Evolutionary rate of SARS-CoV-2 increases during zoonotic infection of farmed mink

Authors: Ashleigh F. Porter[1*], Damian F.J. Purcell[1], Benjamin P. Howden[1,2], Sebastian Duchene[1*]

Affiliations:

[1]Department of Microbiology and Immunology, The University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Melbourne, VIC, Australia

[2]Microbiological Diagnostic Unit Public Health Laboratory, The University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Melbourne, VIC, Australia

*Corresponding author. Email: ashleigh.porter@unimelb.edu.au, sebastian.duchene@unimelb.edu.au

Abstract: To investigate genetic signatures of adaptation to the mink host, we characterized the rate heterogeneity in mink-associated SARS-CoV-2. In 2020, the first detected anthropozoonotic spillover event of SARS-CoV-2 occurred in mink farms throughout Europe and North America. Both spill-back of mink-associated lineages into the human population and spread into surrounding wildlife was reported, highlighting the potential formation of a zoonotic reservoir. Our findings suggest the evolutionary rate of SARS-CoV-2 underwent an episodic increase upon introduction to the mink host before returning to the normal range. Furthermore, SARS-CoV-2 lineages could have circulated in the mink population for a month before detection, and during this period, evolutionary rate estimates of 6.57×10^{-3} could yield an 8-fold increase of mutations compared to the evolutionary rate of SARS-CoV-2 in humans. We suggest that SARS-CoV-2 undergoes a brief, but significant, increase in evolutionary rate in response to greater selective pressures during species jumps, emphasizing the necessity of monitoring zoonotic SARS-CoV-2 infections.

One-Sentence Summary: In 2020, SARS-CoV-2 outbreaks were reported in mink farms. During this period of adaptation, the evolutionary rate of SARS-CoV-2 increased for a short interval.

Main Text:

Coronaviruses (CoV) are zoonotic viruses associated with mammals and avian hosts (*1*) that are known for easily jumping species barriers, due to high mutation rates, a large RNA genome (*2, 3*) and interaction with multiple ACE2 receptors (*4, 5*).In recent decades, there have been 3 major outbreaks of CoV in the human population, causing epidemics: Severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS), and COVID-19. SARS-CoV and MERS-CoV are thought to have originated in bats before spreading to the human population, through an intermediate host (*6*). SARS-CoV-2 likely has zoonotic origins, hypothesised to initially spread from the Huanan seafood market in Wuhan, China (*7-9*). CoV that circulate in the Chinese horseshoe bats (*Rhinolophus affinis* and *Rhinolophus malayanus*) are the closest known relatives to SARS-CoV-2 (although, with estimated divergence from SARS-CoV-2 between 1948 and 1982) (*10-12*). The Malayan pangolin (*Manis javanica)* has been cautiously suggested to be an intermediate host (*13*), with much debate (*8, 14*). Minks are one of many animals (Table 1) that are susceptible to SARS-CoV-2 infection (*15, 16*).

In all zoonotic SARS-CoV-2 cases, human contact is likely the origin of transmission (*17*). Although many sporadic spillover cases have occurred (Table 1), the first detected anthropozoonotic spillover event of SARS-CoV-2 occurred in mink farms (Fig 1), with human-to-mink, mink-to-mink, and mink-to-human transmission networks established (*18*). The first report of SARS-CoV-2 in mink farms occurred late April 2020, in the Netherlands, followed by farms in Denmark during May, and in both countries the outbreaks were sequenced comprehensively (*17, 19-21*). Further outbreaks were seen across Europe (Denmark, France, Poland, Lithuania, Spain, Italy, Sweden and Greece) and North America (USA and Canada) (https://www.oie.int/en/what-we-offer/emergency-and-resilience/covid-19/, Accessed Oct 18). All mink SARS-CoV-2 outbreaks originated from human infections (*18*), with multiple introductions of the virus into the mink population (*17*), along with potential spread between farms (*18*). Mink-associated SARS-CoV-2 forms distinct clades, and the evolutionary rate of the virus is anticipated to increase in mink zoonotic transmission compared to human infections, due to adaptive pressure (*18, 20*).

Mink-associated SARS-CoV-2 has continued to spread, both into surrounding wildlife (*22*) and in "spill-back" cases from infected minks into the human population (*18*). Spill-back events present certain risks to public health – such as the spill-back of mink-associated lineages that acquired mutations in the spike protein receptor binding domain (which can lead to structural and/or functional changes in host receptor binding) (*20, 23-26*). The mink-associated "Cluster 5" lineage in Denmark, which had several mutations in the spike protein, spread widely in the human population surrounding the farms, causing 40% of COVID-19 cases in the region, prompting governments to cull all farmed mink (*14, 15, 27*). The escape of mink-associated SARS-CoV-2 into surrounding wildlife is also a major concern (*22*), particularly as many farms border habitats that have high wild mustelid populations (*28*). An example of a high-risk host for a wildlife reservoir of SARS-CoV-2 is deer mice (*Peromyscus maniculatus*), which have no clinical symptoms but have high SARS-CoV-2 replication levels (*29*). The establishment of a viral reservoir creates issues with controlling pathogens, which has been observed before, with the rabies virus forming a reservoir in wild-living raccoons and skunks (*30*).

Recent debate concerning the origin of the first observed Omicron lineage (BA.1 and BA.2) has also emphasized the possibility of zoonotic spill-back events (*31-33*). Although increased mutational rate is observed during persistent infections in immunocompromised patients receiving antiviral therapy (*34, 35*), the mutations in Omicron were unlikely to have

arisen during evolution in the human host (*13, 31, 32*). Specifically, 45 point mutations distinct to its nearest observed predecessor, lineage (B.1.1), have evolutionary signatures similar to mouse-adapted lineages - particularly in the spike protein sequence, where many mutations overlap with mutations arising from chronic SARS-CoV-2 infection in mice (*33*) that increase binding affinity to mouse ACE2 (*36*).

The spread of zoonotic SARS-CoV-2 to mink farms has highlighted the threat of spill-back events, particularly lineages that have undergone adaptation in animal populations, including the evolution of the virus in unmonitored reservoirs in wildlife populations (*15, 18*). Therefore, to explore the genetic signatures of adaptation to the mink host, we have estimated the evolutionary rate of mink-associated SARS-CoV-2 in comparison to the evolutionary rate observed within the broader SARS-CoV-2 phylogeny. The "evolutionary rate" we are referring to is the combination of substitutions and instantaneous mutations that are occurring in the genome over time, defined as the "evolutionary substitution rate" (*37*). We utilised a range of molecular clock models that have been used to study the pattern of evolutionary rate variation during SARS-CoV-2 variant of concern (VOC) emergence. Based on evidence that, during SARS-CoV-2 VOC emergence, there is an episodic increase of the evolutionary rate (*38*), we expect to see a similar pattern in mink-associated SARS-CoV-2.

Distinguishing mink-associated clades within the SARS-CoV-2 phylogeny

We compiled a dataset of complete genomes with 269 taxa (n=200 from human isolates, n=69 from mink isolates, Table S3) and was 29,839 bp in length. We estimated a maximum likelihood tree (Fig 2a), which revealed two distinct monophyletic mink-associated clades, as observed previously (*19*). A pattern of increased genetic distance was observed in the Netherlands mink-associated SARS-CoV-2 sequences in a root-to-tip regression (Fig 2b) .

An increased rate of evolution observed in stem branches

We used six molecular clock models in a Bayesian framework explore the evolutionary rate heterogeneity within mink-associated SARS-CoV-2 clades (Table 2). The strict clock (SC) assumes that all branches have the same evolutionary rate and is thus a null model. The uncorrelated lognormal clock (UCLC) assumes that branch rates were drawn from a lognormal distribution and is the most liberal of all models compared here. The four fixed local clock (FLC) models represent hypotheses of branch rate variation and separate branch rates as belonging to two categories: 'foreground' and 'background' (*38, 39*). Foreground branches are those assigned a different rate to the rest of the tree. The remaining branches in the tree, the background branches, represent the overall evolutionary dynamics of SARS-CoV-2 and serve as a comparison for the evolutionary rates estimated for the mink-associated clades, termed after their geographical origin, the Netherlands and Denmark.

The FLC stem model considers "foreground branches" as those along the stems leading up to mink lineages (either Netherlands or Denmark) and is consistent with an episodic change in the evolutionary rate (visualised in Fig S1). In the FLC clade & stem model the "foreground branches" includes both the stem branch leading up to mink lineages, and all branches within each independent mink clades, such that any changes in the evolutionary rate are maintained in the mink population (Fig S1). We also specified alternative parameterisation of these two models, but where the rate is shared amongst all mink clades, as in FLC (shared, stem) and FLC (shared, clade & stem).

