

Research



Cite this article: Fischhoff IR, Castellanos AA, Rodrigues JPGLM, Varsani A, Han BA. 2021 Predicting the zoonotic capacity of mammals to transmit SARS-CoV-2. *Proc. R. Soc. B* **288**: 20211651.
<https://doi.org/10.1098/rspb.2021.1651>

Received: 3 August 2021

Accepted: 25 October 2021

Subject Category:

Ecology

Subject Areas:

ecology, computational biology, structural biology

Keywords:

COVID-19, ecological traits, zoonotic, spillback, machine learning, structural modelling

Author for correspondence:

Barbara A. Han

e-mail: hanb@caryinstitute.org

[†]Contributed equally.

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.5690905>.

THE ROYAL SOCIETY
PUBLISHING

Predicting the zoonotic capacity of mammals to transmit SARS-CoV-2

Ilya R. Fischhoff^{1,†}, Adrian A. Castellanos^{1,†}, João P. G. L. M. Rodrigues², Arvind Varsani^{3,4} and Barbara A. Han¹

¹Cary Institute of Ecosystem Studies, Box AB Millbrook, NY 12545, USA

²Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

³The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA

⁴Structural Biology Research Unit, Department of Integrative Biomedical Sciences, University of Cape Town, 7700 Cape Town, Rondebosch, South Africa

IRF, 0000-0001-6956-8284; **AAC**, 0000-0002-3412-0487; **JPGLMR**, 0000-0001-9796-3193; **AV**, 0000-0003-4111-2415; **BAH**, 0000-0002-9948-3078

Back and forth transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) between humans and animals will establish wild reservoirs of virus that endanger long-term efforts to control COVID-19 in people and to protect vulnerable animal populations. Better targeting surveillance and laboratory experiments to validate zoonotic potential requires predicting high-risk host species. A major bottleneck to this effort is the few species with available sequences for angiotensin-converting enzyme 2 receptor, a key receptor required for viral cell entry. We overcome this bottleneck by combining species' ecological and biological traits with three-dimensional modelling of host-virus protein-protein interactions using machine learning. This approach enables predictions about the zoonotic capacity of SARS-CoV-2 for greater than 5000 mammals—an order of magnitude more species than previously possible. Our predictions are strongly corroborated by *in vivo* studies. The predicted zoonotic capacity and proximity to humans suggest enhanced transmission risk from several common mammals, and priority areas of geographic overlap between these species and global COVID-19 hotspots. With molecular data available for only a small fraction of potential animal hosts, linking data across biological scales offers a conceptual advance that may expand our predictive modelling capacity for zoonotic viruses with similarly unknown host ranges.

1. Introduction

The ongoing COVID-19 pandemic has surpassed 4.8 million deaths globally as of 1 October 2021 [1,2]. Like previous pandemics in recorded history, COVID-19 originated from the spillover of a zoonotic pathogen, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a betacoronavirus originating from an unknown animal host [3–6]. The broad host range of SARS-CoV-2 is due in part to its use of a highly conserved cell surface receptor to enter host cells, the angiotensin-converting enzyme 2 receptor (ACE2) [7] found in all major vertebrate groups [8].

The ubiquity of ACE2 coupled with the high prevalence of SARS-CoV-2 in the global human population explains multiple observed *spillback* infections since the emergence of SARS-CoV-2 in 2019 (see natural infections listed in electronic supplementary material, table S1). In *spillback* infection, human hosts transmit SARS-CoV-2 virus to cause infection in non-human animals. In addition to threatening wildlife and domestic animals, repeated *spillback* infections may lead to the establishment of new animal hosts from which SARS-CoV-2 can then pose a risk of *secondary spillover* infection to humans

© 2021 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License <http://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, provided the original author and source are credited.

through bridge hosts (e.g. [9]) or newly established enzootic reservoirs. Indeed, this risk has been realized in Denmark [10] and The Netherlands, where SARS-CoV-2 spilled back from humans to farmed mink (*Neovison vison*) with the secondary spillover of a SARS-CoV-2 variant from mink back to humans [11]. A major concern in such secondary spillover events is the appearance of a mutant strain affecting host range [12] or leading to increased transmissibility in humans [13,14] (but see [15,16]), reduced sensitivity to neutralizing antibodies [10] and reduced vaccine efficacy [17]. Conversely, human-derived variants pose spillback risks to animals. For example, in contrast with previous infection trials [18], two new human variants have overcome the species barrier to infect laboratory mice (*Mus musculus*) [19].

Spillback infections from humans to animals are already occurring worldwide with pets, domesticated animals, zoo animals and wildlife now documented as new hosts of SARS-CoV-2 (figure 1; electronic supplementary material, table S1). SARS-CoV-2 has been found for the first time in wild and escaped mink in multiple states in the United States, with viral sequences identical to SARS-CoV-2 in nearby farmed mink [33–35]. The global scale of human infections and the increasing range of known animal hosts demonstrate that SARS-CoV-2 has the capacity to establish novel infection cycles in animals. In response, recent computational studies predict the susceptibility of particular animal species to SARS-CoV-2 [12,21–28,36] by comparing known sequences of ACE2 orthologues across species (*sequence-based studies*), or by modelling the structure of the viral spike protein bound to ACE2 orthologues (*structure-based studies*) to yield a wide range of predictions with varying degrees of agreement with laboratory animal experiments (figure 1).

Sequence-based studies predict host susceptibility based on amino acid sequence similarity between human (hACE2) and non-human ACE2, and assume that a high degree of similarity correlates with stronger viral binding, especially at amino acid residues where hACE2 interacts with the SARS-CoV-2 spike glycoprotein. For some species, such as rhesus macaques [37], these qualitative predictions are borne out by *in vivo* studies (figure 1), but predictions from these methods do not consistently match real-world outcomes. For example, sequence similarity predicted weak viral binding for minks and ferrets, which have both been confirmed as highly susceptible [11,22,38] (figure 1). Mismatches to *in vivo* outcomes may arise in part because protein three-dimensional structure, the main determinant of the interaction between host ACE2 and the viral spike protein, is incompletely represented by one-dimensional amino acid sequences [39,40].

Structure-based studies model the three-dimensional structure of protein-protein complexes to address some of the limitations of sequence-based approaches. Structural models have proven useful for predicting how different ACE2 orthologues bind to the SARS-CoV-2 viral spike protein receptor-binding domain (RBD) [12,28]. These studies leverage known structures of the hACE2 receptor bound to the SARS-CoV-2 RBD and use powerful simulations to predict how variation across different ACE2 orthologues affects binding with the viral RBD. While these approaches successfully predicted strong binding for species that have been infected (e.g. domestic cat, tiger, dog and ferret) and weak binding for species in which experimental infections have failed (e.g. chicken, duck [38], mouse [18]), the results are also not consistently supported by experiments. For instance,

while guinea pig ACE2 scored favourably in one structure-based study [12], this orthologue was shown experimentally not to bind to the SARS-CoV-2 RBD [29].

Although structural modelling has produced the most accurate results to date, all currently available approaches for predicting the host range of SARS-CoV-2 are fundamentally constrained by the availability of ACE2 sequences across species. ACE2 is ubiquitous across chordates, probably because of its role in highly conserved physiological pathways, for example in regulating blood pressure, salt and water [41]. The vast majority of mammal species (greater than 6000 species) are likely to have ACE2, but sequences are available for only around 300 species. The functional importance of the ACE2 receptor suggests that it has evolved in association with other intrinsic organismal traits for which data are available for many more species. These suites of correlated organismal traits may provide a robust statistical proxy that can be leveraged to predict biologically permissive hosts for SARS-CoV-2. Previous trait-based analyses applied machine learning techniques to accurately distinguish the zoonotic capacity of various organisms [42–44] and predict likely hosts for particular groups of related viruses [45,46], predictions which have subsequently been validated through independent laboratory and field investigations (e.g. [47,48]).

Here, we combine structural modelling of viral binding with machine learning of species ecological and biological traits to predict zoonotic capacity for SARS-CoV-2 across 5400 mammal species, expanding our predictive capacity by an order of magnitude (figure 2). Crucially, this integrated approach enables predictions for the vast majority of species whose ACE2 sequences are currently unavailable by leveraging information from viral binding dynamics and biological traits. In our workflow (figure 2), we first carry out structural modelling to quantify the binding strength of SARS-CoV-2 RBD for vertebrate species using published ACE2 amino acid sequences [49]. We then collate species traits and train a machine learning model to predict the zoonotic capacity for 5400 mammals.

