Furin cleavage sites in the spike proteins of bat and rodent coronaviruses: Implications for virus evolution and zoonotic transfer from rodent species
Alison E. Stout, Jean Kaoru Millet, Michael J. Stanhope, Gary R. Whittaker

To cite this version:
Alison E. Stout, Jean Kaoru Millet, Michael J. Stanhope, Gary R. Whittaker. Furin cleavage sites in the spike proteins of bat and rodent coronaviruses: Implications for virus evolution and zoonotic transfer from rodent species. One Health, Elsevier, 2021, 13, 10.1016/j.onehlt.2021.100282. hal-03487798

HAL Id: hal-03487798
https://hal.archives-ouvertes.fr/hal-03487798
Submitted on 20 Dec 2021

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.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives| 4.0 International License
Furin cleavage sites in the spike proteins of bat and rodent coronaviruses: Implications for virus evolution and zoonotic transfer from rodent species

Alison E. Stout a, Jean K. Millet b, Michael J. Stanhope c, Gary R. Whittaker a,d,*

a Department of Microbiology and Immunology, Cornell University, Ithaca, NY, USA
b Université Paris-Saclay, INRAE, UVSQ, Virologie et Immunologie Moléculaires, 78352, Jouy-en-Josas, France
c Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA
d Master of Public Health Program, Cornell University, Ithaca, NY, USA

A R T I C L E   I N F O
Keywords:
Coronavirus
Furin
Spike protein
Bats
Rats
Rodents

A B S T R A C T

Bats and rodents comprise two of the world’s largest orders of mammals and the order Chiroptera (bats) has been implicated as a major reservoir of coronaviruses in nature and a source of zoonotic transfer to humans. However, the order Rodentia (rodents) also harbors coronaviruses, with two human coronaviruses (HCoV-OC43 and HCoV-HKU1) considered to have rodent origins. The coronavirus spike protein mediates viral entry and is a major determinant of viral tropism; importantly, the spike protein is activated by host cell proteases at two distinct sites, designated as S1/S2 and S2’. SARS-CoV-2, which is considered to be of bat origin, contains a cleavage site for the protease furin at S1/S2, absent from the rest of the currently known betacoronavirus lineage 2b coronaviruses (Sarbecoviruses). This cleavage site is thought to be critical to its replication and pathogenesis, with a notable link to virus transmission. Here, we examine the spike protein across coronaviruses identified in both bat and rodent species, and address the role of furin as an activating protease. Utilizing two publicly available furin prediction algorithms (ProP and PiTou) and based on spike sequences reported in GenBank, we show that the S1/S2 furin cleavage site is typically not present in bat virus spike proteins but is common in rodent-associated sequences, and suggest this may have implications for zoonotic transfer. We provide a phylogenetic history of the Embecoviruses (betacoronavirus lineage 2a), including context for the use of furin as an activating protease for the viral spike protein. From a One Health perspective, continued rodent surveillance should be an important consideration in uncovering novel circulating coronaviruses.

1. Introduction

The Coronaviridae are single-stranded, positive sense, enveloped RNA viruses [1]. They are classified, by genera, as Alpha-, Beta-, Gamma-, and Deltacoronaviruses [2]. The Alpha- and Betacoronaviruses are most often associated with mammalian species, while the Gamma- and Deltacoronaviruses remain largely avian-associated [3]. Across bat and rodent species, both Alpha- and Betacoronaviruses have been identified. Members of the Coronaviridae are further classified into subgenera. Specifically, in bats, this has included the subgenera, Setracoctirus, Myotacovirus, Rhinacovirus, Colacovirus, Pedacovirus, Decacovirus, Minucovirus, and Nyctocovirus, from the Alphacoronavirus genus, in addition to Nobecovirus, Sarbecovirus, Merbecovirus and Hibecovirus from the Betacoronavirus genus. Coronaviruses from rodents, fall within two subgenera: the Luchacoviruses (within the Alphacoronavirus genus) and Embecoviruses (Betacoronavirus genus) [5] (Table 1).

Current classification schemes for the Coronaviridae utilize the replicase domain of the ORF1ab gene; however, the spike protein, is a critical mediator of viral tropism and an important contributor to the viruses natural history and potential for spillover [6]. The rodent Alphacoronaviruses have previously been shown to form a monophyletic group [7]. The Embecovirus subgenera, previously classified as lineage 2a, additionally includes human coronavirus HKU-1, several mammalian coronaviruses such as equine coronavirus and dromedary camel coronavirus HKU23 [8], but lacks any representation of Chiroptera-associated coronaviruses [4]. With discovery of the China Rattus coronavirus (ChRCoV) HKU24, Lau and colleagues have provided evidence that rodent coronaviruses are the ancestors to the Embecovirus subgenus [4,9]. While bats are frequently implicated as the source of new and emerging coronaviruses, rodents need also to be considered as a
potential source of zoonotic spillovers, given previous observations. The natural ability of coronaviruses to recombine adds to the emergence and transmission of new viruses and there are many examples of recombination reported: the prototypic coronavirus mouse hepatitis virus (MHV), an *Embecovirus*, has been noted for its ability to recombine [10,11] and the dromedary camel coronavirus HKU23 appears to have had a recombinant history, based on similarity between portions of HKU23 and RodentCoV-IM2014 [12].