The mean evolutionary rates for the strict, FLC (clade), FLC (stem) and FLC (clade & stem), and all shared FLC models sit within previous estimates of SARS-CoV-2 evolutionary rates (Table 2, Fig 3a) ranging between 7×10^{-4} and 1.1×10^{-3} subs/site/year (*40, 41*). In both the FLC (shared, stem) and FLC (stem) models, the mean background evolutionary rate estimates were respectively 4.86×10^{-4} (95% HPD: 4.34×10^{-4}, 5.2×10^{-4}) and 4.83×10^{-4} (95% HPD: 4.45×10^{-4}, 5.27×10^{-4}). For the mink-associated clades, there was a much faster rate of evolution estimated for FLC (stem) (Table 2, Fig 3b,c) and in the FLC (shared, stem) model, albeit with uncertainties that spanned several orders of magnitude, particularly for the Netherlands clade(Fig 4b). Strikingly the stem only estimates for the mink-associated clades were approximately 20-fold higher in both the FLC (stem), with 12×10^{-2} (95% HPD: 8.22×10^{-3}, 3.58×10^{-1}) and 4.92×10^{-2} (95% HPD: 2.04×10^{-5}, 2×10^{-2}) for the Netherlands and Denmark clades respectively, and 11.5×10^{-2} (95% HPD: 7.26×10^{-3}, 3.47×10^{-1}) in FLC (shared, stem). This pattern is also observed to a lesser extent in FLC (stem & clade) estimates for the Netherlands, however, the Denmark FLC (stem & clade) evolutionary rate was slower than average (Fig 3bc). The evolutionary rate observed in the mink samples in FLC (stem & clade, shared) appeared slightly faster than mean evolutionary rates (Fig 4b).

We also ran FLC models for independent and shared clades, FLC (clade) and FLC (clade, shared) (Supplementary Table 1) where estimates for the evolutionary rate within mink clades appear to have either a similar or slightly slower evolutionary rates when compared to the mean evolutionary rates (Fig 3bc).

Uncertainty in FLC (stems) model estimates

To explore the uncertainty observed in the mink-associated evolutionary rate estimates in the FLC (stems) model, we conducted prior sensitivity analysis on the clock rates in all FLC models, that matches recent estimates of the evolutionary rate of the virus and penalises very high rate values (Supplementary Table 2). We distinguished these FLC models as FLC(stem*), FLC(clade*), FLC(stem & clade*), FLC(shared, stem*), FLC(shared, clade*), and FLC(shared, stem & clade*). There was still an observable increase in the evolutionary rate in mink-associated clades in FLC (stem*) and FLC (shared, stem*) although to a lesser degree than initial estimates (Table 3, Fig 5). In all shared models, the Bayes factor (which is a calculation of posterior odds divided by prior odds, as a measure of statistical support in favour of hypothesis, Table 4) was > 17, with a value of >3.2 representing substantial evidence (*42*). Although in all models tested, the Bayes factor calculated for the Netherlands clade was > 10, this was not the case for the Denmark clade (Table 4).

Divergence of mink-associated clades

In all models, the time to most recent common ancestor (tMRCA) for the whole phylogeny is approximately mid-2019. The Netherlands clade tMRCA estimates ranged from the last few days of 2019 until mid-March 2020, and the Denmark clades following a similar pattern except for the UCLC tMRCA estimate falling much later, in May 2020.

Positive selection during host-adaptation in SARS-CoV-2

It was anticipated that upon introduction to the mink host, SARS-CoV-2 would undergo adaptive evolution, as seen before in SARS-CoV during adaptation to the human

host after jumping from palm civets (*Paradoxurus hermaphroditus*) (*43*). A substantial change in the evolutionary rate (an approximate 8-fold increase of mutations accumulated per month in comparison to previous estimates) was observed along the stem branch leading to the mink-associated clades. We hypothesise that the observed change in the evolutionary rate along the stem branches leading to the mink clades is evidence of positive selection occurring during the strong selective adaptation of SARS-CoV-2 to the mink host. Importantly, this adaptive phase is episodic, as this pattern of increased evolution does not appear to continue within the mink clades when not crossing a species barrier (Fig 3, Fig 4). A similar phenomenon was observed in SARS-CoV-2 variants of concern (VOC), where positive selection was observed along the stem branches (particularly in the case of the Alpha lineage) but not within the VOC clades (*38*).

We suggest that this increased rate of evolution along the stem branch could lead to a dramatic shift in the mutations accumulated in the lineages circulating amongst the farmed mink populations. The estimates produced under the FLC (shared, stem) model (Fig 4b) suggest the evolutionary rate was much more rapid in the mink clades, a trend which continued to be observed even with a more conservative prior in the FLC (shared, stem*) model (Fig 5c), with estimates of the evolutionary rate averaging 6.57×10^{-3} substitutions/site/year (95% HPD: 3.18×10^{-3}, 1.06×10^{-2}, Bayes Factor = ∞). Based on these estimates, the virus could accumulate roughly 16 mutations per month, which is a dramatic increase from the mean evolutionary rate of SARS-CoV-2 (approximately 2 mutations per month). However, we note that it is unexpected that the Denmark clade does not appear to have a strong signal for adaptive selection in comparison to the Netherlands clade (Table 4). There are many underlying factors within the mink farm outbreaks that could contribute to such a difference in results, for example, we do not know how many times SARS-CoV-2 spilled over into the mink population, or for how long it was circulating in the population before detection.

Signatures of animal-adaptation in zoonotic SARS-CoV-2

The broad zoonotic potential (Table 1) and generalist nature of SARS-CoV-2 has been emphasized (*44*), with minimal adaptation required for zoonotic spillovers in novel hosts. In mink populations, identical mutations have arisen independently in the virus (*45*). In SARS-CoV-2 isolates from both mink and white-tailed deer populations, there have been six mutations predicted to be associated with animal-adaptation (*44*) and 23 recurrent mutations (including 3 nonsynonymous mutations in the receptor-binding domain of the spike protein) have arisen at least 4 times in mink-associated SARS-CoV-2 but are rarely seen in human samples (*45*). This is a substantial number of mutations to have accumulated in such a short period, with previous estimates of the evolutionary rate of SARS-CoV-2 (*40, 41*) requiring a year to accumulate 23 mutations. Similarly, under previous evolutionary rate estimates, the 18 mutations observed along the stem branch leading to the Netherlands mink clade would have taken approximately a year to accrue, however, when accounting for a rate increase along the stem branch it is reduced to months, which is more accurate for the timeline of SARS-CoV-2 outbreaks in mink farms.

We assume that this increased rate of evolution would not be unique to the introduction of SARS-CoV-2 in the mink host, and that this phenomenon may be seen in other novel hosts where inter-host transmission is possible. For example, an additional anthropozoonotic spillover event (and potential reservoir of concern) is the wild and captive white-tailed deer (*Odocoileus virginianus*) population in North America (*46, 47*). A third of deer tested in Iowa, and 34 out of 36 deer tested positive at a captive cervid facility in Texas

were positive for SARS-CoV-2 RNA, with evidence of deer-to-deer transmission (*46*), (*47*). Furthermore, the adaption of SARS-CoV-2 to the rodent host, and subsequent spill-back into the human population, could explain the emergence of divergent Omicron lineage in late 2021 (*33, 36*). It has been suggested that, while adapting within a rodent host, SARS-CoV-2 accumulated mutations for approximately 12 months before re-entering the human population as the Omicron lineage (*33*) with increased infectivity (*48*), and higher level of "vaccine-breakthrough" (*49*).

Detection and surveillance of zoonotic SARS-CoV-2

Estimates of divergence for the Netherlands and Denmark clades suggest that they emerged in the first few months of 2020 (Table 2). The outbreaks were first detected in the Netherlands and Denmark during late April and early May respectively, however our estimates suggest that SARS-CoV-2 was circulating in the mink population a month before detection, or perhaps, the ancestral lineages that were present in the "stem branch" period were under-sampled. We note that it is likely that the lack of full diversity of mink-associated SARS-CoV-2 in our dataset means that the tMRCA estimated here is a lower bound.

Due to the magnitude of farmed mink populations (in both population size and geographical reach), in addition with the established transmission pathways (Fig 1) along with the ability of SARS-CoV-2 to accumulate potentially harmful mutations rapidly, zoonotic viral transmission poses a significant threat to global public health (*17, 18, 50*). Furthermore, the formation of a permanent reservoir of SARS-CoV-2 in wildlife populations could lead to spill-back events of animal-adapted lineages of the virus into the human population, and other susceptible animals (*51, 52*). Our work emphasises the necessity of surveillance: to track any zoonotic spread of SARS-CoV-2, identify outbreaks in novel hosts rapidly, and monitor ongoing spread of SARS-CoV-2 after host-switching to prevent the establishment of a viral reservoir. In particular, monitoring "at-risk" animal groups is essential, including farmed and wild-living populations of minks and white-tailed deer, animals that regularly come into contact with humans (*33, 53*), and species that host coronaviruses closely related to SARS-CoV-2 (such as members of the genus *Rhinolophus*) (*54*). This surveillance system relies greatly on whole genome sequencing, which has played a key role in monitoring the emergence and evolution of other variants (*38, 55, 56*). In this study we highlight the power of the plentiful whole genome sequencing of SARS-CoV-2 isolates collected during the 2020 outbreaks in mink farms and recommend that this remains a high priority for future zoonotic spillover events of SARS-CoV-2 so that we can continue to monitor the evolutionary dynamics of zoonotic SARS-CoV-2.