As COVID-19 is primarily a disease affecting humans, spillback infection of SARS-CoV-2 from humans to animals is the most likely mode by which new animal hosts will become established. We therefore identify a subset of species for which the threat of spillback infection appears greatest due to geographic overlaps and opportunities for contact with humans in areas of high SARS-CoV-2 prevalence globally. These approaches underscore the utility of establishing interdisciplinary and iterative processes that join computational modelling, field surveillance and laboratory experiments to more efficiently quantify zoonotic risk [50], and better inform next steps to prevent enzootic SARS-CoV-2 transmission and spread. Our analyses are based on the initial dominant SARS-CoV-2 variant in humans, but these methods can be readily adjusted to enable host range predictions for new variants as their hACE2-RBD crystal structures become available.

2. Methods

(a) Structural modelling of ACE2 orthologues bound to SARS-CoV-2 spike

We assembled ACE2 sequences from the NCBI GenBank and MEROPS databases. The modelling of 326 ACE2 orthologues bound to SARS-CoV-2 spike RBD was carried out using the

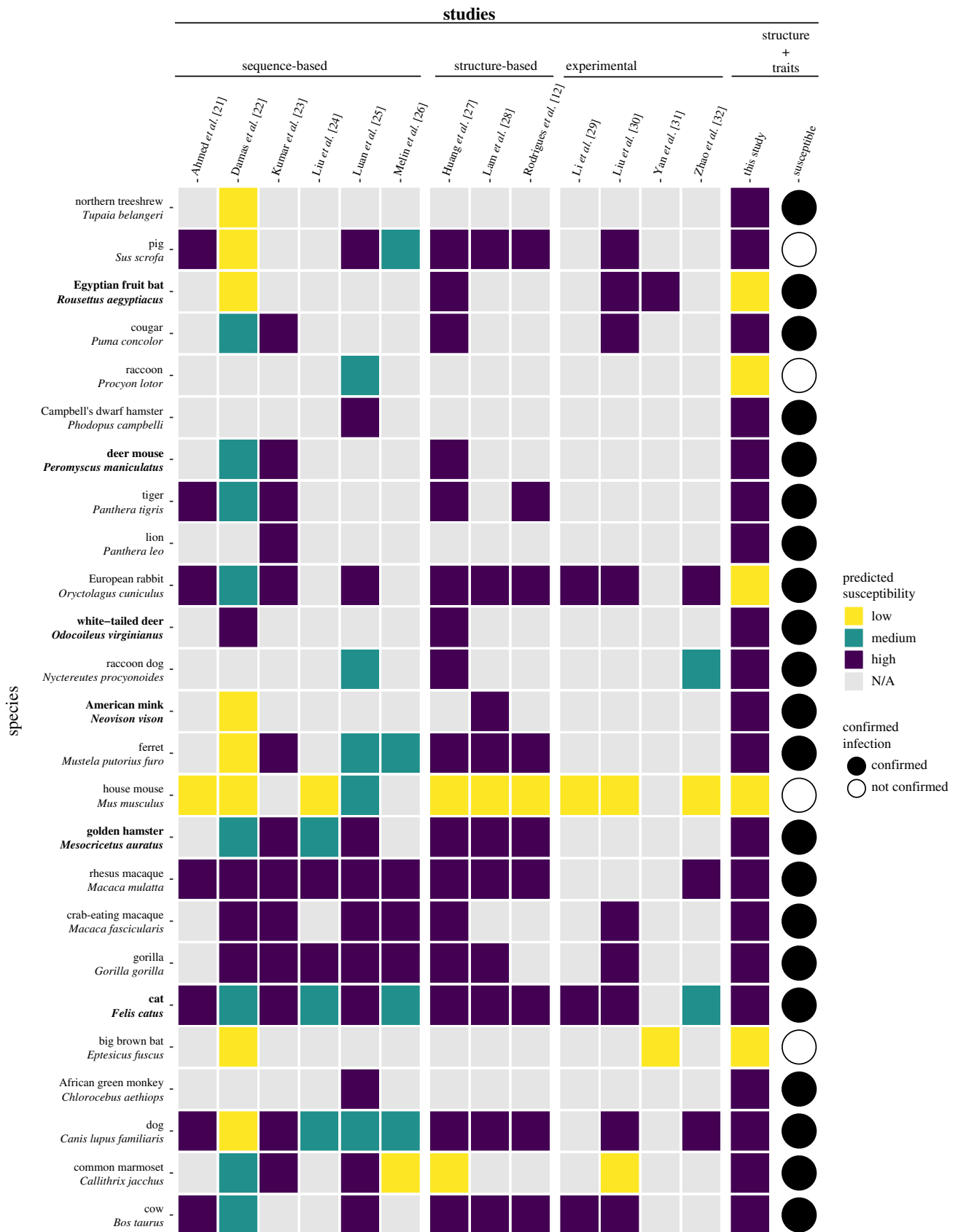


Figure 1. A heatmap summarizing predicted susceptibility to SARS-CoV-2 for species with confirmed infection from *in vivo* experimental studies or from documented natural infections. Studies that make predictions about species susceptibility are shown on the x-axis, organized by the method of prediction (those relying on ACE2 sequences, estimating binding strength using three-dimensional structures, or laboratory experiments). Predictions about zoonotic capacity from this study are listed in the second to last column, with high and low categories determined by zoonotic capacity observed in *Felis catus*. Confirmed infections for species along with the y-axis are depicted as a series of filled or unfilled circles. Bolded species have been experimentally confirmed to transmit SARS-CoV-2 to naive conspecifics. Species predictions range from warmer colours (yellow: low susceptibility or zoonotic capacity for SARS-CoV-2) to cooler colours (purple: high susceptibility or zoonotic capacity). See electronic supplementary material, Methods [20] for detailed methods about how predictions from past studies were categorized as low, medium or high. For a comparison of predictions of species susceptibility from multiple methods, including our study, see electronic supplementary material, figure S1. (Online version in colour.)

