b, NS5) — E (envelope) — M (membrane glycoprotein) — N (nucleocapsid phosphoprotein) — the gene encoding nonstructural protein NS8b.**

p

**Table 2. Genome localization of predicted protein sequences of MOW-BatCoV 15-22 MERS-related coronavirus**

<b>ORF</b>	<b>nt position (start-end)</b>	<b>No. of amino acids</b>	<b>putative leader TRS-L and TRS-B</b>
ORFa	175-13524	4449	TRS-L
ORFab	175-21587	7137	TRS-L
S prot	21529-25629	1367	TRS-B
ORF3	25645-25959	105	TRS-B
ORF4a	25968-26252	95	TRS-B
ORF4b	26173-26925	251	No
ORF 5	26935-27612	226	TRS-B
E prot	27691-27939	83	TRS-B
M prot	27954-28619	222	TRS-B
Nprot	28675-29982	436	TRS-B
ORF 8b	28721-29314	198	no

The size and genomic localization of the nonstructural proteins (NSP 1–16) encoded by ORF1ab were predicted by sequence comparison with MERS CoV (human HCoV-EMC/2012) and other beta-CoV species. Proteins and 15 expected cleavage sites are shown in Table 3. In ORF1ab the sequence “UUUAAAC”, which is conserved throughout all CoVs -is located at nucleotide position 13497–13503. A predicted leader transcription regulatory sequence (TRS-L), and seven putative transcription regulatory sequences body TRS-B, representing signals for the discontinuous transcription of subgenomic mRNAs (sgmRNAs), have been identified (Table 3). All TRS have conserved AACGAA motif forming the conserved TRS core in betacoronaviruses [50]. With only one G/A modification in ORF3 (Figure 3).



p

**Table 3. The list of proteins and 15 expected cleavage sites encoded by ORF1ab of MOW-BatCoV 15-22 MERS-related coronavirus. Superscript numbers indicate positions in polyprotein pp1a/pp1ab or position in available sequence with the supposition of a ribosomal frameshift based on the conserved slippery sequence (UUUAAAC) of Coronaviruses. Localized at nucleotide position 13497–13503 ADRP - ADP-ribose 1-phosphatase, PL2pro - papain-like protease 2, 3CLpro - coronavirus NSP5 protease, Hel - helicase, NTPase - nucleoside triphosphatase, ExoN – exoribonuclease, NMT N7 - methyltransferase, NendoU - endoribonuclease, OMT - 2' O-methyltransferase.**

<b>NSP</b>	<b>Position of the putative cleavage sites</b>	<b>Protein size (no. of amino acids)</b>	<b>Putative functional domain(s)</b>	<b>predicted pp1a/pp1b cleavage site sequences</b>
NSP1	Met <sup>1</sup> -Gly <sup>195</sup>	195		LLGG
NSP2	Asp <sup>196</sup> - Gly <sup>788</sup>	663		LKGG
NSP3	Ala <sup>859</sup> - Gly <sup>2798</sup>	1940	ADRP, PL2pro	IVGG
NSP4	Ala <sup>2799</sup> -Gln <sup>3305</sup>	507		GVLQ
NSP5	Ser <sup>3306</sup> -Gln <sup>3611</sup>	306	3CLpro	LQS
NSP6	Ser <sup>3612</sup> - Gln <sup>3903</sup>	292		VQS
NSP7	Ser <sup>3904</sup> - Gln <sup>3986</sup>	83		LQA
NSP8	Ala <sup>3987</sup> -Gln <sup>4185</sup>	199	Primase	LQN



p

NSP9	Asn <sup>4186</sup> -Gln <sup>4295</sup>	110		LQA
NSP10	Ala <sup>4296</sup> -Gln <sup>4435</sup>	140		PQS
NSP11	Ser <sup>4436</sup> -Ile <sup>4449</sup>	14	Short peptide at the end of ORF1a	RGSI
NSP12	Val <sup>4450</sup> -Gln <sup>5369</sup>	920	RdRp	LQA
NSP13	Ala <sup>5370</sup> -Gln <sup>5967</sup>	598	HEL, NTPase	LQS
NSP14	Ser <sup>5968</sup> -Gln <sup>6491</sup>	524	ExoN, NMT	VQG
NSP15	Gly <sup>6492</sup> -Gln <sup>6834</sup>	343	NendoU	LQA
NSP16	Ala <sup>6835</sup> -Cys <sup>7137</sup>	303	OMT	

ORFab	(1)	-----CGA	ACTTAA	(10)
Sprot	(21471)	CA	AATTAAACGA	ACTAGTT (21488)
ORF3	(25626)	TT	AAGCAACGAG	CTTTAT (25643)
ORF4a	(25952)	AA	AATTAAACGA	ACTCTAT (25969)
ORF5	(26921)	GGT	AAAACGA	ACTATGG (26938)
Eprot	(27677)	TTGGAA	AACGA	ACTATGT (27694)
Mprot	(27932)	GG	ATTAAACGA	ACTCTCT (27949)
Nprot	(28646)	ATC	ATTAAACGA	ATCTTAT (28663)

**Figure 3. Putative leader TRS-L and TRS-B alignment. Strictly conservative and identical residues are highlighted in yellow and blue, respectively.**

p

The ICTV proposed that viruses sharing >90% aa sequence identity in the conserved replicase domains should be considered conspecific [51]. A separate comparison of the amino acid sequences of seven conserved ORF1ab domains, as suggested by the ICTV) for formal CoV species delineation was made and only the NSP3 (ADRP) aa sequence is below the 90% threshold value in comparison with MERS. The ORF1ab possessed 81,5-82,3% aa identities to the ORF1ab of other members of *Merbecovirus*. Comparison of the seven conserved domains in replicase polyprotein pp1ab with other coronaviruses showed that MOW-BatCoV 15-22 possessed for NSP3 (ADRP) - 68,7% of homology, NSP5 (3CLpro) - 90,8%, NSP12 (RdRp) - 96,6%, NSP13 (Hel) - 97,8%, NSP14 (ExoN) - 97,9%, NSP15 (NendoU) - 93,3% and NSP16 (O-MT) - 93,7% aa identities to other members of *Merbecovirus* respectively.