Globally, over 2000 rodent species have been described and encompass a wide range of habitats [13]. Rodents have previously been implicated in transmitting numerous pathogens to humans, including hantavirus, Lassa virus, *Francisella tularensis* and *Yersinia pestis*. In a study involving rodents destined for human consumption, coronaviruses were frequently identified [14]. Among some of the rodents sampled, bat coronavirus SADS-CoV suggest it may have a unique spike protein compared with other alpha coronaviruses, with the bat coronavirus HKU-2 [20]. Coronaviruses in rodents have been identified across several countries and from a number of rodent species, including *Apodemus chevrieri*, *Mus musculus*, *Mus cervicolor*, *Berylmys berdmorei*, *Rattus exulans*, *Rattus tanezumi*, *Niviventer fulvescens*, *Berylmys bowersi*, *Eothenomys miletus*, and *Yersinia pestis*

| α CoV | Order Chiroptera | Subfamily | Genus | Species | Time Period | Sample Size | Ref |
|-------|-----------------|-----------|-------|---------|-------------|-------------|-----|
| Colacovirus | Microchiroptera | Embecovirus | Luchacovirus | Apodemus sp. | 2014–2016 | 16/206 [78] |    |
| Decacovirus | | | | Apodemus agrarius | 2011–2013 | 10/444 [5] |    |
| Duvacovirus | | | | Apodemus chevrieri | 2011 | 21/98 [20] |    |
| Minuscovirus | | | | Apodemus illex | 2014–2015 | 3/150 [18] |    |
| Myotacovirus | | | | Apodemus latorunum | 2014–2015 | 3/6 [17] |    |
| Nyoacovirus | | | | Apodemus sylvaticus | United Kingdom | Not given | [79] |
| Pedacovirus | | | | Articola terrestris | France | 2014–2016 | 0/35 [78] |    |
| Rhinacovirus | | | | Bandicota indica | China | 2014–2015 | 2/5 [17] |    |
| Setraacovirus | | | | Deomys ferrugineus | Democratic Republic of Congo, Republic of Congo | 2006–2018 | 1/1 [80] |    |
| ᵃ CoV | | | | *Eothenomys cachinchus* | China | 2014–2015 | 1/1 [17] |    |
| | | | | *Eothenomys chilensis* | China | 2011 | 1/62 [20] |    |
| | | | | *Eothenomys melita* | China | 2014–2015 | 3/131 [17] |    |
| | | | | *Field rat (Rattus sp. and Bandicota sp.*)* | Viet Nam | 2013–2014 | 239/702*** [14] |    |
| | | | | *Hystrix sp.* | Viet Nam | 2013–2014 | 20/331*** [14] |    |
| | | | | *Malacomyos longipes* | Democratic Republic of Congo, Republic of Congo | 2006–2018 | 1/38 [80] |    |
| | | | | *Microtus minutus* | China | 2011–2013 | 0/2 [5] |    |
| | | | | *Micratosp. sp.* | France | 2014–2016 | 0/9 [78] |    |
| | | | | *Microtus agrestis* | United Kingdom | Not given | 3/11 [79] |    |
| | | | | *Microtus fortis* | China | 2011–2013 | 0/305 [5] |    |
| | | | | *Mus musculus* | United Kingdom | 2014–2015 | 0/10 [17] |    |
| | | | | *Mus musculus* | United Kingdom | 2011–2013 | 0/7 [5] |    |
| | | | | *Mus musculus* | United Kingdom | Not given | 0/394 [79] |    |
| | | | | *Myodes glareolus* | United Kingdom | Not given | 1/1 [79] |    |
| | | | | *Myodes glareolus* | France | 2014–2016 | 5/80 [78] |    |
| | | | | *Niviventer confucianus* | China | 2011–2013 | 1/85 [5] |    |
| | | | | *Niviventer chinensis* | China | 2014–2015 | 0/2 [17] |    |
| | | | | *Niviventer chinensis* | China | 2010–2012 | 0/97**** [9] |    |
| | | | | *Niviventer chinensis* | China | 2014–2015 | 0/2 [17] |    |
| | | | | *Niviventer chinensis* | China | 2013–2014 | 1/1*** [14] |    |
| | | | | *Rattus sp.* | Viet Nam | 2013–2014 | 0/170**** [9] |    |
| | | | | *Rattus sp.* | China | 2010–2012 | 0/170**** [9] |    |
| | | | | *Rattus exulans* | China | 2011–2013 | 0/2 [5] |    |
| | | | | *Rattus exulans* | China | 2011–2013 | 0/2 [5] |    |
| | | | | *Rattus exulans* | China | 2011–2013 | 0/301 [5] |    |
| | | | | *Rattus sp.* | China | 2013–2014 | 0/1*** [14] |    |
| | | | | *Rattus sp.* | China | 2010–2012 | 4/21*** [9] |    |
| | | | | *Rattus norvegicus* | China | 2013–2014 | 1/1*** [14] |    |
| | | | | *Rattus norvegicus* | China | 2010–2012 | 0/170**** [9] |    |
| | | | | *Rattus norvegicus* | China | 2010–2012 | 0/170**** [9] |    |
| | | | | *Rattus norvegicus* | China | 2013–2014 | 2/101 [17] |    |
| | | | | *Rattus norvegicus* | China | 2014–2015 | 0/2 [17] |    |
| | | | | *Rattus norvegicus* | China | 2015 | 0/1 [17] |    |