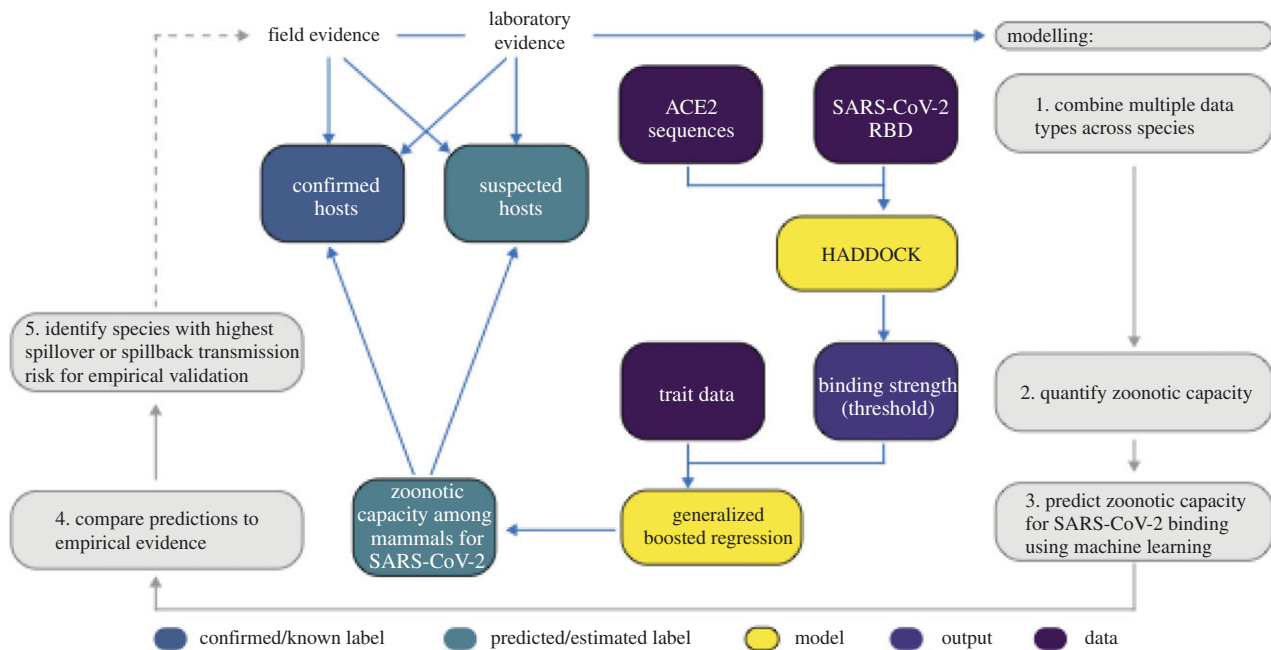


Figure 2. A flowchart showing the progression of our workflow combining evidence from limited laboratory and field studies with additional data types to predict zoonotic capacity across mammals through multi-scale statistical modelling (grey boxes, steps 1–5). For all vertebrates with published ACE2 sequences, we modelled the interface of species' ACE2 bound to the viral RBD using HADDOCK. We then combined the HADDOCK scores, which approximate binding strength, with species' trait data and trained machine learning models (generalized boosted regression) for both mammals and vertebrates. Predictive modelling of host zoonotic capacity focused on mammals only because there are currently no non-mammalian hosts for SARS-CoV-2 and imbalanced ACE2 sequences among non-mammals. Mammal species predicted to have high zoonotic capacity were then compared to results of *in vivo* experiments and *in silico* studies that applied various computational approaches. Based on predictions from our model, we identified a subset of species with particularly high risk of spillback and secondary spillover potential to prioritize additional laboratory validation and field surveillance (dashed line). (Online version in colour.)

HADDOCK software as described previously [13], with a few differences. For details on how we processed sequences, and on our structural modelling of ACE2 orthologues bound to SARS-CoV-2 spike (PDB ID: 6m0j) [51]; see electronic supplementary material, Methods [20]. For each species, we estimated binding strength based on HADDOCK score—a combination of van der Waals, electrostatics and desolvation energies. A lower (more negative) HADDOCK score predicts stronger binding between the two proteins. We hereafter refer to predicted binding strength, or simply binding strength, to indicate HADDOCK score. The HADDOCK server is freely available, and we provide code to reproduce analyses or to aid in the application of this modelling approach to similar problems [52]. Though the effects of multiple simultaneous mutations on binding affinity remain difficult to predict, HADDOCK has been used to explore how minor changes to the structure of the RBD (e.g. from point mutations or deletions in SARS-CoV-2 variants) affect viral binding [12]. These variations are accounted for by the HADDOCK software by estimating and considering their effects on both strong and weak forces that together determine both local and interfacial molecular contacts [53]. As crystal structures of novel variants become available, the modelling pipeline we present here can be applied to predict how differences in binding strengths to ACE2 orthologues affect host range and the risk of spillover transmission across species.

(b) Trait data collection and cleaning

We gathered ecological, life history, phylogenetic and biological trait data from AnAge [54], Amniote Life History Database [55], EltonTraits [56], PanTHERIA [57] and taxonomic databases [58], among other databases. Non-mammal hosts have yet to be confirmed as both susceptible and capable of onward transmission of SARS-CoV-2. Therefore, while we gathered data on certain traits across all vertebrates, we gathered data on additional traits for mammals. For mammal species, we applied boosted regression

(BRT) using the gbm package [59] in R v. 4.0.0 [60] to impute missing trait data (e.g. [44]; see electronic supplementary material, Methods for imputation methods and results). Data and descriptions of each variable can be found in the electronic supplementary material, table S2. For details on data processing, see electronic supplementary material, Methods [20].

(c) Modelling

(i) Quantifying a threshold for zoonotic capacity using HADDOCK

ACE2 binding is necessary for viral entry into host cells, but it is not sufficient for SARS-CoV-2 transmission. Multiple *in vivo* experiments suggest that some species are capable of binding SARS-CoV-2 but not capable of transmitting active infection to other individuals (e.g. cattle, *Bos taurus* [61]; bank voles, *Myodes glareolus* [62]). Viral replication and viral shedding that enables onward transmission are both required to become a suitable bridge or reservoir host for SARS-CoV-2. We constrained our predictions to species with the greatest potential for onward transmission by training our models on a binary label created using a conservative threshold of binding strength (HADDOCK score = −129). This value falls between the scores for two species: the domestic cat (*Felis catus*), currently the species with weakest predicted binding but confirmed conspecific transmission [63], and the pig (*Sus scrofa*), which shows the strongest estimated binding for which experimental inoculation failed to cause detectable infection [38]. Binding strength was binarized according to this threshold, above which it is more likely that both infection and onward transmission will occur following the results of multiple empirical studies (figure 1; electronic supplementary material, table S1). There are susceptible species whose predicted binding strength is weaker than cats, but conspecific transmission has not been confirmed in these species. For additional modelling details, see electronic supplementary material, Methods.

(d) Trait-based modelling to predict zoonotic capacity

We applied generalized BRT [64] to host trait data to predict mammal species' zoonotic capacity (for descriptions and results from additional uninformative models, including models using a vertebrate dataset, see electronic supplementary material, Methods and electronic supplementary material table S3). Code for BRT modelling, along with data for training models and making predictions, can be found at https://github.com/Han-LabDiseaseEcology/zoonotic_capacity.

3. Results

Currently available ACE2 protein sequences came from 326 species spanning eight classes and 87 orders [52]. The majority of sequences belonged to the classes Actinopterygii (22.1%), Aves (23.3%) and Mammalia (46.6%). We predicted binding strength for 299 vertebrates, including 142 mammals (electronic supplementary material, figures S2–S6). Among well-represented mammalian orders (those containing at least 10 species with binding strength predictions), Primates and Carnivora showed predicted mean binding strengths that were stronger than domestic cats (electronic supplementary material, figure S2).

We next constructed a trait-based machine learning model to predict zoonotic capacity (a binarized binding threshold) in mammals. We used the best-performing model to generate predictions of zoonotic capacity among mammal species (corrected test AUC of 0.72; for results of all other model variations see electronic supplementary material, table S3). Citation count in Web of Science, used as a proxy for study effort, had approximately 1% relative importance (electronic supplementary material, figure S7), suggesting that sampling bias across species had little influence on the model.

This model predicted 540 species spanning 13 orders to have zoonotic capacity within the 90th percentile (0.826 or higher, with a total of 2401 mammal species with prediction scores above 0.5; see electronic supplementary material, file S1 for predictions on all 5400 species [20]). Most primates were predicted to have high zoonotic capacity and collectively showed stronger viral binding compared to other mammal groups (figure 3). Additional orders predicted to have high zoonotic capacity (at least 75% of species above 0.5) include Hyracoidea (hyraxes), Perissodactyla (odd-toed ungulates), Scandentia (treeshrews), Pilosa (sloths and anteaters), Pholidota (pangolins) and non-cetacean Artiodactyla (even-toed ungulates) (figure 3). Results of model bootstrap iteration predictions can be found in the electronic supplementary material, file S1 and electronic supplementary material figure S10.

(a) Comparing model predictions to *in vivo* outcomes

These model predictions matched the experimental infection outcomes of several *in vivo* studies (figure 1). For instance, experiments on deer mice (*Peromyscus maniculatus* [65,66]) and raccoon dogs (*Nyctereutes procyonoides* [67]) confirmed SARS-CoV-2 infection and transmission to naive conspecifics. Our model also estimated a high probability of zoonotic capacity of American mink for SARS-CoV-2 (*Neovison vison*, probability = 0.83, 90th percentile), in which farmed individuals present severe infection and demonstrate the capacity to transmit to conspecifics as well as to humans [11]. Our

model also correctly predicted relatively low zoonotic capacity for big brown bats (*Eptesicus fuscus* [68]).

There were notable differences between our model results and currently available experimental studies. Our model estimated high zoonotic capacity for pigs (*Sus scrofa*, probability = 0.72, approximately 80th percentile), but *in vivo* studies report no detectable infection or onward transmission [38,69]. Similarly for cattle (*Bos taurus*), our model estimated a moderately high probability for zoonotic capacity (0.72, approximately 80th percentile), whereas in a live animal experiment, cattle were susceptible but with no onward transmission to conspecifics [61].