MOW-BatCoV 15-22 MERSr-CoV has nucleic acid identity from 81,32% to 82,46% (with coverage 82-85%) to the ten closest MERS or MERSr-CoVs from bats (*Vespertilio sinensis*, *Vespertilio superans*, *Pipistrellus cf. hesperidus*) humans and camels reported between 2013 and 2015 (Table 4). The highest sequence identity was to 2 MERSr-CoVs from bats (*Hypsugo savii* and *Pipistrellus kuhlii*) captured in Italy in 2011 (88% coverage genomes were 81.48% and of 81.45%, respectively).

p

**Table 4. Analysis of full genome identify among MOW-BatCoV 15-22 and other members of Merbecovirus (the top 10 of closest full genome sequences available in GenBank).**

Description	Given Scientific Name (by Author's)	Host	Collection date	Country	Query Cover	Per. ident	Acc. Len	Accession
Bat-CoV/H.savii/Italy/206645-40/2011, complete genome	Middle East respiratory syndrome-related coronavirus	<i>Hypsugo savii</i> (bat)	2011	Italy	88%	81,48%	30048	MG596802.1
Bat-CoV/P.khulii/Italy/206645-63/2011, complete genome	Middle East respiratory syndrome-related coronavirus	<i>Pipistrellus kuhlii</i> (bat)	2011	Italy	88%	81,45%	30039	MG596803.1
Bat coronavirus Vs-CoV-1 genomic RNA, nearly complete genome	Bat coronavirus	<i>Vespertilio sinensis</i> (bat)	undescribed	Japan: Tokyo, Okutama	85%	81,32%	29930	LC469308.1
BtVs-BetaCoV/SC2013, complete genome	BtVs-BetaCoV/SC 2013	<i>Vespertilio superans</i> (bat)	2013	China	82%	81,27%	30423	KJ473821.1
isolate PREDICT/PDF-2180, complete genome	Bat coronavirus	<i>Pipistrellus cf. hesperidus</i> (bat)	2013	Uganda	82%	80,84%	29642	NC_034440.1
isolate Riyadh_9_2013, complete genome	Middle East respiratory syndrome-related coronavirus	<i>Homo sapiens</i>	2013	Saudi Arabia	82%	82,47%	30055	KJ156869.1
isolate Hu/Jordan-201440011123/2014, complete genome	Middle East respiratory syndrome-related coronavirus	<i>Homo sapiens</i>	2014	Jordan	82%	82,46%	30123	MK039552.1

p

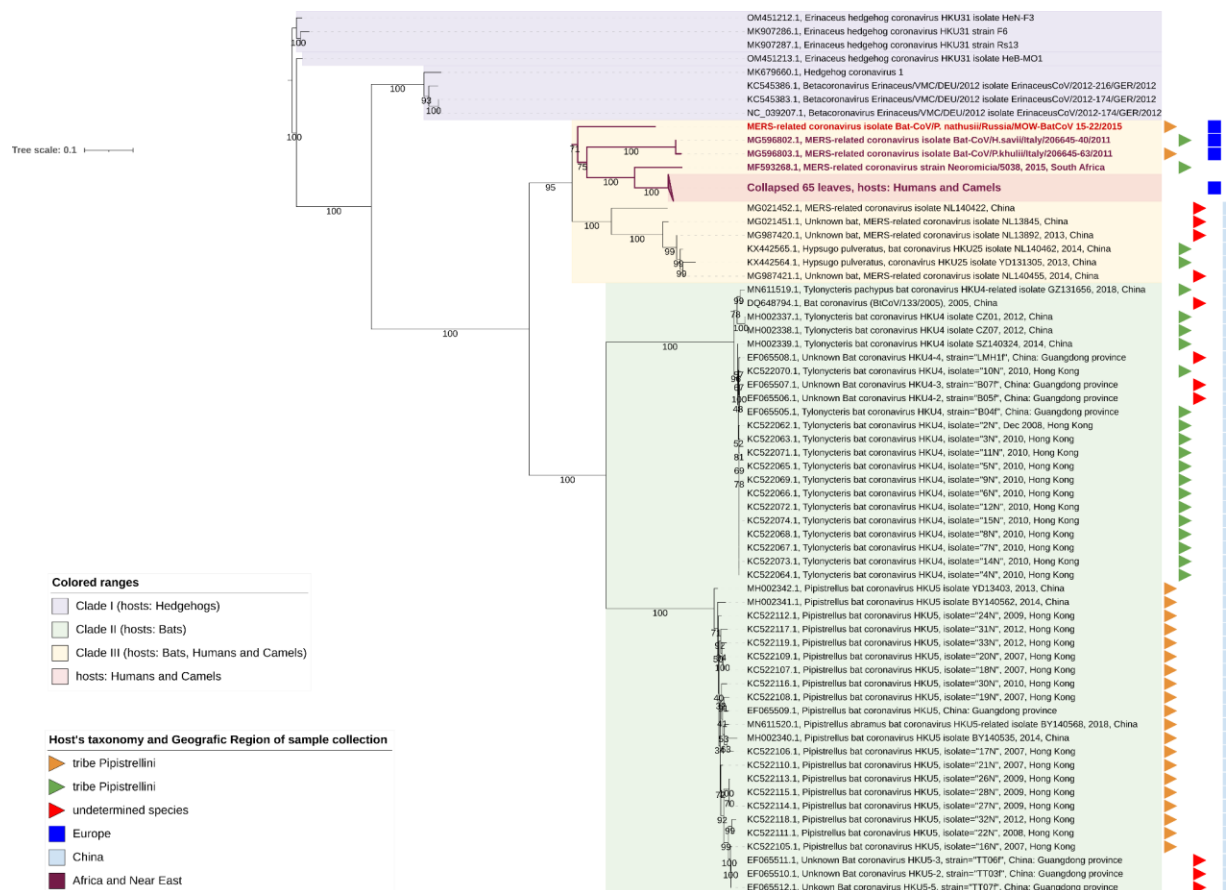
strain Hu/UAE_032_2014 , complete genome	Middle East respiratory syndrome- related coronavirus	<i>Homo sapiens</i>	2014	United Arab Emirate	82%	82,46%	30123	KY581693.1
camel/UAE_B73_2 015, complete genome	Middle East respiratory syndrome- related coronavirus	<i>Camelus dromedarius</i>	2015	United Arab Emirates	82%	82,46%	30123	MF598663.1
isolate Florida/USA- 2_Saudi Arabia_2014, complete genome	Middle East respiratory syndrome- related coronavirus	<i>Homo sapiens</i>	2014	USA	82%	82,46%	30123	KP223131.1

## Phylogenetic analysis

Our phylogenetic analysis of complete MERSr-CoV genomes showed four distinct clades (Figure 4). Clade I consist of nine CoV sequences from hedgehogs. Clade II consists of 17 CoV sequences from bats. Clade III consists of 59 sequences belonging to two subclades: (a) 9 CoV sequences from bats; (b) 49 CoV sequences from humans and camels, as well as one from a bat (strain *Neoromicia*/5038, from virome of the bat *Neoromicia capensis* collected in South Africa, in April of 2015, Genbank accession number MF593268.1). MOW-BatCoV 15-22 belonged to Clade III(a) along with MERSr-CoVs from *H. savii* and *P. kuhlii* from Italy.



