



**HAL**  
open science

## Bioinformatic analysis of the spike protein cleavage sites of coronaviruses in the mammalian order Eulipotyphla

Qinghua Guo, Annette Choi, Jean K. Millet, Gary R Whittaker

► **To cite this version:**

Qinghua Guo, Annette Choi, Jean K. Millet, Gary R Whittaker. Bioinformatic analysis of the spike protein cleavage sites of coronaviruses in the mammalian order Eulipotyphla. 2024. hal-04494499

**HAL Id: hal-04494499**

**<https://hal.uvsq.fr/hal-04494499>**

Preprint submitted on 7 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Bioinformatic analysis of the spike protein cleavage sites of coronaviruses in the mammalian order Eulipotyphla

Qinghua Guo <sup>1</sup>, Annette Choi <sup>2</sup>, Jean K. Millet <sup>3</sup> and Gary R. Whittaker <sup>1,2,4\*</sup>

<sup>1</sup> Master of Public Health Program, Cornell University, Ithaca, NY, USA.

<sup>2</sup> Department of Microbiology & Immunology, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA.

<sup>3</sup> Université Paris-Saclay, INRAE, UVSQ, Virologie et Immunologie Moléculaires, 78352 Jouy-en-Josas, France

<sup>4</sup> Department of Public & Ecosystem Health, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA.

\*corresponding author, [grw7@cornell.edu](mailto:grw7@cornell.edu)

**Abstract:** The mammalian order Eulipotyphla, including hedgehogs and shrews, represent a poorly understood reservoir of coronaviruses with zoonotic potential. Here, we carried out a bioinformatic analyses of these viruses based on the viral spike protein—to illustrate the complexity of coronavirus evolutionary history and the diversity of viruses from these host species, with a focus on the presence of possible furin cleavage sites within the spike protein. We found no evidence for cleavage by furin itself; however, certain strains of Wencheng Sm Shrew coronavirus were shown to have a predicted cleavage site for other member of the proprotein convertases, which are furin family members—suggesting their spillover potential. As the expanding urbanization and the trade of small mammals in the wet markets enhance the wildlife-human interactions, this may increase pathogen spillover risks. Therefore, we should implement broad wild animal surveillance and be vigilant of contact with these small wild mammals in light of one-health perspectives.

## Introduction

The current global pandemic of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), named COVID-19 by WHO, is having a huge impact on global health (Rabaan et al., 2020). It emerged 7 years after an outbreak of Middle East respiratory syndrome coronavirus (MERS-CoV) and 16 years after an outbreak of severe acute respiratory syndrome coronavirus (SARS-CoV) (van Boheemen et al., 2012). The Coronaviridae Study Group (CSG) of the International Committee on Taxonomy of Viruses officially classified SARS-CoV-2 within the family *Coronaviridae* (“The Species Severe Acute Respiratory Syndrome-Related Coronavirus,” 2020).

Coronaviruses are enveloped positive-strand RNA viruses and members of the subfamily *Orthocoronavirinae* in the family *Coronaviridae* within the order *Nidovirales* (Brian & Baric, 2005). Coronaviruses are classified into four genera, *Alphacoronavirus*, *Betacoronavirus*, and *Gammacoronavirus*, and *Deltacoronavirus*. Overall, there are currently seven well-recognized human coronaviruses, HCoV-229E, HCoV-NL63, HCoV-OC43, HCoV-HKU1, MERS-CoV, SARS-CoV, and the most recent SARS-CoV-2 (Decaro & Lorusso, 2020). HCoV-229E and HCoV-NL63 belong to the genus *Alphacoronavirus* while all the rest are classified as *Betacoronavirus*. Specifically, *Betacoronavirus* is further divided into four monophyletic lineages A-D (van Boheemen et al., 2012). Lineage A (embecovirus) includes HCoV-OC43 and HCoV-HKU1; lineage B (sarbecovirus) includes SARS-CoV and SARS-CoV-2; MERS-CoV and related viruses belong to lineage C (merbecovirus).

Interestingly, all the viral strains in lineage A and B are detected in different host species while those in lineage C and D are found in bats. For instance, HCoV-OC43 was transmitted to humans through domestic animals such as cattle or pigs (Corman et al., 2018); HCoV-229E derived from bats is hosted by alpacas (Corman et al., 2015); the ancestors of HCoV-NL63 are circulating in bats, whereas HCoV-HKU1 originated in rodents (Tao et al., 2017). SARS-CoV and MERS-CoV are also originated in bats and are carried by a wide range of wild animals (Decaro & Lorusso, 2020).

The size of coronavirus genome RNA is among the largest known so far. Despite the presence of a proof-reading EndoU enzyme, the diverse genetic features of coronaviruses are due to a high level of recombination and mutation, facilitating the emergence of novel viruses. Horseshoe bats are considered the most likely natural reservoir for SARS-CoV-2. However, the intermediate hosts for SARS-CoV-2 have still not been definitively identified, and these hosts could be a range of highly diversified animals, including mammals, avians, and reptiles (Boni et al., 2020; Hedman et al., 2021). Both domestic animals, such as dogs, cats, and pigs, and wildlife, such as pangolins, mink, and ferrets, can be intermediate hosts for SARS-CoV-2 (Zhao et al., 2020). Therefore, it is extremely important to continue studying coronavirus in wild small mammals and their zoonotic potential.

As mentioned above, the insectivorous bats under the order Chiroptera are important hosts for these alphacoronaviruses and betacoronaviruses. The ability to mutate and recombine enables coronaviruses to jump through different animal species (Cui et al., 2019). Members of the mammalian taxa of the order Eulipotyphla, including the hedgehogs, shrews, moles, and solenodons, have been demonstrated to have a close genetic relationship with the order Chiroptera (Tsagkogeorga et al., 2013) and are also insectivorous. Since the first identification of coronaviruses in European hedgehogs (*Erinaceus europaeus*) in Germany, EriCoV (Erinaceus Coronavirus) was detected in France, the United Kingdom, and Italy (Corman et al., 2014; De Sabato et al., 2020; Monchatre-Leroy et al., 2017). Furthermore, another hedgehog coronavirus was found in Amur hedgehogs (*Erinaceus amurensis*) from China (Lau et al., 2019). Several Wencheng Sm shrew coronaviruses (WESVs) were isolated in the Asian house shrews (*Suncus murinus*) while Shrew coronavirus was discovered in the common shrews (*Sorex araneus*) in China (Wang et al., 2017; Wu et al., 2018).

The Rodentia order is believed to contain the most zoonotic host species, followed by Chiroptera and then two families within the order Eulipotyphla, the shrews (Soricidae) and the moles (Talpidae) (Han et al., 2016). However, there are far fewer studies focusing on the Eulipotyphla mammals compared to the other insectivores like bats, which might cause the gap in understanding these small mammal host species. These small mammals, always living in a complex and densely populated community with rich species, also share similar biological characteristics of high metabolic rates (Bray et al., 2008; Han et al., 2016).

The coronavirus spike (S) glycoprotein on the viral surface plays a vital role in viral infection (Millet et al., 2016; Wang et al., 2013). For instance, it has been demonstrated that MERS-CoV spike protein can adapt to the DPP4 (the cellular receptor for MERS-CoV) variation of different hosts by altering its surface charge (Letko et al., 2018; Wang et al., 2013). The spike protein is composed of the S1 domain in the N-terminal region and is followed by the S2 domain (Wang et al., 2013). The S1 domain helps the viral entry into the target cell by binding to the host cell receptor; the S2 domain mediates membrane fusion (Belouzard et al., 2009). Overall, the spike protein can be proteolytically cleaved and activated at two cleavage sites (Millet & Whittaker, 2014). The cleavage occurs at the boundary between S1 and S2 (S1/S2) during biosynthesis of the spike while at the upstream position of the fusion peptide (S2') during viral entry. The N-terminal domain (NTD) and C-terminal domain (CTD) in S1 are likely to serve as the receptor-binding domain (RBD) (Li, 2016). The majority of S1-NTDs bind carbohydrates, except in MHV, which can recognize the receptor CEACAM1. S1-CTDs are responsible for binding receptors, including ACE2, APN, and DPP4, via the RBD (Wang et al., 2013). As enveloped viruses, these coronaviruses enter host cells by membrane fusion. Therefore, the proteolytic cleavage and the proteases cleaving and activating the spike protein play

significant roles in understanding the viral pathogenesis. Although cellular entry might be different between Eulipotyphla coronaviruses and MERS-CoV/SARS-CoV/SARS-CoV-2 based on their low similarities in the corresponding RBD, it is still worthy to study the zoonotic potential of Eulipotyphla coronaviruses. This research might provide more insights into the viral transmission between the wild small mammals including bats, hedgehogs, and shrews, as well as the evolutionary origins of coronaviruses. We also want to highlight the importance of the wild small mammals as zoonotic host species, especially these insectivores from the order Eulipotyphla.

Here, we studied the phylogenetic relationships within Eulipotyphla coronaviruses and related coronaviruses. We also predicted the cleavage sites in the Eulipotyphla coronaviruses and explored the relationships of furin across different animal species. It was our goal to better understand the zoonotic potential of these wildlife-related coronaviruses and provide some suggestions for the public health field to prevent viral zoonoses.

## Methods

### *Phylogenetic Analysis*

The multiple sequence alignment was constructed in Geneious Prime 2020. The maximum likelihood phylogenetic tree was built in Mega-X 10.2.4 (Kumar et al., 2018). All viral spike protein amino acid sequences were downloaded from the NCBI protein database (<https://www.ncbi.nlm.nih.gov/protein/>). The spike protein accession numbers on the NCBI database were: Shrew-CoV/Tibet2014 (YP\_009755839.1); Wencheng Sm shrew coronavirus isolates: Yudu-76 (ASF90460.1), Yudu-19 (ASF90465.1), Ruian-90 (ASF90470.1), Wencheng-554 (ASF90486.1), Wencheng-578 (ASF90491.1), Wencheng-562 (ASF90496.1), Xingguo-101 (YP\_009389425.1), Xingguo-74 (YP\_009824974.1); Rhinolophus bat coronavirus HKU2 (ABQ57208); Swine acute diarrhea syndrome coronavirus (AXY04083); Erinaceus hedgehog coronavirus cw\_6c (QQQ34381); Betacoronavirus Erinaceus isolates: Italy/116988-1/2018 (QRN68024), Italy/50265-19/2018 (QRN68031), Italy/50265-17/2018 (QRN68055), Italy/50265-1/2018 (QRN68048), Italy/50265-11/2019 (QRN68066), Italy/50265-12/2019 (QRN68078), Italy/50265-13/2019 (QRN68090), Italy/50265-15/2019 (QRN68101), VMC/DEU/2012 (YP\_009513010.1); Hedgehog coronavirus 1 (QCC20713.1); Erinaceus hedgehog coronavirus HKU31 (QGA70702.1); Betacoronavirus England 1 (K9N5Q8); BtVs-BetaCoV/SC2013 (AHY61337.1); Bat coronavirus HKU5-1 (ABN10875.1); Bat coronavirus HKU4-1 (ABN10839.1); Hypsugo bat coronavirus HKU25 (ASL68953.1); Pipistrellus abramus bat coronavirus HKU5-related (QHA24687.1); Tylonycteris pachypus bat coronavirus HKU4-related (QHA24678.1).