References

1. J. O. Wertheim, D. K. Chu, J. S. Peiris, S. L. Kosakovsky Pond, L. L. Poon, A case for the ancient origin of coronaviruses. *J Virol* **87**, 7039-7045 (2013).
2. M. M. Lai, D. Cavanagh, The molecular biology of coronaviruses. *Adv Virus Res* **48**, 1-100 (1997).
3. S. Su *et al.*, Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends Microbiol* **24**, 490-502 (2016).
4. X. Y. Ge *et al.*, Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* **503**, 535-538 (2013).
5. R. L. Graham, R. S. Baric, Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J Virol* **84**, 3134-3146 (2010).
6. J. Cui, F. Li, Z. L. Shi, Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol* **17**, 181-192 (2019).
7. D. S. Hui *et al.*, The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis* **91**, 264-266 (2020).
8. S.-L. Liu, L. Saif. (Multidisciplinary Digital Publishing Institute, 2020).
9. F. Wu *et al.*, A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269 (2020).
10. H. Zhou *et al.*, A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Current biology* **30**, 2196-2203. e2193 (2020).
11. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *nature* **579**, 270-273 (2020).
12. M. F. Boni *et al.*, Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nat Microbiol* **5**, 1408-1417 (2020).
13. T. Zhang, Q. Wu, Z. Zhang, Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Current biology* **30**, 1346-1351. e1342 (2020).
14. R. Frutos, C. Devaux, Mass culling of minks to protect the COVID-19 vaccines: is it rational? *New Microbes and New Infections* **38**, 100816 (2020).
15. C. S. Larsen, S. R. Paludan, Corona's new coat: SARS-CoV-2 in Danish minks and implications for travel medicine. *Travel Med Infect Dis* **38**, 101922 (2020).
16. B. B. Oude Munnink *et al.*, Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science* **371**, 172-177 (2021).
17. N. Oreshkova *et al.*, SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. *Eurosurveillance* **25**, 2-8 (2020).
18. B. B. O. Munnink *et al.*, Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science* **371**, 172-177 (2021).
19. A. S. Hammer *et al.*, SARS-CoV-2 Transmission between Mink (Neovison vison) and Humans, Denmark. *Emerg Infect Dis* **27**, 547-551 (2021).
20. L. Lu *et al.*, Adaptation, spread and transmission of SARS-CoV-2 in farmed minks and associated humans in the Netherlands. *Nat Commun* **12**, 6802 (2021).
21. H. D. Larsen *et al.*, Preliminary report of an outbreak of SARS-CoV-2 in mink and mink farmers associated with community spread, Denmark, June to November 2020. *Euro Surveill* **26**, 2100009 (2021).
22. K. J. Olival *et al.*, Possibility for reverse zoonotic transmission of SARS-CoV-2 to free-ranging wildlife: A case study of bats. *PLoS Pathog* **16**, e1008758 (2020).

23. A. J. Greaney *et al.*, Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host Microbe* **29**, 44-57 e49 (2021).
24. J. Lan *et al.*, Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215-220 (2020).
25. J. Shang *et al.*, Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221-224 (2020).
26. S. Burkholz *et al.*, Paired SARS-CoV-2 spike protein mutations observed during ongoing SARS-CoV-2 viral transfer from humans to minks and back to humans. *Infect Genet Evol* **93**, 104897 (2021).
27. F. Fenollar *et al.*, Mink, SARS-CoV-2, and the Human-Animal Interface. *Front Microbiol* **12**, 663815 (2021).
28. S. A. Shriner *et al.*, SARS-CoV-2 Exposure in Escaped Mink, Utah, USA. *Emerg Infect Dis* **27**, 988-990 (2021).
29. A. Fagre *et al.*, SARS-CoV-2 infection, neuropathogenesis and transmission among deer mice: Implications for reverse zoonosis to New World rodents. *bioRxiv*, 2020.2008.2007.241810 (2020).
30. C. E. Rupprecht, J. S. Smith, M. Fekadu, J. E. Childs, The ascension of wildlife rabies: a cause for public health concern or intervention? *Emerg Infect Dis* **1**, 107-114 (1995).
31. E. Callaway, Heavily mutated Omicron variant puts scientists on alert. *Nature* **600**, 21 (2021).
32. K. Kupferschmidt. (American Association for the Advancement of Science, 2021).
33. C. Wei *et al.*, Evidence for a mouse origin of the SARS-CoV-2 Omicron variant. *J Genet Genomics* **48**, 1111-1121 (2021).
34. B. Choi *et al.*, Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. *New England Journal of Medicine* **383**, 2291-2293 (2020).
35. S. A. Kemp *et al.*, Neutralising antibodies in Spike mediated SARS-CoV-2 adaptation. *medRxiv*, (2020).
36. E. Cameroni *et al.*, Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature*, 1-9 (2021).
37. S. Y. Ho *et al.*, Time-dependent rates of molecular evolution. *Mol Ecol* **20**, 3087-3101 (2011).
38. J. H. Tay, A. F. Porter, W. Wirth, S. Duchene, The emergence of SARS-CoV-2 variants of concern is driven by acceleration of the evolutionary rate. *medRxiv*, 2021.2008.2029.21262799 (2021).
39. M. Worobey, G. Z. Han, A. Rambaut, A synchronized global sweep of the internal genes of modern avian influenza virus. *Nature* **508**, 254-257 (2014).
40. S. Duchene *et al.*, Temporal signal and the phylodynamic threshold of SARS-CoV-2. *Virus Evol* **6**, veaa061 (2020).
41. M. Ghafari *et al.*, Purifying selection determines the short-term time dependency of evolutionary rates in SARS-CoV-2 and pH1N1 influenza. *medRxiv*, 2021.2007.2027.21261148 (2021).
42. R. Kass, A. Raftery, Bayes Factors, in Journal of the American Statistical Association. (1995).
43. J.-F. He *et al.*, Road Z. *Hospital N, Med-F, Hospital R, Road B, Jiang Z, Tech H, Jiao-S*, 1666-1669 (2004).
44. C. C. S. Tan *et al.*, Transmission of SARS-CoV-2 from humans to animals and potential host adaptation. *bioRxiv*, 2020.2011.2016.384743 (2022).

45. L. van Dorp *et al.*, Recurrent mutations in SARS-CoV-2 genomes isolated from mink point to rapid host-adaptation. *bioRxiv*, (2020).
46. S. V. Kuchipudi *et al.*, Multiple spillovers and onward transmission of SARS-Cov-2 in free-living and captive White-tailed deer (Odocoileus virginianus). *bioRxiv*, 2021.2010.2031.466677 (2021).
47. C. M. Roundy *et al.*, High seroprevalence of SARS-CoV-2 in white-tailed deer (Odocoileus virginianus) at one of three captive cervid facilities in Texas. *bioRxiv*, 2022.2001.2005.475172 (2022).
48. J. Chen, R. Wang, N. B. Gilby, G. W. Wei, Omicron (B.1.1.529): Infectivity, vaccine breakthrough, and antibody resistance. *ArXiv*, (2021).
49. N. Andrews *et al.*, Effectiveness of COVID-19 vaccines against the Omicron (B. 1.1. 529) variant of concern. *MedRxiv*, (2021).
50. K. Sharun, R. Tiwari, S. Natesan, K. Dhama, SARS-CoV-2 infection in farmed minks, associated zoonotic concerns, and importance of the One Health approach during the ongoing COVID-19 pandemic. *Vet Q* **41**, 50-60 (2021).
51. R. J. Delahay *et al.*, Assessing the risks of SARS-CoV-2 in wildlife. *One Health Outlook* **3**, 7 (2021).
52. C. Manes, R. Gollakner, I. Capua, Could Mustelids spur COVID-19 into a panzootic? *Veterinaria italiana*, (2020).
53. P. Zhou, Z. L. Shi, SARS-CoV-2 spillover events. *Science* **371**, 120-122 (2021).
54. D. Delaune *et al.*, A novel SARS-CoV-2 related coronavirus in bats from Cambodia. *Nat Commun* **12**, 6563 (2021).
55. C. A. Adlhoch, Erik; Beaute, Julien ; Broberg, Eeva ; Cenciarelli, Orlando ; EinöderMoreno, Margot ; Fleming, Catherine ; Gossner, Céline ; Johansen, Kari; Mirinaviciute, *et al.*, Detection of new SARS-CoV-2 variants related to mink. *Eur. Cent. Dis. Prev. Control*, (2020).
56. E. B. Hodcroft *et al.*, Emergence and spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *medRxiv*, 2020.2010.2025.20219063 (2021).
57. W. Deng *et al.*, Primary exposure to SARS-CoV-2 protects against reinfection in rhesus macaques. *Science* **369**, 818-823 (2020).
58. V. J. Munster *et al.*, Respiratory disease in rhesus macaques inoculated with SARS-CoV-2. *Nature* **585**, 268-272 (2020).
59. B. Rockx *et al.*, Comparative pathogenesis of COVID-19, MERS, and SARS in a nonhuman primate model. *Science* **368**, 1012-1015 (2020).
60. C. Woolsey *et al.*, Establishment of an African green monkey model for COVID-19 and protection against re-infection. *Nature Immunology* **22**, 86-98 (2021).
61. S. Lu *et al.*, Comparison of nonhuman primates identified the suitable model for COVID-19. *Signal Transduct Target Ther* **5**, 157 (2020).
62. P. J. Halfmann *et al.*, Transmission of SARS-CoV-2 in Domestic Cats. *N Engl J Med* **383**, 592-594 (2020).
63. I. Ruiz-Arrondo *et al.*, 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. *Transbound Emerg Dis* **68**, 973-976 (2021).
64. J. Shi *et al.*, Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. *Science* **368**, 1016-1020 (2020).
65. Y. I. Kim *et al.*, Infection and Rapid Transmission of SARS-CoV-2 in Ferrets. *Cell Host Microbe* **27**, 704-709 e702 (2020).
66. K. Schlottau *et al.*, Experimental transmission studies of SARS-CoV-2 in fruit bats, ferrets, pigs and chickens. (2020).