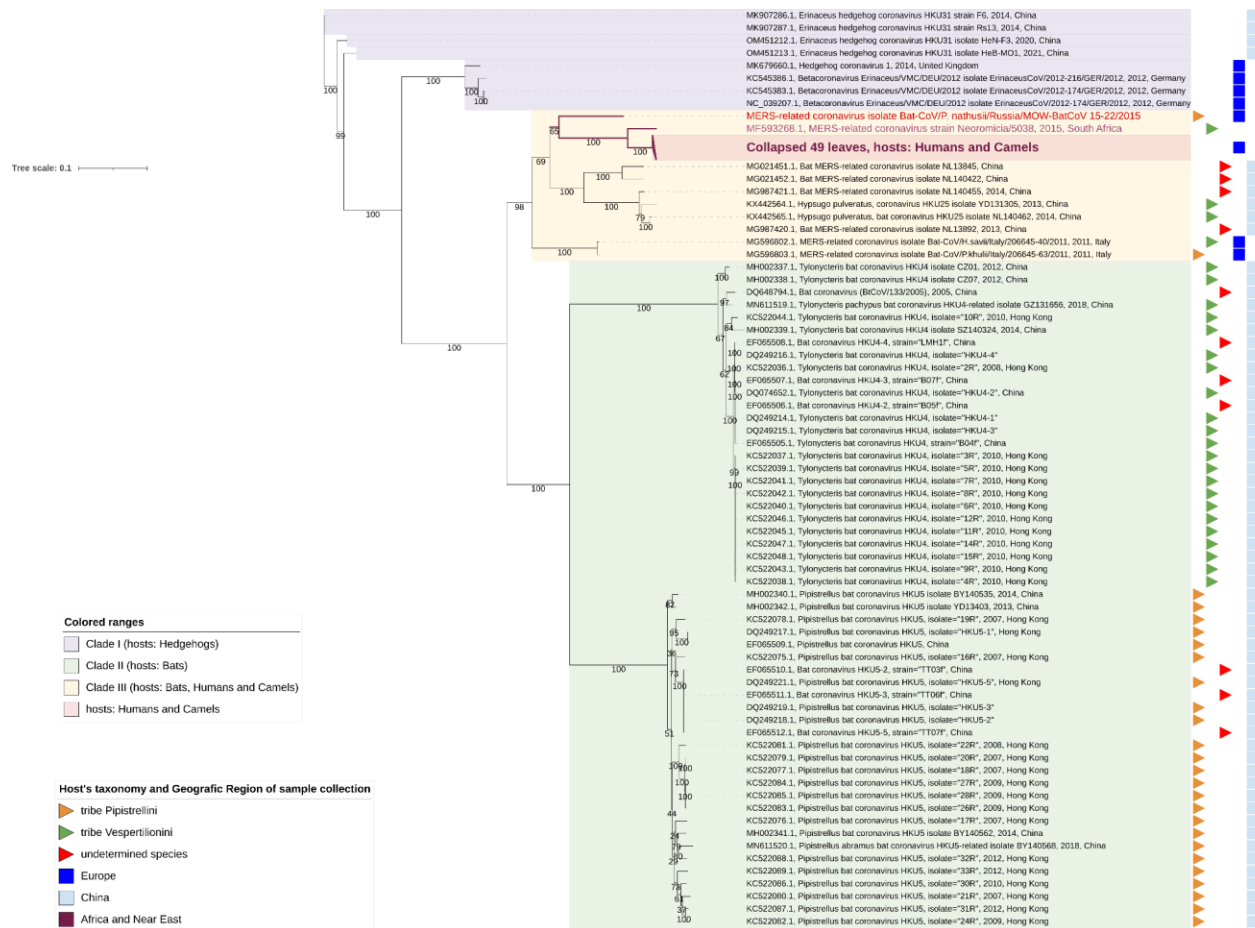
p



**Figure 5a. Phylogenetic trees of the Nucleocapsid protein coding regions of Merbecovirus genomes. The phylogenetic tree was constructed on 65 partial N-gene sequences (1272-1297 bp). Numbers show bootstrap values. Best-fit model of substitution according to BIC: TIM2+F+I+G4. The virus described in this study is labeled in red bold font.**

In the RdRp-tree the MOW-BatCoV 15-22 formed a subclade MOW-BatCoV 15-22 together with the only Neoromicia/5038 and human/camel's viruses, but MERS-related coronaviruses from *H. savii* and *P. kuhlii* (from Italy) fall into different clade (see Figure 5b).

p

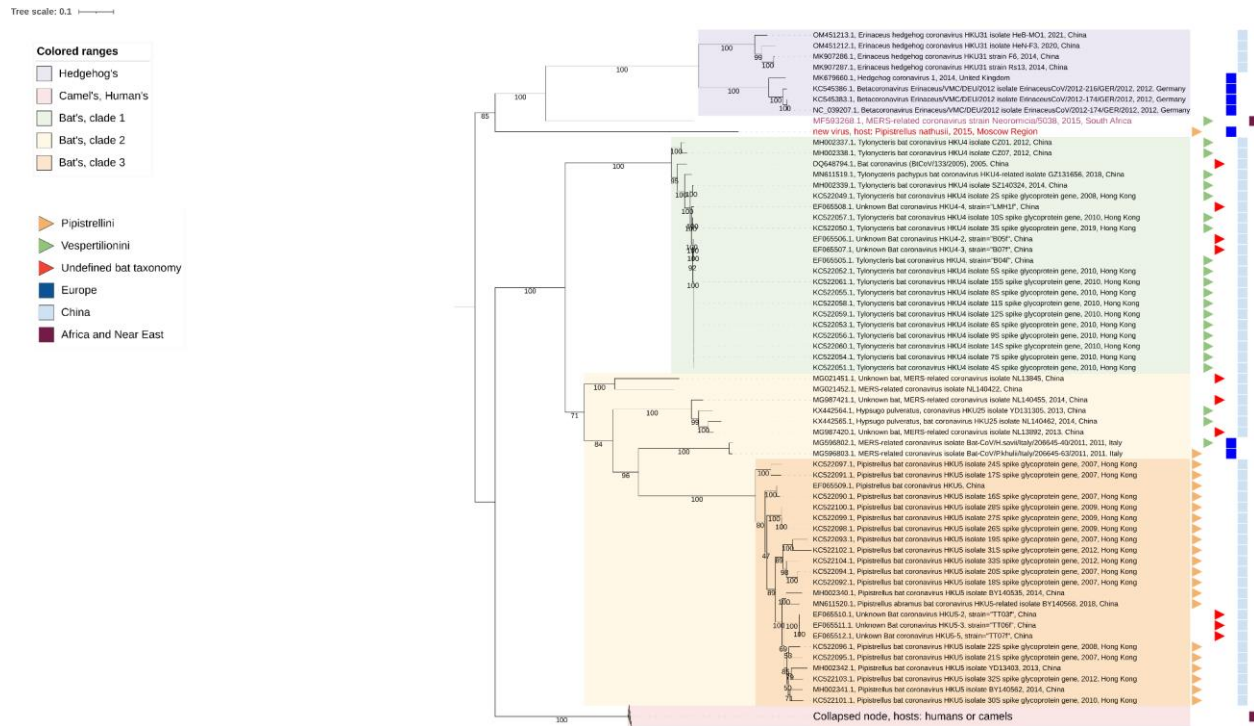


**Figure 5b. Phylogenetic tree based on 2798-2802 of the partial RdRp (NSP12) proteins coding regions of Merbecovirus genomes. Numbers show bootstrap values. Best-fit model of substitution according to BIC: GTR+F+I+G4. The virus described in this study is labeled in red bold font.**

Phylogenetic analysis of the spike protein encoding region of Merbecoviruses (Figure 5c) shows that MOW-BatCoV 15-22 is most closely related to coronavirus strain Neoromicia/5038. Unexpectedly, both MOW-BatCoV 15-22 and Neoromicia/5038 showed closest S-protein sequence identity to CoVs from *Erinaceus* hedgehogs, forming a distinct branch among Merbecovirus (Figure 5c).



p



**Figure 5c. Phylogenetic tree based on the partial Spike glycoprotein protein encoding regions of Merbecovirus genomes. Numbers show bootstrap values. Best-fit model of substitution according to BIC: GTR+F+I+G4. The virus described in this study is labeled in red bold font.**

## Docking

To predict and analyze the interaction of MOW-BatCoV Spike glycoprotein with DPP4 receptors of the different mammalian species, the three-dimensional structures of these proteins were obtained by homologous modeling. The DPP4 proteins of two bats (*Myotis brandtii* and *Pipistrellus kuhlii*), the hedgehog (*Erinaceus europaeus*), domestic cat (*Felis catus*) and mouse (*Mus musculus*) were used for analysis.