### *Geographical Mapping of Coronaviruses from the Mammalian Order Eulipotyphla*

The maps were spatialized using QGIS Desktop 3.16.3. The location information of the shrew coronavirus and hedgehog coronavirus isolates included in the map was retrieved from NCBI viral genome database (<https://www.ncbi.nlm.nih.gov/>).

### *Furin Cleavage Sites Prediction*

Furin cleavage site prediction was generated by the ProP 1.0 Server (<https://services.healthtech.dtu.dk/service.php?ProP-1.0>).

### *Spike Protein Model Construction*

The spike protein models were built using Rosetta Comparative Modeling (Song et al., 2013). The reference spike protein structure from the coronavirus HKU2 (PDB: 6M15) was downloaded from RCSB Protein Data Bank (<https://www.rcsb.org/>).

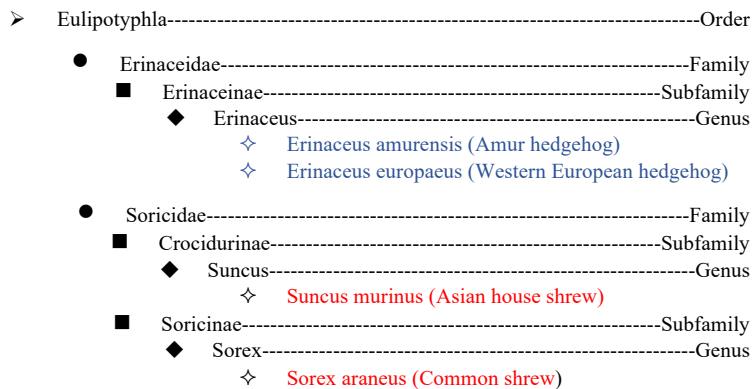
### *Mammalian Furin Sequence Alignment*

The multiple sequence alignment of the mammalian furin was performed by ClustalW (<https://www.genomse.jp/tools-bin/clustalw>) (Thompson et al., 1994).

The mammalian furin amino acid sequences were obtained from NCBI protein database (<https://www.ncbi.nlm.nih.gov/gene/5045/ortholog/?scope=7776>). The protein Accession numbers on NCBI database were: rat (NP\_001074923.1); dog (XP\_850069.2); pig (XP\_001929382.1); cat (XP\_023110662.1); giant panda (XP\_034523748.1); rabbit (XP\_002721548.2); camel (XP\_031296453.1); human (NP\_002560.1); cattle (XP\_024837365.1); horse (XP\_005602832.1); ferret (XP\_004763758.1); fruit bat (XP\_036087816.1); greater horseshoe bat (XP\_032955811.1); dolphin (XP\_004322334.1); shrew (XP\_004617631.1); hedgehog (XP\_016043802.1)

## Results and Discussion

The Eulipotyphlan relationships are depicted in Figure 1.



**Figure 1.** Eulipotyphlan relationships including coronavirus host species. The hedgehog species are colored in blue, and the shrew species are colored in red.

### *Eulipotyphlan Coronavirus Host Species Relationships*

Coronaviruses have been identified in hedgehogs and shrews in the order Eulipotyphla. The two hedgehog species are both classified in the genus *Erinaceus* belonging to the subfamily *Erinaceinae* under the family *Erinaceidae*. Western European hedgehogs included in the study were from Germany, France, Italy, and the UK (Corman et al., 2014; De Sabato et al., 2020; Monchatre-Leroy et al., 2017). Amur hedgehog (*Erinaceus amurensis*) coronaviruses were only detected in Jiangsu Province and Guangdong Province in China (Lau et al., 2019). The two shrew species both belong to the family *Soricidae* but in different subfamilies. Asian house shrews (*Suncus murinus*) are in the genus *Suncus* in the subfamily *Crocidurinae*, whose coronaviruses were discovered in Jiangxi Province and Zhejiang Province in China (Wang et al., 2017). The common shrews are within the genus *Sorex* and the subfamily *Soricinae*, whose coronaviruses were only detected in Tibet Autonomous Region in China (Wu et al., 2018).

### *Phylogenetic Analysis of the Spike Protein Amino Acid Sequences*

To understand the phylogenetic relationship between Eulipotyphla coronaviruses and other related coronaviruses, we constructed a maximum likelihood phylogenetic tree based on the spike amino acid sequences. In doing this, we also wanted to classify those strains with uncertain categories.

The classification represented in Figure 2 was obtained from the current NCBI taxonomy browser under the subfamily *Orthocoronavirinae*. All EriCoVs within the same species *Erinaceus europaeus* found in Italy, Germany, and the UK share an 89.96% to 100% spike amino acid sequence identity. All of the unclassified Italian EriCoVs are similar to the other Eri-CoV isolates from Germany and the UK with 100% bootstrap support. Therefore, these Italian EriCoVs are likely to be classified as *Merbecovirus*. The two Amur hedgehog coronavirus HKU31 strains from China form a single clade

distinct from the cluster of EriCoVs with 77.61% to 79.1% spike amino acid sequence identity. EriCoV and HKU31 spike proteins have an average identity of 55.84% and 56.6% with MERS-CoV spike protein, respectively. Therefore, it remains possible that the hedgehogs could harbor other MERS-like coronaviruses and transmit the virus to different species. The Merbecovirus genome sequences in humans possess around 65%-80% identity on average with other species, which also attributes to the efficient transmission, high diversity, and the significant amount of host species in Merbecovirus (Wong et al., 2019).

The amino acid sequence identities within WESVs are ranged from 90.05% to 99.74% through multiple sequence alignments. The group of WESV isolates forms a highly divergent group in the genus *Alphacoronavirus* based on the spike gene. The highest identity estimated between WESVs and other alphacoronaviruses is 34.68% with Rhinolophus bat coronavirus HKU2 from Rhinacovirus. Shrew-CoV/Tibet2014 itself, currently under unclassified Orthocoronavirinae, is a single clade under alphacoronavirus but distinct from WESV with only 25% identity of spike amino acid sequence. Its identity with other alphacoronaviruses is around 20% on average, which indicates its separate evolution in the genus.

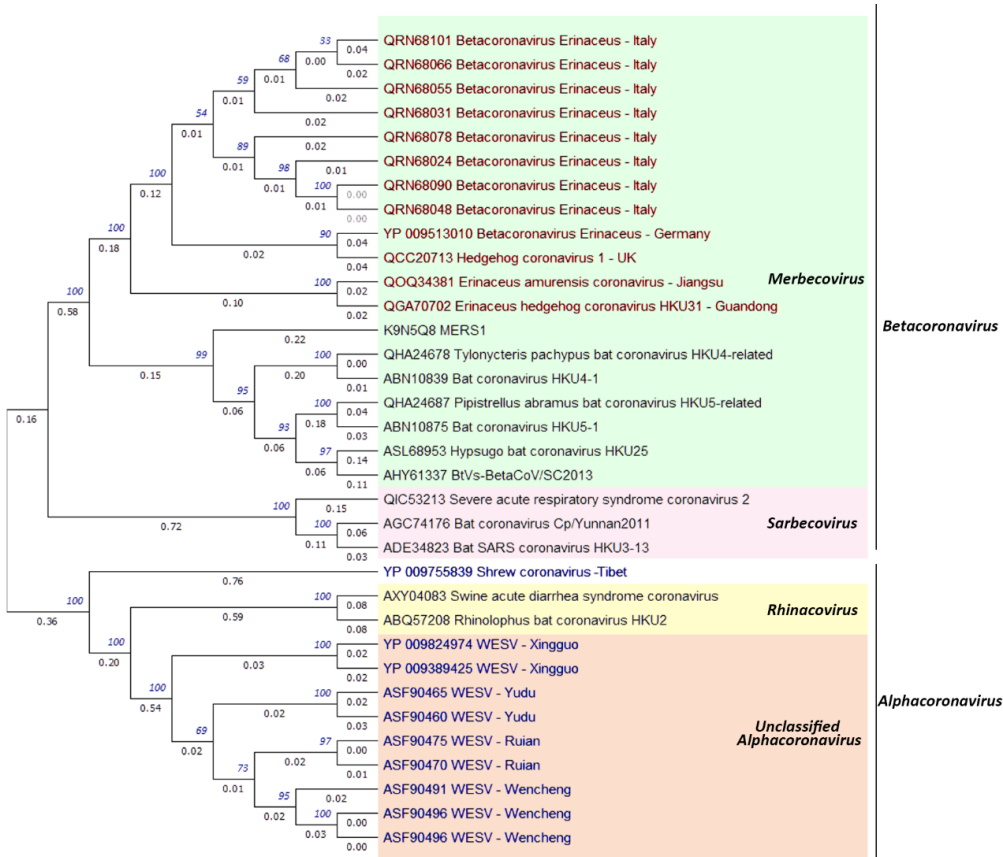
Bats have been considered ancestors of alphacoronaviruses and betacoronaviruses by many studies (Drexler et al., 2014; Huynh et al., 2012; Vijaykrishna et al., 2007; Woo et al., 2012). However, WESVs and Shrew-CoV/Tibet2014 express highly divergent spike proteins against other alphacoronaviruses. Therefore, the evolution and ancestry of alphacoronaviruses might be more complicated than the results from former studies. Furthermore, the discovery of the novel coronavirus, Lucheng Rn rat coronavirus, also indicates the complexity of the evolutionary history of alphacoronaviruses (Wang et al., 2015).

