DISSERTATION

SARS-COV-2 EVOLUTION AND WITHIN-HOST VARIATION IN NONHUMAN ANIMALS

Submitted by

Laura Bashor

Department of Microbiology, Immunology and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2024

Doctoral Committee:

Advisor: Sue VandeWoude

Mark Stenglein
Angela Bosco-Lauth
Dan Sloan
Roderick B. Gagne

ABSTRACT

SARS-COV-2 EVOLUTION AND WITHIN-HOST VARIATION IN NONHUMAN ANIMALS

The COVID-19 pandemic originated following spillover of SARS-CoV-2 from non-human animals into humans. Despite concentrated efforts before and after the pandemic, current research is constrained by the impracticality of witnessing initial host shift events and transmission dynamics that shape infectious disease emergence. SARS-CoV-2 transmission from humans to a range of domestic and wild species has been well documented; furthermore, spillback into humans from white-tailed deer, mink, hamsters, domestic cats, and lions has also been reported. SARS-CoV-2, like other RNA viruses, has the ability to adapt rapidly following host shifts. These cross-species transmission events can accelerate novel variant emergence through selection for genetic variation that improves virus fitness in a novel host environment.

To evaluate the possibility that cross-species transmission accelerates SARS-CoV-2 evolution and variant emergence, we sequenced viral genomes recovered from experimentally and naturally infected animals to characterize within-host virus populations. We demonstrated the use of experimental exposure studies as a controlled system to test hypotheses surrounding SARS-CoV-2 adaptation in cats (*Felis catus*), dogs (*Canis lupus familiaris*), hamsters (*Mesocricetus auratus*), ferrets (*Mustela putorius furo*), deer mice (*Peromyscus maniculatus*), bushy-tailed woodrats (*Neotoma cinerea*), Brazilian free-tailed bats (*Tadarida brasiliensis*), striped skunks (*Mephitis mephitis*), red foxes (*Vulpes vulpes*) and mule deer (*Odocoileus hemionus*). We also evaluated publicly available sequencing data from infected felids, and investigated within-host dynamics in natural infections of Amur tigers (*Panthera tigris altaica*), African lions (*Panthera leo*), and spotted hyenas (*Crocuta crocuta*) in a zoo environment.

Our initial work investigated SARS-CoV-2 evolution across three passages in Vero cells and experimentally infected cats (n = 6), dogs (n = 3), hamsters (n = 3), and a ferret (n = 1). We

observed the rapid selection and fixation of five SARS-CoV-2 mutations in Vero cells, followed by their reversion in dogs, cats and hamsters 1-3 days post-infection. We noted 14 emergent variants across the SARS-CoV-2 genome, including increased variation in the SARS-CoV-2 spike protein. Emergent variants included mutations not detected in the original virus stocks used for inoculation, and several defining mutations of variant lineages of concern in humans. Finally, we noted increased signs of adaptation in dogs, which did not shed infectious virus, including six nonsynonymous mutations in the SARS-CoV-2 open-reading frames (ORFs) encoding proteins for virus replication. In particular, this work underscored the potential for accelerated viral evolution in cell culture systems used commonly in virological research. This work has been published and represents Chapter 2 of this dissertation.

Our next study built upon this work by investigating SARS-CoV-2 evolution in three experimental cohorts of domestic cats (n=23) infected through direct inoculation and cat-to-cat contact transmission. We observed high numbers of within-host variants in SARS-CoV-2 genomes recovered from cats compared to what is documented in humans, over half of which were nonsynonymous changes. The number of variants detected was positively correlated with the experimental dose of virus inoculum, and fewer variants were observed in contact cats. Similar to the previous study, mutations occurring at the same positions as defining VOC mutations, and signatures of positive selection in the viral spike (S) gene were observed. Our concurrent analysis of publicly available SARS-CoV-2 sequences showed no evidence for independent evolutionary trajectories associated with natural infections of domestic cats or other felids, and confirmed susceptibility of felids to the breadth of variants circulating in human populations. This work has also been published and represents Chapter 3 of this dissertation.

We subsequently investigated SARS-CoV-2 evolution in longitudinal samples collected from Amur tigers (n=2), African lions (n=11), and spotted hyenas (n=4) infected during an outbreak at the Denver Zoo. Longitudinal nasal swabs were collected from infected individuals over an approximately three-month sampling period. We determined that the outbreak was caused by a single introduction of the Delta sublineage AY.20, which was a rare variant circulating in human

populations at the time. We inferred a transmission chain from tigers to lions to hyenas, which was consistent with the appearance of clinical signs in infected animals. We observed expansion and diversification of within-host virus populations, and signatures of both purifying and positive selection. The strongest signs of positive selection were evident in the viral nucleocapsid (N) gene, and in viruses recovered from hyenas. Four candidate species-specific adaptive mutations, two of which are in the N gene, were identified in lions and hyenas (N A254V) and hyenas alone (ORF1ab E1724D, S T274I, and N P326). This work is presented in Chapter 4 of this dissertation.

In Chapter 5, we evaluated a large dataset of peridomestic wildlife species experimentally infected with two SARS-CoV-2 variants, WA01 and Delta. Study species included deer mice (n=3), bushy-tailed woodrats (n=3), Brazilian free-tailed bats (n=4), striped skunks (n=5), red foxes (n=9), and mule deer (n=6). Distinct dynamics were observed in within-host virus populations recovered from WA01- and Delta- infected animals. This included increased within-host variation, relative effective population size, and genomic signatures of positive selection in WA01 animals. In contrast to our first study in domestic dogs, Brazilian free-tailed bats, which also did not shed infectious virus, did not show increased signs of adaptation. We also observed a potential host barrier to infection in skunks and one fox, followed by the emergence of potential de novo mutations. Six novel mutations were also detected in contact-exposed mule deer. Our findings suggest that mule deer populations, similar to what has been documented in closely related white-tailed deer, should be investigated for accelerated SARS-CoV-2 evolution.

Collectively, our work reveals the unique dynamics of SARS-CoV-2 evolution and transmission in both naturally- and experimentally- infected felids. We observed rapid viral adaptation both in vitro and in vivo, highlighting advantages and limitations of experimental animal infections for studies of viral evolution. In each study, we used publicly available data to contextualize our experimental data and identify broader patterns. Furthermore, we identified specific SARS-CoV-2 mutations and genomic regions under selective pressures across a range of animal species, setting the groundwork for future mechanistic studies. Our findings underscore the importance of a One

Health approach to understanding SARS-CoV-2 evolution, and the need for surveillance in animal populations.

ACKNOWLEDGEMENTS

I would not be here without the support and contributions of countless friends and colleagues. I would like to thank my advisor, Dr. Sue VandeWoude, for her continual support and mentorship. It is a privilege to call myself a member of the Sue VandeWoude Research Group, and I am thankful for the direct and indirect contributions of SVRG members new and old, near and far. I am eternally indebted to Dr. Erick Gagne for jumpstarting my research journey in the middle of a pandemic. A special thank you to SVRG laboratory manager, Mary Nehring, for her friendship, encouragement and contributions to all aspects of SVRG. The support of the CSU Writes program was essential to writing this dissertation. I am grateful for all of the university and MIP faculty and staff who supported this process, and I want to specifically thank all of the people who worked with animals that contributed samples to this research. I'm also thankful for the camaraderie and solidarity of the members of my PhD cohort, and all of the fantastic graduate students I met along the way. Finally, thank you to my friends and family for accompanying me on this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction: SARS-CoV-2 evolution in the context of its broad host range

1.1 Significance: Why is it necessary to study SARS-CoV-2 evolution in non-human hosts?

The significance of virus evolution to One Health—the intersection of human and animal health and the environment—cannot be overstated. The ability to infect multiple host species and evolve rapidly in response to changing environments, are key characteristics of emerging zoonotic viruses. The emergence of COVID-19 confirmed many scientists' fears and predictions about the pandemic potential of zoonotic viruses. SARS-CoV-2 has adapted to human hosts following its initial spillover into human populations in late 2019, and continues to evolve; in this time, it has retained the ability to infect other animal species (Andersen, Rambaut, Lipkin, Holmes, & Garry, 2020; F. Wu et al., 2020; Holmes et al., 2021). Over the past four years, the evolution of SARS-CoV-2 has developed into a primary global public health concern.

Remarkably, despite focused research efforts, the evolutionary origins of the divergent variant lineages that have swept through human populations are still unknown. Three main hypotheses exist for the evolutionary origins of new variant lineages: (1) they slowly evolved from existing variants while remaining undetected over long periods of time in human populations, (2) the persistent infection of immunocompromised human individuals resulted in the accumulation of a large number of selectively beneficial mutations, and (3) spillover into non-human animals followed by selective pressures in a new host species led to the emergence of a highly divergent variant. In each of these scenarios, variants with enhanced replication and transmission kinetics would be more readily transmitted to new, susceptible hosts, thus perpetuating emergence. Hypothesis (1) is highly unlikely to have occurred for every variant, due to intense global genomic surveillance during the pandemic. However, evidence exists in support of both hypotheses (2) and (3), though this question may not be possible to answer definitively. The aim of this piece is to introduce and review the study of SARS-CoV-2 evolution in the context of its broad host range, and to highlight the fundamental importance of transmission in animals to its origins and continued evolution.

1.2 Methods: What is required to study SARS-CoV-2 evolution in animals?

Many factors need to converge to successfully study virus evolution in non-human animals. Multifaceted collaborations spanning the local, national, and international levels are often required (WOAH, 2024). Successful collaborations engage research scientists, veterinarians and animal care specialists employed by zoos, universities, animal shelters, farms, wildlife management agencies, and a myriad of other professionals (Munnink et al., 2021; Sparrer et al., 2023; McAloose et al., 2020). Once these collaborations are established, researchers employ a range of techniques and technologies to generate the sequencing data that is required for evolutionary analyses. Whole genome sequencing (WGS) of the SARS-CoV-2 virus using next-generation sequencing (NGS) technologies is common, as compared to the targeted sequencing of particular genomic regions. This data can be used for phylogenomic analysis, and has supported the description of SARS-CoV-2 lineages that are characterized by large numbers of mutations and signatures of selection spanning the virus genome.

Although there is variation in the methods employed to obtain SARS-CoV-2 sequence data from a biological sample, they also share many basic steps. In the most common procedure, RNA is isolated and reverse transcribed to cDNA, followed by a multiplex PCR that generates amplicons spanning the full virus genome (Quick et al., 2017). The ARTIC Network has been a leading source of SARS-CoV-2 primer design over the course of the pandemic, although comparable primer schemes have also been developed by private companies (ARTICNetwork, 2020). Non-targeted, or metagenomics sequencing approaches are also in use, but the tiled amplicon protocol is more widely in use. SARS-CoV-2 sequencing libraries have been sequenced on every variety of instrument currently in use, including most commonly Illumina and Oxford Nanopore technologies.

A defining aspect of NGS-based research is quality control, and this has also played a large role in the study of SARS-CoV-2 evolution. Notable limitations include errors introduced by PCR amplification, dropout regions due to mutation or other technical failures, low sequencing depth and incomplete genome coverage (Davis et al., 2021; DeMaio et al., 2020). Ultimately, it

is important to consider the scientific goal of the study in question. If the goal is to investigate evolution at the host population level, the main unit of analysis will be the consensus sequences of the virus genome. However, if the goal is to capture variation in the virus populations within one or more hosts, deeper sequencing will be required. After sequencing reads are generated, next steps include sequence alignment and/or assembly, and variant calling and/or identification of consensus-level mutations (SNPs). In the case of the targeted sequencing data commonly generated for SARS-CoV-2, sequencing reads are aligned to one or more reference sequence to identify changes relative to an ancestral SARS-CoV-2 genome.

Open sharing of genomic sequencing data has played a key role in the study of SARS-CoV-2 evolution in animals and humans to date. In addition to well-established nationally supported genetic databases, the most commonly used platform for sharing SARS-CoV-2 sequence data has been the "Global initiative on sharing all influenza data," or GISAID (Shu & McCauley, 2017). Genomic surveillance in humans has included both the attempt to obtain general, "representative" samples of infected human populations, and also targeted investigations of specific populations or settings. As a result, disparities in the distribution of human SARS-CoV-2 sequencing data have been widely observed (Brito et al., 2022; Chen et al., 2022). Although the breadth and resources directed towards sequencing SARS-CoV-2 from non-human animals has been comparatively limited, public databases like GISAID still provide an indication of the frequency and regularity of non-human animal infections. Furthermore, they facilitate one of the first steps in evaluating SARS-CoV-2 genomes recovered from animals: comparing them to genomic surveillance data in humans from the same time and geographic location.

Evolutionary analyses of SARS-CoV-2 often focus on epidemiology, phylogenetics, and mutation tracking. Importantly, scientists identify and designate emerging SARS-CoV-2 variant lineages using phylogenetic tools (Hadfield et al., 2018; Rambaut et al., 2020; Aksamentov, Roemer, Hodcroft, & Neher, 2021). Other important, but less well-trodden avenues for understanding SARS-CoV-2 evolution include selection, within-host variation, and recombination analyses. Most tools used for evolutionary analyses of SARS-CoV-2 are open source, and many scientists share

their protocols and code in full, enabling reproducibility. Bioinformatic workflows are often organized into publicly available analysis pipelines. For example, the nf-core/viralrecon pipeline, developed with the open-source workflow manager Nextflow, can be used to investigate low frequency variants in SARS-CoV-2 samples (Di Tommaso et al., 2017; Ewels et al., 2020; Patel et al., 2023). Analysis pipelines developed with this approach are highly documented, optimized and validated by the coding community, supporting reproducibility. Furthermore, they are flexible, an extremely valuable characteristic in the world of emerging zoonotic pathogens.

1.3 In practice: How have scientists approached SARS-CoV-2 in animals from an evolutionary perspective?

The scientific literature describing SARS-CoV-2 evolution in animals can be roughly divided into two categories: experimental exposures in laboratory settings, and natural infections in the wider world. Generally speaking, most studies of SARS-CoV-2 evolution in natural infections describe individual cases, and are limited by sampling constraints. On the other hand, most laboratory-based animal studies are not limited by sampling, but can be highly translational, focusing on viral pathogenesis and interventions for human health. A subset of these studies has been dedicated to evaluating the susceptibility of different host species (A. M. Bosco-Lauth et al., 2020; Porter, Hartwig, Bielefeldt-Ohmann, Bosco-Lauth, & Root, 2022; Porter et al., 2024; A. M. Bosco-Lauth et al., 2022, 2021). An even smaller number of studies describe virus evolution in experimentally infected animals, including the research reported in Chapters 2, 3 and 5 of this dissertation (Bashor et al., 2021; Bashor, Gagne, Bosco-Lauth, Stenglein, & VandeWoude, 2022).

The most active initial research into the evolution of SARS-CoV-2 in non-human animals has centered around its evolutionary origins prior to introduction into humans in late 2019. In the early 2000s, the emergence of the first SARS in China, caused a global outbreak. Later research indicated that SARS most likely originated in horseshoe bats (*Rhinolophus*) in China. The spillover into humans occurred following transmission from bats through an intermediate host, likely the civet cat, which was facilitated by trade in live animals. More recently, SARS-CoV-2 emerged

in humans, leading to a global pandemic. Although it may not be possible to prove definitively, existing scientific evidence is consistent with an origin in *Rhinolophus* bats, followed by spillover into humans from wild animals, likely facilitated by live wildlife trade occurring at the Huanan Seafood Market in Wuhan, China (Worobey et al., 2022; Liu et al., 2023). Interestingly, analysis of the earliest cases of SARS-CoV-2 has revealed two distinct SARS-CoV-2 lineages, indicating that multiple spillover events may have transpired (Pekar et al., 2022).

Now that the breadth of the host range of SARS-CoV-2 is better understood, it may not come as a surprise to note that several of the earliest virus genome sequences form non-human animals were recovered from cats and dogs. Two SARS-CoV-2 sequences from infected dogs in Hong Kong were collected in February and March 2020, and the first sequence from an infected cat in Belgium was collected in March 2020 (Shu & McCauley, 2017; Garigliany et al., 2020; Sit et al., 2020). Since then, cases in domestic animals including cats, dogs, ferrets and hamsters have been regularly reported. In these cases, NGS is generally employed to establish high identity between the virus sequences obtained from the closest human contact and the animal, confirming a cross-species transmission event has occurred. Sequences from these cases demonstrate high polyphyly, and transmission events span the breadth of SARS-CoV-2 variant lineages observed in human populations. Furthermore, onward transmission or "spillback" into humans from these closely cohabitating species has also been reported in cats and hamsters (Yen et al., 2022; Sila et al., 2022).

Despite the apparent regularity of human-to-pet cross-species transmission events, more research is needed to confirm species-specific viral adaptation in domestic species. This may be due in part to the rarity of sustained transmission within a species. Cat-to-cat, hamster-to-hamster, and ferret-to-ferret transmission has been established in both natural and experimental settings, whereas the viability of dog-to-dog transmission is uncertain. The limited number of studies in dogs to date do suggest that they are less susceptible to infection than cats and hamsters. Although it may not be completely accurate to call most household cases dead-end infections, research to date shows they are not a source of independent SARS-CoV-2 evolutionary trajectories.

The maintenance and evolution of SARS-CoV-2 in farmed mink and free-ranging white-tailed deer stands in stark contrast to household-based outbreaks reported for domestic animals. Some of the earliest discussions of SARS-CoV-2 evolution in non-human animals outside of pre-spillover hosts were instigated by the widespread infection of mink. These outbreaks were first observed in spring 2020 in the Netherlands and Denmark, but were not restricted to Europe, and also occurred on especially large scales in the United States. Numerous densely cohabitating individuals allowed for sustained mink-to-mink transmission, in addition to cross species transmissions to and from humans in close contact (Munnink et al., 2021; Boklund et al., 2021; Oreshkova et al., 2020; Hammer et al., 2021; Eckstrand et al., 2021; Rabalski et al., 2022; Lu et al., 2021; Larsen et al., 2021). Researchers have used sequence data to provide evidence for these spillover and spillback events, even identifying additional spillovers to domestic cats, dogs and free-ranging wild mink (van Aart et al., 2021; Aguiló-Gisbert et al., 2021). Evolution-driven investigation of SARS-CoV-2 sequences from these outbreaks have identified a number of candidate mutations for mink-specific adaptation (Lassaunière et al., 2021; Hoffmann et al., 2021; Tan et al., 2022). Furthermore, sequencing at one farm in Denmark indicated the evolution of mink-specific SARS-CoV-2 immune escape mutations enabling widespread reinfection of seropositive mink (Rasmussen et al., 2021). The rapidity of SARS-CoV-2 evolution in mink and their potential to act as viral reservoirs led to the culling of millions of farmed mink in 2020 and 2021.

At this time, more than half of US states have reported SARS-CoV-2 in white-tailed deer (APHIS, 2024a). The population size of the species as a whole is estimated at 30 million, with a range extending across the United States and southern Canada (Hanberry, 2021). Serological and viral RNA investigations have uncovered high (up to 40%) infection prevalence in free-ranging deer populations, with circulation beginning as early as September 2020 (Kuchipudi et al., 2022; Marques et al., 2022; Hale et al., 2021; Chandler et al., 2021; Palermo, Orbegozo, Watts, & Morrill, 2021). Sequencing data from these populations indicate that multiple introductions of every major SARS-CoV-2 variant lineage including Alpha, Gamma, Delta, and Omicron, and extensive deer-to-deer transmission have occurred (Kuchipudi et al., 2022; Marques et al., 2022; Hale et al., 2021;

McBride et al., 2023; Vandegrift et al., 2022; Caserta et al., 2023). Virus genomes recovered from infected deer are characterized by numerous mutations, most likely accumulated over time through sustained intraspecies transmission. Remarkably, the evolutionary rate of SARS-CoV-2 in deer in Ohio sampled in 2021 and 2022 was found to be three times higher than the evolutionary rate of the virus in humans (McBride et al., 2023).

Phylogenomic analyses have provided evidence for multiple deer-to-human transmissions of SARS-CoV-2 occurring as early as November 2021 (Pickering et al., 2022; Feng et al., 2023). Moreover, research has also demonstrated the ongoing circulation of Alpha, Gamma, and Delta variants in New York deer long past the period in which these lineages were last observed in human populations (Caserta et al., 2023). Taken together, these findings raise the concern that deer may act as viral reservoirs of divergent variant lineages. A worst-case scenario could involve the spillback of an older, mutated lineage, divergent enough to evade current human immunity, which is based on vaccination against and previous infection with more contemporary lineages.

Predictably, at the same time as free-ranging wildlife have been infected with SARS-CoV-2 through exposure to humans, captive wild animals, both white-tailed deer and other species, have also been exposed at high rates (Roundy et al., 2022). Large felids in zoo settings have been particularly frequently infected throughout the pandemic (Karikalan et al., 2021; Koeppel et al., 2022; Grome et al., 2022; Mishra et al., 2021; Bartlett et al., 2021; Fernández-Bellon et al., 2021; Mitchell et al., 2021; Nagy et al., 2022). The investigation of an outbreak at the Bronx Zoo in April 2020 generated some of the first publicly available SARS-CoV-2 sequences from naturally infected non-human animals (McAloose et al., 2020). At that time, a natural spillover event from a human to another species had not previously been captured. Later, sequence data collected at an earlier date in 2020 from domestic cats and dogs would be made available. However, the sequences from tigers, lions, and their caretakers at the Bronx Zoo initiated an important discussion surrounding the contributions of cross-species transmission to SARS-CoV-2 evolution.

In the case of SARS-CoV-2 infections of domestic and wild felids around the world, NGS has been employed mainly with the goal of identifying the causative virus variant, and characterizing

the infection source, which often points towards close human contacts. Notably, viral spillback from an African lion to humans has also been reported (Siegrist et al., 2023). There are now 139 virus sequences from members of the genus *Panthera* on GISAID (Shu & McCauley, 2017). Including domestic cats and all members of the family Felidae, this number rises to 335. Interestingly, felids have appeared to be more frequently infected than primates in zoo settings, and domestic cats have been more commonly infected than any other domestic species (Nerpel et al., 2022). However, there is no consensus on what factors, if any, might make felids more susceptible to infection than other species, particularly primates, which are much more closely related to humans, and thus have higher ACE2 receptor identity and predicted binding to the SARS-CoV-2 spike protein (Damas et al., 2020). In contrast to outbreaks in mink and white-tailed deer, sustained within-species transmission or maintenance of SARS-CoV-2 in a population of domestic or wild felids has not been observed. This may highlight the potential for felid-specific evolution in feral cat colonies, shelters or larger populations of big cats in zoos or other settings. Chapter 4 of this dissertation describes persistent infection and evidence for SARS-CoV-2 host-adaptation in zoo tigers, lions, and hyenas.

In experimental settings, SARS-CoV-2 evolution in animals has been described indirectly through the use of non-human cell culture systems, and more directly with animal models. Viral adaptation following passage in Vero cells, originally derived from African green monkeys, has been repeatedly documented (Sasaki et al., 2021; Bashor et al., 2021; Chaudhry et al., 2022; Ogando et al., 2020). More specifically, mutations in the furin cleavage region of the SARS-CoV-2 spike gene have been shown to facilitate viral entry into Vero cells despite type II transmembrane serine protease (TMPRSS2)-deficiency, although viruses may also use a TMPRSS2-independent, cathepsin-mediated entry method. Early in the pandemic, mice were found to be resistant to SARS-CoV-2 infection due to the divergence of mouse angiotensin converting enzyme (ACE2). Although many researchers turned to transgenic mouse models expressing humanized ACE2, other studies explored the use of experimental virus evolution to obtain a mouse-adapted SARS-CoV-2 virus (Gu et al., 2020; Huang et al., 2021). SARS-CoV-2 passage in BALB/c mice led to the emergence of

mutations enhancing ACE2 binding, including a critical N501Y mutation in the receptor-binding domain (RBD) of the spike gene.

In addition to mice, domestic animal models have further demonstrated the rapidity of SARS-CoV-2 adaptation following host shifts. SARS-CoV-2 transmitted through direct contact and aerosols between ferrets contained two candidate adaptive mutations in the spike gene, identified through the comparison of sequences from ferrets with the sequence of the virus stock used for inoculation (Richard et al., 2020; Kutter et al., 2021). Of these two, S686G was in the original virus stocks at very low levels, but N501T was not. Experimental SARS-CoV-2 exposures of domestic cats have also highlighted virus host-adaptation. Remarkably, these studies have also reported the emergence of the spike S686G mutation in cats (Bashor et al., 2021, 2022; Braun et al., 2021). In one study, within-host viral variation was transmitted from infected to naïve cats across a transmission bottleneck of just 2-5 viral genomes (Braun et al., 2021). Two other published studies in domestic cats have been reproduced in full as Chapters 2 and 3 of this dissertation.

Laboratory-based studies of SARS-CoV-2 evolution in animals may be specifically advantageous for controlling the uncontrollable factors at play in natural infections, enabling researchers to ask directed evolutionary questions. A fundamental constraint of emergent zoonotic pathogen research is our inability to observe the initial transmission dynamics and host shifts surrounding spillover events. Sequencing of virus samples collected from experimentally exposed animals in addition to the initial inoculum virus used provide snapshots of virus populations throughout every stage of infection. Such experimental studies offer an unparalleled opportunity to carry out rigorous investigations of virus evolution following host shifts.

1.4 Conclusions: What have we learned, what's next and how can we improve?

In humans, SARS-CoV-2 evolution to date has gone through multiple stages (Markov et al., 2023). In early 2020, it was characterized by gradual small changes. For example, the spike D614G mutation, which significantly increases transmissibility in humans, swept through global populations. As the pandemic continued, it shifted towards a more stepwise evolution of variant lineages characterized by sets of new mutations and increased overall substitution rates. As of

May 2024, it continues to evolve from the genetic background of Omicron, a highly transmissible variant lineage. Currently, a key evolutionary driver is human immune pressures resulting from vaccination and prior infection. The trajectory of SARS-CoV-2 evolution in humans can inform our understanding of its evolution in animals, and vice versa. Importantly, the origins of each variant lineage that has emerged in humans are still unknown. Furthermore, although many studies have identified candidate host-species adaptive mutations, further research is needed to elucidate the type and extent of advantage offered by these mutations. Continued study of virus evolution and transmission between humans and animals will provide the essential data required to answer these questions.

The scientific literature describing SARS-CoV-2 evolution in animals can be organized into several different configurations. It embodies a range of animal taxa, environments, and outbreak sizes, and employs a range of methodological approaches, from basic epidemiology to phylogenomics to within-host variation. One of the greatest benefits of this body of research is public data sharing. This has played out as a microcosm of the transformative leap in sharing sequencing data allowing scientists to document SARS-CoV-2 evolution in humans in real time. Future studies should leverage these data to provide context to their work and seek broader evolutionary patterns. For example, the quantification of species-specific virus evolutionary rates requires large sample sizes that are rarely achievable in individual studies.

Altogether, a relatively large number of published studies describe SARS-CoV-2 evolution in animals. Future work may thus benefit from addressing various limitations of earlier studies, with a focus on a priori hypothesis development and testing. Increased use of replication, controls, longitudinal sampling, collection and banking of high-quality samples and metadata, and consistency and transparency of methods will aid in answering a number of outstanding questions in this field. Furthermore, interdisciplinary collaborations will help researchers leverage the wealth of virus genome sequencing data to answer more and more complex questions.

One of the key themes spanning this body of research is that spillover is not unidirectional. The original transmission of SARS-CoV-2 into humans was likely not a single event, and we have since

observed repeated spillovers into a broad range of species described in this piece. Furthermore, spillback from animals into humans has been widely reported, including mink, lions, hamsters, cats and deer. It is important to note that for every species in which transmission is sustained within a population, we see signs of the strong evolutionary pressures host shifts exert on virus evolution. However, SARS-CoV-2 does not transform into a new variant lineage every time it undergoes a host shift. A more concerning possibility may be the maintenance of divergent variant lineages in animal reservoirs, with the potential for spillback into humans in a changing immunity landscape.

By applying a One Health perspective, SARS-CoV-2 evolution in non-human animals might be considered a main feature of, rather than a footnote to this pandemic. Furthermore, the approaches and findings resulting from the studies described in this piece and in the chapters of this dissertation can be directly applied to advancing our understanding of other emerging zoonotic viruses in the future. The following chapters describe findings from experimental and natural infections of non-human animals, offering snapshots of SARS-CoV-2 evolution in action.

Chapter 2

SARS-CoV-2 evolution in animals suggests mechanisms for rapid variant selection[1]

SARS-CoV-2 spillback from humans into domestic and wild animals has been well documented, and an accumulating number of studies illustrate human-to-animal natural transmission is widespread in cats, mink, deer, and other species. Experimental inoculations of cats, mink and ferrets have perpetuated transmission cycles. We sequenced full genomes of Vero-cell expanded SARS-CoV-2 inoculum and virus recovered from cats (n=6), dogs (n=3), hamsters (n=3) and ferret (n=1) following experimental exposure. Five nonsynonymous changes relative to the USA-WA1/2020 prototype strain in nsp12, S, N and M genes were near fixation in the stock used for inoculation, but had reverted to wild-type sequences at these sites in dogs, cats and hamsters within 1-3 days post-exposure. Fourteen emergent variants (six in nonstructural genes, six in spike, one in orf8) were detected in viruses recovered from animals. This included substitutions in spike residues H69, N501, and D614, which also vary in human lineages of concern. Even though live virus was not cultured from dogs, substitutions in replicase genes were readily detected in amplified sequences. The rapid selection of SARS-CoV-2 variants in vitro and in vivo reveals residues with functional significance during host-switching. This also illustrates the potential for spillback from animal hosts to accelerate evolution of new viral lineages, findings of particular concern for dogs and cats living in households with COVID-19 patients. More generally, this glimpse into in vitro and in vivo host switching reveals the unrealized rapidity and plasticity of viral evolution in experimental animal model systems.

[1]This chapter is a reproduction of "Bashor, L., Gagne, R. B., Bosco-Lauth, A., Bowen, R.A., Stenglein, M., & Vande-Woude, S. (2021). SARS-CoV-2 evolution in animals suggests mechanisms for rapid variant selection. *Proceedings of the National Academy of Sciences*, 118(44). doi: 10.1073/PNAS.2105253118". Minimal modifications were made to meet dissertation formatting requirements.

2.1 Significance Statement

SARS-CoV-2 emerged due to viral spillover from animals to humans, and spillback to other animal species has been observed with accelerating frequency. Cross-species transmission generally results in rapid adaptation of the virus to the new host, and repeated transmissions may accelerate viral evolution and novel strain emergence. We report surprisingly rapid selection of numerous SARS-CoV-2 variants in cell culture and following infection of nonhuman mammalian hosts, including dogs and cats. These molecular changes in SARS-CoV-2 provide insight into mechanisms of viral host adaptation, lay the groundwork for additional studies assessing dominant variant fitness and phenotype, and highlight the potential for human reinfection with new viral variants arising in species in close and frequent contact with humans.

2.2 Introduction

Cross-species transmission events, which challenge pathogens to survive in new host environments, typically result in species-specific adaptations (Sauter & Kirchhoff, 2019). These evolutionary changes can determine the pathogenicity and transmissibility of the virus in novel host species (Andersen et al., 2020). Pathogen host-switching resulting in epidemic disease is a rare event that is constrained by interaction between species (Cross, Prosser, Ramey, Hanks, & Pepin, 2019). In contrast to most species, humans move globally and regularly come into contact with domestic and peridomestic animals. Thus, when a novel virus spreads through human populations, there is an incidental risk of exposure to potentially susceptible non-human species.

This scenario has become evident with the SARS-CoV-2 pandemic (Table A.1). Originally resulting from viral spillover into humans (Zhou et al., 2020; Lam et al., 2020), likely from an animal reservoir, spillback into a wide range of companion and wild animals has occurred or been shown to be plausible (A. M. Bosco-Lauth et al., 2020; McAloose et al., 2020; Fernández-Bellon et al., 2021; Oude Munnink et al., 2021; Shereen, Khan, Kazmi, Bashir, & Siddique, 2020), and an increasing number of studies have indicated a high frequency of human-to-animal SARS-CoV-2 spillback transmission (Segalés et al., 2020; Klaus et al., 2021; Neira et al., 2021; Hosie et al., 2021; Garigliany et al., 2020; Newman et al., 2020; Fuentealba et al., 2021; Ruiz-Arrondo et al.,

2021). Given the short duration of viral shedding, serologic analyses present a more accurate characterization of actual animal exposures to SARS-CoV-2. Such studies conducted in a variety of animal species have illustrated surprisingly high levels of seroconversion in cats and dogs, and more recently free-ranging deer (Table A.1). Other well documented spillback events include numerous mink farms (Table A.1). In one of these reports, multiple feral cats living on a mink farm in the Netherlands during a SARS2 outbreak were seropositive, likely from direct transmission of the virus from mink to cats, as owned cats on the same farm were seronegative (van Aart et al., 2022). This further illustrates that cross-species transmission chains are readily achieved. Recent surveys of free-ranging white-tailed deer in Illinois, Michigan, New York, and Pennsylvania revealed 33% seropositivity in free-ranging animals (United States Department of Agriculture Animal and Plant Health Inspection Service (USDA APHIS), 2021a). Active SARS-CoV-2 infection was subsequently confirmed by PCR in a deer in Ohio (United States Department of Agriculture Animal and Plant Health Inspection Service (USDA APHIS), 2021b). Together, these findings suggest the likely establishment of multiple domestic animal and wildlife reservoirs of SARS-CoV-2.

Repeated interspecies transmission of a virus presents the potential for acceleration of viral evolution and a possible source of novel strain emergence. This was demonstrated by reverse zoonosis of SARS-CoV-2 from humans to mink, followed by selection in mink and zoonotic transmission back to humans (Oude Munnink et al., 2021). Given that reverse zoonosis has been reported repeatedly in dogs and cats from households where COVID-19 patients reside, and the fact that up to 50% of households worldwide are inhabited by these companion animals, there is potential for similar transmission chains to arise via humans and their pets (Banerjee, Mossman, & Baker, 2021; Hedman, Krawczyk, Helmy, Zhang, & Varga, 2021). Elucidating viral selection and species-specific adaptation of SARS-CoV-2 in common companion animals is therefore of high interest. Furthermore, understanding viral evolutionary patterns in both companion animals and experimental animal models provides a valuable appraisal of species-specific viral variants that spotlight genomic regions for host-virus interaction.

Significant attention has been directed at substrains evolving from the initial SARS-CoV-2 isolate (F. Wu et al., 2020) and an accumulating number of variant lineages have demonstrated increased transmission potential in humans (Andrew, 2020; Korber et al., 2020). The role, if any, that reverse zoonotic infections of nonhuman species and spillback may have played in the emergence of these novel variants of SARS-CoV-2 remains unknown. Documenting viral evolution following spillover of SARS-CoV-2 into new species is difficult given the unpredictability of timing of these events; therefore, experimental studies can greatly aid understanding of SARS-CoV-2 evolution in animal species. Laboratory-based studies also provide the opportunity to determine how changes that occur during viral expansion in cell culture may influence in vivo infections. This information is highly relevant for interpretation of in vivo and in vitro experiments using inoculum propagated in culture.

We therefore assessed the evolution of SARS-CoV-2 during three rounds of expansion of strain USA-WA1/2020 in Vero E6 cells (Harcourt et al., 2020), followed by measuring variant emergence occurring during primary experimental infection in four mammalian hosts. Specifically, we compared variant proportions, insertions, and deletions occurring in genomes of SARS-CoV-2 recovered from dogs (n=3), cats (n=6), hamsters (n=3) and ferret (n=1).

2.3 Results

2.3.1 SARS-CoV-2 genome sequence recovery

We recovered full genome sequences of SARS-CoV-2 from stocks of three serial passages of USA-WA1/2020 inoculum and virus recovered from five cats with primary exposure, one cat exposed by contact to an infected cat (A. M. Bosco-Lauth et al., 2020), three dogs (A. M. Bosco-Lauth et al., 2020), three hamsters, and one ferret (Table A.3). PCR amplification of the full SARS-CoV-2 genome failed in two additional ferrets due to low viral titers; these samples were subjected to targeted PCR of spike sequences followed by Sanger sequencing. We obtained a median depth of coverage of 2859x (mean = 4380x) using ARTIC version 3 primers as described in methods (Fig. S1). SARS-CoV-2 genomes recovered from Cat 5 and one technical replicate of

Cat 4 were amplified with ARTIC version 2 primers with median depth of coverages of 506x and 1056x respectively.

2.3.2 Identification of single nucleotide and structural variants

Eight-eight unique single nucleotide (SNV) and structural variant (SV; insertions and deletions) were present in 3-100% of viral genome sequences in two technical replicates, and were observed 270 times across all datasets (242 SNV, 28 SV; Figure 2.1A). This included 74 SNVs and 14 SVs. Seventy-one of 270 variants (26.3%) were detected at 25% frequency or greater (Figure 2.1B). Seventy of the 88 unique variants (79.5%) were not detected in any of the three viral inoculum stocks at greater than or equal to 3% frequency. Seven SARS-CoV-2 SNV recovered from one or more species occurred at positions that coincide with variants of concern in humans or other species Table 2.1.

Table 2.1: Seven SARS-CoV-2 SNV recovered from one or more species have been reported as variants of concern in humans or other species. The frequency of each variant detected in Vero E6 cell supernatant at each passage (P1, P2, and P3) represents the average of two replicates; values undetected above 0.1% are reported as 0. Only D215H was detected at high frequency in P3 inoculum.
*S686G was detected in one replicate of P2 at 0.8% and undetected in the second; L37F was detected in one replicate of P3 at 0.5% and undetected in the second.

Position	CDS	Variant	Species	P1/P2/P3 inoculum frequency (%)	Reported variant(s)	Reported species
21768	Spike, NTD	H69R	Ferret, cat, dog, hamster	2.1/1.2/0	69–70del	Human (Andrew, 2020), mink (van Dorp et al., 2020)
22205	Spike, NTD	D215H	Cat, dog, ferret, hamster	4.2/19.4/96.6	D215G	Humans (Tegally et al., 2021), deer mice (Fagre et al., 2021)
22205		D215N	Cat, dog, hamster	1.4/1.5/0		
23064	Spike, RBD	N501T	Ferret	0/0/0	N501T	Ferret (Richard et al., 2020), mink (van Dorp et al., 2020)
					N501Y	Human (Andrew, 2020; Tegally et al., 2021), BALB mice
23403	Spike	D614G	Cat	0/0/0	D614G	Human (Shu & McCauley, 2017)
23525	Spike	H655Y	Cat, dog, ferret, hamster	1.5/4.8/1.1	H655Y	Cat (Braun et al., 2021), human (Shu & McCauley, 2017), hamster (J. F.-W. Chan et al., 2020)
23618	Spike	S686G	Cat	0/0*/0	S686G	Ferret (Richard et al., 2020)
11083	orf1ab, nsp6	L37F	Cat, dog, hamster	1.8/1.8/0*	L37F	Human (Shu & McCauley, 2017)

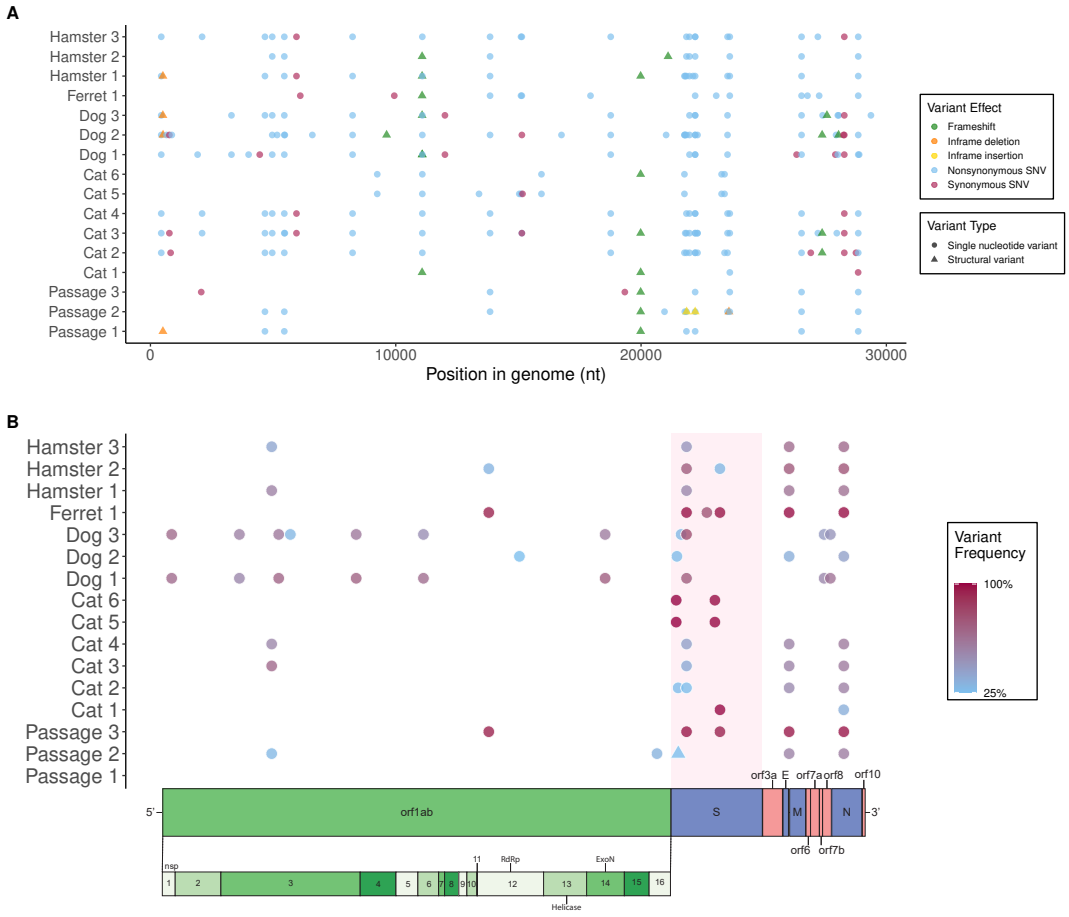


Figure 2.1: Eight-eight unique variants were detected in ≥3% of sequences of in vitro- and in vivo- derived SARS-CoV-2. (A) Position, predicted effect and (B) allele frequency of single nucleotide (SNV) and structural (SV) variants detected across the SARS-CoV-2 genome in sequences obtained from 13 experimentally inoculated animals and three passages of the viral inoculum. Each point represents a SNV (circle) or SV (triangle). (A) All variants detected in ≥3% of sequences, demonstrating a majority of SNVs and a slightly increased occurrence of modifications in the spike protein. (B) All variants detected in ≥25% of sequences, revealing the presence of higher frequency variants in the spike protein of all datasets excluding the Passage 1 stock virus, and the prevalence of higher frequency variants across the entire genome of SARS-CoV-2 recovered from dogs. Frequency indicated by scale from blue (low) to red (high). Schematic of SARS-CoV-2 genome illustrated for orientation.

2.3.3 Cell culture-associated viral variants and reversion in animals

Passage of SARS-CoV-2 in Vero E6 cells resulted in five nonsynonymous amino acid changes that reached fixation or near fixation following three passages (Figure 2.2). These variants reverted to wild type sequences within 1-3 days after inoculation in dogs, cats and hamsters. Reversion frequency was higher in samples recovered three days post-infection (dogs and cats) compared to virus recovered one day post-infection (hamsters, Figure 2.2). Spike variant D215H underwent a second substitution to D215N, which reached near fixation (>72%) in two dogs. This variant was also detected at >6% frequency in the third dog, one hamster and three cats.

2.3.4 Viral variants arising during in vivo passage

Thirteen nonsynonymous variants and one synonymous variant not present in the viral inoculum at $\geq$3% frequency were detected at >50% frequency in one individual, or were detected in all individuals of one species Table 2.2. The default limit of detection for variant calling was set at 3% but it was possible to observe variants with frequencies down to 0.1% with possibly reduced quantitative accuracy (Grubaugh et al., 2019). There were 564 unique variants detected at 0.1% frequency or greater across all 16 datasets in the study. Of these, 174 were detected in the three cell culture-derived samples, and 390 (69%) were not identified in the viral inoculum at all. Variants detected at lower frequencies within the Passage 3 (P3) inoculum in matched replicates would be suggestive of in vivo selection of these pre-existing quasispecies. Using this reasoning, we detected 10 of the 14 emergent variants from animal hosts at 0.1-3% in the viral inoculum, and four were not identified in the viral inoculum at all.

There was no significant difference between the mean number of unique variants detected in viral genomes recovered from the four species (Figure 2.3A). Viral titer did not seem to determine observed variant richness. We were not able to detect viable virus at 3 days post-infection in any of the dogs (<1 log pfu/mL); however, we observed 24, 34 and 22 variants in Dogs 1, 2 and 3 respectively, which is in the same range as Cats 2, 3 and 4, for which we assessed viral titers of 5.3, 6 and 5.4 log pfu/mL and 22, 30 and 19 variants respectively. In addition, Hamster 1 had 20

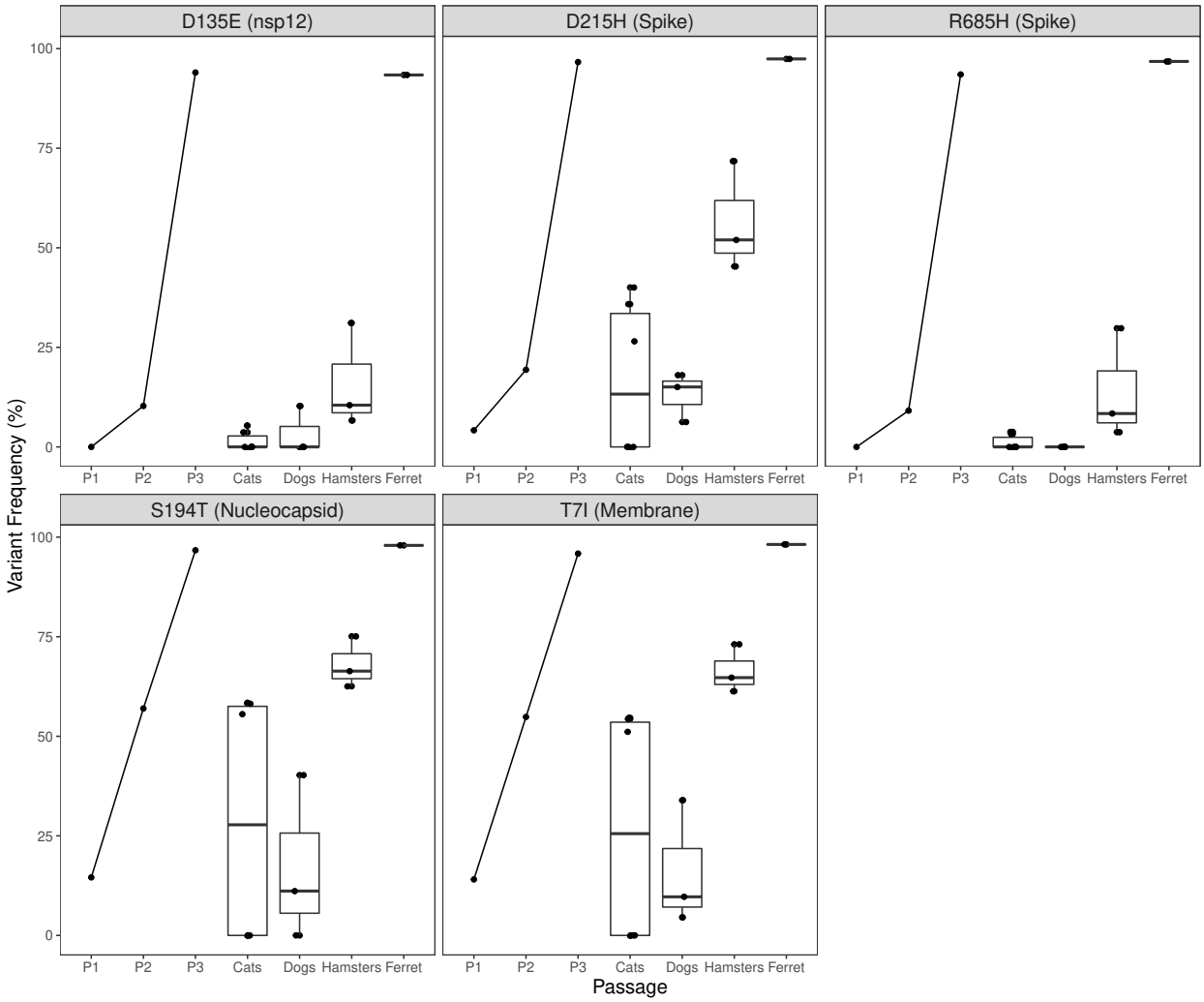


Figure 2.2: SARS-CoV-2 cell culture variants revert rapidly during in vivo experimental infection. SARS-CoV-2 isolate USA-WA1/2020 was passaged three times in Vero E6 cell line. Five single nucleotide variant substitutions across the genome reached >93% frequency; variant proportion recovered from each supernatant stock indicated by P1, P2, P3. Cats, dogs, hamsters and ferret (n=6, 3, 3, and 1, respectively) were inoculated with 10^5–10^6 pfu intranasally. Virus was recovered 1-3 days PI, sequenced using a tiled amplicon technique and analyzed with a pipeline for calling single nucleotide and structural variants in viral populations. Cell culture variants decreased in frequency in all animals with the exception of the ferret (n=1). Variants are indicated in reference to the consensus residue in USA-WA1/2020 and their position within the coding sequence of a SARS-CoV-2 protein. Each point represents the mean of two technical replicates, aside from one cat for which a replicate was not sequenced.

variants and a titer of 3.3 log pfu/mL, whereas Hamster 3 had 22 variants and a titer of <1 log pfu/mL.

SARS-CoV-2 sequences also did not cluster by species (Fig. S2). Excluding variants that were detected in the original inoculum sequences at $\geq$3% frequency, 69.6% of novel variants were found in a single species, 11.6% were found in two species, and 17.4% were found in three species. Only one variant that was not detectable in the inoculum at $\geq$3% frequency was detected in all four animal species, a single nucleotide insertion at L37 in nsp6 causing a frameshift.

Variants were unevenly distributed among individuals of a species. When variants detected in the viral inoculum were excluded from the analysis, a large proportion of variants detected in cats, dogs and hamsters were observed in only one individual of the species (Figure 2.3B). The SARS-CoV-2 spike gene contained a larger proportion of variants relative to its gene length (Figure 2.3C), and the majority of variant positions corresponding to human variants of concern were in the spike protein Table 2.1, Figure 2.3D).

Six nonsynonymous and one synonymous amino acid variant were found in all dogs and were present in >50% of sequences recovered from two of three dogs Table 2.2. These seven variants were found in cats and hamsters at lower frequency. A single novel amino acid variant (D138Y) was detected in all three hamsters. Cats 1-5 were experimentally inoculated; Cat 6 was infected through co-housing with Cat 5 one day post-infection (A. M. Bosco-Lauth et al., 2020). Cats were inoculated in two cohorts (Cohort 1 = 1, 2, 3; Cohort 2 = 4, 5, 6). Virus recovered from cats in Cohort 1 had significantly more variants than cats in Cohort 2, but variants were at higher frequency (near fixation) in cats in Cohort 2 (P<0.02). This observed variation is likely due to the difference in sample collection timeline between the two experimental cohorts. Two amino acid variants (H69R and D614G) were found at >98% in Cats 5 and 6, and S686G was found in 99% of sequences from Cat 1. Six of the ten variants detected in Cat 5, including spike variants H69R and D614G, were also detected in Cat 6, which was infected through contact with Cat 5 (Table B.3). The single unique variant detected in the contact cat caused a frameshift at D124 in nsp15. Depth

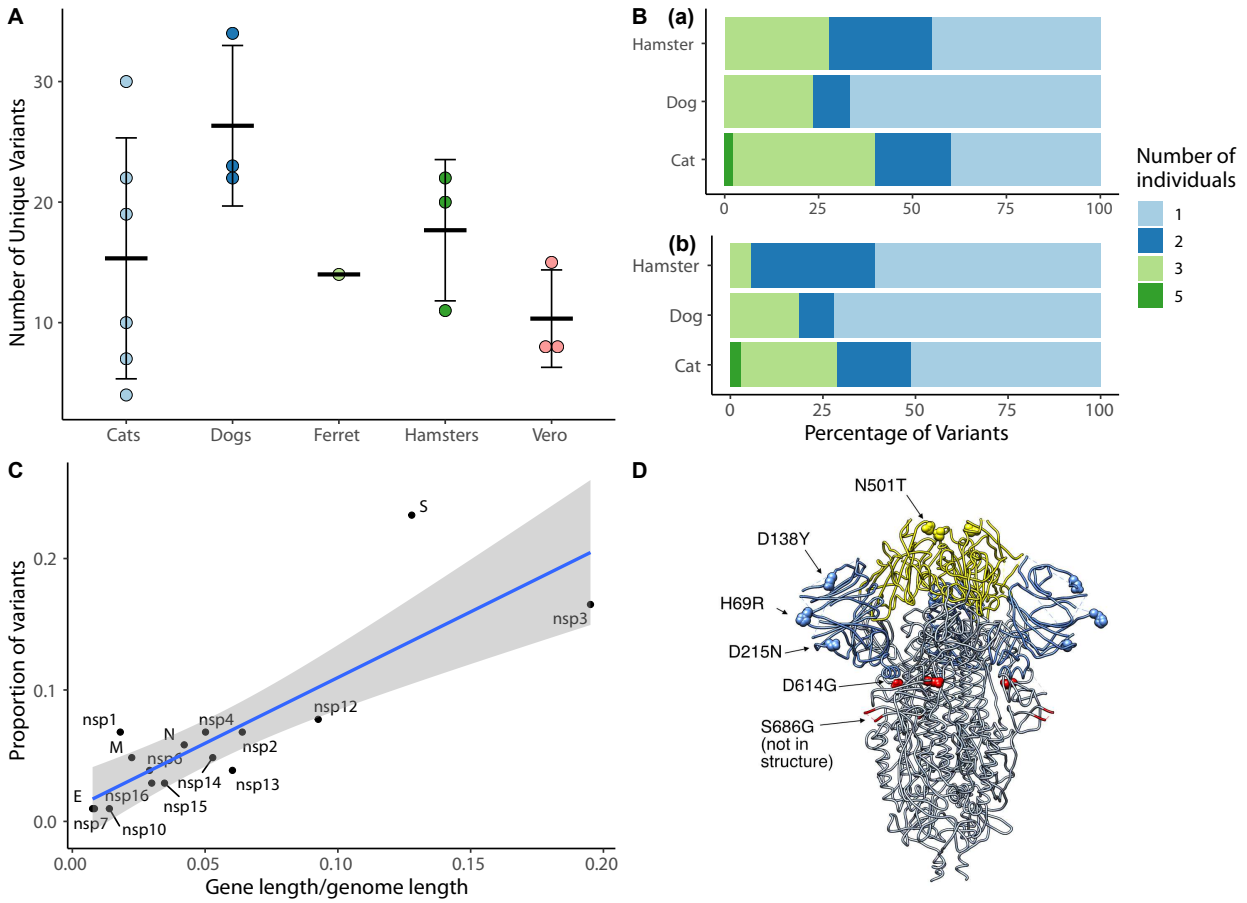


Figure 2.3: SARS-CoV-2 viral evolution differs across species, gene regions, and individuals. (A) Each point indicates the number of unique variants detected at ≥3% frequency in SARS-CoV-2 genomes recovered from individual animals. There is no significant difference in the number of unique variants detected in different species (ANOVA, p=0.22). (B) Analysis of variant distribution within species reveals that the majority of variants were detected in just one individual within each species. Subplot (a) shows distribution for all variants, while (b) illustrates only variants not occurring in the P3 inoculum at ≥3%. (C) Variants are distributed across viral genes in relation to each gene's length as a proportion of the entire genome length (linear model, R^2=0.69, p<0.0001). The spike protein contained a notably higher proportion of all variants in comparison to its share of genome length. Grey shading represents the 95% confidence interval for the slope of the regression line. (D) SARS-CoV-2 spike protein variant "residues of concern" are in NTD, RBD and furin cleavage site. Residues described in text and Table 2.1 in the SARS-CoV-2 spike trimer are highlighted on structure 6VXX. Blue indicates NTD, Yellow indicates RBD. The furin cleavage site and adjacent residue 686 are in the indicated loop, which was not resolved in this structure.

of coverage for Cat 5 was <40x at this position, so no variants were called. No amino acid variants from SARS-CoV-2 sequences were shared across all six cats.