4. Discussion

We combined structure-based models of viral binding with species-level data on biological and ecological traits to predict the capacity of mammal species to become zoonotic hosts of SARS-CoV-2 (zoonotic capacity). Importantly, this approach extends our predictive capacity beyond the limited number of species for which ACE2 sequences are currently available. Numerous mammal species were predicted to have zoonotic capacity that meets or exceeds the viral susceptibility and transmissibility observed in experimental infections with SARS-CoV-2 (figure 1; electronic supplementary material, table S1). Many species with high model-predicted zoonotic capacity also live in human-associated habitats and overlap geographically with global COVID-19 hotspots (figure 4). Below we discuss predictions of zoonotic capacity for a number of ecologically and epidemiologically relevant categories of mammalian hosts.

(a) Captive, farmed or domesticated species

Given that contact with humans fundamentally underlies transmission risk, it is notable that our model predicted high zoonotic capacity for multiple captive species that have also been confirmed as susceptible to SARS-CoV-2. These include numerous carnivores, such as large cats from multiple zoos and pet dogs and cats. Our model also predicted high SARS-CoV-2 zoonotic capacity for many farmed and domesticated species. The water buffalo (*Bubalus bubalis*), widely kept for dairy and plowing, had the highest probability of zoonotic capacity among livestock (0.91). Model predictions in the 90th percentile also included American mink (*Neovison vison*), red fox (*Vulpes vulpes*), sika deer (*Cervus nippon*), white-lipped peccary (*Tayassu pecari*), nilgai (*Boselaphus tragocamelus*) and raccoon dogs (*Nyctereutes procyonoides*), all of which are farmed. The escape of farmed individuals into wild populations has implications for the enzootic establishment of SARS-CoV-2 [33]. These findings also have implications for vaccination strategies, for instance, prioritizing people in contact with potential bridge species (e.g. slaughterhouse workers, farmers, veterinarians).

(b) Live traded or hunted wildlife species

The *Macaca* genus comprised the majority of live-traded primates. Our model predicted high zoonotic capacity for all *Macaca* species (20/21 species, with all species within the top 10% of predictions except *M. assamensis*). Several live-traded carnivores and pangolins were also assigned high zoonotic capacity, including the Asiatic black bear (*Ursus*

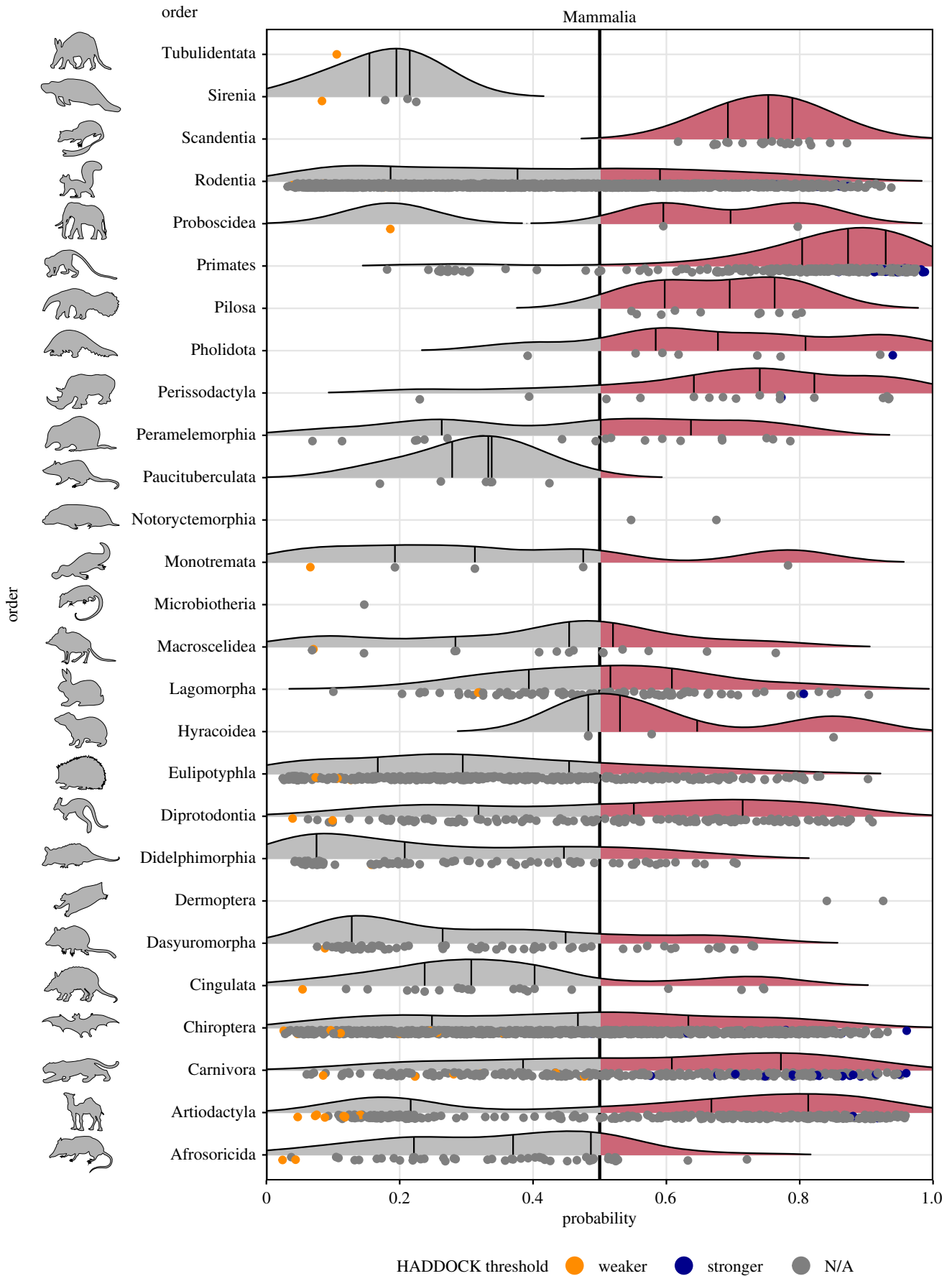


Figure 3. Ridgeline plots showing the distribution of predicted zoonotic capacity across mammals. Predicted probabilities for zoonotic capacity across the x -axis range from 0 (likely not susceptible) to 1 (zoonotic capacity predicted to be the same or greater than *Felis catus*), with the vertical line representing 0.5. The y -axis depicts all mammalian orders represented by our predictions. Density curves represent the distribution of the predictions, with those parts of the curve over 0.5 coloured pink and lines representing distribution quartiles. The predicted values for each order are shown as points below the density curves. Points that were used to train the model are coloured: orange represents species with weaker predicted binding, blue represents species with stronger predicted binding. Selected family-level distributions are shown in the electronic supplementary material, figures S8 and S9 [20]. (Online version in colour.)