The spillover of coronaviruses from animal species to humans has the potential for major global health implications. Following the 2002–2003 severe acute respiratory syndrome coronavirus (SARS-CoV) outbreak, surveillance studies began to recognize the role that bat species play as reservoirs for novel coronaviruses [22,23]. Ten years later, Middle East respiratory coronavirus (MERs-CoV) emerged in humans and clusters phylogenetically with two bat viruses, *Tylonycteris* bat coronavirus HKU4 (*BtCoV-HKU4*) and *Pipistrellus* bat coronavirus HKU5 (*BtCoV-HKU5*) [24] in the *Merbecovirus* clade (group 2c coronaviruses). Further investigation identified NeoCoV from South African *Neoromicia capensis* and *Neoromicia capensis*...
bats and suggested it as the sister group to the MERS-CoV-camel clade [25]. MERS-CoV involved the dromedary camel as an intermediate host in the spillover to humans [26–28]. The most recent SARS-CoV-2 outbreak has been traced to bat origins, with a high similarity to a previously identified bat coronavirus RaTG13 (96% genome-wide sequence identity) [29], as well as several other bat coronaviruses, RpYN06 (94.5% sequence identity) [30] and bat-SL-CoVZC45 and bat-SL-CoVZXC21 [31]. The role of an intermediate host in the spread of SARS-CoV-2 to humans remains an open area of investigation.

A major factor for understanding coronavirus tropism lies with the spike protein, consisting of S1 and S2 domains that carry out the receptor binding and membrane fusion functions of the virus, respectively [1]. SARS-CoV-2 and SARS-CoV both utilize ACE-2 as their receptor, while MERS-CoV utilizes DPP4 [32–34]. Equally important as a usable receptor is the presence of a functional protease for S activation [35]. Numerous proteases have been investigated across the Coronavirus, including TMPRSS2, which acts to activate SARS-CoV-2, as well as furin, trypsin, and cathepsins [36]. While receptor binding is well recognized as a factor in host tropism, it is becoming more recognized that host cell proteases can also influence coronavirus species susceptibility and “spillover” events [37,38].

A notable difference between SARS-CoV-2 and RaTG13 (and all other Sarbecoviruses) is the insertion of a predicted cleavage site for a furin-like protease (P-R-R-A-R) between the S1/S2 regions of the spike protein of SARS-CoV-2 [39–41] with loss of this site associated with virus attenuation and decreased transmission [42,43]. However, the P-R-R-A-R motif is sub-optimal for furin cleavage, indicating it may not be fully adapted for humans. RmYN02, identified in R. malayanus has 93.3% genomic nucleotide identity with SARS-CoV-2 and while the identity is only 71.9% (nucleotide sequence) in the spike protein, RmYN02 has a closely homologous sequence (P-A-A-R) at the S1/S2 cleavage site; however, this is lacking basic amino acids needed for cleavage by furin-like proteases [44]. Comparatively, MERS-CoV is highly unusual in that it contains both an S1/S2 furin cleavage site and an S2’ furin cleavage site [35]. Among two closely related bat coronaviruses, BatCoV-HKU5 and BatCoV-HKU4, HKU5 is also cleaved by furin, though HKU4 may be more closely related to MERS-CoV and can utilize the same receptor, DPP4 [45,46].

Furin is a ubiquitously expressed serine protease that functions across numerous physiological processes [47]. In addition to several other proprotein convertases, furin has the ability to cleave and activate viral proteins [48]. Differences in furin-like activity across animal hosts may impact viral processing or alter viral tropism. Furin cleavage sites are not unique to the Coronaviridae. The highly pathogenic influenza strain, H5N1, for instance, also contains a polybasic furin cleavage site in the hemagglutinin H5 protein, with major implications for disease outcome [49]. Other viral proteins that contain furin cleavage sites include human immunodeficiency virus (HIV) envelope glycoprotein gp160, herpesvirus glycoprotein B, tick-borne encephalitis virus envelope protein pMr, and Ebola virus GP [47,50,51]. Additionally, within the Nidovirales (the order within which the Coronaviridae are placed), a furin cleavage site has been predicted in an insect nidovirus isolated from Culex mosquitoes [52].

In the case of coronaviruses, it is known that furin cleavage sites often occur naturally and that this is highly dependent on the specific virus family or lineage [57]. While uncommon for coronaviruses of bats, furin cleavage sites are commonly found in coronaviruses of rodents and it is perhaps fitting to note that proteolytic processing of the coronavirus spike protein was first recognized in the model rodent coronavirus, murine hepatitis virus, MHV-AS9 [53], with later analyses demonstrating the importance of furin for the proteolytic cleavage and function of its spike protein [54]. Here, we provide a comprehensive analysis of furin cleavage sites found in rodent coronaviruses, with a focus on the Embecovirus subgenus within the Betacoronavirus family, and discuss implications for virus evolution and zoonotic transfer from rodent species from a One Health perspective.