67. J. F. Chan *et al.*, 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. *Clin Infect Dis* **71**, 2428-2446 (2020).
68. B. L. Haagmans *et al.*, SARS-CoV-2 Neutralizing Human Antibodies Protect Against Lower Respiratory Tract Disease in a Hamster Model. *J Infect Dis* **223**, 2020-2028 (2021).
69. S. F. Sia *et al.*, Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. *Nature* **583**, 834-838 (2020).
70. A. Z. Mykytyn *et al.*, Susceptibility of rabbits to SARS-CoV-2. *Emerg Microbes Infect* **10**, 1-7 (2021).
71. B. D. Griffin *et al.*, SARS-CoV-2 infection and transmission in the North American deer mouse. *Nat Commun* **12**, 3612 (2021).
72. T. H. Sit *et al.*, Canine SARS-CoV-2 infection. *Nature* **586**, 776 (2020).
73. L. Wang *et al.*, Complete genome sequence of SARS-CoV-2 in a tiger from a US zoological collection. *Microbiology resource announcements* **9**, e00468-00420 (2020).
74. K. N. Koeppel *et al.*, SARS-CoV-2 Reverse Zoonoses to Pumas and Lions, South Africa. *Viruses* **14**, 120 (2022).
75. Y. Zhao *et al.*, Susceptibility of tree shrew to SARS-CoV-2 infection. *Sci Rep* **10**, 16007 (2020).
76. S. Elbe, G. Buckland-Merrett, Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob Chall* **1**, 33-46 (2017).
77. Y. Shu, J. McCauley, GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill* **22**, (2017).
78. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772-780 (2013).
79. K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059-3066 (2002).
80. B. Q. Minh *et al.*, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol* **37**, 1530-1534 (2020).
81. A. Rambaut, T. T. Lam, L. Max Carvalho, O. G. Pybus, Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* **2**, vew007 (2016).
82. A. J. Drummond, A. Rambaut, BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214 (2007).

Acknowledgments: This work was performed using the LIEF HPC-GPGPU Facility hosted by the University of Melbourne which was established via assistance from LIEF Grant LE170100200. We are grateful to the contributors and moderators of GISAID EpiCoV, especially the laboratories who contributed the sequences that were used in this analysis.

Funding:

Australian Research Council (DE190100805) (SD)

Medical Research Future Fund (MRF9200006) (SD, BH)

Author contributions:

Conceptualization: AFP

Methodology: AFP, SD

Investigation: AFP, SD

Visualization: AFP, SD

Funding acquisition: SD, BPH

Writing – original draft: AFP, SD

Writing – review & editing: AFP, SD, DP, BPH

Competing interests: Authors declare that they have no competing interests

Data and materials availability: All data are available in the main text or the supplementary materials.

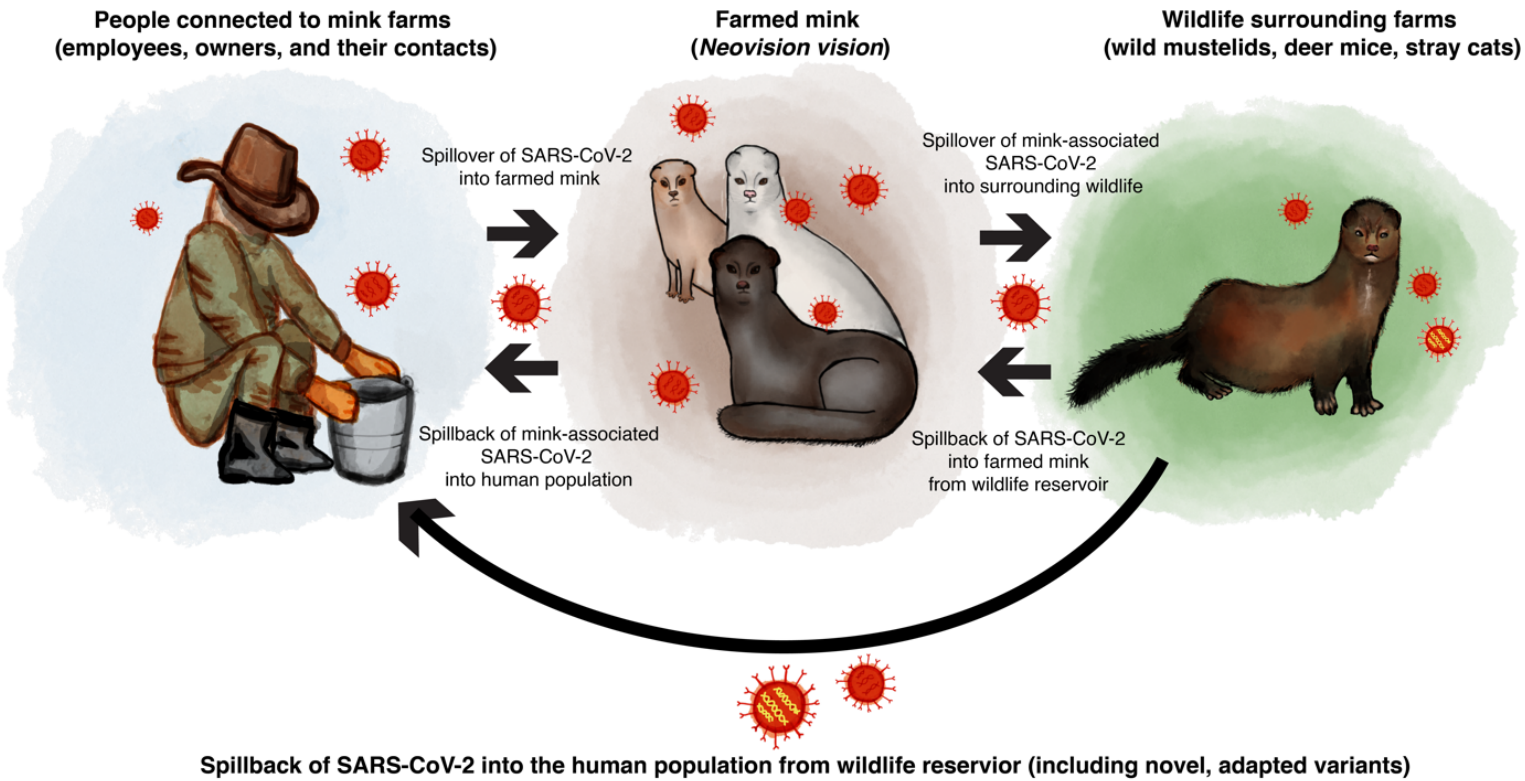

Fig 1. Illustration of the anthropozoonotic dynamics of mink-associated SARS-CoV-2, including human-to-mink, mink-to-mink, and mink-to-human transmission networks. To highlight the risk of viral reservoirs, we have shown a potential scenario of spill-back of mink-adapted lineages into the human population from infected wildlife.