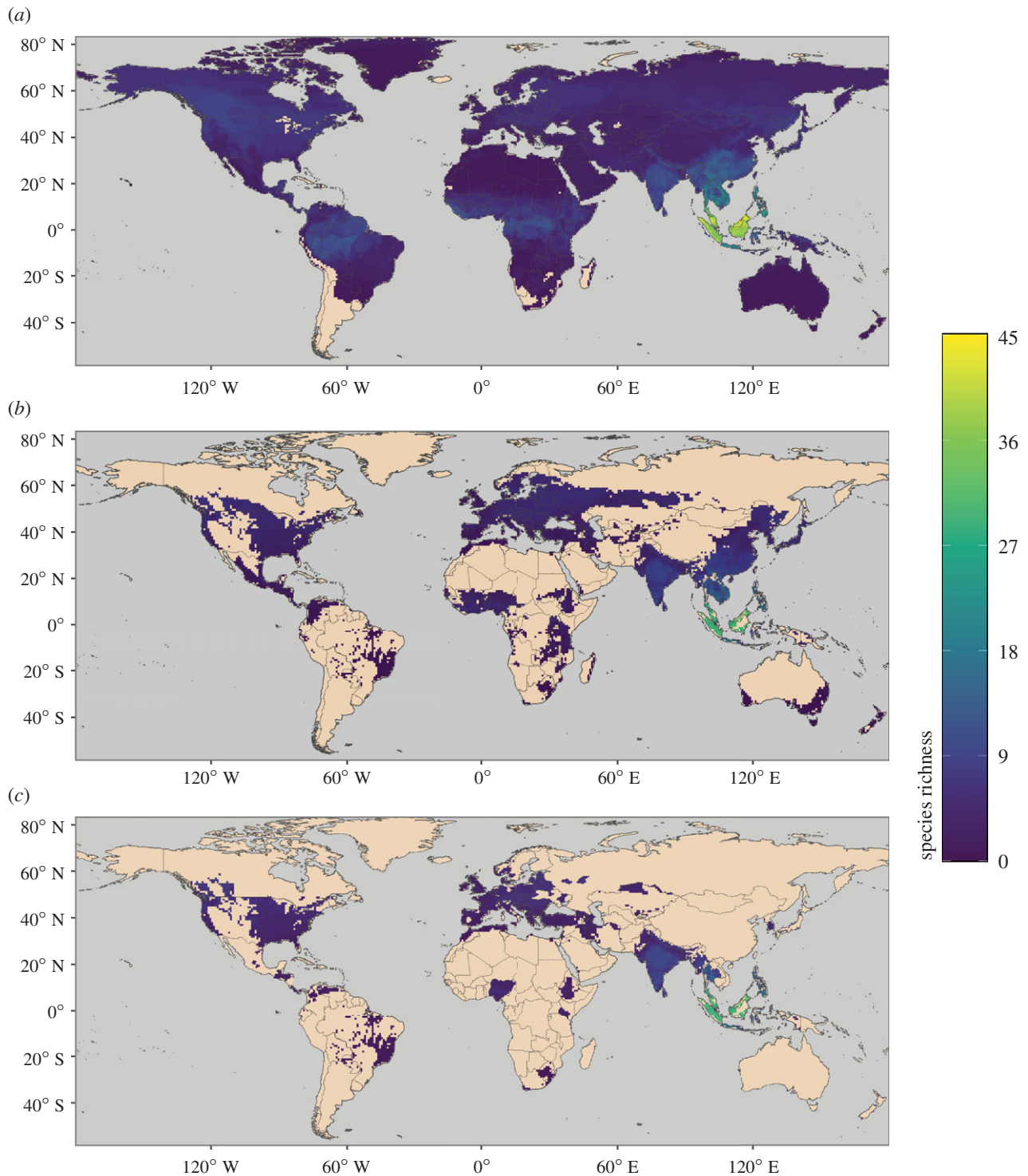


Figure 4. Maps showing the global distribution of species with the predicted capacity to transmit SARS-CoV-2; (a) depicts global species richness of the top 10 per cent of model-predicted zoonotic capacity. Geographic ranges of this subset of species were filtered to those associated with human-dominated or human-altered habitats (b), and further filtered to show the subset of species that overlaps with areas of high human SARS-CoV-2 positive case counts (over 100 000 cumulative cases as of 17 May 2021) (c). For a full list of model-predicted zoonotic capacity of species by country, see electronic supplementary material, file S2 [20]. (Online version in colour.)

thibetanus), grey wolf (*Canis lupus*) and jaguar (*Panthera onca*), and the Philippine pangolin (*Manis culionensis*) and Sunda pangolin (*M. javanica*). One of the betacoronaviruses with the highest sequence similarity to SARS-CoV-2 was isolated from Sunda pangolins [70]. Interestingly, pangolin burrows are known to be occupied by other animal species, including numerous bats [71].

Commonly hunted species in the top 10% of predictions include duiker (*Cephalophus zebra*, West Africa), warty pig (*Sus celebes*, Southeast Asia) and two deer (*Odocoileus*

hemionus and *O. virginianus*, Americas). The white-tailed deer (*O. virginianus*) was recently confirmed to transmit SARS-CoV-2 to conspecifics via aerosolized virus particles [72].

(c) Bats

Our model identified 35 bat species within the 90th percentile of zoonotic capacity. Within the genus *Rhinolophus*, our model identified the large rufous horseshoe bat (*Rhinolophus*

rufus) as having the highest probability of zoonotic capacity (0.89). *Rhinolophus rufus* is a known natural host for bat betacoronaviruses [73] and a congener to three other horseshoe bats harbouring betacoronaviruses with high nucleotide sequence similarity to SARS-CoV-2 (approx. 92–96%) [6,74,75]. For these other three species, our model assigned a range of probabilities for SARS-CoV-2 zoonotic capacity (*Rhinolophus affinis* (0.58), *R. malayanus* (0.70) and *R. shameli* (0.71)) and also predicted relatively high probabilities for two congeners, *Rhinolophus acuminatus* (0.84) and *R. macrotis* (0.70). These predictions agree with recent experiments demonstrating efficient viral binding of SARS-CoV-2 RBD for *R. macrotis* [76] and confirmation of SARS-CoV-2-neutralizing antibodies in field-caught *R. acuminatus* harbouring a closely related betacoronavirus [77].

Our model also identified 17 species in the genus *Pteropus* (flying foxes) with high probabilities of zoonotic capacity for SARS-CoV-2. Some of these species are confirmed reservoirs of other zoonotic viruses (e.g. henipaviruses in *P. lylei*, *P. vampyrus*, *P. conspicillatus* and *P. alecto*), with Southeast Asia also having the most mammal species with a high predicted zoonotic capacity (figure 4). Annual outbreaks attributed to spillover transmission from bats illustrate a persistent epizootic risk to humans [78–80] and confirm that gaps in systematic surveillance of zoonotic viruses, including betacoronaviruses, remain an urgent priority (e.g. [81]).

(d) Rodents

Our model identified 76 rodent species with high zoonotic capacity. Among these are the deer mouse (*Peromyscus maniculatus*) and white-footed mouse (*P. leucopus*), which are reservoirs for multiple zoonotic pathogens and parasites in North America [82–84]. Experimental infection, viral shedding and sustained intraspecific transmission of SARS-CoV-2 were recently confirmed for *P. maniculatus* [65,66]. Also in the top 10% were two rodents considered to be human commensals whose geographic ranges are expanding due to human activities: *Rattus argentiventer* (0.84) and *R. tiomanicus* (0.79) (electronic supplementary material, file S1) [85–87]. It is notable that many of these rodent species are preyed upon by carnivores, such as the red fox (*Vulpes vulpes*) or domestic cats (*Felis catus*) who themselves were predicted to have high zoonotic capacity by our model.

(e) Strengthening predictive capacity for zoonoses

While there was a wide agreement between our model predictions and empirical studies, examining mismatches between experimental results and model-generated predictions focuses attention on characterizing what factors underlie these disconnects. For instance, this study and others predicted that pigs (*Sus scrofa*) would be susceptible to SARS-CoV-2 (figure 1), but these predictions have not been supported by whole-animal inoculations [38,69]. Similarly, SARS-CoV-2 replicated in adult cattle but onward transmission has not been observed *in vivo* [61].

Disconnects between real-world observations, *in vivo* experimental results, and *in silico* predictions of zoonotic capacity offer insight upon which to iterate further study. For instance, mismatches may arise because host susceptibility and transmission capacity are necessary but not sufficient for zoonotic risk to be realized in natural settings. These processes are embedded in a broader ecological context that impacts host susceptibility, intra-host infection dynamics

(latency, recrudescence, tolerance) and viral persistence that collectively determine where and when viral shedding and spillover will occur [88–91]. Infection processes also depend strongly on the cellular environments in which cell entry and viral replication take place (e.g. the presence of key proteases [7]), and on host immunogenicity [91], factors which are themselves influenced by the environment [92]. Insofar as data limitations preclude perfect computational predictions of zoonotic capacity (e.g. limited ACE2 sequences, crystal structures or trait data), laboratory experiments are also limited. For SARS-CoV-2 and other host–pathogen systems, animals that are readily infected in the laboratory may be less susceptible elsewhere (ferrets in the laboratory versus mixed results in ferrets as pets [69,93,94]; rabbits in the laboratory versus rabbits as pets [95,96]). Moreover, wild-life hosts confirmed to shed multiple zoonotic viruses in nature (e.g. bats [97]) can be much less tractable for whole-animal laboratory investigations (for instance, requiring high biosecurity containment and very limited sample sizes). While laboratory experiments are critical for understanding mechanisms of pathogenesis and disease, without field surveillance they offer imperfect reflections of zoonotic capacity realized in natural settings.