1.1. Furin cleavage across bat and rodent species

Based on the known ability of bats and rodents to harbor coronaviruses, comparing the spike protein and furin cleavage sites of viruses across these two hosts may help elucidate origins of novel coronaviruses and guide future surveillance efforts. Through the use of two publicly available programs used to accurately predict furin cleavage sites, PiTou and ProP [55,56], we analyzed furin cleavage across spike sequences in rodent and chiropteran species. PiTou utilizes a hidden Markov model, specifically targeting 20 amino acid residues surrounding furin cleavage sites and important for binding and solvent accessibility [56]. The final score in PiTou is based on log-odds probability. ProP utilizes an artificial neural network to predict furin cleavage [55]. In ProP, sites are given a score of between 0 and 1, with a score above 0.5 being a predicted furin cleavage site.

The protease furin cleaves at a distinct multi-basic motif containing paired arginine residues; furin requires a minimal motif of R-X-X-R, with a preference for an addition basic residue; i.e., R-X-B-R [48]. While most studies have focused on human furin, previous work has indicated furin-like proteases in the megabat Pteropus [57]. The presence of suitable protease activators to enable viral infections further adds to the mystery that enables bats to act as viral reservoirs without seemingly showing clinical signs. Many, but not all, coronaviruses contain a cleavage site at S1/S2, which primes the spike protein for fusion and may increase its affinity for the viral receptor but may also make it structurally unstable. As such S1/S2 is considered dispensable for virus infection. The presence of S1/S2 furin cleavage sites in viruses uniquely identified in bats is shown in Fig. 1A (see also Supplementary Table 2). Across the Coronaviridae, the S2’ cleavage site is also present, with cleavage here considered a required event for viral infection [58]. Bat sequences which have predicted furin cleavage sites at S2’ are additionally shown in Fig. 1B.

The S1/S2 cleavage site was first noticed in coronaviruses infecting laboratory rodents; for example, the mouse hepatitis virus (MHV-1) strain JHM contains a strong S1/S2 cleavage site in addition to a second predicted furin cleavage site just distal. The presence of S1/S2 furin cleavage sites in viruses uniquely identified in rodents is shown in Fig. 2A. Most often, a second, overlapping, S1/S2 furin cleavage site was predicted by PiTou and not by ProP, though an exception was rodent coronavirus RBl-CoV/FJ2015, in which both PiTou and ProP predicted adjacent furin cleavage sites. This redundancy may indicate a needed feature for rodent-associated Coronaviridae. Unsurprisingly, the majority of the rodent coronaviruses with S1/S2 furin cleavage sites were classified as Embecoviruses, a subgenus known for this feature [59]. A weakly predicted cleavage site was also detected in the Alphacoronavirus, AcCoV-JC34. A total of 21 unique potential S1/S2 furin cleavage sites are shown in Fig. 2A. It is important to note, however, that isolate specific differences exist; for example, two isolates of HKU24 possess disparate predicted furin cleavage sites. Lastly, the virus with the strongest predicted furin cleavage site, based on PiTou, was RTR-CoV-T21006A. ProP also gave a strong furin cleavage prediction score (0.88), the highest score observed from ProP, but shared by several other viruses. S2’ furin cleavage sites additionally occur naturally in rodent coronaviruses and are shown in Fig. 2B. It is worthwhile to note, the rodent coronaviruses with S2’ furin cleavage sites also possess S1/S2 furin cleavage sites. Compared to S1/S2 furin cleavage sites in rodent coronaviruses, S2’ furin cleavage sites are less common. Given the well-conserved nature of the S2’ cleavage site, it is unsurprising that the predicted scores in the rodent coronaviruses are relatively similar to the predicted scores from bat associated coronaviruses.

1.2. Phylogenetic analysis

The majority of rodent coronaviruses investigated here are from the Embecovirus clade. This clade additionally includes two human coronaviruses, HCoV-OC43 and HCoV-HKU1 [15]. As previously reported and supported by our analysis, HCoV-HKU1 likely emerged through a rodent
The phylogeny depicts a basal split in the \textit{Embecoviruses} into a predominately rodent clade and a predominately ungulate clade, with spike sequences from horses at the base of that ungulate group. Historically, rodents have been considered major sources of pathogens and our evidence support their role in the emergence of two human coronaviruses. Closest relatives of the \textit{Embecovirus} family are the \textit{Betacoronavirus}, which includes the \textit{SARS-CoV} and \textit{MERS-CoV} and are closely related to the \textit{Embecoviruses}.

### Table of Predicted Furin Cleavage Sites

| Isolate/strain | Classification | PiTou Score | Furin Score |
|----------------|----------------|-------------|-------------|
| Murine hepatitis virus strain JHM | EmeCoV | 0.50 | 0.45 |
| Longmann Rl rat coronavirus Longmann-189 | EmeCoV | 0.46 | 0.43 |
| Betacoronavirus HEMD, Longmann-723 | EmeCoV | 0.45 | 0.42 |
| Longmann Rl rat coronavirus Nulli-66 | EmeCoV | 0.38 | 0.35 |
| Betacoronavirus HEMD, Lijiang-4 | EmeCoV | 0.39 | 0.36 |
| Longmann Rl rat coronavirus Nulli-67 | EmeCoV | 0.37 | 0.34 |
| Betacoronavirus HEMD, Nulli-67 | EmeCoV | 0.38 | 0.36 |
| MHV strain A9 | EmeCoV | 0.46 | 0.43 |
| Coronavirus AcoV-20124 | EmeCoV | 0.51 | 0.48 |
| KLU--G0V2018 | Unclassified | 0.22 | 0.20 |
| Redon coronavirus RTm-CoV202015 | Unclassified | 0.55 | 0.50 |
| Redon coronavirus RTm-CoV2014 | Unclassified | 0.61 | 0.58 |
| Redon coronavirus RTm-CoV2014 | Unclassified | 0.54 | 0.51 |
| Redon coronavirus RTm-CoV202015 | Unclassified | 0.55 | 0.50 |
| Redon coronavirus RTm-CoV2014 | Unclassified | 0.54 | 0.51 |
| Redon coronavirus RTm-CoV202015 | Unclassified | 0.55 | 0.50 |
| Redon coronavirus RTm-CoV202015 | Unclassified | 0.54 | 0.51 |