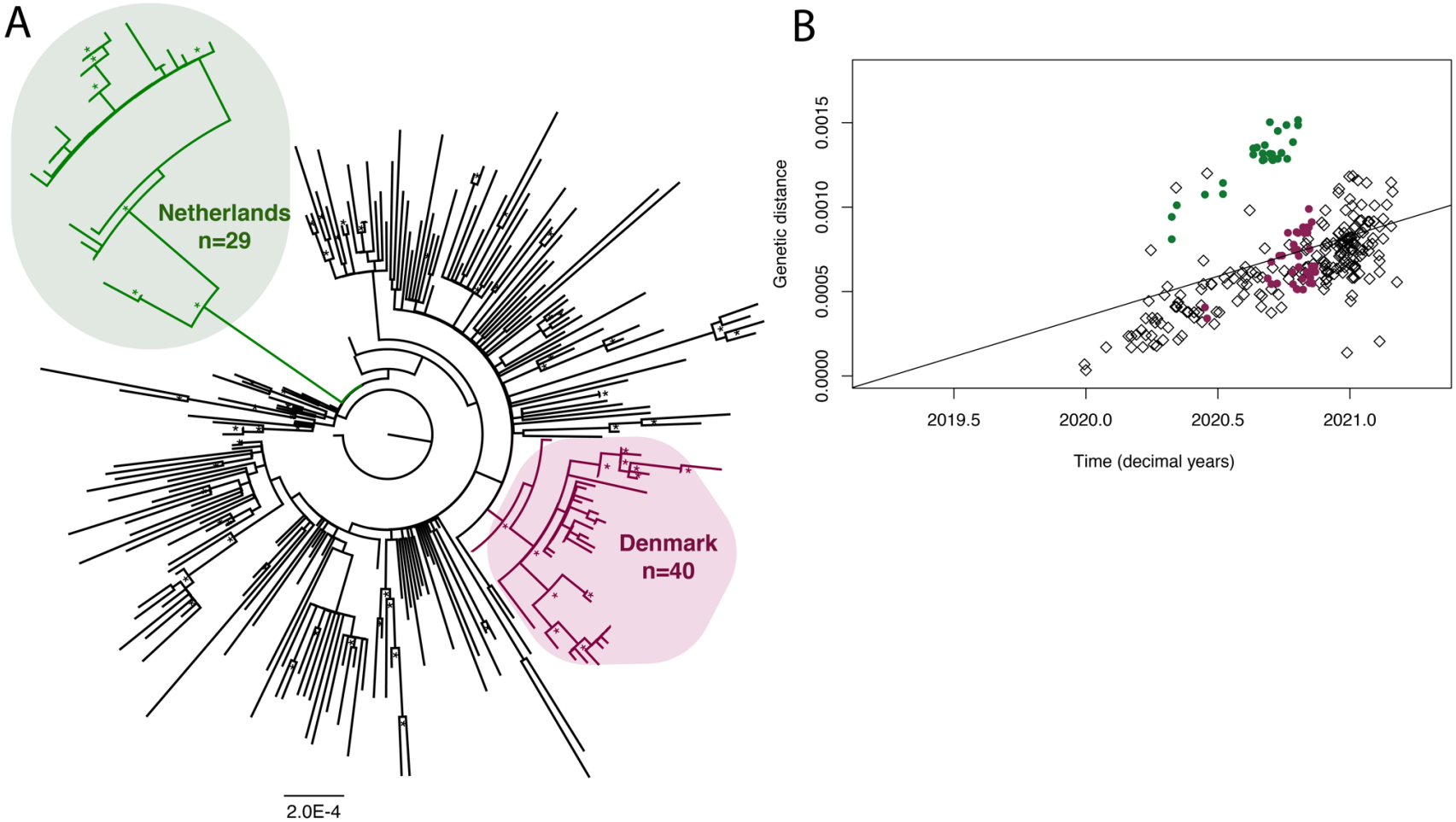

Fig 2. Phylogenetic analysis of mink-associated SARS-CoV-2 genomes, focusing on two geographical outbreaks from the Netherlands and Denmark. (A) Maximum likelihood tree of SARS-CoV-2 sequences (n=269). The two mink associated clades are colored: the Netherlands clade (n=29) is highlighted in green and Denmark clade (n=40) in purple, with remaining human isolates in black (n=200). The tree is rooted with SARS-CoV-2 reference sequence Wuhan/IVDC-HB-04/EPI_ISL_402120. Bootstrap replicates (n=1000) >70% are marked with an

asterisk. The scale bar represents substitutions per site (subs/site). (B) A root-to-tip regression plot of the genetic distance vs time (in decimal years) of the 269 SARS-CoV-2 genomes used in this analysis. Sequences in the Netherlands group (n=29) are represented by green dots, and Denmark sequences (n=40) are represented by purple dots, with remaining human isolates represented by black diamonds (n=200).

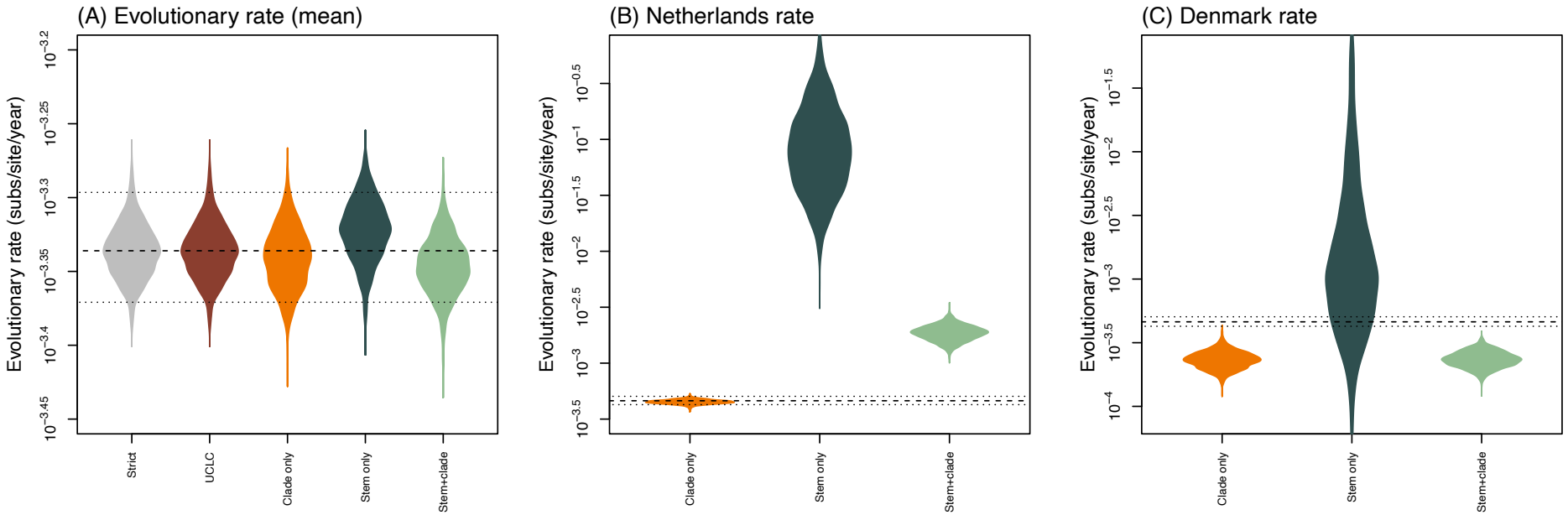

Fig 3. Violin plots of posterior statistics for the evolutionary rates (substitutions/site/year) estimated during model testing. (A) The mean evolutionary rates from the strict clock (grey), uncorrelated relaxed clock (brown), and FLC models: clade only (orange), stem only (dark green) and stem and clade (light green). Evolutionary rates estimated for the Netherlands clade (B) and Denmark clade (C) using FLC models: clade only (orange), stem only (dark green) and stem and clade (light green). Dashed lines represent the mean evolutionary rate (and 95% HPD intervals) from the strict clock.

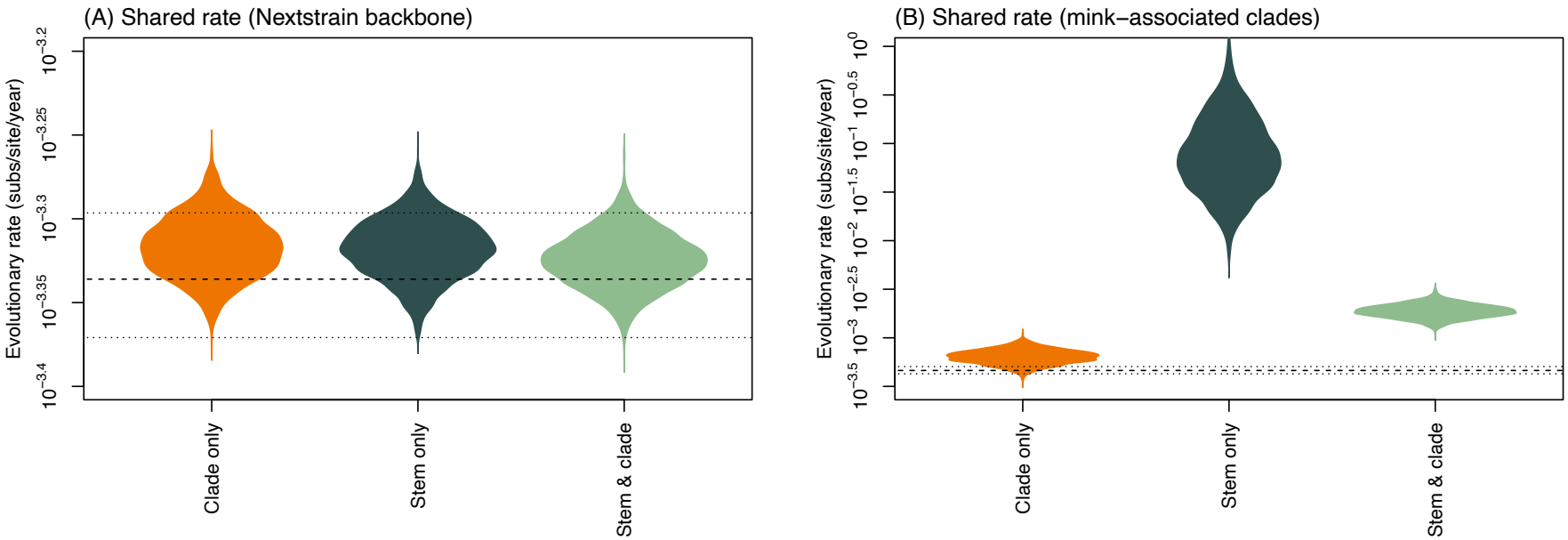

Fig 4. Violin plots of posterior statistics for the shared evolutionary rates (substitutions/site/year) estimated during model testing. (A) The evolutionary rates from FLC models using a shared rate for all background clades: clade only (orange), stem only (dark green) and stem and clade (light green). (B) Evolutionary rates from FLC models using a shared rate for all mink clades: clade only (orange), stem only (dark green) and stem and clade (light green). Dashed lines represent the mean evolutionary rate (and 95% HPD intervals) from the strict clock.

