bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

SARS-CoV-2 mutations among minks show reduced lethality and infectivity to humans

Tomokazu Konishi

Graduate School of Bioresource Sciences, Akita Prefectural University

Abstract

SARS-CoV-2 infection in minks has become a serious problem, as the virus may mutate and reinfect humans; some countries have decided to cull minks. Here, the virus sequencing data in minks were analysed and compared to those of human-virus. Although the mink-virus maintained the characteristics of human-virus, some variants rapidly mutated, adapting to minks. Some mink-derived variants infected humans, which accounted for 40% of the total SARS-CoV-2 cases in the Netherlands. These variants appear to be less lethal and infective compared to those in humans. Variants that have mutated further among minks were not found in humans. Such mink-virus might be suitable for vaccination for humans, such as in the case of the smallpox virus, which is less infective and toxic to humans.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the cause of coronavirus disease (COVID-19), infects not only humans but also several other animal species [1]. Infection in minks has become a particularly serious problem [2-4]; symptoms in

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

minks appear to be lethal in the USA and Denmark, [5,6] but are milder in Spain [7]. Mink livestock have been farmed for a long time as their fur has a commercial value [8]. Although mink farming has been declining recently, up to 50 million animals are farmed worldwide, mainly in Europe. Minks are culled because of the suspicion that the mink-virus can mutate and infect humans again [3,4]. Here, we report the results of principal component analysis (PCA) [9] of the mink- and human-virus sequences (for basics of the methodology, see Materials and Methods section). Many of the mink-virus were identical to that of humans; however, some variants mutated rapidly. One such variant that was closer to the human-virus was prevalent in humans of the Netherlands, amounting to approximately 40% of the total cases. This variant was probably less lethal. Other variants that have mutated further among minks are unlikely to infect humans. If mink farming continues, more variants that have a low affinity to humans will become available. It is possible that such mink-virus could be used for vaccination in humans, such as in the case of the smallpox virus, which is less infective and toxic to humans.

The two highest axes of principal component (PC) are shown in Figure 1A. As is apparent, viral variants in human and mink appeared into four groups that are temporarily numbered as 0–3. Mink-virus in the Netherlands consisted of all groups, whereas those in Denmark belonged only to group 1. The highly infective variants of the second wave [10] were not found. In the sequence data from both countries, variations in mink-virus that appeared in those axes were few, forming concentrated points stacked with each other (Fig. 1A). The limited number of variations reflects that viral transmission from humans to minks is rare and that the migration of the virus is also rare. The contribution of the viral samples from minks to these axes was small, and the viral samples from minks appeared in exactly

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the same PC as one of the humans. The axes represent the process by which SARS-CoV-2 adapts to humans [10]; group 3 includes the earliest variants, group 0 includes the first variant transferred to Europe [11], and groups 1 and 2 were derived from group 0. The routes of virus migration and mutation presented in these axes were completed in April, and all groups could be found on all continents.

However, mutations also occurred in a mink-virus in the Netherlands (Fig. 1B). The directions of mutations among mink- and human-virus are completely different from one another as shown in Fig. 1A and 1B. This difference suggests that the parental strain of SARS-CoV-2 did not originate from minks, otherwise the reversion mutation would return to the direction that occurs among humans that are presented in Fig. 1A. None of the bases or amino acids were unique to mink-virus. They can be recognised as characteristic rearrangements that collect parts of human-virus. Such directed mutations were only found in group 2 variants of the Netherlands and were not found in mink-virus in Denmark.

These mutations may be the result of adaptation during viral transmission between minks, and the mutation rates were high (Extended Data Table 1). The differences are comparable to half of those between peaks of the seasonal H1N1 influenza virus, which usually takes a few years among humans to gain enough cumulative mutations that allow escaping herd immunity [10]. In addition, the rates of the missense mutations were high [12]. These phenomena were also observed in mutations in human-virus [10]. The direction strongly suggests that the mutations are meaningful rather than random. These properties could be due to the positive selection of variants that are more infectious among new hosts, minks.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Viruses that mutate and leave the human-virus population can also reinfect humans. As shown in Fig. 1C, variants that were mainly prevalent in the Netherlands (Fig. 2A) at one time, accounted for approximately 40% of total virus present (Fig. 2B). However, these are apart from other human-virus, and no such direction of mutations has been observed among (Fig. 1C). Rather, the seventeen thousand human samples formed a concentrated mass in Fig. 1B and 1C. Furthermore, human-to-human infectivity seemed to decrease and become sporadic (Fig. 1C, variants indicated by IDs). Variants further away than those from humans' mass have not yet been identified in humans (Fig. 1B). Conversely, highly infective viral variants of the second wave in humans [10] have not yet been identified in minks. These human and mink variants could have become too adapted for the host animals.