### Notes

- Predicted cleavage sites are commonly found in rodent associated coronaviruses. Examples of predicted S1/S2 furin cleavage sites in rodent associated coronaviruses include: Longmann Rl rat coronavirus Longmann-189 (ABM43191), Betacoronavirus HKU15 (AGP04938), Betacoronavirus HKU24 (AYR18647), and Coronavirus AcoV-20124 (ADY17911).
- Our analysis supports classification as an \textit{Embecovirus}.

### References

1. A. E. Stout et al. (2021). "One Health 13 (2021) 100282".
2. BetaCoV/GX2018 (QX53858), Rousetts bat coronavirus HKU 9 (AVP25406), Bat coronavirus 1B (ACA52157), 229E-related bat coronavirus (ALK28781), BtMf-ACoV/GD2015 (ATP66752), RttR-CoV/Tn2018 (QYM73848).

### Fig. 3

Examples of predicted furin cleavage sites identified in coronaviruses associated with chiropteran species. Over 150 spike sequences from bat associated CoVs were screened for furin cleavage sites, including those in both the Alpha- and Beta-coronaviruses genera, using the programs PiTou and ProP. A. Unique S1/S2 furin cleavage sites predicted in bat associated coronaviruses. B. Unique S2 furin cleavage sites predicted across bat associated coronaviruses. Associated NCBI accession numbers are as follows: Bat Hp-beta-coronavirus/Zhejiang2013 (YP_009072440), (Putative) Zaria bat coronavirus (ADY17911), Bat coronavirus HKU5-1 (ABN10875), Bat-beta-coronavirus/GD202013 (AAU62185), Pipistrellus abramus bat coronavirus HKU5-related (QHA24687), Bat coronavirus HKU5-2 (ABN10884), Pipistrellus bat coronavirus HKU5 (AGP04938), Coronavirus Neoromicia/PML-PHE1/RSA/2011 (AGY29650), Bat coronavirus A434/2005 (ABG11962), Bat coronavirus PREDICT/PDF-2180 (YP_009361857), Betacoronavirus HKU24 Lijiang-53 (QOE77307) and Betacoronavirus HKU24 Lijiang-41 (QOE77297).

### Fig. 2

Examples of predicted furin cleavage sites identified in coronaviruses associated with chiropteran species. Over 150 spike sequences from bat associated CoVs were screened for furin cleavage sites, including those in both the Alpha- and Beta-coronaviruses genera, using the programs PiTou and ProP. A. Unique S1/S2 furin cleavage sites predicted in bat associated coronaviruses. B. Unique S2 furin cleavage sites predicted across bat associated coronaviruses. Associated NCBI accession numbers are as follows: Bat Hp-beta-coronavirus/Zhejiang2013 (YP_009072440), (Putative) Zaria bat coronavirus (ADY17911), Bat coronavirus HKU5-1 (ABN10875), Bat-beta-coronavirus/GD202013 (AAU62185), Pipistrellus abramus bat coronavirus HKU5-related (QHA24687), Bat coronavirus HKU5-2 (ABN10884), Pipistrellus bat coronavirus HKU5 (AGP04938), Coronavirus Neoromicia/PML-PHE1/RSA/2011 (AGY29650), Bat coronavirus A434/2005 (ABG11962), Bat coronavirus PREDICT/PDF-2180 (YP_009361857), Betacoronavirus HKU24 Lijiang-53 (QOE77307) and Betacoronavirus HKU24 Lijiang-41 (QOE77297).
Almost all *Embecoviruses* contain furin cleavage sites at the S1/S2 interface, with two notable exceptions (shown with black boxes in Fig. 3B). MHV-2 (AF201929) is the cause of a laboratory-acquired infection, first identified in 1952 as acute hepatitis associated with mouse leukemia [61]. It is a distinct virus in cell culture, with a spike protein known to be cleaved by cathepsins [62] and differs significantly from the more widely studied MHV-1 and may represent a tissue-specific variant (S1/S2 site H**R**A**R** for MHV-2 vs. H**R**A**R** for MHV-1; basic residues in bold, predicted mutation underlined). As such, MHV-2 may show parallels with a tropism shift associated with a loss of furin cleavage seen with feline coronaviruses [63]. The other notable exception is Longquan Aa mouse coronavirus (LAMV) (AID16631). Comparatively HKU24 (ATP66750) contains an S1/S2 furin cleavage and aligns with specific residues LL-GAP**R**E for LAMV. In this case, the S1/S2 site seems to represent an indel at that region, and would be of high interest for further follow up with regard to its evolution.