### ***Geographical Distribution of the Coronaviruses from the Mammalian Order Eulipotyphla***

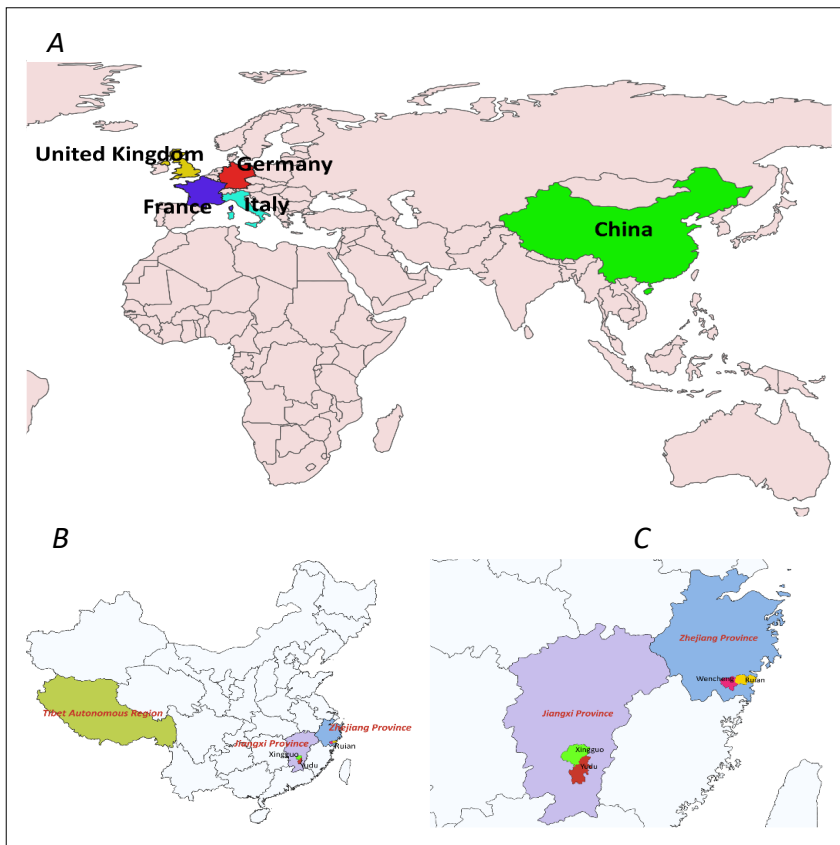
As the phylogenetic analysis illustrates, the WESV isolates are clustered based on their geographic origins, which indicates the potential of in situ evolution of coronaviruses in Asian house shrews (Figure 3). The Western European hedgehog coronaviruses are centered in western Europe because of their host distributions. The Eulipotyphla coronaviruses have only been discovered in western Europe and China for now, which might be attributed to these animal species distributions and the limited research. Importantly, Europe, mainland Southeast Asia, Central Africa, as well as Central and South America, have been considered regions with high biodiversity of zoonotic infectious disease mammalian host species (Han et al., 2016). More studies should focus on these small



mammals to shed light on the wildlife coronavirus evolution and zoonotic potential.



**Figure 2.** Maximum likelihood phylogenetic analysis of the hedgehog coronavirus stains, shrew coronavirus strains, and related alphacoronavirus and betacoronavirus strains based on spike protein amino acid sequences. The isolates colored in red are hedgehog coronaviruses and the isolates colored in blue are shrew coronaviruses

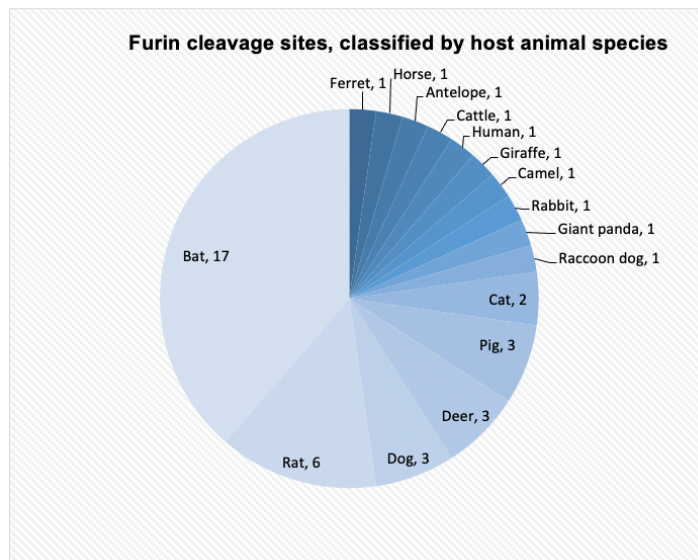


**Figure 3.** (A) Geographical distribution of the hedgehog and shrew (the mammalian order *Eulipotyphla*) coronavirus. The coronavirus strains are classified by countries: China (Wencheng Sm shrew coronavirus, Shrew coronavirus, Erinaceus hedgehog coronavirus HKU31, Erinaceus amurensis coronavirus); Germany (Betacoronavirus Erinaceus/VMC/DEU/2012); France (Erinaceus europaeus Alphacoronavirus); Italy (Betacoronavirus Erinaceus); UK (Betacoronavirus Erinaceus). (B) Geographical distribution of shrew coronaviruses. The shrew coronavirus strains are classified by provinces or regions: Tibet Autonomous Region (Shrew coronavirus); Zhejiang Province (Wencheng Sm shrew coronavirus); Jiangxi Province (Wencheng Sm shrew coronavirus). (C) Geographical distribution of Wencheng Sm shrew coronavirus. Wencheng Sm shrew coronavirus isolates are classified by counties: Yudu County (Yudu-76, Yudu-19); Xingguo County (Xingguo-101, Xingguo-74); Ruian County (Ruian-90, Ruian-133); Wencheng County (Wencheng-554, Wencheng-578, Wencheng-133).

## ***Furin cleavage Prediction***

As furin cleavage is often associated with spike protein activation and may be a key factor in zoonotic spillover and virus transmission, we analyzed 201 alphacoronavirus and betacoronavirus strains from the NCBI protein database for furin cleavage sites present on their spike proteins. Cleavage sites are found in 44 of them, with 17 hosted by bats and 27 by other mammals. The other 157 strains do not contain furin cleavage sites, with 143 hosted by bats and 14 by other mammals. The host animals and the number of coronaviruses they harbor are shown in Figure 4. The pie chart also includes some domestic animals that people have a big chance to contact, such as cats, dogs, pigs, horses, cattle, and rabbits. Therefore, we should pay attention to these animals as they have the potential to host transmissible coronaviruses.

We also examined the sequence of furin itself, as this can vary across animal species. Table 1 shows the pairwise sequence alignment result for furins across different animal species to see if there is any



**Figure 4.** The number of alphacoronaviruses and betacoronaviruses with furin cleavage sites corresponding to their host animals. Most viruses are hosted by bats (17 strains).

difference. The animals involved in the study are shrew, hedgehog, and the mammals shown in figure 4. Dolphin and greater horseshoe bat, whose furin proteases do not show any evidence to cleave the spike protein based on the surveillance study, are also included. The results show a high degree of conservation of furins across different animal species, which aligns with the previous study (El Najjar et al., 2015). The identities lower than 90% are between the camel and greater horseshoe bat, shrew and greater horseshoe bat, camel and horseshoe bat, and camel and hedgehog. Furins are ubiquitously expressed in various sites in all cells and tissues in eukaryotes, but little is known about whether the difference in furins across different species might affect their ability or efficiency to cleave the coronavirus spike protein (Hoffmann et al., 2018). Hence, more animals should be included, and more experiments need to be carried out to draw the conclusions (El Najjar et al., 2015; Huang et al., 2018).

**Table 1. Pairwise sequence alignment result of furin across different animal species represented by percentage.** Values in red are identities higher than 98%. Values in green are identities lower than 90%.

	human	camel	horse	dolphin	giant panda	ferret	cat	pig	cattle	greater horseshoe bat	fruit bat	rabbit	shrew	hedgehog
human		92.8	96.6	96.2	95.7	95.6	96.5	94.7	95.1	91.4	95.1	95.0	93.5	92.7
camel	92.8		93.0	92.6	91.7	91.4	92.0	91.4	92.4	87.6	91.7	90.5	90.4	88.7
horse	96.6	93.0		96.7	96.0	95.7	96.5	95.1	95.2	91.0	95.3	94.8	94.3	93.2
dolphin	96.2	92.6	96.7		96.0	95.4	96.5	95.5	95.6	90.8	94.6	94.2	93.7	93.2
giant panda	95.7	91.7	96.0	96.0		98.2	97.5	94.5	94.5	90.8	94.1	94.3	93.3	92.3
ferret	95.6	91.4	95.7	95.4	98.2		97.0	94.5	94.6	90.8	94.7	94.4	93.1	92.3
cat	96.5	92.0	96.5	96.5	97.5	97.0		95.0	95.2	91.2	94.7	94.7	93.6	93.5
pig	94.7	91.4	95.1	95.5	94.5	94.5	95.0		95.5	90.2	94.0	93.6	93.2	92.7
cattle	95.1	92.4	95.2	95.6	94.5	94.6	95.2	95.5		90.7	94.9	93.7	93.2	92.6
greater horseshoe bat	91.4	87.6	91.0	90.8	90.8	90.8	91.2	90.2	90.7		91.9	90.7	89.7	89.4
fruit bat	95.1	91.7	95.3	94.6	94.1	94.7	94.7	94.0	94.9	91.9		93.7	93.5	92.5
rabbit	95.0	90.5	94.8	94.2	94.3	94.4	94.7	93.6	93.7	90.7	93.7		92.9	92.5
shrew	93.5	90.4	94.3	93.7	93.3	93.1	93.6	93.2	93.2	89.7	93.5	92.9		92.1
hedgehog	92.7	88.7	93.2	93.2	92.3	92.3	93.5	92.7	92.6	89.4	92.5	92.5	92.1	