We detected spike variant N501T in Ferret 1 at a frequency of 76%. Two additional ferret nasal wash samples from the same experimental cohort were subjected to targeted PCR, cloning and direct sequencing to ascertain the presence of this variant in additional ferrets. N501T (A>C at position 23064 in USA-WA1/2020) was present in four of six (66.7%) clones from Ferret 2 and two of three (66.7%) clones from Ferret 3.

2.3.5 Signatures of selection

Nonsynonymous (πN) and synonymous (πS) nucleotide diversity and πN/πS for SARS-CoV-2 populations was evaluated as a measure of within-host diversity and signatures of selection (Figure 2.4, Table A.5, Table A.5). At the population level, πN was greater than πS for 11 out of 13 SARS-CoV-2 samples (paired t-test, $p<0.01$), indicating positive selection across hosts (Figure 2.4A). At the species level, selection of SARS-CoV-2 in dogs and hamsters was most significant (Figure 2.4A). Further, we recorded a significant difference between πN and πS in orf1ab, spike and membrane proteins ($p<0.02$, Figure 2.4B), indicating positive selection at the level of these gene products. In particular, spike, πN was significantly greater than πS for viral genomes recovered from all 16 host and cell culture samples ($p<0.0001$).

2.4 Discussion

2.4.1 Cell culture-associated mutation

Variants that reached near fixation in cell culture but reverted to wild type in animal infections suggests a role for these residues in facilitating in vivo versus in vitro infection. Rapid reversion of variants arising in Vero E6 cells (originally derived from African green monkeys) occurred in dogs, cats, and hamsters, suggesting that selective advantages of cell culture variants are due to differences between in vivo and in vitro environments rather than specific host-virus adaptation. Alternatively, changes may reflect host-specific adaptations to the monkey and not in vitro infections. This explanation is less likely as the variants reverted in multiple species that are more

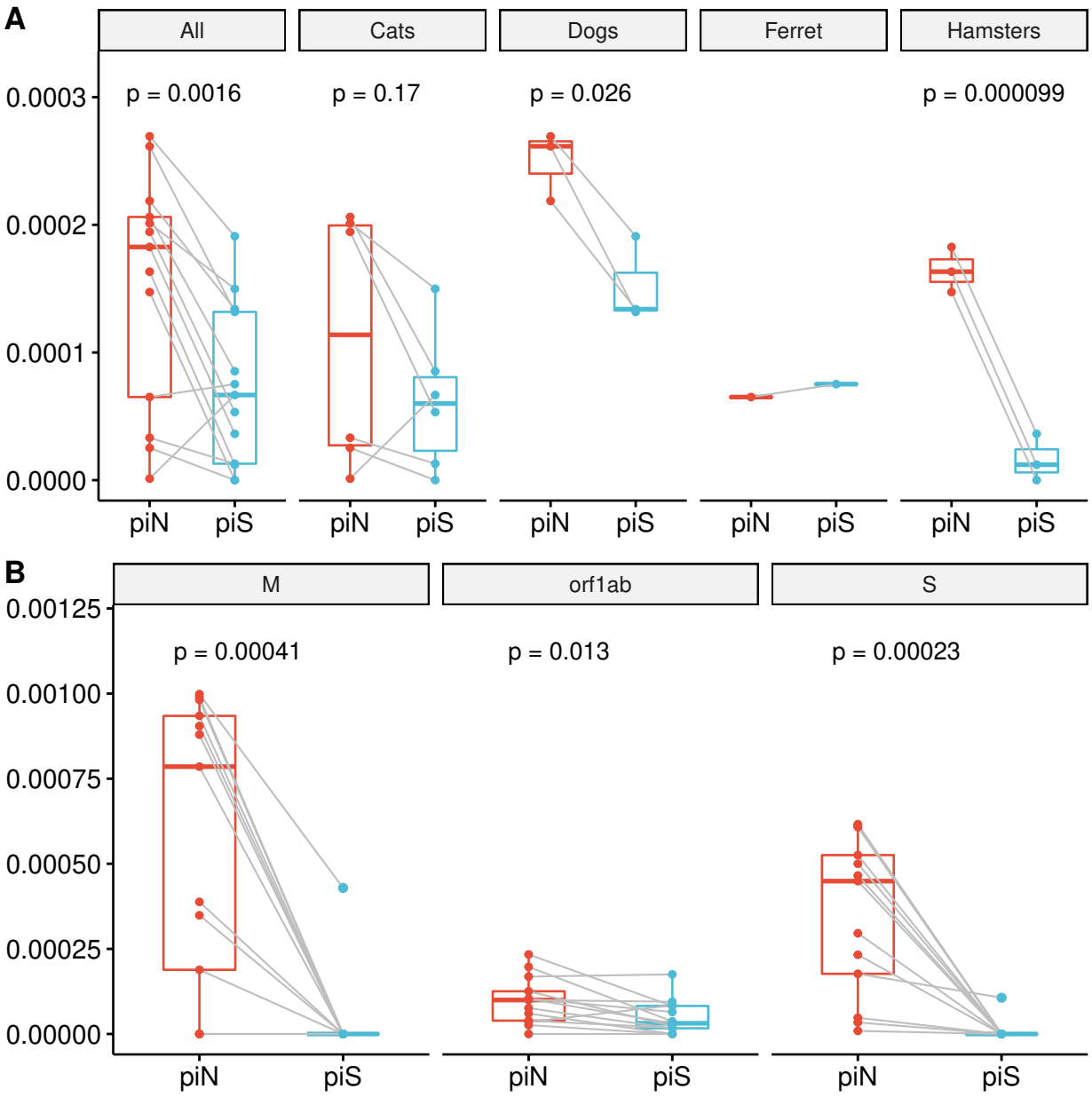


Figure 2.4: Signatures of positive selection are detected in SARS-CoV-2 genome sequences recovered from experimentally inoculated cats, dogs, hamsters and ferret. (A) Comparison of nonsynonymous (πN) and synonymous (πS) nucleotide diversity reveals that πN is significantly greater than πS, indicating positive selection. Each point represents a measurement πN (red) or πS (blue) calculated for the entire SARS-CoV-2 genome from sequences recovered from an individual animal, relative to the reference sequence of USA-WA1/2020. Analysis of the same measures within each species reveals $\pi N > \pi S$ for viral genomes is greater in dogs and hamsters. (B) Orf1ab, S, and M are undergoing positive selection in mammalian hosts. Each point represents πN or πS calculated for a specific SARS-CoV-2 gene or open reading frame from sequences recovered from an individual animal.

distantly related to humans than African green monkeys, and distinct differences exist between in vivo and in vitro environments. For example, in vivo viral attachment and entry in vivo requires circumvention of respiratory epithelial defenses such as mucus and cilia; residues which play roles in overcoming these physical features would not be highly selected in cell culture.

Amino acid R685 is located in the furin cleavage site, a motif unique to SARS-CoV-2 relative to SARS-CoV and other SARS-related coronaviruses (Coutard et al., 2020). The PRRAR insertion at this location provides a motif for serine proteases to cleave the spike protein into S1 and S2 subunits, a key step initiating viral entry via membrane fusion. Deletion of the furin cleavage site has been shown to reduce SARS-CoV-2 replication in human respiratory cells, hamster and hACE2 mice (Johnson et al., 2021). Previous work has demonstrated rapid emergence of spike variants at the furin-cleavage motif, including R685H, following five passages in type II transmembrane serine protease (TMPRSS2)-deficient cells (Sasaki et al., 2021). This explains the emergence of this variant in our study following propagation in Vero E6 cells, which do not express high levels of TMPRSS2 (Ou et al., 2021). The distribution of variants arising in vitro (nsp12, spike, N, M; Figure 2.2) indicates functions beyond receptor binding and entry as drivers of selection, which would be predicted to occur in spike.

2.4.2 Patterns in viral evolution within and across species

We observed the emergence of low, intermediate and high frequency variants across the SARS-CoV-2 genome in viral RNA recovered from cats, dogs, hamsters and ferrets. The abundance of variants detected in our study is in contrast to the few variants observed in large datasets of SARS-CoV-2 genomes recovered from humans (Lythgoe et al., 2021). Low frequency variants (<25%) may represent mutations occurring during population expansion following infection of a new individual. The majority of variants arising in vivo were in the spike protein, which was under the strongest selective pressure, illustrating the importance of this residue in shaping cross-species transmission events. Variants detected in animals were likely to arise either by rapid selection of low frequency inoculum variants with structural features that favored in vivo attachment, entry, and replication, or as de novo mutations arising from errors during viral replication.

The magnitude of cell culture variant reversion was greater in dogs than other species. All five cell culture-associated variants decreased from >93% to <41%; in particular, D135E and R685H decreased to $\leq$10%. SARS-CoV-2 was not culturable from canine samples, consistent with previous reports (Shi et al., 2020), and was below the limit of detection by qPCR, though low level seroconversion suggested some viral replication had occurred (A. M. Bosco-Lauth et al., 2020). These experimental results contrast with widespread global reports of canine companions becoming infected through contact with their SARS-CoV-2-infected owners (Sit et al., 2020; Perisé-Barrios et al., 2021; Patterson et al., 2020). Many variants detected in dogs were not found in the inoculum, reached high frequencies Table 2.2 and were under strong selective pressure (Figure 2.4). The majority of canine variants were found in nonstructural replicase genes Table 2.2, providing strong evidence of selective pressure and host-viral molecular interactions in dogs directed towards selection for viral replication function. The differences between reports of natural reverse-zoonotic infections and laboratory exposures in dogs may stem from disparate doses, dose frequency, or strains (human-adapted or cell culture-propagated virus).

Both domestic and nondomestic cats appear to be highly susceptible to SARS-CoV-2 infection, and readily transmit virus to other cats (A. M. Bosco-Lauth et al., 2020; McAloose et al., 2020; Fernández-Bellon et al., 2021). The similarity between sequences recovered from Cat 5 and Cat 6 (exposed via direct contact with Cat 5), demonstrates the ability of novel mutations to be horizontally transmitted (Table A.4). In contrast to dogs, high viral titers were recovered from cats (A. M. Bosco-Lauth et al., 2020), and fewer consensus and shared variants arose in cats than dogs, suggesting that SARS-CoV-2 viral pathogenesis in cats is more similar to humans than dogs (Rudd et al., 2021). SARS-CoV-2 RNA recovered from hamsters one-day post inoculation had relatively low frequency variants compared to dogs and cats Table 2.2, but was under strong selective pressure (Figure 2.4). Virus recovered from one ferret did not demonstrate reversion of Vero E6 cell-passaged variants (Figure 2.2). However, spike residue 501 was under selection as has been reported in other mustelids (Richard et al., 2020; van Dorp et al., 2020; Sawatzki et al., 2021).

2.4.3 Variants of concern

We detected variation at positions that are implicated in the adaptive evolution of SARS-CoV-2 in humans. Some variants, like spike D614G, are identical to variants from human infections. Others, like spike D215N and N501T, represent different substitutions at the same sites as "variants of concern," which are termed as such due to their potential effects on viral fitness, transmission, replication, or immune evasion. The rapid emergence of these variants in RNA collected 1-3 days post-inoculation illustrates substantial adaptive pressures shaping viral evolution in early stages of host-switching.

Spike N-terminal domain (NTD) variants H69R, D215N, D138Y were detected in cats, dogs and hamsters (Figure 2.3D). D215N reached highest frequencies in dogs and H69R went to fixation in Cats 5 and 6. In contrast, D215H was present at fixation in the inoculum (Figure 2.2). In humans the spike mutation D215G is characteristic of the B.1.351 lineage that emerged in South Africa in October 2020 (Tegally et al., 2021). Position 215 is also adjacent to an insertion in virus recovered from deer mice (*Peromyscus maniculatus*; (Fagre et al., 2021)). D138Y is a defining mutation of the P.1 lineage that emerged in Brazil in December 2020 (Faria et al., 2021).

The frequency of NTD variants reported in humans, and in animals in this study, suggests this domain is a hotspot for variants with transmission or replication advantages to emerge. It is somewhat surprising that the number of variants identified here exceeded those detected in receptor binding domain (RBD, Figure 2.3D), as RBD-ACE2 interactions resulting in viral entry have been projected to be key determinants for species susceptibility (Damas et al., 2020; Yang et al., 2020). At least one neutralizing human antibody that binds to the NTD has been identified (Chi et al., 2020) and this domain has been intensely scrutinized as a vaccine epitope. NTD variants are thus of concern as adaptive changes in this region may facilitate immune evasion.

Amino acid 501, asparagine (N) in initial isolates of SARS-CoV-2, is located in the RBD of the spike protein (Figure 2.3D). This position is a threonine (T) in SARS-CoV and in the closely related coronavirus isolated from a pangolin (Tang et al., 2020). We observed the rapid emergence of N501T in one ferret, and uncovered its occurrence at similarly high frequencies in samples from

two additional ferrets through targeted PCR and direct sequencing. The identical variant arose in 11/11 ferrets following experimental infection (Richard et al., 2020), and has also emerged repeatedly on mink farms in association with 69-70del (van Dorp et al., 2020). Thus, these amino acid changes may contribute to adaptation of SARS-CoV-2 to ferret and mink hosts (Sawatzki et al., 2021; Welkers, Han, Reusken, & Eggink, 2021). The co-occurrence of spike 69-70del and N501Y has been recorded in emergent SARS-CoV-2 lineages in humans (Andrew, 2020; Meng et al., 2021). The N501Y substitution has been shown to increase binding affinity to the human ACE2 (K. K. Chan, Tan, Narayanan, & Procko, 2021).

H655Y, a feature of the Brazilian P.1 lineage in humans, was detected in RNA from three of six cats, all three dogs, and two hamsters, and was present at low frequency (1.1%) in the P3 viral inoculum Table 2.1. Spike amino acid 655 is located near the S1/S2 cleavage site of the spike protein between the RBD and the fusion peptide. This variant has been reported to be under positive selection in experimentally inoculated cats (Braun et al., 2021) and has been shown to emerge following one passage of pseudotyped SARS-CoV-2 in cell culture in the presence of anti-spike human antibodies (Baum et al., 2020).

Spike variant D614G was the first widely-recognized emergent mutation in human SARS-CoV-2 lineages in early 2020 and has rapidly increased in prevalence to become the dominant sequence worldwide. By the end of February 2021, D614G represented 94.8% SARS-CoV-2 sequences publicly available in the GISAID database (Shu & McCauley, 2017). The growth advantage of this variant is ascribed to replication fitness advantage (Korber et al., 2020; Grubaugh, Hanage, & Rasmussen, 2020). We identified D614G in two inoculated cats. This variant was not detectable in the inoculum, and thus likely arose and was under selection in cats. The variant reached fixation in Cat 5, and was transmitted to Cat 6 via contact infection, indicating this variant could become established in a cat-to-cat transmission chain. While it is important to consider each amino acid change in SARS-CoV-2 in conjunction with all changes across the viral genome, the detection of specific substitutions in experimentally inoculated animals identical to changes in viral lineages

infecting humans may suggest convergent evolution or the potential for variants to originate in animals followed by spillover back into humans.

Variant L37F in nsp6, detected in RNA from cats, dogs and hamsters, also emerged early during the COVID-19 pandemic as a defining mutation of the GISAID Clade V; however, it was not as ultimately successful in spreading in humans as D614G (Shu & McCauley, 2017). We detected L37F resulting from a single nucleotide change in all four species. Two frameshift mutations (nucleotide insertion and nucleotide deletion) were also detected at L37. Although there is no longer sustained public attention on L37F in human SARS CoV-2, previous research has linked the presence of the L37F G>T SNV to asymptomatic infection in humans (R. Wang, Chen, Hozumi, Yin, & Wei, 2020).

Given the susceptibility of companion animals to SARS-CoV-2 infection and reports of SARS-CoV-2 transmission from humans to animals and then back to humans (Table A.1), the rapid adaptation we document illustrates the potential of reverse zoonosis to accelerate variant emergence for SARS-CoV-2 and other viruses. The rapidity of in vitro and in vivo adaptation underscores the extraordinary plasticity and adaptive potential of SARS-CoV-2. Pathogens are under strong selective pressure to propagate in the host environment, while host defenses are aligned to prevent pathogen replication. This host-pathogen arms race results in varied outcomes that can lead to increases or decreases in virulence and transmission (Longdon et al., 2015). Thus, monitoring SARS-CoV-2 evolution in hosts with the potential to infect humans should be a high priority, as de novo novel variants or recombinants between human and animal strains could result in altered transmission pathways and vaccine efficacy (Prévost & Finzi, 2021).

Our work additionally illustrates that virus evolution and adaptation following passage in cell culture and experimental infection is far more complex than has typically been acknowledged. The advent of new technologies that afford the opportunity to assess viral quasispecies within inocula and biological samples provides an exciting opportunity for future studies of viral evolution and pathogenesis in both humans and animal hosts.

2.5 Materials and Methods

2.5.1 Cell culture passage in vitro

SARS-CoV-2 strain USA-WA1/2020 (Genbank MN985325.1) was obtained (Harcourt et al., 2020) and passaged in Vero E6 cells a total of three times. 100 μL of viral stock was inoculated onto a flask of confluent Vero E6 cells, allowed to adsorb for 30 minutes, and incubated in cell culture media (Dulbecco's Modified Eagle Media supplemented with 10% fetal bovine serum and antibiotics) for 3-4 days, harvested and frozen in aliquots.

2.5.2 Infections in vivo

Intranasal inoculations of cats, dogs, and ferrets with SARS-CoV-2 infection were conducted in ABSL3 facilities at Colorado State University. We conducted a comprehensive analysis of infection outcomes, including assessment of virus neutralization, seroconversion, cat-to-cat transmission, and resistance to reinfection (A. M. Bosco-Lauth et al., 2020). SARS-CoV-2 strain USA-WA1/2020 was expanded in Vero E6 cells, and all four species were inoculated with the same "Passage 3" viral stock. Cats 2, 3 and 4 were in Cohort 1, and Cats 1 and 5 were in Cohort 2. Cat 6 was co-housed with Cat 5 for contact transmission of the virus (A. M. Bosco-Lauth et al., 2020). Animals were lightly anesthetized and between 10^5–10^6 pfu SARS-CoV-2 was instilled via the nares. Oropharyngeal and/or nasal samples collected for up to 10 days. Oral swabs were placed in BA-1 medium (Tris-buffered MEM containing 1% BSA) supplemented with gentamicin, amphotericin B and penicillin/streptomycin. Nasal flushes were performed by instilling 1 mL BA-1 dropwise into the nares of awake or lightly anesthetized cats, dogs and ferrets and collecting nasal discharge into a sterile petri dish by allowing the wash fluid to be sneezed out or dripped onto the dish. Viral titers (expressed as plaque-forming units [pfu]) were assessed for all infected individuals (A. M. Bosco-Lauth et al., 2020).

2.5.3 RNA extraction and sequencing

100 μL of lavage fluid was added into 900 μL of trizol to inactivate the virus and prepare the sample for RNA extraction. RNA was extracted using a modified Zymo RNA Clean and Con-

centrator 5 kit (Zymo Research, Irvine, CA, USA). Samples were thawed on ice and 200 μL of chloroform was added. Samples were agitated by hand for 15 seconds, incubated at room temperature for at least two minutes and centrifuged for 10 minutes at 12,000g. 450 μL of the aqueous phase was removed and placed in 1.5 mL microcentrifuge tubes containing 900 μL of 1:1 Zymo RNA binding buffer and 100% EtOH (450 μL each). Samples were vortexed and then 750 μL of each was transferred to RNA CC-5 columns and centrifuged for 1 minute at 12,000g. Each column was washed with 400 μL of Zymo RNA wash buffer, centrifuged again for 1 minute at 12,000g. A mixture of 24 μL Zymo RNA wash buffer, 3 μL DNase I buffer, and 3 μL DNase I (both New England Biolabs, Ipswich, MA, USA) was added to 30 μL of the mixture and incubated for 22 minutes at room temperature. Following incubation, columns were spun in collection tubes for 30 seconds at 9000g. RNA prep buffer (400 μL) was added to each column and centrifuged again. Flow through was discarded and columns washed with 800 μL RNA wash buffer, followed immediately by a second wash with 400 μL. Wash buffer was discarded and samples were centrifuged for one minute at 16,000g to dry the membrane. We transferred columns to labeled sterile microcentrifuge tubes and to elute RNA. Immediately following RNA extraction cDNA was generated by adding 1 μL of 3μg/μL of random primers (Thermo Fisher, Waltham, MA, USA) and 1 μL of 10 mM dNTPs (New England Biolabs, Ipswich, MA, USA) to 10 μL RNA and incubating at 65°C for 5 min on a C1000 Touch Thermal Cycler (Bio-Rad, USA) followed by chilling on ice. We added 4 μL 5x First-Strand Buffer, 2 μL 0.1M DTT, and 1 μL RNase OUT (40 units/μL) to each sample, pipette mixed, and incubated at 25°C for 2 minutes. 1 μL (200 units) of SuperScript II RT was added pipette mixed, and incubated at 25°C for 10 min, 42°C for 50 min, and 70°C for 15 min. cDNA was frozen at -20°C until use in library preparation.

2.5.4 Amplicon generation and library preparation

We employed an amplicon based next generation sequencing approach using the primers and protocols developed and optimized for SARS-CoV-2 by the ARTIC Network (Itokawa, Sekizuka, Hashino, Tanaka, & Kuroda, 2020; Quick et al., 2017). All samples were sequenced with ARTIC version 3 primers except for Cat 5 and one replicate of Cat 4, which were sequenced with the

previous version 2 primers. This sequencing method generates high coverage of viral genomes from even low template samples by utilizing an initial PCR step resulting in 400bp amplicons and primers that span the coding region (Quick et al., 2017). Briefly, two primer pools were used for the initial PCR amplification and then pooled prior to library preparation. We then visualized PCR products on agarose gels and quantified them with the Qubit Broad Range kit (Thermo Fisher, Waltham, MA, USA). Samples from Ferret 2 and Ferret 3 were not sufficiently amplified for downstream next generation sequencing. We normalized samples to 540ng of cDNA. We prepared sequencing libraries using the New England Biolabs Ultra II kit with the only modification being the use of Ampure XP beads for cleanup and size selection (Beckman Coulter, Brea, CA, USA). Following adapter ligation a single size selection was conducted at 0.65X. We then pooled libraries for sequencing on Illumina MiSeq (Illumina, San Diego, CA, USA) using the v2 500 cycle 2 by 250bp kit.

2.5.5 Targeted PCR and cloning for spike variant N501T

We designed a targeted PCR for the region of the spike gene containing N501T to amplify cDNA from the two additional ferrets in our study that did not adequately amplify with ARTIC primers. PCR amplification was carried out according to the manufacturer's protocol for PlatinumTM SuperFiTM PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA) using the following primer set: Forward: AGGCTGCGTTATAGCTTGGA; Reverse: CTGTGGATCACGGACAGCAT. Products were visualized on agarose gels and purified using the MEGAquickspinTM Plus Total Fragment DNA Purification Kit (iNtRON Biotechnology, Gyeonggi-do, South Korea), prior to being ligated into pJET1.2/blunt vectors and cloned using a CloneJet PCR cloning kit (Thermo Fisher Scientific, Waltham, MA, USA). Plasmids were purified from Ferret 2 (n=6 clones) and Ferret 3 (n=3 clones) with the DNA-Spin Plasmid Purification Kit (iNtRON Biotechnology, Seongnam-si, Gyeonggi-do, South Korea) according to the manufacturer's protocol, and direct sequenced in both directions (Psomagen, Rockville, MD, USA). Sequences were aligned to the USA-WA1/2020 reference sequence and interrogated at position 23064 for the presence of the A>C single nucleotide change that causes spike N501T.

2.5.6 Bioinformatics and data analysis

Raw sequencing data were input into a comprehensive Nextflow pipeline for processing next generation sequencing data and single nucleotide and structural variant calling. Nextflow is a bioinformatics workflow manager that facilitates the development of complex and reproducible computational pipelines (Di Tommaso et al., 2017). Briefly, data were trimmed for adapters and low quality using Cutadapt (Martin, 2011), followed by aligning reads to the viral reference sequence. Data were pre-processed for quality with GATK (McKenna et al., 2010) prior to calling single nucleotide and structural variants with LoFreq (Wilm et al., 2012a). SnpEff and SnpSift were used to annotate variants and predict their functional effects (Cingolani, Platts, et al., 2012; Cingolani, Patel, et al., 2012). We designed our variant calling pipeline to ensure that we could differentiate between variants that were not detected due to their absence or presence at a frequency below our detection threshold as compared to inadequacy of coverage depth. The outputs of these analyses were tabulated, processed and visualized in R.

UCSF Chimera software was used to visualize the location of variants in the three-dimensional structure of the SARS-CoV-2 spike protein (Pettersen et al., 2004). SNPGenie was used to calculate nonsynonymous and synonymous nucleotide diversity (Nelson & Hughes, 2015). SNPGenie takes a population of sequences and estimates the mean number of pairwise differences per site, or the nucleotide diversity. Estimates are weighted by allele frequencies. The SNPGenie snpgenie.pl takes as input variant call information contained in .vcf files, a reference genome and genome annotations, and outputs estimates of nucleotide diversity at nonsynonymous and synonymous coding sites at both a genome/population level and by gene product. One advantage of using the nucleotide diversity statistic (π) to compare intrahost viral diversity is that it is not biased by variations in sequencing depth (Zhao & Illingworth, 2019).

2.5.7 Data and code availability

All SARS-CoV-2 raw sequence data used in this study will be publicly available upon publication in the NCBI SRA database under BioProject PRJNA704947. The bioinformatics pipeline used to analyze the data is publicly available at `https://github.com/stenglein-lab/`

viral_variant_caller. The code used to create the three-dimensional visualization of the spike protein is available at: https://github.com/stenglein-lab/highlight_residues_on_spike_structure. Additional data including variant summary tables output by the pipeline and R scripts for data processing and visualization are available at https://github.com/laurabashor/SARS-CoV-2-evolution.

2.6 Funding

This work was funded by the Colorado State University College of Veterinary Medicine and Biomedical Sciences Research Council Award. Computational resources were supported by National Institute of Health/National Center for Advancing Translational Science (NCATS) Colorado Clinical and Translational Science Awards (CTSA) grant UL1 TR002535.

Table 2.2: Nonsynonymous SARS-CoV-2 variants emerge following experimental inoculation in cats, dogs, hamsters, and one ferret. SARS-CoV-2 isolate USA-WA1/2020 was expanded for three passages in Vero E6 cells and 10^5 to 10^6 pfu was inoculated intranasally in cats, dogs, hamsters, and the ferret (n = 6, 3, 3, and 1). Virus was recovered 1- to 3-d post-inoculation (dog and cat infections described in (A. M. Bosco-Lauth et al., 2020)), sequenced using a tiled amplicon technique and analyzed with a pipeline for calling SNVs and SVs in viral populations. Variants representing >50% of recovered genomes in at least one individual or that were found in all individuals of a species are indicated here. Bold indicates variants occurring in >50% of genomes; boxes indicate variants detected in all sampled individuals of that species. Only one synonymous mutation met the criteria and is displayed. Variants that emerged following the passage in Vero E6 cells (Fig. 2) are not listed. Each point represents the mean of two technical replicates, aside from Cat 5, for which a replicate was not sequenced.

Variant	Position	CDS	N/S	C1	C2	C3	C4	C5	C6	D1	D2	D3	F1	H1	H2	H3
G59D	441	nsp1	N		0.15	0.06	0.1			0.66	0.14	0.66		0.04		0.07
E195G	3303	nsp3	N							0.53		0.56				
L37F	11083	nsp6	N		0.15	0.06	0.09	0.05	0.06	0.64	0.1	0.49		0.04		0.06
I242V	18763	nsp14	N		0.18	0.06	0.1			0.71	0.15	0.58				0.06
H69R	21768	S	N		0.06			1	0.99		0.14			0.06		
D138Y	21974	S	N		0.07	0.06	0.05			0.06	0.07	0.29		0.1	0.12	0.1
D215N	22205	S	N		0.17	0.07	0.1			0.72	0.14	0.78				0.06
N501T	23064	S	N										0.76			
D614G	23403	S	N		0.16			0.99	0.98							
S686G	23618	S	N	0.99												
S43Y	28021	orf8	N							0.53		0.44				
N4N	28285	N	S		0.15	0.06	0.11			0.62	0.1	0.52				0.05

Chapter 3

Rapid evolution of SARS-CoV-2 in domestic cats[2]

SARS-CoV-2 infection of a novel permissive host species can result in rapid viral evolution. Data suggest that felids are highly susceptible to SARS-CoV-2 infection, and species-specific adaptation following human-to-felid transmission may occur. We employed experimental infection and analysis of publicly available SARS-CoV-2 sequences to observe variant emergence and selection in domestic cats. Three cohorts of cats (N=23) were inoculated with SARS-CoV-2 USA-WA1/2020 or infected via cat-to-cat contact transmission. Full viral genomes were recovered from RNA obtained from nasal washes 1-3 days post-infection, and analyzed for within-host viral variants. We detected 118 unique variants at $\geq$3% allele frequency in two technical replicates. Seventy of these (59%) were nonsynonymous single nucleotide variants; the remainder were synonymous single nucleotide variants or structural variants. On average, we observed 12 variants per cat, nearly 10-fold higher than what is commonly reported in human patients. We observed signatures of positive selection in the spike protein and the emergence of 11 within-host variants located at the same genomic positions as mutations in SARS-CoV-2 variant lineages that have emerged during the pandemic. Fewer variants were noted in cats infected from contact with other cats and in cats exposed to lower doses of cultured inoculum. An analysis of 93 publicly available SARS-CoV-2 consensus genomes recovered from naturally infected domestic cats reflected variant lineages circulating in the local human population at the time of sampling, illustrating that cats are susceptible to SARS-CoV-2 variants that have emerged in humans, and suggesting human-to-felid transmission occurring in domestic settings is typically unidirectional. These experimental results underscore the rapidity of SARS-CoV-2 adaptation in felid hosts, representing a theoretical potential origin for variant lineages in human populations. Further, cats should be considered susceptible hosts capable of shedding virus during infections occurring within households.

[2]This chapter is a reproduction of "Bashor, L., Gagne, R. B., Bosco-Lauth, A., Stenglein, M., & VandeWoude, S. (2022). Rapid evolution of SARS-CoV-2 in domestic cats. *Virus Evolution*, 8(2). doi: 10.1093/ve/veac092". Minimal modifications were made to meet dissertation formatting requirements.

3.1 Introduction

The devastating global effects of the COVID-19 pandemic have drawn attention to the relatively broad host species range of the SARS-CoV-2 (SARS2) virus. In addition to humans, known susceptible hosts include a wide range of wild and domestic animals that have been naturally or experimentally exposed to the virus. The variety of animal species infected by SARS2 highlights the importance of studying the virus in other susceptible hosts, with the goal of evaluating the potential for animals to act as disease reservoirs and mediate further viral evolution and to characterize the changes associated with cross-species coronavirus transmission.

Many species within the carnivore family Felidae are susceptible to SARS2, and transmission from human-to-felid and felid-to-felid have been established (Garigliany et al., 2020; Gaudreault et al., 2020; Halfmann et al., 2020; Shi et al., 2020; A. M. Bosco-Lauth et al., 2020; Braun et al., 2021). Experimental exposures of domestic cats (*Felis catus*) to SARS2 have also been previously described (Halfmann et al., 2020; A. M. Bosco-Lauth et al., 2020; Gaudreault et al., 2020; Braun et al., 2021; Shi et al., 2020). Although most experimentally infected cats remained asymptomatic throughout infection, a range of respiratory and gastrointestinal signs have been reported in domestic cats naturally infected by their COVID-19 positive owners (Garigliany et al., 2020; Segalés et al., 2020; Curukoglu et al., 2021; Ferasin et al., 2021; Hosie et al., 2021; Keller et al., 2021; Klaus et al., 2021; Neira et al., 2021; Barroso-Arévalo, Sánchez-Morales, Pérez-Sancho, Domínguez, & Sánchez-Vizcaíno, 2022; Zoccola et al., 2021). Infections of large numbers of captive wild felids have been documented throughout the pandemic (of Agriculture Animal & Service, 2022; Shu & McCauley, 2017; McAloose et al., 2020; Bartlett et al., 2021; Fernández-Bellon et al., 2021; Karikalan et al., 2021; Mishra et al., 2021), with occasionally fatal consequences for some individual animals (Thebault, 2021).

Recurring reports of SARS2-infected felids suggest that human-to-felid transmissions occur frequently, and infections of domestic household cats are likely unreported due to the lack of recognizable symptoms in most individuals. Several studies have aimed to estimate the incidence of domestic cat infections using serology. All but one study (Temmam et al., 2020) of cats living in

households with known COVID-19 infection have detected antibodies in some proportion of cats, ranging from 3% to over 50% of cats tested (Barrs et al., 2020; Calvet et al., 2021; Fritz et al., 2020; Hamer et al., 2021). An early study conducted in Wuhan, China reported 14.7% seropositivity in cats (Zhang et al., 2020) while other surveys of domestic cat populations have observed rates of infection, ranging from 0 to 5.8% (Patterson et al., 2020; Villanueva-Saz et al., 2022), collectively, revealing the regularity of felid infections.

SARS2 genome sequences recovered from experimentally or naturally infected felids have led to the identification of some single nucleotide polymorphisms (SNPs), or consensus variants (present at >50% frequency within a host) but no felid-adapted variants have been substantiated to date (Barrs et al., 2020; Braun et al., 2021; Hosie et al., 2021; Klaus et al., 2021; Bashor et al., 2021). In general, sequences recovered from naturally infected felids have closely resembled sequences recovered from their owners or zoo keepers. A smaller subset of experimental studies have characterized within-host viral diversity in domestic cats and reported cat-to-cat transmission of viral variants (Braun et al., 2021; Bashor et al., 2021). Cat-to-cat transmission was characterized in one study by a narrow bottleneck, with only 2-5 viral genomes transmitted (Braun et al., 2021), which is consistent with what occurs between humans (Lythgoe et al., 2021). However, both low frequency and consensus variants were reliably transmitted, including variants under positive selection (Braun et al., 2021; Bashor et al., 2021). We have recently reported the rapid emergence of numerous variants in experimentally exposed cats, including mutations identical to or at the same genomic position as mutations characteristic of human SARS2 variant lineages (Bashor et al., 2021).

It is likely that SARS2 faces unique selective pressures when it “spills over” into new hosts, which results in viral variant emergence and virus-host adaptation. The prevalence of SARS2 in felids draws attention to the possibility for spillback infections and novel variant transmission from felids to humans. Among species known to be highly susceptible to SARS2 and capable of shedding virus following infections, domestic cats have the most intimate contact with humans, with approximately 25% of US households supporting an average of 1.8 cats (A.V.M.A, 2017).

Companion cats thus represent a unique concern among SARS2-susceptible nonhuman animals relating to their role in perpetuating infections or resulting in propagation of viral variants that may re-infect humans. Furthermore, repeated natural transmissions and host adaptation would support the potential establishment of a felid reservoir of SARS2.

In this study, we assessed SARS2 adaptation in felids through experimental exposure of three experimental cohorts of domestic cats to SARS2 and analysis of existing SARS2 sequences from cats. Experimental infection systems offer a unique opportunity to study controlled transmission and observe viral variant emergence and adaptation. Eighteen cats were directly inoculated, and five others were exposed through contact with one or more inoculated cats. A cohort of juvenile cats was exposed to low, medium, or high doses of SARS2. Nasal lavage samples were recovered 1-3 days post infection, and full SARS2 genomes were sequenced at a high depth of coverage to enable within-host viral variant identification. Experimental results were assessed in conjunction with all publicly available genome sequences from SARS2-infected felids for evidence of felid-specific adaptation. Our results demonstrate the rapidity of SARS2 adaptation within felid hosts, and highlight the importance of studying susceptible animal populations in close contact with humans, and cross-species transmission as a mechanism for novel viral variant emergence.

3.2 Results

3.2.1 Sample collection & viral titer

Twenty-three cats from three experimental cohorts were experimentally exposed to SARS2 through direct intranasal inoculation or cat-to-cat contact transmission (Figure 3.1). Cat-to-cat transmission was facilitated by cohousing cats inoculated 24 hours prior with naïve cats. Nasal and oral lavage samples were obtained one- or three- days post-exposure and viral titers were assessed in plaque forming units per milliliter (pfu/mL) (Table B.1). Viral titers from nasal lavage ranged from 0-6.2 log pfu/mL. For all samples with 0 log pfu/mL recovered from nasal lavage, $\geq$1.6 log pfu/mL was detected in the oral lavage. There was a significant positive relationship between dose received and titer recovered via nasal lavage ($p=0.023$, $R^2=0.24$).

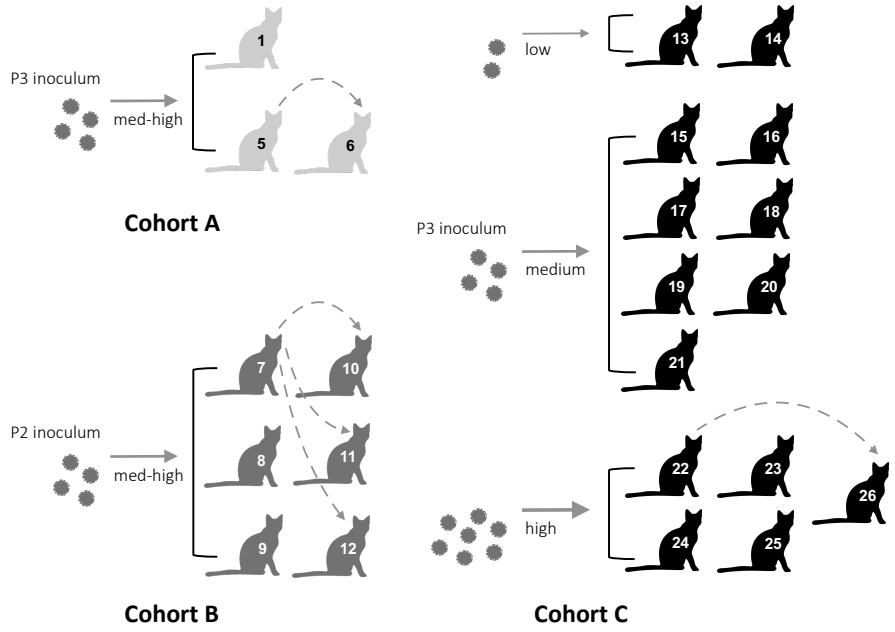


Figure 3.1: Twenty-three cats experimentally exposed to SARS2 were included in this study. Cohorts A and B were inoculated with a standard "medium-high" dose of $1x10^6$ plaque-forming units (pfu). Cohort A included two directly inoculated cats and a contact cat; Cohort B included three directly inoculated and three contact animals; Cohort C included seven cats inoculated with a "medium dose" of $4.6x10^5$ pfu; two cats inoculated with $1.0x10^4$ pfu ("low" dose), and four cats inoculated with $2.1x10^7$ pfu ("high" dose). One contact cat was exposed to the "high" dose animals in Cohort C. All contact cats were co-housed with directly inoculated cats one day post- infection. In Cohorts B and C, samples were collected from directly inoculated and contact cats three days post-exposure. In Cohort A, samples were collected from directly inoculated cats one day post-infection and from the contact cat six days post-exposure. Images from PhyloPic available under Public Domain Mark 1.0 license: Cat (`http://phylopic.org/image/23cd6aa4-9587-4a2e-8e26-de42885004c9/`) by David Orr and SARS-CoV-2 (`http://phylopic.org/image/81bc7804-4940-4fa5-a1ca-dd6c3a26aaa2/`) by Alissa Eckert and Dan Higgins.

3.2.2 Genome sequencing & variant calling

Full viral genome sequences were recovered from RNA isolated from 23 cat-derived nasal lavage samples and three cell culture-derived viral stock samples (P1, P2, and P3). Two technical replicates were sequenced for each sample. Median and mean depths of coverage of the SARS2 genome were 2500x and 4100x respectively (Figure S1). Genomes were analyzed for single nucleotide (SNVs) and structural variants (SVs) present in both technical replicates and $\geq$3% of genome sequences (Figure 3.2). There was no significant relationship between mean or median sequencing coverage and the viral titer measured for an individual cat on the date of sampling (linear models; p=0.076, R^2=0.14 and p=0.56, R^2=0.016 respectively) or between the mean or median depth of coverage and the number of within-host variants in a sample (linear model; p=0.59, R^2=0.012 and p=0.93, R^2=0.00035 respectively).

Across the 23 cat-derived samples, 118 SNVs and SVs in coding regions of the SARS2 genome were observed at $\geq$3% allele frequency in two technical replicates within individual hosts (see Materials and Methods for more detail). Eighteen of these variants were also detected in the P1, P2 and P3 viral stock samples. Therefore, 85% of variants detected in cats (N=100) were not detected in the inoculum stocks at $\geq$3% frequency. Sequence data were also analyzed for variants present at as low as 0.1% frequency (Figure S2) to enable identification of variants that were initially present in the inoculum at low frequencies. Of the 118 variants detected in cats at $\geq$3% allele frequency, 66 (56%) were not detectable in the stock samples above 0.1% frequency. Five hundred fifty-seven variants were identified in cat-derived samples at $\geq$0.1% frequency. Four hundred forty of these (79%) were not observed in the cell culture-derived viral stock samples at $\geq$0.1% frequency.

3.2.3 Variant characteristics

We analyzed 118 variants occurring at $\geq$3% frequency for type, predicted effect, and position in the SARS2 genome (Figure 3.2). One hundred of 118 variants (85%) were SNVs, and 18 (15%) were SVs. The 100 SNVs and 18 SVs were observed 246 and 65 times throughout the dataset respectively, indicating that some were detected in more than one animal. The majority of SNVs (N=70, 70%) were nonsynonymous. SVs included nine frameshift variants, six disruptive

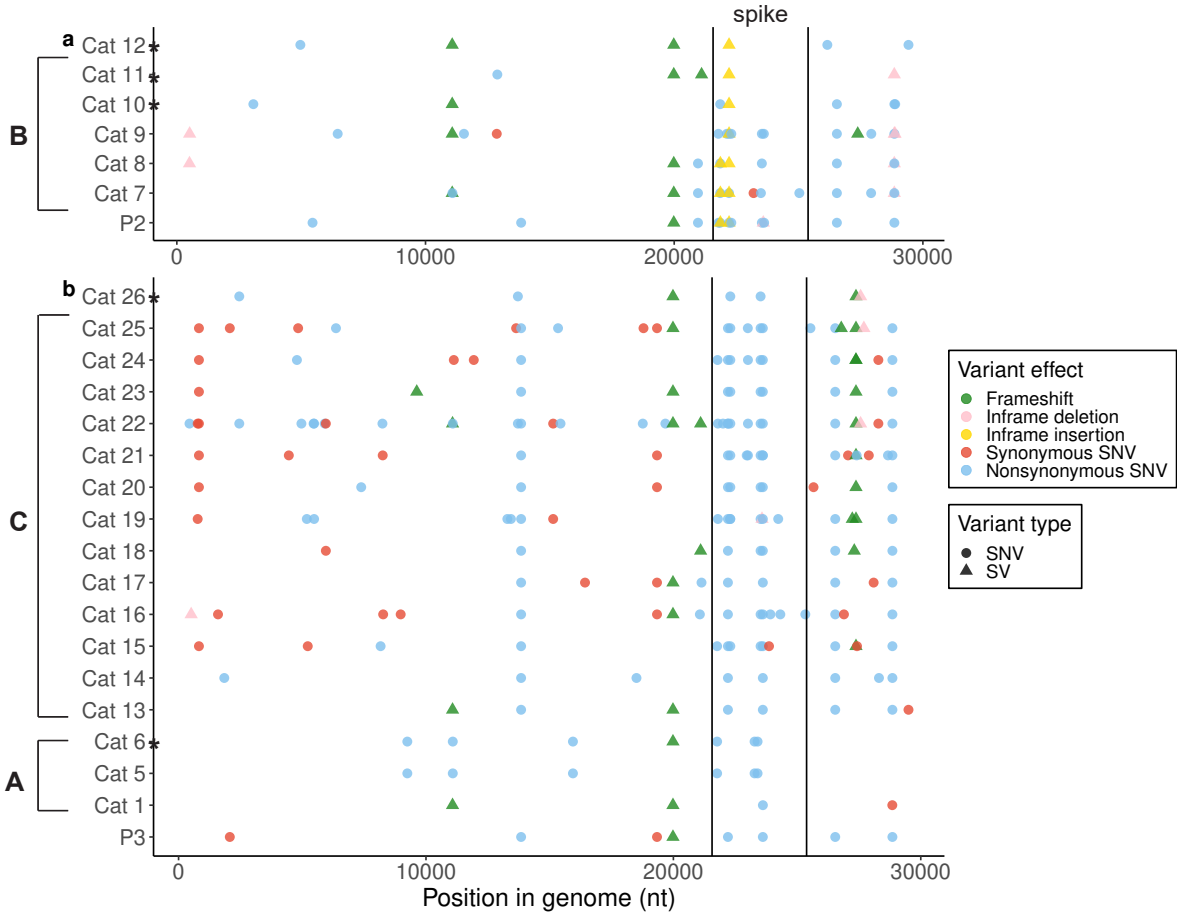


Figure 3.2: SARS2 within-host single nucleotide (SNV) and structural (SV) variants are detected at high frequency following infection of domestic cats. One hundred and eighteen variants were observed 311 times at $\geq$3% allele frequency in 23 experimentally inoculated cats (median = 10 variants per cat). Cats were exposed to (a) Passage 2 (P2) or (b) Passage 3 (P3) cell culture-derived viral stocks. Each point represents a variant and point shape and color indicate variant type and predicted functional effect respectively. Cohorts are denoted on the left-hand side of the plot, and contact cats are indicated with an asterisk after their name along the y-axis. Cats 10, 11 and 12 were cohoused with Cats 7, 8 and 9. Cat 26 was cohoused with Cats 22, 23, 24 and 25. Cat 6 was cohoused with Cat 5. The location of the spike gene is delineated between two vertical lines.

in-frame deletions, one disruptive in-frame insertion, and a single conservative in-frame insertion and deletion respectively. We observed 33 variants in the spike, representing 26% of all variants in a gene segment representing less than 13% of the SARS2 genome. The next largest number of variants was in nonstructural protein 3 (nsp3). Nineteen variants (16%) were detected in this segment which spans approximately 19% of the genome.

Variants were further assessed for convergent emergence among cats, their relation to mutations present in human variant lineages, and the frequency of specific single nucleotide substitutions Table 2.1, Table 2.2. Eleven of these variants were not detected at any level (>0.1% allele frequency) in the viral stock inoculum, including three (D614G, S686G, and F59_Q76del) that are shared between two or more cats not connected by a transmission chain Table 2.1. Fourteen variants originally derived from the viral stock inoculum were shared among three or more cats Table 2.2. Eleven variants occur at the same genomic positions as mutations characteristic of human SARS2 lineages labeled as variants of concern (VOC) by the World Health Organization (WHO) and United States Centers for Disease Control and Prevention (CDC) Table 2.1, Table 2.2. Eight are in the spike gene and occur at the same positions as mutations in Alpha, Beta, Gamma and Omicron variant lineages: H69R, F79L, D215H, D215N, D215_L216insKLRS, E484D, D614G, H655Y; three others are in the nucleocapsid gene and share genomic positions with mutations in the Beta and Omicron variant lineages: P13S, G204_R209del and T205I. We noted that C>T substitutions (C>U changes in the SARS2 genome) constituted the largest proportion (31%) of single nucleotide changes observed in cats in this dataset (Figure S3).

3.2.4 Drivers of variant richness

The mean ± SD number of SARS2 variants per cat was 12 ± 1.4 and the median was 10 (N=23 samples; Figure 3.3a). SARS2 genomes recovered from directly inoculated cats had significantly more variants than those from contact cats, which averaged 6.4 ± 0.4 variants per cat (Figure 3.3b; t-test, p=0.00027). There were also more variants in samples collected 3 days post-infection (dpi) from directly inoculated cats than in samples collected 1 dpi (t-test, p=0.050). There was no significant difference between the mean number of variants detected in cat-derived samples compared

to cell culture-derived samples (t-test, p=0.59). There was not a significant difference among the number of variants detected in directly inoculated cats belonging to the three infection cohorts (linear model; p=0.15, R^2=0.12). There was also no significant difference between the number of variants in female and male cats or between juvenile and adult cats (t-tests, p=0.48 and p=0.23 respectively).

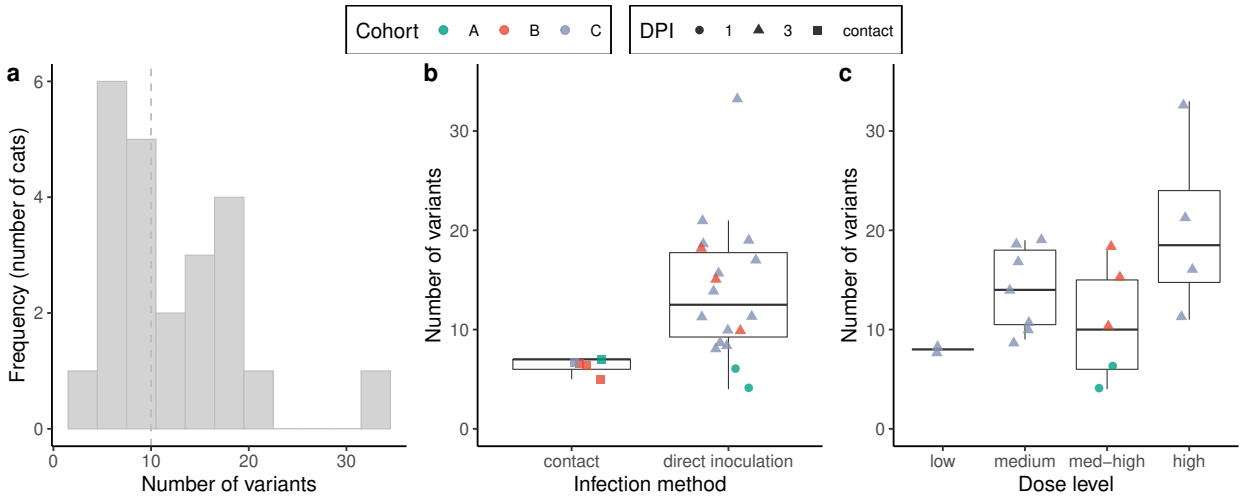


Figure 3.3: Method of infection and infection dose impact within-host variant detection. Variants detected in RNA recovered from SARS2-infected cats varied (a) among cats, (b) as a function of infection method and (c) by dose level. (a) Distribution of variant richness ranges from 4-33. Dashed line indicates the median (10). (b) More SARS2 variants were detected in directly inoculated cats (N=18) compared to contact-infected cats (N=5; t-test). (c) The mean number of variants was higher in samples from cats inoculated with higher doses of SARS2 virus (N=18 directly inoculated cats; linear model). In plots (b) and (c) each point represents an individual cat, and point shape and color indicate days post-infection (dpi) and infection cohort respectively.

We also observed that the number of variants was higher in cats inoculated with higher doses of SARS2 (Figure 3.3c), and a linear model of number of variants as a function of dose revealed a significant positive relationship between dose category and variant richness (p=0.022, R^2=0.38). A linear model assessing log pfu/ml recovered from nasal lavage of directly inoculated cats as a function of both inoculation dose and dpi of sample collection was highly significant (p=0.0073, R^2=0.41). However, there was no significant relationship between this measurement of viral titer

and the number of viral variants detected in RNA extracted from the same samples (Figure S4, linear model, p=0.30, R^2=0.0079).

3.2.5 Contact transmission of SARS2 infection and within-host variants

SARS2 was transmitted to five cats (Cats 6, 10, 11, 12, and 26) following contact with cats experimentally inoculated one day prior to co-housing. In Cohort A, Cat 6 was cohoused with one other cat, Cat 5, whereas in Cohorts B and C contact cats were cohoused with multiple directly inoculated cats. Cats 10, 11 and 12 were cohoused with three other cats (Cats 7, 8 and 9). Cat 26 was cohoused with cats receiving a "high" dose of SARS2 (Cats 22, 23, 24 and 25; Table B.1, Figure 3.1). It was possible to deduce the infection donor for each contact cat by matching SARS2 variants detected in potential donors (at any level above 0.1% allele frequency) with recipients (Figure 3.4). In Cohort B, contact cats 10, 11 and 12 each shared a larger proportion of within-host variants with Cat 7 (6/7, 4/5 and 3/6 variants respectively) than with Cat 8 (4/7, 4/5 and 2/6 respectively) or Cat 9 (5/7, 3/5 and 2/6 respectively). In Cohort C, all seven variants in Cat 26 were also observed in Cat 22.

Putative donor cats 7 and 22 also each had the highest viral titer assessed from nasal lavage of all directly inoculated cats in their cohorts, suggesting transmission was related to viral shedding competency. The viral titer recovered from contact cats varied, and was in some cases lower than that of the infection donor or other directly inoculated cats, and in other cases higher (Table B.1).