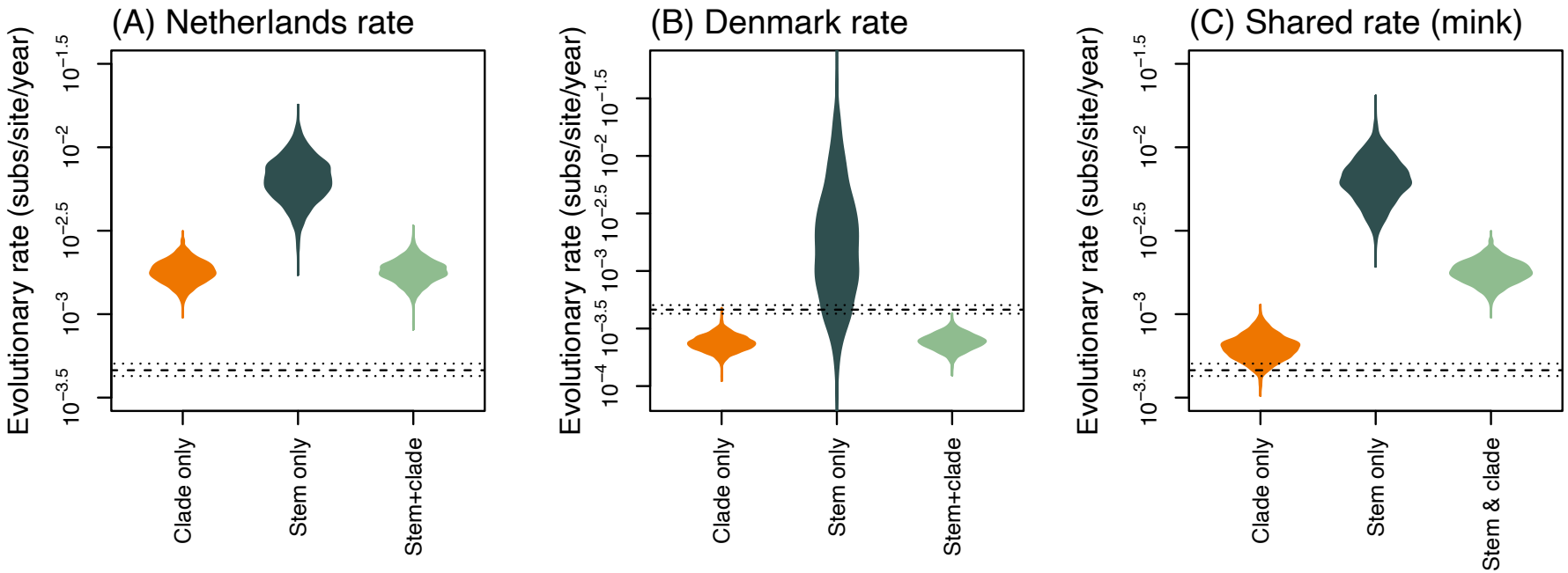

Fig 5. Violin plots of posterior statistics for the evolutionary rates (substitutions/site/year) estimated from FLC models using a prior on clock rate (termed FLC(stem*), FLC(clade*), FLC(stem & clade*), FLC(shared, stem*), FLC(shared, clade*), and FLC(shared, stem & clade*)). (A) The evolutionary rates from individual analysis of the Netherlands clade estimated from FLC clade only (orange), stem only (dark green) and stem and clade (light green). (B) The evolutionary rates from individual analysis of the Denmark clade estimated from FLC clade only (orange), stem only (dark green) and stem and clade (light green). (C) Evolutionary rates from shared analysis of mink clades estimated from FLC models clade only (orange), stem only (dark green) and stem and clade (light green). Dashed lines represent the mean evolutionary rate (and 95% HPD intervals) from the strict clock.

Table 1. Common and species names of animals susceptible to SARS-CoV-2 infection.

Common name (scientific name)	Reference
Mink (*Neovision vision*)	(*15, 16*)
Rhesus macaques (*Macaca mulatta*)	(*57, 58*)
Cynomolgus macaques (*Macaca fascicularis*)	(*59*)
African green monkeys (*Chlorocebus aethiops*)	(*60*)
Common marmosets *(Callithrix jacchus*)	(*61*)
Domestic cats (*Felis catus*)	(*62-64*)
Domestic ferrets *(Mustela putoruis*)	(*65, 66*)
Golden Syrian hamsters (*Mesocricetus auratus*)	(*67-69*)
Domestic rabbits (*Oryctolagus cuniculus domesticus*)	(*70*)
Deer mice (*Peromyscus spp.*)	(*29, 71*)
Domestic dogs (*Canis lupis familiaris)*	(*64, 72*)
Malayan tiger (*Panthera tigris jacksoni*)	(*73*)
African lion (*Panthera leo leo*)	(*74*)
Tree shrews (*Tupaia belangeris*)	(*75*)
White tailed deer (*Odocoileus virginianus*)	(*46, 47*)
Fruit bats (*Rousettus aegyptiacus)*	(*64, 66*)

Table 2. Estimates generated from each molecular clock model (SC = strict clock, UCLC = uncorrelated lognormal clock, FLC = fixed local clock). Estimates include time to most common recent ancestor (tMRCA) of the whole phylogeny, tMCRA of the Netherlands and Denmark clades, and the evolutionary rates (substitution/site/year) estimated for the whole phylogeny, and the Netherlands and Denmark clades. The 95% HPD interval is shown in brackets.

Model	tMRCA	Netherlands tMRCA	Denmark tMRCA	Estimated evolutionary rate (mean)	Netherlands evolutionary rate	Denmark evolutionary rate
SC	**17-07-2019** [02-07-2019, 01-10-2019]	**03-03-2020** [13-02-2020, 21-03-2020]	**24-02-2020** [04-01-2020, 12-04-2020]	**4.66** $\times 10^{-4}$ [4.25×10^{-4}, 5.07×10^{-4}]		
UCLC	**28-06-2019** [22-03-2019, 09-10-2019]*	**24-12-2019** [13-09-2019, 06-03-2020]	**04-05-2020** [28-03-2020, 06-06-2020]	**7.1** $\times 10^{-4}$ [6.35×10^{-4}, 7.87×10^{-4}]		
FLC (stem)	**20-07-2019** [02-07-2019, 18-08-2019]	**17-02-2020** [26-01-2020, 06-03-2020]*	**17-02-2020** [28-12-2019, 08-04-2020]	**4.77** $\times 10^{-4}$ [4.34×10^{-4}, 5.2×10^{-4}]*	**1.2** $\times 10^{-1}$ [8.22×10^{-3}, 3.58×10^{-1}]	**4.85** $\times 10^{-3}$ [2.04×10^{-5}, 2×10^{-2}]
FLC (stem & clade)	**17-07-2019** [02-07-2019, 17-09-2019]	**14-03-2020** [24-02-2020, 28-03-2020]	**01-01-2020** [10-11-2019, 13-02-2020]*	**4.5** $\times 10^{-4}$ [4.08×10^{-4}, 4.91×10^{-4}]*	**1.86** $\times 10^{-3}$ [1.3×10^{-3}, 2.45×10^{-3}]	**2.37** $\times 10^{-4}$ [1.67×10^{-4}, 3.1×10^{-4}]
FLC (shared, stem)	**24-07-2019** [02-07-2019, 22-08-2019]	**17-02-2020** [30-01-2020, 06-03-2020]*	**28-02-2020** [15-01-2020, 16-04-2020]	**4.86** $\times 10^{-4}$ [4.45×10^{-4}, 5.27×10^{-4}]*	**1.15** $\times 10^{-1}$ [7.26×10^{-3}, 3.47×10^{-1}]	
FLC (shared, stem & clade)	**17-07-2019** [02-07-2019, 20-09-2019]	**17-03-2020** [28-02-2020, 01-04-2020]	**21-02-2020** [01-01-2020, 05-04-2020]	**4.74** $\times 10^{-4}$ [4.35×10^{-4}, 5.14×10^{-4}]	**1.91** $\times 10^{-3}$ [1.33×10^{-3}, 2.5×10^{-3}]	

* ESS < 200

Table 3. Estimates generated from local clock (FLC) models with a gamma prior on the clock rate. Estimates include the evolutionary rates (substitution/site/year) estimated for the whole phylogeny, and the Netherlands and Denmark clades. The 95% HPD interval is shown in brackets.

Model	Estimated evolutionary rate (mean)	Netherlands evolutionary rate	Denmark evolutionary rate
FLC (stem*)	$\mathbf{4.67 \times 10^{-4}}$ [4.35×10^{-4}, 5.17×10^{-4}]*	$\mathbf{6.59 \times 10^{-3}}$ [2.96×10^{-3}, 1.06×10^{-2}]	$\mathbf{3.16 \times 10^{-3}}$ [4.77×10^{-5}, 1.14×10^{-2}]
FLC (shared, stem*)	$\mathbf{4.73 \times 10^{-4}}$ [4.33×10^{-4}, 5.14×10^{-4}]*	$\mathbf{6.57 \times 10^{-3}}$ [3.18×10^{-3}, 1.06×10^{-2}]	

* ESS < 200

Table 4. Bayes factor for each molecular clock model used in this study.

Model	Bayes factor	
	Netherlands	Denmark
FLC(stem)	∞	4.8
FLC(stem & clade)	∞	1
FLC(clade)	10.2	1
FLC(stem*)	∞	8
FLC(stem & clade*)	∞	1
FLC(clade*)	10.6	1
FLC(shared, stem)	∞	
FLC(shared, stem & clade)	∞	
FLC(shared, clade)	17.24	
FLC(shared, stem*)	∞	
FLC(shared, stem & clade*)	∞	
FLC(shared, clade*)	18.41	

Science
AAAS

Supplementary Materials for

Evolutionary rate of SARS-CoV-2 increases during zoonotic infection of farmed mink

Ashleigh F. Porter, Damian F.J. Purcell, Benjamin P. Howden & Sebastian Duchene
Correspondence to: ashleigh.porter@unimelb.edu.au, sebastian.duchene@unimelb.edu.au

This PDF file includes:

Materials and Methods

Data collection

We downloaded a subset of SARS-CoV-2 isolates that were collected from minks (n=100) and along with a subset (n=200) of the Nextstrain backbone diversity (Tay et al., 2021) collected from GISAID (Table S3) (*76, 77*). We generated an alignment in MAFFT v7 (*78, 79*) and a maximum likelihood phylogenetic tree using IQ-TREE2 (*80*). From this tree, we only retained mink genomes that clustered in large monophyletic clades (>20 sequences, >99% bootstrap support) and that did not contain any human isolates (n=69). Specifically, the mink sequences retained were defined into two geographically distinct clades, from the Netherlands (n=29) and Denmark (n=40). We used TempEst (*81*) to explore the temporal signal of the dataset.