In some cases, there is discrepancy between the ProP and PiTou scores, notably for sequences YP_009113025 and QEY10649. YP_009113025 is a variant of HKU24 (ATP66750) containing an S1/S2 furin cleavage and aligns with specific residues LL-GAP**R**E for LAMV. In this case, the S1/S2 site seems to represent an indel at that region, and would be of high interest for further follow up with regard to its evolution.

2. Discussion

Continued surveillance in bat and rodent species, concomitant with laboratory experimentation and computational analysis, is essential for understanding the overall disease ecology of coronaviruses in these two host groups. Understanding species cross-over potential for coronaviruses often focuses on known features, such as mutation within the receptor-binding domains or in genomic recombination events, but other factors, such as alterations in protease cleavage sites within spike, may also be important.

The Chiroptera and Rodentia encompass vast lineages and inhabit unique ecological niches from bats roosting in caves to mice living in apartment buildings. The spillover of coronaviruses between these species and into humans, livestock, and other species remains a threat to One Health and global health. Across species, it is also important to consider host adaptations that may facilitate their ability to act as viral reservoirs. In bats, the ability to harbor and shed coronaviruses, with...
limited or inapparent clinical signs has remained an ongoing question [38,64]. By comparison, less has been investigated in wild rodent species harboring coronaviruses, though in laboratory settings it is apparent that rodents can show disease associated with coronavirus infection, including Golden Syrian Hamsters (*Mesocricetus auratus*) infected with SARS-CoV-2, which have proven a robust and effective model for SARS-CoV-2 infection [65,66].

SARS-CoV-2 has remained a global challenge and like most diseases, rodent models are helpful in guiding our understanding. In laboratory mice, transgenic humanized ACE2 mice better recapitulate SARS-CoV-2 infection and pathology [67,68]. However, new SARS-CoV-2 variants of concern (VOCs) have expanded the host range of SARS-CoV-2, allowing infection of laboratory mice [69]. Challenges of deer mice (*Peromyscus maniculatus*) with SARS-CoV-2 have demonstrated infection and transmission [70,71]. Additionally, experimentally challenged bank voles (*Myodes glareolus*) appear susceptible to SARS-CoV-2, despite not showing clinical signs [72]. Understanding coronavirus adaptation and transmission among rodent species is an important area of investigation both from the One Health perspective as well as understanding spillover between rodents. Additional important areas of study include how/why some rodents harbor unique viruses, as well as what are the underlying host differences that previously allowed one rodent species (laboratory mice) to be largely unaffected by SARS-CoV-2 while another rodent species (hamsters) readily develop disease. In mice infected with MHV, disease has been associated with encephalomyelitis, wasting, and mortality in naïve animals [73,74]. Further, a case report has described a wasting syndrome in guinea pigs with a presumed coronavirus infection [75].

The activation of the spike protein is a complex process and furin is a commonly used protease to activate fusion machinery. Among the *Alphacoronaviruses*, the presence of an S1/S2 furin cleavage site is infrequently observed, though a notable example is in feline coronavirus, which normally possesses a furin cleavage site at the S1/S2 boundary, but loss of basic residues is often associated with the systemic disease feline infectious peritonitis [63]. A second *Alphacoronavirus*, canine coronavirus 23/03, also possesses a furin cleavage site, but the role in disease is less well defined [76]. In bat spike sequences from *Alphacoronaviruses*, no obvious S1/S2 furin cleavage site was identified based on ProP and PiTou predictions. Two proposed alphacoronavirus from rodents, AcCoV-JC34 and Lucheng Rn rat coronavirus Lijiang-170, had a shared, weakly predicted furin cleavage site (S-R-R-A-R), based on the PiTou program. A caveat to using bioinformatics programs lies in the correlation between a software prediction and biological plausibility. A predicted site, for example, may not be accessible to furin and thus, non-functional and the action of other proteases, needs also to be considered. It is also presently unclear how the presence or absence of a cleavage site and its relative cleavability score translates into a given biological process, such a transmissibility, cross-species tropism or pathogenicity. Nevertheless, our study provides a predictor that can be integrated with experimental validation and further surveillance.

Based on our studies reported here, it is possible that viruses with S1/S2 furin cleavage sites are more commonly found in *Coronaviridae* circulating in rodents. Over three times as many spike sequences were publicly available and screened from bats (179) compared to rodents (55). While MHV was included in our screening, 41 rodent viruses were from previous surveillance work and can be considered naturally occurring. Of these 41 rodent sequences, 32 (78%) had potential S1/S2 furin cleavage sites. Of the 41 spike sequences from bats, only 11 of 179 (6%) sequences had predicted S1/S2 furin cleavage sites. For some CoVs, such as HKU5, numerous accessions are publicly available through sources such as NCBI, so it is challenging to strictly compare the frequency of furin cleavage sites across the species that sequences come from, but it does seem apparent that rodents regularly harbor viruses with strong S1/S2 furin cleavage sites. With regard to the S2’ furin cleavage site, the proportion of rodent versus bat sequences possessing this site is relatively similar, being found in 3 of 41 (7.3%) naturally occurring rodent spike sequences and in 13 of 179 (7.3%) bat sequences. Host differences may additionally contribute to these observed differences. Both rodent and bat associated coronaviruses have been associated with human disease and both must continue to be investigated, including in regular surveillance studies. Recently, a web-based, risk assessment tool for evaluating viruses of wildlife origin designated Lassa virus, a rodent-associated virus, as the virus most likely to spillover into human populations (ranked #1); notably, several rodent-related coronaviruses were ranked in the top 50, specifically Murine CoV (#18), Longquan Aa CoV (#23), Rodent CoV (#27), and HKU-1 (37) [77].