MOW-BatCoV Spike protein structure was determined using the 6Q04 reference structure of human MERS-CoV Spike protein. We selected the protein

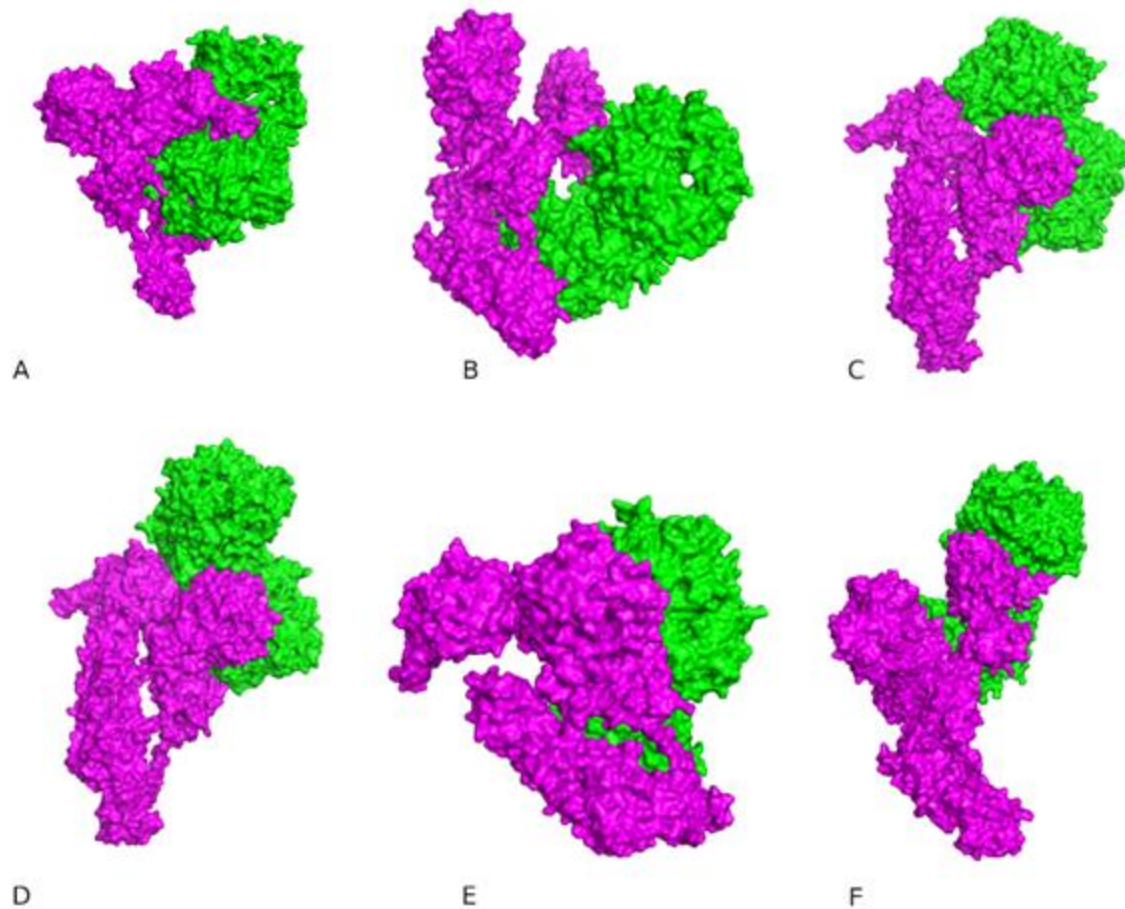
p

structure of the amino acid sequence with the highest homology (61.41%) to the input sequence as a template. The resulting model had a GMQE of 0.61.

For the DPP4 structure models of mammals, the reference human FAPalpha was used. The best 6Y0F structure models were determined for *M. brandtii* and *P. kuhlii* DPP4 with 92.14% identity (GMQE = 0.91) and 92.31% identity (GMQE = 0.93), respectively. The best reference structure for DPP4 structure models of *E. europaeus*, *F. catus*, and *M. musculus* was 2QT9 with identities of 84.97% (GMQE = 0.9), 87.97% (GMQE = 0.89) and 84.15% (GMQE = 0.9), respectively.

Molecular docking allowed us to determine the best binding cluster for the Spike-protein of MOW-BatCoV and DPP4 proteins of the organisms studied. Highest binding was predicted between MOW-BatCoV spike protein and DPP4 of *M. brandtii* (docking score -320.15). The second highest binding was predicted for *E. europaeus* (docking score -294.51), which is consistent with phylogenetic analyses. Docking results predicted higher binding of the MOW-BatCoV S to *H. sapiens* DPP4 (docking score -290.79) than to *P. kuhlii* DPP4 (-274.21), *M. musculus* (docking score -262.74) and *F. catus* (docking score -248.18), see Figure 6 (a-f).

p



**Figure 6. Protein–protein binding interactions of MOW-BatCoV Spike protein and DPP4: A - *M. brandtii*, B - *P. kuhlii*, C - *E. europaeus*, D - *F. catus*, E - *M. musculus*, F - *H. sapiens*. The MOW-BatCoV Spike proteins are marked in magenta, DPP4 are marked in green.**

Analysis of the results of molecular docking of the Spike-protein of MOW-BatCoV and DPP4 protein of *M. brandtii* and *E. europaeus* showed that the Spike-protein of MOW-BatCoV shares binding sites with DPP4 of these two species. Thus, 43 and 42 binding sites were shown for *M. brandtii* and *E. europaeus* in the Spike-protein of MOW-BatCoV, among which 31 sites were identical (Table 5).

p

**Table 5. Predicted MOW-BatCoV-DPP4 binding sites of *M. brandtii* and *E. europaeus***

	<i>Myotis brandtii</i>	<i>Erinaceus europaeus</i>
binding sites	36, 163, 165, 166, 188, 189, 191, 230, 232, 233, 370, 389, 417, 419, 485, 486, 487, 488, 489, 490, 585, 611, 613, 617, 618, 619, 620, 637, 696, 697, 698, 701, 703, 704, 705, 707, 716, 718, 719, 720, 722, 724, 734	36, 38, 39, 154, 156, 157, 158, 159, 163, 165, 166, 188, 189, 190, 191, 197, 230, 232, 233, 370, 389, 416, 417, 419, 485, 486, 487, 488, 489, 490, 585, 613, 617, 618, 619, 620, 697, 698, 699, 703, 704, 718
overlapping binding sites	36, 163, 165, 166, 188, 189, 191, 230, 232, 233, 370, 389, 417, 419, 485, 486, 487, 488, 489, 490, 585, 613, 617, 618, 619, 620, 697, 698, 703, 704, 718	

## Discussion

A large number of  $\beta$ -CoVs has been identified from bats globally. The MERSr-CoV described here is the first to be identified in Russia. Twenty-six animals of six different bat species which wide distributed in the central European part of the Russian Federation (five of *Myotis dasycneme*, three of *Myotis brandtii*, five of *Myotis daubentoniid*, four of *Nyctalus noctule*, six of *Pippistrellus nathusii*, two of *Plecotus auritus* and single *Vespertilio murinus*) were analysed in the study. The products of amplificaton of RdRp genes was detected in 50% RNA samples extracted from rectal swabs of animals. Sequencing confirms the presence of merbecoviruses in three of six analysed samples from *P. nathusii* only (semi-adult animals, one of them was female and two were males). The animals were caught at the same time, in the same geographic location, so we believe they were from the same colony. BLAST against GenBank records as well as topology of the

p

phylogenetic tree based on the fragments of the RdRp genes revealed that the bats were infected with same virus, MOW-BatCoV/15-22.