Molecular clock model testing

We used a range of Bayesian molecular clock models (Table S4) to examine the patterns the evolutionary rate variation between the mink and human clades. The models range in describing the evolutionary rate along specific branches within phylogenetic trees (Supplementary Figure 1). We set up these models in BEAST 1.10 (*82*) using a Markov chain Monte Carlo of length 5×10^6 and sampling every 5000 steps.

Of the models tested, the strict and relaxed (uncorrelated lognormal distribution) clock models range respectively from the simplest model (parameter n=1) to the most complex (parameter n= number of branches+2). The other models we applied, fixed local clock (FLC) models, enable evolutionary hypotheses as they require the definition of which branches will share an evolutionary rate *a priori.* This definition is usually based on a biological assumption, for example, a variant of concern (VOC) lineage of SARS-CoV-2 to have a higher evolutionary rate than other lineages (*38*). We used six FLC models which have been described in detail in Tay et al. The first allows the evolutionary rate to vary within the mink clades, FLC (clade), or clades and along the stem, FLC (clade & stem), or only along the stem, FLC (stem). Additionally, these configurations were repeated where these rates could be shared with all mink clades, as in FLC (shared, stem), FLC (shared, clade) and FLC (shared, clade & stem). The biological theory behind models restricted to rate variability along the stem branches of mink clades is that the evolutionary rate is likely to increase over a short period of time during adaptation to a new host, then returns to the background rate, as in FLC (stem). The rate could also vary along the stem and within the clade of a new host, such as the FLC (clade & stem) model, or only within the clade, as in FLC (clade). In all models, we included *a priori* knowledge that the tMRCA of SARS-CoV-2 is estimated to be November 2019 (*40, 41*), with a prior distribution on the age of the root (Table S5).

To assess prior sensitivity of the FLC models, we specified a more informative rate prior for the mink-associated clades (Table S5). We used a gamma distribution with shape=1 and scale=10^{-2} such that the 95% percentile range was 2.5×10^{-5} to 3.7×10^{-3}, and thereby imposing a stronger penalty on high rates than the default prior in BEAST1.10 (*82*). This was used to determine if the increased evolutionary rate observed in mink clades was not an artefact of the relatively wide distribution of the initial clock rate prior (discussed in Tay et al. 2022). To estimate the marginal likelihood of each model, we randomly subsampled the total dataset (n=169) and used path sampling. We also calculated the Bayes factor for each model

by dividing the prior odds (the average of how much the rate of the clade of interest was greater than the clock rate) by the posterior odds.

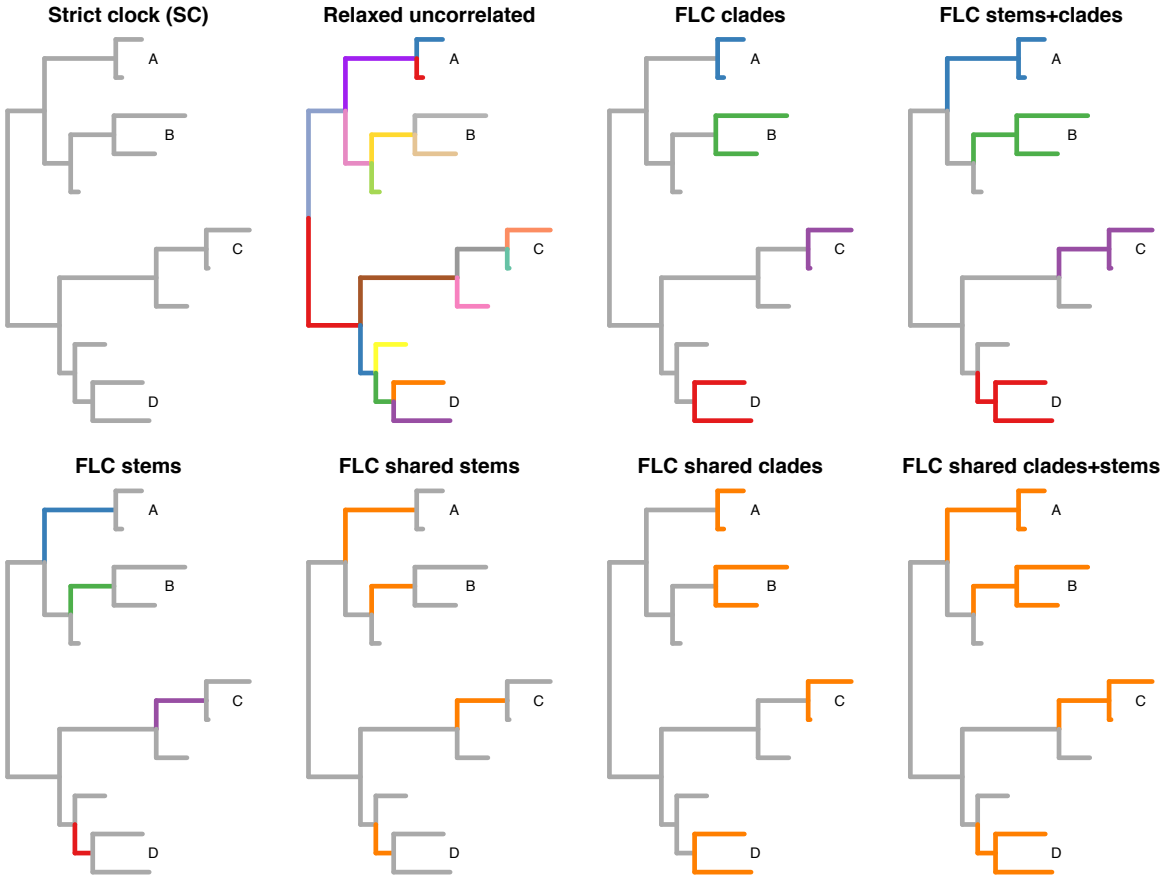

Fig. S1. **Visualization of the molecular model testing methodology.** Methods have been modified from Tay et al. (2022), with eight molecular clock models applied in hypothetical trees with 4 clades of interest (A, B, C and D). Figure adapted from Tay et al. (2022).

Table S1. Results of model selection of clade only individual and shared FLC models including estimates tMRCA of phylogeny, tMCRA of the Netherlands and Denmark clades, and the evolutionary rates for the whole phylogeny, and the Netherlands and Denmark clades. The 95% HPD interval is shown in brackets.

Model	tMRCA	Netherlands tMRCA	Denmark tMRCA	Estimated evolutionary rate (mean)	Netherlands evolutionary rate	Denmark evolutionary rate
FLC (clade)	**20-07-2019** [02-07-2019, 24-09-2019]	**14-03-2020** [17-02-2020, 25-03-2020]	**28-12-2019** [10-11-2019,13-02-2020]*	$\mathbf{4.54 \times 10^{-4}}$ [4.13×10^{-4}, 4.91×10^{-4}]*	$\mathbf{6.06 \times 10^{-4}}$ [4.12×10^{-4}, 8.23×10^{-3}]	$\mathbf{2.3 \times 10^{-4}}$ [1.65×10^{-4}, 3×10^{-4}]
FLC (shared, clade)	**20-07-2019** [02-07-2019, 15-08-2019]	**06-03-2020** [17-02-2020, 25-03-202]	**24-02-2020** [01-01-2020, 08-04-2020]	$\mathbf{4.77 \times 10^{-4}}$ [4.36×10^{-4}, 5.17×10^{-4}]*	$\mathbf{6.33 \times 10^{-4}}$ [4.28×10^{-4}, 8.45×10^{-4}]	

* ESS < 200

Table S2. Estimates generated from FLC clock models with a gamma prior on the clock rate. Estimates include the evolutionary rates (substitution/site/year) estimated for the whole phylogeny, and the Netherlands and Denmark clades. The 95% HPD interval is shown in brackets.

Model	Estimated evolutionary rate (mean)	Netherlands evolutionary rate	Denmark evolutionary rate
FLC (clade*)	$\mathbf{4.75 \times 10^{-4}}$ [4.32×10^{-4}, 5.15×10^{-4}]*	$\mathbf{6.08 \times 10^{-4}}$ [4.08×10^{-4}, 8.2×10^{-4}]	$\mathbf{2.36 \times 10^{-4}}$ [1.69×10^{-4}, 3.1×10^{-4}]
FLC (stem & clade*)	$\mathbf{4.64 \times 10^{-4}}$ [4.26×10^{-4}, 5.06×10^{-4}]*	$\mathbf{1.81 \times 10^{-3}}$ [1.29×10^{-3}, 2.38×10^{-3}]	$\mathbf{2.38\times 10^{-4}}$ [1.7×10^{-4}, 3.12×10^{-4}]
FLC (shared, clade*)	$\mathbf{4.76 \times 10^{-4}}$ [4.4×10^{-4}, 5.17×10^{-4}]*	$\mathbf{6.37 \times 10^{-4}}$ [4.4×10^{-4}, 8.53×10^{-4}]	
FLC (shared, stem & clade*)	$\mathbf{4.71 \times 10^{-4}}$ [4.33×10^{-4}, 5.09×10^{-4}]	$\mathbf{1.83 \times 10^{-3}}$ [1.29×10^{-3}, 2.4×10^{-3}]	
* ESS < 200			

Table S3. SARS-CoV-2 sequences used in this study.