3.2.6 Signatures of selection

We measured the nucleotide diversity, or the average number of pairwise changes at nonsynonymous and synonymous sites throughout the genome, to identify signatures of genomic selection. There was not a significant difference between nonsynonymous (πN) and synonymous (πS) nucleotide diversity (paired t-test, p=0.86) across individual full viral genomes. However, signatures of selection emerged across our datasets at the gene level (Figure 3.5). Notably, πN was significantly greater than πS in the spike gene, indicating positive or diversifying selection (paired

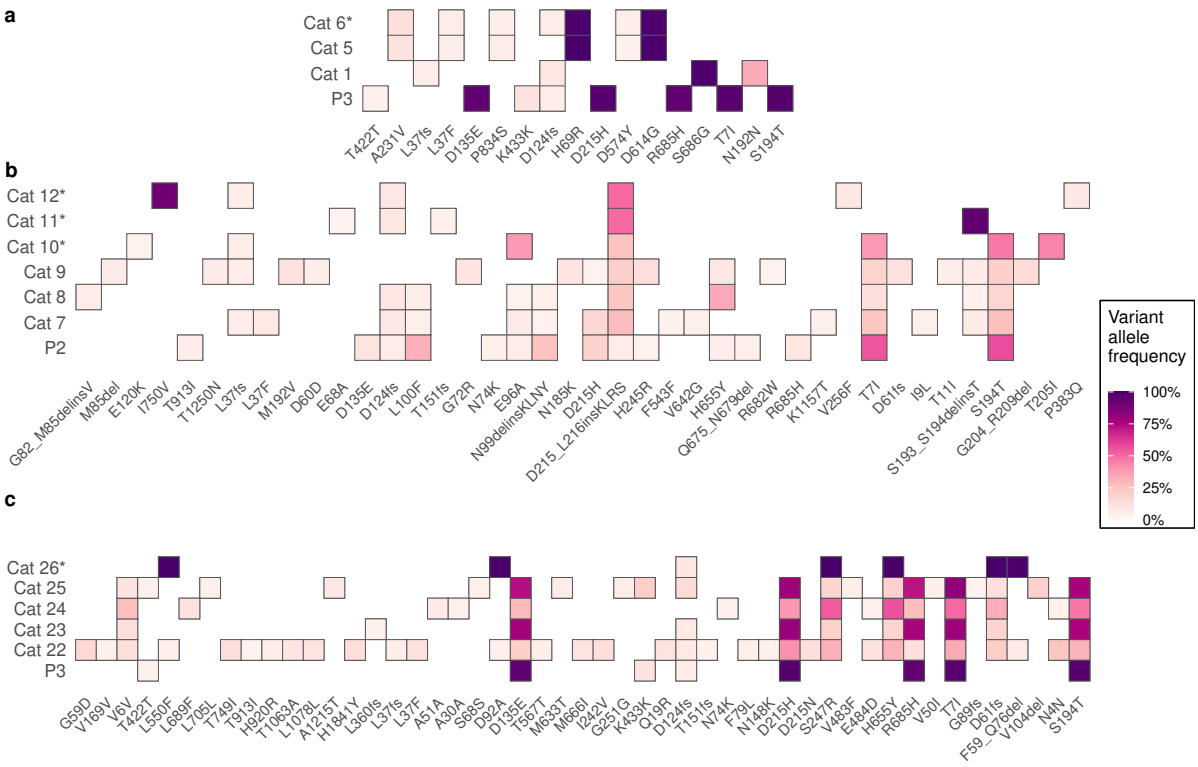


Figure 3.4: SARS2 within-host variants are transmitted between cats in three experimental cohorts. Each plot presents variants detected at $\geq$3% allele frequency in contact cats, directly inoculated cats with which they were cohoused, and the viral stock used for inoculation. In plot (a) contact Cat 6 was infected by contact with Cat 5 which was inoculated with Passage 3 (P3) cell culture-derived viral stock. In plot (b) contact Cats 10, 11 and 12 were cohoused with Cats 7, 8 and 9 which were originally inoculated with Passage 2 (P2) viral stock. In plot (c), contact Cat 26 was cohoused with Cats 22, 23, 24 and 25, originally inoculated with P3 viral stock. Each tile represents a variant and color indicates variant allele frequency. Contact cats are denoted with an asterisk after their name along the y-axes. Code to prepare plots was modified from the R package for outbreak.info: (`https://github.com/outbreak-info/R-outbreak-info`).

t-test, p=0.00024). In contrast, πS was significantly greater than πN in orf1ab, suggesting negative or purifying selection in this region (paired t-test, p=0.012).

Global signatures of felid-specific adaptation of SARS2 One hundred and seventeen full-length SARS2 genomes isolated from domestic cats and 128 genomes from other felid species were submitted to the GISAID EpiCoV public database between March 2020 and the end of March 2022. Ninety-three domestic cat and 29 other felid sequences passed our selection criteria and quality control parameters and were used to generate time-measured phylogenies ((Figure 3.6), Figure S5). Custom Nextstrain builds were prepared separately for sequences from domestic cats alone (Figure 3.6) and all felids (Figure S5). Relationships between SARS2 genomes did not appear to be defined by the emergence of shared, felid-specific SNPs. Instead, sequences clustered based on date of sample collection and the predominant SARS2 human lineage at that time. Almost all sequences after spring 2020 carried the spike D614G mutation. Only one felid (a lion infected in April 2020) did not have this variant sequence. Although one snow leopard (infected in California in July 2021) appeared to diverge from the surrounding sequences, there was one matched sequence from an infected human in the same location and time period assigned to the same clade, suggesting that that particular felid sequence was still representative of some human infections in that area at that time.

3.3 Discussion

In response to the widespread, frequent infection of felids with SARS2, we investigated viral evolution and variant emergence in 23 cats from three experimental cohorts aimed at defining SARS2 infection characteristics in domestic cats. Cats were infected with SARS2 USA-WA1/2020, which represents one of the earliest culture-derived isolates (from a patient infected in January 2020) used in laboratory studies. We observed large numbers of within-host variants in viral genomes recovered from infected cats, with an average of 12 variants per cat sample. This stands in stark contrast to a reported mean of 1.4 variants per human sample (Lythgoe et al., 2021). This result is consistent with the rapid adaptive response of the virus to the cellular environment and immune defenses of a new host species in an experimental setting. Many variants observed

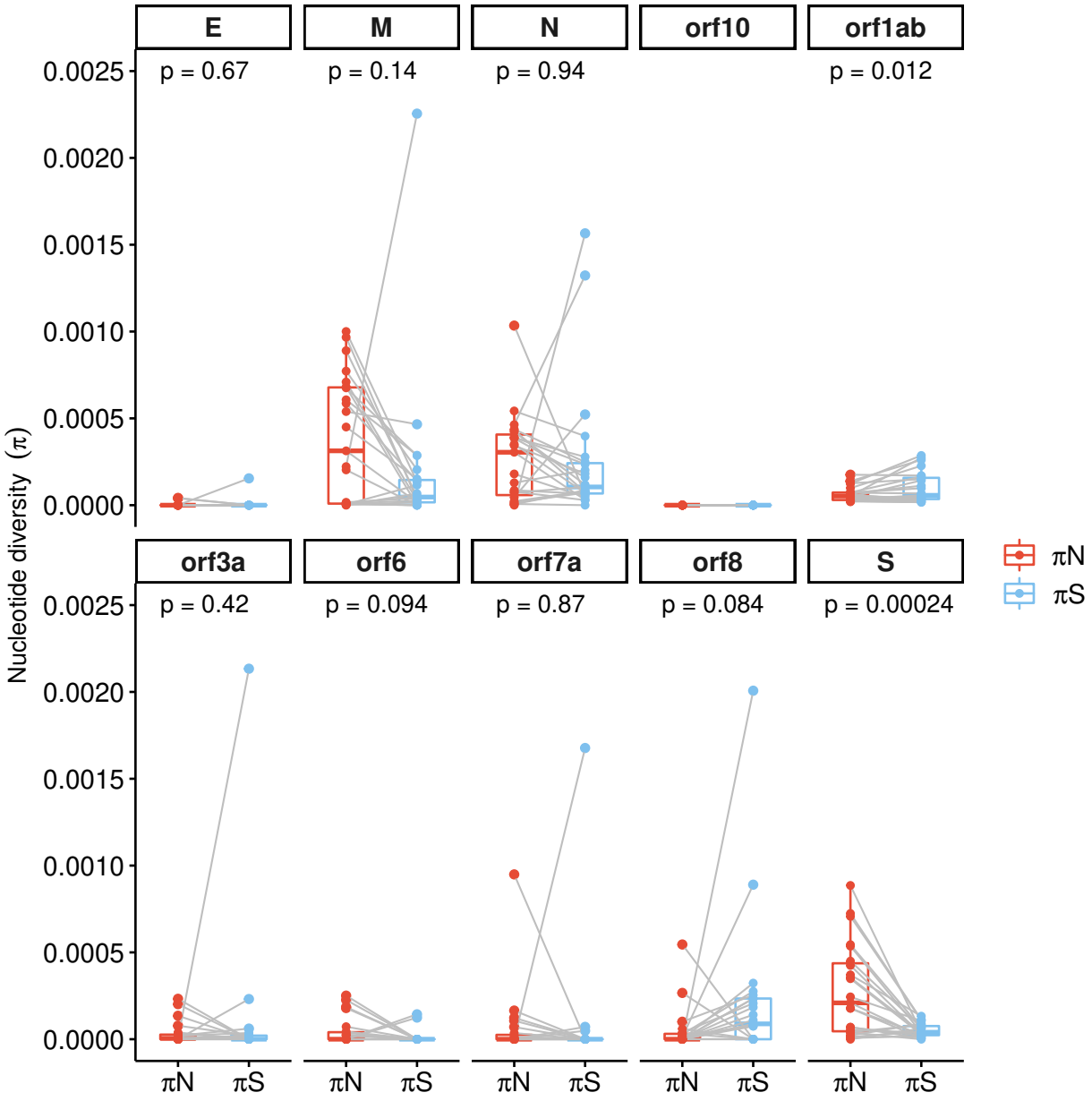


Figure 3.5: The SARS2 spike gene is under strong positive selection compared to other gene segments. Values of nonsynonymous (πN) and synonymous (πS) nucleotide diversity were calculated for each SARS2 gene or open reading frame across a dataset of 23 experimentally infected cats. Each point represents mean πN (red) or πS (blue) for two technical replicates from an individual cat in the indicated SARS2 gene region. P-values are indicated for paired t-tests between mean πN and mean πS.

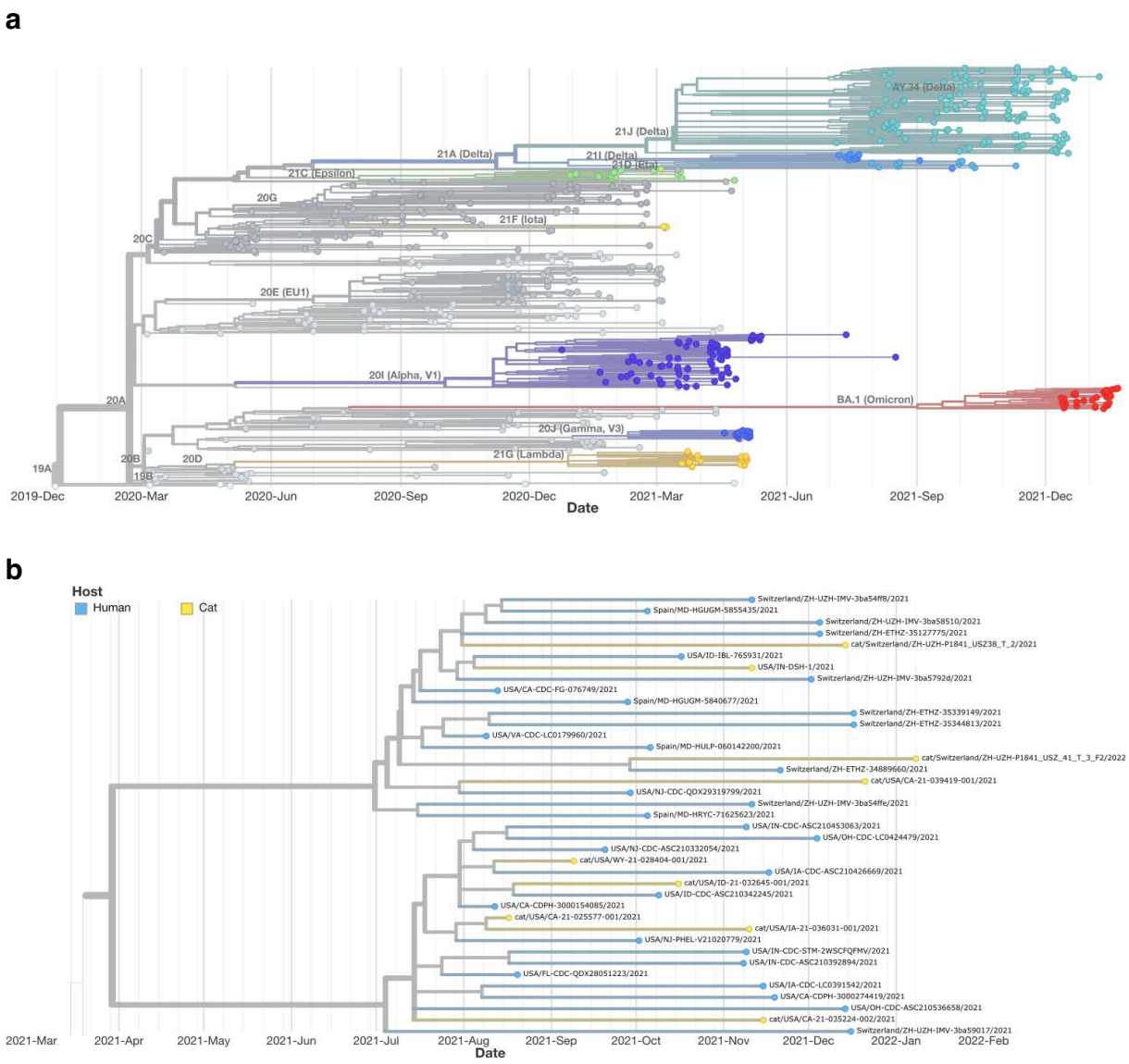


Figure 3.6: Domestic cat-derived SARS2 genomes reflect lineages of SARS2 circulating in humans during the same time period. SARS2 sequences recovered from (A) naturally infected domestic cats (N=93) and humans (N= 543) between March 2020 and March 2022. Sequences are colored by clade as assigned by Nextstrain. The interactive version of this phylogeny is hosted at: `https://nextstrain.org/community/laurabashor/SARS2felids/catshumans`. (B) A highlight from the same phylogeny showing a subset of sequences assigned to the Delta clade demonstrating the lack of speciesspecific patterns. Sequences are colored by host species (blue=human; yellow=cat). Data were obtained from the GISAID database on April 4, 2022, analyzed with the Augur pipeline and visualized with the Auspice web application.

in our study were nonsynonymous, and predominated in the spike protein, indicating potential selection for improved binding and entry of feline host cells, host immune evasion, and other vital species-specific functions.

We witnessed the emergence of 11 SARS2 within-host variants identical to or within the same residue as mutations characteristic of known variant lineages that have emerged since SARS2 USA-WA1/2020 was isolated. The fact that these variants arose or were maintained after just one passage in cats signifies the strength of selective pressures in the feline hosts. Furthermore, it points to a potential mechanism for the introduction of sets of new SARS2 mutations into human populations. Spillback from an animal is a leading theory for the emergence of the Omicron variant lineage in December 2021 (Wei et al., 2021), notable for its large number of mutations, particularly in the spike, with the predominant alternative theory being evolution in an immunocompromised human (Li, de Vries, Kamar, Peppelenbosch, & Pan, 2022). Further, SARS2 spillback from mink, hamsters and deer to people has been documented (Hammer et al., 2021; Munnink et al., 2021; Pickering et al., 2022; Yen et al., 2022), illustrating the potential for animal-to-human transmission of SARS2, even in animals that are not typically in close proximity to humans. Our finding of an increase in SARS2 within-host variants in the viral spike after experimental infection of domestic cats highlights the potential for viral diversification and evolution resulting in a SARS2 variant more adapted for spread in cats.

We also found that cat-derived variants can be transmitted through contact transmission from one cat to another, notwithstanding the potential random effects of bottlenecks and genetic drift on variant allele frequencies during natural transmission. Shared within-host variants allowed us to confirm the donor-recipient relationship despite multiple potential donors in Cohorts B and C, and illustrated that contact cats tended to be infected by donor animals with a higher viral titer measured in nasal flush. The observed decrease in variant richness in contact cats is consistent with previous reporting of a narrow bottleneck during cat-to-cat transmission (Braun et al., 2021). The finding that higher number of variants occurred in animals exposed to higher viral titers supports the potential for variant emergence when viral exposure levels and viral shedding are high.

Although experimental infections cannot completely recapitulate natural infections and the stochastic processes that affect viral populations as they spread between and within hosts, they still provide valuable insights into the dynamics of viral evolution. They also provide the advantage of a controlled environment in which input and output viruses can be characterized. We used the same approach to sequence and call variants in samples of the viral inoculum and samples recovered from cats. This experimental design allowed us to identify what fraction of variants were present in low frequencies in the inoculum and increased in frequency in cats as compared to variants that were not detected in the inoculum and may have arisen de novo. There were consistently more of the latter group of putative de novo variants than low-frequency inoculum variants in our dataset, although proportions of each group varied as a function of the stringency of allele frequency cutoff used. Out of the 118 within-host variants present in cats at >3% frequency, 52 (44%) were detectable in viral stock samples and 66 (56%) were not. It is also worth noting that the cell culture-derived viral stocks used to inoculate the cats contained variants that have been described previously as SARS2 adaptations to Vero cell culture (Bashor et al., 2021).

Cat-to-cat transmission findings demonstrate the general transmissibility of within-host variants, and led us to question what might result from the onward transmission of cat-adapted SARS2 among cats. We analyzed publicly available SARS2 genomes from domestic cats (N=93) matched with SARS2 genomes recovered from temporally and geographically representative human patients (N=543). We found that naturally occurring feline SARS2 genomes were most closely associated with dominant SARS2 lineages present in human populations at the time of sample collection. This is consistent with frequent transmission of SARS2 among humans and felids, although we were not able to trace individual spillover-spillback events with limited information about sample origin. Our analysis documents that felids have continued to be susceptible to newly emergent and increasingly transmissible human variant lineages over the course of the pandemic, as has been previously reported for several felid species (Shu & McCauley, 2017; Curukoglu et al., 2021; Ferasin et al., 2021; Karikalan et al., 2021; Keller et al., 2021; Mishra et al., 2021; Barroso-Arévalo et al., 2022; Zoccola et al., 2021).

Ultimately, our experimental infection system provided us with unique information that would be difficult or impossible to obtain in a natural setting. In particular, we were able to sequence both the input and output viruses, for both direct inoculations and contact transmissions, and we were able to assess viral titer at relevant timepoints. This allowed us to determine that a higher viral titer was more likely to result in cat-to-cat transmission when a naïve cat was cohoused with a group of infected cats. This is consistent with reported SARS2 transmission dynamics in humans (Goyal, Reeves, Cardozo-Ojeda, Schiffer, & Mayer, 2021; Marks et al., 2021). A limitation of our approach was that we sequenced samples at a single timepoint, so we were unable to document changes in frequency of within-host variants over time. However, we observed shared variants transmitted from infection donor to recipient cats in five independent contact transmission events, affirming that our samples were collected at a biologically relevant timepoint.

It is widely accepted that SARS2 originally spilled over from an animal host (F. Wu et al., 2020; Zhou et al., 2020), and viral spillback from farmed mink (Hammer et al., 2021; Munnink et al., 2021), hamsters (Yen et al., 2022) and deer (Pickering et al., 2022) to humans has been reported. In our phylogenetic analysis of naturally occurring felid sequences, we were not able to discern specific consensus variants occurring in cats prior to widespread circulation in people, and felid-to-human SARS2 transmission has not been confirmed by sequence analysis from intensive and coincident animal and human sampling. However, as evidenced by our and other analyses of reported SARS2 genomes, human-to-felid transmissions are frequent (Barrs et al., 2020; Patterson et al., 2020; Temmam et al., 2020; Zhang et al., 2020; Calvet et al., 2021; Fritz et al., 2020; Hamer et al., 2021; Villanueva-Saz et al., 2022). Of all animals with documented SARS2 susceptibility, humans are most likely to be in direct and intimate contact with the domestic cats living in 25% of households in the United States (A.V.M.A, 2017). Methods for documenting cat-to-human transmission have been described and will require intentional and continued research (Totton, Sargeant, & O'Connor, 2021). Epidemiological and next-generation sequencing methods to assess cat-to-human transmissions are well within the capabilities of researchers globally, and the continued prevalence of SARS2 points to the urgency for ongoing monitoring of domestic cats and other

animals in close contact with humans for evidence of SARS2 variant emergence and reservoir persistence.

3.4 Methods

3.4.1 Experimental exposures and sample collection

Directly inoculated cats in this study were exposed to viral stocks derived from an original SARS2 isolate obtained from the Biodefense and Emerging Infections Research Resources Repository (`https://www.beiresources.org/`). The USA-WA1/2020 isolate (Genbank MN985325) was expanded in Vero cells over three passages (P1, P2, P3) as previously described (Bashor et al., 2021). Three cohorts of adult (5-8 years) and juvenile (12 weeks) domestic cats (N=23) were intranasally inoculated with P2 or P3 virus stock in Animal Biosafety Level 3 (ABSL-3) facilities at Colorado State University. Cat cohorts are depicted in Figure 3.1 and additional relevant metadata are provided in Table B.1. Cohort A was part of a study to assess domestic cat susceptibility to SARS2 with previously reported results (A. M. Bosco-Lauth et al., 2020). Cats 1 and 5 were inoculated with a "med-high" dose of $1x10^6$ plaque-forming units (pfu) of P3 virus stock, and Cat 6 was cohoused with Cat 5 one day post-infection (dpi) to facilitate contact transmission. For both direct- and contact-exposed cats, dpi refers to the number of days following exposure of the directly inoculated cats in the cohort. Infection outcomes, including viral titer at 1 dpi (Cats 1 and 5) and 7 dpi (Cat 6), virus neutralization, seroconversion, and cat-to-cat transmission were assessed as previously described (A. M. Bosco-Lauth et al., 2020).

Sequencing and within-host variant analysis for Cats 1, 5 and 6 has been previously reported (Bashor et al., 2021). Data were reanalyzed in this study in light of the increased information available from two additional cohorts of experimentally exposed cats. Cohort B was inoculated to assess age related susceptibility to SARS2 (Bosco-Lauth, unpublished data). Cats 7-9 were inoculated with a "med-high" dose of $1x10^6$ pfu of P2 virus stock and were cohoused with three naïve cats (Cats 10-12) at 1 dpi to facilitate contact transmission. Viral titer at 3 dpi (Cats 7-9) and 4 dpi (Cats 10-12) was assessed for all six Cohort B cats. Cohort C was inoculated as part of a dose-response study SARS2 (Bosco-Lauth, unpublished data). Cats 13-25 were inoculated with

three doses of P3 virus stock: low (Cats 13 and 14; $1.05x10^4$ pfu), medium (Cats 15-21; $4.65x10^5$ pfu) or high (Cats 22-25; $2.1x10^7$ pfu). Cat 26 was cohoused with Cats 22-25 at 1 dpi to facilitate contact transmission. Viral titer was assessed at 3 dpi for Cats 13-25, and 4 dpi for Cat 26. Within-host variant analysis for Cohort B and C animals was conducted opportunistically following viral recovery in order to learn more about SARS2 adaptation in domestic cats.

In all studies, infections were performed as described previously(A. M. Bosco-Lauth et al., 2020). Briefly, animals were lightly anesthetized prior to inoculation via the nares. Nasal lavage and oral swab samples were collected 1-3 days post-exposure. Nasal lavages were performed through the dropwise instillation of 1 mL BA-1 into the nares of awake or lightly anesthetized cats. Discharge was collected in a sterile petri dish. Oral swabs were obtained and stored in BA-1 medium (Tris-buffered Minimum Essential Medium with 1% Bovine Serum Albumin) supplemented with gentamicin, amphotericin B and penicillin/streptomycin. Viral titers (measured in pfu/mL) were assessed from nasal lavage samples from each individual.

3.4.2 RNA isolation and sequencing

Nasal lavage samples used to assess viral titer (collected at 1, 3, 4 or 7 dpi as indicated above) were then subjected to RNA extraction, complementary DNA generation and library preparation performed as previously described (Bashor et al., 2021). Briefly, nasal lavage fluid was inactivated in Trizol (100 μL fluid in 900 μL Trizol) and frozen at -80°C prior to RNA extraction. RNA was isolated using a modified Zymo RNA Clean and Concentrator 5 kit (Zymo Research) with DNAse I (New England Biolabs). Complementary DNA (cDNA) was generated with 10 μL RNA and SuperScript II RT (Thermo Fisher Scientific) as previously described and frozen at -20°C prior to tiled amplicon enrichment and library preparation.

Samples were enriched for SARS2 genomic material using a PCR protocol designed by the ARTIC Network (ARTICNetwork, 2020). This method employs a pooled primer scheme that generates 400bp overlapping ("tiled") amplicons that span the full viral genome. All samples were amplified in technical duplicates with ARTIC version 3 (V3) primers except for Cat 5 which was amplified singly with the previous ARTIC version (V2) primers. ARTIC primer sequences are

publicly available here: https://github.com/artic-network/artic-ncov2019/tree/master/primer_schemes/nCoV-2019. PCR products were visualized on a 1% agarose gel and quantified with a Qubit dsDNA Broad Range Assay kit (Thermo Fisher Scientific). Sequencing libraries were prepared with 540 ng input DNA per sample using a NEBNext Ultra II DNA Library Prep kit and replicate samples were individually barcoded with NEBNext Multiplex Oligos for Illumina (New England Biolabs). Ampure XP beads were used for cleanup and size selection (Beckman Coulter), and a single 0.65X size selection was carried out following adapter ligation. Libraries were pooled for sequencing on an Illumina MiSeq instrument (Illumina) with v2 500 cycle 2 x 250 bp kits. All libraries were sequenced over the course of three sequencing runs on the same MiSeq instrument at the Colorado State University Next Generation Sequencing Facility. Cat 5 in Cohort A and its technical replicate were both sequenced on the same run as part of a pilot project in May 2020. For the three viral stock inoculum Passages (P1, P2 and P3), all cats in Cohort C, and Cats 1 and 6 in Cohort A, one technical replicate was sequenced in a run in September 2020 and the second technical replicate was sequenced as part of a second run in October 2020. Both technical replicates for all six cats in Cohort B were sequenced on the October 2020 run.

3.4.3 Bioinformatics & variant calling

Raw sequencing data in fastq file format were analyzed with a custom Nextflow (Di Tommaso et al., 2017) bioinformatics pipeline designed to call single nucleotide and structural variants in viral populations. The pipeline is publicly available on Github at https://github.com/stenglein-lab/viral_variant_caller. Sequencing reads were trimmed for quality and adapters with Cutadapt (Martin, 2011), then aligned to the USA-WA1/2020 reference sequence (GenBank MN985325). Data were pre-processed with GATK (der Auwera et al., 2013) followed by LoFreq (Wilm et al., 2012b) variant calling at a minimum depth of coverage of 40x and a minimum allele frequency of 0.1% (variants were later filtered to $\geq$3% frequency for further analysis). Variants were annotated and functional effects were predicted with SnpEff and SnpSift

(Cingolani, Platts, et al., 2012; Cingolani, Patel, et al., 2012). Pipeline output was further analyzed and visualized in R statistical software (Team, 2021).

We took the additional approach of running iVar (Grubaugh et al., 2019), a bioinformatics pipeline designed for within-host viral variant calling from amplicon sequencing data to validate our pipeline. Any variants that were not also detected in iVar were filtered out unless they were identified to be an artifact of the 3% allele frequency cutoff, which iVar applies before averaging technical duplicates and viral_variant_caller applies after. Eleven variants detected in cats were filtered out at this step. Nonsynonymous and synonymous nucleotide diversity were calculated with the SNPGenie pipeline (Nelson & Hughes, 2015). Nucleotide diversity (π) is calculated as the mean number of pairwise differences per nonsynonymous or synonymous site in the genome for a population of sequences. These estimates are weighted by allele frequencies. This analysis was done at both the population level (the population of full viral genomes within a host) and by gene product.

3.4.4 Phylogenetic analysis

Consensus genome sequences used for phylogenetic reconstruction of felid-derived SARS2 were obtained from GISAID (Shu & McCauley, 2017). Sequence identifiers, metadata and acknowledgments are listed in Tables S2 and S3. All complete SARS2 genome sequences from host species belonging to the family Felidae available on GISAID as of April 4, 2022 were downloaded. Species included cat (*Felis catus*), tiger (*Panthera tigris*), lion (*Panthera leo)*, snow leopard (*Panthera uncia*), cougar (*Puma concolor*), fishing cat (*Prionailurus viverrinus*) and leopard cat (*Prionailurus bengalensis*). Sequences that were incomplete, low coverage (>5% Ns) or passaged through cell culture were removed prior to analysis, which included the only cougar and fishing cat sequences. For each felid-derived sequence, a random sample of up to 10 human-derived SARS2 genome sequences were also obtained from the full GISAID EpiCoV database. These human-derived sequences were subsampled in the same country and division during a two-week period surrounding the collection date of each felid sequence. Subsampling employed metadata available from GISAID and the Nextstrain (Hadfield et al., 2018) Augur filter subcommand. Maxi-

mum likelihood time-resolved phylogenies were reconstructed and visualized from these sequence data using open source Nextstrain bioinformatic tools. Briefly, sequences and metadata were input into the Augur pipeline, where they were filtered, aligned to the Wuhan-Hu-1 reference (Genbank NC_045512), and a tree was inferred with TreeTime. Results were visualized with Auspice software (Figure 3.6, Figure S5). Publicly available interactive versions of these custom Nextstrain builds are hosted online. To query the presence and location of specific consensus variants throughout the phylogeny, we used the interactive tools provided by Nextstrain to filter the displayed data by genotype.

3.5 Funding

The research reported in this publication was supported by the Colorado State University College of Veterinary Medicine and Biomedical Sciences Research Council Award and Colorado State University's Office of the Vice President for Research's "Accelerating Innovations in Pandemic Disease" initiative, made possible through support from The Anschutz Foundation. The content is solely the responsibility of the authors. Computational resources were supported by NIH/National Center for Advancing Translational Science Colorado Clinical and Translational Science Awards grant UL1 TR002535.

3.6 Data and code availability

All SARS2 raw next-generation sequencing data used in this study are publicly available in the NCBI SRA database under BioProject PRJNA704947 (previously published sequencing data for P1, P2 and P3 of the viral stock and Cats 1, 5 and 6)32 and BioProject PRJNA842572 (all remaining cats). The bioinformatics pipeline used to analyze the data is publicly available at `https://github.com/stenglein-lab/viral_variant_caller`. Additional data including complete variant summary tables output by the pipeline listing all within-host variants detected within samples from cats and the P1, P2, and P3 stocks used for inoculations, and R scripts for data processing, statistical analysis and visualization are available at `https://github.com/laurabashor/SARS2felids`. Custom Nextstrain builds for publicly avail-

able sequence data from domestic cats and felids are hosted at: https://nextstrain.org/community/laurabashor/SARS2felids/catshumans and https://nextstrain.org/community/laurabashor/SARS2felids/felidshumans.

Chapter 4

SARS-CoV-2 within-host population expansion, diversification and adaptation in zoo tigers, lions and hyenas[3]

SARS-CoV-2 rapidly adapts to a new host environment following transmission from one species to another. This is notable, as many non-human animal species are susceptible to SARS-CoV-2, and the emergence of novel within-host viral variants has been observed in both domestic and wild species. Furthermore, transmission back to humans (spillback) has been reported from white-tailed deer, mink, hamsters, domestic cats, and lions. We examined transmission and evidence of host-specific adaptation following an outbreak of SARS-CoV-2 in Amur tigers (*Panthera tigris altaica*), African lions (*Panthera leo*), and spotted hyenas (*Crocuta crocuta*) at the Denver Zoo between October and December 2021. Longitudinal nasal swabs from 16 individuals tested positive for SARS-CoV-2 viral RNA multiple times during a two-month sampling period. Full SARS-CoV-2 genomes were sequenced to identify variant lineage, within-host variation and genomic signatures of selection. Sequence analysis indicated that the outbreak was initiated by a single spillover of a rare SARS-CoV-2 Delta sublineage, AY.20, concurrently circulating in humans at very low levels, that was subsequently transmitted from tigers to lions to hyenas. Within-host virus populations rapidly expanded and diversified over time, and signatures of both purifying and positive selection were evident. Strong positive selection was detected across the SARS-CoV-2 genome in samples from hyenas and in the nucleocapsid (N) gene in all animals. Four candidate species-specific adaptive mutations were identified: one in lions and hyenas (N A254V) and three in hyenas (open reading frame (ORF)1a E1724D, spike (S) T274I, and N P326L). This study describes cross-species transmission dynamics and within-host evolution of SARS-CoV-2 in three species in a natural setting, revealing viral adaptation following host shifts.

[3]This chapter was co-authored by Bashor, L., Gallichotte, E., Stenglein, M., Galvan, M., Erbeck, K., Croft, L., Stache, K., Johnson, J., Pabilonia, K., & VandeWoude, S.

4.1 Introduction

The evolution of SARS-CoV-2, and the emergence of successive new variant lineages, moved to the forefront of global public health awareness during the COVID-19 pandemic. In addition to humans, the virus infects a wide range of animal species (Barrs et al., 2020; Segalés et al., 2020; Hale et al., 2021; Munnink et al., 2021; Shi et al., 2020; APHIS, 2024a; WOAH, 2024). It is widely accepted that SARS-CoV-2 spilled over from an animal host into human populations in late 2019 (Andersen et al., 2020). Although thousands of studies have been conducted since the start of this pandemic, key questions remain unanswered - including identifying the drivers of SARS-CoV-2 evolution and transmission dynamics in non-human animals, and the origin of variant lineages that infect human and animal populations.

In March 2020, SARS-CoV-2 sequence analysis revealed human-to-animal transmission in five tigers and three lions at the Bronx Zoo, making nondomestic felids among the first documented non-human animals infected through viral spillover (Bartlett et al., 2021; McAloose et al., 2020). There have since been many other cases in felids, including reports of potential cat-to-human and lion-to-human spillback (Koeppel et al., 2022; Fernández-Bellon et al., 2021; L. Wang et al., 2022; Nagy et al., 2022; Karikalan et al., 2021; Mishra et al., 2021; Siegrist et al., 2023; Sila et al., 2022). Reported infections most likely represent a fraction of the true number of infections in these animals. It is therefore important to study SARS-CoV-2 cross-species transmission in zoos to promote animal and human health, and to understand risk factors and potential for variant emergence in these settings.

SARS-CoV-2 within-host evolution is shaped by both deterministic effects of selection and the stochastic effects of genetic drift. At this level, selection acts to maximize the fitness of the population of viruses, often referred to as a quasispecies, within an individual host. Accordingly, viruses experience novel and strong selective pressures following host shifts. To evaluate this phenomenon, we investigated SARS-CoV-2 host adaptation following a unique spillover event with an unprecedented sample set: an outbreak in large cats and hyenas living at the Denver Zoo in Colorado, USA.

Nasal swab samples from two Amur tigers (*Panthera tigris altaica*), eleven African lions (*Panthera leo*), and two spotted hyenas (*Crocuta crocuta*) tested positive by RT-PCR for SARS-CoV-2 between October and December 2021, with some lions testing positive throughout the two-month period. Two additional hyenas were tested multiple times during this period, with inconclusive results. Infectious virus was isolated from a subset of samples, and robust neutralizing antibodies were detected in all serum samples tested. We evaluated SARS-CoV-2 genomes recovered from infected animals over a period of approximately three months for variant lineage, phylogenomic epidemiology, within-host variation and genomic signatures of selection. We documented a significant impact of host shifts on virus within-host evolution.

4.2 Results

4.2.1 SARS-CoV-2 RNA was detected in tigers, lions, and hyenas, and most persistently in lions

On October 7th, 2021, both tigers (Tigers A and B) were observed with coughing and nasal discharge. Nasal swab samples were collected and transferred to the Colorado State University (CSU) Veterinary Diagnostic Laboratories where they tested positive for SARS-CoV-2 viral RNA (vRNA) by RT-PCR (Figure 4.1). Both tigers again tested RT-PCR positive on October 13th, but were negative thereafter. On October 14th, after several lions exhibited similar respiratory signs, all eleven lions in the collection also tested positive by RT-PCR. All lions tested positive at least four times over the next 15 days. After the initial positive tests in lions, swabs were collected regularly for approximately three months (Figure 4.1). The lions in pride 1 (Lions A-D, all aged 6 years old) were positive over the longest interval, with some individuals testing positive for 60 days (Lion B). A lion cub in pride 2 (Lion I) also tested positive intermittently over a 46-day period (Figure 4.1).

Mild clinical signs were observed in two cohabitating hyenas (Hyenas A and B), on October 28th, and samples collected that day were positive for SARS-CoV-2 by RT-PCR, constituting the first and only documented cases of SARS-CoV-2 infection in *C. Crocuta*. Hyenas A and B were exhibited separately from the lions but rotated through the same areas. Two younger hyenas were

kept separately (clan 2, Hyenas C and D). These animals rotated through the same areas as Hyenas A and B, but did not interact directly with the older individuals. Swabs from Hyenas C and D were occasionally inconclusive through December, but never tested conclusively RT-PCR positive (Figure 4.1). Detailed descriptions of all infections and complete datasets of RT-PCR cycle threshold (Ct) values have been previously described.

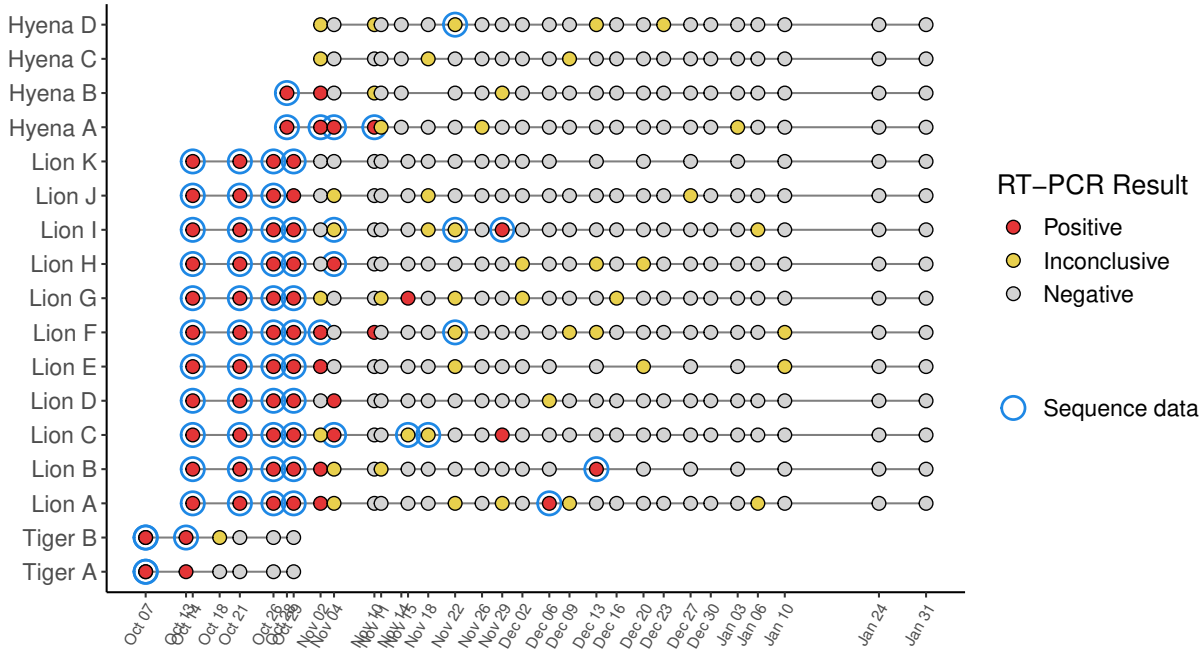


Figure 4.1: SARS-CoV-2 RNA was persistently detected by RT-PCR in nasal swabs from tigers (N=2), lions (N=11) and hyenas (N=4) between October 2021 and January 2022, and next-generation sequencing (NGS) datasets were generated from 63 of these samples. Each point represents a nasal swab sample tested for three SARS-CoV-2 gene targets, colored by result determined by diagnostic software. Points with a blue border indicate samples from which high-quality NGS data was generated.

4.2.2 High quality full genome SARS-CoV-2 sequences were obtained from positive and inconclusive samples from zoo tigers, lions and hyenas

SARS-CoV-2 vRNA recovered from nasal swab samples was sequenced in technical duplicate to identify SARS-CoV-2 lineage, within-host variation and genomic signatures of selection. SARS-CoV-2 sequencing reads were aligned to the Wuhan-1 reference sequence (GenBank

MN908947.3). High-quality SARS-CoV-2 genome sequencing datasets derived from 63 biological samples were obtained from two tigers, eleven lions, and three hyenas at multiple timepoints between October 7th and December 13th, 2021 (Figure 4.1). This corresponded to one or more timepoints for each tiger and hyena, and three to seven timepoints for each lion. Although the majority of NGS datasets were obtained from positive swab samples, the 50 RT-PCR-inconclusive samples were screened for potential NGS, and six yielded high-quality sequencing data (Figure 4.1).

4.2.3 SARS-CoV-2 Delta variant sublineage AY.20 was transmitted from tigers to lions to hyenas

To obtain a focused phylogenomic epidemiology of this outbreak, a time-based tree was generated using sequence data from zoo animals and humans in Colorado September 23rd and November 4th, 2021 (Figure 4.2A). All sequences from animals were classified as the same Delta variant sublineage, Pangolin lineage AY.20, and generally clustered by species. Virus sequences from lions and hyenas were more closely related to each other than to those from tigers, and sequences from the two prides of lions were intermixed throughout the phylogeny. The AY.20 sublineage was circulating at low rates (1%) in humans globally between May 2021 and February 2022. During the period of the zoo outbreak, only 461 out of the 79,455 (0.58%) SARS-CoV-2 sequences in the GISAID database from Colorado (between October and December 2021) were classified as AY.2019. Excepting twelve sequences from the Denver Zoo animals generated by USDA APHIS following the outbreak, only three other AY.20 sequences recovered from non-human animals have been submitted to the GISAID database: two from white-tailed deer (collected in New York and North Carolina in November 2021; EPI_ISL_13610631, EPI_ISL_16464980), and one from a domestic cat (collected in California in December 2021; EPI_ISL_10656104). To examine the potential direction of transmission among animals, we generated a haplotype network (Figure 4.2B). The species-specific clusters evident in this network are consistent with the trajectory of observed disease onset, supporting the conclusion that SARS-CoV-2 originated in tigers, then moved to lions, then finally to hyenas. We inferred a second phylogeny from all publicly available felid- and

hyena- derived SARS-CoV-2 sequences in addition to the sequences reported here (Figure S1). Importantly, we saw no evidence for felid-driven viral adaptation occurring at the global scale.

4.2.4 SARS-CoV-2 sequence variation

We identified 63 unique sequence variants relative to the Wuhan-1 reference sequence (Table C.1). Thirty of these were defining mutations of the Pangolin AY.20 lineage, of which the majority (n=27) are missense, or nonsynonymous single nucleotide variants (SNVs) (Table C.2). The remaining 33 non-AY.20 variants were SNVs distributed throughout the virus genome, including synonymous (n=11), nonsynonymous (n=12), and noncoding (n=10) mutations. Sub-consensus sequence variants (present at <50% allele frequency) were identified very rarely, making up 16 (0.58%) of the 2769 observations of these 63 variants. Variants at <99% allele frequency were somewhat more common, making up 4.2% (n = 116).

Sequences clustered due to mutations shared by more than one individual of the same species (Figure 4.2, Table C.1). The A254V mutation in the nucleocapsid (N) gene differentiated all hyena and lion sequences from the three tiger-derived sequences and was supported by robust sequencing coverage at that genomic position across the board (Figure 4.2C). Furthermore, all hyena sequences except Hyena D had the mutations P326L in N, T274I in the spike (S) gene, and E1724D in open reading frame (ORF) 1a (nonstructural protein (nsp) 3 E906D), mutations not detected in lions or tigers (Figure 4.2C).

Interestingly, the sequence obtained from Hyena D differed from Hyenas A and B, and was more related to sequences recovered from lions. Hyenas C and D were cohoused separately from the significantly older hyenas A and B, and the younger pair did not test positive during the period in which samples were collected. The sequence for Hyena D was therefore generated from an RT-PCR inconclusive sample. The first SARS-CoV-2 sequences from Hyenas A and B derived from swabs collected on October 28th were identical at the consensus level. Sequences obtained from three subsequent Hyena A samples share the ORF1b mutation G2068E (nsp15 endoribonuclease G17E). Investigation of within-host variation reveals that this mutation was also present at a sub-

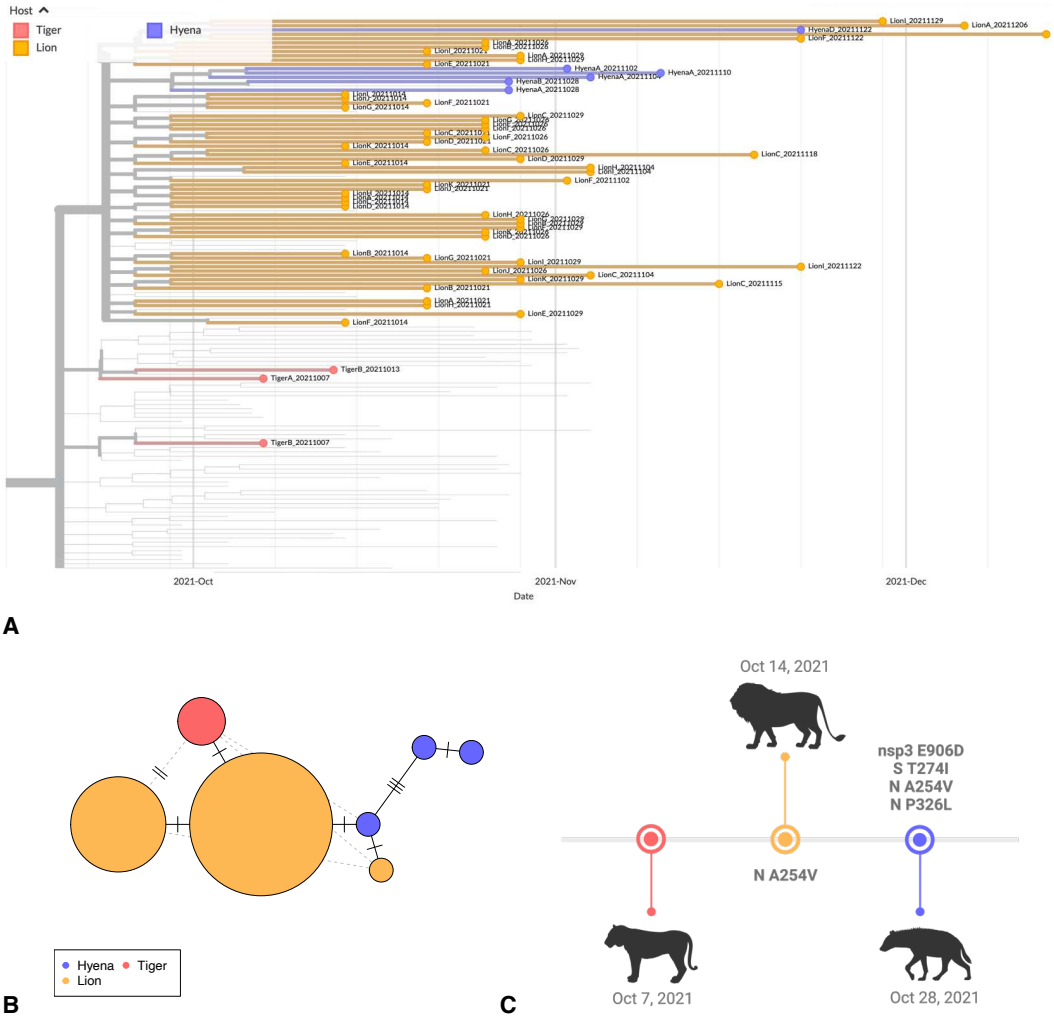


Figure 4.2: A single introduction of SARS-CoV-2 lineage AY.20 was transmitted from tigers to lions to hyenas. (**A**) Time-based phylogenetic tree of SARS-CoV-2 sequences from tigers, lions and hyenas, with collection date indicated along the x-axis. Tree was generated with Nextstrain and an interactive version is available at `https://nextstrain.org/community/laurabashor/DZSARS2`. Sequences are colored by host species and scaled by time. Three data inputs: (1) sixty-three SARS-CoV-2 consensus sequences generated from the zoo outbreak, (2) all human-derived sequences classified as AY.20 in the state of Colorado between September 23rd and November 4th, 2021 (N = 198 sequences), and (3) a random subsample of human-derived sequences in Colorado from the days leading up to the zoo outbreak (100 sequences/day from September 23rd to October 7th, 2021; N=1500 sequences). Sequence data from humans were obtained from the GISAID database, and are listed in GISAID EPI_SET_240407tz. (**B**) A haplotype network of the latest high-quality sequence obtained from each individual (N=16 sequences) is consistent with the trajectory of cases in the outbreak starting with tigers before transmitting to lions and hyenas. Haplotype network was generated using the 'pegas' package in R. (**C**) One mutation detected in SARS-CoV-2 genomes from lions and hyenas (N A254V) and three mutations detected in hyenas (N P326L, S T274I, and ORF1a E1724D) represent candidate species-specific adaptations.

consensus allele frequency in Hyena A on October 28th, but was not observed in the other two hyenas (Figure 4.3D).

Lions in pride 2 had more viral variation compared to pride 1, and Lions H and I in particular shared multiple mutations in their samples collected on November 4th. Unique mutations were also detected in a handful of lion samples which changed in allele frequency over time, and differentiated sequences from each other. For example, ORF1a P1497S (nsp3 P679S) was detected above consensus-level in Lion F on October 14th, sub-consensus-level October 21st, and then not observed in any later samples. Similarly, ORF3a V50I appears initially in Lion E on October 14th at relatively high allele frequency, then gradually decreased in frequency in subsequent samples, and was not observed in the final sample from Lion E on October 29th.

4.2.5 Within-host SARS-CoV-2 variation revealed species-specific adaptation in lions and hyenas

Based on the conclusion that the two tigers were the first to be infected, we interrogated our NGS datasets for within-host variation relative to the consensus sequence of the first SARS-CoV-2 samples recovered from tigers ("tiger reference sequence"). Fourteen unique sequence variants remained following this approach (Figure 4.3C, D). This within-host variation was dominated by missense, or nonsynonymous, SNVs, constituting nine (64%) of fourteen unique variants, and 87 (93%) of 94 observations of these variants. The nine missense SNVs were located in ORF1a (n = 2 in nsp3 papain-like protease), ORF1b (n = 2 in nsp15, endoribonuclease), S (n = 2), ORF3a (n = 1) and N (n = 2). The two synonymous SNVs were both in ORF1a (nsp1 leader protein and nsp3), and two noncoding variants located at the same nucleotide position in the 5' UTR. Sub-consensus variants made up a larger proportion of sequence variation relative to the tiger consensus sequence as compared to the Wuhan-1 reference. Of the 94 variant observations, 14 (14.9%) were detected at <99%, and five (5.3%) of were observed at <50% allele frequency (Figure 4.3D).

Hyenas had more within-host variants overall (Figure 4.3A, D). In general, within-host variation remained consistent within an individual across repeated samples (Figure 4.3D). All samples obtained from Lions B, C, D, J and Hyena D had only one within-host variant relative to the tiger

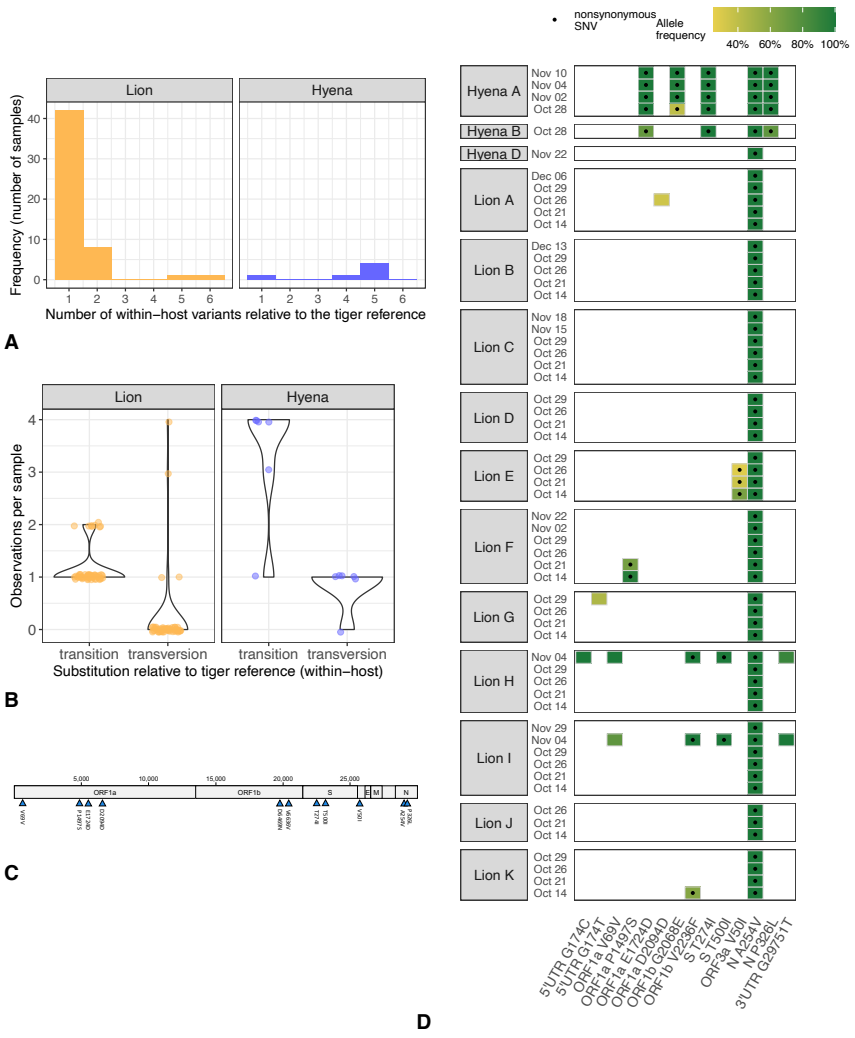


Figure 4.3: Species-specific patterns of SARS-CoV-2 within-host variation were observed in lions and hyenas. Only sequence variation relative to the consensus sequence of the two infected tigers is shown (N=58 samples). (**A**) One within-host variant (N A254V) was detected in the majority of samples. Plot shows the frequency distribution of the number of within-host variants detected per sample. (**B**) Transition substitutions were more common than transversions, particularly in hyenas. (**C**) Within-host variants observed in lions and hyenas relative to the tiger reference sequence were detected throughout the SARS-CoV-2 genome. The location of each variant located in SARS-CoV-2 coding regions is indicated with a triangle. The three variants in the 5' and 3' UTRs are not shown. (**D**) Fourteen within-host variants were identified in lions and hyenas but not tigers. Each row represents one nasal swab sample, and sample collection date and animal identifier are indicated on the y-axis. Each tile represents a mutation detected in that sample relative to the consensus sequence of SARS-CoV-2 genomes recovered from infected tigers and is colored by allele frequency. Black dots indicate nonsynonymous single nucleotide variants. A SARS-CoV-2 genome schematic at the bottom indicates the genomic loci of the eleven mutations. Code to prepare plot (**D**) was adapted from the R package 'outbreakinfo'.

reference sequence, N A254V, that was shared across every sample from lions and hyenas (Figure 4.3D). There was more within-host variation in samples from hyenas compared to lions (lion mean = 1.33, median = 1 within-host variant per sample; hyena mean = 4.17, median = 5). Interestingly, the majority of substitutions observed in our dataset were transitions across all three species (Figure 4.3B). This pattern held whether considering variants relative to the Wuhan-1 reference sequence or the tiger reference sequence, and was more distinct in hyena samples.