The complete genome of MOW-BatCoV 15-22 showed the highest similarity (88% identity) to MERS-related viruses from bats (Bat-CoV/H.savii/Italy/206645-40/2011 and Bat-CoV/P.khulii/Italy/206645-63/2011) which has been reported in Italy [17]. *P. nathusii* is a migratory bat, which habitats the major parts of Europe: from Fennoscandia and British Isles in the north to the Mediterranean areas in the south. The breeding areas of this species are regions of north-eastern Europe. As a result of low abundance of aerial insects during winter, Nathusius' pipistrelles from Central European (Germany and Poland) and northeastern populations (Fennoscandia, the Baltic countries, Belarus, and Russia) perform long-distance flights migration in the late summer (during approximately two months with stopping for mating) in the Switzerland, Benelux countries, France, Spain, Italy, and Croatia [52,53]. The longest migration record of this species was documented at 2224 km, between S Latvia and N Spain [53]. During migration, *P. nathusii* may come into contact with bats of the same species when mating. And they can also come into contact with bats of other species in roosting areas. Migration routes explain the fact that very similar viruses have been found in bats from Italy and Russia.

According to phylogenetic analysis of the complete genomes the MOW-BatCoV 15-22 falls into clade of human/camel's MERS viruses together with a few bat viruses and due to its distinct phylogenetic position and amino acid differences should be considered as a novel MERSr-CoV. The replicase polyprotein of new virus (MOW-BatCoV 15-22) showed more than 90% homology for six of seven domains, the only NSP3 (ADRP) is shown 68,7% of homology to sequences of the other members of Merbecovirus. According to demarcation criteria of ICTV [51],



p

we believe MOW-BatCoV 15-22 represents the same species of Merbecovirus as NeoCoV because of phylogenetic analysis of RdRp showed, MOW-BatCoV 15-22 and NeoCoV together are closest to MERS-CoV. Of known bat viruses, the NeoCoV from *Neoromicia capensis* (S. Africa) is the closest to Middle East respiratory syndrome coronavirus (MERS-CoV) which could infect humans and dromedary camels and it is considered that NeoCoV shares sufficient genetic similarity in the replicase genes to be part of the same viral species with MERS-CoV [20,54].

Phylogenetic analysis of complete genome sequences as well as N- and RdRp-sequences suggest that ten viruses from bats found in distinct geographic regions (MOW-BatCoV 15-22 from Russia, MG596802.1 and MG596803.1 from Italy, NeoCoV *Neoromicia*/5038 from South Africa, and multiple strains from bats in China - MG021452.1, MG021451.1, MG987420.1, MG987421.1, KX442565.1, KX442564.1) form a distinct phylogenetic clade with MERS-CoVs from humans and camels, with high bootstrap support. However, phylogenetic analysis of spike protein encoding genes demonstrated similarity of two of these ten bat viruses - the novel MOW-BatCoV 15-22 and NeoCoV - to CoVs from *Erinaceus europaeus* (the European hedgehog).

We believe that the proven close relationship between Spike genes of viruses from bats and hedgehogs which live in the same geographic regions (namely Europe) raises the question of the possibility of interspecies transmission in the present time. Middle East syndrome Coronavirus (MERS-CoV) likely originated in bats and passed to humans through dromedary camels. Previous work suggests that MERS-CoV originated from an ancestral virus in a bat reservoir and spilled over into dromedary camels around 40 decades ago, where it circulated endemically before emerging in humans in 2012 [6,7]. Presently camels play an important role as a

p

constantly reservoir of MERS-CoV and transmit virus to people [55–57], while the bats are widely considered to be the evolutionary, disposable source of the virus [20]. But, besides dromedaries which are the proven source of human MERS and bats, MERS-related viruses have also been discovered in hedgehogs (Erinaceus). Hedgehog carriers of betacoronaviruses have been found in China [16], Germany [58], France [59] and Poland [60]. Phylogenetic analysis carried previously showed that betacoronaviruses from Chinese and German hedgehogs (Ea-HedCoV HKU31 and BetaCoV Erinaceus/VMC/DEU/2012) were closely related to NeoCoV and BatCoV from African bats in the spike region. Therefore, authors suggested that the bat viruses arose as a result of recombination between hedgehogs and bat viruses [16]. Our independent finding of one more virus from European bat *Pipistrellus nathusii* which was found closely related to viruses from hedgehogs in the Spike region but not in N- and RdRp regions support the idea of recombination between ancestral viruses of bats and hedgehogs. Our independent finding of a novel CoV from a European bat with spike protein encoding sequences closely related to those from hedgehog MERSr-CoVs also suggests recombination between ancestral viruses of bats and hedgehogs. The overlap in geographic range between *P. nathusii* and the natural distribution of *E. europaeus* raises the possibility that this recombination represents an ancestral interspecies transmission. Our findings support the need for wider surveillance of MERSr-CoVs in both bats and hedgehogs

MERS-CoV targets a cell-surface receptor, the Dipeptidyl peptidase 4 (DPP4, also known as CD26). A receptor-binding domain (RBD) on the viral spike glycoprotein (S) mediates this interaction and is bound to the extracellular domain of human DPP4 [31,61]. The HKU4 and HKU5 merbecoviruses from Chinese bats are closely related to MERS-CoV in spike protein genes. The HKU4 viruses can use the MERS-CoV receptor DPP4, but not HKU5. Another MERS-CoV-related



p

betacoronavirus, Hp-BatCoV HKU25 occupies a phylogenetic position between that of HKU4 and HKU5 and can binding of DPP4 protein for entry to DPP4-expressing cells, although with lower efficiency than that of MERS and HKU4 viruses [26]. At the same time, at least in some merbecoviruses (namely, Bat-CoV-PREDICT/PDF-2180 and NeoCoV) the domain RBDs of S-protein have been shown to use the ACE2 receptor, their RBDs amino acid composition and sequence differ greatly from RBD of SARS-CoV-2 (slightly more than 18% of identical aa) [62]. In the novel MERSr-CoV (MOW-BatCoV 15-22), the RBD region appears likely to interact with DPP4 across amino acids 366–624 within the S1 subunit. The amino acid composition of the RBD domain of the MOW-BatCoV/15-22 virus differs by approximately the same range from the RBD domains of those viruses that interact with DPP4 receptors (32-36,6% the same amino acids) and of viruses that interact ACE2 (33-33,7% the same amino acids). Thus, while it cannot be ruled out that MOW-BatCoV/15-22 binds to other cell receptors (e.g. ACE2), it is more likely it binds to DPP4.