### *Prediction of Cleavage Sites in Hedgehog and Shrew Coronavirus Spike Proteins*

With the presence of the furin cleavage sites, the fusion activation of the coronavirus is likely to be enhanced with increased infectivity, which broadens the host tropism range and enables the virus to be more transmissible. We estimated the furin cleavage sites of all Eulipotyphla coronaviruses by two programs, ProP and PiTou. The results run by ProP are displayed in Table 2. However, the analysis from PiTou (Tian et al., 2012) indicated that none of the viral strains have furin cleavage sites (not shown). Considering the different algorithms between these two programs, we followed up the results by ProP, based on predicted cleavage for furin family members. The general PC prediction stands for the general proprotein convertase cleavage sites prediction. The typical furin cleavage sites occur around S1/S2 (containing multiple arginines around site 700, such as **RXXR**) and S2' (containing one or two arginines and the second **R** would be essential around 890) (Hoffmann et al., 2020; Kleine-Weber et al., 2018). For WESVs, six of them have positive ProP prediction results around site 513 and site 1121, which are not the typical S1/S2 or S2' sites, so these sites might be cleaved by other proteases or be false positives. Yudu-76, Yudu-19, Ruian-90, Ruian-133, Xingguo-101, and Xingguo-74 isolates have a common potential cleavage site (**KR|S** around site 512) cleaved by proprotein convertases based on the general PC protection. The high scores in Yudu and Xingguo isolates suggest a high potential for the presence of these cleavage sites. They are likely to be cleaved by furin, but at atypical sites or by other furin-like proprotein convertases (PCs). Around 490 potential furin cleavage protein candidates were identified in the human proteome (Shiryayev et al., 2013). The PC family contains 7 members in human, comprising furin, PC1/3, PC2, PC4, PACE4,

PC5/6, and PC7 (Remacle et al., 2008). PC5, PC7, and PACE4 are considered furin-like PCs. PACE4 is less likely to be the candidate for WESV cleavage sites because it cleaves at **KR** motifs less frequently than **RXK/RR** or **RXRR** (Gordon et al., 1997). PC5 has been demonstrated to cleave the SARS-CoV spike protein efficiently and it can cleave at **KR motifs**, indicating its potential to cleave WESV spike protein at **KR|S** (Basak et al., 2007). Though no studies have displayed the PC7 ability in cleaving coronavirus spike protein, it can cleave the unusual site **KR|S** (Corman et al., 2014). The endoproteases, cathepsin L and TMPRSS2, which can activate the SARS-CoV, MERS-CoV, and SARS-CoV-2 spike proteins, also have a likelihood to activate the WESV spike protein (Hoffmann et al., 2018; Matsuyama et al., 2020; Sacco et al., 2020). The studies on atypical furin cleavage sites and the corresponding PCs to specific viral proteins are still limited. Therefore, the spike protein model for WESV was further constructed to examine their furin cleavage site possibilities.

**Table 2. Prediction of cleavage sites in hedgehog and shrew coronavirus spike proteins.** basic amino acids corresponding to positive scores above threshold (> 0.6) are shown in bold text

<i>Virus Strain</i>	<i>ProP furin prediction</i>	<i>General PC prediction<sup>1</sup></i>
WESV Yudu-76	513 NQIDLKR SV 0.505 1121 CIRRSR DS 0.564	513 NQIDL <b>KR</b>  SV <b>0.946</b>
WESV Yudu-19	513 NQIDLKR SV 0.583 1121 CIRRSR DS 0.564	513 NQIDL <b>KR</b>  SV <b>0.940</b>
WESV Ruian-90	No cleavage site	512 NQIDTKR SV 0.593
WESV Ruian-133	No cleavage site	512 NQIDTKR SV 0.593
WESV Wencheng-554	512 TQIDTKR SV 0.506	321 NNNLQLR GP 0.504
WESV Wencheng-578	No cleavage site	321 NNNLQLR GP 0.504
WESV Wencheng-562	512 TQIDTKR SV 0.506	321 NNNLQLR GP 0.504
WESV Xingguo-101	512 NQIDTKR SV 0.543	512 NQIDT <b>KR</b>  SV <b>0.885</b>
WESV Xingguo-74	512 NQIDTKR SV 0.543 1120 CIRRSR DS 0.564	512 NQIDT <b>KR</b>  SV <b>0.885</b>
Shrew coronavirus Shrew-CoV/Tibet2014	No cleavage site	139 VNKTLTR TT 0.561
Erinaceus hedgehog coronavirus HKU31	No cleavage site	No cleavage site
Erinaceus hedgehog coronavirus cw_6c	No cleavage site	No cleavage site
Betacoronavirus	No cleavage site	No cleavage site
Erinaceus/VMC/DEU/2012	No cleavage site	No cleavage site
Hedgehog coronavirus 1	No cleavage site	1216 ELEEFYR NI 0.632
ErinaceusCoV/Italy/116988-1/2018	No cleavage site	323 VNGYITR AV 0.561 867 TSNSQYR SA 0.571
ErinaceusCoV/Italy/50265-19/2018	No cleavage site	324 VNGYITR AV 0.561
ErinaceusCoV/Italy/50265-1/2018	No cleavage site	323 VNGYITR AV 0.561
ErinaceusCoV/Italy/50265-17/2018	No cleavage site	324 VNGYITR AV 0.561 1219 ELEDFYR NI 0.703
ErinaceusCoV/Italy/50265-11/2019	No cleavage site	324 VNGYITR AV 0.561
ErinaceusCoV/Italy/50265-12/2019	No cleavage site	323 VNGYITR AV 0.561
ErinaceusCoV/Italy/50265-13/2019	No cleavage site	323 VNGYITR AV 0.561
ErinaceusCoV/Italy/50265-15/2019	No cleavage site	323 VNGYITR AV 0.561

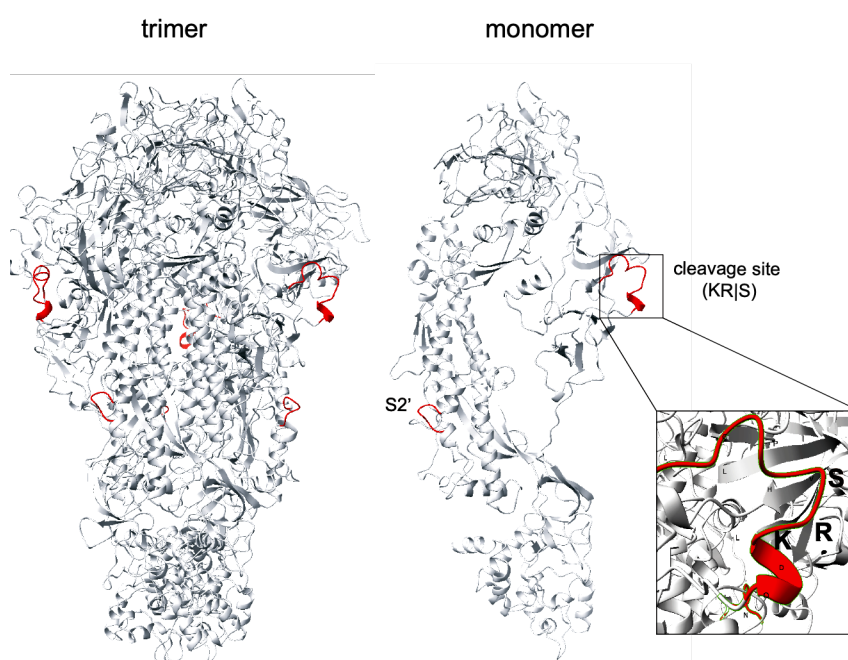
<sup>1</sup> PC: proprotein convertase;

## ***Spike Protein Multiple Sequence Alignment***

To better define the position of the putative cleavage sites, two representative WESV spikes (Ruian-90 and Yudu-76) showing high PC consensus sequence scores were aligned using MUSCLE alignment, and compared to MERS-CoV as a representative zoonotic coronavirus (see Appendix). Based on this, it is apparent that the putative PC cleavage sites do not align with the expected S1/S2 spike cleavage site, but rather occur in an upstream location within S1—being positioned broadly in line with a cleavage position previously identified in an unrelated rodent coronavirus, AcCoV-JC34 (discussed in Choi et al, 2021). As expected the conserved S2' cleavage site was adjacent to the fusion peptide.

## ***Spike Protein Model Construction***

To further examine the putative PC cleavage sites identified based on amino acid sequences, a WESV Yudu 76 spike protein structural model was constructed based on the HKU2 spike, which shares relatively high identity with WESVs (Figure 5). The cleavage site depicted in the structural model is clearly an exposed loop, which indicates it is in an appropriate environment to be accessible by proteases. One limitation of our work is that the amino acid identity of the reference model (*Rhinolophus* ba coronavirus HKU2) with WESV is relatively low (36%); however, we consider this to be a reliable model to illustrate the presence of a cleavage site—which, while it does not represent a typical furin cleavage site, may be cleaved by other furin-like proteases or trypsin-like proteases. Further biochemical and in vivo assays are needed to determine the presence of functional cleavage sites for these viruses.



**Figure 5. Wencheng Sm shrew coronavirus spike protein structure model, as a trimer and as a monomer. The model was constructed based on HKU2 spike glycoprotein (6M15). The predicted cleavage loop (QIDLKRSVHPVGL) is shown in red, with the proposed KR|S cleavage site indicated in the inset. The S2' site and cleavage loop is also indicated.**

## ***One Health Perspectives to Understand the Zoonotic Potential of Eulipotyphla Coronaviruses***

Increasing zoonotic infectious diseases are affecting humans, domestic animals, and wildlife, in which case One Health approaches should be considered to reduce the risk of infectious diseases (Cunningham et al., 2017; Daszak et al., 2000). Here we integrated the factors involving the environment, wild animals, domestic animals, humans, and pathogens to discuss their interactions and the corresponding strategies to prevent the emerging coronaviruses.

Shrews are widely distributed in Asia, Africa, and Europe, sharing the same habitats with other insectivorous mammals and rodents (Bown et al., 2011). They tend to live in the rural and forested areas as well as the urban sewer drains (Rahman et al., 2018). Western European hedgehogs are endemic to the UK and Western Europe and mostly distributed in the gardens, woodlands, and grasslands (Hof et al., 2019). Amur hedgehogs have been found in China, Korea, and Russia so far (Ai et al., 2018; Lau et al., 2019). Their habitats include woodlands, grasslands, and also forest edges (Smith et al., 2010). Though many of these small mammals live in wild areas, the number of people who raise exotic companion mammals as pets is increasing all over the world (Rahman et al., 2018). Therefore, people need to be aware of the potential risk that exotic companion mammals like hedgehogs might carry a range of viruses when adopting them (Keeble & Koterwas, 2020).