In England, France, the USA, and Australia, the variants mutated before the second wave [10]. This phenomenon has also been confirmed in the Netherlands and Denmark, where similar variants were prevalent and waves occurred similarly (Fig. 2 and Extended Data Fig. 1). However, there is one difference; the variant that seems to be derived from minks appeared mainly in the Netherlands (Fig. 2A, B). This variant has disappeared to be replaced by the variant belonging to group 0 that caused the second wave. This is also the case in Denmark (Extended Data Fig. 1B). This shows that the mink-derived variant is less infective than the second wave variant.

After the mink-derived variant disappeared, the fatality rate in the Netherlands increased (Fig. 2B). In the first wave, many of the cases might have been ignored; hence, the rate became very high, and then decreased. The confidence intervals showed that the increase was substantial (Fig. 2C). This difference may correspond to the proportion of the mink-

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

derived human-virus present; if the variant is not lethal to humans, this difference can be explained by its disappearance. Since lower infectivity suggests a slow spread in the human body, it could allow sufficient time for the immune system to function, achieving lower lethality. Unfortunately, no data are available regarding individual human medical conditions; hence, we were unable to verify this lethality directly. However, data from 360 human samples with these mink-derived variants are shown in Figshare [13] to enable clinical verification. Additionally, to check the toxicity of other variants, PCs of other human samples from each continent are available on the same page in Figshare.

Low lethality can also be expected, as the virus was derived from minks maintained in dense populations [2,3]. Lethal viruses, such as those in the USA and Denmark [5,6], eventually die with the host. Therefore, a pathogenicity that is excessively strong becomes a selective pressure. This could be the reason why adaptation was not observed among mink-virus in Denmark, where only limited variants were found. This pressure is not as effective for humans as it is for minks, as some patients do not show symptoms. However, because most of farmed minks are of the same age and have similar genetic backgrounds, the disease status can be expected to be fairly uniform. Minks showed limited pathological traits in some countries, [7] which suggested that this selection and attenuation is progressing within the mink population, or attenuated variants of human-derived virus are becoming apparent in minks.

The mink-derived human-virus in the Netherlands differ by only eight amino acids and 14 nucleotides from the closest human-virus. Although the mink-derived variant seemed to be disappeared in Europe, if the variant remained in humans, the difference would not be safe enough for reverse mutations to re-adapt to humans.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Mutations of the group 2 mink-virus (Fig. 1B) appeared to adapt to the new host, minks. Accumulating mutations will further reduce infectivity to humans. Variations among these strains can be increased by mink farming. Such strains could be maintained in minks or in Vero cells; among them, we could identify a combination of a strain and dosage that infects, but does not cause symptoms in humans. If this is achieved, vaccination will become possible by infecting the intestinal tract via oral administration. Therefore, local governments should encourage farmers to maintain their minks, rather than culling them. However, as minks can also carry human-virus as is, sequence analysis would be required to avoid this risk. It may be advisable to artificially infect minks with a known safe strain to restart farming. Human SARS-CoV-2 variants continue to change independently in each region [10], as shown in Fig. 2A. Although some of the mink-virus from Denmark appear to be highly toxic to minks [6], the same group 1 variants would be the most adapted to humans, and they continue to mutate among humans in regions such as South Africa and Brazil. Some of these may have lower toxicity in minks.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