Previous work demonstrated that only two mutations in the HKU4 coronavirus spike protein encoding region can make it infectious for human cells [61]. These are changes in one amino acid in two motifs each hPPC (recognized by furin proprotein convertase) and hECP (recognized by endosomal cysteine protease Cathepsin L). In MERS-CoV which causes the Middle East respiratory syndrome hPPC is Arg748- Ser749-Val750-Arg751-Ser760, hECP MERS Ala763-Phe764-Asn765 while in HKU4 hPPC is Ser746-Thr747-Phe748-Arg749-Ser750 hECP Asn762-Tyr763-Thr764. When Ser746 was changed on Arg (to make motifs recognizable by protease) and Asn762 on Ala (to destroy a potentially existing N-linked glycosylation site) fully mediates viral entry into human cells. In the MOW-BatCoV/15-22 virus hPPC motif is Pro758-His759-Ser760-Arg761 (based on MERS and HKU4). If compared with the ACE2 interacting S proteins of Human

p

Sars-CoV-2, Bat CoV PREDICT/PDF-2180 and NeoCoV, then it is probably Ser760-Arg761-Thr762-Asn763. It is possible that substitution of either Pro758 or Asn763 the mentioned constitutions can lead to the furin cleavage site formation and it's subsequent recognition by furin and could result in an increased ability to infect human cells.

hEPC in MOW-BatCoV/15-22 virus is Ala772-Tyr773-Pro774. Here, as in HKU4, there is no aa containing nitrogen, which removes the possibility of an N-linked glycosylation site. The second aa is Tyr773, which corresponds to the conservative aa at this site (Tyr or Phe) in coronaviruses. That is, there is probably nothing that needs to be changed. Thus, just one mutation in MOW-BatCoV/15-22 virus could well lead to infection of human cells.

Computer molecular docking modeling identified the DPP4 receptors of species that MOW-BatCoV/15-22 spike glycoprotein is likely able to bind to. All of these have overlapping habitats with people, providing opportunity for spillover in a natural setting. The lowest protein-protein binding was predicted in the interaction of MOW-BatCoV/15-22 Spike-protein and DPP4 of mouse and cat. This finding is supported by the lack of ability of MERS-CoV to infect cell lines derived from mice and cats [10,27]. The highest protein-protein binding was predicted in the interaction of MOW-BatCoV/15-22 Spike-protein and DPP4 of the bat *M. brandtii*, then of the hedgehog, *E. europaeus*. This is consistent with phylogenetic analyses and may indicate evolutionary relationships and recombination between spike protein sequences of bat and hedgehog CoVs. These, and previous reports of a wide distribution of MERSr-CoVs in hedgehogs suggest they could represent the natural reservoir of this clade of novel betacoronaviruses, subgenus Merbecovirus, for example, see [16,58–60] and that the pathway of emergence of MERS-CoV may be

p

more complicated than currently known. Additionally, hedgehogs are increasingly kept as pets, with substantial numbers bred and shipped internationally, including in the Americas where MERSr-CoVs have not yet been reported. We suggest that hedgehogs should be considered, tentatively, as potential intermediate hosts for spillover of MERSr-CoVs between bats and humans, and that screening of captive hedgehogs should be conducted to rule out potential for future zoonotic spillover.

## **Funding and support**

ASS, AIV, SAE, KEV, YAP, SMV: The zoology, molecular virology, sequencing and bioinformatics works, supported by RFBR Grant № 20-04-60561.

ASD, ASG, VGD: The biochemistry analysis and comparative virology analysis works supported by RSF Grant № 20-64-46014

## **Declaration of Competing Interest**

The authors declare that they have no competing interests.

## **References**

1. WHO team. Weekly epidemiological update on COVID-19 - 11 May 2022. Weekly epidemiological update on COVID-19. 2022 [cited 17 May 2022]. Available: <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---11-may-2022>
2. de Wit E, van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: recent insights into emerging coronaviruses. *Nature Reviews Microbiology*. 2016;14: 523–534. doi:10.1038/nrmicro.2016.81
3. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *New England Journal of Medicine*. 2012;367: 1814–1820. doi:10.1056/NEJMoa1211721
4. Dighe A, Jombart T, van Kerkhove MD, Ferguson N. A systematic review of MERS-CoV seroprevalence and RNA prevalence in dromedary camels: Implications for animal vaccination. *Epidemics*. 2019;29: 100350. doi:10.1016/j.epidem.2019.100350

p

5. WHO team. WHO MERS-CoV Global Summary and Assessment of Risk (WHO/MERS/RA/August18). In: [www.who.int/publications/i/item/who-mers-cov-global-summary-and-risk-assessment---august-2018](http://www.who.int/publications/i/item/who-mers-cov-global-summary-and-risk-assessment---august-2018) [Internet]. 1 Aug 2018 [cited 17 May 2022]. Available: [www.who.int/publications/i/item/who-mers-cov-global-summary-and-risk-assessment---august-2018](http://www.who.int/publications/i/item/who-mers-cov-global-summary-and-risk-assessment---august-2018)
6. Ramadan N, Shaib H. Middle East respiratory syndrome coronavirus (MERS-CoV): A review. *Germs*. 2019;9: 35–42. doi:10.18683/germs.2019.1155
7. Aljasim TA, Almasoud A, Aljami HA, Alenazi MW, Alsagaby SA, Alsaleh AN, et al. High Rate of Circulating MERS-CoV in Dromedary Camels at Slaughterhouses in Riyadh, 2019. *Viruses*. 2020;12: 1215. doi:10.3390/v12111215
8. Smith I, Wang L-F. Bats and their virome: an important source of emerging viruses capable of infecting humans. *Current Opinion in Virology*. 2013;3: 84–91. doi:10.1016/j.coviro.2012.11.006
9. Shi Z. Emerging infectious diseases associated with bat viruses. *Science China Life Sciences*. 2013;56: 678–682. doi:10.1007/s11427-013-4517-x
10. Hu B, Ge X, Wang L-F, Shi Z. Bat origin of human coronaviruses. *Virology Journal*. 2015;12: 221. doi:10.1186/s12985-015-0422-1
11. Woo PCY, Lau SKP, Lam CSF, Lau CCY, Tsang AKL, Lau JHN, et al. Discovery of Seven Novel Mammalian and Avian Coronaviruses in the Genus Deltacoronavirus Supports Bat Coronaviruses as the Gene Source of Alphacoronavirus and Betacoronavirus and Avian Coronaviruses as the Gene Source of Gammacoronavirus and Deltacoronavirus. *Journal of Virology*. 2012;86: 3995–4008. doi:10.1128/JVI.06540-11
12. Calisher CH, Childs JE, Field HE, Holmes K v., Schountz T. Bats: Important Reservoir Hosts of Emerging Viruses. *Clinical Microbiology Reviews*. 2006;19: 531–545. doi:10.1128/CMR.00017-06
13. de Groot RJ, Cowley JA, Enjuanes L, Faaberg KS, Perlman S, Rottier PJM, et al. *Virus Taxonomy Classification and Nomenclature of Viruses*. 1st Edition. King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ, editors. Elsevier Academic Press; 2011.
14. Lelli D, Papetti A, Sabelli C, Rosti E, Moreno A, Boniotti M. Detection of Coronaviruses in Bats of Various Species in Italy. *Viruses*. 2013;5: 2679–2689. doi:10.3390/v5112679
15. Wong A, Li X, Lau S, Woo P. Global Epidemiology of Bat Coronaviruses. *Viruses*. 2019;11: 174. doi:10.3390/v11020174