The prevalence of coronaviruses in wildlife is not very high, due to the low density of animals caused by the dilution effect and restricted interspecies transmissions (Johnson & Thielges, 2010; Zappulli et al., 2020). Wild animals require a certain receptor to harbor the virus, limiting the spread of the virus (Ji et al., 2020). The majority of the coronavirus wildlife host populations are rodents, insectivores like bats, and other mammals such as raccoon dogs, camels, pangolins, etc. (Irving et al., 2021). Typically, there are two wildlife-human interaction pathways. One pathway is that wild small mammals such as rodents might leave feces on crops, and farmers might get infected while handling these contaminated crops; the other is through the wildlife food supply trade, which is more prevalent (Huong et al., 2020). The initial SARS-CoV-2 transmissions were closely related to Wuhan wet food market (Tiwari et al., 2020). Various kinds of small mammals were sold either directly to the customers or to the restaurants, including hedgehogs, rodents, bats, pangolins and ferrets, which are all highly potential intermediate coronavirus hosts (Mahdy et al., 2020). Although shrews have not been demonstrated to be sold in the wet market, evidence has showed their presence with the rodents in the wet food market because of the poor sanitation and regulation (Tung et al., 2013).

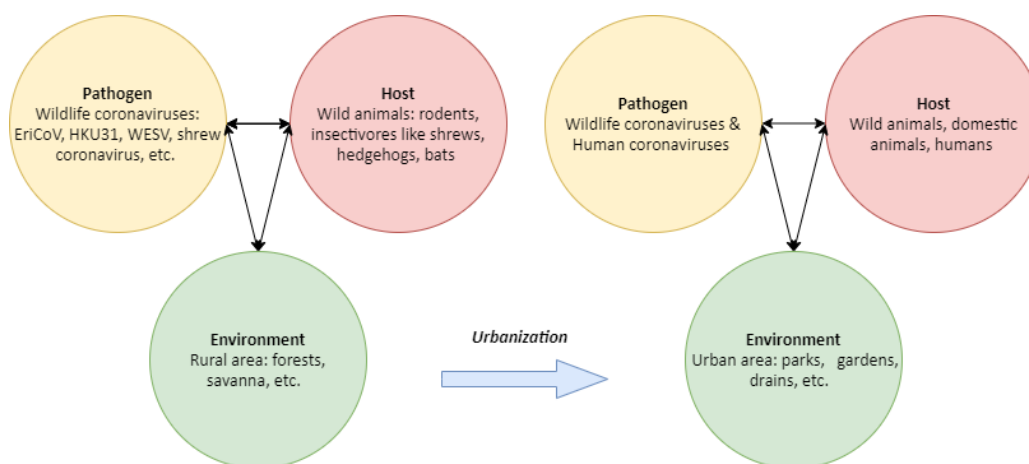
A study on the viral infection persistence in temperate bats during their hibernation has shown that the coronavirus can persist for up to four months in laboratory setting (Subudhi et al., 2017). The reason might be that the bats conserve their energy and enter torpor with a low metabolic rate as well as low levels of inflammation during hibernation (Calisher et al., 2006; Subudhi et al., 2017).

However, whether the wild bats have the equivalent ability of viral shedding is still not fully understood (Plowright et al., 2015). Hedgehogs also have the habits of hibernation during winter and share the nocturnal habits and insectivorous diets with bats (South et al., 2020). Therefore, we can speculate a high risk in hedgehogs to harbor and maintain the coronavirus during their hibernation, which raises the probability to transmit and circulate the virus among their living populations. Although there is no evidence that hedgehog and shrew coronaviruses are linked to any diseases, these mammals are considered hosts and potential natural reservoirs for coronaviruses. They might hold an important position in the virus evolution and have the ability to host newly emerging coronaviruses due to recombination and mutation of the coronaviruses.

Human activities, such as deforestation and conversion from land to agricultural use, drive the wild animals to move from rural areas to urban areas (McMahon et al., 2018). Climate change might also increase the insectivore populations feeding on insects due to the expanded insects (Morueta-Holme et al., 2010; Vega & Castro, 2019). More and more zoonotic epidemics occur due to these ecosystem changes, including the most recent SARS-CoV-2 and Ebola (Cunningham et al., 2017).

As a consequence of urbanization, humans are closer to wild animals than ever. For instance, Western European hedgehogs have been mostly found in backyards, urban grasslands, and woodlands (Hof et al., 2019). The interaction between domestic animals and wild animals is also becoming more frequent. For example, cats (the predators of the rodents and shrews), are likely to carry these wildlife coronaviruses during predation (Tsoleridis et al., 2016). The risk of zoonotic spillover is also elevated, leading to the spread of zoonotic epidemics (Hassani & Khan, 2020; Wang & Anderson, 2019). In Figure 6, an epidemiology triad was constructed to describe the process of how wildlife animal coronaviruses jump to humans due to urbanization.

**Figure 6. Epidemiology triad model before and after urbanization.**



## Conclusions and Future Directions

This study conducted a bioinformatic analysis of the coronaviruses from the mammalian order Eulipotyphla, including hedgehogs and shrews. Shrew coronavirus and WESVs might diverge from a different ancestry compared to other alphacoronaviruses, revealing a more complicated alphacoronavirus evolution history than was previously thought. The hedgehog coronaviruses EriCoV and HKU31 are within the same subgenus Merbecovirus as MERS-CoV. WESV shows a PC cleavage site at **KR|S** around 512 site, indicating its zoonotic potential. The potential cleavage proteases could be PC5, PC7, and endoproteases other than furin. The animals, whose furins were demonstrated to have the ability to cleave the spike protein, are at a higher risk to host and transmit the coronaviruses, including bats, rats, dogs, deer, pig, cats, raccoon dogs, giant pandas, rabbits, camels, giraffes, cattle, horses, antelopes, and ferrets. No significance was found within the furins from a wide range of animals, but more studies are needed to demonstrate furin cleavage abilities. In conclusion, hedgehogs and shrews are natural reservoirs for coronaviruses, which might be able to recombine with other wildlife coronaviruses and generate new emerging viruses. As novel coronaviruses causes a heavy burden to global health, One Health perspectives and approaches are necessary to understand the evolution and transmission of the these viruses. Growing urbanization increases the wildlife to human interactions, and the small mammals sold in wet market might also bring the pathogen spillover risks. Therefore, people should be vigilant of small wild mammals, and take precautions to control and prevent contact with these animals.

## ACKNOWLEDGEMENTS

We thank Michael Stanhope for invaluable help and advice in generating and analysing the data presented in this manuscript. A.C. is supported by grant T32EB023860 from the National Institute of Biomedical Imaging and Bioengineering. Work on coronavirus entry in the Whittaker lab is funded in part by the National Institute of Health research grant R01AI35270.



## References

- Ai, H.-S., He, K., Chen, Z.-Z., Li, J.-Q., Wan, T., Li, Q., Nie, W.-H., Wang, J.-H., Su, W.-T., & Jiang, X.-L. (2018). Taxonomic revision of the genus *Mesechinus* (Mammalia: Erinaceidae) with description of a new species. *Zoological Research*, 39(5), 335–347. <https://doi.org/10.24272/j.issn.2095-8137.2018.034>
- Basak, A., Mitra, A., Basak, S., Pasko, C., Chrétien, M., & Seaton, P. (2007). A Fluorogenic Peptide Containing the Processing Site of Human SARS Corona Virus S-Protein: Kinetic Evaluation and NMR Structure Elucidation. *Chembiochem*, 8(9), 1029–1037. <https://doi.org/10.1002/cbic.200700007>
- Belouzard, S., Chu, V. C., & Whittaker, G. R. (2009). Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proceedings of the National Academy of Sciences*, 106(14), 5871–5876. <https://doi.org/10.1073/pnas.0809524106>
- Boni, M. F., Lemey, P., Jiang, X., Lam, T. T.-Y., Perry, B. W., Castoe, T. A., Rambaut, A., & Robertson, D. L. (2020). Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nature Microbiology*, 5(11), 1408–1417. <https://doi.org/10.1038/s41564-020-0771-4>
- Bown, K. J., Lambin, X., Telford, G., Heyder-Bruckner, D., Ogden, N. H., & Birtles, R. J. (2011). The Common Shrew (*Sorex araneus*): A Neglected Host of Tick-Borne Infections? *Vector-Borne and Zoonotic Diseases*, 11(7), 947–953. <https://doi.org/10.1089/vbz.2010.0185>
- Bray, D. P., Bennett, M., Stockley, P., Hurst, J. L., & Kipar, A. (2008). Composition and function of haemolymphatic tissues in the European common shrew. *PloS One*, 3(10), e3413. <https://doi.org/10.1371/journal.pone.0003413>
- Brian, D. A., & Baric, R. S. (2005). Coronavirus Genome Structure and Replication. In L. Enjuanes (Ed.), *Coronavirus Replication and Reverse Genetics* (Vol. 287, pp. 1–30). Springer Berlin Heidelberg. [https://doi.org/10.1007/3-540-26765-4\\_1](https://doi.org/10.1007/3-540-26765-4_1)
- Calisher, C. H., Childs, J. E., Field, H. E., Holmes, K. V., & Schountz, T. (2006). Bats: Important Reservoir Hosts of Emerging Viruses. *Clinical Microbiology Reviews*, 19(3), 531–545. <https://doi.org/10.1128/CMR.00017-06>
- Choi, A., Singleton, D.T., Stout, A.E., Millet, J.K. and Whittaker, G.R. (2021) In vitro and computational analysis of the putative furin cleavage site (RRARS) in the divergent spike protein of the rodent coronavirus AcCoV-JC34 (sub-genus luchacovirus). *bioRxiv*, <https://www.biorxiv.org/content/10.1101/2021.12.16.473025v1>
- Corman, Victor M., Muth, D., Niemeyer, D., & Drosten, C. (2018). Hosts and Sources of Endemic Human Coronaviruses. *Advances in Virus Research*, 100, 163–188. <https://doi.org/10.1016/bs.aivir.2018.01.001>
- Corman, Victor Max, Baldwin, H. J., Tateno, A. F., Zerbinati, R. M., Annan, A., Owusu, M., Nkrumah, E. E., Maganga, G. D., Oppong, S., Adu-Sarkodie, Y., Vallo, P., da Silva Filho, L. V. R. F., Leroy, E. M., Thiel, V., van der Hoek, L., Poon, L. L. M., Tschapka, M., Drosten, C., & Drexler, J. F. (2015). Evidence for an Ancestral Association of Human Coronavirus 229E with Bats. *Journal of Virology*, 89(23), 11858–11870. <https://doi.org/10.1128/JVI.01755-15>
- Corman, Victor Max, Kallies, R., Philipps, H., Göpner, G., Müller, M. A., Eckerle, I., Brünink, S., Drosten, C., & Drexler, J. F. (2014). Characterization of a novel betacoronavirus related to middle East respiratory syndrome coronavirus in European hedgehogs. *Journal of Virology*, 88(1), 717–724. <https://doi.org/10.1128/JVI.01600-13>
- Cui, J., Li, F., & Shi, Z.-L. (2019). Origin and evolution of pathogenic coronaviruses. *Nature Reviews. Microbiology*, 17(3), 181–192. <https://doi.org/10.1038/s41579-018-0118-9>
- Cunningham, A. A., Daszak, P., & Wood, J. L. N. (2017). One Health, emerging infectious diseases and wildlife: Two decades of progress? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1725). <https://doi.org/10.1098/rstb.2016.0167>
- Daszak, P., Cunningham, A. A., & Hyatt, A. D. (2000). Emerging Infectious Diseases of Wildlife—Threats to Biodiversity and Human Health. *Science*, 287(5452), 443–449. <https://doi.org/10.1126/science.287.5452.443>
- De Sabato, L., Di Bartolo, I., De Marco, M. A., Moreno, A., Lelli, D., Cotti, C., Delogu, M., & Vaccari, G. (2020). Can Coronaviruses Steal Genes from the Host as Evidenced in Western European Hedgehogs by EriCoV Genetic Characterization? *Viruses*, 12(12). <https://doi.org/10.3390/v12121471>
- Decaro, N., & Lorusso, A. (2020). Novel human coronavirus (SARS-CoV-2): A lesson from animal coronaviruses. *Veterinary Microbiology*, 244, 108693. <https://doi.org/10.1016/j.vetmic.2020.108693>
- Drexler, J. F., Corman, V. M., & Drosten, C. (2014). Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. *Antiviral Research*, 101, 45–56. <https://doi.org/10.1016/j.antiviral.2013.10.013>