4.2.6 Within-host virus populations were characterized by expansion, diversification, and selection

To characterize selective pressures on within-host SARS-CoV-2 populations, we calculated nucleotide diversity (π) across the virus genome at nonsynonymous and synonymous sites (Tables C.5 and C.6). Although virus populations within each individual animal followed their own trajectories, we observed increasing within-host nucleotide diversity over time (linear mixed model; Figure 4.4A, Table 4.1). We measured time as the number of days since an individual animal's first positive test and accounted for repeated measures of the same individuals. Our model revealed that host species also had a small but significant effect on within-host nucleotide diversity. Synonymous nucleotide diversity also increased over time, but this metric was not impacted by species (linear mixed model; Figure 4.4B, Table 4.1). In other words, we observed SARS-CoV-2 within-host populations expand and diversify over time, and we found that host species had an impact on virus genetic diversity, but not effective population size.

There were genomic signatures of positive selection (πN > πS) in 49% (31 of 63) of virus samples had, whereas 51% (32 of 63) had genomic signatures of purifying selection (πS > πN). Positive selection at the genome level was observed in all six samples from hyenas (Hyenas A, B and D). Differences in the strength of selection acting on SARS-CoV-2 genomes among host species was particularly evident in the final virus samples obtained from each individual (Figure 4.5A). When nucleotide diversity was calculated separately for each SARS-CoV-2 gene product, interesting patterns in nonsynonymous and synonymous nucleotide diversity were noted in ORF1ab, S and N genes. Strong positive selection was detected in the N gene in most samples. Many samples also

Table 4.1: Summary of linear mixed models of nucleotide diversity (π) as a function of time and host species. Time was measured as days since first positive RT-PCR test. Time and species were fixed effects, with animal ID as a random effect to account for repeated measures of the same individual over time. The model was fit with the lmer() function in the 'lme4' package in R.

	Nucleotide diversity		**Synonymous nucleotide diversity**	
Predictors	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.000082 **	**0.001**	0.000050 *	**0.049**
Time (days after first positive test)	0.000002 ***	**<0.001**	0.000003 ***	**<0.001**
Species [lion]	-0.000051 *	**0.048**	-0.000020	0.465
Species [tiger]	0.000012	0.771	0.000045	0.298
Random Effects				
σ^2	3.2 x 10^{-9}		3.6 x 10^{-9}	
τ_{00}	3.1 x 10^{-11} animal ID		0.00 animal ID	
ICC	0.0096			
N	16 animal ID		16 animal ID	
Observations	63		63	
Marginal R^2 / Conditional R^2	0.225 / 0.233		0.229 / NA	

** p<0.05 ** p<0.01 *** p<0.001*

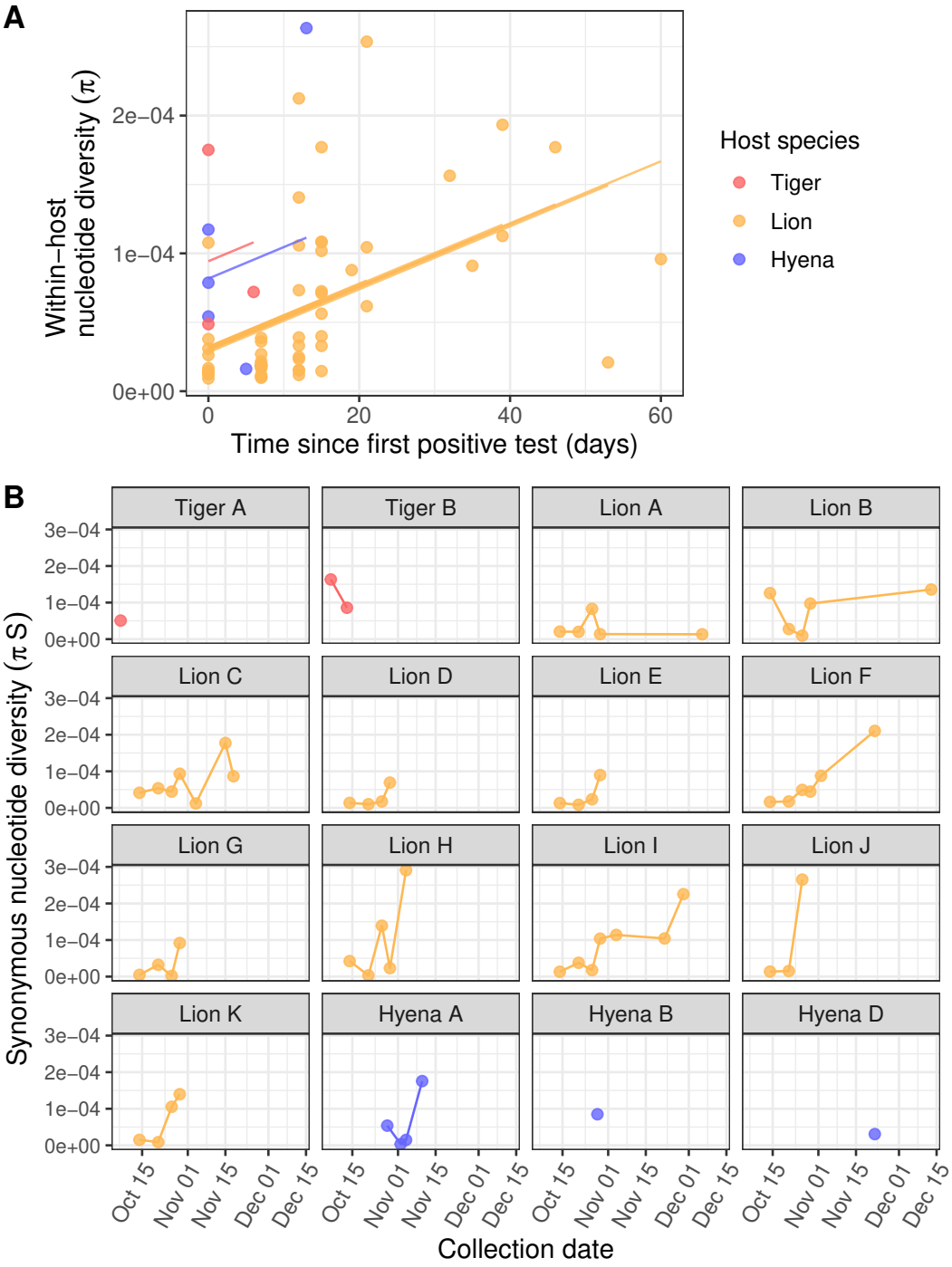


Figure 4.4: SARS-CoV-2 within-host populations in infected zoo animals underwent (**A**) expansion and (**B**) diversification over time. Each point indicates the nucleotide diversity of a SARS-CoV-2 virus population observed in one nasal swab sample, calculated as the mean of two sequencing replicates. (**A**) Within-host nucleotide diversity (π) increased over time as measured by the number of days after an individual's first positive SARS-CoV-2 test. Lines show the predicted fit of a linear mixed model by species. (**B**) Increasing within-host synonymous nucleotide diversity (πS) over time is evident in most lions.

had signs of positive selection in the S gene, whereas both positive and purifying selection were detected in ORF1ab (Figure 4.5B).

4.3 Discussion

Longitudinal nasal swab samples initially collected for diagnostic purposes were subjected to next-generation sequencing to generate a unique dataset for detailed analysis of virus within-host evolution following cross-species transmission. We conclude that a single introduction of a rare SARS-CoV-2 Delta sublineage spread from tigers to lions to hyenas. Repeated longitudinal sampling of individual animals provided us with snapshots of virus population expansion and diversification within hosts. Signatures of both purifying and positive selection were evident across the virus genome, and four potentially species-specific adaptive mutations were identified in lions and hyenas. While no concerning SARS-CoV-2 variant lineage evolved in infected animals, uniquely strong signatures of positive selection detected in the N gene and in samples from hyenas highlight the combined impacts of mutation and selection following virus host species shifts.

Sample collection was initiated immediately after individual animals displayed clinical signs. Tigers were the first species to display respiratory symptoms, followed by lions and hyenas. Nasal swabs were continuously collected until no individuals of a species remained positive. This symptom-driven testing scheme introduced some selection bias into the dataset, as it is not possible to identify the timeline of exposure or true duration of each infection. Positive RT-PCR tests for certain lions span more than a month, whereas positive tests for tigers and hyenas span 1–2-week periods. Despite this limitation, similarly high levels of neutralizing antibodies were detected in multiple individuals, including in Hyena B, for whom testing spanned only four days.

All nasal swabs that were either RT-PCR positive or inconclusive were subjected to NGS. Although the majority of sequencing datasets were obtained from RT-PCR positive samples, six inconclusive samples yielded high-quality NGS datasets. Viral RNA present in these samples may or may not correspond to the presence of infectious virus; however, infectious virus isolated from swabs from two lions and Hyena A over a range of time points during the outbreak (including a sample from Lion A on December 6th) confirmed active infection among these animals.

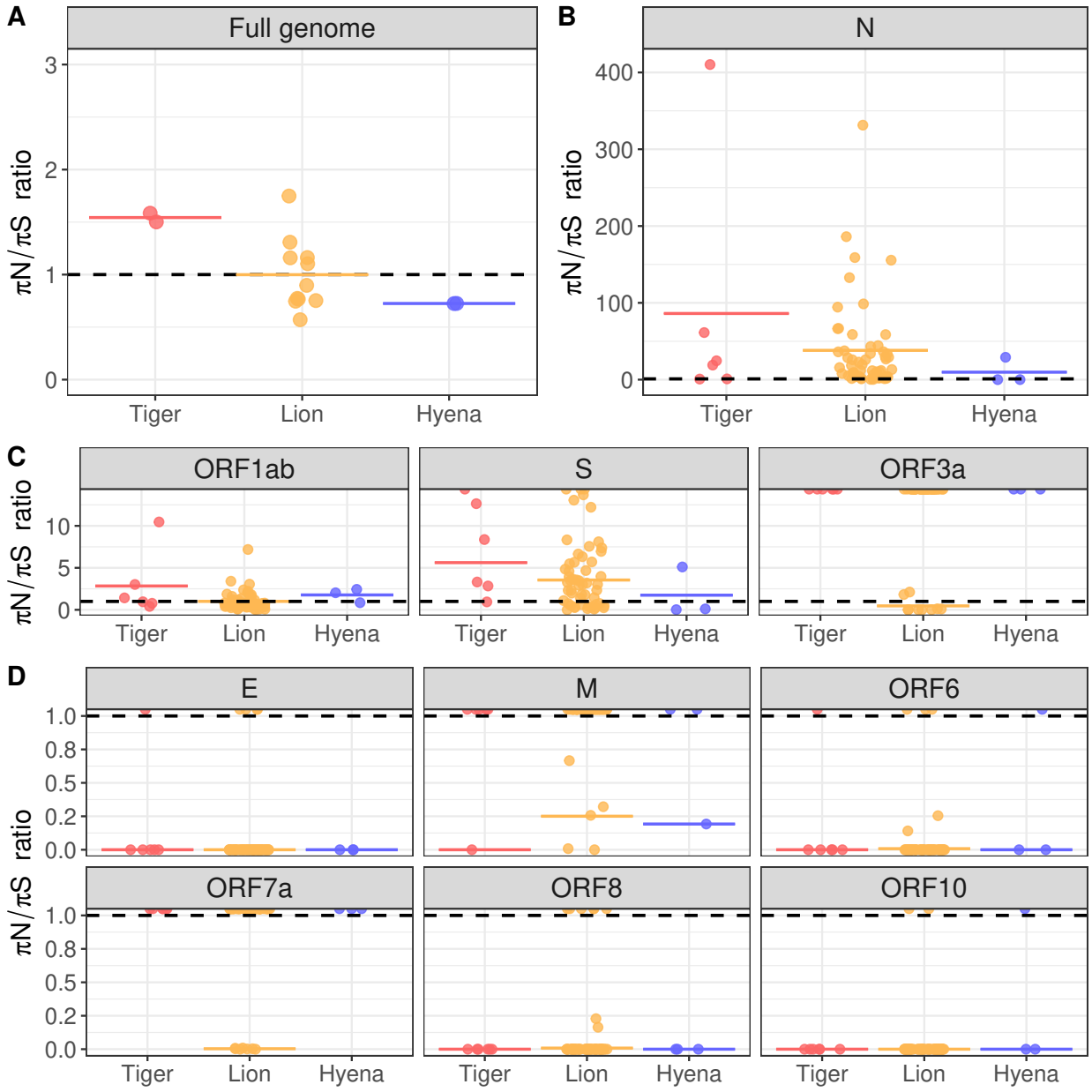


Figure 4.5: Strong signatures of positive selection were observed, most notably in full SARS-CoV-2 genomes recovered from (**A**) infected hyenas and (**B**) the nucleocapsid (N) gene. The remaining gene segments showed (**C**) mixed signatures of positive and purifying selection, and (**D**) no signatures or exclusively purifying signatures of selection. The ratio of nonsynonymous (πN) to synonymous (πS) nucleotide diversity was calculated as a measure of the type and strength of selection acting on SARS-CoV-2 virus populations in zoo animals. Black dashed lines indicate where the πN/πS ratio = 1. (**A**) Each point represents the πN/πS ratio of a SARS-CoV-2 virus population observed in the last nasal swab sample collected for each individual (n = 16 animals), calculated as the mean of two technical sequencing replicates. (**B, C, D**) Each point represents the πN/πS ratio of a SARS-CoV-2 gene segment observed in one nasal swab sample, calculated as the mean of two technical sequencing replicates. Infinite values (where πS = 0) are indicated with half points along plot borders. The abbreviated gene segment is indicated at the top of each plot (nucleocapsid (N), open reading frame (ORF), spike (S), envelope (E), membrane (M).

The consensus sequence of every virus sample in this outbreak was classified as the same Delta lineage AY.20, supporting the hypothesis that the zoo outbreak was initiated by a single spillover event, presumably from an infected human. The rarity of this lineage among humans around this time further supports the single introduction hypothesis. Investigation of other SARS-CoV-2 outbreaks that have occurred at zoos indicates that zoo animals have been infected by variant lineages both rare and dominant in humans. It is possible that AY.20 is more transmissible in non-human animals (e.g., felids), than other lineages, contributing to this multi-species outbreak.

This outbreak occurred during one of the peaks of the COVID-19 pandemic, but only approximately 1% of human infections occurring in Colorado at this time were caused by the AY.20 sublineage. If the zoo animals were exposed to one of these rare human cases, the defining AY.20 mutations would already be present in the original virus that spilled over. However, we cannot rule out that infection may have originated in another captive or feral species present at the zoo. Free-ranging peridomestic animals are present in most outdoor zoo environments, and susceptible species like deer mice and domestic cats could potentially have moved through large felid habitat.

During the period of positivity in lions, one animal care specialist wearing full PPE reported nasal discharge from a lion directed towards their face (unpublished direct communication). This individual tested positive for COVID-19 within one week of this incident and communications from zoo staff present during this outbreak suggest that the virus was transmitted to a human from an infected lion. A sample was collected from this individual at this time and sequenced by the Colorado Department of Public Health and the Environment (CDPHE). There are 17 SARS-CoV-2 sequences recovered from humans in the GISAID database that are classified as AY.20 and have the N A254V mutation that we identified in lions and hyenas, but not tigers (Shu & McCauley, 2017). All of these except one were collected between October 14th and 29th, 2021 in Colorado.

This putative lion to human transmission event therefore may have led to fifteen more cases in humans; however, the AY.20 lineage quickly disappeared from human populations. The relative transmissibility (among animals and between animals and humans) of this lineage carrying this mutation or the candidate hyena-adapted mutations, is still unknown. Ultimately, this event pro-

vides evidence that despite the viral within-host adaptation reported here (mainly, the N A254V mutation that was fixed in all lions and hyenas), the virus retained the ability to infect humans.

The potentially species-specific mutations identified in our study have been rarely observed in human sequences and are not associated with any particular variant lineage. Out of more than 15 million SARS-CoV-2 genome sequences in the GISAID database, N A254V is present in just 494 (complete, low coverage excluded) sequences recovered from humans (0.0033%) (Shu & McCauley, 2017). S T274I has been observed in 1,785 sequences, N P326L in 6,281 sequences and ORF1a E1724D in 42,275 sequences.

Interestingly, the three hyena-specific mutations have been observed independently in other non-human animal species. The ORF1a E1724D mutation has been identified in six sequences from white-tailed deer, one from a cat collected in July 2021, and one from an Amur tiger in Pennsylvania in December 2021 (GISAID accessions: EPI_ISL_13610651, EPI_ISL_16297302, EPI_ISL_16297304, EPI_ISL_16384469, EPI_ISL_16384478, EPI_ISL_9347710, EPI_ISL_3128536, EPI_ISL_8145733) (Shu & McCauley, 2017; Caserta et al., 2023). The S T274I mutation was reported in a white-tailed deer submitted to GISAID on the same date by the same laboratory as one of the deer sequences with E1724D, indicating these mutations may have been present in the same population of deer in New York (GISAID accession EPI_ISL_13658808). Finally, the N P326L mutation was detected in a mink sample collected in August 2020 (GISAID accession EPI_ISL_577784).

When we examined within-host viral variation in lions and hyenas relative to the tiger index cases, we observed variation at the species level in addition to the unique trajectory of each virus population in an individual host. The N A254V mutation was fixed in all lions and hyenas from the first samples collected on October 14th to the last positive test on December 13th. However, it was not observed in any tiger samples despite robust sequencing depth. This indicates it emerged in animals during the outbreak and was maintained by positive selection.

Viruses recovered from hyenas had more genetic variation than those from lions. Three potentially adaptive mutations were identified in hyenas: ORF1a E1724D, S T274I, and N P326L.

The sequence obtained from an RT-PCR inconclusive swab from Hyena D collected on November 22nd did not have these mutations and was indistinguishable from many lion sequences. In Hyena D, sequencing coverage was very good at the genomic position corresponding to the S T274I mutation but was mixed for N P326L and ORF1a E1724D. It is possible that younger Hyenas C and D were exposed prior to Hyenas A and B, along a similar timeline to lions, and quickly cleared their infections. If this were the case, the sequence from Hyena D would represent residual vRNA.

It is not a coincidence that the strongest signatures of positive selection were observed in the SARS-CoV-2 N gene, where we observed two potentially adaptive mutations in lions and hyenas, A254V and P326L. This gene encodes the RNA-binding nucleocapsid protein, which plays a crucial structural role in packaging the viral genome. Mutations in the N gene have been associated with the evolution of SARS-CoV-2 variants of concern in humans; most notably, R203K/G204R have been shown to increase viral replication and fitness (Johnson et al., 2022; H. Wu et al., 2021). Both A254V and P326L are located in the C-terminal, or dimerization domain of N. A recent study found that introducing the P326L mutation did not destabilize the dimer, and by one measurement, was more stable than the wildtype structure (Muradyan et al., 2024). Increased stability of the nucleocapsid protein could improve virion survival in the environment, which would directly enhance SARS-CoV-2 transmissibility. In some samples, signatures of positive selection were also observed in the spike gene. The location of the hyena-specific T274I mutation in this gene thus suggests potential adaptation to the hyena ortholog of angiotensin-converting enzyme 2 (ACE2), which mediates host cell entry. Interestingly, however, a recent review examining structural impacts of S mutations suggested that mutation at residue T274 would contribute to CD8+ T-cell immune escape in humans (Kombe, Biteghe, Ndoutoume, & Jin, 2022).

Overall, the within-host variation in lions and hyenas was indicative of the plasticity of within-host virus populations experiencing selection and genetic drift over time. The limited below consensus-level variation in lions is comparable to what is seen in humans (Lythgoe et al., 2021), while variation in hyenas resembles experimental studies demonstrating significant increases in within-host variation in domestic cats following experimental infection and cat-to-cat transmission

(Bashor et al., 2022). Virus populations often face substantial challenges to the maintenance of genetic variation due to transmission bottlenecks. In humans, SARS-CoV-2 infections are often initiated with just 1-2 virions (Markov et al., 2023), and this phenomenon likely played a role in limiting the size of the virus population transmitted from one individual animal to another during the outbreak. We found that SARS-CoV-2 virus populations in our samples were characterized by expansion and diversification over time. This finding is consistent with the long periods of infection seen in lions, and with the fundamental behavior of RNA viruses featuring rapid, but error-prone replication following the establishment of infection. Importantly, as a virus population expands, the potential strength of selection increases as well.

This dataset offered an unparalleled opportunity to document these phenomena in the context of multiple host shifts. We found that host species had a significant effect on the overall nucleotide diversity of our virus populations, but not on synonymous nucleotide diversity, which is a relative measure of virus effective population size. This could reflect a host-mediated species effect on mutation, as opposed to other aspects of the viral life cycle, like overall replication rate, assembly, or cell entry. As one possible example, members of the apolipoprotein B mRNA-editing enzyme catalytic polypeptide (APOBEC) family of proteins vary across vertebrates, and can directly introduce mutations into viral genomes (Harris & Dudley, 2015).

Signature of selection identified in virus populations in tigers, lions and hyenas were evenly split between positive and purifying selection. All six samples recovered from hyenas showed strong signs of positive selection at the genome level in comparison to other species. The patterns of mutation and selection in hyenas certainly points towards the possibility of an increased SARS-CoV-2 evolutionary rate in this species, as has previously been identified in white-tailed deer (McBride et al., 2023). However, a larger sample size representing sustained within-species transmission over time, in combination with Bayesian phylogenetic techniques, would be required to effectively evaluate this hypothesis. Ultimately, the within-host signatures of selection observed in our study can only suggest directions of SARS-CoV-2 host species-directed evolution, if the virus were given the opportunity in hyenas or another species.

In the case of this outbreak, no new, concerning variant lineage emerged in animals, but a handful of potentially adaptive mutations became fixed. It is important to note, however, that viral spillback from infected zoo felids to humans has been observed (Siegrist et al., 2023). Animal-to-human transmission has also been observed in a wide range species and environments, including pet hamsters, free-ranging deer, and farmed mink (Yen et al., 2022; Pickering et al., 2022; Feng et al., 2023; Munnink et al., 2021; Rabalski et al., 2022; Hammer et al., 2021). It is generally thought that after an initial period of evolution for transmissibility in humans, the current evolution of SARS-CoV-2 is driven by human immune pressure from both vaccination and prior infection26. Though both drivers act at the between-host level on a global scale, the variation upon which selection acts is still generated within individual hosts. The outbreak described in this study offered an exceptional opportunity to witness the mechanisms underlying virus within-host evolution following host shifts in action.

4.4 Methods

4.4.1 Animals and sample collections

The Denver Zoo exhibits African lions and spotted hyenas in one location, and Amur tigers in another. These two locations are approximately 200 m apart. At the time of the outbreak, eleven lions, four hyenas, and two tigers were housed at the Denver Zoo. The lions and hyenas did not have direct contact but rotated through shared spaces indoors and outdoors. Animal care specialists working with tigers, lions and hyenas did not overlap.

Both tigers were 11 years old at the time of the COVID-19 outbreak. The four male lions in pride 1 were siblings (all aged 6 years old) and were housed and exhibited communally. The four other adult lions and three lion cubs in pride 2 were housed and exhibited together. The members of pride 2 ranged in age from 1-2 and 6-9 years old. The hyenas were co-housed and exhibited in two clans: a pair of younger individuals (both 7 years old at the time of the outbreak) and a pair of older individuals (both 22 years old).

Following the onset of clinical signs in each species, voluntary nasal swabs were obtained from all individuals of a species regularly between October 7th, 2021, and January 31st, 2022

(Figure 4.1). Personal protective equipment was worn by animal care specialists before, during and after the outbreak. All animals made full recoveries following SARS-CoV-2 infection. Animals were later vaccinated with the animal-specific Zoetis vaccine. Tiger B was humanely euthanized one year after the outbreak due to chronic health conditions. Lion G and Lion H were both humanely euthanized in October and November 2023 respectively. Lion G developed a severe infection on top of several chronic health conditions; Lion H suffered from hydrocephalus, kidney failure, and severe arthritis.

4.4.2 Real-time PCR for diagnosis of SARS-CoV-2 infection

Swabs were transported to the CSU Veterinary Diagnostic Laboratories. Real-time PCR (RT-PCR) was carried out as previously described (Miller et al., 2024). Nasal swab samples were stored at -80°C for 6-8 months, at which point 200-500 μL aliquots were prepared for NGS evaluation.

4.4.3 Next-generation sequencing of RT-PCR positive nasal swab samples

Aliquots of nasal swab samples that tested positive for SARS-CoV-2 via RT-PCR were thawed and inactivated with Trizol reagent (Invitrogen) for five minutes at room temperature prior to RNA isolation using a modified protocol for the Zymo RNA Clean & Concentrator-5 kit with New England Biolabs DNase I. RNA was eluted in 20-24 μL nuclease-free water with a one minute and 9,000 x g centrifugation step. RNA was kept on ice or stored at -80°C before continuing with reverse transcription. Complementary DNA (cDNA) was generated using SuperScript™ IV Reverse Transcriptase (Invitrogen) and stored at -20°C prior to tiled amplicon PCR. Complete protocols for RNA and cDNA are available in the project Github repository at `https://github.com/laurabashor/DZSARS2/`. To control for false-positives resulting from PCR amplification, all biological samples were processed in two technical replicates for cDNA generation onwards.

Replicate samples were enriched for SARS-CoV-2 genomic material using a tiled amplicon PCR protocol developed by the ARTIC Network. The ARTIC version 4 (V4) primer scheme was used to generate a pool of overlapping amplicons that cover the entire SARS-CoV-2 genome. Pooled PCR products were visualized on 1% agarose gels, quantified using a Qubit dsDNA 1x High

Sensitivity Assay kit (Thermo Fisher Scientific), and 100 ng of DNA was input into library preparation. Libraries were prepared using NEBNext Ultra II DNA Library Prep kits (New England Biolabs) and individual replicates were uniquely dual indexed with NEBNext Multiplex Oligos for Illumina (New England Biolabs). All bead cleanups were performed using Ampure XP beads (Beckman Coulter).

All RT-PCR positive samples were sequenced on the same Illumina MiSeq instrument at the Colorado State University Next Generation Sequencing Facility using v2 500 cycle 2 x 250 bp kits (Illumina). The technical replicate libraries for Lion F collected on October 21st were pooled and sequenced as a test run with a separate project. All other individual libraries from positive samples were pooled evenly into three sequencing libraries corresponding to three separate sequencing runs, each containing a blank RNA extraction negative control library that was subjected to all of the processing steps. A spreadsheet detailing the samples and index sequences in each library is available in the project Github repository at `https://github.com/laurabashor/DZSARS2`.

4.4.4 Next-generation sequencing of RT-PCR inconclusive nasal swab samples

Aliquots of nasal swab samples that tested inconclusive for SARS-CoV-2 by RT-PCR (N= 1 tiger, 33 lion, 13 hyena samples) were screened for the quantity and quality of SARS-CoV-2 genomic material by processing through the ARTIC PCR and gel visualization step of the above protocol. Samples that appeared to be candidates for successful library preparation were then prepared in technical duplicate according to the above protocol, in addition to one blank RNA extraction negative control. All RT-PCR inconclusive samples were sequenced on one run of an Illumina MiSeq instrument at the University of Colorado Boulder Center for Microbial Exploration Sequencing Center using a v2 500 cycle 2 x 250 bp kit (Illumina). Sequencing datasets from six of these passed quality control and are included in this report, including the sample for Lion I on November 4th, Lion C on November 15th and 18th, and Lion F, Lion I and Hyena D on November 22nd.

4.4.5 Variant calling and bioinformatics

Sequencing data in FASTQ file format were input into the nf-core/viralrecon bioinformatics analysis pipeline version 2.6.0 for detailed quality control metrics, lineage assignment, within-host viral variant calling, variant annotation, and the generation of consensus sequences. The pipeline is designed to perform low-frequency variant calling for viral samples. For initial analyses, we used the nf-core/viral recon default as a SARS-CoV-2 reference sequence, GenBank MN908947.3. However, we used a modified GFF for MN908947.3 that separates ORF1a and ORF1b to account for the -1 ribosomal frameshift. This is consistent with the approach used by Nextstrain.

Briefly, nf-core/viral recon merges paired FASTQ files and read quality is assessed (FastQC) prior to trimming for quality and adapters (fastp) and removal of host reads (Kraken 2). Reads are aligned to a reference sequence, and variant calling is performed (Bowtie 2, SAMtools, iVar, picard) followed by variant annotation (SnpEff, SnpSift). Finally, consensus sequences are generated and assessed (BCFtools, QUAST, Pangolin, Nextclade). Coverage and quality metrics and a complete variant table, modified minimally for clarity, are available in Tables C.3 and C.4.

At this stage sequence data were subjected to a manual round of quality control prior to downstream analyses, using the metrics reported in the nf-core/viralrecon pipeline multiqc report (Table C.3). Replicate NGS datasets that had greater than or equal to 75% 1X and 10X coverage and less than 25,000 Ns per 100kb were retained for within-host variant analysis. This removed three datasets generated from blank RNA extraction negative control samples, 16 datasets corresponding to both technical replicates for eight biological samples, and two single replicate datasets corresponding to two other biological samples (Lion I on November 22nd and Lion C on November 4th). After quality control, 124 next-generation sequencing datasets corresponding to 63 biological samples from 16 individuals (three hyenas, 11 lions and two tigers) were retained. This included 1 or 2 timepoints for each tiger, 3-7 timepoints for each lion, and 1-4 timepoints for each hyena. The mean sequencing depth of a dataset ranged from 2012X to 6997X. The overall mean and median sequencing depth of coverage across the full SARS-CoV-2 genome were 3754X and 2826X respectively.

For analysis requiring all consensus sequences in our dataset (the phylogenetic tree) one of the two technical replicates was selected using the quality control metrics generated by the nf-core/viralrecon pipeline. Of the two technical replicates, the sequence with a higher value for "99% coverage > 10x" was selected. If this metric was the same, the sequence with fewer Ns per 100kb was selected. If the number of Ns was the same, the sequence with a higher overall median coverage value was selected.

For analyses involving just one consensus sequence per individual (the haplotype network), we aimed to select the highest quality sequence collected at the latest date. To do this, the sequence generated at the latest timepoint with high coverage metrics ($\geq$ "99% coverage > 1X" and $\geq$ "99% coverage > 10X" and < "1225 Ns per 100kb") was selected. The dataset generated from the sample collected on Nov. 22nd from Hyena D passed the initial manual quality control step after NGS, but did not meet this more stringent cutoff. However, it was still included because it was the only sequence generated from this individual.

For a subset of additional analyses, a second bioinformatics pipeline for viral variant analysis was employed: `https://github.com/stenglein-lab/viral_variant_caller`. This pipeline outputs raw sequencing depth across all datasets at all genomic positions which we used to calculate summary statistics. It was also used to generate VCF files in the appropriate format for input in the SNPGenie pipeline (see below).

Based on our conclusion that the two tigers were infected first, we also assessed within-host variation relative to a "tiger reference sequence:" the consensus SARS-CoV-2 sequence of the virus samples recovered from the two tigers tested on October 7th. The two tiger sequences were aligned with MUSCLE, and a consensus sequence was generated with emboss. This consensus was then used as the reference sequence for running the nf-core/viralrecon pipeline with a custom configuration, using the remaining NGS datasets from lions and hyenas (only those with two high quality technical replicates) as input.

4.4.6 Mutation naming conventions and validation

For all mutations described and visualized in this report, the first naming convention used for mutations in the region that codes for the ORF1ab polyprotein is described as its position within either ORF1a or ORF1b. To help provide additional context, for mutations in this region we also list the specific nonstructural protein affected. We also interrogated sequencing coverage at the genomic position corresponding to each mutation discussed in this report to and confirmed >100X coverage in both sequencing replicates.

To identify mutations in our dataset that are characteristic of SARS-CoV-2 lineage AY.20, a list of 32 characteristic mutations was obtained from `https://outbreak.info/situation-reports?pango=AY.20`. This list is available as Table C.2. We modified this table with information about the nonstructural protein impacted by mutations in ORF1ab, and with alternate names for mutations based on the annotation and naming convention output by the nf-core/viralrecon pipeline. Briefly, the outbreak.info mutation list includes E156G and del157/158 in Spike, whereas our dataset has the mutation E156_R158delinsG. This is the same change with a different name. Similarly, outbreak.info lists del119/120 in ORF8, whereas our dataset has D119_F120del. Finally, our dataset describes a C>T change at genomic position 27874, which would cause a T to an I amino acid change. In our annotation schema, this effect is described as upstream of ORF8, and in outbreak.info it is listed as T40I occurring in ORF7b. These are also the same change with a different name. Finally, S84L in ORF8 is listed as a characteristic mutation of AY.20 in the outbreak.info list. However, this mutation was already present in the reference sequence used in our analysis (GenBank MN908947.3), and all sequences we generated also had the L residue at this position. Therefore, after accounting for these annotation disparities, all mutations characteristic of AY.20 were identified in our sequences.

Some mutations that appear in the Nextstrain phylogeny are not discussed in our results, mainly due to insufficient sequencing depth or failure to be detected in both technical replicates. Various mutations observed at residue S2287 in ORF1a in Lion I on October 29th, Lion J on October 26th, Lion I on November 22nd, and Lion C on November 4th were not supported by robust se-

quencing depth, indicating potential error. A number of other discrepancies between the Nextstrain phylogeny and the variant analysis reported here are noted below. The sequence from Lion F on November 2nd also had the unique mutation T632I in S, but this was not supported by both technical replicates despite robust sequencing depth. The sequence from Lion C on November 18th had the S3376A in ORF1a and T1009I in S but S3376A is not supported by sequencing depth and T1009I was not supported by both technical replicates despite robust sequencing depth. The sequences from Lion J on October 26th, Lion C on November 4th, and Lion I collected on November 22nd share the reversion to root ORF1a S2287P in nsp3 (nucleotide mutation T7124C). However, this genomic position is not supported by robust sequencing depth in either of these samples. The sequence from Lion B on December 13th had the unique mutation ORF1a L1375F and T1854I; both were not supported by both technical replicates despite robust sequencing depth. The sequence from Lion C on November 4th also has three other unique mutations: L3201F and C4213F in ORF1a and Q56L in ORF6. Because we were not able to generate a technical replicate for this sample, we did not include this sample in our within-host variation analysis. As no other samples from Lion C had these mutations, they may represent sequencing errors.

4.4.7 Public SARS2-CoV-2 data, phylogenetic trees and haplotype network

A time-based phylogenetic tree was generated with Nextstrain software following the ncov SARS-CoV-2 workflow (Hadfield et al., 2018). All sequence data not generated in this study were obtained from GISAID (Shu & McCauley, 2017). The tree (Figure 4.2A) included three data inputs: (1) sixty-three SARS-CoV-2 consensus sequences generated from zoo animals, (2) all human-derived sequences classified as AY.20 in the state of Colorado between September 23rd and November 4th, 2021 (N = 198 sequences), and (3) a random subsample of human-derived sequences in Colorado from the days leading up to the zoo outbreak (100 sequences/day from September 23rd to October 7th, 2021; N=1500 sequences). The first tree is publicly available at `https://nextstrain.org/community/laurabashor/DZSARS2/COAY20`. A second phylogenetic tree (Figure S1) was generated with Nextstrain using the above method, and including four data inputs: (1) sixteen SARS-CoV-2 consensus sequences from the Denver Zoo out-

break corresponding to the highest quality sequence collected at the latest date from each individual animal, (2) a random subsample of human-derived sequences in Colorado from the days leading up to the zoo outbreak (100 sequences/day from September 23rd to October 7th, 2021; N=1500 sequences), (3) all felid- and hyena- derived SARS-CoV-2 sequences available in the GISAID database prior to December 2021, and (4) contextual sequences collected in the same time and place as each animal sequence (100 human-derived sequence per animal-derived sequence from the United States; 10 sequences for animal-derived sequences from other locations). Felids include domestic cats, lions, tigers, snow leopards, a fishing cat, and a leopard cat. The second tree is publicly available at `https://nextstrain.org/community/laurabashor/DZSARS2/felidshumans`. Detailed acknowledgements for both trees including the authors, originating and submitting laboratories of all sequence and metadata obtained from GISAID are available in GISAID EPI_SET_240407tz and EPI_SET_240407wv (see Supplementary Information). A haplotype network was generated with the 'pegas' package in R using an alignment (MUSCLE) of the latest high-quality sequence obtained from each individual.

SARS-CoV-2 genomes from nasal swab samples from all eleven lions collected on October 14th and one hyena collected on October 28th were also sequenced by USDA APHIS at the time of the outbreak (accessions included in GISAID EPI_SET_240407tz). A comparison of the consensus sequences obtained in our study from lions and hyenas on these dates reveals high overall identity, with two SNPs observed in the APHIS sequences that were not detected in our datasets. Based on shared SNPs, the hyena sequence (GISAID accession EPI_ISL_6810900) was likely generated from a sample from Hyena A. All other GISAID sequences discussed in this report are included in GISAID EPI_SET_240407tz.

4.4.8 Within-host virus population demographics and signatures of selection

Within-host virus population demographics and signatures of selection were calculated with SNPGenie (Nelson, Moncla, & Hughes, 2015). The SNPGenie pipeline calculates nucleotide diversity as the mean number of pairwise differences per nonsynonymous or synonymous site in the genome from next-generation virus sequencing data, and estimates are weighted by allele

frequencies at these sites. This analysis was done at both the population, or virus sample level, and by SARS-CoV-2 gene product.

Linear mixed models were used to evaluate statistical differences in πN and πS over time (measured as days since first positive test) and among host species, using the lmer() function in the 'lme4' package in R. Time and species were fixed effects, with animal ID as a random effect to account for repeated measures of the same individual over time. The residuals of both models were normally distributed, and the statistical summary table was generated with the tab_model() function in the 'sjPlot' package in R. In this table, the conditional R^2 for the πS model on the righthand side is "NA" because zero variance was explained by the animal ID grouping factor. No formal analysis was done to distinguish endpoint πN/πS ratios for Figure 4.5A due to insufficient sample size.

4.5 Funding

The research reported in this chapter was supported by the National Institute Of Allergy And Infectious Diseases of the National Institutes of Health under Award Number T32AI162691 and the Colorado State University College of Veterinary Medicine and Biomedical Sciences Research Council Award and Colorado State University's Office of the Vice President for Research's 'Accelerating Innovations in Pandemic Disease' initiative, made possible through support from The Anschutz Foundation. A MARC Scholar was funded by a grant from the National Institute of General Medical Sciences of the National Institutes of Health: T34GM140958. Computational resources were supported by NIH/National Center for Advancing Translational Science Colorado Clinical and Translational Science Awards grant UL1 TR002535.

Chapter 5

Distinct patterns of within-host evolution detected in experimental exposures of wildlife species with two SARS-CoV-2 variants[4]

SARS-CoV-2 infects a wide range of non-human animal species, and the risk of virus maintenance and continued evolution, particularly in peridomestic wildlife populations, is still poorly understood. Furthermore, SARS-CoV-2 has continued to change as it adapts to humans, evolving into successive new variant lineages. We demonstrate the use of experimental exposures as a controlled system to test hypotheses surrounding SARS-CoV-2 adaptation following host shifts. We evaluated viral genomes recovered from six species of wildlife experimentally exposed to two SARS-CoV-2 variants, WA01 and Delta, that were isolated from humans and propagated in Vero cells prior to infection. Oral and nasal swab samples were collected between 2-7 days post-infection from deer mice (*Peromyscus maniculatus*, n=3), bushy-tailed woodrats (*Neotoma cinerea*, n=3), Brazilian free-tailed bats (*Tadarida brasiliensis*, n=4), striped skunks (*Mephitis mephitis*, n=5), red foxes (*Vulpes vulpes*, n=9), and mule deer (*Odocoileus hemionus*, n=6), and subjected to next-generation sequencing of full viral genomes at a high depth of coverage, using a tiled amplicon approach. Consistent with our expectations, we observed within-host variant emergence and genomic signatures of selection in viruses recovered from infected animals. Notably, we observed significantly more within-host variation (WA01 mean: 5.25 variants, Delta mean: 2.67 variants), higher relative effective population sizes, and genomic signatures of positive selection in viruses from animals infected with the more ancestral WA01 variant lineage compared the Delta variant. We also noted a potential host species barrier to infection in striped skunks exposed to WA01. Our findings did not support our hypothesis that we would observe increased adaptation in a relatively less SARS-CoV-2-susceptible species, Brazilian free-tailed bats. However, we observed six apparently de novo mutations in two mule deer infected following deer-to-deer contact. This confirmed our expectation that free-ranging mule deer, similar to white-tailed deer populations, pose the risk of acting

[4]This chapter was co-authored by Bashor, L., Porter, S., Marano, J.M., Root, J., Bosco-Lauth, A., & VandeWoude, S.

as viral reservoirs and supporting rapid SARS-CoV-2 evolution. The results of this study highlight the continued need for SARS-CoV-2 surveillance in peridomestic wildlife species, and successfully demonstrate the use of experimental infections to investigate viral evolution following host shifts.

5.1 Introduction

SARS-CoV-2 is the major contemporary example of the origins of a pandemic: zoonotic spillover of a pathogen from a non-human animal host into humans, followed by adaptation to humans. However, spillover is not unidirectional; a broad range of animal species including domestic animals, wild animals living in captivity, and peridomestic wildlife have been exposed through human-to-animal transmission of SARS-CoV-2 (APHIS, 2024a). Furthermore, humans have also been infected again through spillback (Sila et al., 2022; Yen et al., 2022; Munnink et al., 2021; Lu et al., 2021; Hammer et al., 2021; Siegrist et al., 2023; Pickering et al., 2022; Feng et al., 2023).

Every cross-species transmission event is noteworthy due to the potential for strong selective pressures exerted by a novel host. The frequency with which these host shifts have occurred and the role that they have played in SARS-CoV-2 evolution to date can only be estimated. However, it is clear that the virus has undergone strong selection in both humans and animals over the course of the pandemic (Kemp et al., 2021; McBride et al., 2023; Tan et al., 2022). Of particular note, high infection prevalence, rapid virus evolution and at least one spillback into humans have been documented in free-ranging white-tailed deer (WTD; *Odocoileus virginianus*) populations in North America (Pickering et al., 2022; Feng et al., 2023; McBride et al., 2023; Chandler et al., 2021; Hale et al., 2021; Kuchipudi et al., 2022; Palermo et al., 2021). Virus genomes recovered from WTD have spanned the breadth of SARS-CoV-2 variant lineages, including ancestral variants that are no longer circulating in humans (Feng et al., 2023; Marques et al., 2022; Caserta et al., 2023; Vandegrift et al., 2022). Furthermore, they carried many divergent mutations, leading to the finding that SARS-CoV-2 evolves at a faster rate in WTD than in humans (McBride et al., 2023).

Although the most extensive SARS-CoV-2 surveillance of wildlife has focused on WTD populations, there have also been surveillance efforts in other species (Aguiló-Gisbert et al., 2021; Earnest et al., 2023; Despres et al., 2023; Italiya et al., 2023; APHIS, 2024b). Pairing disease surveillance with investigation of virus evolution in free-ranging animals poses additional challenges. An important first step is to identify species that pose a greater risk or are of particular scientific interest. Experimental exposures offer an opportunity to evaluate virus evolution following host shifts across multiple species in a controlled environment. Previous studies by our research group have highlighted interesting patterns in within-host virus evolution in domestic animals, including increased signs of virus adaptation in dogs, a relatively less susceptible host species (Bashor et al., 2021, 2022).

Evaluating SARS-CoV-2 evolution and transmission in additional species is thus important for two reasons: (1), furthering scientific understanding of virus evolution following host shifts in the face of both SARS-CoV-2 and future pandemics, and (2), assessing the potential of specific animal populations to support the evolution of divergent SARS-CoV-2 variants. The susceptibility of many peridomestic wildlife species to SARS-CoV-2 has previously been established. Furthermore, the ecological factors that enable spillover and spillback are already in place for these animals living in close proximity to humans. However, species-specific drivers of virus evolution, in particular the within-host dynamics immediately following a cross-species transmission event, remain to be determined.

In this study, we tested the hypothesis that SARS-CoV-2 evolution following a host shift would be characterized by within-host variant emergence and genomic signatures of selection using virus samples collected from six species of experimentally inoculated animals: deer mice (*Peromyscus maniculatus*), bushy-tailed woodrats (*Neotoma cinerea*), Brazilian free-tailed bats (*Tadarida brasiliensis*), striped skunks (*Mephitis mephitis*), red foxes (*Vulpes vulpes*) and mule deer (*Odocoileus hemionus*). We observed distinct patterns of within-host evolution based on the SARS-CoV-2 variant lineage used for infection (WA01 or Delta). In particular, WA01 virus populations in striped skunks were impacted by host barrier. Although we did not confirm our hypoth-

esis that SARS-CoV-2 would evolve more rapidly in Brazilian free-tailed bats, a less susceptible species, we did confirm rapid virus evolution following deer-to-deer transmission.

5.2 Results

5.2.1 Next-generation sequencing of SARS-CoV-2 genomes from experimentally infected wildlife

Deer mice, bushy-tailed woodrats, Brazilian free-tailed bats, red foxes, striped skunks, and mule deer were previously determined to be susceptible to SARS-CoV-2 by experimental exposure (A. M. Bosco-Lauth et al., 2021, 2022; Porter et al., 2022, 2024). A brief summary of the findings from these studies, including the dose and variant lineage of the inoculum virus, detection of virus shedding, seroconversion, and clinical signs is presented in Table 5.1. In this study, next-generation sequencing (NGS) was used to further evaluate viral RNA isolated from nasal and oral swab samples collected during these studies.

Relative differences in the levels of viral RNA present in inoculum and swab samples included in this study were detected by RT-PCR (Figure 5.1a,b). There were significant differences in viral RNA levels measured by RT-PCR cycle threshold (Ct) between animal-derived (mean: 24.6) and inoculum-derived (mean: 14.7) samples (t-test, t = 2.6, df = 29, p-value = 0.013). Furthermore, there were significant differences in viral RNA levels by host species (ANOVA, df = 5, f-value = 7.3, p-value = 0.00030). Pairwise comparisons indicated significantly more viral RNA (lower mean Ct values) was detected in mule deer and red fox samples (Table D.1).

To evaluate SARS-CoV-2 variation and evolution in susceptible wildlife species, we sequenced complete virus genomes from six species experimentally exposed to two variant lineages of SARS-CoV-2: WA1/2020WY96 and Delta hCoV-19/USA/MD-HP05647/2021, hereafter referred to as WA01 and Delta (Figure 5.1c,d). Sequencing data were also generated from the viral stocks used for inoculation. Sampling collection method and timepoint or passage varied among animal and inoculum samples (Table 5.2). High-quality NGS datasets were obtained for 31 samples, including two inoculum samples and 29 animal-derived samples (Figure 5.1d, Table D.2). Although all samples were sequenced at a high depth of coverage, Delta samples were sequenced more deeply

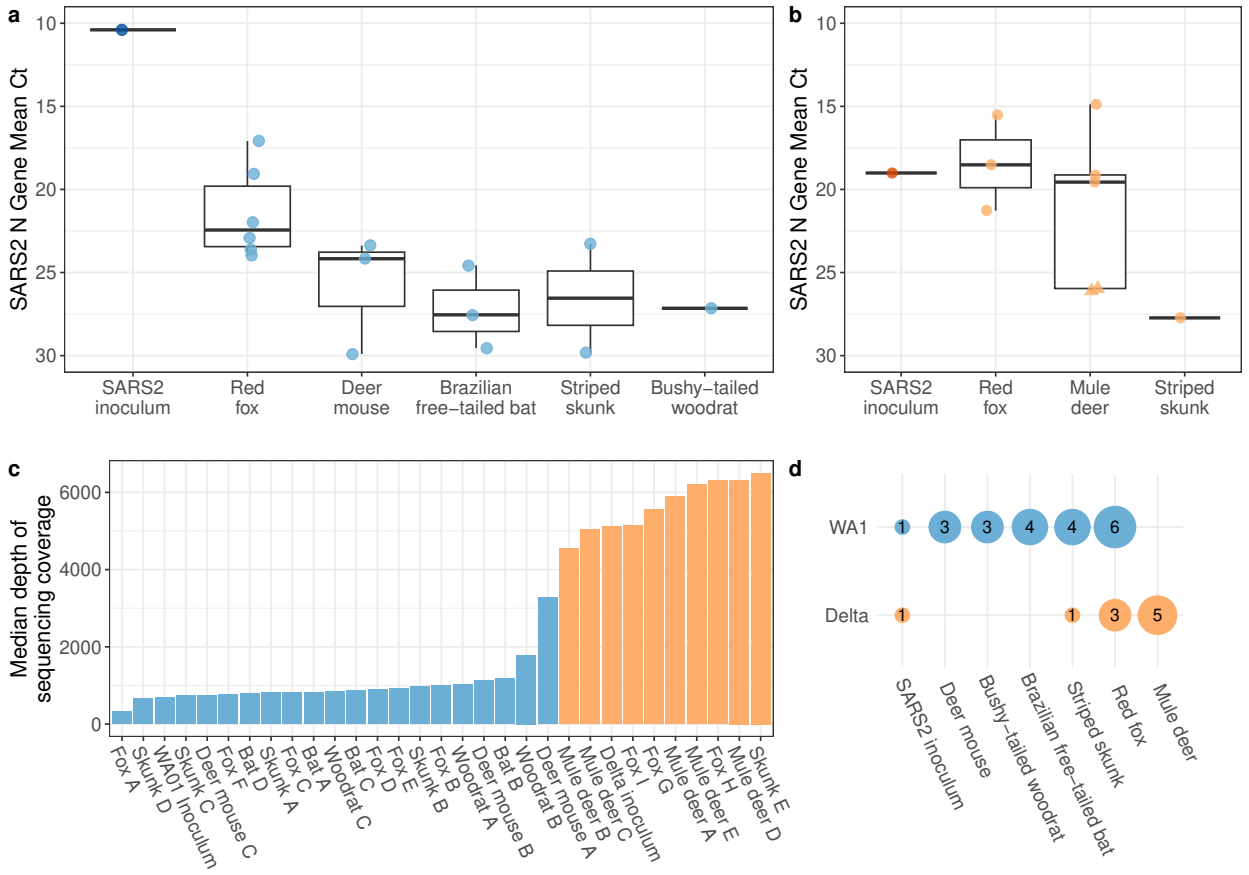


Figure 5.1: Detection of SARS-CoV-2 RNA and NGS coverage depth varied among virus samples. Viral RNA was isolated from SARS-CoV-2 variant inoculums (n=2), and nasal and oral swabs collected from experimentally infected deer mice (n=3), bushy-tailed woodrats (n=3), striped skunks (n=5), Brazilian free-tailed bats (n=4), red foxes (n=9), and mule deer (n=5). In each plot, samples are colored according to SARS-CoV-2 variant (WA01=blue, Delta=orange). (**a,b**) Levels of SARS-CoV-2 (SARS2) viral RNA were measured with an RT-PCR targeting the viral nucleocapsid (N) gene. Each point indicates the mean cycle threshold (Ct) value of two technical replicates for one viral RNA sample. Triangles indicate mule deer infected through contact transmission. RT-PCR was performed with the EXPRESS One-Step SuperScript qRT-PCR Kit (ThermoFisher Scientific). (**c**) Median depth of sequencing coverage across the SARS-CoV-2 genome for each viral RNA sample subjected to NGS. All WA01 samples were sequenced in technical duplicate except for Woodrat B, and the Delta inoculum was sequenced in triplicate; in these cases, each bar represents the median of the replicate datasets combined. Delta samples from animals were sequenced singly. (**d**) Total number of samples subjected to NGS grouped by species and SARS-CoV-2 variant inoculum.

(Figure 5.1c). All WA01 and Delta samples, including the inoculums, were classified as Pangolin lineages A and AY.106 respectively.

5.2.2 Within-host and consensus-level SARS-CoV-2 genetic variation

Across all SARS-CoV-2 samples from infected animals and the WA01 and Delta inoculums, we observed 68 unique mutations in the SARS-CoV-2 genome relative to the ancestral Wuhan-1 reference sequence (GenBank MN908947). Twenty-eight of these were found at least once in an animal or inoculum sample at variable allele frequencies (<100%), or were fixed in an animal and not detected in the inoculum. In this report, we will refer to these mutations as within-host variants. Using this classification, four within-host variants were present in the WA01 inoculum, and two in the Delta inoculum sample (Figure 5.2a). Notably, there was a significant difference between the average number of within-host variants in WA01 (mean: 5.25 variants) and Delta (mean: 2.67 variants) samples from infected animals (t-test, t = -4.1, df = 27, p-value = 0.00037).

There was variation in the number of within-host variants observed both between and within species. However, a significant association between the number of within-host variants and host species was not observed, even taking into account the effect of variant lineage (ANOVA, species: df = 6, f-value = 2.1, p-value = 0.094; variant lineage: df = 1, f-value = 23.2, p-value = 0.000074). Notably, the three directly inoculated deer had the same sequence as the Delta inoculum, whereas samples from the two contact-infected mule deer each had 2-4 more within-host variants. We also tested the hypothesis that the amount of viral RNA in the original sample could impact the number of within-host variants detected, but found no significant effect of RT-PCR Ct on the number of variants (linear model, R^2 = 0.00086 , f-value = 0.016, df = 19 , p-value = 0.90).

5.2.3 Specific SARS-CoV-2 mutations identified in wildlife and their predicted effects

More than half (n=17) of the 28 SARS-CoV-2 within-host variants in our dataset were classified as missense (nonsynonymous) single nucleotide variants (SNVs) (Figure 5.2c,d). Predicted effects of the remaining variants included synonymous (n=7), conservative in-frame deletion (n=1),

Table 5.1: Summary of relevant findings from published experimental exposure studies revealing that the six wildlife species in this study are susceptible to SARS-CoV-2. Dose is represented by the range of virus plaque-forming units (pfu) received by individuals of a species via intranasal inoculation, measured by back-titration on Vero cells. Two mule deer were also infected through contact transmission. Shedding of infectious virus was assessed by plaque assay on Vero cells. Samples collected between 2-7dpi from each species were obtained for further evaluation by NGS in this study.