p

16. Lau SKP, Luk HKH, Wong ACP, Fan RYY, Lam CSF, Li KSM, et al. Identification of a Novel Betacoronavirus (Merbecovirus) in Amur Hedgehogs from China. *Viruses*. 2019;11: 980. doi:10.3390/v11110980
17. Moreno A, Lelli D, de Sabato L, Zaccaria G, Boni A, Sozzi E, et al. Detection and full genome characterization of two beta CoV viruses related to Middle East respiratory syndrome from bats in Italy. *Virology Journal*. 2017;14: 239. doi:10.1186/s12985-017-0907-1
18. Ithete NL, Stoffberg S, Corman VM, Cottontail VM, Richards LR, Schoeman MC, et al. Close Relative of Human Middle East Respiratory Syndrome Coronavirus in Bat, South Africa. *Emerging Infectious Diseases*. 2013;19: 1697–1699. doi:10.3201/eid1910.130946
19. Anthony SJ, Ojeda-Flores R, Rico-Chávez O, Navarrete-Macias I, Zambrana-Torrel CM, Rostal MK, et al. Coronaviruses in bats from Mexico. *Journal of General Virology*. 2013;94: 1028–1038. doi:10.1099/vir.0.049759-0
20. Anthony SJ, Gilardi K, Menachery VD, Goldstein T, Ssebide B, Mbabazi R, et al. Further Evidence for Bats as the Evolutionary Source of Middle East Respiratory Syndrome Coronavirus. *mBio*. 2017;8. doi:10.1128/mBio.00373-17
21. Reusken CBEM, Lina PHC, Pielat A, de Vries A, Dam-Deisz C, Adema J, et al. Circulation of Group 2 Coronaviruses in a Bat Species Common to Urban Areas in Western Europe. *Vector-Borne and Zoonotic Diseases*. 2010;10: 785–791. doi:10.1089/vbz.2009.0173
22. Yang L, Wu Z, Ren X, Yang F, Zhang J, He G, et al. **MERS–Related Betacoronavirus in *Vespertilio superans* Bats, China**. *Emerging Infectious Diseases*. 2014;20. doi:10.3201/eid2007.140318
23. Woo PCY, Lau SKP, Li KSM, Poon RWS, Wong BHL, Tsoi H, et al. Molecular diversity of coronaviruses in bats. *Virology*. 2006;351: 180–187. doi:10.1016/j.virol.2006.02.041
24. Woo PC, Lau SK, Li KS, Tsang AK, Yuen K-Y. Genetic relatedness of the novel human group C betacoronavirus to *Tylonycteris* bat coronavirus HKU4 and *Pipistrellus* bat coronavirus HKU5. *Emerging Microbes & Infections*. 2012;1: 1–5. doi:10.1038/emi.2012.45
25. Lau SKP, Li KSM, Tsang AKL, Lam CSF, Ahmed S, Chen H, et al. Genetic Characterization of Betacoronavirus Lineage C Viruses in Bats Reveals Marked Sequence Divergence in the Spike Protein of *Pipistrellus* Bat Coronavirus HKU5 in Japanese *Pipistrellus*: Implications for the Origin of the Novel Middle East Respiratory Syndrome Coronavirus. *Journal of Virology*. 2013;87: 8638–8650. doi:10.1128/JVI.01055-13

p

26. Lau SKP, Zhang L, Luk HKH, Xiong L, Peng X, Li KSM, et al. Receptor Usage of a Novel Bat Lineage C Betacoronavirus Reveals Evolution of Middle East Respiratory Syndrome-Related Coronavirus Spike Proteins for Human Dipeptidyl Peptidase 4 Binding. *The Journal of Infectious Diseases*. 2018;218: 197–207. doi:10.1093/infdis/jiy018
27. Letko M, Miazgowicz K, McMinn R, Seifert SN, Sola I, Enjuanes L, et al. Adaptive Evolution of MERS-CoV to Species Variation in DPP4. *Cell Reports*. 2018;24: 1730–1737. doi:10.1016/j.celrep.2018.07.045
28. Ohnuma K, Haagmans BL, Hatano R, Raj VS, Mou H, Iwata S, et al. Inhibition of Middle East Respiratory Syndrome Coronavirus Infection by Anti-CD26 Monoclonal Antibody. *Journal of Virology*. 2013;87: 13892–13899. doi:10.1128/JVI.02448-13
29. Raj VS, Mou H, Smits SL, Dekkers DHW, Müller MA, Dijkman R, et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature*. 2013;495: 251–254. doi:10.1038/nature12005
30. Cui J, Eden J-S, Holmes EC, Wang L-F. Adaptive evolution of bat dipeptidyl peptidase 4 (dpp4): implications for the origin and emergence of Middle East respiratory syndrome coronavirus. *Virology Journal*. 2013;10: 304. doi:10.1186/1743-422X-10-304
31. Wang Q, Qi J, Yuan Y, Xuan Y, Han P, Wan Y, et al. Bat Origins of MERS-CoV Supported by Bat Coronavirus HKU4 Usage of Human Receptor CD26. *Cell Host & Microbe*. 2014;16: 328–337. doi:10.1016/j.chom.2014.08.009
32. Burimuah V, Sylverken A, Owusu M, El-Duah P, Yeboah R, Lampitey J, et al. Molecular-based cross-species evaluation of bovine coronavirus infection in cattle, sheep and goats in Ghana. *BMC Veterinary Research*. 2020;16: 405. doi:10.1186/s12917-020-02606-x
33. da Silva Filho LVRF, Zerbinati RM, Tateno AF, Boas L v., de Almeida MB, Levi JE, et al. The Differential Clinical Impact of Human Coronavirus Species in Children With Cystic Fibrosis. *Journal of Infectious Diseases*. 2012;206: 384–388. doi:10.1093/infdis/jis274
34. Osang LKM. Charakterisierung des viralen Erregerspektrums bei akuten Atemwegsinfekten von Kindern und Erwachsenen mittels Tiefsequenzierung. Inaugural-Dissertation zur Erlangung des Doktorgrades der Hohen Medizinischen Fakultät der Rheinischen-Friedrich-Wilhelms-Universität Bonn. München; 2017. pp. 1–127.
35. Dedkov VG, Lukashev AN, Deviatkin AA, Kuleshov KV, Safonova MV, Poleshchuk EM, et al. Retrospective diagnosis of two rabies cases in humans by high throughput sequencing. *Journal of Clinical Virology*. 2016;78: 74–81. doi:10.1016/j.jcv.2016.03.012