GISAID Accession ID
EPI_ISL_1001001
EPI_ISL_1001002
EPI_ISL_849377
EPI_ISL_430818
EPI_ISL_856766
EPI_ISL_1139148
EPI_ISL_1139151
EPI_ISL_1013818
EPI_ISL_522686
EPI_ISL_678316
EPI_ISL_1121039
EPI_ISL_812423
EPI_ISL_854745
EPI_ISL_885139
EPI_ISL_806729
EPI_ISL_794718
EPI_ISL_667780
EPI_ISL_944741
EPI_ISL_654807
EPI_ISL_467989
EPI_ISL_516550
EPI_ISL_740874
EPI_ISL_767859
EPI_ISL_849747
EPI_ISL_854748
EPI_ISL_693297
EPI_ISL_812436
EPI_ISL_451613
EPI_ISL_530231
EPI_ISL_414414
EPI_ISL_755571
EPI_ISL_513911
EPI_ISL_544964
EPI_ISL_577607
EPI_ISL_592769
EPI_ISL_596712
EPI_ISL_854247
EPI_ISL_437942
EPI_ISL_636973
EPI_ISL_681308
EPI_ISL_483623
EPI_ISL_977658
EPI_ISL_722930
EPI_ISL_837553
EPI_ISL_955142
EPI_ISL_1016884
EPI_ISL_1181405
EPI_ISL_1240642
EPI_ISL_480298
EPI_ISL_918371
EPI_ISL_845548
EPI_ISL_1258234
EPI_ISL_1272402
EPI_ISL_1167761

EPI_ISL_746644
EPI_ISL_1168770
EPI_ISL_591278
EPI_ISL_402120
EPI_ISL_434534
EPI_ISL_1220065
EPI_ISL_956289
EPI_ISL_770029
EPI_ISL_1272023
EPI_ISL_1232316
EPI_ISL_1014556
EPI_ISL_817398
EPI_ISL_877555
EPI_ISL_481245
EPI_ISL_907112
EPI_ISL_1138530
EPI_ISL_995915
EPI_ISL_887175
EPI_ISL_788934
EPI_ISL_1197071
EPI_ISL_860184
EPI_ISL_1283938
EPI_ISL_862039
EPI_ISL_568689
EPI_ISL_582030
EPI_ISL_1254794
EPI_ISL_895035
EPI_ISL_686924
EPI_ISL_568746
EPI_ISL_1039223
EPI_ISL_968815
EPI_ISL_968827
EPI_ISL_1027645
EPI_ISL_1027652
EPI_ISL_1005230
EPI_ISL_904957
EPI_ISL_1191600
EPI_ISL_424667
EPI_ISL_1060683
EPI_ISL_979330
EPI_ISL_1040605
EPI_ISL_913984
EPI_ISL_1279462
EPI_ISL_516936
EPI_ISL_516931
EPI_ISL_768840
EPI_ISL_1159697
EPI_ISL_887472
EPI_ISL_964945
EPI_ISL_833041
EPI_ISL_461267
EPI_ISL_1210475
EPI_ISL_728656
EPI_ISL_579068
EPI_ISL_1016861
EPI_ISL_582022
EPI_ISL_794622

EPI_ISL_755622
EPI_ISL_843198
EPI_ISL_1250712
EPI_ISL_755631
EPI_ISL_456327
EPI_ISL_456301
EPI_ISL_456168
EPI_ISL_456193
EPI_ISL_548059
EPI_ISL_730010
EPI_ISL_825714
EPI_ISL_1013453
EPI_ISL_944786
EPI_ISL_954301
EPI_ISL_635116
EPI_ISL_493358
EPI_ISL_493351
EPI_ISL_666593
EPI_ISL_417444
EPI_ISL_548942
EPI_ISL_596543
EPI_ISL_873165
EPI_ISL_1225364
EPI_ISL_496772
EPI_ISL_693478
EPI_ISL_693479
EPI_ISL_693480
EPI_ISL_729891
EPI_ISL_1111176
EPI_ISL_536510
EPI_ISL_536489
EPI_ISL_1093165
EPI_ISL_833336
EPI_ISL_941340
EPI_ISL_1138775
EPI_ISL_960670
EPI_ISL_912388
EPI_ISL_815391
EPI_ISL_738343
EPI_ISL_807154
EPI_ISL_925905
EPI_ISL_751217
EPI_ISL_678072
EPI_ISL_678158
EPI_ISL_1167139
EPI_ISL_1167161
EPI_ISL_1167135
EPI_ISL_754100
EPI_ISL_875532
EPI_ISL_1063695
EPI_ISL_831245
EPI_ISL_517662
EPI_ISL_510868
EPI_ISL_934389
EPI_ISL_1199586
EPI_ISL_648160
EPI_ISL_594152

EPI_ISL_693840
EPI_ISL_796181
EPI_ISL_738065
EPI_ISL_534336
EPI_ISL_1010730
EPI_ISL_1074019
EPI_ISL_707707
EPI_ISL_717697
EPI_ISL_717700
EPI_ISL_1118884
EPI_ISL_707700
EPI_ISL_803120
EPI_ISL_1073806
EPI_ISL_1096069
EPI_ISL_954297
EPI_ISL_1112329
EPI_ISL_859932
EPI_ISL_407071
EPI_ISL_1009808
EPI_ISL_1017213
EPI_ISL_1159050
EPI_ISL_1171100
EPI_ISL_1220967
EPI_ISL_1234033
EPI_ISL_873201
EPI_ISL_876142
EPI_ISL_983856
EPI_ISL_979505
EPI_ISL_873262
EPI_ISL_527660
EPI_ISL_812342
EPI_ISL_812672
EPI_ISL_626344
EPI_ISL_626349
EPI_ISL_641501
EPI_ISL_641409
EPI_ISL_641514
EPI_ISL_683012
EPI_ISL_683020
EPI_ISL_683026
EPI_ISL_683036
EPI_ISL_683055
EPI_ISL_683069
EPI_ISL_641415
EPI_ISL_683079
EPI_ISL_683088
EPI_ISL_683098
EPI_ISL_683114
EPI_ISL_683126
EPI_ISL_683139
EPI_ISL_683147
EPI_ISL_683159
EPI_ISL_683169
EPI_ISL_683179
EPI_ISL_641428
EPI_ISL_683189
EPI_ISL_683193

EPI_ISL_641437
EPI_ISL_683208
EPI_ISL_683224
EPI_ISL_683246
EPI_ISL_683261
EPI_ISL_683269
EPI_ISL_683286
EPI_ISL_683294
EPI_ISL_683299
EPI_ISL_683310
EPI_ISL_683322
EPI_ISL_641453
EPI_ISL_641471
EPI_ISL_641479
EPI_ISL_641488
EPI_ISL_523012
EPI_ISL_523055
EPI_ISL_523058
EPI_ISL_577778
EPI_ISL_577780
EPI_ISL_577788
EPI_ISL_577807
EPI_ISL_577808
EPI_ISL_577810
EPI_ISL_577815
EPI_ISL_577819
EPI_ISL_577821
EPI_ISL_577822
EPI_ISL_577825
EPI_ISL_577828
EPI_ISL_632452
EPI_ISL_632380
EPI_ISL_632457
EPI_ISL_632458
EPI_ISL_632459
EPI_ISL_632436
EPI_ISL_632334
EPI_ISL_632489
EPI_ISL_722414
EPI_ISL_722416
EPI_ISL_722418
EPI_ISL_447628
EPI_ISL_447630
EPI_ISL_447633

Table S4. Molecular clock models used in model testing.

Clock model	*Abbreviation*	*Parameters*
Strict clock	SC	Clock rate *(n=1)*
Relaxed clock	UCLN	Total branch number, mean, sd *(n=269+1+1)*
Fixed local clock (clades)	FLC (clade)	Background rate, clade number (1+3)
Fixed local clock (stems)	FLC (stem)	Background rate, stem number (1+3)
Fixed local clock (clades & stems)	FLC (stem & clade)	Background rate, clade/stem number (1+3)
Fixed local clock (shared clades)	FLC (shared, clade)	Background rate, clade rate (1+1)
Fixed local clock (shared stems)	FLC (shared, stem)	Background rate, stem rate (1+1)
Fixed local clock (shared clades & stems)	FLC (shared, stem & clade)	Background rate, clade/stem rate (1+1)

Table S5. Specific priors used in molecular clock models.

Models with prior	*Parameters*
All models	Uniform distribution on age(root) (2019.5, 2020)
FLC(stem), FLC(clade*), FLC(stem & clade*), FLC(shared, stem*), FLC(shared, clade*), FLC(shared, stem & clade*)*	CTMC gamma distribution for clock.rate (shape =1, scale = 0.001)