These examples illustrate that there is no single methodology sufficient to understand and predict zoonotic transmission, for SARS-CoV-2 or any zoonotic pathogen. They also demonstrate the need to capitalize on underused or disconnected data sources, such as natural history collections, which are well-positioned to fill knowledge gaps about the spatial and temporal extents of animal hosts and their pathogens [98,99]. Integration of methods and data across biological scales creates avenues to more efficient iteration between computational predictions, laboratory experiments and targeted animal surveillance necessary to connect transmission mechanisms to the broader conditions underpinning zoonotic disease emergence in nature.

Data accessibility. All data, code and models required for recreating our structural modelling analysis with HADDOCK are available from Zenodo: <https://doi.org/10.5281/zenodo.4517509> [52]. All data and code required for BRT modelling using species trait data are available from GitHub: https://github.com/HanLabDiseaseEcology/zoonotic_capacity. Additional methods, tables, figures and files of our model predictions can be found in the supplemental material of this article and on Figshare: <https://doi.org/10.25390/caryinstitute.c.5293339.v5> [20].

Authors' contributions. I.R.F., A.A.C. and B.A.H. were responsible for conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, and writing, reviewing and editing the manuscript. J.P.G.L.M.R. contributed additional analysis. J.P.G.L.M.R., A.V. and B.A.H. contributed funding. A.V. and A.A.C. created all figures. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the NSF EEID program (grant no. DEB 1717282), DARPA PREEMPT program (grant no. D18AC00031), CREATE-NEO, a member of the NIH NIAID CREID program (grant no. 1U01 AI151807-01) and the NVIDIA Corporation GPU grant program (B.A.H.); by the NSF Polar program (grant no. OPP 1935870 and 1947040) (A.V.); and by NIH NIGMS (grant no. R35GM122543) (J.P.G.L.M.R.).

Acknowledgements. We are grateful for discussions with Drs. Alexandre Bonvin, Dennis Bente, Susan Hafenstein, Kathryn Hanley, Hyunwook Lee, Colin Parrish and John Paul Schmidt about various components of this project.

1. Dong E, Du H, Gardner L. 2020 An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534. (doi:10.1016/S1473-3099(20)30120-1)
2. World Health Organization. 2021 WHO coronavirus disease (COVID-19) dashboard. See <https://covid19.who.int>.
3. Keele BF *et al.* 2006 Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* **313**, 523–526. (doi:10.1126/science.1126531)
4. Gage KL, Kosoy MY. 2005 Natural history of plague: perspectives from more than a century of research. *Annu. Rev. Entomol.* **50**, 505–528. (doi:10.1146/annurev.ento.50.071803.130337)
5. Taubenberger JK, Reid AH, Lourens RM, Wang R, Jin G, Fanning TG. 2005 Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**, 889. (doi:10.1038/nature04230)
6. Zhou P *et al.* 2020 A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273. (doi:10.1038/s41586-020-2012-7)
7. Letko M, Marzi A, Munster V. 2020 Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol.* **5**, 562–569. (doi:10.1038/s41564-020-0688-y)
8. Chou CF *et al.* 2006 ACE2 orthologues in non-mammalian vertebrates (*Danio*, *Gallus*, *Fugu*, *Tetraodon* and *Xenopus*). *Gene* **377**, 46–55. (doi:10.1016/j.gene.2006.03.010)
9. Guth S, Visher E, Boots M, Brook CE. 2019 Host phylogenetic distance drives trends in virus virulence and transmissibility across the animal-human interface. *Phil. Trans. R. Soc. B* **374**, 20190296. (doi:10.1098/rstb.2019.0296)
10. World Health Organization. 2020 SARS-CoV-2 mink-associated variant strain—Denmark. See www.who.int/home/cms-decommissioning.
11. Oude Munnink BB *et al.* 2020 Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science* **37**, 172–177. (doi:10.1126/science.abe5901)
12. Rodrigues JPGLM, Barrera-Vilmarau S, M C Teixeira J, Sorokina M, Seckel E, Kastitis PL, Levitt M. 2020 Insights on cross-species transmission of SARS-CoV-2 from structural modeling. *PLoS Comput. Biol.* **16**, e1008449. (doi:10.1371/journal.pcbi.1008449)
13. Davies NG *et al.* 2021 Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **372**, eabg3055. (doi:10.1126/science.abg3055)
14. Volz E *et al.* 2021 Transmission of SARS-CoV-2 lineage B.1.1.7 in England: insights from linking epidemiological and genetic data. *medRxiv*, 2020.12.30.20249034.
15. Rambaut A *et al.* 2020 Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. *Virological*. See <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563> (accessed on 28 January 2021).
16. Tegally H *et al.* 2020 Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv*, 2020.12.21.20248640.
17. Van Egeren D *et al.* 2021 Risk of rapid evolutionary escape from biomedical interventions targeting SARS-CoV-2 spike protein. *PLoS ONE* **16**, e0250780. (doi:10.1371/journal.pone.0250780)
18. Bao L *et al.* 2020 The pathogenicity of SARS-CoV-2 in hACE2 transgenic mice. *Nature* **583**, 830–833. (doi:10.1038/s41586-020-2312-y)
19. Montagutelli X *et al.* 2021 The B.1.351 and P.1 variants extend SARS-CoV-2 host range to mice. *bioRxiv*. 2021.03.18.436013. (doi:10.1101/2021.03.18.436013)
20. Fischhoff IR, Castellanos AA, Rodrigues JPGLM, Varsani A, Han BA. 2021 Predicting the zoonotic capacity of mammals to transmit SARS-CoV-2. Figshare.
21. Ahmed R, Hasan R, Siddiki AMAMZ, Islam MS. 2021 Host range projection of SARS-CoV-2: South Asia perspective. *Infect. Genet. Evol.* **87**, 104670. (doi:10.1016/j.meegid.2020.104670)
22. Damas J *et al.* 2020 Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. *Proc. Natl Acad. Sci. USA* **117**, 22 311–22 322. (doi:10.1073/pnas.2010146117)
23. Kumar A, Pandey SN, Pareek V, Narayan RK, Faiq MA, Kumari C. 2020 Predicting susceptibility for SARS-CoV-2 infection in domestic and wildlife animals using ACE2 protein sequence homology. *Zoo Biol.* **40**, 79–85. (doi:10.1002/zoo.21576)
24. Liu Z *et al.* 2020 Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. *J. Med. Virol.* **92**, 595–601. (doi:10.1002/jmv.25726)
25. Luan J, Jin X, Lu Y, Zhang L. 2020 SARS-CoV-2 spike protein favors ACE2 from Bovidae and Cricetidae. *J. Med. Virol.* **92**, 1649–1656. (doi:10.1002/jmv.25817)
26. Melin AD, Janiak MC, Marrone III F, Arora PS, Higham JP. 2020 Comparative ACE2 variation and primate COVID-19 risk. *Commun. Biol.* **3**, 641. (doi:10.1038/s42003-020-01370-w)
27. Huang X, Zhang C, Pearce R, Omenn GS, Zhang Y. 2020 Identifying the zoonotic origin of SARS-CoV-2 by modeling the binding affinity between the spike receptor-binding domain and host ACE2. *J. Proteome Res.* **19**, 4844–4856. (doi:10.1021/acs.jproteome.0c00717)
28. Lam SD *et al.* 2020 SARS-CoV-2 spike protein predicted to form complexes with host receptor protein orthologues from a broad range of mammals. *Sci. Rep.* **10**, 16471. (doi:10.1038/s41598-020-71936-5)
29. Li Y *et al.* 2020 SARS-CoV-2 and three related coronaviruses utilize multiple ACE2 orthologs and are potently blocked by an improved ACE2-Ig. *J. Virol.* **94**, e01283–20. (doi:10.1128/JVI.01283-20)
30. Liu Y *et al.* 2021 Functional and genetic analysis of viral receptor ACE2 orthologs reveals a broad potential host range of SARS-CoV-2. *Proc. Natl Acad. Sci. USA* **118**, e2025373118. (doi:10.1073/pnas.2025373118)
31. Yan H *et al.* 2021 ACE2 receptor usage reveals variation in susceptibility to SARS-CoV and SARS-CoV-2 infection among bat species. *Nature Ecol. Evol.* **5**, 600–608. (doi:10.1038/s41559-021-01407-1)
32. Zhao X *et al.* 2020 Broad and differential animal angiotensin-converting enzyme 2 receptor usage by SARS-CoV-2. *J. Virol.* **94**, e01283–20. (doi:10.1128/JVI.01283-20)
33. DeLiberto T, Shriner S. 2020 PRO/AH/EDR> COVID-19 update (536): animal, USA (UT) wild mink, 1st case. ProMED. See <https://promedmail.org/promed-post/?id=8015608>.
34. ODA. 2020 Mink at affected Oregon farm negative for SARS-CoV-2, wildlife surveillance continues. 23 December. See <https://odanews.wpengine.com/mink-at-affected-oregon-farm-negative-for-sars-cov-2-wildlife-surveillance-continues/>.
35. Shriner SA *et al.* 2021 SARS-CoV-2 exposure in escaped mink, Utah, USA. *Emerg. Infect. Dis.* **27**, 988–990. (doi:10.3201/eid2703.204444)
36. Mathavarajah S, Stoddart AK, Gagnon GA, Dellaire G. 2021 Pandemic danger to the deep: the risk of marine mammals contracting SARS-CoV-2 from wastewater. *Sci. Total Env.* **760**, 143346. (doi:10.1016/j.scitotenv.2020.143346)
37. Deng W *et al.* 2020 Ocular conjunctival inoculation of SARS-CoV-2 can cause mild COVID-19 in rhesus macaques. *Nature Communications* **11**, 4400. (doi:10.1038/s41467-020-18149-6)
38. Shi J *et al.* 2020 Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. *Science* **368**, 1016–1020. (doi:10.1126/science.abb7015)
39. Rodrigues JPGLM *et al.* 2013 Defining the limits of homology modeling in information-driven protein docking. *Proteins* **81**, 2119–2128. (doi:10.1002/prot.24382)
40. Sander C, Schneider R. 1991 Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* **9**, 56–68. (doi:10.1002/prot.340090107)
41. Fournier D, Luft FC, Bader M, Ganten D, Andrade-Navarro MA. 2012 Emergence and evolution of the