p

36. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30: 2114–2120. doi:10.1093/bioinformatics/btu170
37. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. *Current Protocols in Bioinformatics*. 2020;70. doi:10.1002/cpbi.102
38. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of Molecular Biology*. 1990;215: 403–410. doi:10.1016/S0022-2836(05)80360-2
39. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 2012;9: 357–359. doi:10.1038/nmeth.1923
40. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*. 2017;14: 587–589. doi:10.1038/nmeth.4285
41. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*. 2015;32: 268–274. doi:10.1093/molbev/msu300
42. Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast Approximation for Phylogenetic Bootstrap. *Molecular Biology and Evolution*. 2013;30: 1188–1195. doi:10.1093/molbev/mst024
43. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*. 2021;49: W293–W296. doi:10.1093/nar/gkab301
44. Schwede T. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research*. 2003;31: 3381–3385. doi:10.1093/nar/gkg520
45. Yan Y, Zhang D, Zhou P, Li B, Huang S-Y. HDock: a web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Research*. 2017;45: W365–W373. doi:10.1093/nar/gkx407
46. The PyMOL Molecular Graphics System. Schrödinger, LLC; Available: <http://pymol.sourceforge.net/faq.html#:~:text=external%20GUI%20window%3F-,How%20does%20one%20cite%20PyMOL%3F,1.2r3pre%2C%20Schr%C3%B6dinger%2C%20LLC>
47. Annan A, Baldwin HJ, Corman VM, Klose SM, Owusu M, Nkrumah EE, et al. Human Betacoronavirus 2c EMC/2012–related Viruses in Bats, Ghana and Europe. *Emerging Infectious Diseases*. 2013;19: 456–459. doi:10.3201/eid1903.121503



p

48. [www.eurobats.org](http://www.eurobats.org). [cited 22 May 2022]. Available: [https://www.eurobats.org/about\\_eurobats/protected\\_bat\\_species/pipistrellus\\_nathusii](https://www.eurobats.org/about_eurobats/protected_bat_species/pipistrellus_nathusii)
49. [www.eol.org](http://www.eol.org). [cited 31 May 2022]. Available: <https://eol.org/pages/289882>
50. van Boheemen S, de Graaf M, Lauber C, Bestebroer TM, Raj VS, Zaki AM, et al. Genomic Characterization of a Newly Discovered Coronavirus Associated with Acute Respiratory Distress Syndrome in Humans. *mBio*. 2012;3. doi:10.1128/mBio.00473-12
51. de Groot RJ, Baker SC, Baric R, Enjuanes L, Gorbalenya AE, Holmes KV, et al. Coronaviridae, ICTV 9th Report (2011) - Virus Taxonomy: 2021 Release, EC 53, Online. 2022 Mar. Available: [https://talk.ictvonline.org/ictv-reports/ictv\\_9th\\_report/positive-sense-rna-viruses-2011/w/posrna\\_viruses/222/coronaviridae](https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/222/coronaviridae)
52. Petersons G. Seasonal migrations of north-eastern populations of Nathusius' bat *Pipistrellus nathusii* (Chiroptera). *Myotis*. 2004;41–42: 29–56. Available: [https://www.researchgate.net/profile/Gunars-Petersons/publication/313473217\\_Seasonal\\_migrations\\_of\\_north-eastern\\_populations\\_of\\_Nathusius'\\_bat\\_Pipistrellus\\_nathusii\\_Chiroptera/links/59bfb25a6fdcca8e56fb26f/Seasonal-migrations-of-north-eastern-populations-of-Nathusius-bat-Pipistrellus-nathusii-Chiroptera.pdf](https://www.researchgate.net/profile/Gunars-Petersons/publication/313473217_Seasonal_migrations_of_north-eastern_populations_of_Nathusius'_bat_Pipistrellus_nathusii_Chiroptera/links/59bfb25a6fdcca8e56fb26f/Seasonal-migrations-of-north-eastern-populations-of-Nathusius-bat-Pipistrellus-nathusii-Chiroptera.pdf)
53. Alcalde JT, Jiménez M, Brila I, Vintulis V, Voigt CC, Petersons G. Transcontinental 2200 km migration of a Nathusius' pipistrelle ( *Pipistrellus nathusii* ) across Europe. *Mammalia*. 2021;85: 161–163. doi:10.1515/mammalia-2020-0069
54. Corman VM, Ithete NL, Richards LR, Schoeman MC, Preiser W, Drosten C, et al. Rooting the Phylogenetic Tree of Middle East Respiratory Syndrome Coronavirus by Characterization of a Conspecific Virus from an African Bat. *Journal of Virology*. 2014;88: 11297–11303. doi:10.1128/JVI.01498-14
55. Azhar EI, Hashem AM, El-Kafrawy SA, Sohrab SS, Aburizaiza AS, Farraj SA, et al. Detection of the Middle East Respiratory Syndrome Coronavirus Genome in an Air Sample Originating from a Camel Barn Owned by an Infected Patient. *mBio*. 2014;5. doi:10.1128/mBio.01450-14
56. Müller MA, Meyer B, Corman VM, Al-Masri M, Turkestani A, Ritz D, et al. Presence of Middle East respiratory syndrome coronavirus antibodies in Saudi Arabia: a nationwide, cross-sectional, serological study. *The Lancet Infectious Diseases*. 2015;15: 559–564. doi:10.1016/S1473-3099(15)70090-3
57. Woo P, Lau S, Fan R, Lau C, Wong E, Joseph S, et al. Isolation and Characterization of Dromedary Camel Coronavirus UAE-HKU23 from Dromedaries of the Middle East: Minimal Serological Cross-Reactivity between MERS Coronavirus and Dromedary Camel

p

- Coronavirus UAE-HKU23. *International Journal of Molecular Sciences*. 2016;17: 691. doi:10.3390/ijms17050691
58. Corman VM, Kallies R, Philipps H, Göpner G, Müller MA, Eckerle I, et al. Characterization of a Novel Betacoronavirus Related to Middle East Respiratory Syndrome Coronavirus in European Hedgehogs. *Journal of Virology*. 2014;88: 717–724. doi:10.1128/JVI.01600-13
59. Monchatre-Leroy E, Boué F, Boucher J-M, Renault C, Moutou F, Ar Gouilh M, et al. Identification of Alpha and Beta Coronavirus in Wildlife Species in France: Bats, Rodents, Rabbits, and Hedgehogs. *Viruses*. 2017;9: 364. doi:10.3390/v9120364
60. Pomorska-Mól M, Ruszkowski JJ, Gogulski M, Domanska-Blicharz K. First detection of Hedgehog coronavirus 1 in Poland. *Scientific Reports*. 2022;12: 2386. doi:10.1038/s41598-022-06432-z
61. Yang Y, Liu C, Du L, Jiang S, Shi Z, Baric RS, et al. Two Mutations Were Critical for Bat-to-Human Transmission of Middle East Respiratory Syndrome Coronavirus. *Journal of Virology*. 2015;89: 9119–9123. doi:10.1128/JVI.01279-15
62. Xiong Q, Cao L, Ma C, Liu C, Si J, Liu P, et al. Close relatives of MERS-CoV in bats use ACE2 as their functional receptors. *bioRxiv*. 2022. pp. 1–60.