Species	Susceptibility category	SARS-CoV-2 Variant	Dose (log_{10} pfu)	Shedding of infectious virus	Seroconversion	Clinical signs	Reference
Deer mouse	susceptible	WA01	4.5–4.9	shed orally	neutralizing Abs	none	A. M. Bosco-Lauth et al., 2021
Bushy-tailed woodrat	susceptible	WA01	4.5–4.9	shed orally	neutralizing Abs	none	A. M. Bosco-Lauth et al., 2021
Brazilian free-tailed bat	minimally permissive	WA01	3.5-3.6	no shedding	neutralizing Abs in 25% of bats tested	none	A. M. Bosco-Lauth et al., 2022
Striped skunk	susceptible	WA01 Delta	4.5–4.9 4.5-4.8	shed orally and nasally	neutralizing Abs	none	A. M. Bosco-Lauth et al., 2021
Red fox	susceptible	WA01 Delta	5.1-6.0 4.5-4.8	shed orally and nasally	neutralizing Abs	lethargy and sneezing	Porter et al., 2022
Mule deer	susceptible	Delta	4.5	shed orally and nasally	neutralizing Abs	none	Porter et al., 2024

Table 5.2: Sampling collection method and timepoint or passage varied among animal and inoculum samples. For experimentally exposed animals, timepoint is presented as days post-infection (dpi). For contact mule deer, dpi references the number of days after the direct inoculation of the other deer. For inoculum samples, timepoint is presented as passage number (P1 or P2) in Vero cells.

Variant lineage	Species	Animal ID	Infection type	Sample timepoint and type
WA01	WA01	SARS-CoV-2	inoculum	P2 cell supernatant
	Deer mouse	Deer mouse A-C	direct	3dpi oral swab
	Bushy-tailed woodrat	Woodrat A-C	direct	3dpi oral swab
	Brazilian free-tailed bat	Bat A-C	direct	3dpi oral swab
		Bat D		2dpi oral swab
	Striped skunk	Skunk A-D	direct	3dpi oral swab
	Red fox	Fox A-F	direct	3dpi oral swab
Delta	Delta	SARS-CoV-2	inoculum	P1 cell supernatant
	Striped skunk	Skunk E	direct	2dpi oral swab
	Red fox	Fox G-I	direct	2dpi oral swab
	Mule deer	Mule deer A	direct	2dpi nasal swab
	Mule deer	Mule deer B	contact	3dpi nasal swab
	Mule deer	Mule deer C	contact	7dpi nasal swab
	Mule deer	Mule deer D	direct	5dpi nasal swab
	Mule deer	Mule deer E	direct	3dpi nasal swab

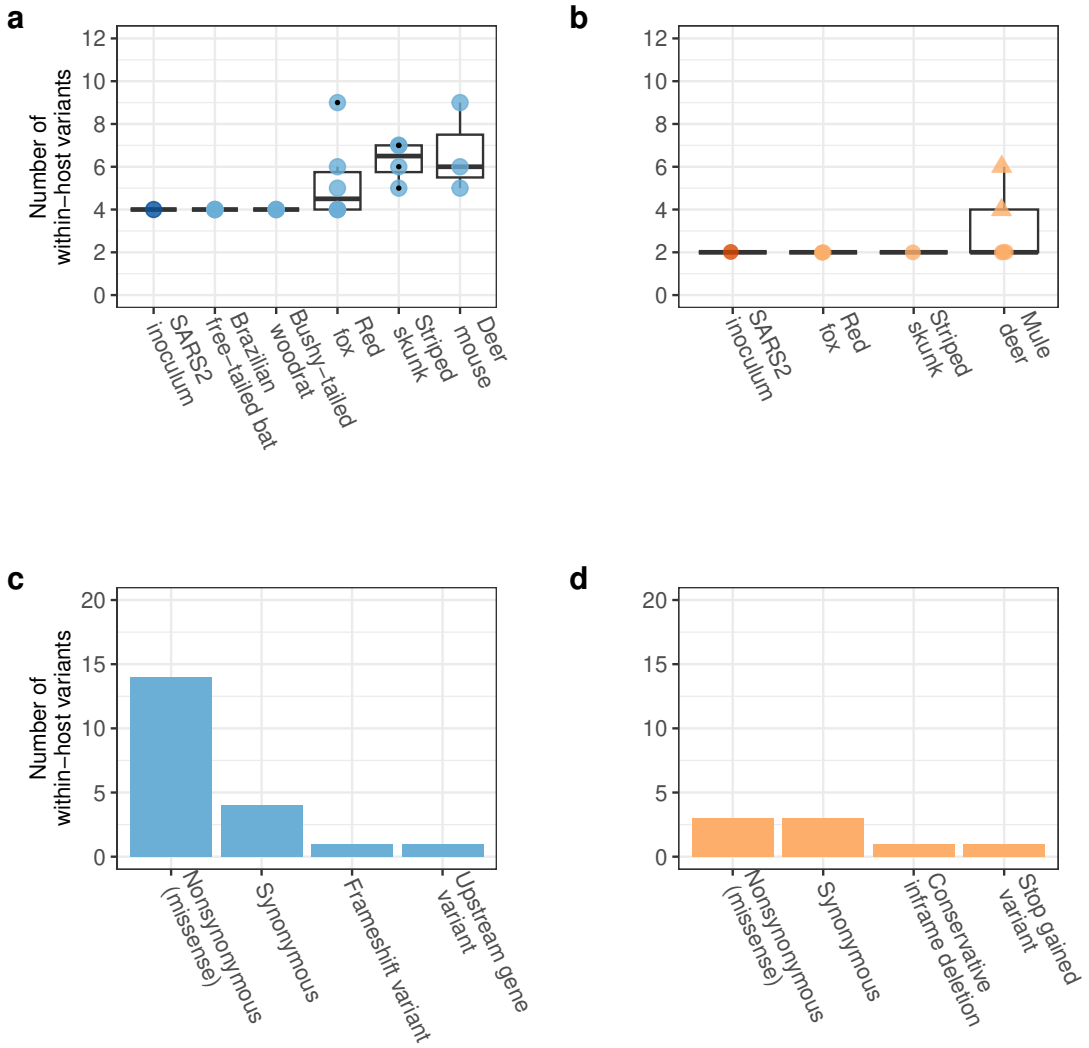


Figure 5.2: Virus genomes from animals infected with the SARS-CoV-2 WA01 variant had (**a,b**) more within-host variants than Delta variant samples, although there were more in contact-infected deer than other Delta animals; (**c,d**) nonsynonymous single nucleotide variants (SNVs) were the most common type of within-host variant in both WA01 and Delta samples. In each plot, samples are colored according to SARS-CoV-2 variant (WA01=blue, Delta=orange). Within-host variants include mutations relative to the Wuhan-1 reference sequence that were detected at least once in an animal or inoculum sample at variable allele frequency, or mutations that were fixed in an animal and not detected in the inoculum. (**a,b**) The number of within-host variants detected in (**a**) WA01 and (**b**) Delta variant inoculum and animal virus samples. The mean number of within-host variants in WA01 animals was significantly greater than in Delta animals (t-test, t = -4.1, df = 27, p-value = 0.00037). All animals were exposed by direct intranasal inoculation except for two mule deer infected through contact transmission, indicated with triangles. A black circle (•) in the center of a point indicates a virus sample from a WA01 red fox or striped skunk that showed signs of a host barrier to infection. (**c,d**) The predicted effect of the 28 unique within-host variants detected in (**c**) WA01 and (**d**) Delta virus samples.

frameshift, stop lost, & splice region variant (n=1), stop gained (n=1), and upstream gene variant (n=1).

Eighteen within-host and two fixed variants were observed just once in one individual animal, and not in the relevant viral inoculum (Table 5.3). These variants may have emerged de novo during infection. Interestingly, these potential de novo variants were only observed in three of the five WA01-infected species: deer mice, red foxes, and striped skunks. Furthermore, potential de novo variants were only seen in the contact-infected Delta mule deer. The six single-observation variants seen in Mule deer B and C included three mutations in open reading frame (ORF) 1b: A831A, P864P, and D1388Y, the first two of which are synonymous changes in the nonstructural protein (nsp) 12, which encodes the SARS-CoV-2 RNA-dependent RNA polymerase (RdRp), and the third a nonsynonymous change in nsp13, which encodes the SARS-CoV-2 helicase. The three other mule deer mutations include envelope (E) L51F and ORF7a Q62* and Y97Y.

Based on our observation of significant differences between the WA01 and Delta infections, we chose to examine specific mutations identified in these datasets separately. In the original WA01 inoculum sample, passaged twice in Vero cells, six mutations were identified relative to the Wuhan-1 reference sequence ranging from 47-100% allele frequency (AF). Three of these mutations were fixed in the inoculum sample and essentially remained fixed in all infected animals: ORF1a S2839S, ORF1b L1531L, and ORF8 L84S (Figure 5.3a). ORF8 L84S decreased to 99.5% AF in one bat, and ORF1a S2839S and ORF1b L1531L were not observed in some animal samples due to low sequencing coverage at those genomic positions. Three other mutations, spike (S) D215H, membrane (M) T7I, and nucleocapsid (N) S194T, were detected in the inoculum at 47-66% AF, and saw further variation in animals.

Changes in variant AF between the inoculum and animal samples revealed the impacts of a bottleneck or selective sweep reflecting a host barrier to infection (Figure 5.3a). The three variants present at varying allele frequencies in the WA01 inoculum, S D215H, M T7I, and N S194T, reached near fixation in samples from all four striped skunks, and one red fox, Fox C. This suggests that all three are linked and were transmitted as part of the same viral haplotype. Notably, potential

Table 5.3: Twenty within-host variants may have emerged de novo in experimentally infected animals. The listed mutations were observed once in an animal sample but were not detected in the viral inoculum. Mutations are grouped by species and inoculum variant, including relevant annotations and observed allele frequency (AF). Both mule deer B and C were infected through contact transmission from another directly inoculated mule deer (indicated with an *). Samples from Fox C and Skunks A-D showed signs of a host barrier to infection (indicated with a •)

Species (inoculum variant)	Animal ID	Genome position	Amino acid variant	AF	Reference base	Variant base	Predicted effect
Deer mouse (WA01)	A	22301	S S247R	0.45	A	C	missense
	A	23525	S H655Y	0.49	C	T	missense
	A	27381	ORF6 D61fs	0.495	TGATTA	T	frameshift, stop lost & splice region variant
Mule deer (Delta)	C*	15960	ORF1b A831A (nsp12 A840A)	0.62	C	T	synonymous
	B*	16059	ORF1b P864P (nsp12 P873P)	0.33	T	C	synonymous
	C*	17632	ORF1b D1388Y (nsp13 D465Y)	0.57	G	T	missense
	B*	26395	E L51F	0.42	C	T	missense
	B*	27577	ORF7a Q62*	0.28	C	T	stop gained
	B*	27684	ORF7a Y97Y	0.28	C	T	synonymous
Red fox (WA01)	A	78	T78G	0.415	T	G	upstream gene variant
	C •	1401	ORF1a G379E (nsp2 G199E)	0.9	G	A	missense
	C •	12755	ORF1a T4164A (nsp9 T24A)	0.98	A	G	missense
	D	21849	S E96A	0.285	A	C	missense
	C •	23536	S N658N	0.975	C	T	synonymous
	C •	29118	N T282I	0.935	C	T	missense
Striped skunk (WA01)	C •	9985	ORF1a D3240D (nsp4 D477D)	0.705	C	T	synonymous
	A •	22458	S T299I	1	C	T	missense
	D •	22661	S V367F	0.49	G	T	missense
	A •	24259	S A899A	1	T	C	synonymous
	C •	28133	ORF8 T80T	0.45	A	G	synonymous

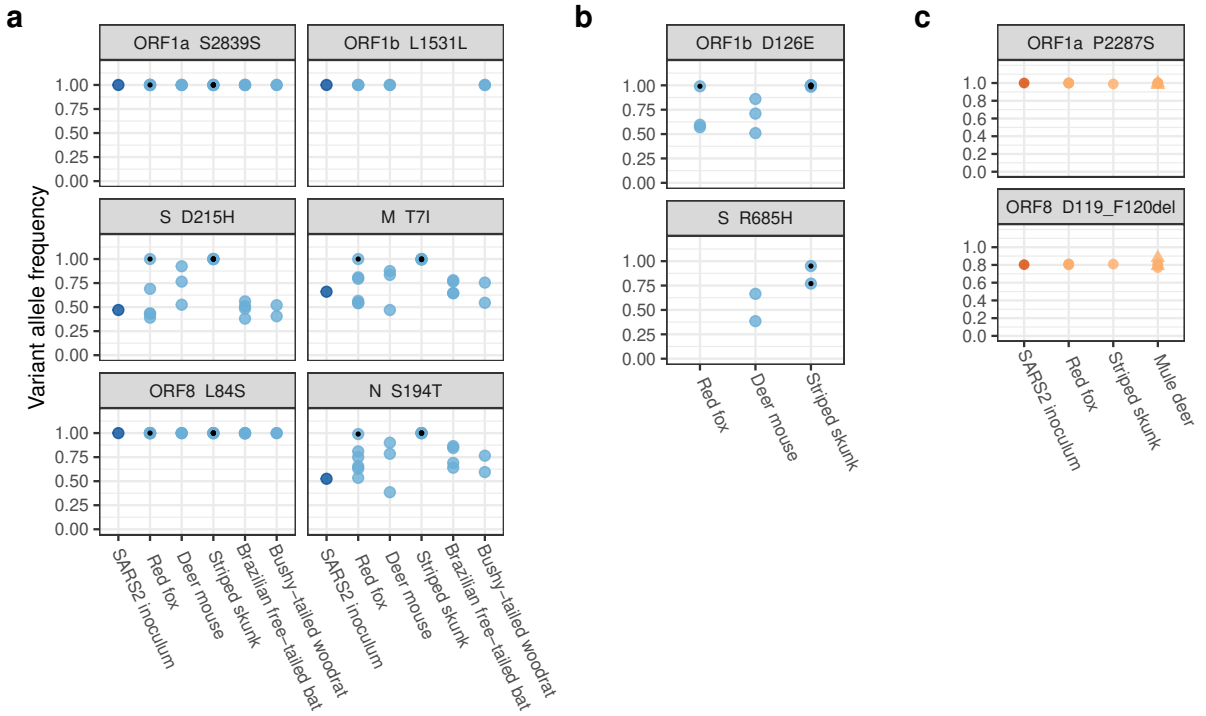


Figure 5.3: SARS-CoV-2 within-host variants were shared between and within experimentally infected wildlife species; the fixation of a set of variants in WA01 striped skunks and one red fox implies a host barrier to infection. In all plots, each plot facet is a SARS-CoV-2 mutation relative to the Wuhan-1 reference sequence, and each point represents one observation of that mutation in an inoculum or animal sample. Points are colored according to SARS-CoV-2 variant used in infection (WA01=blue, Delta=orange). Triangles indicate mule deer infected through contact transmission. A black circle (•) in the center of a point indicates a virus sample from a WA01 striped skunk or red fox that showed signs of a host barrier to infection. (**a**) The six within-host variants detected in the WA01 inoculum were also detected in directly inoculated animals. Although S D215H, M T7I, and N S194T were detected at lower frequency in the WA01 inoculum and many animal samples, they became fixed in Fox C and Skunks A-D. (**b**) Two within-host variants were found in multiple WA01 animals but were not detected in the inoculum. (**c**) Two within-host variants were observed in the Delta variant inoculum and animals.

de novo variants were noted more often in the virus samples from these four striped skunks and Fox C (Table 5.3). Fox C had four de novo variants, all detected at >90% AF and the striped skunks altogether had five de novo variants, two of which were fixed in Skunk A.

Two additional variants were observed in multiple WA01-infected animal samples, but we were unable to detect in the inoculum at any level (Figure 5.3b). The S R685H mutation was observed in two deer mice and two striped skunks. The ORF1b D126E mutation was observed in all three deer mice, three red foxes, and all four striped skunks. The fixation of this mutation in all four striped skunks and Fox C suggests that it may have been linked to the three inoculum mutations that became fixed in those individuals. If this were the case, this mutation may have been present at an undetectably low frequency in the original inoculum sample.

Overall, there was more genetic variation relative to the Wuhan-1 reference sequence in Delta samples as compared to WA01 samples, but that variation remained almost identical after passage of the Delta virus through animals. In the original Delta inoculum sample, passaged only once in Vero cells, 40 mutations were identified relative to the Wuhan-1 reference sequence. All but two of these were fixed in the inoculum and all animal samples (Figure 5.3c). The ORF8 deletion mutation D119_F120del was present at 80.3% AF in the Delta inoculum, and ranged from 77-88% in Delta animals. The ORF1a P2287S mutation was fixed in the inoculum but was decreased to 98% and 99% AF in one mule deer and one striped skunk (Skunk E).

5.2.4 Within-host population dynamics and signatures of selection

To characterize within-host virus populations, we calculated nucleotide diversity (π) for WA01- and Delta-infected wildlife samples and their respective inoculums. Overall, the samples from WA01 animals had significantly higher measures of π as compared to Delta samples (Figure 5.4a; t-test, t = -8.1, df = 26, p-value = 1.5x10-8). This pattern held true for synonymous nucleotide diversity, or πS, as well (Figure 5.4b; t-test, t = -6.8, df = 26, p-value = 3.2x10-7).

In general, much more variation in measurements of nucleotide diversity was observed in WA01 samples, both within and between species. Many WA01 animal samples also demonstrated slight increases in π and πS as compared to the WA01 inoculum. The most variation in π in Delta

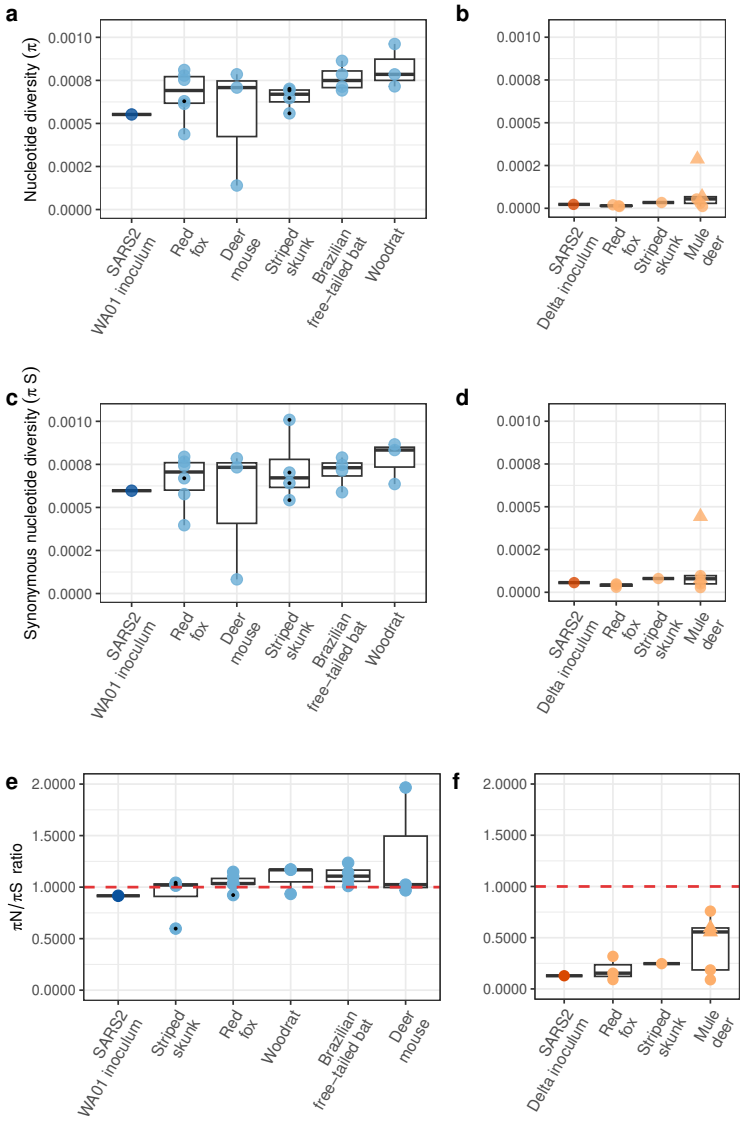


Figure 5.4: Higher baselines for (**a,b**) genomic nucleotide diversity (π), (c,d) relative effective population size (πS), and (**e,f**) selection (πN/πS ratio) were observed in SARS-CoV-2 genomes recovered from WA01-infected as compared to Delta-infected wildlife species; all Delta samples showed signatures of purifying selection. In all plots, each point indicates a value of π, πS or πN/πS obtained from NGS data for SARS-CoV-2 inoculum or animal samples. Measurements of nucleotide diversity were calculated with the SNPGenie pipeline. Points are colored according to SARS-CoV-2 variant used for infection (WA01=blue, Delta=orange). A black circle (•) in the center of a point indicates a virus sample from a WA01 striped skunk or red fox that showed signs of a host barrier to infection. Triangles indicate mule deer infected through contact transmission. For WA01 samples (**a,c,e**), points represent the mean of two technical replicates; for the Delta inoculum (see **b,d,f**), each point represents the mean of three technical replicates. (**a, b**) Genomic nucleotide diversity for (**a**) WA01 and (**b**) Delta virus samples. (**c, d**) Nucleotide diversity at synonymous sites, a relative measure of effective population size, for (**c**) WA01 and (**d**) Delta virus samples. (**e, f**) The ratio of nonsynonymous (πN) to synonymous (πS) nucleotide diversity for (**e**) WA01 and (**f**) Delta virus samples. A πN/πS ratio of 1 is shown with a red dotted line. A πN/πS ratio >1 indicates positive selection; a πN/πS ratio < 1 indicates purifying selection.

samples was observed in virus samples from deer, explained in large part by the potential de novo mutations identified in the two contact deer. Measures of πS, representing relative virus effective population size, also varied more in deer, which is consistent with a varied timing of sample collection ranging from 2-7dpi. However, no correlation was observed between π and dpi (linear model, R^2 = 0.0030, f-value = 0.089, df = 29, p-value = 0.77) or πS and dpi (linear model, R^2 = 0.0016, f-value = 0.045, df = 29, p-value = 0.83).

To assess the type and strength of selection acting on within-host virus populations, we calculated a πN/πS ratio for each sample (Figure 5.4c). All samples with genomic signatures of positive selection (πN/πS > 1) came from the WA01 dataset (n=16 samples). This represented 80% of all WA01 infections in this study (n=20 WA01-infected animals). However, the WA01 inoculum itself had an overall weak signature of purifying selection (πN/πS < 1), along with four WA01 samples from a deer mouse, a bushy-tailed woodrat, a red fox, and a striped skunk. All Delta animal samples (n=10 samples) and the viral inoculum showed signatures of purifying selection.

5.2.5 Comparison to previous passage of WA01 isolate three times in in Vero E6 cells

To further interrogate the differences observed between WA01 and Delta lineage infections, we re-examined data generated by our laboratory in which another isolate of USA-WA01/2020 (GenBank: MN985325.1) was passaged three times in Vero E6 cells (Chapter 2 of this dissertation)26. Samples were collected following each passage (P1, P2, and P3) and SARS-CoV-2 genomes were sequenced at a high depth of coverage. Interestingly, after just one passage we observed a total of eight mutations relative to the MN985325.1 reference sequence, which increased to 15 in P2, then back to eight in P3.

Three of these variants were WA01 inoculum variants also observed in our study: S D215H, M T7I, and N S194T, which increased from 4.2-14.6% AF in P1, to near fixation (95.9-96.7% AF) in P3. Notably, both ORF1b D126E and S R685H, detected in our study only in animal samples, were also detected in our previous study in P2 at 9.1-10.3% AF, increasing to 93.5-94.0% AF in P3, but neither were detected in P1. This provides further evidence that both ORF1b D126E and S R685H were present in low levels in the WA01 viral inoculum used in this study.

The difference between measurements of π and πS for the WA01 and Delta inoculum samples in our study may be explained in part by virus passage and expansion in cell culture. WA01 was passaged twice in Vero cells, whereas Delta was passaged once. In our previous work, we observed an increase in π from P1 to P2 in Vero E6, although this did not hold through P3 (Figure 5.5a). We also saw an increase in πS from P1 to P3, consistent with an expanding virus population size (Figure 5.5b). Finally, we observed positive selection in P1 and P2, but purifying selection in P3 (Figure 5.5c). This stands in contrast to the current study, in which the πN/πS ratio was less than one in both WA01 (P2) and Delta (P1), with the strongest signature of purifying selection in the Delta inoculum (Figure 5.4c). This provides further support for the conclusion that differences observed in our study result from biological differences between variant lineages.

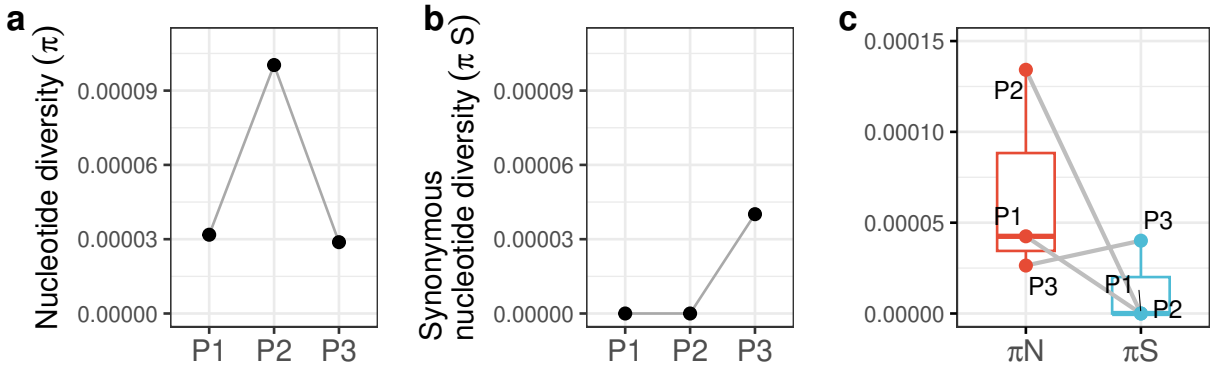


Figure 5.5: A SARS-CoV-2 WA01 isolate showed (**a**) varying nucleotide diversity (π), (**b**) increased relative effective population size (πS), and (**c**) signatures of positive and then purifying selection following three passages (P1, P2, P3) in Vero E6 cells. Data shown in these plots were previously obtained and published by the authors, and re-analyzed for this study(Bashor et al., 2021). Measurements of nucleotide diversity were calculated from SARS-CoV-2 NGS datasets using the SNPGenie pipeline. Each point represents the mean of two technical sequencing replicates. (**a**) Genomic nucleotide diversity (π) increased from P1 to P2, then decreased in P3. (**b**) Synonymous nucleotide diversity (πS), a measurement of the relative effective population size, increased from P1 to P3. (**c**) Because πS = 0 for P1 and P2, πN and πS were shown separately (πN > πS indicates positive selection; πN < πS indicates purifying selection).

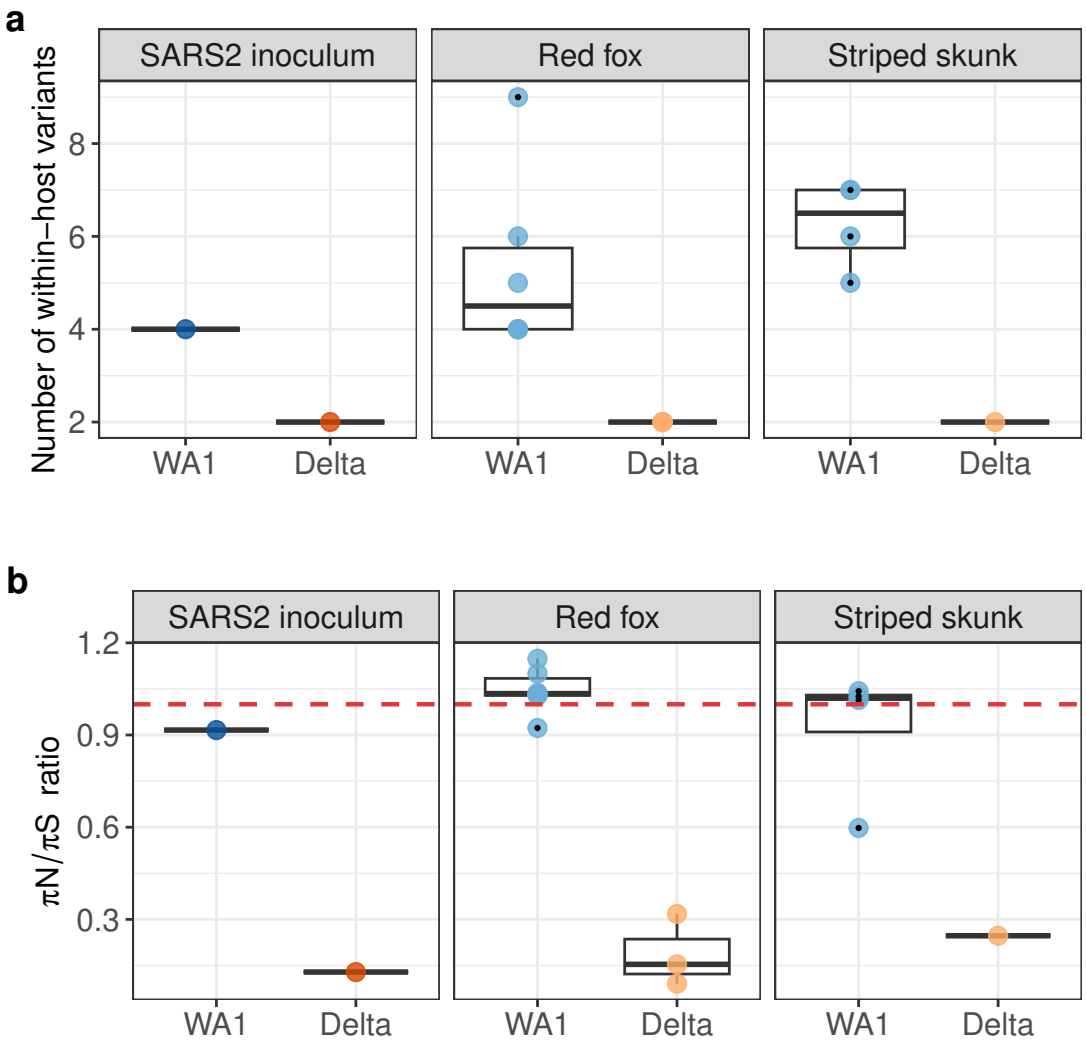


Figure 5.6: A direct comparison of SARS-CoV-2 genomes recovered from WA01- and Delta-infected striped skunks and red foxes reveals variant lineage-driven differences in (**a**) within-host variation and (**b**) the type and strength of selection (πN/πS ratio). In all plots, points are colored according to SARS-CoV-2 variant used for inoculation (WA01=blue, Delta=orange). A black circle (•) in the center of a point indicates a virus sample from a WA01 striped skunk or red fox that showed signs of a host barrier to infection. (**a**) Each point indicates the number of within-host variants in a WA01 and Delta inoculum and animal samples. Within-host variants include mutations relative to the Wuhan-1 reference sequence that were detected at least once in an animal or inoculum sample at variable allele frequency, or mutations that were fixed in an animal and not detected in the inoculum. (**b**) Each point indicates the ratio of nonsynonymous (πN) to synonymous (πS) nucleotide diversity calculated from NGS datasets for SARS-CoV-2 inoculum or animal samples using the SNPGenie pipeline. For all WA01 samples, data shown are the mean of two technical replicates; for the Delta inoculum, data shown are the mean of three technical replicates. A πN/πS ratio of 1 is shown with a red dotted line. A πN/πS ratio >1 indicates positive selection; a πN/πS ratio < 1 indicates purifying selection.

5.2.6 Direct comparison of WA01- and Delta-infected striped skunks and red foxes

A direct comparison of virus sequencing data from striped skunks and red foxes, which were infected with both variant lineages, provides further evidence that the baseline characteristics of the inoculum variant lineage was a driver of virus within-host population dynamics. Between WA01- and Delta-infected animals, no mutations were shared among members of a species. Virus samples from WA01-infected striped skunks and red foxes had more within-host variants compared to Delta-infected animals of the same species (Figure 5.6a). Similarly, both π, πS and πN/πS ratios were higher in WA01 striped skunks and red foxes (Figure 5.6b).

Interestingly, signs of a bottleneck or selective sweep following WA01 infection were present in both species. Furthermore, both striped skunks and red foxes infected with the WA01 lineage saw the emergence of new mutations, but Delta-infected animals did not (Table 5.3). Skunks A and C had two potential de novo mutations each, Skunk D had one; Foxes A and D had one mutation each, Fox C had four).

5.2.7 Potential effect of host species body temperature on virus within-host population dynamics

Based on our observation of variation in SARS-CoV-2 within-host population dynamics among different wildlife species, we carried out preliminary investigations based on the hypothesis that host body temperature can impact several parts of the viral life cycle. The average body temperature of the species in this study ranges from 36.0 to 40.0°C (Figure 5.7). Notably, higher body temperature was significantly associated with higher levels of viral RNA (lower RT-PCR Ct), and decreasing values of π, πS and πN/πS ratio in our sequencing dataset (linear models, Ct: R^2 = 0.32, f-value = 13.6, df = 29, p-value = 0.00091; π: R^2 = 0.37, f-value = 16, df = 27, p-value = 0.00044; πS: R^2 = 0.33, f-value = 13.2, df = 27, p-value = 0.0012; πN/πS: R^2 = 0.28, f-value = 10.4, df = 27, p-value = 0.0032).

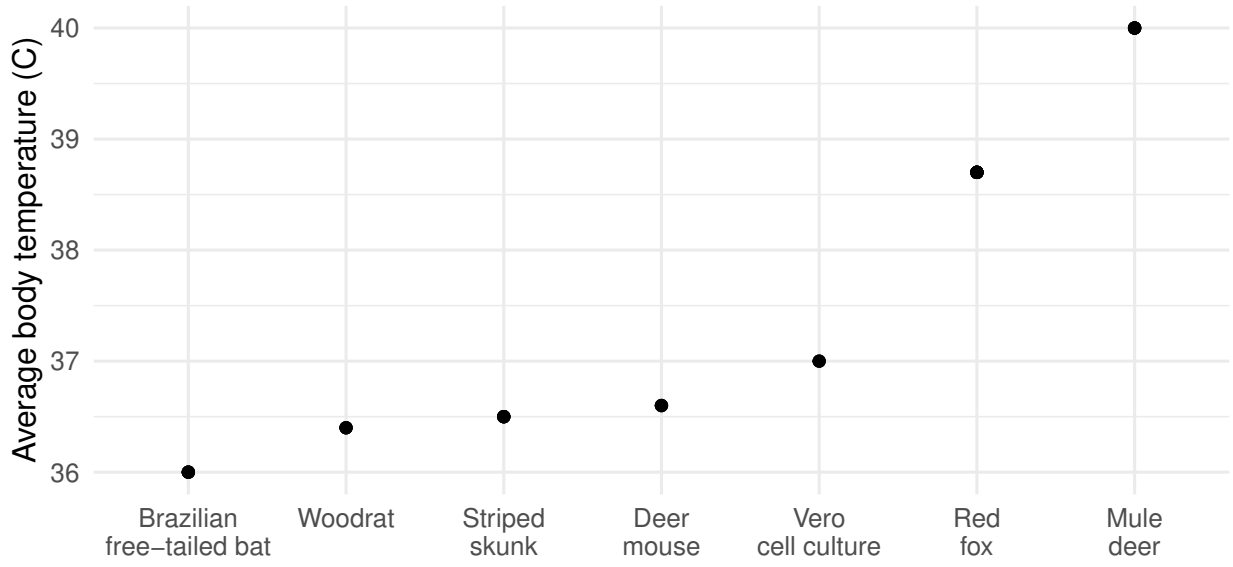


Figure 5.7: Host body temperature may have an impact on SARS-CoV-2 within-host variation and evolution. Each point indicates the typical body temperature of the six wildlife species in this study or the temperature of Vero cell culture. Body temperatures were obtained from the AnAge database: `https://genomics.senescence.info/species/index.html`. The temperature indicated for woodrats represents the mean body temperature of white-throated and desert woodrats, two *Neotoma* species in the AnAge database whose geographic ranges overlap with the bushy-tailed woodrat. Supporting references are available on the AnAge database page for each species.

5.3 Discussion

In this study, we demonstrated the application of experimental infection systems to study virus adaptation following spillover into novel host species. Twenty-nine individuals belonging to six North American wildlife species were experimentally exposed to one of two SARS-CoV-2 variants: WA01 and Delta. We observed virus host-adaptation, host barriers to infection, and the emergence of six novel mutations in mule deer following onward deer-to-deer transmission (Figure 5.8).

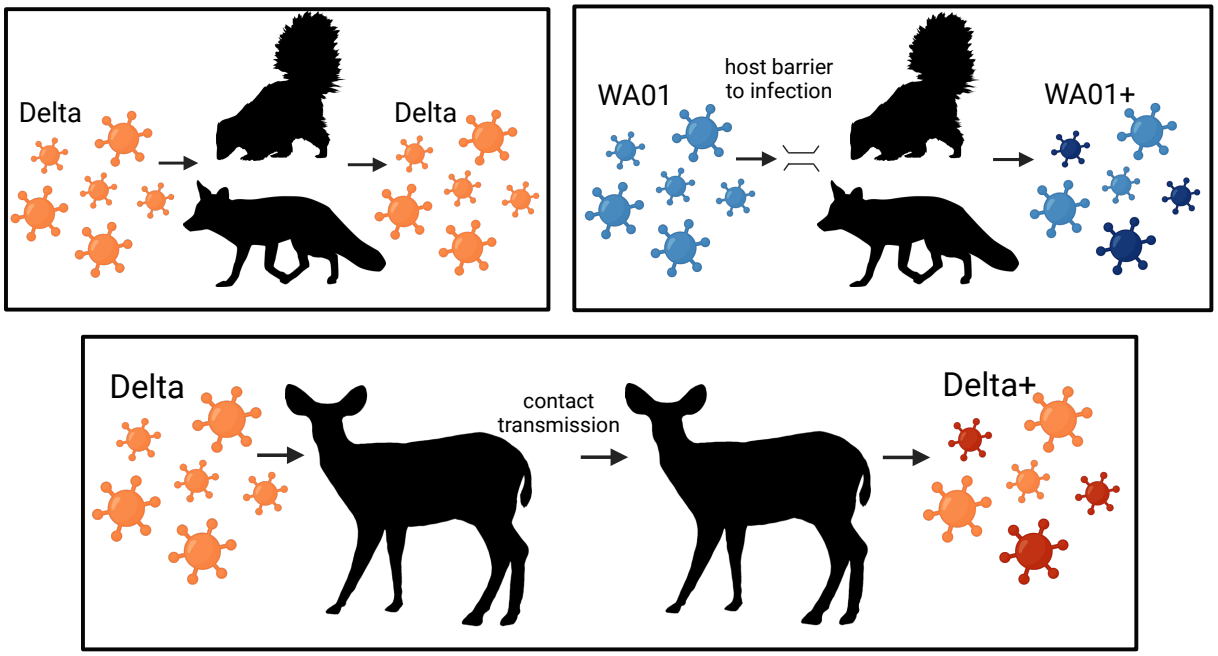


Figure 5.8: Virus adaptation and within-host variation differed based on the SARS-CoV-2 variant used in experimental exposures of wildlife, potentially due in part to host barriers to infection. Six novel SARS-CoV-2 mutations emerged in mule deer following deer-to-deer transmission. A graphical model demonstrates how experimental exposures can be used to reveal dynamics of SARS-CoV-2 evolution following cross-species transmission. Graphic was created with `BioRender.com`. Silhouette images were obtained from `https://www.phylopic.org/`. *Vulpes vulpes* by Anthony Caravaggi and *Odocoileus hemionus* by Gabriela Palomo-Munoz are available under CC BY-NC 3.0 Deed license: `https://creativecommons.org/licenses/by-nc/3.0/`.

Although the maintenance and evolution of SARS-CoV-2 in WTD populations in North America have been extensively documented, there are limited reports of natural infections in mule deer

populations. Our findings support the conclusion that mule deer are competent SARS-CoV-2 reservoir hosts, and virus maintenance in mule deer populations results in the emergence of novel genetic variation. We identified six de novo viral variants emerging in mule deer infected through contact with directly inoculated deer. Of 603 SARS-CoV-2 sequences derived from *Odocoileus* species in the GISAID database, only two are from mule deer (collected in California in September 2021 and Arizona in October 2022) (Shu & McCauley, 2017). These two mule deer sequences do not share any novel mutations with the sequences generated in this study, but one of them is classified as a Delta variant (AY.122 lineage).

SARS-CoV-2 sequences recovered from WTD are often highly divergent from each other and from sequences from humans; furthermore, the virus may have a higher evolutionary rate in WTD as compared to humans (McBride et al., 2023). We noted the rapid emergence of 2-4 de novo mutations per mule deer following deer-to-deer transmission, and a change in the virus genomic signature of selection relative to directly inoculated animals. Interestingly, two of these mutations were synonymous changes in the gene encoding the viral RdRp, the most highly conserved protein in RNA viruses. These findings suggest free-ranging mule deer populations exposed to SARS-CoV-2 could be a source of novel viral variants.

We hypothesized that the inoculum virus would undergo rapid change in response to strong selective pressures following establishment in a less permissive host environment. Brazilian free-tailed bats are considered only minimally susceptible to SARS-CoV-2 infection based on the absence of detectable shedding of infectious virus or neutralizing antibody response following experimental exposure. We thus predicted that we would detect increased SARS-CoV-2 adaptation, characterized by the emergence of within-host viral variation and genomic signatures of selection, in this species. Our findings refuted this hypothesis; we did not observe novel within-host variation in viruses recovered from bats, and signs of selection were comparable to other WA01-infected species.

Limitations to this study included variation in experimental exposures, sample collection methods, and sequencing approach. Importantly, the extent to which an NGS dataset is representative

of the actual virus population sampled is determined in part by the depth of sequencing coverage. Although we aimed for a high depth of coverage, WA01 samples in particular were sequenced at a lower depth of coverage than Delta samples. However, WA01 samples were sequenced in technical duplicates, allowing us to control for false positive variants resulting from PCR amplification. Across the variables measured in our dataset, the difference in sequencing coverage may confound differences observed between the two variant lineages. However, it may be relevant to consider that within the WA01 dataset, there were two outliers in sequencing coverage: Deer mouse A had more than three times higher, whereas Fox A had less than half the sequencing coverage of most other WA01 samples respectively. If sequencing coverage was driving the differences observed between WA01 and Delta variant infections, we would expect that Deer mouse A would look very different from Fox A, and more similar to Delta samples. However, Deer mouse A had high within-host variation and a strong signature of positive selection, very unlike Delta samples. Furthermore, together, Deer mouse A and Fox A had the lowest measures of π and πS across all WA01 samples.

Our analysis revealed clear and unanticipated differences in infection dynamics based on inoculum, which may have resulted from biological differences between the WA01 and Delta variant lineages. Infections of deer mice, striped skunks, and red foxes were characterized by the emergence of increased within-host variation relative to the WA01 inoculum. In contrast, the only novel within-host variation in Delta-infected animals emerged after deer-to-deer contact transmission. Overall, virus samples from WA01-infected animals had significantly more within-host variants than those from Delta-infected animals. Furthermore, the majority (80%) of virus samples from WA01-infected animals showed signs of positive selection, whereas all Delta samples showed signs of purifying selection. Some differences in replication kinetics in cell culture between SARS-CoV-2 ancestral variants and newer variants of concern (VOC) have previously been noted (Mlcochova et al., 2021; Mautner et al., 2022). In general, the Delta variant has a distinct replication advantage in human airway epithelial systems 32. However, in a direct comparison among variants using Vero E6 cells, a non-VOC (B.1.1 lineage) variant replicated faster, and the number of infectious

particles increased more quickly than any VOCs tested, including the Delta (B.1.617.2 lineage) and Omicron variants (Mautner et al., 2022).

In addition to the disparity in sequencing coverage, there was also a difference in culture conditions between the two variants. WA01 was passaged twice in Vero cells, whereas Delta was passaged once, and this was reflected in the higher levels of viral RNA detected by RT-PCR in the WA01 inoculum. This may account in part for the larger relative effective population size in WA01 compared to Delta. However, overall, more or less viral RNA in the inoculum didn't correspond directly to higher or lower levels of viral RNA in an infected animal. This supports our conclusion that the difference in baseline population metrics between WA01 and Delta infections may be due in part to biological differences between the two variant lineages, not only to expansion in cell culture.

Future studies are required to determine how our findings extend to the Omicron variant. At the time of this publication, SARS-CoV-2 continues to evolve off of the genetic backbone of Omicron, which swept to dominance at the end of 2021, and thus future exposures of naïve wildlife will likely involve this variant. Of note, preliminary experimental data indicates that striped skunks, red foxes and deer mice are less susceptible to the Omicron and do not shed infectious virus (A. Bosco-Lauth, 2024). However, pre-Omicron SARS-CoV-2 variants have been detected in WTD populations long after they were no longer circulating in human populations (Feng et al., 2023; Marques et al., 2022; Caserta et al., 2023; Vandegrift et al., 2022). The findings of this study may therefore remain relevant, and it is of particular importance to determine if mule deer populations have the same reservoir capacity. Increased disease surveillance of mule deer and other periodomestic wildlife populations, including NGS analysis, is of key importance. Although sampling free-ranging deer populations poses some challenges, annual harvests provide a unique opportunity for disease surveillance. This has already been extensively utilized for SARS-CoV-2 testing of WTD collected as part of statewide chronic wasting disease (CWD) surveillance programs (Pickering et al., 2022; Chandler et al., 2021; Kuchipudi et al., 2022; Caserta et al., 2023).

Furthermore, with appropriate handling and storage, samples collected for PCR-based diagnostic testing are also usable for NGS.

Of all species investigated in this study, mule deer had the highest average body temperature at 40°C. Conversely, striped skunks are documented to enter daily torpor, during which time their body temperatures drops lower than any other carnivore species, as low as 26°C (Hwang, Larivière, & Messier, 2007). In vitro, temperature increases to 40°C have been shown to restrict SARS-CoV-2 replication, and SARS-CoV-2 appears to replicate best at lower temperatures that are more reflective of the human upper respiratory airway (Herder et al., 2021; V'kovski et al., 2021). Ultimately, future investigations may benefit from focusing on the temperatures of the oral and nasal cavities and airways of different species, including striped skunks during and out of torpor, as possible determinants of SARS-CoV-2 replication and evolution.

Another unanticipated finding of this study was an apparent selective sweep or bottleneck following direct inoculation of all four striped skunks and one red fox with the WA01 variant, leading to the complete fixation of all inoculum variants in these animals. This was somewhat unexpected, given that all animals received a considerable intranasal dose of virus. It is not as obvious whether a similar phenomenon affected the striped skunk and three red foxes infected with the Delta variant. There was little-to-no within-host variation in the Delta inoculum for selection or a bottleneck to act upon in the first place. However, the Delta within-host variant ORF8 D119_F120del was observed at a similar AF (80%) in the inoculum and infected animals. This finding suggests that striped skunks may possess a host species-specific barrier to natural SARS-CoV-2 infection and transmission. As this occurred in an experimental rather than natural setting, it suggests that the barrier may be genetic, rather than ecological in nature. Further investigations of the type and extent of host barrier to natural infection are of high importance and relevance to understanding fundamental aspects of SARS-CoV-2 biology, in addition to its evolution and transmission in non-human animals generally.

5.4 Methods

5.4.1 Animal exposures, inoculums, and sample collection

Experimental exposures of animals were carried out as previously described (A. M. Bosco-Lauth et al., 2021, 2022; Porter et al., 2022, 2024). Briefly, deer mice, bushy-tailed woodrats, and striped skunks were exposed to an ancestral variant of SARS-CoV-2, WA01 as part of a larger study of the susceptibility of nine species of peridomestic wildlife (A. M. Bosco-Lauth et al., 2021). Red foxes were also exposed to the WA01 variant along with coyotes (Porter et al., 2022). Brazilian free-tailed bats were exposed to the WA01 and Delta variants, but only WA01-infected bat samples were included in this study (A. M. Bosco-Lauth et al., 2022). Mule deer were exposed to the Delta variant along with elk (Porter et al., 2024). The exposures of three red foxes and one striped skunk to the Delta variant have not yet been published (A. Bosco-Lauth, 2024). Swab samples from cottontail rabbits, fox squirrels, Wyoming ground squirrels, black-tailed prairie dogs, house mice, racoons, and coyotes were screened for potential for SARS-CoV-2 genome sequencing by tiled amplicon PCR (see below), but were not included in downstream steps due to insufficient virus genetic material. The animal IDs used in this study and their corresponding IDs in published studies are indicated in Table D.3.

The virus stocks used for inoculation were obtained from the Biodefense and Emerging Infections Research Resources Repository (BEI Resources). The WA1/2020WY96 isolate used for inoculation was passaged twice in Vero cells, whereas the Delta hCoV-19/USA/MD-HP05647/2021 isolate was passaged once in Vero cells prior to inoculation. All animals in this study were infected by intranasal inoculation except for two mule deer, who were infected through contact transmission as previously described (Porter et al., 2024). Briefly, mule deer were cohoused in two groups of three animals. Two deer in each group were directly inoculated, with a third animal as a contact. Contact Mule deer B was co-housed with Mule deer E, and the second directly inoculated deer in this group did not shed infectious virus. Mule deer C was co-housed with Mule deer A and D. Following intranasal inoculations, SARS-CoV-2 dose was determined by back-titration of the inoculum on Vero cells.

All samples included in this study were oral swabs except for the samples from mule deer, which were nasal swabs. Oral swab samples from all animals infected with WA01 were collected at 3 days post-infection (dpi) except for Bat D collected at 2 dpi. Oral swab samples from red foxes and one striped skunk infected with Delta were collected at 2 dpi, and nasal swab samples from three directly inoculated deer were collected between 2-5 dpi, and two contact-exposed deer at 3 and 7 dpi. Swab samples were stored in BA-1 medium (Tris-buffered minimum essential media containing 1% bovine serum albumin) supplemented with 10% fetal bovine serum, gentamicin, amphotericin B, and penicillin/streptomycin at -80°C prior to RNA isolation.

5.4.2 RNA isolation and RT-PCR detection of SARS-CoV-2 RNA

All WA01 samples were inactivated by adding 100 μl TRIzol reagent (ThermoFisher Scientific) to 300 μl sample prior to RNA isolation with a modified Zymo RNA Clean & Concentrator-5 kit (Zymo Research) as previously described 25,26. RNA isolation of Delta samples was carried out according to the manufacturer's instructions with QiaAmp™ Viral RNA mini kits (Qiagen).

RT-PCR detection of SARS-CoV-2 RNA was performed with the EXPRESS One-Step SuperScript qRT-PCR Kit (ThermoFisher Scientific), following a protocol developed by the Ebel lab at Colorado State University targeting the N gene using US Centers for Disease Control (CDC) 2019-nCoV N1 primers/probe and positive control (IDT) (Gallichotte et al., 2021). RNA samples were diluted 1:2 in nuclease-free water prior to RT-PCR. Cycle threshold (Ct) values for each sample are reported as the mean of two technical replicates.

5.4.3 Tiled amplicon PCR, library preparation and NGS

Samples collected from WA01 studies were prepared for sequencing in technical duplicate, whereas Delta samples were prepared without replication, except for the Delta inoculum, which was prepared in triplicate. First, complementary DNA (cDNA) was generated in 10 μl reaction volumes according to the manufacturer's instructions with SuperScript IV Reverse Transcriptase with 50 μM random primers, 10 mM dNTP mix, RNase OUT (40 units/μl) and 5 μl RNA as input (ThermoFisher Scientific).

SARS-CoV-2 genomic material was amplified with a tiled amplicon PCR using the ARTIC V4.1 NCOV-2019 Panel (IDT) according to the following protocol: dx.doi.org/10.17504/protocols.io.j8nlk4b36g5r/v2. Conditions were modified for 12.5 μl reaction volumes and decreased to 30 cycles of amplification. PCR products from ARTIC V4.1 primer pools were combined in equal volumes, subjected to a 1X bead cleanup Ampure XP beads (Beckman Coulter), and quantified with the Qubit 1X dsDNA HS Assay kit (Qubit).

Sequencing libraries were prepared with the NEBNext Ultra II DNA Library Prep kit using 100ng DNA as input, and unique dual indexing with NEBNext Multiplex Oligos for Illumina (New England Biolabs) according to the manufacturer's instructions. All bead cleanups were performed with Ampure XP beads (Beckman Coulter). Sequencing libraries for WA01 samples were pooled evenly and sequencing was carried out in single run on a MiSeq instrument at the Colorado State University NGS facility. A MiSeq Reagent kit v2 (500 cycles) was used with a loading concentration of 7pM and 4% PhiX. The target sequencing depth was not achieved with WA01 samples due to an unsatisfactory MiSeq run (55.81% clusters passed filter, cluster density: 561). Sequencing libraries for Delta samples were pooled evenly and MiSeq 2x250bp sequencing was carried out in a single run by CU-Boulder Center for Microbial Exploration Sequencing Center.

High quality next-generation sequencing datasets were obtained from oral and nasal swab samples from deer mice (n=3), bushy-tailed woodrats (n=3), Brazilian free-tailed bats (n=4), red foxes (n=6), and striped skunks (n=4) exposed to an ancestral variant of SARS-CoV-2, WA01, and red foxes (n=3), striped skunk (n=1), and mule deer (n=5) exposed to the Delta variant. All WA01 samples were sequenced in duplicate except the second technical replicate for Woodrat B failed sequencing. The mean and median depths of coverage across the SARS-CoV-2 genome for WA01 samples was 1151X and 900X respectively. For Delta samples, the mean and median depths of coverage were 5832X and 5115X respectively (Table D.2).

5.4.4 Bioinformatics and statistical analyses

Sequencing datasets were input into the nf-core/viralrecon pipeline (version 2.6.0) for variant lineage classification and low-frequency variant calling (Patel et al., 2023; Ewels et al., 2020). Al-

though the experimental infection system allowed us to have a known sequence for the input virus for all animal infections, we used GenBank MN908947.3 (Wuhan-1) as a reference sequence. Using this reference facilitated cross-compatibility with the broader scientific literature on SARS-CoV-2 phylogenetics and mutation tracking, and also allowed us to make direct comparisons between the WA01 and Delta inoculums. GenBank MN908947.3 is the default reference sequence for nf-core/viralrecon. However, we input a lightly modified annotation file for MN908947.3, because the default GFF file does not distinguish between ORF1a and ORF1b and account for the -1 ribosomal frameshift. This file is available at `https://github.com/laurabashor/`.

Detailed quality control metrics and variant lineage assignments output by the pipeline are included in Table D.2. A complete, annotated low-frequency variant table is included as Table D.4. For all samples with technical replicates, variant data were further processed to keep only variants that were detected in both sequencing replicates, and to calculate the mean allele frequency of the two technical replicates.

More stringent quality control of low-frequency variants resulted in the removal of eight additional variants called by nf-core/viralrecon and reported in Table D.4. A28271C, S E673G, and ORF8 F120L were not supported by a high depth of sequencing coverage in our datasets. Furthermore, ORF1b D1914E, R1915G, R1915R, Y1916C, and P1917H appeared to be sequencing artefacts. All five mutations only appeared together between genomic positions 19209-19217, and all five were detected only in Fox A, Skunk B, Bat C, and Bat D. They were all located only in the same subset of forward sequencing reads adjacent to the ARTIC V4.1 SARS-CoV-2_64_LEFT primer, but not observed on the overlapping reverse reads.