- renin-angiotensin-aldosterone system. *J. Mol. Med.* **90**, 495–508. (doi:10.1007/s00109-012-0894-z)
42. Han BA, Schmidt JP, Bowden SE, Drake JM. 2015 Rodent reservoirs of future zoonotic diseases. *Proc. Natl Acad. Sci. USA* **112**, 7039–7044. (doi:10.1073/pnas.1501598112)
43. Yang LH, Han BA. 2018 Data-driven predictions and novel hypotheses about zoonotic tick vectors from the genus *Ixodes*. *BMC Ecol.* **18**, 7. (doi:10.1186/s12898-018-0163-2)
44. Han BA, O'Regan SM, Paul Schmidt J, Drake JM. 2020 Integrating data mining and transmission theory in the ecology of infectious diseases. *Ecol. Lett.* **23**, 1178–1188. (doi:10.1111/ele.13520)
45. Han BA, Schmidt JP, Alexander LW, Bowden SE, Hayman DTS, Drake JM. 2016 Undiscovered bat hosts of Filoviruses. *PLoS Negl. Trop. Dis.* **10**, e0004815. (doi:10.1371/journal.pntd.0004815)
46. Han BA *et al.* 2019 Confronting data sparsity to identify potential sources of Zika virus spillover infection among primates. *Epidemics* **27**, 59–65. (doi:10.1016/j.epidem.2019.01.005)
47. Yang X-L *et al.* 2017 Genetically diverse Filoviruses in *Rousettus* and *Eonycteris* spp. bats, China, 2009 and 2015. *Emerg. Infect. Dis.* **23**, 482–486. (doi:10.3201/eid2303.161119)
48. Goldstein T *et al.* 2018 The discovery of Bombali virus adds further support for bats as hosts of ebolaviruses. *Nat. Microbiol.* **3**, 1084–1089. (doi:10.1038/s41564-018-0227-2)
49. Sorokina M, Teixeira J, Barrera-Vilmarau S, Paschke R, Papatotiriou I, Rodrigues JPGLM, Kastirits PL. 2020 Structural models of human ACE2 variants with SARS-CoV-2 spike protein for structure-based drug design. *Sci. Data* **7**, 309. (doi:10.1038/s41597-020-00652-6)
50. Restif O *et al.* 2012 Model-guided fieldwork: practical guidelines for multidisciplinary research on wildlife ecological and epidemiological dynamics. *Ecol. Lett.* **15**, 1083–1094. (doi:10.1111/j.1461-0248.2012.01836.x)
51. Lan J *et al.* 2020 Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220. (doi:10.1038/s41586-020-2180-5)
52. Fischhoff I, Castellanos AA, Rodrigues J, Varsani A, Han B. 2021 Models of ACE2 orthologs interacting with SARS-CoV-2 Spike RBD (canonical variant). (doi:10.5281/zenodo.4517509)
53. Bonvin A, Rodrigues J. 2021 HADDOCK2.4 manual. See www.bonvinlab.org/software/haddock2.4/ introduction (accessed on 19 October 2021).
54. de Magalhães JP, Costa J. 2009 A database of vertebrate longevity records and their relation to other life-history traits. *J. Evol. Biol.* **22**, 1770–1774. (doi:10.1111/j.1420-9101.2009.01783.x)
55. Myhrvold NP, Baldridge E, Chan B, Sivam D, Freeman DL, Ernest SKM. 2015 An amniote life-history database to perform comparative analyses with birds, mammals, and reptiles: ecological archivesE096-269. *Ecology* **96**, 3109. (doi:10.1890/15-0846R.1)
56. Wilman H, Belmaker J, Simpson J, de la Rosa C, Rivadeneira MM, Jetz W. 2014 EltonTraits 1.0: species-level foraging attributes of the world's birds and mammals. *Ecology* **95**, 2027. (doi:10.1890/13-1917.1)
57. Jones KE *et al.* 2009 PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals: Ecological Archives E090-184. *Ecology* **90**, 2648. (doi:10.1890/08-1494.1)
58. Wilson DE, Reeder DM. 2005 *Mammal species of the world: a taxonomic and geographic reference*. Baltimore, MD: JHU Press.
59. Greenwell B, Boehmke B, Cunningham J, GBM Developers. 2020 Generalized boosted regression models. See <https://cran.r-project.org/web/packages/gbm/index.html>.
60. R Core Team. 2020 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See www.R-project.org.
61. Ulrich L, Wernike K, Hoffmann D, Mettenleiter TC, Beer M. 2020 Experimental infection of cattle with SARS-CoV-2. *Emerg. Infect. Dis.* **26**, 2979–2981. (doi:10.3201/eid2612.203799)
62. Ulrich L, Michelitsch A, Halwe N, Wernike K, Hoffmann D, Beer M. 2021 Experimental SARS-CoV-2 infection of bank voles. *Emerg. Infect. Dis.* **27**, 1193–1195. (doi:10.3201/eid2704.204945)
63. Bosco-Lauth AM *et al.* 2020 Experimental infection of domestic dogs and cats with SARS-CoV-2: pathogenesis, transmission, and response to reexposure in cats. *Proc. Natl Acad. Sci. USA* **117**, 26 382–26 388. (doi:10.1073/pnas.2013102117)
64. Elith J, Leathwick JR, Hastie T. 2008 A working guide to boosted regression trees. *J. Anim. Ecol.* **77**, 802–813. (doi:10.1111/j.1365-2656.2008.01390.x)
65. Fagre A *et al.* 2021 SARS-CoV-2 infection, neuropathogenesis and transmission among deer mice: Implications for spillback to New World rodents. *PLoS Pathog.* **17**, e1009585. (doi:10.1371/journal.ppat.1009585)
66. Griffin BD *et al.* 2021 SARS-CoV-2 infection and transmission in the North American deer mouse. *Nat. Commun.* **12**, 3612. (doi:10.1038/s41467-021-23848-9)
67. Freuling CM *et al.* 2020 Susceptibility of raccoon dogs for experimental SARS-CoV-2 infection. *Emerg. Infect. Dis.* **26**, 2982–2985. (doi:10.3201/eid2612.203733)
68. Hall JS *et al.* In press. Experimental challenge of a North American bat species, big brown bat (*Eptesicus fuscus*), with SARS-CoV-2. *Transbound. Emerg. Dis.* (doi:10.1111/tbed.13949)
69. Schlottau K *et al.* 2020 SARS-CoV-2 in fruit bats, ferrets, pigs, and chickens: an experimental transmission study. *Lancet Microbe* **1**, e218–e225. (doi:10.1016/S2666-5247(20)30089-6)
70. Lam TTY *et al.* 2020 Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* **583**, 282–285. (doi:10.1038/s41586-020-2169-0)
71. Lehmann D, Halbwax ML, Makaga L, Whytock R, Ndindiwe Malata L, Bombenda Mouele W, Momboua BR, Koumba Pambo AF, White LJ. 2020 Pangolins and bats living together in underground burrows in Lopé National Park, Gabon. *Afr. J. Ecol.* **58**, 540–542. (doi:10.1111/aje.12759)
72. Palmer MV *et al.* 2021 Susceptibility of white-tailed deer (*Odocoileus virginianus*) to SARS-CoV-2. *J. Virol.* **95**, e00083-21. (doi:10.1128/JVI.00083-21)
73. Tsuda S *et al.* 2012 Genomic and serological detection of bat coronavirus from bats in the Philippines. *Arch. Virol.* **157**, 2349–2355. (doi:10.1007/s00705-012-1410-z)
74. Zhou H *et al.* 2020 A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Curr. Biol.* **30**, 2196–2203.e3. (doi:10.1016/j.cub.2020.05.023)
75. Hul V *et al.* 2021 A novel SARS-CoV-2 related coronavirus in bats from Cambodia. *bioRxiv.*, 2021.01.26.428212. (doi:10.1101/2021.01.26.428212)
76. Mou H *et al.* 2021 Mutations derived from horseshoe bat ACE2 orthologs enhance ACE2-Fc neutralization of SARS-CoV-2. *PLoS Pathogens* **17**, e1009501. (doi:10.1371/journal.ppat.1009501)
77. Wacharapluesadee S *et al.* 2021 Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. *Nat. Commun.* **12**, 972. (doi:10.1038/s41467-021-21240-1)
78. Pulliam JRC *et al.* 2012 Agricultural intensification, priming for persistence and the emergence of Nipah virus: a lethal bat-borne zoonosis. *J. R. Soc. Interface* **9**, 89–101. (doi:10.1098/rsif.2011.0223)
79. Plowright RK *et al.* 2015 Ecological dynamics of emerging bat virus spillover. *Proc. R. Soc. B* **282**, 20142124. (doi:10.1098/rspb.2014.2124)
80. Kessler MK *et al.* 2018 Changing resource landscapes and spillover of henipaviruses. *Ann. N. Y. Acad. Sci.* **1429**, 78–99. (doi:10.1111/nyas.13910)
81. Peel AJ, Field HE, Aravena MR, Edson D, McCallum H, Plowright RK, Prada D. 2020 Coronaviruses and Australian bats: a review in the midst of a pandemic. *Aust. J. Zool.* **67**, 346–360. (doi:10.1071/ZO20046)
82. Bordes F, Blasdel K, Morand S. 2015 Transmission ecology of rodent-borne diseases: new frontiers. *Integr. Zool.* **10**, 424–435. (doi:10.1111/1749-4877.12149)
83. Ostfeld RS, Canham CD, Oggenfuss K, Winchcombe RJ, Keesing F. 2006 Climate, deer, rodents, and acorns as determinants of variation in lyme-disease risk. *PLoS Biol.* **4**, e145. (doi:10.1371/journal.pbio.0040145)
84. Machtinger ET, Williams SC. 2020 Practical guide to trapping *Peromyscus leucopus* (Rodentia: Cricetidae) and *Peromyscus maniculatus* for vector and vector-borne pathogen surveillance and ecology. *J. Insect Sci.* **20**, 5. (doi:10.1093/jisesa/ieaa028)