- Duckert, P., Brunak, S., & Blom, N. (2004). Prediction of proprotein convertase cleavage sites. *Protein Engineering, Design and Selection*, 17(1), 107–112. <https://doi.org/10.1093/protein/gzh013>
- El Najjar, F., Lampe, L., Baker, M. L., Wang, L.-F., & Dutch, R. E. (2015). Analysis of cathepsin and furin proteolytic enzymes involved in viral fusion protein activation in cells of the bat reservoir host. *PLoS One*, 10(2), e0115736. <https://doi.org/10.1371/journal.pone.0115736>
- Gordon, V. M., Rehemtulla, A., & Leppla, S. H. (1997). A role for PACE4 in the proteolytic activation of anthrax toxin protective antigen. *Infection and Immunity*, 65(8), 3370–3375. <https://doi.org/10.1128/IAI.65.8.3370-3375.1997>
- Han, B. A., Kramer, A. M., & Drake, J. M. (2016). Global Patterns of Zoonotic Disease in Mammals. *Trends in Parasitology*, 32(7), 565–577. <https://doi.org/10.1016/j.pt.2016.04.007>
- Hassani, A., & Khan, G. (2020). Human-Animal Interaction and the Emergence of SARS-CoV-2. *JMIR Public Health and Surveillance*, 6(4). <https://doi.org/10.2196/22117>
- Hedman, H. D., Krawczyk, E., Helmy, Y. A., Zhang, L., & Varga, C. (2021). Host Diversity and Potential Transmission Pathways of SARS-CoV-2 at the Human-Animal Interface. *Pathogens*, 10(2). <https://doi.org/10.3390/pathogens10020180>
- Hillis, D. M., & Bull, J. J. (1993). An Empirical Test of Bootstrapping as a Method for Assessing Confidence in Phylogenetic Analysis. *Systematic Biology*, 42(2), 182–192. <https://doi.org/10.1093/sysbio/42.2.182>
- Hof, A. R., Allen, A. M., & Bright, P. W. (2019). Investigating the Role of the Eurasian Badger (*Meles meles*) in the Nationwide Distribution of the Western European Hedgehog (*Erinaceus europaeus*) in England. *Animals : An Open Access Journal from MDPI*, 9(10). <https://doi.org/10.3390/ani9100759>
- Hoffmann, M., Hofmann-Winkler, H., & Pöhlmann, S. (2018). Priming Time: How Cellular Proteases Arm Coronavirus Spike Proteins. In E. Böttcher-Friebertshäuser, W. Garten, & H. D. Klenk (Eds.), *Activation of Viruses by Host Proteases* (pp. 71–98). Springer International Publishing. [https://doi.org/10.1007/978-3-319-75474-1\\_4](https://doi.org/10.1007/978-3-319-75474-1_4)
- Huang, Y., Long, Y., Li, S., Lin, T., Wu, J., Zhang, Y., & Lin, Y. (2018). Investigation on the processing and improving the cleavage efficiency of furin cleavage sites in *Pichia pastoris*. *Microbial Cell Factories*, 17. <https://doi.org/10.1186/s12934-018-1020-x>
- Huong, N. Q., Nga, N. T. T., Long, N. V., Luu, B. D., Latinne, A., Pruvot, M., Phuong, N. T., Quang, L. T. V., Hung, V. V., Lan, N. T., Hoa, N. T., Minh, P. Q., Diep, N. T., Tung, N., Ky, V. D., Robertson, S. I., Thuy, H. B., Long, N. V., Gilbert, M., ... Olson, S. H. (2020). Coronavirus testing indicates transmission risk increases along wildlife supply chains for human consumption in Viet Nam, 2013-2014. *PLoS ONE*, 15(8). <https://doi.org/10.1371/journal.pone.0237129>
- Huynh, J., Li, S., Yount, B., Smith, A., Sturges, L., Olsen, J. C., Nagel, J., Johnson, J. B., Agnihothram, S., Gates, J. E., Frieman, M. B., Baric, R. S., & Donaldson, E. F. (2012). Evidence Supporting a Zoonotic Origin of Human Coronavirus Strain NL63. *Journal of Virology*, 86(23), 12816–12825. <https://doi.org/10.1128/JVI.00906-12>
- Irving, A. T., Ahn, M., Goh, G., Anderson, D. E., & Wang, L.-F. (2021). Lessons from the host defences of bats, a unique viral reservoir. *Nature*, 589(7842), 363–370. <https://doi.org/10.1038/s41586-020-03128-0>
- Ji, W., Wang, W., Zhao, X., Zai, J., & Li, X. (2020). Cross-species transmission of the newly identified coronavirus 2019-nCoV. *Journal of Medical Virology*, 92(4), 433–440. <https://doi.org/10.1002/jmv.25682>
- Johnson, P. T. J., & Thielges, D. W. (2010). Diversity, decoys and the dilution effect: How ecological communities affect disease risk. *Journal of Experimental Biology*, 213(6), 961–970. <https://doi.org/10.1242/jeb.037721>
- Keeble, E., & Koterwas, B. (2020). Selected Emerging Diseases of Pet Hedgehogs. *Veterinary Clinics of North America: Exotic Animal Practice*, 23(2), 443–458. <https://doi.org/10.1016/j.cvex.2020.01.010>
- Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, 35(6), 1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Lau, S. K. P., Luk, H. K. H., Wong, A. C. P., Fan, R. Y. Y., Lam, C. S. F., Li, K. S. M., Ahmed, S. S., Chow, F. W. N., Cai, J.-P., Zhu, X., Chan, J. F. W., Lau, T. C. K., Cao, K., Li, M., Woo, P. C. Y., & Yuen, K.-Y. (2019). Identification of a Novel Betacoronavirus (Merbecovirus) in Amur Hedgehogs from China. *Viruses*, 11(11). <https://doi.org/10.3390/v11110980>
- Letko, M., Miazgowiec, K., McMinn, R., Seifert, S. N., Sola, I., Enjuanes, L., Carmody, A., van Doremalen, N., & Munster, V. (2018). Adaptive Evolution of MERS-CoV to Species Variation in DPP4. *Cell Reports*, 24(7), 1730–1737.

<https://doi.org/10.1016/j.celrep.2018.07.045>

Li, F. (2016). Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annual Review of Virology*, 3(1), 237–261. <https://doi.org/10.1146/annurev-virology-110615-042301>

Mahdy, M. A. A., Younis, W., & Ewaida, Z. (2020). An Overview of SARS-CoV-2 and Animal Infection. *Frontiers in Veterinary Science*, 7. <https://doi.org/10.3389/fvets.2020.596391>

Matsuyama, S., Nao, N., Shirato, K., Kawase, M., Saito, S., Takayama, I., Nagata, N., Sekizuka, T., Katoh, H., Kato, F., Sakata, M., Tahara, M., Kutsuna, S., Ohmagari, N., Kuroda, M., Suzuki, T., Kageyama, T., & Takeda, M. (2020). Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proceedings of the National Academy of Sciences of the United States of America*, 117(13), 7001–7003. <https://doi.org/10.1073/pnas.2002589117>

McMahon, B. J., Morand, S., & Gray, J. S. (2018). Ecosystem change and zoonoses in the Anthropocene. *Zoonoses and Public Health*, 65(7), 755–765. <https://doi.org/10.1111/zph.12489>

Millet, J. K., & Whittaker, G. R. (2014). Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proceedings of the National Academy of Sciences of the United States of America*, 111(42), 15214–15219. <https://doi.org/10.1073/pnas.1407087111>

Monchatre-Leroy, E., Boué, F., Boucher, J.-M., Renault, C., Moutou, F., Ar Gouilh, M., & Umhang, G. (2017). Identification of Alpha and Beta Coronavirus in Wildlife Species in France: Bats, Rodents, Rabbits, and Hedgehogs. *Viruses*, 9(12). <https://doi.org/10.3390/v9120364>

Morueta-Holme, N., Fløjgaard, C., & Svenning, J.-C. (2010). Climate Change Risks and Conservation Implications for a Threatened Small-Range Mammal Species. *PLoS ONE*, 5(4). <https://doi.org/10.1371/journal.pone.0010360>

Plowright, R. K., Eby, P., Hudson, P. J., Smith, I. L., Westcott, D., Bryden, W. L., Middleton, D., Reid, P. A., McFarlane, R. A., Martin, G., Tabor, G. M., Skerratt, L. F., Anderson, D. L., Crameri, G., Quammen, D., Jordan, D., Freeman, P., Wang, L.-F., Epstein, J. H., ... McCallum, H. (2015). Ecological dynamics of emerging bat virus spillover. *Proceedings of the Royal Society B: Biological Sciences*, 282(1798). <https://doi.org/10.1098/rspb.2014.2124>

Rabaan, A. A., Al-Ahmed, S. H., Haque, S., Sah, R., Tiwari, R., Malik, Y. S., Dhama, K., Yattoo, M. I., Bonilla-Aldana, D. K., & Rodriguez-Morales, A. J. (2020). SARS-CoV-2, SARS-CoV, and MERS-COV: A comparative overview. *Le Infezioni in Medicina*, 28(2), 174–184.