The *Coronaviridae* lend themselves well to zoonotic spillover. The unique spike protein determines host tropism and is largely a balance between receptor binding ability and the presence of an acceptable host cell protease for spike protein activation. Bats and rodents have both been implicated as reservoirs for ancestral coronavirus species. The common presence of predicted furin cleavage sites across rodent coronaviruses, suggests rodents warrant further consideration as a putative source of new viral emergences.

**Acknowledgments**

We thank all members of the Whittaker Lab for helpful discussions during the preparation of this manuscript.

**Funding**

AES is supported by NIH Comparative Medicine Training Program T32OD011000.

**Declaration of interest**

The authors declare no conflict of interest.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.omega.2021.100282.

**References**

[1] D.M. Knipe, P. Howley, Fields Virology, Wolters Kluwer, Philadelphia, United States, 2013. http://ebookcentral.proquest.com/lib/cornell/detail.action?docID=3418302 (accessed November 20, 2019).

[2] S. Payne, Family coronaviridae, Viruses (2017) 149–158, https://doi.org/10.3390/v9090116.

[3] P.C.Y. Woo, S.K.P. Lau, C.S.F. Lam, C.C.Y. Lau, A.K.L. Tsang, J.H.N. Lau, R. Bai, J.L. Teng, C.C.C. Tsang, M. Wang, B.J. Zheng, K.-H. Chan, K.-Y. Yuen, 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, J. Virol. 86 (2012) 3995–4008, https://doi.org/10.1128/JVI.0540–11.

[4] A.C.P. Wong, X. Li, S.K.P. Lau, P.C.Y. Woo, Global epidemiology of bat coronaviruses, Viruses 11 (2019), https://doi.org/10.3390/v11020174.

[5] W. Xiang, D.-X. Lin, W.-P. Guo, R.-H. Zhou, M.-R. Wang, C.-Q. Wang, S. Ge, S.-H. Mei, M.-H. Li, M. Shi, E.C. Holmes, Y.-Z. Zhang, Discovery, diversity and evolution of novel coronaviruses sampled from rodents in China, Virology 474 (2015) 19–27, https://doi.org/10.1016/j.virol.2014.10.017.

[6] G.R. Whitaker, N.M. André, J.K. Millet, Improving virus taxonomy by recontextualizing sequence-based classification with biologically relevant data: the case of the alphacoronavirus 1 species, MSphere 3 (2018), e00463–17, https://doi.org/10.1128/msphere.00463–17.

[7] T. Tsoelirid, J.G. Chappell, O. Osianwa, D.A. Marston, A.R. Fooks, E. Monchatre-Leroy, G. Umbang, M.A. Müller, J.F. Drexler, C. Drosten, R.E. Tarlinton, C. P. McClure, E.C. Holmes, J.K. Ball, Shared common ancestry of rodent alphacoronaviruses sampled globally, Viruses 11 (2019), https://doi.org/10.3390/v11020125.

[8] P.C.Y. Woo, S.K.P. Lau, U. Wernery, E.Y.M. Wong, A.K.L. Tsang, B. Johnson, C.C.C. Yip, C.C.Y. Lau, S. Sivakumar, J.-P. Cai, R.Y.Y. Fan, K.-H. Chan, R. Mareena, K.-Y. Yuen, Novel betacoronavirus in dromedaries of the Middle East, 2013, Emerg. Infect. Dis. 20 (2014) 560–572, https://doi.org/10.3201/eid2004.131769.

[9] S.K.P. Lau, P.C.Y. Woo, K.S.M. Li, A.K.L. Tsang, R.Y.Y. Fan, H.K.H. Luk, J.-P. Cai, K.-H. Chan, B.-J. Zheng, M. Wang, K.-Y. Yuen, Discovery of a novel Coronavirus, China Rattus Coronavirus HKU24, from Norway rats supports the murine origin of
J.F.-W. Chan, A.J. Zhang, S. Yuan, V.K.-M. Poon, C.C.-S. Chan, A.C.-Y. Lee, W.-A.T. Irving, M. Ahn, G. Goh, D.E. Anderson, L.-F. Wang, Lessons from the host.

C.A.M. de Haan, K. Stadler, G.-J. Godeke, B.J. Bosch, P.J.M. Rottier, Cleavage.

J. Castresana, Selection of conserved blocks from multiple alignments for their use.

Y. Wu, S. Zhao, Furin cleavage sites naturally occur in coronaviruses, Stem Cell 8 (2021) 370-371, https://doi.org/10.1016/j.stem.2020.05.028.