NGS datasets were also analyzed with the viral_variant_caller pipeline: `https://github.com/stenglein-lab/viral_variant_caller`. The output from this pipeline used in this study included a file with the raw sequencing depth at every genomic position, and VCF files for each dataset. VCF files were input into the SNPGenie pipeline to calculate within-host population metrics (Nelson et al., 2015). Data were further analyzed and visualized using R Statistical Software (Team, 2024). R scripts used for analysis are available at `https://github.com/`

`laurabashor/`. T-tests, ANOVA, and linear models were performed in base R, and the 'emmeans' package was used to obtain Tukey-adjusted pairwise comparisons for modeling viral RNA levels as a function of host species.

Chapter 6

Conclusions and future directions

Non-human animals have played a central role in the origins and evolution of SARS-CoV-2. The studies presented in this dissertation highlight the complex dynamics of viral evolution in the context of host shifts, utilizing next-generation sequencing to observe SARS-CoV-2 evolution in action in a range of non-human animal species. Furthermore, a key innovation of this work is the integration of experimental, natural, and public datasets. Our findings suggest that frequent spillover and spillback events have the potential to contribute to the emergence of new viral variants in human populations.

Chapters 2 and 3 reveal rapid selection for SARS-CoV-2 mutations following experimental infection of cats, dogs, hamsters and ferrets. In the first study, selection for mutations in the spike gene demonstrates virus host-adaptation both in vivo and in vitro. In particular, the changes in viral stocks following expansion in cell culture seen in our study have wide-ranging implications for experimental virological research. Furthermore, mutations in the SARS-CoV-2 replicase associated with domestic dog infections suggest strong selective pressures associated with viral adaptation to a less permissive host. In Chapter 3 we focused on felid hosts, uncovering the emergence and onward transmission of within-host viral variation in experimentally infected domestic cats, in contrast to previously documented natural infections of cats. This study identifies specific variables associated with experimental infections that impact virus transmission and evolutionary dynamics, including both the dose and route of infection. More broadly, both chapters highlight the potential for domestic animals living in close contact with humans to contribute to the SARS-CoV-2 diversity seen in human populations. Given the certain prospect of future emerging zoonoses, our findings emphasize the importance of considering societal norms surrounding within-household transmissions of pathogens between humans and companion animals.

In Chapters 4 and 5, we examine SARS-CoV-2 within-host evolution in naturally and experimentally infected wildlife species. Chapter 4 presents the trajectory of SARS-CoV-2 within-host

evolution in tigers, lions and hyenas at the Denver Zoo following a single spillover event. This study represents a unique dataset of longitudinal samples enabling direct interspecific comparisons of within-host viral dynamics, and our findings reveal virus population expansion, diversification and species-specific adaptation within and between hosts. In Chapter 5, we note distinct patterns of within-host evolution following experimental exposures of six species of wildlife with two different SARS-CoV-2 variants. The findings of this study emphasize functional differences between ancestral SARS-CoV-2 and more contemporary variants, in addition to the complexity of within-host virus evolution in experimental animal infections. Both Chapters 4 and 5 underscore the need for increased surveillance and research in both free-ranging and captive wildlife populations.

Our findings lay the groundwork for several future research directions. First, mechanistic studies of specific SARS-CoV-2 mutations identified in our studies are needed to confirm functional impacts and species-specific adaptation. For example, these studies could detect changes in the intrinsic transmissibility of SARS-CoV-2 in a host species by measuring variables related to host cell entry, replication, infectivity and infectiousness, and immune escape. These studies may further identify how a specific mutation helps the virus overcome a host barrier to infection. Second, enhanced SARS-CoV-2 surveillance in domestic and peridomestic animal species is needed to detect viral reservoirs and track virus evolution in non-human animals. To date, research has focused on mink and white-tailed deer populations to the exclusion of other susceptible host species. Our work identifies companion animals, members of Felidae, mule deer and zoo animals as surveillance priorities due to their proximity to humans and potential to support rapid virus evolution.

Collectively, the research presented in this dissertation highlights the plasticity of SARS-CoV-2 evolution in response to host shifts, and highlights the potential for zoonotic virus transmission in non-human animals. Our work reveals that a range of host species possess the ability to contribute to the novel variant emergence. From the broadest perspective, our findings emphasize the importance of a One Health approach to increase our understanding and resilience to emerging infectious diseases. This body of work advances scientific understanding of both virus evolution and public health resilience to future pandemics.

References

Aguiló-Gisbert, J., Padilla-Blanco, M., Lizana, V., Maiques, E., Muñoz-Baquero, M., Chillida-Martínez, E., ... Rubio-Guerri, C. (2021, 5). First description of sars-cov-2 infection in two feral american mink (neovison vison) caught in the wild. *Animals*, *11*, 1422. doi: 10.3390/ani11051422

Aksamentov, I., Roemer, C., Hodcroft, E., & Neher, R. (2021, 11). Nextclade: clade assignment, mutation calling and quality control for viral genomes. *Journal of Open Source Software*, *6*, 3773.

Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., & Garry, R. F. (2020). The proximal origin of sars-cov-2. *Nature medicine*, *26*(4), 450–452.

Andrew, R. (2020). Preliminary genomic characterisation of an emergent sars-cov-2 lineage in the uk defined by a novel set of spike mutations. *Virological*.

APHIS, U. (2024a, 3). *One health: White tailed deer.* Retrieved from `https://www.aphis.usda.gov/sars-cov-2/wtd`

APHIS, U. (2024b, 3). *Sars-cov-2 projects and studies.* Retrieved from `https://www.aphis.usda.gov/sars-cov-2/projects-studies`

ARTICNetwork. (2020). *Artic network.* Retrieved from `https://artic.network/ncov-2019`

A.V.M.A. (2017). *Pet ownership and demographics sourcebook.* American Veterinary Medical Association. Retrieved from `https://www.avma.org`

Banerjee, A., Mossman, K., & Baker, M. L. (2021). Zooanthroponotic potential of sars-cov-2 and implications of reintroduction into human populations. *Cell Host & Microbe*, *29*(2), 160–164.

Barroso-Arévalo, S., Sánchez-Morales, L., Pérez-Sancho, M., Domínguez, L., & Sánchez-Vizcaíno, J. M. (2022). First detection of sars-cov-2 b. 1.617. 2 (delta) variant of concern in a symptomatic cat in spain. *Frontiers in veterinary science*, *9*, 841430.

Barrs, V. R., Peiris, M., Tam, K. W., Law, P. Y., Brackman, C. J., To, E. M., ... Sit, T. H. (2020, 12). Sars-cov-2 in quarantined domestic cats from covid-19 households or close contacts, hong kong, china - volume 26, number 12—december 2020 - emerging infectious diseases journal - cdc. *Emerging Infectious Diseases*, *26*, 3071-3074. Retrieved from https://wwwnc.cdc.gov/eid/article/26/12/20-2786_article doi: 10.3201/EID2612.202786

Bartlett, S. L., Diel, D. G., Wang, L., Zec, S., Laverack, M., Martins, M., ... Calle, P. P. (2021, 1). Sars-cov-2 infection and longitudinal fecal screening in malayan tigers (panthera tigris jacksoni), amur tigers (panthera tigris altaica), and african lions (panthera leo krugeri) at the bronx zoo, new york, usa. *https://doi.org/10.1638/2020-0171*, *51*, 733-744. doi: 10.1638/2020-0171

Bashor, L., Gagne, R. B., Bosco-Lauth, A., Stenglein, M., & VandeWoude, S. (2022, 11). Rapid evolution of sars-cov-2 in domestic cats. *Virus Evolution*, *8*. doi: 10.1093/ve/veac092

Bashor, L., Gagne, R. B., Bosco-Lauth, A. M., Bowen, R. A., Stenglein, M., & VandeWoude, S. (2021, 11). Sars-cov-2 evolution in animals suggests mechanisms for rapid variant selection. *Proceedings of the National Academy of Sciences*, *118*, e2105253118. Retrieved from https://www.pnas.org/content/118/44/e2105253118 doi: 10.1073/PNAS.2105253118

Baum, A., Fulton, B. O., Wloga, E., Copin, R., Pascal, K. E., Russo, V., ... others (2020). Antibody cocktail to sars-cov-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science*, *369*(6506), 1014–1018.

Boklund, A., Hammer, A. S., Quaade, M. L., Rasmussen, T. B., Lohse, L., Strandbygaard, B., ... others (2021). Sars-cov-2 in danish mink farms: course of the epidemic and a descriptive analysis of the outbreaks in 2020. *Animals*, *11*(1), 164.

Bosco-Lauth, A. (2024). *Unpublished data.*

Bosco-Lauth, A. M., Hartwig, A. E., Porter, S. M., Gordy, P. W., Nehring, M., Byas, A. D., ... Bowen, R. A. (2020). Experimental infection of domestic dogs and cats with sars-cov-2:

Pathogenesis, transmission, and response to reexposure in cats. *Proceedings of the National Academy of Sciences*, *117*(42), 26382–26388.

Bosco-Lauth, A. M., Porter, S. M., Fox, K. A., Wood, M. E., Neubaum, D., & Quilici, M. (2022, 8). Experimental infection of brazilian free-tailed bats (tadarida brasiliensis) with two strains of sars-cov-2. *Viruses*, *14*, 1809. doi: 10.3390/v14081809

Bosco-Lauth, A. M., Root, J. J., Porter, S. M., Walker, A. E., Guilbert, L., Hawvermale, D., ... Bowen, R. A. (2021, 8). Peridomestic mammal susceptibility to severe acute respiratory syndrome coronavirus 2 infection - volume 27, number 8—august 2021 - emerging infectious diseases journal - cdc. *Emerging Infectious Diseases*, *27*, 2073-2080. Retrieved from `https://wwwnc.cdc.gov/eid/article/27/8/21-0180_article` doi: 10.3201/EID2708.210180

Braun, K. M., Moreno, G. K., Halfmann, P. J., Hodcroft, E. B., Baker, D. A., Boehm, E. C., ... others (2021). Transmission of sars-cov-2 in domestic cats imposes a narrow bottleneck. *PLoS pathogens*, *17*(2), e1009373.

Brito, A. F., Semenova, E., Dudas, G., Hassler, G. W., Kalinich, C. C., Kraemer, M. U. G., ... Faria, N. R. (2022, 11). Global disparities in sars-cov-2 genomic surveillance. *Nature Communications*, *13*, 7003.

Calvet, G. A., Pereira, S. A., Ogrzewalska, M., Pauvolid-Corrêa, A., Resende, P. C., de Souza Tassinari, W., ... others (2021). Investigation of sars-cov-2 infection in dogs and cats of humans diagnosed with covid-19 in rio de janeiro, brazil. *PloS one*, *16*(4), e0250853.

Caserta, L. C., Martins, M., Butt, S. L., Hollingshead, N. A., Covaleda, L. M., Ahmed, S., ... Diel, D. G. (2023, 2). White-tailed deer (odocoileus virginianus) may serve as a wildlife reservoir for nearly extinct sars-cov-2 variants of concern. *Proceedings of the National Academy of Sciences*, *120*. doi: 10.1073/pnas.2215067120

Chan, J. F.-W., Zhang, A. J., Yuan, S., Poon, V. K.-M., Chan, C. C.-S., Lee, A. C.-Y., ... others (2020). Simulation of the clinical and pathological manifestations of coronavirus disease 2019 (covid-19) in a golden syrian hamster model: implications for disease pathogenesis

and transmissibility. *Clinical infectious diseases*, *71*(9), 2428–2446.

Chan, K. K., Tan, T. J., Narayanan, K. K., & Procko, E. (2021). An engineered decoy receptor for sars-cov-2 broadly binds protein s sequence variants. *Science Advances*, *7*(8), eabf1738.

Chandler, J. C., Bevins, S. N., Ellis, J. W., Linder, T. J., Tell, R. M., Jenkins-Moore, M., ... Shriner, S. A. (2021, 11). Sars-cov-2 exposure in wild white-tailed deer (odocoileus virginianus). *Proceedings of the National Academy of Sciences of the United States of America*, *118*. Retrieved from `https://www.pnas.org/content/118/47/e2114828118` doi: 10.1073/PNAS.2114828118/

Chaudhry, M. Z., Eschke, K., Hoffmann, M., Grashoff, M., Abassi, L., Kim, Y., ... Cicin-Sain, L. (2022, 3). Rapid sars-cov-2 adaptation to available cellular proteases. *Journal of Virology*, *96*. doi: 10.1128/jvi.02186-21

Chen, Z., Azman, A. S., Chen, X., Zou, J., Tian, Y., Sun, R., ... Yu, H. (2022, 4). Global landscape of sars-cov-2 genomic surveillance and data sharing. *Nature Genetics*, *54*, 499-507.

Chi, X., Yan, R., Zhang, J., Zhang, G., Zhang, Y., Hao, M., ... others (2020). A neutralizing human antibody binds to the n-terminal domain of the spike protein of sars-cov-2. *Science*, *369*(6504), 650–655.

Cingolani, P., Patel, V. M., Coon, M., Nguyen, T., Land, S. J., Ruden, D. M., & Lu, X. (2012). Using drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, snpsift. *Frontiers in genetics*, *3*, 35.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., ... Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, snpeff: Snps in the genome of drosophila melanogaster strain w1118; iso-2; iso-3. *fly*, *6*(2), 80–92.

Colitti, B., Bertolotti, L., Mannelli, A., Ferrara, G., Vercelli, A., Grassi, A., ... others (2021). Cross-sectional serosurvey of companion animals housed with sars-cov-2–infected owners, italy. *Emerging Infectious Diseases*, *27*(7), 1919.

Coutard, B., Valle, C., De Lamballerie, X., Canard, B., Seidah, N. G., & Decroly, E. (2020).

The spike glycoprotein of the new coronavirus 2019-ncov contains a furin-like cleavage site absent in cov of the same clade. *Antiviral research*, *176*, 104742.

Cross, P. C., Prosser, D. J., Ramey, A. M., Hanks, E. M., & Pepin, K. M. (2019). Confronting models with data: The challenges of estimating disease spillover. *Philosophical Transactions of the Royal Society B*, *374*(1782), 20180435.

Curukoglu, A., Ergoren, M. C., Ozgencil, F. E., Sayiner, S., Ince, M. E., & Sanlidag, T. (2021, 11). First direct human-to-cat transmission of the sars-cov-2 b.1.1.7 variant. *Australian Veterinary Journal*, *99*, 482-488. doi: 10.1111/AVJ.13109

Damas, J., Hughes, G. M., Keough, K. C., Painter, C. A., Persky, N. S., Corbo, M., ... others (2020). Broad host range of sars-cov-2 predicted by comparative and structural analysis of ace2 in vertebrates. *Proceedings of the National Academy of Sciences*, *117*(36), 22311–22322.

Davis, J. J., Long, S. W., Christensen, P. A., Olsen, R. J., Olson, R., Shukla, M., ... Musser, J. M. (2021, 12). Analysis of the artic version 3 and version 4 sars-cov-2 primers and their impact on the detection of the g142d amino acid substitution in the spike protein. *Microbiology Spectrum*, *9*. doi: 10.1128/Spectrum.01803-21

DeMaio, N., Walker, C., Borges, R., Weilguny, L., Slodkowicz, G., & Goldman, N. (2020). *Issues with sars-cov-2 sequencing data.* Retrieved from `https://virological.org/t/issues-with-sars-cov-2-sequencing-data/473`

Deng, J., Liu, Y., Sun, C., Bai, J., Sun, J., Hao, L., ... Tian, K. (2020). Sars-cov-2 serological survey of cats in china before and after the pandemic. *Virologica Sinica*, *35*(6), 846–848.

der Auwera, G. A. V., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., ... DePristo, M. A. (2013). From fastq data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*. Retrieved from `https://pubmed.ncbi.nlm.nih.gov/25431634/` doi: 10.1002/0471250953.bi1110s43

Despres, H. W., Mills, M. G., Schmidt, M. M., Gov, J., Perez, Y., Jindrich, M., ... Bruce, E. A.

(2023, 9). Surveillance of vermont wildlife in 2021–2022 reveals no detected sars-cov-2 viral rna. *Scientific Reports*, *13*, 14683. doi: 10.1038/s41598-023-39232-0

Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nature biotechnology*, *35*(4), 316–319.

Earnest, R., Hahn, A. M., Feriancek, N. M., Brandt, M., Filler, R. B., Zhao, Z., ... Grubaugh, N. D. (2023, 12). Survey of white-footed mice (peromyscus leucopus) in connecticut, usa reveals low sars-cov-2 seroprevalence and infection with divergent betacoronaviruses. *npj Viruses*, *1*, 10. doi: 10.1038/s44298-023-00010-4

Eckstrand, C. D., Baldwin, T. J., Rood, K. A., Clayton, M. J., Lott, J. K., Wolking, R. M., ... Baszler, T. (2021, 11). An outbreak of sars-cov-2 with high mortality in mink (neovison vison) on multiple utah farms. *PLOS Pathogens*, *17*, e1009952. doi: 10.1371/journal.ppat.1009952

Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., ... Nahnsen, S. (2020, 3). The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology*, *38*, 276-278.

Fagre, A., Lewis, J., Eckley, M., Zhan, S., Rocha, S. M., Sexton, N. R., ... others (2021). Sars-cov-2 infection, neuropathogenesis and transmission among deer mice: Implications for spillback to new world rodents. *PLoS Pathogens*, *17*(5), e1009585.

Faria, N. R., Claro, I. M., Candido, D., Franco, L. M., Andrade, P. S., Coletti, T. M., ... others (2021). Genomic characterisation of an emergent sars-cov-2 lineage in manaus: preliminary findings. *Virological*, *372*, 815–821.

Feng, A., Bevins, S., Chandler, J., DeLiberto, T. J., Ghai, R., Lantz, K., ... Wan, X.-F. (2023, 7). Transmission of sars-cov-2 in free-ranging white-tailed deer in the united states. *Nature Communications*, *14*, 4078. doi: 10.1038/s41467-023-39782-x

Ferasin, L., Fritz, M., Ferasin, H., Becquart, P., Corbet, S., Gouilh, M. A., ... Leroy, E. M. (2021, 11). Infection with sars-cov-2 variant b.1.1.7 detected in a group of dogs and cats with

suspected myocarditis. *Veterinary Record*, *189*, no-no. doi: 10.1002/VETR.944

Fernández-Bellon, H., Rodon, J., Fernández-Bastit, L., Almagro, V., Padilla-Solé, P., Lorca-Oró, C., ... others (2021). Monitoring natural sars-cov-2 infection in lions (panthera leo) at the barcelona zoo: Viral dynamics and host responses. *Viruses*, *13*(9), 1683.

Fritz, M., Rosolen, B., Krafft, E., Becquart, P., Elguero, E., Vratskikh, O., ... others (2020). High prevalence of sars-cov-2 antibodies in pets from covid-19+ households. *One Health*, *11*, 100192.

Fuentealba, N. A., Moré, G., Bravi, M. E., Unzaga, J. M., De Felice, L., Salina, M., ... others (2021). First detection and molecular analysis of sars-cov-2 from a naturally infected cat from argentina. *Veterinary Microbiology*, *260*, 109179.

Gallichotte, E. N., Quicke, K. M., Sexton, N. R., Fitzmeyer, E., Young, M. C., Janich, A. J., ... Ehrhart, N. (2021, 12). Early adoption of longitudinal surveillance for sars-cov-2 among staff in long-term care facilities: Prevalence, virologic and sequence analysis. *Microbiology Spectrum*, *9*. doi: 10.1128/Spectrum.01003-21

Garigliany, M., Van Laere, A.-S., Clercx, C., Giet, D., Escriou, N., Huon, C., ... Desmecht, D. (2020). Sars-cov-2 natural transmission from human to cat, belgium, march 2020. *Emerging infectious diseases*, *26*(12), 3069.

Gaudreault, N. N., Trujillo, J. D., Carossino, M., Meekins, D. A., Morozov, I., Madden, D. W., ... Richt, J. A. (2020, 1). Sars-cov-2 infection, disease and transmission in domestic cats. *Emerging Microbes & Infections*, *9*, 2322-2332. doi: 10.1080/22221751.2020.1833687

Giner, J., Villanueva-Saz, S., Tobajas, A. P., Pérez, M. D., González, A., Verde, M., ... others (2021). Sars-cov-2 seroprevalence in household domestic ferrets (mustela putorius furo). *Animals*, *11*(3), 667.

Goyal, A., Reeves, D. B., Cardozo-Ojeda, E. F., Schiffer, J. T., & Mayer, B. T. (2021, 2). Viral load and contact heterogeneity predict sars-cov-2 transmission and super-spreading events. *eLife*, *10*, 1-63. doi: 10.7554/ELIFE.63537

Grome, H. N., Meyer, B., Read, E., Buchanan, M., Cushing, A., Sawatzki, K., ... Dunn, J. (2022,

4). Sars-cov-2 outbreak among malayan tigers and humans, tennessee, usa, 2020. *Emerging infectious diseases*, *28*, 833-836. doi: 10.3201/eid2804.212219

Grubaugh, N. D., Gangavarapu, K., Quick, J., Matteson, N. L., De Jesus, J. G., Main, B. J., ... others (2019). An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using primalseq and ivar. *Genome biology*, *20*(1), 1–19.

Grubaugh, N. D., Hanage, W. P., & Rasmussen, A. L. (2020). Making sense of mutation: what d614g means for the covid-19 pandemic remains unclear. *Cell*, *182*(4), 794–795.

Gu, H., Chen, Q., Yang, G., He, L., Fan, H., Deng, Y.-Q., ... others (2020). Adaptation of sars-cov-2 in balb/c mice for testing vaccine efficacy. *Science*, *369*(6511), 1603–1607.

Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., ... Neher, R. A. (2018, 12). Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics*, *34*, 4121-4123. Retrieved from https://academic.oup.com/bioinformatics/article/34/23/4121/5001388 doi: 10.1093/bioinformatics/bty407

Hale, V. L., Dennis, P. M., McBride, D. S., Nolting, J. M., Madden, C., Huey, D., ... Bowman, A. S. (2021, 12). Sars-cov-2 infection in free-ranging white-tailed deer. *Nature 2021*, 1-8. Retrieved from https://www.nature.com/articles/s41586-021-04353-x doi: 10.1038/s41586-021-04353-x

Halfmann, P. J., Hatta, M., Chiba, S., Maemura, T., Fan, S., Takeda, M., ... Kawaoka, Y. (2020, 5). Transmission of sars-cov-2 in domestic cats. *https://doi.org/10.1056/NEJMc2013400*, *383*, 592-594. Retrieved from https://www.nejm.org/doi/full/10.1056/nejmc2013400 doi: 10.1056/NEJMC2013400

Hamer, S. A., Pauvolid-Corrêa, A., Zecca, I. B., Davila, E., Auckland, L. D., Roundy, C. M., ... others (2021). Sars-cov-2 infections and viral isolations among serially tested cats and dogs in households with infected owners in texas, usa. *Viruses*, *13*(5), 938.

Hammer, A. S., Quaade, M. L., Rasmussen, T. B., Fonager, J., Rasmussen, M., Mundbjerg, K., ... Bøtner, A. (2021, 2). Sars-cov-2 transmission between mink (neovison vison) and humans, denmark. *Emerging Infectious Diseases*, *27*, 547. doi: 10.3201/EID2702.203794

Hanberry, B. B. (2021, 10). Addressing regional relationships between white-tailed deer densities and land classes. *Ecology and Evolution*, *11*, 13570-13578. doi: 10.1002/ece3.8084

Harcourt, J., Tamin, A., Lu, X., Kamili, S., Sakthivel, S. K., Murray, J., ... others (2020). Severe acute respiratory syndrome coronavirus 2 from patient with coronavirus disease, united states. *Emerging infectious diseases*, *26*(6), 1266.

Harris, R. S., & Dudley, J. P. (2015, 5). Apobecs and virus restriction. *Virology*, *479-480*, 131-145. doi: 10.1016/j.virol.2015.03.012

Hedman, H. D., Krawczyk, E., Helmy, Y. A., Zhang, L., & Varga, C. (2021). Host diversity and potential transmission pathways of sars-cov-2 at the human-animal interface. *Pathogens*, *10*(2), 180.

Herder, V., Dee, K., Wojtus, J. K., Epifano, I., Goldfarb, D., Rozario, C., ... Boutell, C. (2021, 12). Elevated temperature inhibits sars-cov-2 replication in respiratory epithelium independently of ifn-mediated innate immune defenses. *PLOS Biology*, *19*, e3001065. doi: 10.1371/journal.pbio.3001065

Hoffmann, M., Zhang, L., Krüger, N., Graichen, L., Kleine-Weber, H., Hofmann-Winkler, H., ... Pöhlmann, S. (2021, 4). Sars-cov-2 mutations acquired in mink reduce antibody-mediated neutralization. *Cell Reports*, *35*, 109017. doi: 10.1016/j.celrep.2021.109017

Holmes, E. C., Goldstein, S. A., Rasmussen, A. L., Robertson, D. L., Crits-Christoph, A., Wertheim, J. O., ... Rambaut, A. (2021, 9). The origins of sars-cov-2: A critical review. *Cell*, *184*, 4848-4856. doi: 10.1016/J.CELL.2021.08.017

Hosie, M. J., Epifano, I., Herder, V., Orton, R. J., Stevenson, A., Johnson, N., ... others (2021). Detection of sars-cov-2 in respiratory samples from cats in the uk associated with human-to-cat transmission. *Veterinary Record*, *188*(8), no–no.

Huang, K., Zhang, Y., Hui, X., Zhao, Y., Gong, W., Wang, T., ... Jin, M. (2021, 5). Q493k and q498h substitutions in spike promote adaptation of sars-cov-2 in mice. *EBioMedicine*, *67*, 103381. doi: 10.1016/j.ebiom.2021.103381

Hwang, Y. T., Larivière, S., & Messier, F. (2007, 1). Energetic consequences and ecological sig-

nificance of heterothermy and social thermoregulation in striped skunks (mephitis mephitis). *Physiological and Biochemical Zoology*, *80*, 138-145. doi: 10.1086/509211

Italiya, J., Vacek, V., Matějů, P., Dering, C., Celina, S. S., Ndiaye, A., & Černý, J. (2023, 8). First detection of sars-cov-2 in white rhinoceros during a small-scale coronavirus surveillance in the bandia reserve, senegal. *Animals*, *13*, 2593. doi: 10.3390/ani13162593

Itokawa, K., Sekizuka, T., Hashino, M., Tanaka, R., & Kuroda, M. (2020). Disentangling primer interactions improves sars-cov-2 genome sequencing by multiplex tiling pcr. *PloS one*, *15*(9), e0239403.

Johnson, B. A., Xie, X., Bailey, A. L., Kalveram, B., Lokugamage, K. G., Muruato, A., ... others (2021). Loss of furin cleavage site attenuates sars-cov-2 pathogenesis. *Nature*, *591*(7849), 293–299.

Johnson, B. A., Zhou, Y., Lokugamage, K. G., Vu, M. N., Bopp, N., Crocquet-Valdes, P. A., ... Menachery, V. D. (2022, 6). Nucleocapsid mutations in sars-cov-2 augment replication and pathogenesis. *PLOS Pathogens*, *18*, e1010627. doi: 10.1371/journal.ppat.1010627

Karikalan, M., Chander, V., Mahajan, S., Deol, P., Agrawal, R. K., Nandi, S., ... Sharma, G. K. (2021). Natural infection of delta mutant of sars-cov-2 in asiatic lions of india. *Transboundary and Emerging Diseases*. doi: 10.1111/TBED.14290

Keller, M., Hagag, I. T., Balzer, J., Beyer, K., Kersebohm, J. C., Sadeghi, B., ... Groschup, M. H. (2021, 11). Detection of sars-cov-2 variant b.1.1.7 in a cat in germany. *Research in Veterinary Science*, *140*, 229-232. doi: 10.1016/J.RVSC.2021.09.008

Kemp, S. A., Collier, D. A., Datir, R. P., Ferreira, I. A. T. M., Gayed, S., Jahun, A., ... Gupta, R. K. (2021, 2). Sars-cov-2 evolution during treatment of chronic infection. *Nature 2021 592:7853*, *592*, 277-282. Retrieved from `https://www.nature.com/articles/s41586-021-03291-y` doi: 10.1038/s41586-021-03291-y

Klaus, J., Meli, M. L., Willi, B., Nadeau, S., Beisel, C., Stadler, T., ... others (2021). Detection and genome sequencing of sars-cov-2 in a domestic cat with respiratory signs in switzerland. *Viruses*, *13*(3), 496.

Koeppel, K. N., Mendes, A., Strydom, A., Rotherham, L., Mulumba, M., & Venter, M. (2022, 1). Sars-cov-2 reverse zoonoses to pumas and lions, south africa. *Viruses*, *14*. doi: 10.3390/v14010120

Kombe, A. J. K., Biteghe, F. A. N., Ndoutoume, Z. N., & Jin, T. (2022, 10). Cd8+ t-cell immune escape by sars-cov-2 variants of concern. *Frontiers in Immunology*, *13*. doi: 10.3389/fimmu.2022.962079

Korber, B., Fischer, W. M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., ... others (2020). Tracking changes in sars-cov-2 spike: evidence that d614g increases infectivity of the covid-19 virus. *Cell*, *182*(4), 812–827.

Kuchipudi, S. V., Surendran-Nair, M., Ruden, R. M., Yon, M., Nissly, R. H., Vandegrift, K. J., ... Kapur, V. (2022, 2). Multiple spillovers from humans and onward transmission of sars-cov-2 in white-tailed deer. *Proceedings of the National Academy of Sciences of the United States of America*, *119*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/35078920 doi: 10.1073/PNAS.2121644119

Kutter, J. S., de Meulder, D., Bestebroer, T. M., Lexmond, P., Mulders, A., Richard, M., ... Herfst, S. (2021, 3). Sars-cov and sars-cov-2 are transmitted through the air between ferrets over more than one meter distance. *Nature Communications*, *12*, 1653. doi: 10.1038/s41467-021-21918-6

Lam, T. T.-Y., Jia, N., Zhang, Y.-W., Shum, M. H.-H., Jiang, J.-F., Zhu, H.-C., ... others (2020). Identifying sars-cov-2-related coronaviruses in malayan pangolins. *Nature*, *583*(7815), 282–285.

Larsen, H. D., Fonager, J., Lomholt, F. K., Dalby, T., Benedetti, G., Kristensen, B., ... Mølbak, K. (2021, 2). Preliminary report of an outbreak of sars-cov-2 in mink and mink farmers associated with community spread, denmark, june to november 2020. *Eurosurveillance*, *26*, 2100009. doi: 10.2807/1560-7917.ES.2021.26.5.210009/

Lassaunière, R., Fonager, J., Rasmussen, M., Frische, A., Polacek, C., Rasmussen, T. B., ... Fomsgaard, A. (2021, 6). In vitro characterization of fitness and convalescent antibody

neutralization of sars-cov-2 cluster 5 variant emerging in mink at danish farms. *Frontiers in Microbiology*, *12*. doi: 10.3389/fmicb.2021.698944

Li, P., de Vries, A. C., Kamar, N., Peppelenbosch, M. P., & Pan, Q. (2022, 5). Monitoring and managing sars-cov-2 evolution in immunocompromised populations. *The Lancet Microbe*, *3*, e325-e326. doi: 10.1016/S2666-5247(22)00061-1/ATTACHMENT/179754E6-6A40-4F97-8A99-84BDE2F9F0F7/MMC1.PDF

Liu, W. J., Liu, P., Lei, W., Jia, Z., He, X., Shi, W., ... Wu, G. (2023, 4). Surveillance of sars-cov-2 at the huanan seafood market. *Nature*. doi: 10.1038/s41586-023-06043-2

Longdon, B., Hadfield, J. D., Day, J. P., Smith, S. C., McGonigle, J. E., Cogni, R., ... Jiggins, F. M. (2015). The causes and consequences of changes in virulence following pathogen host shifts. *PLoS pathogens*, *11*(3), e1004728.

Lu, L., Sikkema, R. S., Velkers, F. C., Nieuwenhuijse, D. F., Fischer, E. A., Meijer, P. A., ... Koopmans, M. P. (2021, 11). Adaptation, spread and transmission of sars-cov-2 in farmed minks and associated humans in the netherlands. *Nature Communications 2021 12:1*, *12*, 1-12. Retrieved from https://www.nature.com/articles/s41467-021-27096-9 doi: 10.1038/s41467-021-27096-9

Lythgoe, K. A., Hall, M., Ferretti, L., de Cesare, M., MacIntyre-Cockett, G., Trebes, A., ... others (2021). Sars-cov-2 within-host diversity and transmission. *Science*, *372*(6539), eabg0821.

Markov, P. V., Ghafari, M., Beer, M., Lythgoe, K., Simmonds, P., Stilianakis, N. I., & Katzourakis, A. (2023, 6). The evolution of sars-cov-2. *Nature Reviews Microbiology*, *21*, 361-379. doi: 10.1038/s41579-023-00878-2

Marks, M., Millat-Martinez, P., Ouchi, D., h. Roberts, C., Alemany, A., Corbacho-Monné, M., ... Mitjà, O. (2021, 5). Transmission of covid-19 in 282 clusters in catalonia, spain: a cohort study. *The Lancet Infectious Diseases*, *21*, 629-636. doi: 10.1016/S1473-3099(20)30985-3/

Marques, A. D., Sherrill-Mix, S., Everett, J. K., Adhikari, H., Reddy, S., Ellis, J. C., ... Anis, E. (2022, 10). Multiple introductions of sars-cov-2 alpha and delta variants into white-tailed deer in pennsylvania. *mBio*, *13*. doi: 10.1128/mbio.02101-22

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, *17*(1), 10–12.

Mautner, L., Hoyos, M., Dangel, A., Berger, C., Ehrhardt, A., & Baiker, A. (2022, 12). Replication kinetics and infectivity of sars-cov-2 variants of concern in common cell culture models. *Virology Journal*, *19*, 76. doi: 10.1186/s12985-022-01802-5

McAloose, D., Laverack, M., Wang, L., Killian, M. L., Caserta, L. C., Yuan, F., ... others (2020). From people to panthera: Natural sars-cov-2 infection in tigers and lions at the bronx zoo. *MBio*, *11*(5), 10–1128.

McBride, D. S., Garushyants, S. K., Franks, J., Magee, A. F., Overend, S. H., Huey, D., ... Bowman, A. S. (2023, 8). Accelerated evolution of sars-cov-2 in free-ranging white-tailed deer. *Nature Communications*, *14*, 5105. doi: 10.1038/s41467-023-40706-y

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... others (2010). The genome analysis toolkit: a mapreduce framework for analyzing next-generation dna sequencing data. *Genome research*, *20*(9), 1297–1303.

Meng, B., Kemp, S. A., Papa, G., Datir, R., Ferreira, I. A., Marelli, S., ... others (2021). Recurrent emergence of sars-cov-2 spike deletion h69/v70 and its role in the alpha variant b. 1.1. 7. *Cell reports*, *35*(13).

Michelitsch, A., Hoffmann, D., Wernike, K., & Beer, M. (2020). Occurrence of antibodies against sars-cov-2 in the domestic cat population of germany. *Vaccines*, *8*(4), 772.

Miller, M. R., Tkachenko, A., Guag, J., Alexander, S., Webb, B. T., & Stenger, B. L. S. (2024, 3). Comparative evaluation of assay performance for sars-cov-2 detection in animal oral samples, lung homogenates, and phosphate-buffered saline using the taqpath covid-19 combo kit. *Journal of Veterinary Diagnostic Investigation*, *36*, 229-237. doi: 10.1177/10406387241230315

Mishra, A., Kumar, N., Bhatia, S., Aasdev, A., Kanniappan, S., Sekhar, A. T., ... Singh, V. P. (2021, 10). Sars-cov-2 delta variant among asiatic lions, india. *Emerging infectious diseases*, *27*, 2723-2725. doi: 10.3201/eid2710.211500

Mitchell, P. K., Martins, M., Reilly, T., Caserta, L. C., Anderson, R. R., Cronk, B. D., ... Diel, D. G. (2021, 12). Sars-cov-2 b.1.1.7 variant infection in malayan tigers, virginia, usa. *Emerging Infectious Diseases*, *27*, 3171. doi: 10.3201/EID2712.211234

Mlcochova, P., Kemp, S. A., Dhar, M. S., Papa, G., Meng, B., Ferreira, I. A. T. M., ... Gupta, R. K. (2021, 11). Sars-cov-2 b.1.617.2 delta variant replication and immune evasion. *Nature*, *599*, 114-119. doi: 10.1038/s41586-021-03944-y

Munnink, B. B., Sikkema, R. S., Nieuwenhuijse, D. F., Molenaar, R. J., Munger, E., Molenkamp, R., ... Koopmans, M. P. (2021, 1). Transmission of sars-cov-2 on mink farms between humans and mink and back to humans. *Science*, *371*, 172-177. Retrieved from `https://pubmed.ncbi.nlm.nih.gov/33172935/` doi: 10.1126/science.abe5901

Muradyan, N., Arakelov, V., Sargsyan, A., Paronyan, A., Arakelov, G., & Nazaryan, K. (2024, 3). Impact of mutations on the stability of sars-cov-2 nucleocapsid protein structure. *Scientific Reports*, *14*, 5870. doi: 10.1038/s41598-024-55157-8

Nagy, A., Stará, M., Vodička, R., Černíková, L., Jiřincová, H., Křivda, V., & Sedlák, K. (2022, 8). Reverse-zoonotic transmission of sars-cov-2 lineage alpha (b.1.1.7) to great apes and exotic felids in a zoo in the czech republic. *Archives of virology*, *167*, 1681-1685. doi: 10.1007/s00705-022-05469-9

Neira, V., Brito, B., Agüero, B., Berrios, F., Valdés, V., Gutierrez, A., ... others (2021). A household case evidences shorter shedding of sars-cov-2 in naturally infected cats compared to their human owners. *Emerging Microbes & Infections*, *10*(1), 376–383.

Nelson, C. W., & Hughes, A. L. (2015). Within-host nucleotide diversity of virus populations: insights from next-generation sequencing. *Infection, Genetics and Evolution*, *30*, 1–7.

Nelson, C. W., Moncla, L. H., & Hughes, A. L. (2015, 11). Snpgenie: estimating evolutionary parameters to detect natural selection using pooled next-generation sequencing data. *Bioinformatics*, *31*, 3709-3711. doi: 10.1093/bioinformatics/btv449

Nerpel, A., Yang, L., Sorger, J., Käsbohrer, A., Walzer, C., & Desvars-Larrive, A. (2022, 7). Sars-ani: a global open access dataset of reported sars-cov-2 events in animals. *Scientific data*, *9*,

438. doi: 10.1038/s41597-022-01543-8

Newman, A., Smith, D., Ghai, R. R., Wallace, R. M., Torchetti, M. K., Loiacono, C., ... others (2020). First reported cases of sars-cov-2 infection in companion animals—new york, march–april 2020. *Morbidity and Mortality Weekly Report*, *69*(23), 710.

of Agriculture Animal, U. S. D., & Service, P. H. I. (2022). *Confirmed cases of sars-cov-2 in animals in the united states.* United States Department of Agriculture. Retrieved from `https://www.aphis.usda.gov/aphis/dashboards/tableau/sars-dashboard`

Ogando, N. S., Dalebout, T. J., Zevenhoven-Dobbe, J. C., Limpens, R. W., van der Meer, Y., Caly, L., ... Snijder, E. J. (2020). Sars-coronavirus-2 replication in vero e6 cells: replication kinetics, rapid adaptation and cytopathology. *The Journal of General Virology*, *101*, 925. doi: 10.1099/JGV.0.001453

Oreshkova, N., Molenaar, R. J., Vreman, S., Harders, F., Munnink, B. B. O., Hakze-van Der Honing, R. W., ... others (2020). Sars-cov-2 infection in farmed minks, the netherlands, april and may 2020. *Eurosurveillance*, *25*(23), 2001005.

Ou, T., Mou, H., Zhang, L., Ojha, A., Choe, H., & Farzan, M. (2021). Hydroxychloroquine-mediated inhibition of sars-cov-2 entry is attenuated by tmprss2. *PLoS pathogens*, *17*(1), e1009212.

Oude Munnink, B. B., Sikkema, R. S., Nieuwenhuijse, D. F., Molenaar, R. J., Munger, E., Molenkamp, R., ... others (2021). Transmission of sars-cov-2 on mink farms between humans and mink and back to humans. *Science*, *371*(6525), 172–177.

Palermo, P. M., Orbegozo, J., Watts, D. M., & Morrill, J. C. (2021, 12). Sars-cov-2 neutralizing antibodies in white-tailed deer from texas. *Vector-Borne and Zoonotic Diseases*. doi: 10.1089/vbz.2021.0094

Patel, H., Monzon, S., Varona, S., Espinosa-Carrasco, J., Garcia, M., Heuer, M. L., ... Kelly, S. (2023). *nf-core/viralrecon: nf-core/viralrecon v2.6.0 - rhodium raccoon.*

Patterson, E. I., Elia, G., Grassi, A., Giordano, A., Desario, C., Smith, S., ... others (2020).

Evidence of exposure to sars-cov-2 in cats and dogs from households in italy. *Nature communications*, *11*(1), 6231.

Pekar, J. E., Magee, A., Parker, E., Moshiri, N., Izhikevich, K., Havens, J. L., ... Wertheim, J. O. (2022, 8). The molecular epidemiology of multiple zoonotic origins of sars-cov-2. *Science*, *377*, 960-966. doi: 10.1126/science.abp8337

Perisé-Barrios, A. J., Tomeo-Martín, B. D., Gómez-Ochoa, P., Delgado-Bonet, P., Plaza, P., Palau-Concejo, P., ... others (2021). Humoral responses to sars-cov-2 by healthy and sick dogs during the covid-19 pandemic in spain. *Veterinary research*, *52*(1), 1–11.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). Ucsf chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, *25*(13), 1605–1612.

Pickering, B., Lung, O., Maguire, F., Kruczkiewicz, P., Kotwa, J. D., Buchanan, T., ... others (2022). Divergent sars-cov-2 variant emerges in white-tailed deer with deer-to-human transmission. *Nature Microbiology*, *7*(12), 2011–2024.

Porter, S. M., Hartwig, A. E., Bielefeldt-Ohmann, H., Bosco-Lauth, A. M., & Root, J. J. (2022, 9). Susceptibility of wild canids to sars-cov-2. *Emerging Infectious Diseases*, *28*, 1852-1855.

Porter, S. M., Hartwig, A. E., Bielefeldt-Ohmann, H., Marano, J. M., Root, J. J., & Bosco-Lauth, A. M. (2024, 2). Experimental sars-cov-2 infection of elk and mule deer. *Emerging Infectious Diseases*, *30*.

Prévost, J., & Finzi, A. (2021). The great escape? sars-cov-2 variants evading neutralizing responses. *Cell Host & Microbe*, *29*(3), 322–324.

Quick, J., Grubaugh, N. D., Pullan, S. T., Claro, I. M., Smith, A. D., Gangavarapu, K., ... others (2017). Multiplex pcr method for minion and illumina sequencing of zika and other virus genomes directly from clinical samples. *Nature protocols*, *12*(6), 1261–1276.

Rabalski, L., Kosinski, M., Mazur-Panasiuk, N., Szewczyk, B., Bienkowska-Szewczyk, K., Kant, R., ... Grzybek, M. (2022, 3). Zoonotic spill-over of sars-cov-2: mink-adapted virus in humans. *Clinical Microbiology and Infection*, *28*, 451.e1-451.e4. doi: 10.1016/j.cmi.2021

.12.001

Rambaut, A., Holmes, E. C., O'Toole, Á., Hill, V., McCrone, J. T., Ruis, C., ... Pybus, O. G. (2020, 7). A dynamic nomenclature proposal for sars-cov-2 lineages to assist genomic epidemiology. *Nature Microbiology 2020 5:11*, *5*, 1403-1407.

Rasmussen, T. B., Fonager, J., Jørgensen, C. S., Lassaunière, R., Hammer, A. S., Quaade, M. L., ... Bøtner, A. (2021, 11). Infection, recovery and re-infection of farmed mink with sars-cov-2. *PLOS Pathogens*, *17*, e1010068. doi: 10.1371/journal.ppat.1010068

Richard, M., Kok, A., de Meulder, D., Bestebroer, T. M., Lamers, M. M., Okba, N. M., ... others (2020). Sars-cov-2 is transmitted via contact and via the air between ferrets. *Nature communications*, *11*(1), 3496.

Roundy, C. M., Nunez, C. M., Thomas, L. F., Auckland, L. D., Tang, W., Richison, J. J., ... Hamer, S. A. (2022, 4). High seroprevalence of sars-cov-2 in white-tailed deer (odocoileus virginianus) at one of three captive cervid facilities in texas. *Microbiology Spectrum*, *10*. doi: 10.1128/spectrum.00576-22

Rudd, J. M., Tamil Selvan, M., Cowan, S., Kao, Y.-F., Midkiff, C. C., Narayanan, S., ... Miller, C. A. (2021). Clinical and histopathologic features of a feline sars-cov-2 infection model are analogous to acute covid-19 in humans. *Viruses*, *13*(8), 1550.

Ruiz-Arrondo, I., Portillo, A., Palomar, A. M., Santibáñez, S., Santibáñez, P., Cervera, C., & Oteo, J. A. (2021). 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. *Transboundary and emerging diseases*, *68*(2), 973–976.

Sailleau, C., Dumarest, M., Vanhomwegen, J., Delaplace, M., Caro, V., Kwasiborski, A., ... others (2020). First detection and genome sequencing of sars-cov-2 in an infected cat in france. *Transboundary and emerging diseases*, *67*(6), 2324–2328.

Sasaki, M., Uemura, K., Sato, A., Toba, S., Sanaki, T., Maenaka, K., ... Sawa, H. (2021). Sars-cov-2 variants with mutations at the s1/s2 cleavage site are generated in vitro during propagation in tmprss2-deficient cells. *PLoS pathogens*, *17*(1), e1009233.

Sauter, D., & Kirchhoff, F. (2019). Key viral adaptations preceding the aids pandemic. *Cell host & microbe*, *25*(1), 27–38.

Sawatzki, K., Hill, N. J., Puryear, W. B., Foss, A. D., Stone, J. J., & Runstadler, J. A. (2021). Host barriers to sars-cov-2 demonstrated by ferrets in a high-exposure domestic setting. *Proceedings of the National Academy of Sciences*, *118*(18), e2025601118.

Segalés, J., Puig, M., Rodon, J., Avila-Nieto, C., Carrillo, J., Cantero, G., ... others (2020). Detection of sars-cov-2 in a cat owned by a covid-19- affected patient in spain. *Proceedings of the National Academy of Sciences*, *117*(40), 24790–24793.

Shereen, M. A., Khan, S., Kazmi, A., Bashir, N., & Siddique, R. (2020). Covid-19 infection: Emergence, transmission, and characteristics of human coronaviruses. *Journal of advanced research*, *24*, 91–98.

Shi, J., Wen, Z., Zhong, G., Yang, H., Wang, C., Huang, B., ... others (2020). Susceptibility of ferrets, cats, dogs, and other domesticated animals to sars–coronavirus 2. *Science*, *368*(6494), 1016–1020.

Shriner, S. A., Ellis, J. W., Root, J. J., Roug, A., Stopak, S. R., Wiscomb, G. W., ... DeLiberto, T. J. (2021). Sars-cov-2 exposure in escaped mink, utah, usa. *Emerging infectious diseases*, *27*(3), 988.

Shu, Y., & McCauley, J. (2017). Gisaid: Global initiative on sharing all influenza data–from vision to reality. *Eurosurveillance*, *22*(13), 30494.

Siegrist, A. A., Richardson, K. L., Ghai, R. R., Pope, B., Yeadon, J., Culp, B., ... Boyer, L. V. (2023, 6). Probable transmission of sars-cov-2 from african lion to zoo employees, indiana, usa, 2021. *Emerging Infectious Diseases*, *29*. doi: 10.3201/eid2906.230150

Sila, T., Sunghan, J., Laochareonsuk, W., Surasombatpattana, S., Kongkamol, C., Ingviya, T., ... Chusri, S. (2022, 7). Suspected cat-to-human transmission of sars-cov-2, thailand, july–september 2021. *Emerging Infectious Diseases*, *28*, 1485-1488. doi: 10.3201/eid2807.212605

Sit, T. H., Brackman, C. J., Ip, S. M., Tam, K. W., Law, P. Y., To, E. M., ... others (2020).

Infection of dogs with sars-cov-2. *Nature*, *586*(7831), 776–778.

Sparrer, M. N., Hodges, N. F., Ragan, I., Yamashita, T., Reed, K. J., Sherman, T. J., ... Mayo, C. (2023, 10). Sars-cov-2 surveillance in a veterinary health system provides insight into transmission risks. *Journal of the American Veterinary Medical Association*, 1-7.

Stevanovic, V., Vilibic-Cavlek, T., Tabain, I., Benvin, I., Kovac, S., Hruskar, Z., ... others (2021). Seroprevalence of sars-cov-2 infection among pet animals in croatia and potential public health impact. *Transboundary and emerging diseases*, *68*(4), 1767–1773.

Tan, C. C. S., Lam, S. D., Richard, D., Owen, C. J., Berchtold, D., Orengo, C., ... Balloux, F. (2022, 5). Transmission of sars-cov-2 from humans to animals and potential host adaptation. *Nature communications*, *13*, 2988. doi: 10.1038/s41467-022-30698-6

Tang, X., Wu, C., Li, X., Song, Y., Yao, X., Wu, X., ... others (2020). On the origin and continuing evolution of sars-cov-2. *National science review*, *7*(6), 1012–1023.

Team, R. C. (2021). *R: A language and environment for statistical computing.*

Team, R. C. (2024). *R: A language and environment for statistical computing.*

Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., ... others (2021). Detection of a sars-cov-2 variant of concern in south africa. *Nature*, *592*(7854), 438–443.

Temmam, S., Barbarino, A., Maso, D., Behillil, S., Enouf, V., Huon, C., ... others (2020). Absence of sars-cov-2 infection in cats and dogs in close contact with a cluster of covid-19 patients in a veterinary campus. *One health*, *10*, 100164.

Thebault, R. (2021, Nov). *A zoo's three 'beloved' snow leopards die of covid- 19.* Retrieved from `https://www.washingtonpost.com/science/2021/11/14/snow-leopard-death-covid/`

Totton, S. C., Sargeant, J. M., & O'Connor, A. M. (2021, 2). How could we conclude cat-to-human transmission of sars-cov-2? *Zoonoses and Public Health*, *68*, 67-68. doi: 10.1111/ZPH.12788

United States Department of Agriculture Animal and Plant Health Inspection Service (USDA

APHIS). (2021a, July). *Department of agriculture animal and plant health inspection service (usda aphis) surveillance data shows white-tailed deer exposed to sars-cov-2.* https://www.aphis.usda.gov/aphis/newsroom/stakeholder-info/stakeholder-messages/wildlife-damage-news/deer-sars.

United States Department of Agriculture Animal and Plant Health Inspection Service (USDA APHIS). (2021b, August). *United states department of agriculture animal and plant health inspection service (usda aphis) confirmation of covid-19 in deer in ohio.* https://www.aphis.usda.gov/aphis/newsroom/stakeholder-info/sa_by_date/sa-2021/sa-08/covid-deer.

van Aart, A. E., Velkers, F. C., Fischer, E. A., Broens, E. M., Egberink, H., Zhao, S., ... Smit, L. A. (2021). Sars-cov-2 infection in cats and dogs in infected mink farms. *Transboundary and Emerging Diseases*. doi: 10.1111/TBED.14173

van Aart, A. E., Velkers, F. C., Fischer, E. A., Broens, E. M., Egberink, H., Zhao, S., ... others (2022). Sars-cov-2 infection in cats and dogs in infected mink farms. *Transboundary and Emerging Diseases*, *69*(5), 3001–3007.

Vandegrift, K. J., Yon, M., Nair, M. S., Gontu, A., Ramasamy, S., Amirthalingam, S., ... Kuchipudi, S. V. (2022, 12). Sars-cov-2 omicron (b.1.1.529) infection of wild white-tailed deer in new york city. *Viruses*, *14*, 2770. doi: 10.3390/v14122770

van Dorp, L., Tan, C. C., Lam, S. D., Richard, D., Owen, C., Berchtold, D., ... Balloux, F. (2020). Recurrent mutations in sars-cov-2 genomes isolated from mink point to rapid host-adaptation. *bioRxiv*. Retrieved from https://www.biorxiv.org/content/early/2020/11/16/2020.11.16.384743 doi: 10.1101/2020.11.16.384743

Villanueva-Saz, S., Giner, J., Tobajas, A. P., Pérez, M. D., González-Ramírez, A. M., Macías-León, J., ... others (2022). Serological evidence of sars-cov-2 and co-infections in stray cats in spain. *Transboundary and Emerging Diseases*, *69*(3), 1056–1064.

V'kovski, P., Gultom, M., Kelly, J. N., Steiner, S., Russeil, J., Mangeat, B., ... Dijkman, R. (2021, 3). Disparate temperature-dependent virus–host dynamics for sars-cov-2 and sars-

cov in the human respiratory epithelium. *PLOS Biology*, *19*, e3001158. doi: 10.1371/journal.pbio.3001158

Wang, L., Gyimesi, Z. S., Killian, M. L., Torchetti, M., Olmstead, C., Fredrickson, R., & Terio, K. A. (2022, 9). Detection of sars-cov-2 clade b.1.2 in three snow leopards. *Transboundary and emerging diseases*, *69*, e3346-e3351. doi: 10.1111/tbed.14625

Wang, R., Chen, J., Hozumi, Y., Yin, C., & Wei, G.-W. (2020). Decoding asymptomatic covid-19 infection and transmission. *The journal of physical chemistry letters*, *11*(23), 10007–10015.

Wei, C., Shan, K.-J., Wang, W., Zhang, S., Huan, Q., & Qian, W. (2021, 12). Evidence for a mouse origin of the sars-cov-2 omicron variant. *Journal of Genetics and Genomics*, *48*, 1111-1121. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S1673852721003738 doi: 10.1016/J.JGG.2021.12.003

Welkers, M. R., Han, A. X., Reusken, C. B., & Eggink, D. (2021). Possible host-adaptation of sars-cov-2 due to improved ace2 receptor binding in mink. *Virus evolution*, *7*(1), veaa094.

Wilm, A., Aw, P. P. K., Bertrand, D., Yeo, G. H. T., Ong, S. H., Wong, C. H., ... Nagarajan, N. (2012a). Lofreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic acids research*, *40*(22), 11189–11201.

Wilm, A., Aw, P. P. K., Bertrand, D., Yeo, G. H. T., Ong, S. H., Wong, C. H., ... Nagarajan, N. (2012b, 12). Lofreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Research*, *40*, 11189-11201. Retrieved from https://pubmed.ncbi.nlm.nih.gov/23066108/ doi: 10.1093/nar/gks918

WOAH. (2024). *Sars-cov-2.* World Organisation for Animal Health. Retrieved from https://www.woah.org/en/disease/sars-cov-2/

Worobey, M., Levy, J. I., Serrano, L. M., Crits-Christoph, A., Pekar, J. E., Goldstein, S. A., ... Andersen, K. G. (2022, 8). The huanan seafood wholesale market in wuhan was the early epicenter of the covid-19 pandemic. *Science*, *377*, 951-959. doi: 10.1126/science.abp8715

Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., ... others (2020). A new coronavirus associated with human respiratory disease in china. *Nature*, *579*(7798), 265–269.