Rahman, M., Islam, S., Masuduzzaman, Md., Alam, M., Chawdhury, M. N. U., Ferdous, J., Islam, Md. N., Hassan, M. M., Hossain, M. A., & Islam, A. (2018). Prevalence and diversity of gastrointestinal helminths in free-ranging Asian house shrew (*Suncus murinus*) in Bangladesh. *Veterinary World*, 11(4), 549–556. <https://doi.org/10.14202/vetworld.2018.549-556>

Remacle, A. G., Shiryayev, S. A., Oh, E.-S., Cieplak, P., Srinivasan, A., Wei, G., Liddington, R. C., Ratnikov, B. I., Parent, A., Desjardins, R., Day, R., Smith, J. W., Lebl, M., & Strongin, A. Y. (2008). Substrate Cleavage Analysis of Furin and Related Proprotein Convertases. *The Journal of Biological Chemistry*, 283(30), 20897–20906. <https://doi.org/10.1074/jbc.M803762200>

Sacco, M. D., Ma, C., Lagarias, P., Gao, A., Townsend, J. A., Meng, X., Dube, P., Zhang, X., Hu, Y., Kitamura, N., Hurst, B., Tarbet, B., Marty, M. T., Kolocouris, A., Xiang, Y., Chen, Y., & Wang, J. (2020). Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against Mpro and cathepsin L. *Science Advances*, 6(50). <https://doi.org/10.1126/sciadv.abe0751>

Shiryayev, S. A., Chernov, A. V., Golubkov, V. S., Thomsen, E. R., Chudin, E., Chee, M. S., Kozlov, I. A., Strongin, A. Y., & Cieplak, P. (2013). High-Resolution Analysis and Functional Mapping of Cleavage Sites and Substrate Proteins of Furin in the Human Proteome. *PLoS ONE*, 8(1). <https://doi.org/10.1371/journal.pone.0054290>

Smith, A. T., Xie, Y., Hoffmann, R. S., Lunde, D., MacKinnon, J., Wilson, D. E., & Wozencraft, W. C. (2010). *A Guide to the Mammals of China*. Princeton University Press.

Song, Y., DiMaio, F., Wang, R.Y., Kim, D., Miles, C., Brunette, T.J., Thompson, J. and Baker, D. (2013) High resolution comparative modeling with RosettaCM. *Structure* 21:1735-42.

South, K. E., Haynes, K., & Jackson, A. C. (2020). Hibernation Patterns of the European Hedgehog, *Erinaceus europaeus*, at a Cornish Rescue Centre. *Animals : An Open Access Journal from MDPI*, 10(8). <https://doi.org/10.3390/ani10081418>

Subudhi, S., Rapin, N., Bollinger, T. K., Hill, J. E., Donaldson, M. E., Davy, C. M., Warnecke, L., Turner, J. M., Kyle,

- C. J., Willis, C. K. R., & Misra, V. (2017). A persistently infecting coronavirus in hibernating *Myotis lucifugus*, the North American little brown bat. *The Journal of General Virology*, *98*(9), 2297–2309. <https://doi.org/10.1099/jgv.0.000898>
- Tao, Y., Shi, M., Chommanard, C., Queen, K., Zhang, J., Markotter, W., Kuzmin, I. V., Holmes, E. C., & Tong, S. (2017). Surveillance of Bat Coronaviruses in Kenya Identifies Relatives of Human Coronaviruses NL63 and 229E and Their Recombination History. *Journal of Virology*, *91*(5). <https://doi.org/10.1128/JVI.01953-16>
- Tian, S., Huajun, W., & Wu, J. (2012). Computational prediction of furin cleavage sites by a hybrid method and understanding mechanism underlying diseases. *Scientific Reports*, *2*. <https://doi.org/10.1038/srep00261>
- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*(22), 4673–4680. <https://doi.org/10.1093/nar/22.22.4673>
- Tiwari, R., Dhama, K., Sharun, K., Iqbal Yattoo, Mohd., Malik, Y. S., Singh, R., Michalak, I., Sah, R., Bonilla-Aldana, D. K., & Rodriguez-Morales, A. J. (2020). COVID-19: Animals, veterinary and zoonotic links. *The Veterinary Quarterly*, *40*(1), 169–182. <https://doi.org/10.1080/01652176.2020.1766725>
- Tsagkogeorga, G., Parker, J., Stupka, E., Cotton, J. A., & Rossiter, S. J. (2013). Phylogenomic Analyses Elucidate the Evolutionary Relationships of Bats. *Current Biology*, *23*(22), 2262–2267. <https://doi.org/10.1016/j.cub.2013.09.014>
- Tsoleridis, T., Onianwa, O., Horncastle, E., Dayman, E., Zhu, M., Danjitrong, T., Wachtl, M., Behnke, J. M., Chapman, S., Strong, V., Dobbs, P., Ball, J. K., Tarlinton, R. E., & McClure, C. P. (2016). Discovery of Novel Alphacoronaviruses in European Rodents and Shrews. *Viruses*, *8*(3). <https://doi.org/10.3390/v8030084>
- Tung, K.-C., Hsiao, F.-C., Wang, K.-S., Yang, C.-H., & Lai, C.-H. (2013). Study of the endoparasitic fauna of commensal rats and shrews caught in traditional wet markets in Taichung City, Taiwan. *Journal of Microbiology, Immunology, and Infection = Wei Mian Yu Gan Ran Za Zhi*, *46*(2), 85–88. <https://doi.org/10.1016/j.jmii.2012.01.012>
- van Boheemen, S., de Graaf, M., Lauber, C., Bestebroer, T. M., Raj, V. S., Zaki, A. M., Osterhaus, A. D. M. E., Haagmans, B. L., Gorbalenya, A. E., Snijder, E. J., & Fouchier, R. A. M. (2012). Genomic Characterization of a Newly Discovered Coronavirus Associated with Acute Respiratory Distress Syndrome in Humans. *MBio*, *3*(6). <https://doi.org/10.1128/mBio.00473-12>
- Vega, A., & Castro, L. (2019). Impact of climate change on insect–human interactions. *Current Opinion in Allergy and Clinical Immunology*, *19*(5), 475–481. <https://doi.org/10.1097/ACI.0000000000000565>
- Vijaykrishna, D., Smith, G. J. D., Zhang, J. X., Peiris, J. S. M., Chen, H., & Guan, Y. (2007). Evolutionary Insights into the Ecology of Coronaviruses. *Journal of Virology*, *81*(8), 4012–4020. <https://doi.org/10.1128/JVI.02605-06>
- Wang, L.-F., & Anderson, D. E. (2019). Viruses in bats and potential spillover to animals and humans. *Current Opinion in Virology*, *34*, 79–89. <https://doi.org/10.1016/j.coviro.2018.12.007>
- Wang, N., Shi, X., Jiang, L., Zhang, S., Wang, D., Tong, P., Guo, D., Fu, L., Cui, Y., Liu, X., Arledge, K. C., Chen, Y.-H., Zhang, L., & Wang, X. (2013). Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. *Cell Research*, *23*(8), 986. <https://doi.org/10.1038/cr.2013.92>
- Wang, W., Lin, X.-D., Guo, W.-P., Zhou, R.-H., Wang, M.-R., Wang, C.-Q., Ge, S., Mei, S.-H., Li, M.-H., Shi, M., Holmes, E. C., & Zhang, Y.-Z. (2015). Discovery, diversity and evolution of novel coronaviruses sampled from rodents in China. *Virology*, *474*, 19–27. <https://doi.org/10.1016/j.virol.2014.10.017>
- Wang, W., Lin, X.-D., Liao, Y., Guan, X.-Q., Guo, W.-P., Xing, J.-G., Holmes, E. C., & Zhang, Y.-Z. (2017). Discovery of a Highly Divergent Coronavirus in the Asian House Shrew from China Illuminates the Origin of the Alphacoronaviruses. *Journal of Virology*, *91*(17). <https://doi.org/10.1128/JVI.00764-17>
- Webb, B., & Sali, A. (2016). Comparative Protein Structure Modeling Using MODELLER. *Current Protocols in Bioinformatics*, *54*, 5.6.1–5.6.37. <https://doi.org/10.1002/cpbi.3>
- Wong, A. C. P., Li, X., Lau, S. K. P., & Woo, P. C. Y. (2019). Global Epidemiology of Bat Coronaviruses. *Viruses*, *11*(2). <https://doi.org/10.3390/v11020174>
- Woo, P. C. Y., Lau, S. K. P., Lam, C. S. F., Lau, C. C. Y., Tsang, A. K. L., Lau, J. H. N., Bai, R., Teng, J. L. L., Tsang, C. C. C., Wang, M., Zheng, B.-J., Chan, K.-H., & Yuen, K.-Y. (2012). 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*, *86*(7), 3995–4008. <https://doi.org/10.1128/JVI.06540-11>

Wu, Z., Lu, L., Du, J., Yang, L., Ren, X., Liu, B., Jiang, J., Yang, J., Dong, J., Sun, L., Zhu, Y., Li, Y., Zheng, D., Zhang, C., Su, H., Zheng, Y., Zhou, H., Zhu, G., Li, H., ... Jin, Q. (2018). Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases. *Microbiome*, 6. <https://doi.org/10.1186/s40168-018-0554-9>

Zappulli, V., Ferro, S., Bonsembiante, F., Brocca, G., Calore, A., Cavicchioli, L., Centelleghes, C., Corazzola, G., De Vreese, S., Gelain, M. E., Mazzariol, S., Moccia, V., Rensi, N., Sammarco, A., Torrigiani, F., Verin, R., & Castagnaro, M. (2020). Pathology of Coronavirus Infections: A Review of Lesions in Animals in the One-Health Perspective. *Animals : An Open Access Journal from MDPI*, 10(12). <https://doi.org/10.3390/ani10122377>

Zhao, J., Cui, W., & Tian, B. (2020). The Potential Intermediate Hosts for SARS-CoV-2. *Frontiers in Microbiology*, 11. <https://doi.org/10.3389/fmicb.2020.580137>