F. El Najjar, L. Lampe, M.L. Baker, L.-F. Wang, R.E. Dutch, Analysis of cathepsin.

S. Tian, W. Huajun, J. Wu, Computational prediction of furin cleavage sites by a.

R. Kuwata, T. Satho, H. Isawa, N.T. Yen, T.V. Phong, T. Kurashige, G. Izaguirre, The proteolytic regulation of virus cell entry by furin and other.

R. Althage, T. Satho, H. Isawa, N.T. Yen, T.V. Phong, T. Kurashige, G. Izaguirre, The proteolytic regulation of virus cell entry by furin and other.

J.K. Lee, L. Shen, Y. Zhou, X. Yang, T.-Y. Zhao, C.-F. Fan, Y.-S. Zhou, C.-F. Qin, Y.-C. Wang, A mouse model of SARS-CoV-2 infection and pathogenesis, Cell Host Microbe 28 (2020) 124–133, e4, https://doi.org/10.1016/j.chom.2020.05.030.

E.S. Winkler, A.L. Bailey, N.M. Rafai, S. Nair, B.T. McCormack, J. Yu, J.M. Fox, R. Chen, J.T. Earnest, S.P. Keeler, J.H. Ritter, L.-I. Kang, S. Dort, A. Robichaud, R. Head, M.J. Holtzman, M.S. Diamond, SARS-COV-2 infection of human ACE2.

X. Montagutelli, M. Prot, L. Levillayer, E.B. Salazar, G. Jouvion, L. Conquet, F. Donati, M. Albert, F. Gambardo, S. Behillil, V. Enouf, D. Rousset, J. Jauber, F. Rey, S. van der West, E. Simon-Loriere, The B1.351 and P.1 variants extend SARS-CoV-2 host range to mice, BioRxiv (2021), https://doi.org/10.1101/2021.03.18.436013. 2021.03.18.436013.

A. Fagre, J. Lewis, M. Eckley, S. Zhan, S.M. Rocha, N.R. Sexton, B. Burke, B. Geiss, O. Peeressen, R. Kading, J. Roovink, G.D. Ebel, R.B. Tjaklenis, T. Abboeljil, T. Schoutens, SARS-CoV-2 infection, neuropathogenesis and transmission among Deer Mice: implications for reverse zoonosis to New World Rodents, BioRxiv (2020), https://doi.org/10.1101/2020.08.07.241810.

B.D. Griffin, M. Chan, N. Tailor, E.J. Mendoza, A. Leung, B.M. Warner, A. T. Duggan, E. Moffat, S. H. Lindert, S. Madany, A. Alibert, K. Tierney, J. Aued, A. Bello, R. Vrendamelli, A.S. Boese, L. Fernando, L. R. Lindsay, C.M. Jardine, H. Wood, G. Polliquin, J.E. Strong, M. Drebout, D. Safonetz, C. Embury-Haytt, D. Kobasa, North American deer mice are susceptible to SARS-CoV-2, BioRxiv (2020), https://doi.org/10.1101/2020.07.25.221291.

W. Liu, A. Michelitsch, N. Halwe, K. Wernike, D. Hoffmann, M. Beer, Experimental SARS-CoV-2 Infection of Bank Voles - Volume 27, Number 4 – 2021 - Emerging Infectious Diseases Journal - CDC, (n.d.). doi: https://doi.org/10.3201/eid2704.210418

B. Duggan, E. Moffat, H. Druelle, K. Tietjen, J. Aued, A. Bello, R. Vrendamelli, A.S. Boese, L. Fernando, L. R. Lindsay, C.M. Jardine, H. Wood, G. Polliquin, J.E. Strong, M. Drebout, D. Safonetz, C. Embury-Haytt, D. Koba, North American deer mice are susceptible to SARS-CoV-2, BioRxiv (2020), https://doi.org/10.1101/2020.07.25.221291.

B. MPassi, J.J.M. Tamfum, C.E. Lange, Coronavirus surveillance in Congo basin forest ecosystems, Arch. Virol. 158 (2013) 2273–2284, https://doi.org/10.1007/s00705-013-1741-4.

K. Sawabe, Characterization of Dak Nong virus, an insect nidovirus isolated from Culex mosquitoes in Vietnam, Arch. Virol. 158 (2013) 2273–2284, https://doi.org/10.1007/s00705-013-1741-4.

A. Fagre, J. Lewis, M. Eckley, S. Zhan, S.M. Rocha, N.R. Sexton, B. Burke, B. Geiss, O. Peeressen, R. Kading, J. Roovink, G.D. Ebel, R.B. Tjaklenis, T. Abboeljil, T. Schoutens, SARS-CoV-2 infection, neuropathogenesis and transmission among Deer Mice: implications for reverse zoonosis to New World Rodents, BioRxiv (2020), https://doi.org/10.1101/2020.08.07.241810.

B.D. Griffin, M. Chan, N. Tailor, E.J. Mendoza, A. Leung, B.M. Warner, A. T. Duggan, E. Moffat, S. H. Lindert, S. Madany, A. Alibert, K. Tierney, J. Aued, A. Bello, R. Vrendamelli, A.S. Boese, L. Fernando, L. R. Lindsay, C.M. Jardine, H. Wood, G. Polliquin, J.E. Strong, M. Drebout, D. Safonetz, C. Embury-Haytt, D. Koba, North American deer mice are susceptible to SARS-CoV-2, BioRxiv (2020), https://doi.org/10.1101/2020.07.25.221291.