85. Morand S *et al.* 2015 Global parasite and *Rattus* rodent invasions: the consequences for rodent-borne diseases. *Integr. Zool.* **10**, 409–423. (doi:10.1111/1749-4877.12143)
86. Hamdan NES, Ng YL, Lee WB, Tan CS, Khan FAA, Chong YL. 2017 Rodent species distribution and Hantavirus Seroprevalence in residential and forested areas of Sarawak, Malaysia. *Trop. Life Sci. Res.* **28**, 151–159. (doi:10.21315/tlsr2017.28.1.11)
87. Louys J, Herrera MB, Thomson VA, Wiewel AS, Donnellan SC, O'Connor S, Aplin K. 2020 Expanding population edge craniometrics and genetics provide insights into dispersal of commensal rats through Nusa Tenggara, Indonesia. *Rec. Aust. Mus.* **72**, 287–302. (doi:10.3853/j.2201-4349.72.2020.1730)
88. Becker DJ, Washburne AD, Faust CL, Pulliam JRC, Mordecai EA, Lloyd-Smith JO, Plowright RK. 2019 Dynamic and integrative approaches to understanding pathogen spillover. *Phil. Trans. R. Soc. B* **374**, 20190014. (doi:10.1098/rstb.2019.0014)
89. Plowright RK, Parrish CR, McCallum H, Hudson PJ, Ko AI, Graham AL, Lloyd-Smith JO. 2017 Pathways to zoonotic spillover. *Nat. Rev. Microbiol.* **15**, 502–510. (doi:10.1038/nrmicro.2017.45)
90. Morris DH *et al.* 2020 The effect of temperature and humidity on the stability of SARS-CoV-2 and other enveloped viruses. *bioRxiv*. (doi:10.1101/2020.10.16.341883)
91. Bean AGD, Baker ML, Stewart CR, Cowled C, Deffrasnes C, Wang LF, Lowenthal JW. 2013 Studying immunity to zoonotic diseases in the natural host—keeping it real. *Nat. Rev. Immunol.* **13**, 851–861. (doi:10.1038/nri3551)
92. Letko M, Seifert SN, Olival KJ, Plowright RK, Munster VJ. 2020 Bat-borne virus diversity, spillover and emergence. *Nat. Rev. Microbiol.* **18**, 461–471. (doi:10.1038/s41579-020-0394-z)
93. OIE. 2021 Events in animals. See <https://www.oie.int/en/scientific-expertise/specific-information-and-recommendations/questions-and-answers-on-2019-novel-coronavirus/events-in-animals/> (accessed on 28 January 2021).
94. Sawatzki K, Hill NJ, Puryear WB, Foss AD, Stone JJ, Runstadler JA. 2021 Host barriers to SARS-CoV-2 demonstrated by ferrets in a high-exposure domestic setting. *Proc. Natl Acad. Sci. USA* **118**, e2025601118. (doi:10.1073/pnas.2025601118)
95. Mykytyn AZ *et al.* 2021 Susceptibility of rabbits to SARS-CoV-2. *Emerg. Microbes Infect.* **10**, 1–7. (doi:10.1080/22221751.2020.1868951)
96. Ruiz-Arrondo I, Portillo A, Palomar AM, Santibanez S, Santibanez P, Cervera C, Oteo JA. 2020 Detection of SARS-CoV-2 in pets living with COVID-19 owners diagnosed during the COVID-19 lockdown in Spain: a case of an asymptomatic cat with SARS-CoV-2 in Europe. *bioRxiv*. (doi:10.1101/2020.05.14.20101444)
97. Peel AJ *et al.* 2019 Synchronous shedding of multiple bat paramyxoviruses coincides with peak periods of Hendra virus spillover. *Emerg. Microbes Infect.* **8**, 1314–1323. (doi:10.1080/22221751.2019.1661217)
98. Cook JA *et al.* 2020 Integrating biodiversity infrastructure into pathogen discovery and mitigation of emerging infectious diseases. *Bioscience* **70**, 531–534. (doi:10.1093/biosci/biaa064)
99. Thompson CW *et al.* 2021 Preserve a voucher specimen! The critical need for integrating natural history collections in infectious disease studies. *MBio* **12**, e02698–20. (doi:10.1128/mBio.02698-20)