Wu, H., Xing, N., Meng, K., Fu, B., Xue, W., Dong, P., ... Zhu, Z. (2021, 12). Nucleocapsid mutations r203k/g204r increase the infectivity, fitness, and virulence of sars-cov-2. *Cell Host & Microbe*, *29*, 1788-1801.e6. doi: 10.1016/j.chom.2021.11.005

Yang, J., Petitjean, S. J., Koehler, M., Zhang, Q., Dumitru, A. C., Chen, W., ... Alsteens, D. (2020). Molecular interaction and inhibition of sars-cov-2 binding to the ace2 receptor. *Nature communications*, *11*(1), 4541.

Yen, H.-L., Sit, T. H., Brackman, C. J., Chuk, S. S., Cheng, S. M., Gu, H., ... Poon, L. L. (2022, 1). Transmission of sars-cov-2 (variant delta) from pet hamsters to humans and onward human propagation of the adapted strain: A case study. *SSRN Electronic Journal*. Retrieved from `https://papers.ssrn.com/abstract=4017393` doi: 10.2139/SSRN.4017393

Zhang, Q., Zhang, H., Gao, J., Huang, K., Yang, Y., Hui, X., ... others (2020). A serological survey of sars-cov-2 in cat in wuhan. *Emerging microbes & infections*, *9*(1), 2013–2019.

Zhao, L., & Illingworth, C. J. (2019). Measurements of intrahost viral diversity require an unbiased diversity metric. *Virus Evolution*, *5*(1), vey041.

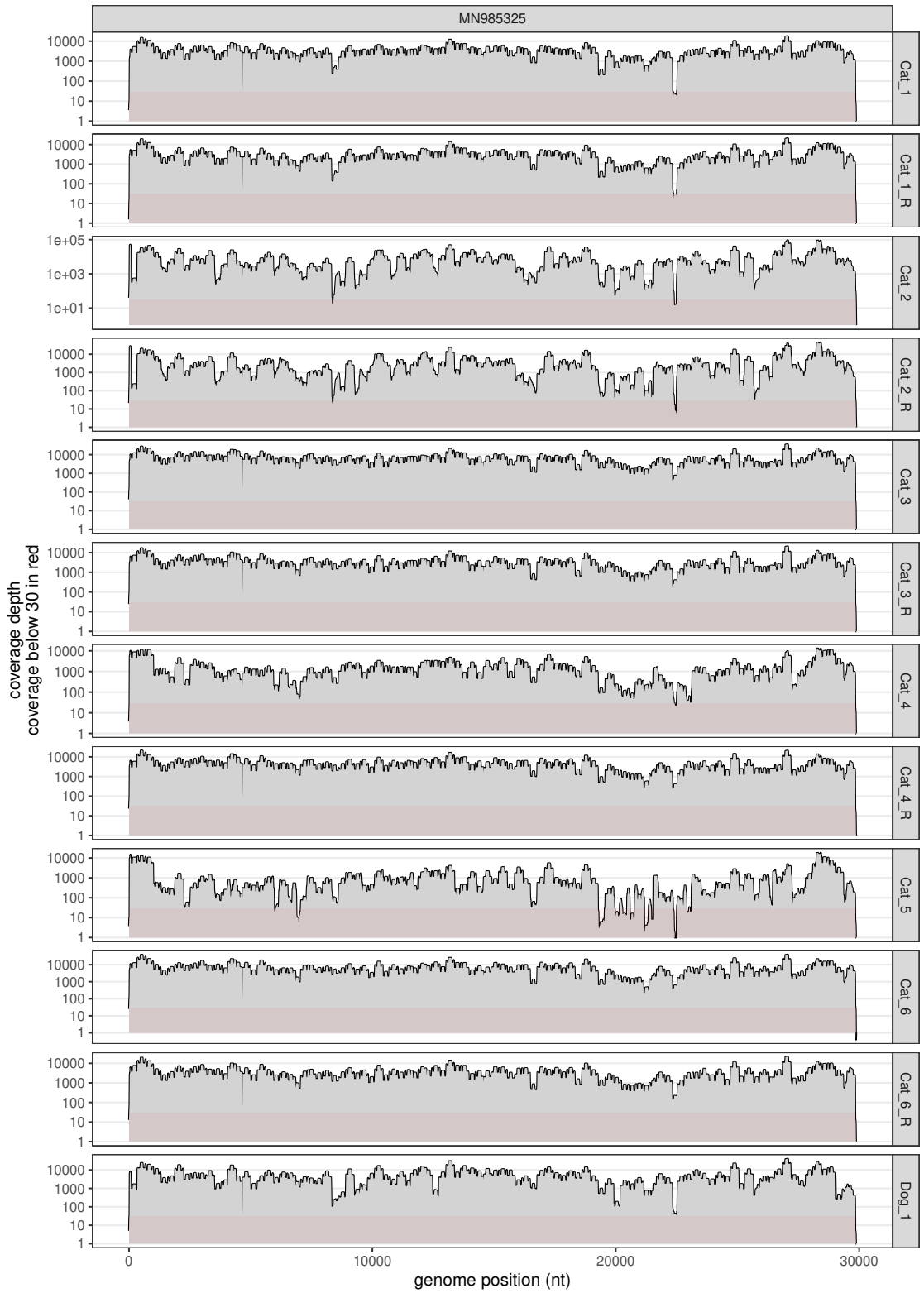
Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., ... others (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *nature*, *579*(7798), 270–273.

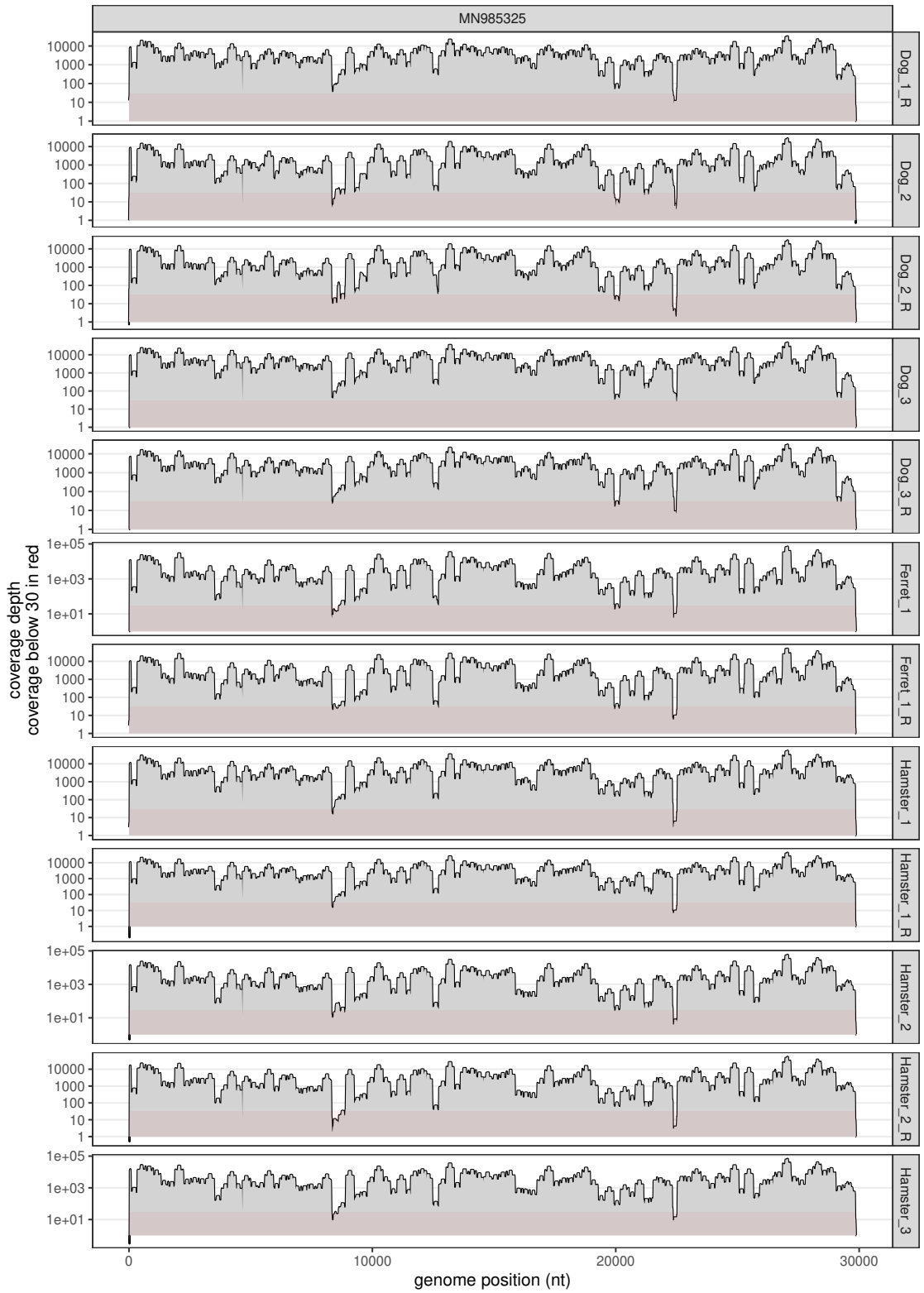
Zoccola, R., Beltramo, C., Magris, G., Peletto, S., Acutis, P., Bozzetta, E., ... Goria, M. (2021, 12). First detection of an italian human-to-cat outbreak of sars-cov-2 alpha variant – lineage b.1.1.7. *One Health*, *13*, 100295. doi: 10.1016/J.ONEHLT.2021.100295

Appendix A

Chapter 2 Supplementary Information[5]

[5]This chapter is a reproduction of "Bashor, L., Gagne, R. B., Bosco-Lauth, A., Bowen, R.A., Stenglein, M., & Vande-Woude, S. (2021). SARS-CoV-2 evolution in animals suggests mechanisms for rapid variant selection. *Proceedings of the National Academy of Sciences*, 118(44). doi: 10.1073/PNAS.2105253118". Minimal modifications were made to meet dissertation formatting requirements.


MN985325
Cat_1
Cat_1_R
Cat_2
Cat_2_R
Cat_3
Cat_3_R
Cat_4
Cat_4_R
Cat_5
Cat_6
Cat_6_R
Dog_1
coverage depth
coverage below 30 in red
10000
1000
100
10
1
1e+05
1e+03
1e+01
0
10000
20000
30000
genome position (nt)


MN985325
Dog_1_R
Dog_2
Dog_2_R
Dog_3
Dog_3_R
Ferret_1
Ferret_1_R
Hamster_1
Hamster_1_R
Hamster_2
Hamster_2_R
Hamster_3
coverage depth
coverage below 30 in red
10000
1000
100
10
1
1e+05
1e+03
1e+01
0
10000
20000
30000
genome position (nt)

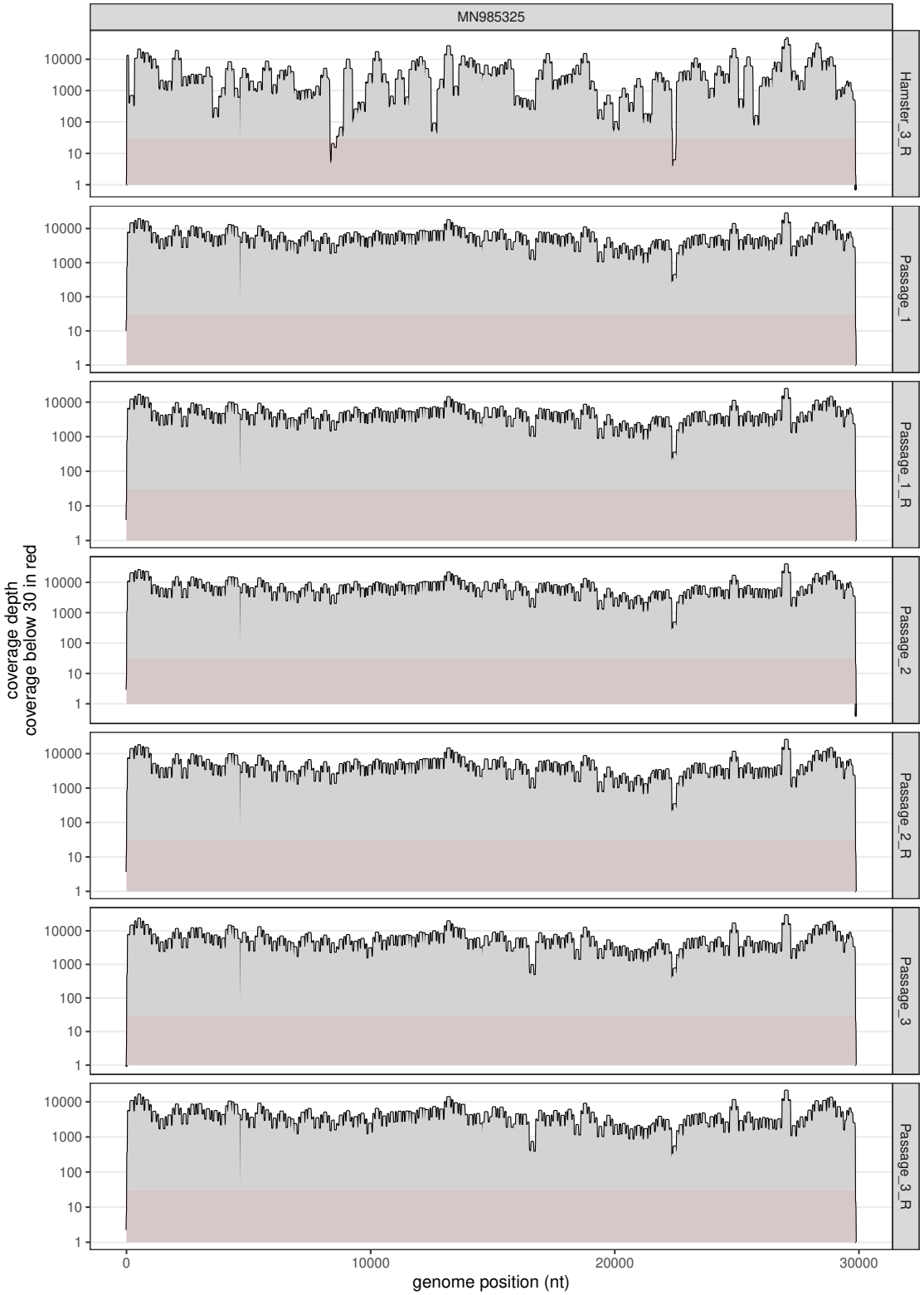


Figure A.1: Depth of coverage across the SARS-CoV-2 genome.

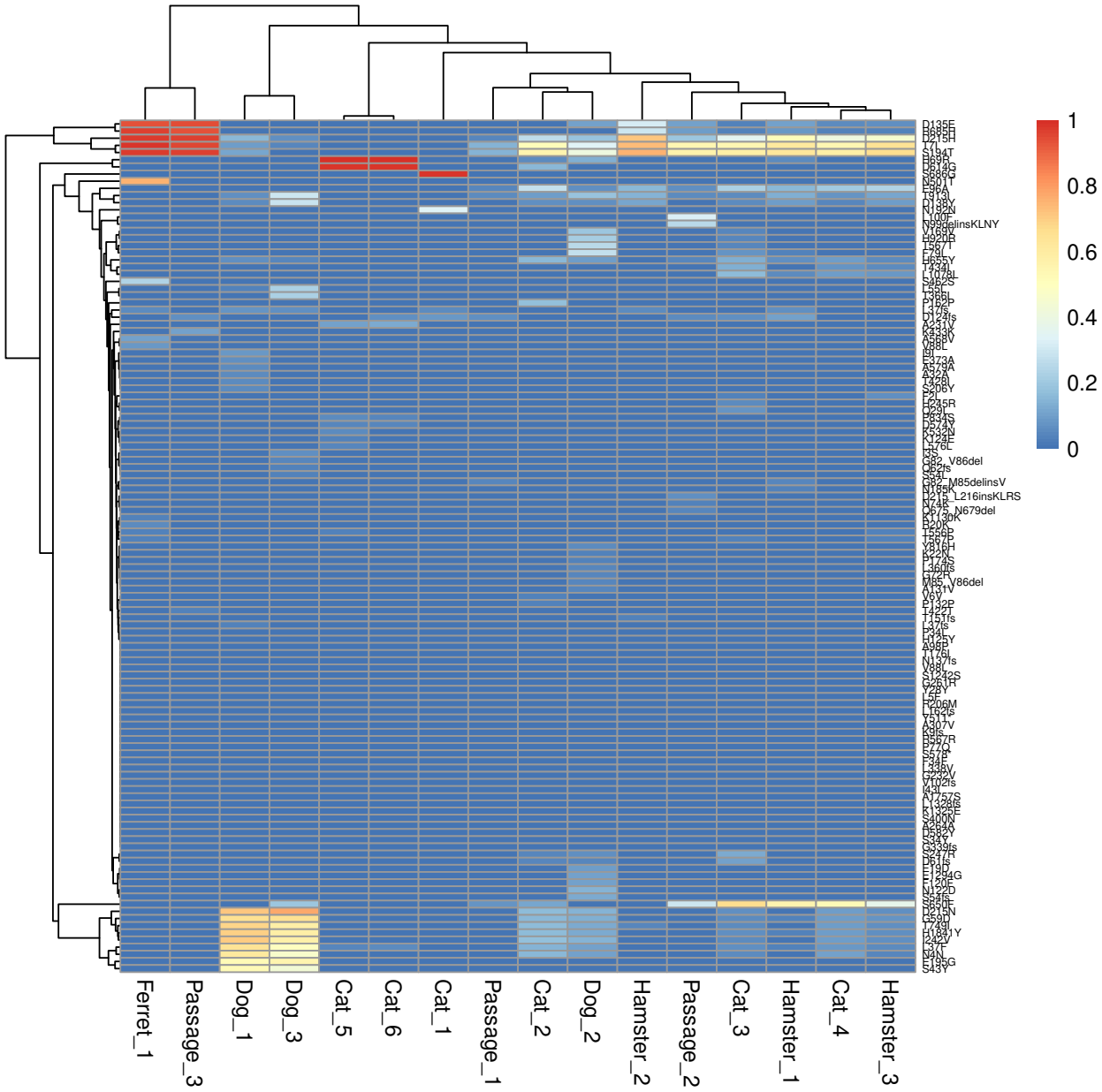


Figure A.2: Heatmap of SARS-CoV-2 variants across samples recovered from cell culture and experimentally inoculated animals.

Table A.1: Global serological surveys of domestic and nondomestic animals reveal the extent of human-to-animal transmission of SARS-CoV-2 in 2020 and 2021. Assays used to measure seropositivity include Luciferase Immuno-Precipitation System (LIPS), plaque reduction neutralization test (PRNT), virus neutralization test (VN), combined microsphere immunoassay (MIA) and retrovirus-based pseudoparticle assay, enzyme-linked immunosorbent assay (ELISA), microneutralisation test (MNT), indirect immunofluorescence test (iIFT), and Luminex xMAP microsphere-based assay (xMAP).
*Reported as "a commercially available SARS-CoV-2 antibody screening test."

Study scope	Location	Dates of sample collection	Species	Assay	Seropositivity rate	Ref
COVID-19 household close contacts	France	Mar-20	cat	LIPS	0% (0/9)	(Temmam et al., 2020)
COVID-19 household	Brazil	May-Oct 2020	cat	PRNT90	20% (2/10)	(Calvet et al., 2021)
COVID-19 household	USA	Jun-Sep 2020	cat	VN	43.8% (7/16)	(Hamer et al., 2021)
COVID-19 household	France	Jun-20	cat	MIA and pseudoparticle assay	23.5% (8/34)	(Fritz et al., 2020)
Confirmed suspected COVID-19 household	France	Apr-20	cat	ELISA	4.5% (1/22)	(Sailleau et al., 2020)
COVID-19 household, stray, vet hospital patient	China	Apr-Jun 2020	cat	ELISA	14.7% (15/102)	(Zhang et al., 2020)
General population	China	Feb-Apr 2020	cat	ELISA	0% (0/423)	(Deng et al., 2020)
General population	Croatia	Feb-Jun 2020	cat	MNT	0.76% (1/131)	(Stevanovic et al., 2021)
General population	Germany	Apr-Sep 2020	cat	ELISA and iIFT	0.69% (6/920)	(Michelitsch et al., 2020)
General population COVID-19 household	Italy	Mar-May 2020	cat	PRNT80	5.8% (11/191); 4.5% (1/22 from COVID-19 households)	(Patterson et al., 2020)
General population COVID-19 household	Italy	Mar-Jun 2020	cat	xMAP	16.2% (11/68)	(Colitti et al., 2021)
Feral cats	Spain	Jan-Oct 2020	cat	ELISA	3.51% (4/115)	(Villanueva-Saz et al., 2022)
COVID-19 outbreak on mink farm (house feral cats)	Netherlands	May-Oct 2020	cat	ELISA and VN	0% (0/12 house cats); 18% (11/89 feral cats)	(van Aart et al., 2022)
COVID-19 outbreak on mink farm (feral cats)	Netherlands	Apr-May 2020	cat	MNT	29.2% (7/24)	(Oreshkova et al., 2020)
COVID-19 outbreaks on 215 mink farms	Denmark	Jun-Nov 2020	cat	ELISA	0% (0/21)	(Boklund et al., 2021)
COVID-19 household close contacts	France	Mar-20	dog	LIPS	0% (0/12)	(Temmam et al., 2020)
COVID-19 household	Brazil	May-Oct 2020	dog	PRNT90	3.4% (1/29)	(Calvet et al., 2021)
COVID-19 household	USA	Jun-Sep 2020	dog	VN	11.9% (7/59)	(Hamer et al., 2021)

Table A.2: Table A.1 continued. Global serological surveys of domestic and nondomestic animals reveal the extent of human-to-animal transmission of SARS-CoV-2 in 2020 and 2021. Assays used to measure seropositivity include Luciferase Immuno-Precipitation System (LIPS), plaque reduction neutralization test (PRNT), virus neutralization test (VN), combined microsphere immunoassay (MIA) and retrovirus-based pseudoparticle assay, enzyme-linked immunosorbent assay (ELISA), microneutralisation test (MNT), indirect immunofluorescence test (iIFT), and Luminex xMAP microsphere-based assay (xMAP).
*Reported as "a commercially available SARS-CoV-2 antibody screening test."

Study scope	Location	Dates of sample collection	Species	Assay	Seropositivity rate	Ref
COVID-19 household	France	Jun-20	dog	MIA and pseudoparticle assay	15.4% (2/13)	(Fritz et al., 2020)
Confirmed suspected COVID-19 household	France	Apr-20	dog	ELISA	0% (0/11)	(Sailleau et al., 2020)
COVID-19 household, vet hospital patient	Spain	Apr-Jun 2020	dog	ELISA	25% (5/20 from COVID-19 households); 5.88% (1/17 patients)	(Perisé-Barrios et al., 2021)
COVID-19 household	Hong Kong	Mar-20	dog	PRNT90	13.3% (2/15)	(Sit et al., 2020)
General population	Croatia	Feb-Jun 2020	dog	MNT and ELISA	0.31% (2/654) neutralizing; 7.56% (13/172)	(Stevanovic et al., 2021)
General population COVID-19 household	Italy	Mar-May 2020	dog	PRNT80	3.3% (15/451); 12.8% (6/47 from COVID-19 households)	(Patterson et al., 2020)
General population COVID-19 household	Italy	Mar-Jun 2020	dog	xMAP	2.3% (3/130)	(Colitti et al., 2021)
COVID-19 outbreak on mink farm (house feral cats)	Netherlands	May-Oct 2020	dog	ELISA and VN	15% (2/13)	(van Aart et al., 2022)
COVID-19 outbreaks on 215 mink farms	Denmark	Jun-Nov 2020	dog	ELISA	75% (3/4)	(Boklund et al., 2021)
COVID-19 household	USA	Mar-Apr 2020	ferret	ELISA	0% (0/29)	(Sawatzki et al., 2021)
General population	Spain	Jan-Nov 2020	ferret	ELISA	1.57% (2/127)	(Giner et al., 2021)
COVID-19 outbreak on mink farm	Netherlands	Apr-May 2020	mink	MNT	98.3% (59/60 randomly selected)	(Oreshkova et al., 2020)
Escaped mink captured outside mink farm with COVID outbreak	USA	Aug-20	mink	VN	100% (11/11)	(Shriner et al., 2021)
Free-ranging deer populations	USA	Jan 2020-Mar 2021	white-tailed deer	*see legend	33% (159/481)	(United States Department of Agriculture Animal and Plant Health Inspection Service (USDA APHIS), 2021a)

Table A.3: Selected metadata for experimentally inoculated animals included in this study.

Sample ID	Age	Sex	Inoculation	Sample	log pfu/mL recovered	dpi
Dog 1	adult (5-8 years)	female	nasal	nasal	1	3
Dog 2	adult (5-8 years)	female	nasal	nasal	1	3
Dog 3	adult (5-8 years)	female	nasal	oral	1	3
Ferret 1	juvenile		nasal	nasal	3.6	3
Ferret 2	juvenile		nasal	nasal	1.9	3
Ferret 3	juvenile		nasal	nasal	1	3
Hamster 1	juvenile		nasal	oral	3.3	1
Hamster 2	juvenile		nasal	oral	3.5	1
Hamster 3	juvenile		nasal	oral	1	1
Cat 1	adult (5-8 years)	female	nasal	nasal	6.6	3
Cat 2	adult (5-8 years)	male	nasal	nasal	5.3	3
Cat 3	adult (5-8 years)	female	nasal	nasal	6	3
Cat 4	adult (5-8 years)	female	nasal	nasal	5.4	3
Cat 5	adult (5-8 years)	female	nasal	nasal	4.4	3
Cat 6	adult (5-8 years)	female	contact	nasal	5.1	

Table A.4: SARS-CoV-2 variants detected in Cats 5 and 6.

Variant	Position	CDS	N or S	Cat 5	Cat 6
A231V	9246	nsp4	missense	0.11	0.13
L37F	11083	nsp6	missense	0.05	0.06
K124E	13394	nsp10	missense	0.04	0
K532N	15036	nsp12	missense	0.06	0
T556P	15106	nsp12	missense	0.03	0
L576L	15168	nsp12	synonymous	0.05	0
P834S	15940	nsp12	missense	0.06	0.05
D124fs	19983	nsp15	frameshift variant stop gained	–	0.06
H69R	21768	S	missense	1	0.99
D574Y	23282	S	missense	0.04	0.05
D614G	23403	S	missense	0.99	0.98

Table A.5: Measurements of nucleotide diversity for each sample at the population level and associated statistical analyses.

Sample ID	pi	piN	piS	piN/piS	p-value	t-value
Cat 1	1.60E-05	1.19E-06	6.67E-05	1.78E-02	2.64E-02	6.04E+00
Cat 2	1.84E-04	2.01E-04	1.50E-04	1.34E+00		
Cat 3	1.73E-04	2.06E-04	8.53E-05	2.42E+00		
Cat 4	1.57E-04	1.95E-04	5.33E-05	3.65E+00		
Cat 5	2.77E-05	3.32E-05	1.29E-05	2.57E+00		
Cat 6	1.89E-05	2.53E-05	0.00E+00			
Dog 1	1.94E-04	2.19E-04	1.34E-04	1.63E+00	1.71E-01	1.60E+00
Dog 2	2.45E-04	2.69E-04	1.91E-04	1.41E+00		
Dog 3	2.25E-04	2.62E-04	1.32E-04	1.98E+00		
Ferret 1	6.57E-05	6.51E-05	7.52E-05	8.66E-01		
Hamster 1	1.25E-04	1.63E-04	1.21E-05	1.35E+01	9.90E-05	1.01E+02
Hamster 2	1.10E-04	1.47E-04	0.00E+00			
Hamster 3	1.45E-04	1.83E-04	3.63E-05	5.03E+00		

Table A.6: Measurements of nucleotide diversity for each sample at the gene product level and associated statistical analyses.

Sample ID	Gene product	piN	piS	piN/piS	p-value	t-value
Cat 1	E	0.00E+00	0.00E+00		3.37E-01	-1.00E+00
Cat 2	E	0.00E+00	0.00E+00			
Cat 3	E	0.00E+00	0.00E+00			
Cat 4	E	0.00E+00	0.00E+00			
Cat 5	E	0.00E+00	0.00E+00			
Cat 6	E	0.00E+00	0.00E+00			
Dog 1	E	0.00E+00	1.59E-03	0.00E+00		
Dog 2	E	0.00E+00	0.00E+00			
Dog 3	E	0.00E+00	0.00E+00			
Ferret 1	E	0.00E+00	0.00E+00			
Hamster 1	E	0.00E+00	0.00E+00			
Hamster 2	E	0.00E+00	0.00E+00			
Hamster 3	E	0.00E+00	0.00E+00			
Cat 1	M	0.00E+00	0.00E+00		4.10E-04	4.83E+00
Cat 2	M	9.98E-04	4.29E-04	2.33E+00		
Cat 3	M	9.90E-04	0.00E+00			
Cat 4	M	9.82E-04	0.00E+00			
Cat 5	M	0.00E+00	0.00E+00			
Cat 6	M	0.00E+00	0.00E+00			
Dog 1	M	3.49E-04	0.00E+00			
Dog 2	M	8.79E-04	0.00E+00			
Dog 3	M	1.89E-04	0.00E+00			
Ferret 1	M	3.88E-04	0.00E+00			
Hamster 1	M	9.35E-04	0.00E+00			
Hamster 2	M	7.85E-04	0.00E+00			
Hamster 3	M	9.05E-04	0.00E+00			
Cat 1	N	0.00E+00	1.51E-03	0.00E+00	9.54E-02	-1.81E+00
Cat 2	N	5.16E-04	2.10E-03	2.46E-01		
Cat 3	N	5.08E-04	3.72E-04	1.36E+00		
Cat 4	N	5.05E-04	8.23E-04	6.14E-01		
Cat 5	N	0.00E+00	0.00E+00			
Cat 6	N	0.00E+00	0.00E+00			
Dog 1	N	3.17E-04	1.60E-03	1.98E-01		
Dog 2	N	5.11E-04	6.76E-04	7.55E-01		
Dog 3	N	3.79E-04	1.63E-03	2.32E-01		
Ferret 1	N	5.16E-05	0.00E+00			
Hamster 1	N	4.90E-04	0.00E+00			
Hamster 2	N	3.91E-04	0.00E+00			
Hamster 3	N	4.62E-04	3.04E-04	1.52E+00		
Cat 1	orf10	0.00E+00	0.00E+00			
Cat 2	orf10	0.00E+00	0.00E+00			
Cat 3	orf10	0.00E+00	0.00E+00			
Cat 4	orf10	0.00E+00	0.00E+00			
Cat 5	orf10	0.00E+00	0.00E+00			
Cat 6	orf10	0.00E+00	0.00E+00			
Dog 1	orf10	0.00E+00	0.00E+00			
Dog 2	orf10	0.00E+00	0.00E+00			
Dog 3	orf10	0.00E+00	0.00E+00			
Ferret 1	orf10	0.00E+00	0.00E+00			
Hamster 1	orf10	0.00E+00	0.00E+00			
Hamster 2	orf10	0.00E+00	0.00E+00			
Hamster 3	orf10	0.00E+00	0.00E+00			

Table A.7: Table A.6 continued. Measurements of nucleotide diversity for each sample at the gene product level and associated statistical analyses.

Sample ID	Gene product	piN	piS	piN/piS	p-value	t-value
Cat 1	orf1ab	0.00E+00	0.00E+00		1.31E-02	2.91E+00
Cat 2	orf1ab	1.04E-04	6.47E-05	1.60E+00		
Cat 3	orf1ab	9.98E-05	9.47E-05	1.05E+00		
Cat 4	orf1ab	1.25E-04	2.34E-05	5.36E+00		
Cat 5	orf1ab	3.92E-05	1.78E-05	2.20E+00		
Cat 6	orf1ab	2.60E-05	0.00E+00			
Dog 1	orf1ab	1.97E-04	3.71E-05	5.31E+00		
Dog 2	orf1ab	1.68E-04	1.75E-04	9.62E-01		
Dog 3	orf1ab	2.34E-04	8.23E-05	2.84E+00		
Ferret 1	orf1ab	3.53E-05	8.41E-05	4.20E-01		
Hamster 1	orf1ab	7.61E-05	1.67E-05	4.56E+00		
Hamster 2	orf1ab	6.01E-05	0.00E+00			
Hamster 3	orf1ab	1.03E-04	3.15E-05	3.27E+00		
Cat 1	orf3a	0.00E+00	0.00E+00		3.37E-01	1.00E+00
Cat 2	orf3a	0.00E+00	0.00E+00			
Cat 3	orf3a	0.00E+00	0.00E+00			
Cat 4	orf3a	0.00E+00	0.00E+00			
Cat 5	orf3a	0.00E+00	0.00E+00			
Cat 6	orf3a	0.00E+00	0.00E+00			
Dog 1	orf3a	0.00E+00	0.00E+00			
Dog 2	orf3a	5.94E-04	0.00E+00			
Dog 3	orf3a	0.00E+00	0.00E+00			
Ferret 1	orf3a	0.00E+00	0.00E+00			
Hamster 1	orf3a	0.00E+00	0.00E+00			
Hamster 2	orf3a	0.00E+00	0.00E+00			
Hamster 3	orf3a	0.00E+00	0.00E+00			
Cat 1	orf6	0.00E+00	0.00E+00		9.82E-02	1.79E+00
Cat 2	orf6	0.00E+00	0.00E+00			
Cat 3	orf6	4.11E-04	0.00E+00			
Cat 4	orf6	0.00E+00	0.00E+00			
Cat 5	orf6	0.00E+00	0.00E+00			
Cat 6	orf6	0.00E+00	0.00E+00			
Dog 1	orf6	0.00E+00	0.00E+00			
Dog 2	orf6	0.00E+00	0.00E+00			
Dog 3	orf6	0.00E+00	0.00E+00			
Ferret 1	orf6	7.88E-04	0.00E+00			
Hamster 1	orf6	0.00E+00	0.00E+00			
Hamster 2	orf6	0.00E+00	0.00E+00			
Hamster 3	orf6	9.13E-04	0.00E+00			
Cat 1	orf7a	0.00E+00	0.00E+00		3.37E-01	1.00E+00
Cat 2	orf7a	0.00E+00	0.00E+00			
Cat 3	orf7a	0.00E+00	0.00E+00			
Cat 4	orf7a	0.00E+00	0.00E+00			
Cat 5	orf7a	0.00E+00	0.00E+00			
Cat 6	orf7a	0.00E+00	0.00E+00			
Dog 1	orf7a	0.00E+00	0.00E+00			
Dog 2	orf7a	0.00E+00	0.00E+00			
Dog 3	orf7a	4.68E-04	0.00E+00			
Ferret 1	orf7a	0.00E+00	0.00E+00			
Hamster 1	orf7a	0.00E+00	0.00E+00			
Hamster 2	orf7a	0.00E+00	0.00E+00			
Hamster 3	orf7a	0.00E+00	0.00E+00			

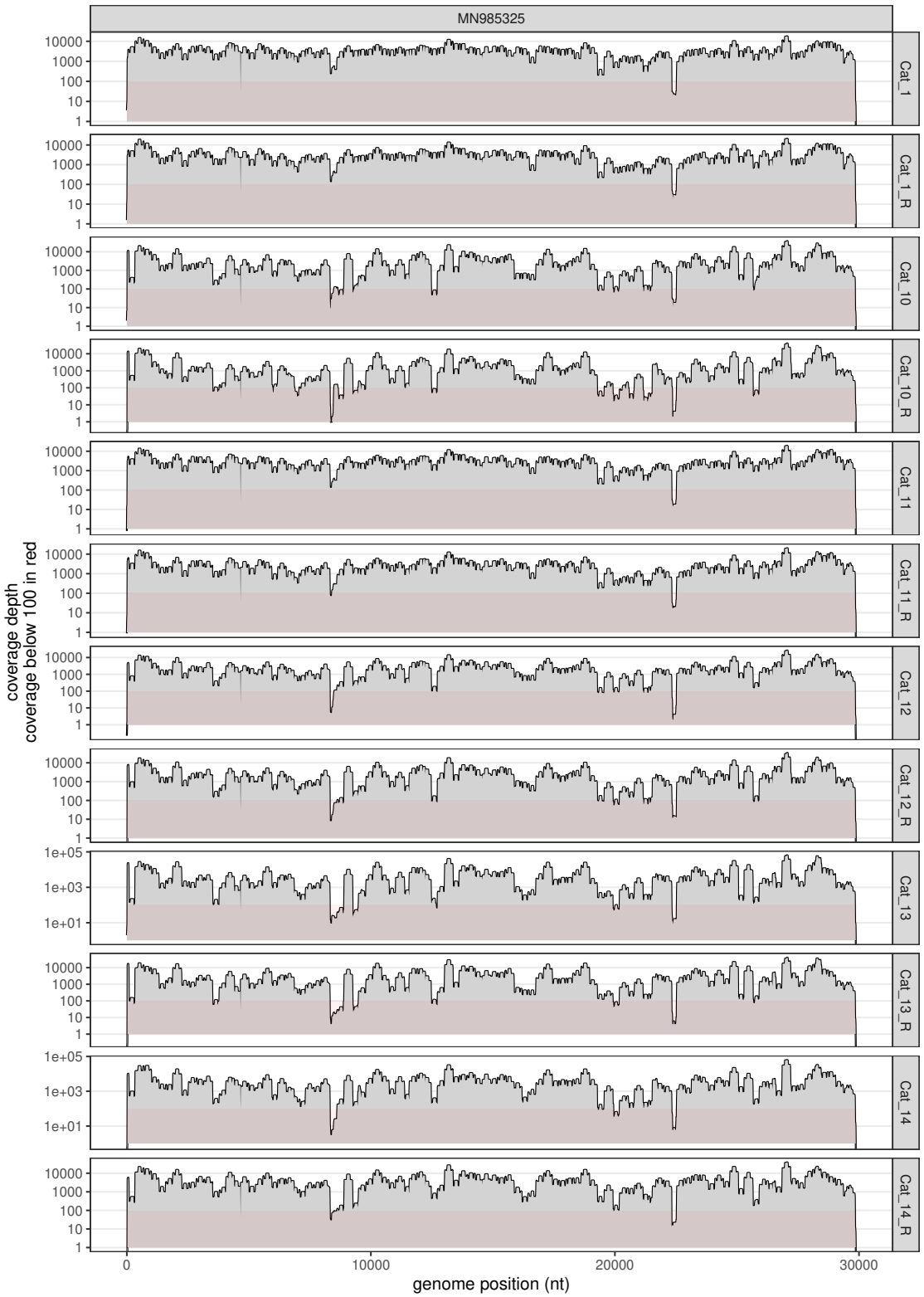
Table A.8: Table A.6 continued. Measurements of nucleotide diversity for each sample at the gene product level and associated statistical analyses.

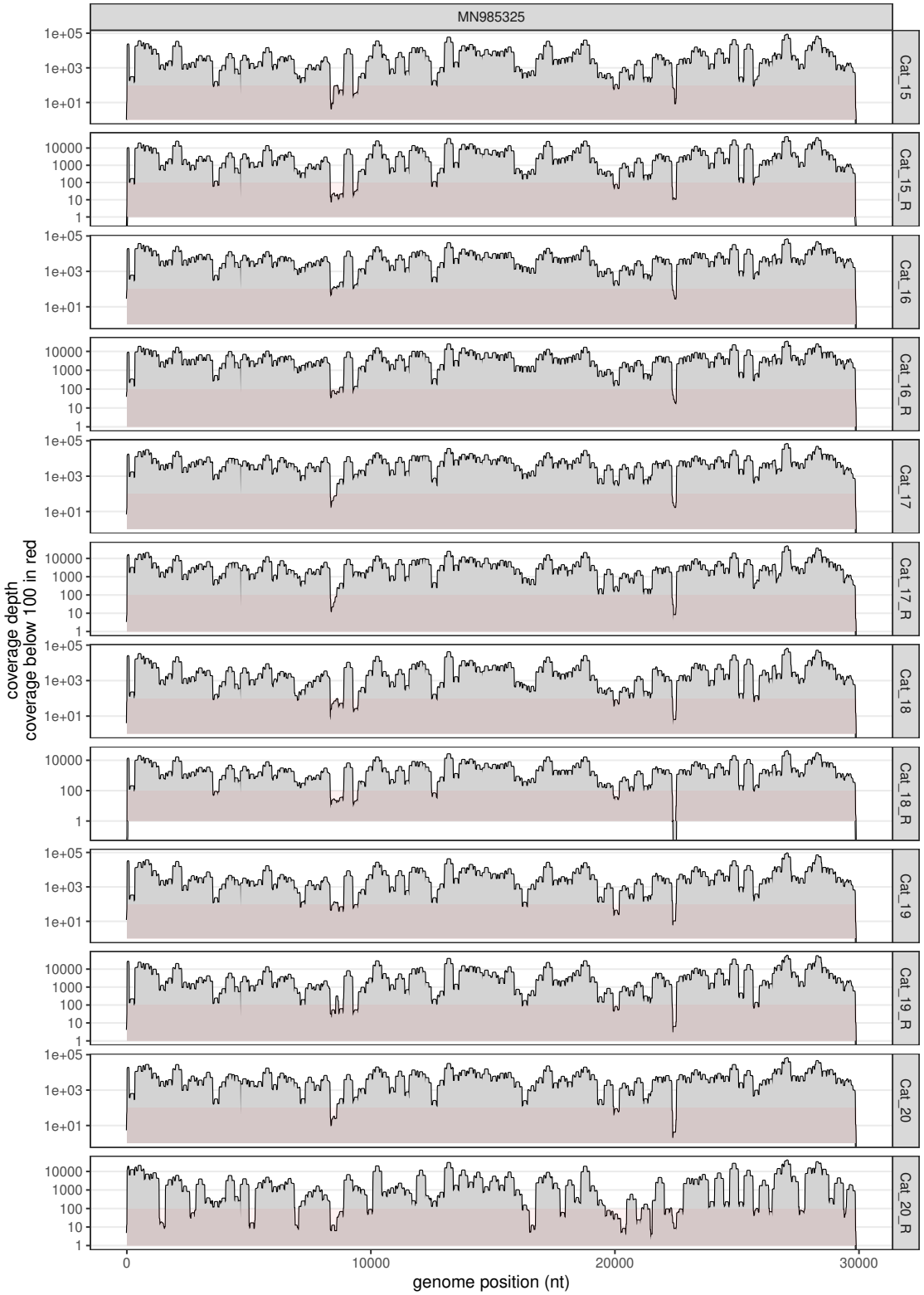
Sample ID	Gene product	piN	piS	piN/piS	p-value	t-value
Cat 1	orf8	0.00E+00	0.00E+00		9.18E-01	1.05E-01
Cat 2	orf8	0.00E+00	0.00E+00			
Cat 3	orf8	4.81E-04	0.00E+00			
Cat 4	orf8	0.00E+00	0.00E+00			
Cat 5	orf8	0.00E+00	0.00E+00			
Cat 6	orf8	0.00E+00	0.00E+00			
Dog 1	orf8	1.78E-03	1.81E-03	9.83E-01		
Dog 2	orf8	6.28E-04	2.77E-03	2.27E-01		
Dog 3	orf8	2.02E-03	0.00E+00			
Ferret 1	orf8	0.00E+00	0.00E+00			
Hamster 1	orf8	0.00E+00	0.00E+00			
Hamster 2	orf8	0.00E+00	0.00E+00			
Hamster 3	orf8	0.00E+00	0.00E+00			
Cat 1	S	9.08E-06	0.00E+00		2.27E-04	5.19E+00
Cat 2	S	6.16E-04	0.00E+00			
Cat 3	S	6.12E-04	0.00E+00			
Cat 4	S	4.49E-04	0.00E+00			
Cat 5	S	3.40E-05	0.00E+00			
Cat 6	S	4.76E-05	0.00E+00			
Dog 1	S	2.33E-04	0.00E+00			
Dog 2	S	6.08E-04	0.00E+00			
Dog 3	S	2.96E-04	0.00E+00			
Ferret 1	S	1.77E-04	1.07E-04	1.66E+00		
Hamster 1	S	5.00E-04	0.00E+00			
Hamster 2	S	5.25E-04	0.00E+00			
Hamster 3	S	4.66E-04	0.00E+00			

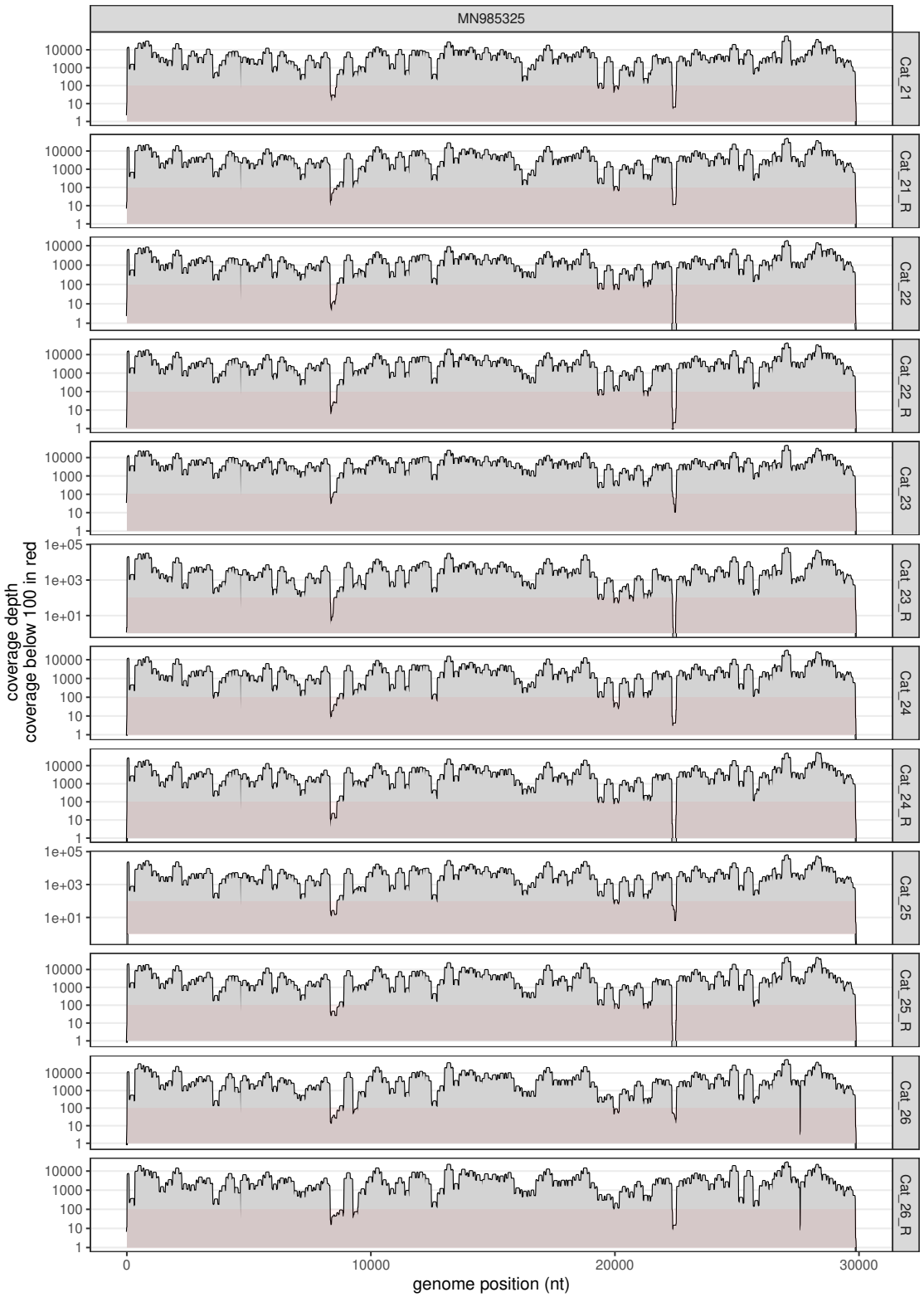
Appendix B

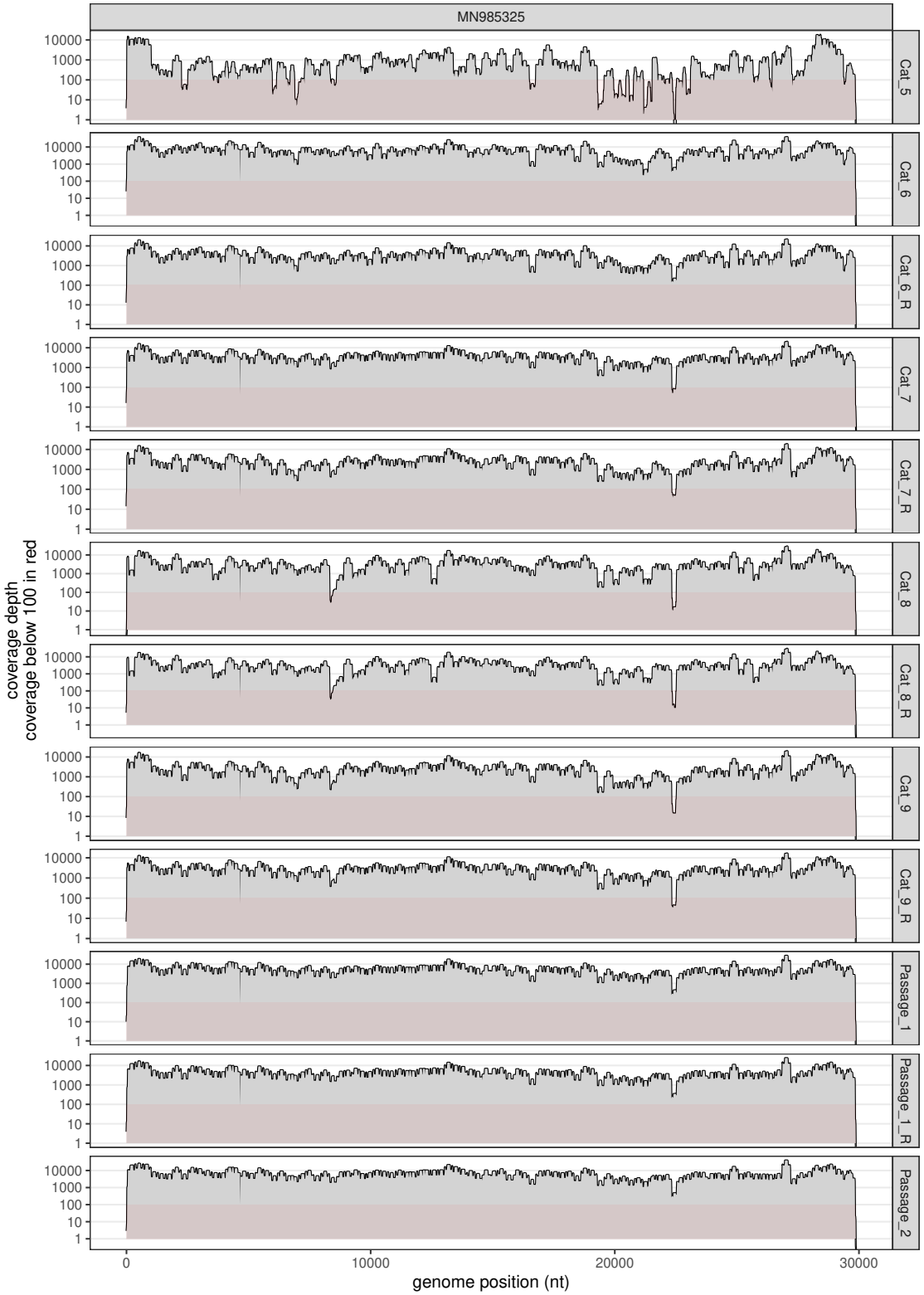
Chapter 3 Supplementary Information[6]

[6]This chapter is a reproduction of "Bashor, L., Gagne, R. B., Bosco-Lauth, A., Stenglein, M., & VandeWoude, S. (2022). Rapid evolution of SARS-CoV-2 in domestic cats. *Virus Evolution*, 8(2). doi: 10.1093/ve/veac092". Minimal modifications were made to meet dissertation formatting requirements.


MN985325
coverage depth
coverage below 100 in red
Cat_1
Cat_1_R
Cat_10
Cat_10_R
Cat_11
Cat_11_R
Cat_12
Cat_12_R
Cat_13
Cat_13_R
Cat_14
Cat_14_R
10000
1000
100
10
1
1e+05
1e+03
1e+01
0
10000
20000
30000
genome position (nt)


MN985325
Cat_15
Cat_15_R
Cat_16
Cat_16_R
Cat_17
Cat_17_R
Cat_18
Cat_18_R
Cat_19
Cat_19_R
Cat_20
Cat_20_R
coverage depth
coverage below 100 in red
1e+05
1e+03
1e+01
10000
1000
100
10
1
0
10000
20000
30000
genome position (nt)


MN985325
Cat_21
Cat_21_R
Cat_22
Cat_22_R
Cat_23
Cat_23_R
Cat_24
Cat_24_R
Cat_25
Cat_25_R
Cat_26
Cat_26_R
coverage depth
coverage below 100 in red
10000
1000
100
10
1
1e+05
1e+03
1e+01
0
10000
20000
30000
genome position (nt)


MN985325
Cat_5
Cat_6
Cat_6_R
Cat_7
Cat_7_R
Cat_8
Cat_8_R
Cat_9
Cat_9_R
Passage_1
Passage_1_R
Passage_2
10000
1000
100
10
1
coverage depth
coverage below 100 in red
0
10000
20000
30000
genome position (nt)

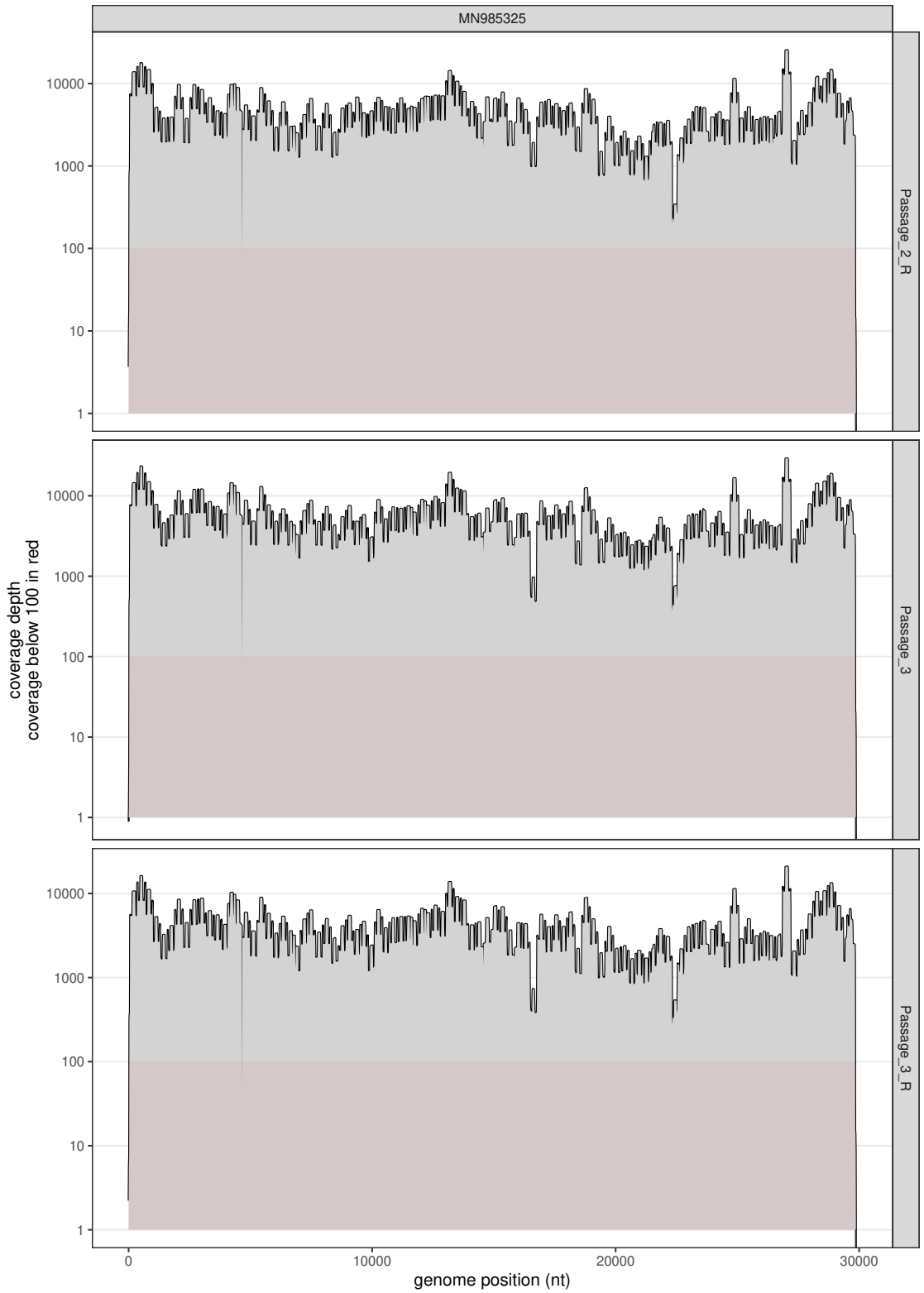


Figure B.1: Sequencing coverage at all positions across the SARS2 genome for each dataset.