## Appendix

WESV Ruian-90	MK-F--ILLFLII-----SCVECAVY-IKCDTLLNSLDLDRITYSLFGLGDVSVNSTTLIN	53
WESV Yudu-76	MK-L--IFLLLI-----SCVDCAEY-IKCDKTLKLNLDLDRITYSLFGLGDVSVNSTTLIN	53
MERS-CoV	MIHSVFLMFLLTPTESYVDVGPDSVKSACIEVDIQQTFFDKTWRPRPIDVSKADGI IYPQGRYTSNITITYQ	72
WESV Ruian-90	DYFPSKDI-NNWLCEKQ-----GDSVLYYQAFYTIILNKVNDASLKVLIQPIVYSNSSYVI	108
WESV Yudu-76	DYFPSKDI-NNWLCEKN-----GAAVLYHGQAFYTIILNKVNSASLKFLLIQPLVYTNSSYVI	108
MERS-CoV	GLFFPYQGDHGDYVYSAGHATGTPQKLFVANYSDVKQFANGFVVRIGAAANSTGTVIISPSTSA-TIRKI	143
WESV Ruian-90	YMVYKSGGL--SFSI-----CYRNYFEFNWNLEPSSPISC-----	141
WESV Yudu-76	YMVYKVGSL--SFSI-----CYRNPFFSFSWNSDPPSPISC-----	141
MERS-CoV	YPAFMLGSSVGNFSDGKMGRRFNHTLVLLPDGCGTLLRAFYCILEPRSNHCPAGNSYTSFATYHTPATDCS	215
WESV Ruian-90	--AFNRNIYF-----GYRY-----VYLGIIWRHGYSVKIITPNFVQQFELDNMYLEVIRVA	189
WESV Yudu-76	--AFNRNIYF-----GDRY-----VYLGIIWRHGYSVKIITPNFVQQFELDNMYLEVIRVA	189
MERS-CoV	DGNYNRNASLNSFKKEYFNLRNCTFMYTYNITEDEILEWFGITQTAQGVHLFSSRYVDLYGG-NMF----QFA	282
WESV Ruian-90	TGRFGTTGSG-YNIVMPFQIVD-----KRSTFLVNYDQSGYVSNLLNCSESIEAQLKCKSR	244
WESV Yudu-76	TGRFGTAGSEMLSVVMPPFQIVD-----KRSTFLVNYNQSGYVSNLLNCSESVEAQLKCKSR	245
MERS-CoV	TLPVYDTI-K-YYSIIPHSIRSIQSDRKAWAAFVYKQLPLTFLFLDFSVGDYIRRAIDCGFNDSLQLHCYSY	352
WESV Ruian-90	SFELQSGVYEIAL-EEKVGN-----YTVVTNVTSDCS-NIKYLFNGFTGQKGL	291
WESV Yudu-76	SFELQSGVYEIAL-EERTGGY-----YTVVTNVTSDCN-NIKYLFNSFTGEKGL	292
MERS-CoV	SFDVESGVYSVSSFEAKPSGSVVEQAEGVECDFSPLLSGTTPQVYNFKRLVFTNCNYNLTKLLSLSF-----	419
WESV Ruian-90	YAGNFASMSCIVPFNFISDKDLLNPNPQLRGPNGEFKPFIDCKDINPIGLFGTNGTCYAGMEWRYWTRRFSG	363
WESV Yudu-76	YAGNFVSMSCIVPFNFISNDMLNPNPQLRGPNGEFKPFIDCKDINPIGLFGSNGTCYAGMEWRYWTRRFSG	364
MERS-CoV	--VN-----DFTCSQISPAIA---SNCYSSLILDYFSYPLSM	452
WESV Ruian-90	---YNMSIEGTAESF-YDDGVGENECYGLLKGMQYNNIFKFTVLYYK-----	407
WESV Yudu-76	---YNMSIEGTAESF-YEDGVGENECYGLLKGRYNNIFRFTEVLYYK-----	408
MERS-CoV	KSDLSVSSAGPISQFNYSQSFNSPTCLIL--ATVPHNLTITKPLKYSYINKSRLSDDRTEVPQLVNAQ	522
WESV Ruian-90	-----VRFMEKGLDVCYGN-----L	422
WESV Yudu-76	-----VRFMEKGLDVCYGN-----L	423
MERS-CoV	YSPCVSIVPSTVWEDGDYRQKLSPLEGGWLVASGSTVAMTEQLQMGFGITVQYGTDTNSVCPKLEFANDT	594
WESV Ruian-90	PQAPQVNKCNVWTFNDITFEGVLETKTFNPQLKVLYKGS--MVIMIRVFGKVEVVPICIKTRLSVIYNAN	492
WESV Yudu-76	PPAPQVNKCNVWTFNDITFEGVLETKTFNPQLKVLYKGS--MVIMIRIFGIVYEVVPCIKTRLSVIYNAN	493
MERS-CoV	KIASQLGNCVEYSLYGVSGRGVFNQCTAVGVRQRFVYDAYQNLVGYYSDDGNYCLRACVSPVSVIYDKE	666
WESV Ruian-90	NNNYSTLYRVANCNQIDTKRSVYPAGL---L-SRD---GINTPCGCLRNAQYLPNITVDSYSCKTLLSDGY	556
WESV Yudu-76	NNNYSTLHRVANCNQIDLKRSVHPVGL---L-SRD---GINTPCGCLRNAQYSPNITVDSYSCKTLLSDGY	557
MERS-CoV	TKTHATLFGSVACEHISSTMSQYSFSTRSMLKRDSTYGFPLQTPVGCVLGL---VNSSLFVEDCKLPLGQSL	735
S1/S2		
WESV Ruian-90	CVVFNNNGSTE-----VVPIGIVGANTTYMMP-----LEESFVELALDH--VLVSESEYFQTEMPVFDIDCE	616
WESV Yudu-76	CVVFNNNGSTE-----VVPIGIVGANTTYMMP-----LEESFVELALDH--VLVSESEYFQTEMPVFDIDCE	617
MERS-CoV	CALPDTPTSLTPRSVRSVPGEMRLASIAFNHPIQVDQLNSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDC	807
WESV Ruian-90	KYICDNSESCRMLLEQYGGFCSKII SDVKGNSQLLNANTLAMYKSIVPSSDSIKPIDNF-GDFNMSMFI----	683
WESV Yudu-76	KYVCDNSESCRMLLEQYAGFCSKITSDVKGNSQLLNANTLAMYKSIVPSGSIKPIESF-GDFNMSMFI----	684
MERS-CoV	QYVCGNFQKCEQLLREYGFQFCSKINQALHGANLQDDSVRNLFASVKSSQSSPIPGFGGDFNLTLLEPVSI	879
S2' fusion peptide (FP)		
WESV Ruian-90	---DKKGRSVLEDLLFDKIVTTGPGFYDDYK---RRTTVKDLVCAQYNGIMVPIPIQDA-----ETIGM	744
WESV Yudu-76	---GKKGRSVLEDLLFDKIVTNGPGFYDDYK---RATTVKDLVCAQYNGIMVPIPIQDA-----ETIGM	745
MERS-CoV	STGSRSAIAIEDLLFDKVTIADPGYMQYDDCMQGGPASARDLCAQYVAGYKVLPPMLDVNMEAAAYTSSL	951

WESV Ruian-90	YGGIVAAGMTMGLFGGQVTMGWTAMTARSARLNAGVVTQGMIMEDVQKLAQGFNTLATSFSKMASITADALSR	816
WESV Yudu-76	YGGIVAAGMTMGLFGGQATMGWTAMTARSARLNAGVVTQGMIMEDVQKLAQGFNTLATSFSKMASVTADALSR	817
MERS-CoV	LGSIAAGVGTAGLSSFAA--IPFAQSI FYRLNGVGTITQQVLSQKLIANKFNQALGAMQTGF TTTNEAFQK	1021
WESV Ruian-90	IQGVVNANALQVELLAKATMDNFGAISSNFEIIEKRLKQVQAEQQMDRVINGKMNALLNFVSNYKLVSELT	888
WESV Yudu-76	IQGVVNANALQVELLAKATMDNFGAISSNFEIIEKRLKQVQAEQQMDRVINGKMNALLNFVSNYKLVSELT	889
MERS-CoV	VQDAVNNNAQALSKLASELSNTFGAISASIGDIQRDLVLEQDAQIDRLINGRLTTLNLFVAQQQLVRSESAA	1093
WESV Ruian-90	SIQQLVKQKQVQECVHAQSLRNGFCGSGLVHMSIPQLVPNGIMFLHFSLVKNNSIIVKQTPGLCKE-DKINCI	959
WESV Yudu-76	SIQQLVKQKQVQECVHAQSLRNGFCGSGLVHMSIPQLVSNVIMFLHFSLVKNNSIIVKQTPGLCKE-DKVNCI	960
MERS-CoV	LSAQLAKDKVNECVKAQSKRSRSGFCGQGHIVSFVFNAPNGLYFMHVGYYPNSNHIEVVSAYGLCDAANPTNCI	1165
WESV Ruian-90	VPKSGIFVSKNDS----NWYITPRNKFPKPNITNSDIIFTTYSENFTIINNTIQFP-----GIEFPQDGDF	1021
WESV Yudu-76	VPKSGIFVSKNDS----NWYITPRNKFPKPNITNSDIIFTTYSENFTIINNTIQFP-----GIEFPQDGDF	1022
MERS-CoV	APVNGYFIKTNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQNI STNLPPPLLGNSTGIDFQD--EL	1235
WESV Ruian-90	DGQIKNVTLELQSLKDVV-ANMSYLDLSNHTALLKNI SEEVAKMNITVSQFKQYVKYVWKPWYVWLLIVMAL	1092
WESV Yudu-76	DGQIKNVTLELQSLKDVV-ANMSYLDLSNHTALLKNI SEEVAKMNITVSQFKQYVKYVWKPWYVWLLIVMAL	1093
MERS-CoV	DEFFKNVSTSI PNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYNKWPWYIWLGFIAGL	1307
WESV Ruian-90	IAFGLVCFWLFMCTGCCGLCGC-IR-RSCCDN-TYE-PIEKIHI-Q	1134
WESV Yudu-76	IAFGLLCFVWFVFMCTGCCGLCDC-IR-RSCRDSN-TYE-PIEKIHI-Q	1135
MERS-CoV	VALALCVFFILCCTGCGTNCMGKLCNRCCDRYEEYDLEPHKVHVH*	1354