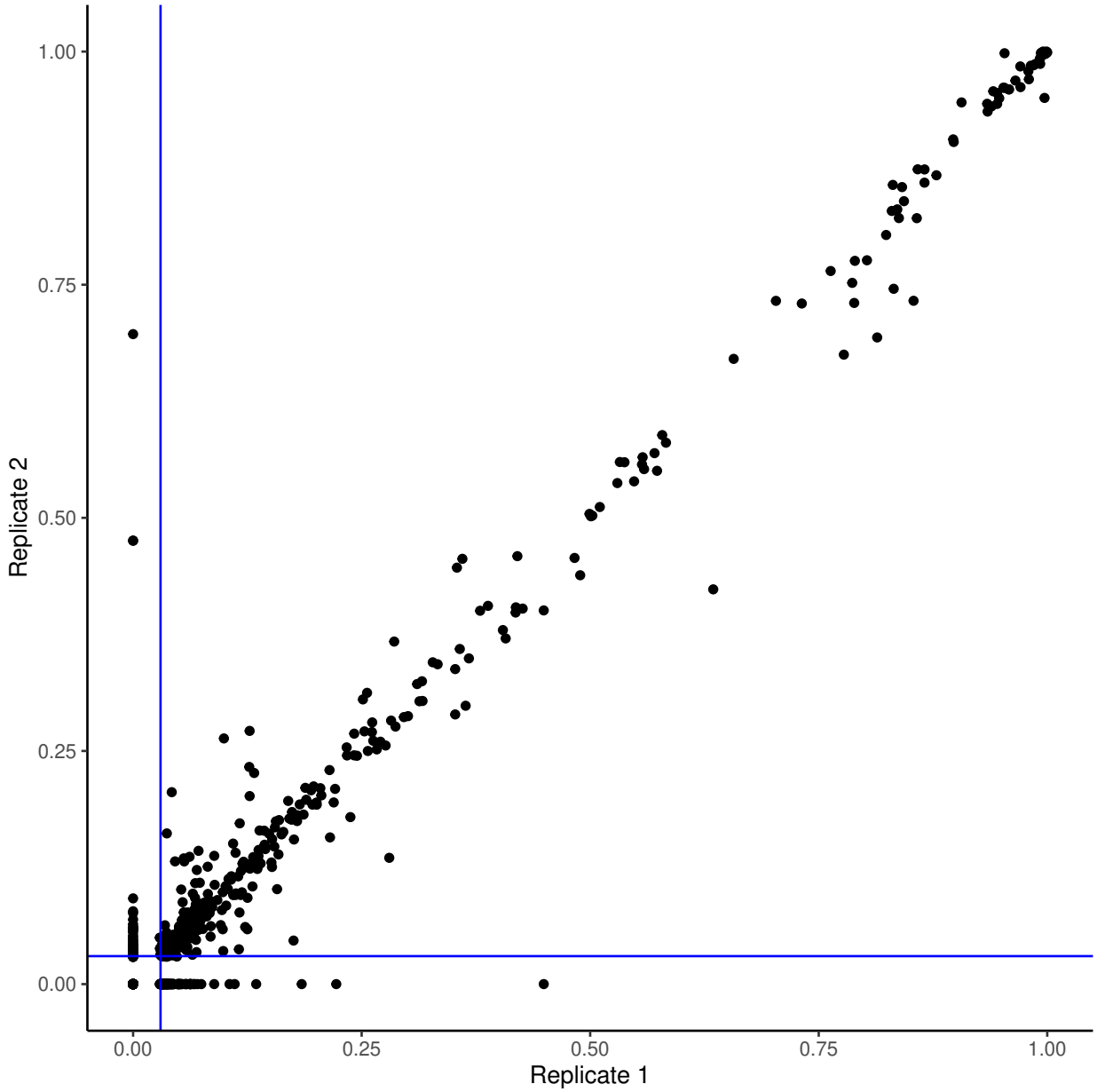


Figure B.2: Correlation between variant allele frequencies in technical replicates. Each point represents one cell culture- or cat-derived sample. Blue lines indicate the 3% allele frequency cutoff.

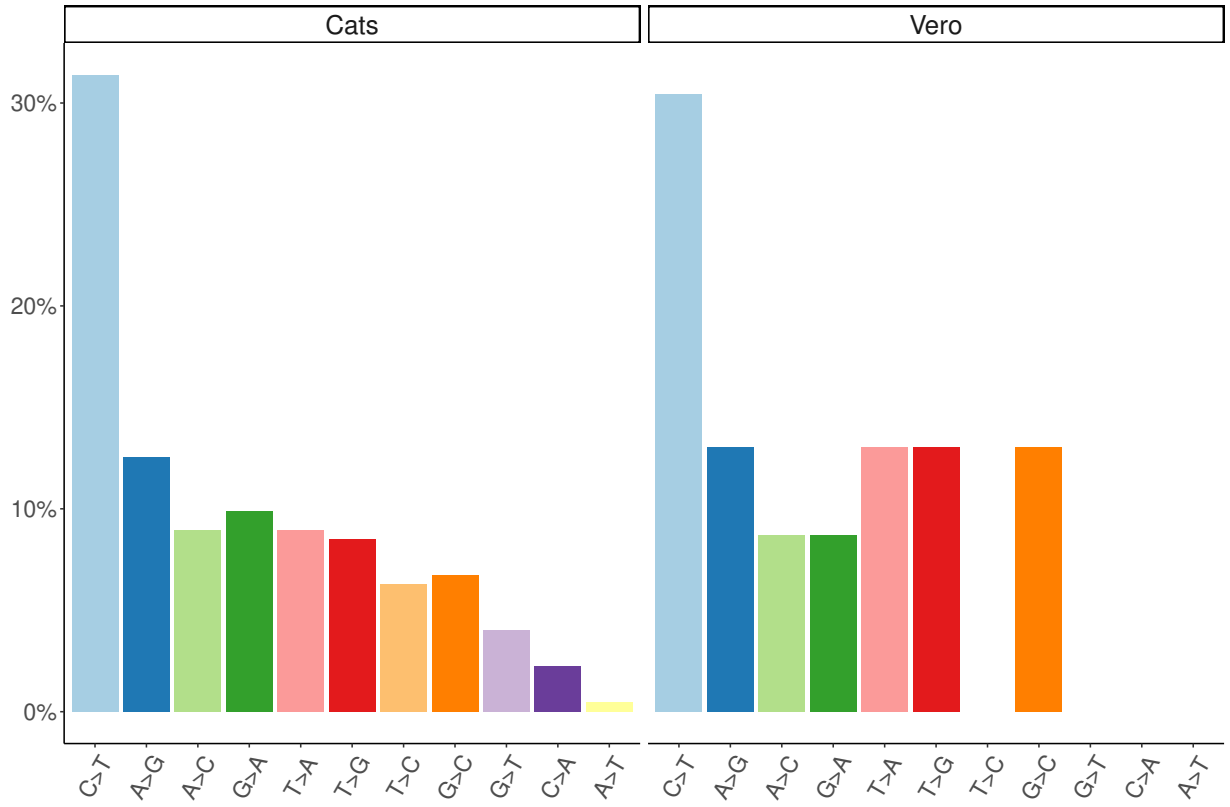


Figure B.3: C>T substitutions predominate in SARS2 genomes recovered from cats. Proportions of single nucleotide changes relative to the USA-WA1/2020 reference sequence detected in SARS2 recovered from experimentally exposed cats and viral inoculum stocks passaged in Vero cells.

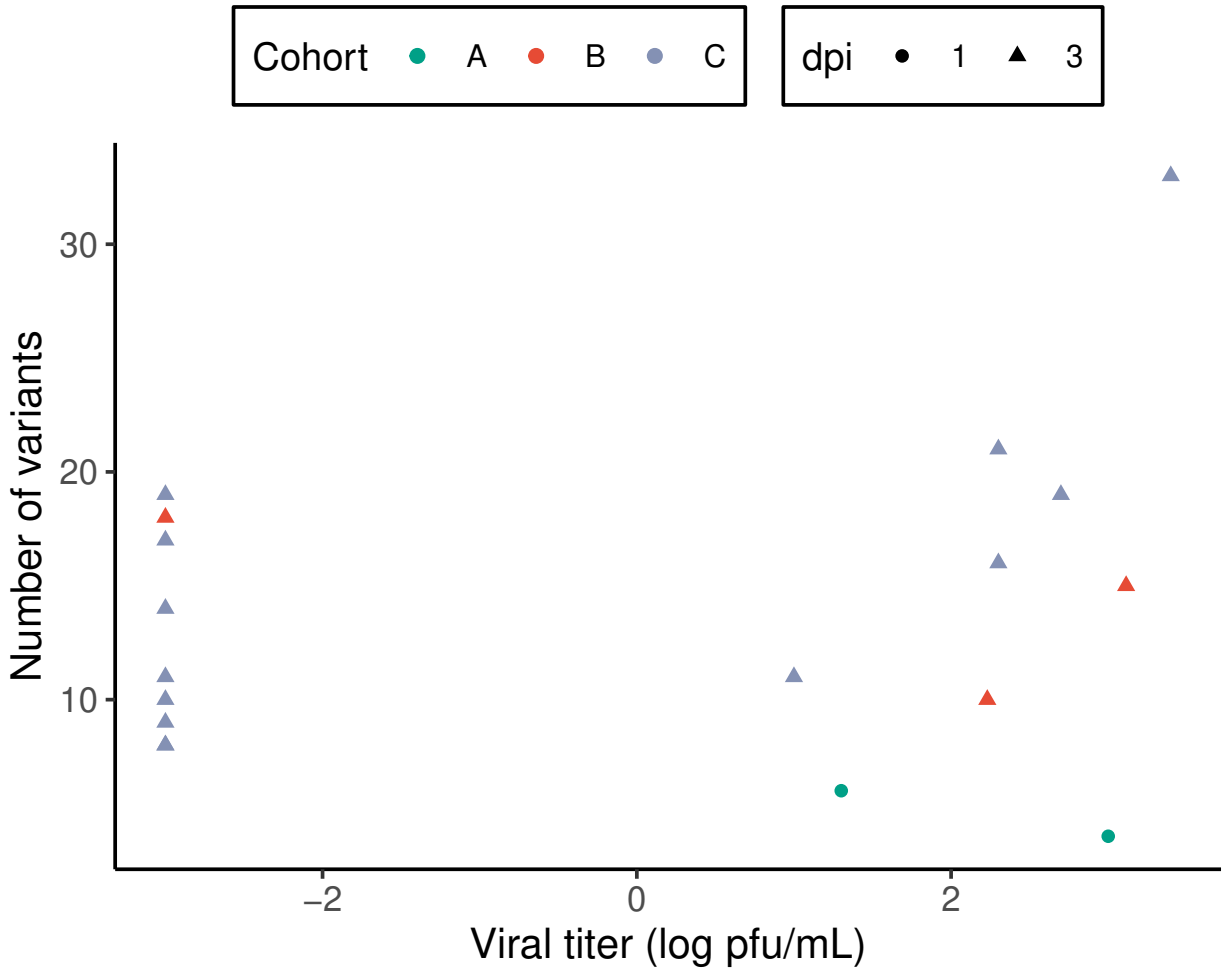


Figure B.4: The relationship between viral titer measured as log pfu/ml and the number of viral variants detected in viral RNA from directly inoculated cats (N=18). Each point represents one cat; color indicates cohort and shape indicates dpi of sample collection. We did not find evidence for a significant association between viral titer and number of variants (linear model).

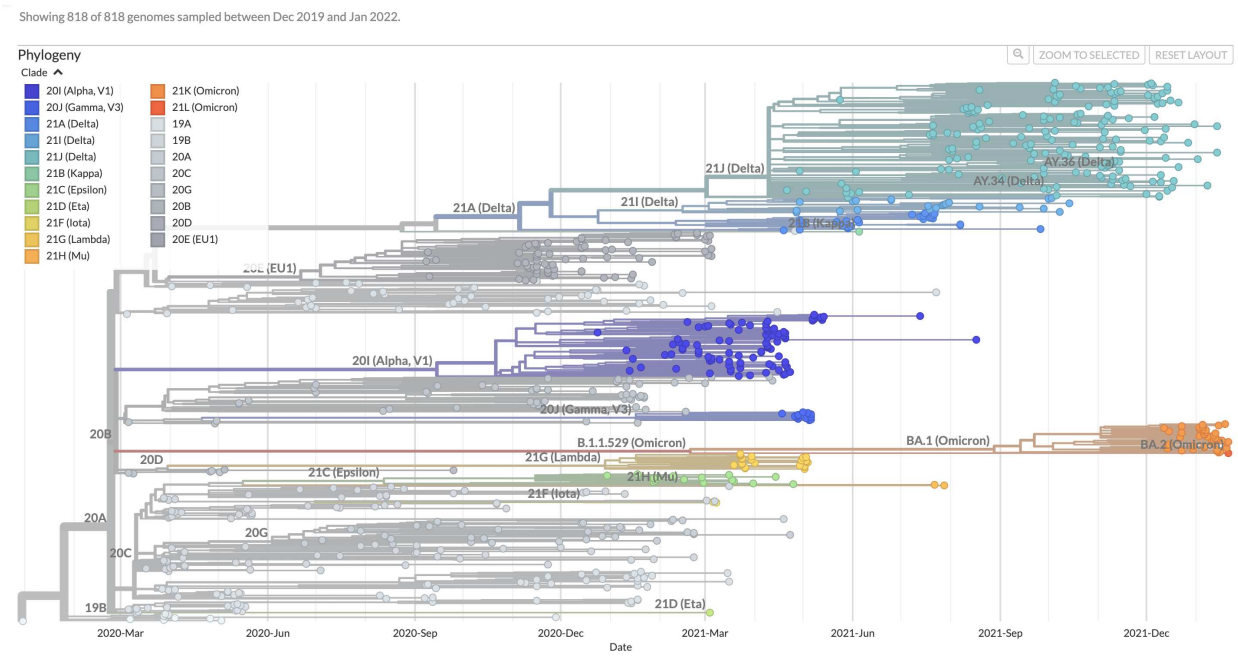


Figure B.5: Domestic cat- and felid- derived SARS2 genomes reflect lineages of SARS2 circulating in humans during the same time period. SARS2 sequences recovered from (a) naturally infected domestic cats (N=93) (b) all felids (N=29, including 93 cats (*Felis catus*), 9 tigers (*Panthera tigris*), 17 lions (*Panthera leo*), 2 snow leopards (*Panthera uncia*), and one leopard cat (*Prionailurus bengalensis*). Data were obtained from the GISAID database on April 4, 2022, analyzed with the Augur pipeline and visualized with the Auspice web application.

Table B.1: Twenty-three domestic cats were exposed to SARS2 in three experimental cohorts. Contact animals are indicated in bold. An asterisk indicates the potential infection donor in each cohort. The number of days post-infection (dpi) at which nasal lavage samples were collected for analysis refers to the number of days following the initial exposure of directly inoculated cats in each cohort. Doses included "low" ($1.05x10^4$ plaque-forming units (pfu)), "medium" ($4.65x10^5$ pfu), "med-high" ($1x10^6$ pfu), and "high" ($2.1x10^7$ pfu).

Cohort	ID	Age	Sex	Dose category	Inoculation stock passage	Sample collection dpi	Nasal log pfu/mL	Number of variants
A	Cat 1	adult	F	med-high	P3	1	3.0	4
A	Cat 5*	adult	F	med-high	P3	1	1.3	6
A	**Cat 6**	adult	F	contact		7	6.2	7
B	Cat 7*	adult	M	med-high	P2	3	3.1	15
B	Cat 8	adult	F	med-high	P2	3	2.2	10
B	Cat 9	adult	F	med-high	P2	3	0.0	18
B	**Cat 10**	adult	F	contact		4	3.6	7
B	**Cat 11**	adult	M	contact		4	2.9	5
B	**Cat 12**	adult	F	contact		4	1.8	6
C	Cat 13	juvenile	M	low	P3	3	0.0	8
C	Cat 14	juvenile	M	low	P3	3	0.0	8
C	Cat 15	juvenile	M	medium	P3	3	0.0	14
C	Cat 16	juvenile	F	medium	P3	3	0.0	17
C	Cat 17	juvenile	F	medium	P3	3	0.0	10
C	Cat 18	juvenile	M	medium	P3	3	0.0	9
C	Cat 19	juvenile	F	medium	P3	3	2.7	19
C	Cat 20	juvenile	F	medium	P3	3	0.0	11
C	Cat 21	juvenile	M	medium	P3	3	0.0	19
C	Cat 22*	juvenile	F	high	P3	3	3.4	33
C	Cat 23	juvenile	M	high	P3	3	1.0	11
C	Cat 24	juvenile	M	high	P3	3	2.3	16
C	Cat 25	juvenile	F	high	P3	3	2.3	25
C	**Cat 26**	juvenile	F	contact		4	0.0	7

Table B.2: Sequence identifiers, metadata and acknowledgments for virus genome sequences obtained from GISAID and used for phylogenetic reconstruction in Figure 3.6

We gratefully acknowledge the following Authors from the Originating laboratories responsible for obtaining the specimens, as well as the Submitting laboratories where the genome data were generated and shared via GISAID, on which this research is based.

All Submitters of data may be contacted directly via www.gisaid.org

Authors are sorted alphabetically.

Acknowledgement EPI_SET Identifier: EPI_SET_20220524fh

Accession ID	Originating Laboratory	Submitting Laboratory	Authors
EPI_ISL_3857106	1 - VetAgro Sup, Université de Lyon, 69280 Marcy-l'Etoile, France 2 - CIRI-Centre International de Recherche en Infectiologie, Team EVIR, Université Claude Bernard Lyon 1, Inserm, U111, CNRS, UMR5308, ENS Lyon	1 - Université de Caen, Groupe de Recherche sur l'Adaptation Microbienne (GRAM 2.0), Normandie Université, UNICAEN, UNIROUEN, EA2656 2 - Laboratoire de Virologie, Centre Hospitalo-Universitaire de Caen.	*; 2; 4; 5; Angeli Kodjo 1; Bertrand Boson 2; Bryce Leterrier 3; Emilie Krafft 1; Eric M. Leroy 5; François-Loïc Cosset 2; Matthieu Fritz 5; Meriadeg Ar Gouilh 3; Nicolas Nesi 3; Sandrine Corbet 4; Solène Denolly 2; Sophie Angelloz-Pessey 1; Sophie Levaltier 3; †; ‡; ‡ and Vincent Legros 1
EPI_ISL_3838696	1 - VetAgro Sup, Université de Lyon, 69280 Marcy-l'Etoile, France 2 - CIRI—Centre International de Recherche en Infectiologie, Team EVIR, Université Claude Bernard Lyon 1, Inserm, U111, CNRS, UMR5308, ENS Lyon	1 - Université de Caen, Groupe de Recherche sur l'Adaptation Microbienne (GRAM 2.0), Normandie Université, UNICAEN, UNIROUEN, EA2656 2 - Laboratoire de Virologie, Centre Hospitalo-Universitaire de Caen.	*; 2; 4; 5; Angeli Kodjo 1; Bertrand Boson 2; Bryce Leterrier 3; Emilie Krafft 1; Eric M. Leroy 5; François-Loïc Cosset 2; Matthieu Fritz 5; Meriadeg Ar Gouilh 3; Nicolas Nesi 3; Sandrine Corbet 4; Solène Denolly 2; Sophie Angelloz-Pessey 1; Sophie Levaltier 3; †; ‡; ‡ and Vincent Legros 1
EPI_ISL_1998018, EPI_ISL_8258718, EPI_ISL_8357620	Aegis Sciences Corporation	Centers for Disease Control and Prevention Division of Viral Diseases, Pathogen Discovery	Adrian Paskey; Alec Vest; Benjamin Rambo-Martin; Christopher Gulvick; Clinton Paden; Clinton R. Paden; Cyndi Clark; Dakota Howard; Darlene Wagner; Dhwani Batra; Dillon Nall; Duncan MacCannell; Erisa Sula; Ethan Sanders; Holly Houdeshell; Jason Caravas; Kara Moser; Kristine Lacek; Matthew Hardison; Matthew Schmerer; Ola Kvalvaag; Patrick Campbell; Peter Cook; Peter W. Cook; Rob Case; Scott Sammons; Shatavia Morrison; Shaun Westlund; Tymeckia Kendall; Victoria Caban Figueroa; Vikramsinha Ghorpade; Yvette Unoarumhi
EPI_ISL_738835	Alameda County Public Health Lab	Chan-Zuckerberg Biohub	CZB Cilahub Consortium
EPI_ISL_2958982, EPI_ISL_3128551, EPI_ISL_3152900	Antech Diagnostics	Diagnostic Virology Laboratory, National Veterinary Services Laboratories, USDA1920 Dayton Avenue, Ames, IA 50010, USA	Emily R. Love; Kerrie M. Franzen; Mary L. Killian; Mia Torchetti; Suelee Robbe-Austerman; Tod P. Stuber
EPI_ISL_3255522	Arizona State University	Arizona State University	Ajeet Bains; Efrem S. Lim; Joshua LaBaer; LaRinda A. Holland; Matthew F. Smith; Nathaniel Johnson; Nicholas J. Mellor; Peter T. Skidmore; Rabia Maqsood; Vel Murugan
EPI_ISL_2179607	BIOMNIS EUROFINS IVRY	Department of Virology, Henri Mondor University Hospital, Assistance Publique Hôpitaux de Paris, Université Paris-Est Créteil, INSERM U955	Alexandre Soulier; Christophe Rodriguez; Elisabeth Trawinski; Guillaume Gricourt; Jean-Michel Pawlotsky; Melissa N'Debi; Slim Fourati; Vanessa Demontant
EPI_ISL_1219363, EPI_ISL_1219617, EPI_ISL_1219619, EPI_ISL_1312768, EPI_ISL_6204713, EPI_ISL_6204714, EPI_ISL_8099231			
see above	BIOR	Latvian Biomedical Research and Study Centre	Alise Jakovele; Daina Pule; Davids Fridmanis; Elina Dimina; Guntars Zarins; Irena Meistere; Ivars Silamikelis; Janis Klovins; Janis Pjalkovskis; Jurijs Perevoscikovs; Kaspars Megnis; Laila Silamikele; Lauma Freimane; Laura Ansone; Liga Birzniece; Monta Briviba; Monta Ustinova; Nikita Zrelovs; Uga Dumpis; Una Krumina; Vita Rovite
EPI_ISL_1707190, EPI_ISL_1707193	Biogroup Bio Lam-LCD Saint-Denis	Department of Virology, Henri Mondor University Hospital, Assistance Publique Hôpitaux de Paris, Université Paris-Est Créteil, INSERM U955	Alexandre Soulier; Christophe Rodriguez; Elisabeth Trawinski; Guillaume Gricourt; Jean-Michel Pawlotsky; Melissa N'Debi; Slim Fourati; Vanessa Demontant
EPI_ISL_1749246	Biology Lab, HIA STE-ANNE	IRBA, 2MI	CHAPUS C.; DEPEILLE A.; GORGE O.; GRANDPERRET V.; JANVIER F.; JARJAVAL F.; NOLENT F.; SARILAR V.; VERGUET N.
EPI_ISL_2797625, EPI_ISL_2826155, EPI_ISL_2826213, EPI_ISL_9272958	CDPH VBL	California Department of Public Health	CDPH COVIDNet; CDPH-COVIDNet; SwabSeq/UCLA; UCSF Center for Advanced Technology
EPI_ISL_2226281	CHNDS	CHU Poitiers	Agnes BEBY-DEFAUX; Birama N'DIAYE; Caroline MICHAUD; Magali GARCIA; Manon PRAT; Maxime PICHON; Nicolas LEVEQUE; Valentin BON-BARET
EPI_ISL_1433871, EPI_ISL_1457757	CHU BORDEAUX - GH ST ANDRE	CNR Virus des Infections Respiratoires - France SUD	Antonin Bal; Bruno Lina; Bruno Simon; Gregory Destras; Gwendolyne Burfin; Hadrien Regue; Laurence Josset; Martine Valette; Quentin Semanas
EPI_ISL_1577349	CHU Lille	CHU Lille - Laboratoire de Virologie	AIT YAHYA Emilie; ALIDJINOU Enagnon Kazali; BOCKET Laurence; CREPIN Michel; DEMAY Christophe; ENGELMANN Ilka; GEFFROY Sandrine; GUIGON Aurélie; LAMBERT Valérie; LAZREK Mouna; NOBILLIAUX Florian; PREVOST Brigitte; TCHANTCHOU NJOSSE YANICK; THUILLIER Caroline; TINEZ Claire
EPI_ISL_663243	CHU Nîmes	CNR Virus des Infections Respiratoires - France SUD	Antonin Bal; Bruno Lina; Gregory Destras; Gwendolyne Burfin; Hadrien Règue; Jean-Philippe Lavigne; Laurence Josset; Marie-Josée Carles; Martine Valette; Maxence Lotellier; Quentin Semanas; Stephan Robin
EPI_ISL_663236, EPI_ISL_663240	CHU Poitiers	CNR Virus des Infections Respiratoires - France SUD	Agnès Beby-Defaux; Antonin Bal; Bruno Lina; Clément Jousselin; Gregory Destras; Gwendolyne Burfin; Hadrien Règue; Laurence Josset; Magali Garcia; Martine Valette; Nicolas Lévêque; Quentin Semanas
EPI_ISL_641534	CHU de Nice - Hôpital Archet 16	CNR Virus des Infections Respiratoires - France SUD	Antonin Bal; Bruno Lina; Gregory Destras; Gwendolyne Burfin; Géraldine Gonfrier; Hadrien Règue; Laurence Josset; Martine Valette; Quentin Semanas; Valérie Giordanengo
EPI_ISL_678508, EPI_ISL_678545, EPI_ISL_1461160, EPI_ISL_1526524	CNR Virus des Infections Respiratoires - France SUD	CNR Virus des Infections Respiratoires - France SUD	Antonin Bal; Bruno Lina; Bruno Simon; Gregory Destras; Gwendolyne Burfin; Hadrien Regue; Laurence Josset; Martine Valette; Quentin Semanas; Solenne Brun
EPI_ISL_451935	CUB Hopital Erasme Laboratoire d'Anatomie Pathologique	CUB Hopital Erasme Laboratoire d'Anatomie Pathologique	Dr. Nicky D'Haene; Prof. Isabelle Salmon
EPI_ISL_824564	Cedars-Sinai Medical Center, Department of Pathology & Laboratory Medicine, Molecular Pathology Laboratory	Cedars-Sinai Medical Center, Molecular Pathology Laboratory of Department of Pathology & Laboratory Medicine and Genomic Core	Brian Davis; Eric Vail; Jasmine T Plummer; Jorge Mario Sincuir Martinez; Stephanie Chen; Wenjuan Zhang
EPI_ISL_7508612	Cedars-Sinai Medical Center, Molecular Pathology Laboratory of Department of Pathology & Laboratory Medicine and Genomic Core	Cedars-Sinai Medical Center, Molecular Pathology Laboratory of Department of Pathology & Laboratory Medicine and Genomic Core	Brian Davis; Eric Vail; Jasmine T Plummer; Jorge Mario Sincuir Martinez; Stephanie Chen; Wenjuan Zhang
EPI_ISL_5320246	Center of excellence for emerging and re-emerging diseases in animals, Faculty of Veterinary Science, Chulalongkorn University	Center of excellence for emerging and re-emerging diseases in animals, Faculty of Veterinary Science, Chulalongkorn University	Alongkorn Amonsin; Ekkapat Chamsai; Kamol Suwannakarn; Kamonpan Charoenkul; Kitikhun Udom; Navapon Techakriengkrai; Ratanaporn Tangwangvivat; Supassama Chaiyawong; Waleemas Jairak
EPI_ISL_3255069, EPI_ISL_3255076	Central Public Health Laboratory - LACEN - Bahia, Salvador, Brazil	Central Public Health Laboratory - LACEN - Bahia, Salvador, Brazil	Arabela Leal; Breno Dominguez; Felicidade Pereira; Jaqueline Gomes; Luciana Oliveira; Luiz Alcantara; Marcela Gómez; Marta Giovanetti; Patrícia Cajado; Stephane Tosta; Vagner Fonseca; Vanessa Nardy
EPI_ISL_1706379	Centre Hospitalier Universitaire de Rouen Laboratoire de Virologie	Centre Hospitalier Universitaire de Rouen Laboratoire de Virologie	Alice Moisan; Fabienne De Oliveira; Marie Leoz
EPI_ISL_1590375, EPI_ISL_1590753	Centrālā Laboratorija; Eurofins Genomics Europe Sequencing GmbH	Riga East University Hospital-National Microbiology Reference Laboratory; Eurofins Genomics Europe Sequencing GmbH	Arzu Algulieva; Diāna Dusacka; Dārta Pūpola; Ilva Pole; Jana Osīte; Reinis Vangravs; Reinis Zeltmatis; Sergejs Nikisins; Stella Lapiņa; Girts Šķenders
EPI_ISL_1643114	Chemische und VeterinÃ¤runtersuchungsamt MÃ¼nsterland-Emscher-Lippe (CVUA-MEL)	Robert Koch Institute	
EPI_ISL_1005699, EPI_ISL_2521766, EPI_ISL_2521767, EPI_ISL_2521769	Clinical Laboratory, Vetsuisse Faculty, University of Zurich	Department of Biosystems Science and Engineering, ETH Zürich	Chaoran Chen; Christian Beisel; David Dreifuss; Elodie Burcklen; Ina Nissen; Ivan Topolsky; Julia Klaus; Katharina Jahn; Lara Fuhrmann; Marina Luisa Meli; Natascha Santacroce; Niko Beerenwinkel; Philipp Jablonski; Rebecca Denes; Regina Hofmann-Lehmann; Sarah Nadeau; Tanja Stadler
EPI_ISL_1273729	Clinical Virology	Clinical Bacteriology	Adrian Egli; Alfredo Mari; Hans Hirsch; Helena MB Seth-Smith; Julia Bielicki; Karoline Leuzinger; Madlen Stange; Manuel Battegay; Tim Roloff
EPI_ISL_1029973	Cliniques universitaires Saint-Luc	UCLouvain/IREC/MBLG	Benoit Kabamba Mukadi; Jean Ruelle; Lysa Pinsmaye
EPI_ISL_4505903	Colmar - SMIT	IAME UMR1137 Inserm, Université de Paris, Hôpital Bichat	Alexandre Storto; Anna Maria Franco Yusti; Antoine Bridier-Nahmias; Benoit Visseaux; Charlotte Charpentier; Diane Descamps; Gilles Collin; Lena Daniel; Matilda Berkell; Mélanie Bertine; Nadhira Houhou-Fidouh; Quentin Le Hingrat; Romain Coppée; Samuel Lebourgeois; Surbhi Malhotra
EPI_ISL_5636798, EPI_ISL_9400652	Color Health Inc.	California Department of Public Health	CDPH-COVIDNet and UCSF Center for Advanced Technology; Emily Smith on behalf of CDPH-COVIDNet and UCB/IGI SARS-CoV-2 Sequencing Lab
EPI_ISL_3010050	Cornell Diagnostic Laboratory	Diagnostic Virology Laboratory, National	Emily R. Love; Kerrie M. Franzen; Mary L. Killian; Mia Torchetti; Suelee Robbe-Austerman; Tod P. Stuber

Table B.3: Sequence identifiers, metadata and acknowledgments for virus genome sequences obtained from GISAID and used for phylogenetic reconstruction in Figure B.5

We gratefully acknowledge the following Authors from the Originating laboratories responsible for obtaining the specimens, as well as the Submitting laboratories where the genome data were generated and shared via GISAID, on which this research is based.

All Submitters of data may be contacted directly via www.gisaid.org

Authors are sorted alphabetically.

Acknowledgement EPI_SET Identifier: EPI_SET_20220524vd

Accession ID	Originating Laboratory	Submitting Laboratory	Authors
EPI_ISL_5515824, EPI_ISL_5516482, EPI_ISL_5773338, EPI_ISL_5859590, EPI_ISL_6071417, EPI_ISL_6785656, EPI_ISL_6785925, EPI_ISL_6786912, EPI_ISL_7230533, EPI_ISL_7237744, EPI_ISL_7438093, EPI_ISL_7920213, EPI_ISL_8030284, EPI_ISL_8150216, EPI_ISL_8179117, EPI_ISL_8357951			
see above	Aegis Sciences Corporation	Centers for Disease Control and Prevention Division of Viral Diseases, Pathogen Discovery	Alec Vest; Benjamin Rambo-Martin; Christopher Gulvick; Clinton Paden; Cyndi Clark; Dakota Howard; Dhwani Batra; Dillon Nall; Duncan MacCannell; Erisa Sula; Ethan Sanders; Holly Houdeshell; Jason Caravas; Kristine Lacek; Matthew Hardison; Matthew Schmerer; Ola Kvalvaag; Patrick Campbell; Peter Cook; Rob Case; Scott Sammons; Shatavia Morrison; Shaun Westlund; Tymeckia Kendall; Victoria Caban Figueroa; Vikramsinha Ghorpade; Yvette Unoarumhi
EPI_ISL_11348757	Animal Disease Diagnostic Laboratory, Purdue University	Animal Disease Diagnostic Laboratory, Purdue University College of Veterinary Medicine	Jobin Kattoor and Rebecca Wilkes
EPI_ISL_8145732	Antech Diagnostics	Diagnostic Virology Laboratory, National Veterinary Services Laboratories, USDA1920 Dayton Avenue, Ames, IA 50010, USA	Emily R. Love; Kerrie M. Franzen; Mary L. Killian; Mia Torchetti; Suelee Robbe-Austerman; Tod P. Stuber
EPI_ISL_5761534	Bronson Animal Disease Diagnostic Laboratory	Diagnostic Virology Laboratory, National Veterinary Services Laboratories, USDA1920 Dayton Avenue, Ames, IA 50010, USA	Emily R. Love; Kerrie M. Franzen; Mary L. Killian; Mia Torchetti; Suelee Robbe-Austerman; Tod P. Stuber
EPI_ISL_5321558, EPI_ISL_7130704, EPI_ISL_7859491	CDPH VBL	California Department of Public Health	CDPH COVIDNet; Emily Smith on behalf of CDPH-COVIDNet
EPI_ISL_7974436	California Animal Health and Food Safety Laboratory - University of California-Davis	Diagnostic Virology Laboratory, National Veterinary Services Laboratories, USDA1920 Dayton Avenue, Ames, IA 50010, USA	Emily R. Love; Kerrie M. Franzen; Mary L. Killian; Mia Torchetti; Suelee Robbe-Austerman; Tod P. Stuber
EPI_ISL_9461296, EPI_ISL_11583324, EPI_ISL_11583327, EPI_ISL_11583331, EPI_ISL_11583332, EPI_ISL_11583333	Clinical Laboratory, Vetsuisse Faculty, University of Zurich	Department of Biosystems Science and Engineering, ETH Zürich	Cecilia Valenzuela Agüí; Chaoran Chen; Christian Beisel; David Dreifuss; Elodie Burcklen; Evelyn Kuhlmeier; Ina Nissen; Ivan Topolsky; Katharina Jahn; Lara Fuhrmann; Marina Luisa Meli; Natascha Santacroce; Niko Beerenwinkel; Philipp Jablonski; Rebecca Denes; Regina Hofmann-Lehmann; Tanja Stadler; Tatjana Chan
EPI_ISL_4645894	Department of Infectious Diseases, Kobe Institute of Health	Department of Infectious Diseases, Kobe Institute of Health	Kentaro Itokawa; Makoto Kuroda; Masanori Hashino; Noriko Nakanishi; Rina Tanaka; Ryohei Nomoto; Tomotada Iwamoto; Tsuyoshi Sekizuka
EPI_ISL_5104522	Department of Veterinary Science, National Institute of Infectious Diseases	Department of Veterinary Science, National Institute of Infectious Diseases	Keita Ishijima; Ken Maeda; Tsukasa Yamamoto; Yudai Kuroda
EPI_ISL_6858845	Florida Bureau of Public Health Laboratories	Florida Bureau of Public Health Laboratories	Jason Blanton; Namratha Tarigopula; Sarah Schmedes; Tiffany Splatt
EPI_ISL_4641091	Fukuoka Institute of Health and Environmental Sciences	Fukuoka Institute of Health and Environmental Sciences	Akira Ohishi; Asako Nakamura; Chiharu Katamune; Hiroaki Shigemura; Ikumi Sato; Kentaro Itokawa; Makoto Kuroda; Masanori Hashino; Mitsuhiro Hamasaki; Rina Tanaka; Saori Ueda; Takayuki Kobayashi; Tsuyoshi Sekizuka; Yoshiki Etoh; Yuki Ashizuka; Yuki Carl
EPI_ISL_3659918, EPI_ISL_4027091, EPI_ISL_6654971, EPI_ISL_7799954	Fulgent Genetics	Centers for Disease Control and Prevention Division of Viral Diseases, Pathogen Discovery	Adrian Paskey; Becky Tsai; Benafsh Sapra; Benjamin Rambo-Martin; Christopher Gulvick; Clinton Paden; Dakota Howard; Darlene Wagner; Dhwani Batra; Doreen Ng; Duncan MacCannell; Erisa Sula; Harry Gao; James Xie; Jason Caravas; John Gao; Joseph Fierro; Kara Moser; Kristine Lacek; Matthew Schmerer; Mickey Li; Peter Cook; Scott Sammons; Shatavia Morrison; Tymeckia Kendall; Victoria Caban Figueroa; Yan Meng; Yvette Unoarumhi
EPI_ISL_5947577, EPI_ISL_6404981, EPI_ISL_6405730, EPI_ISL_6787920, EPI_ISL_6902201	Helix	Centers for Disease Control and Prevention Division of Viral Diseases, Pathogen Discovery	Benjamin Rambo-Martin; Christopher Gulvick; Clinton Paden; Dakota Howard; Dhwani Batra; Duncan MacCannell; Erisa Sula; Helix CA; Jason Caravas; Kristine Lacek; Matthew Schmerer; Peter Cook; Scott Sammons; Shatavia Morrison; Tymeckia Kendall; Victoria Caban Figueroa; Yvette Unoarumhi
EPI_ISL_4986963, EPI_ISL_4987045, EPI_ISL_5328344	Hospital General Universitario Gregorio Marañón	Hospital General Universitario Gregorio Marañón	Cristina Rodríguez-Grande; Darío García de Viedma; Julia Suárez; Laura Pérez-Lago; Marta Herranz Martin; Patricia Muñoz; Pedro Sola Campoy; Pilar Catalán; Sergio Buenestado Serrano; Víctor Manuel de la Cueva
EPI_ISL_5761527, EPI_ISL_5761535, EPI_ISL_5781756, EPI_ISL_6088086, EPI_ISL_6088089, EPI_ISL_10656103, EPI_ISL_10656104			
see above	IDEXX Reference Laboratories	Diagnostic Virology Laboratory, National Veterinary Services Laboratories, USDA1920 Dayton Avenue, Ames, IA 50010, USA	Emily R. Love; Kerrie M. Franzen; Mary L. Killian; Mia Torchetti; Suelee Robbe-Austerman; Tod P. Stuber
EPI_ISL_6340423	Idaho Bureau of Laboratories	Idaho Bureau of Laboratories	"R. Beukelman; Aimee Ceniseros; Christian Loera; Christopher Ball"; Matthew Charles Burns; Robert L. Voermans
EPI_ISL_4955419	Infinity Biologix	Centers for Disease Control and Prevention Division of Viral Diseases, Pathogen Discovery	Benjamin Rambo-Martin; Chirayu Goswami; Christian Bixby; Christopher Gulvick; Clinton Paden; Dakota Howard; Dhwani Batra; Duncan MacCannell; Erisa Sula; Jason Caravas; Jonathan Schultz; Kristine Lacek; Matthew Schmerer; Peter Cook; Robin Grimwood; Russ Hager; Scott Sammons; Shatavia Morrison; Tymeckia Kendall; Victoria Caban Figueroa; Yihe Wang; Yvette Unoarumhi
EPI_ISL_3419381	Institute for Urban Disease Control and Prevention	Division of Genomic Medicine and Innovation support,Department of Medical Sciences, Ministry of Public Health, Thailand	Archawin Rojanawiwat; Jirapha Pakdee; Natthakul Bunneang; Nuanjun Wichukchinda; Penpitcha Thawong; Pilailuk Akkapaiboon Okada; Pundharika Piboonsiri; Surakameth Mahasirimongkol; Waritta Sawaengdee
EPI_ISL_7974437	Iowa State University Veterinary Diagnostic Laboratory	Diagnostic Virology Laboratory, National Veterinary Services Laboratories, USDA1920 Dayton Avenue, Ames, IA 50010, USA	Emily R. Love; Kerrie M. Franzen; Mary L. Killian; Mia Torchetti; Suelee Robbe-Austerman; Tod P. Stuber
EPI_ISL_9537091	Kaiser Permanente Southern California	Helix	Helix; Kaiser Permanente Southern California
EPI_ISL_6652122	Kantonsspital Winterthur	Institute of Medical Virology	Alexandra Trkola; Annette Audigé; Catharine Aquino; Cyril Shah; Daniel Ehrsam; Gabriela Ziltener; Guido Bloemberg; Hubert Rehrauer; Isabel Stürmer; Joel Wirz; Jon Huder; Jürg Böni; Kevin Steiner; Maria Grünberg; Maryam Zaheri; Michael Huber; Riccarda Capaul; Stefan Schmutz; Verena Kufner; Weihong Qi
EPI_ISL_8015366, EPI_ISL_8120134	Kantonsspital Winterthur	Institute of Medical Virology, University of Zurich	Alexandra Trkola; Annette Audigé; Catharine Aquino; Cyril Shah; Daniel Ehrsam; Gabriela Ziltener; Guido Bloemberg; Hubert Rehrauer; Isabel Stürmer; Joel Wirz; Jon Huder; Jürg Böni; Kevin Steiner; Maria Grünberg; Maryam Zaheri; Michael Huber; Riccarda Capaul; Stefan Schmutz; Verena Kufner; Weihong Qi
EPI_ISL_4235160, EPI_ISL_5065420, EPI_ISL_5065495	Klang Hospital	Division of Genomic Medicine and Innovation support,Department of Medical Sciences, Ministry of Public Health, Thailand	Archawin Rojanawiwat; Jirapha Pakdee; Natthakul Bunneang; Nuanjun Wichukchinda; Penpitcha Thawong; Pilailuk Akkapaiboon Okada; Pundharika Piboonsiri; Surakameth Mahasirimongkol; Waritta Sawaengdee

Appendix C

Chapter 4 Supplementary Information

Chapter 4 Supplementary Information:

GISAID sequence data

Figure C.1

Tables C.1-C.6 (Captions included below. Tables available in a supplementary excel file.)

GISAID sequence data

Sequence data from GISAID included in this study:

We prepared two GISAID EPI_SETs to acknowledge publicly available SARS-CoV-2 sequences included in this study. EPI_SET_240407tz includes all sequences used to generate the phylogenetic tree in Figure 2A, and additional individual GISAID sequences discussed in the text. EPI_SET_240407wv includes all additional sequences used to generate the phylogenetic tree in Supplementary Figure 1.

EPI_SET_240407tz Data Availability

GISAID Identifier: EPI_SET_240407tz doi: `10.55876/gis8.240407tz` All genome sequences and associated metadata in this dataset are published in GISAID's EpiCoV database. To view the contributors of each individual sequence with details such as accession number, Virus name, Collection date, Originating Lab and Submitting Lab and the list of Authors, visit `10.55876/gis8.240407tz`

EPI_SET_240407tz Data Snapshot

EPI_SET_240407tz is composed of 1,525 individual genome sequences. The collection dates range from 2019-12-31 to 2021-12-21; Data were collected in 4 countries and territories; All sequences in this dataset are compared relative to hCoV-19/Wuhan/WIV04/2019 (WIV04), the

official reference sequence employed by GISAID (EPI_ISL_402124). Learn more at https://gisaid.org/WIV04.

EPI_SET_240407wv Data Availability

GISAID Identifier: EPI_SET_240407wv doi: 10.55876/gis8.240407wv All genome sequences and associated metadata in this dataset are published in GISAID's EpiCoV database. To view the contributors of each individual sequence with details such as accession number, Virus name, Collection date, Originating Lab and Submitting Lab and the list of Authors, visit 10.55876/gis8.240407wv

EPI_SET_240407wv Data Snapshot

EPI_SET_240407wv is composed of 11,999 individual genome sequences. The collection dates range from 2019-12-31 to 2021-11-22; Data were collected in 26 countries and territories; All sequences in this dataset are compared relative to hCoV-19/Wuhan/WIV04/2019 (WIV04), the official reference sequence employed by GISAID (EPI_ISL_402124). Learn more at https://gisaid.org/WIV04.

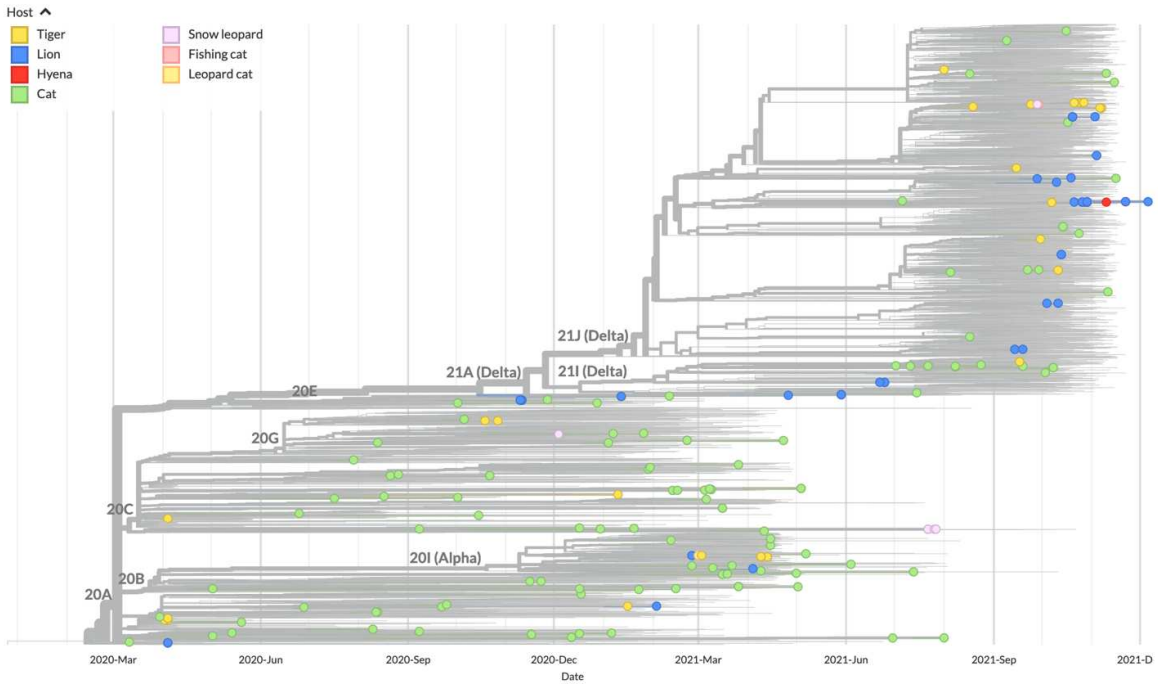


Figure C.1: No evidence at the global level for species-specific adaptation of SARS-CoV-2 in felids or hyenas. A time-based phylogenetic tree was generated from from all felid- and hyena-derived SARS-CoV-2 sequences available in the GISAID database prior to December 2021, with contextual human-derived sequences collected in the same time and place. Tree was generated with Nextstrain and an interactive version is available at https://nextstrain.org/community/laurabashor/DZSARS2. Sequences are colored by host species and scaled by time, with collection date indicated along the x-axis. four data inputs: (1) sixteen SARS-CoV-2 consensus sequences from the Denver Zoo outbreak corresponding to the highest quality sequence collected at the latest date from each individual animal, (2) a random subsample of human-derived sequences in Colorado from the days leading up to the zoo outbreak (100 sequences/day from September 23rd to October 7th, 2021; N=1500 sequences), (3) all felid-derived SARS-CoV-2 sequences available in the GISAID database prior to December 2021, and (4) contextual sequences collected in the same time and place as each felid sequence (10-100 human-derived sequence per felid-derived sequence) with enriched sampling within the United States. All data obtained from GISAID were processed with the Augur pipeline for input into Nextstrain, and acknowledgements are available in GISAID EPI_SET_240407wv.

Table C.1: Unique mutations identified relative to the Wuhan-1 reference sequence. Defining mutations of the AY.20 lineage and within-host (relative to the tiger reference) mutations are indicated.

Table C.2: Defining mutations of the AY.20 lineage listd by outbreak.info. Mutation table accessed here: `https://outbreak.info/situation-reports?pango=AY.20`

Table C.3: Summary of the quality and classification of SARS-CoV-2 genome sequencing datasets obtained from SARS-CoV-2 samples output by the nf-core/viralrecon pipeline.

Table C.4: Low frequency variant calling table for virus populations recovered from Denver Zoo animals, output by the nf-core/viralrecon pipeline.

Table C.5: Genome-level measures of nucleotide diversity in virus populations recovered from Denver Zoo animals, output by the SNPGenie pipeline.

Table C.6: Gene-level measures of nucleotide diversity in virus populations recovered from Denver Zoo animals, output by the SNPGenie pipeline.

Appendix D

Chapter 5 Supplementary Information

Chapter 5 Supplementary Information:

Tables D.1 and **D.3**

Tables D.2 and **D.4** (Captions included below. Tables available in a supplementary excel file.)

Table D.1: Significant differences in viral RNA levels by host species among SARS-CoV-2-infected wildlife samples subjected to NGS variant analysis. Estimates and Tukey-adjusted pairwise comparisons were obtained with the 'emmeans' R package.

	Viral RNA (Mean Ct)	
Predictors	*Estimates*	*p-value*
(Intercept)	28.4 ***	**<0.001**
species [Bushy-tailed woodrat]	1.70	0.529
species [Deer mouse]	-2.63	0.334
species [Mule deer]	-7.33 **	**0.005**
species [Red fox]	-8.01 ***	**0.001**
species [Striped skunk]	0.0208	0.993
Observations	29	
R^2 / R^2 adjusted	0.615 / 0.531	

Pairwise contrasts	*Estimates*	*SE*	*df*	*t-ratio*	*p-value*
bat - woodrat	-1.7	2.67	23	-0.64	0.987
bat - deer mouse	2.63	2.67	23	0.99	0.917
bat - mule deer	7.33 *	2.34	23	3.13	**0.048**
bat - red fox	8.01 *	2.1	23	3.82	**0.01**
bat - striped skunk	-0.02	2.34	23	-0.01	1
woodrat - deer mouse	4.34	2.85	23	1.52	0.655
woodrat - mule deer	9.03 *	2.55	23	3.54	**0.019**
woodrat - red fox	9.72 **	2.33	23	4.17	**0.004**
woodrat - striped skunk	1.68	2.55	23	0.66	0.985
deer mouse - mule deer	4.69	2.55	23	1.84	0.462
deer mouse - red fox	5.38	2.33	23	2.31	0.23
deer mouse - striped skunk	-2.65	2.55	23	-1.04	0.899
mule deer - red fox	0.69	1.95	23	0.35	0.999
mule deer - striped skunk	-7.35 *	2.21	23	-3.33	**0.031**
red fox - striped skunk	-8.03 **	1.95	23	-4.12	**0.005**

** p<0.05 ** p<0.01 *** p<0.001*

Table D.2: Assessment of the quality and classification of SARS-CoV-2 genome sequencing datasets obtained from SARS-CoV-2 samples from experimentally infected wildlife.

Table D.3: A list of animal IDs used in this study (Study ID) alongside the corresponding ID from published studies (Paper ID) and citation.

Study ID	Paper ID	Citation
Deer mouse A	Deer mouse B1	Bosco-Lauth et al. 2021
Deer mouse B	Deer mouse B2	
Deer mouse C	Deer mouse B3	
Woodrat A	Woodrat 1B	
Woodrat B	Woodrat 2B	
Woodrat C	Woodrat 3B	
Skunk A	Skunk 1	
Skunk B	Skunk 2	
Skunk C	Skunk 4	
Skunk D	Skunk 5	
Bat A	WA01 Bat 1	Bosco-Lauth et al. 2022
Bat B	WA01 Bat 5	
Bat C	WA01 Bat 6	
Bat D	WA01 Bat 7	
Fox A	Fox 1	Porter et al. 2022
Fox B	Fox 2	
Fox C	Fox 3	
Fox D	Fox 4	
Fox E	Fox 5	
Fox F	Fox 6	
Mule deer A	Mule deer 4	Porter et al. 2024
Mule deer B	Mule deer 3	
Mule deer C	Mule deer 6	
Mule deer D	Mule deer 5	
Mule deer E	Mule deer 2	

Table D.4: Low frequency variant calling table output by the nf-core/viralrecon pipeline for SARS-CoV-2 samples from experimentally infected wildlife, slightly modified for clarity.