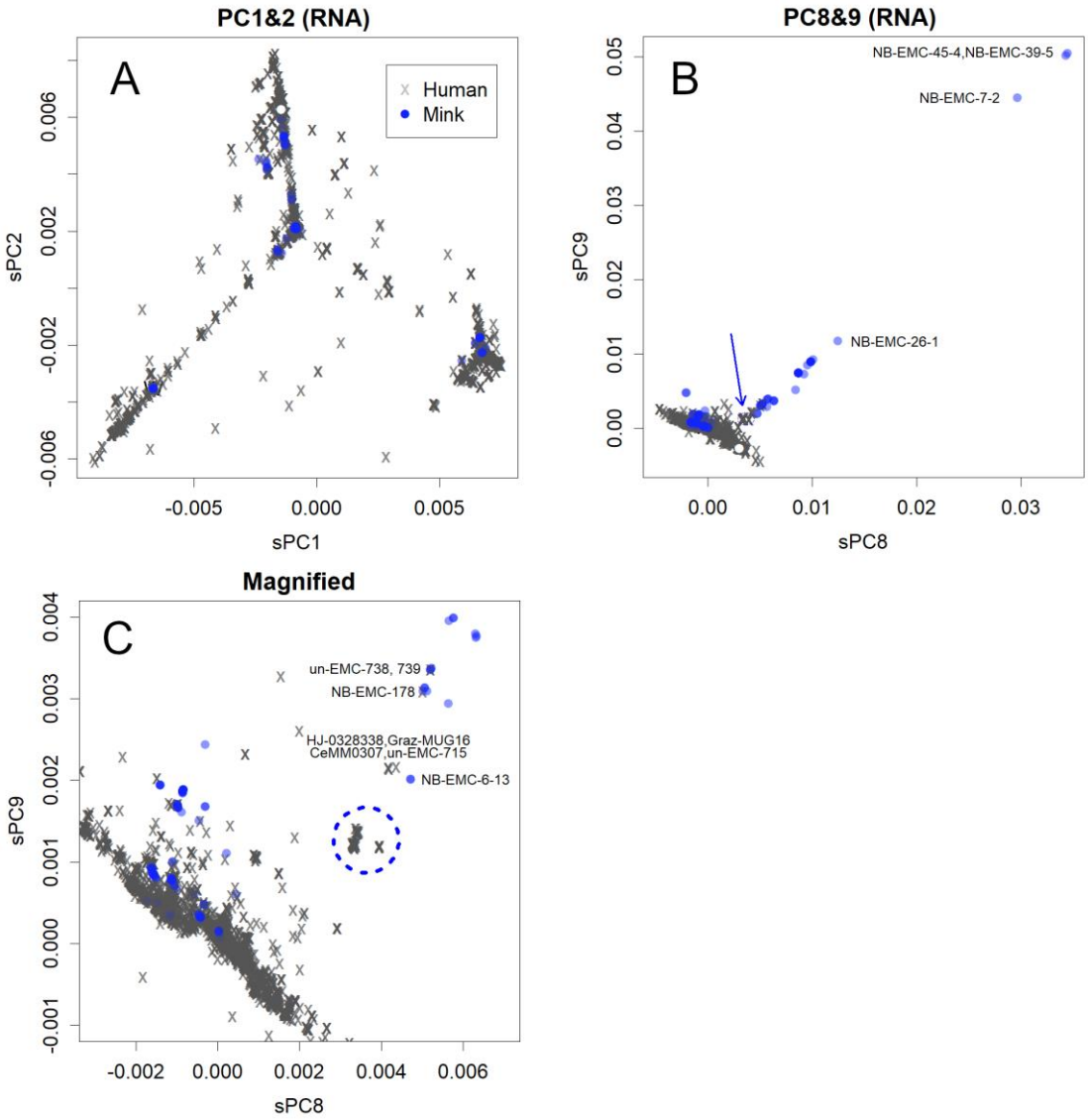


Figure 1. PCA for samples. **A**. PC1 and PC2. These axes show the adaptation process of SARS-CoV-2 to humans [10]. Seven thousands of mink-virus are coloured in transparent blue; concentrated blue show multiple stackings. This is the same in 17,000 human-virus. The white dot indicates the position of NB-EMC-35-3, a group 2 mink-virus (Extended Data Table 1). **B**. PC8 and PC9. Mutations found in group 2 mink-virus in the Netherlands are presented on these axes. The presented IDs are those for mink-virus. The downward blue arrow indicates the mink-derived human-virus. **C**. Part of panel B enlarged. All IDs

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

indicate human-virus, except for NB-EMC-6-13. The mink-derived human-virus are circled with a blue dotted line.

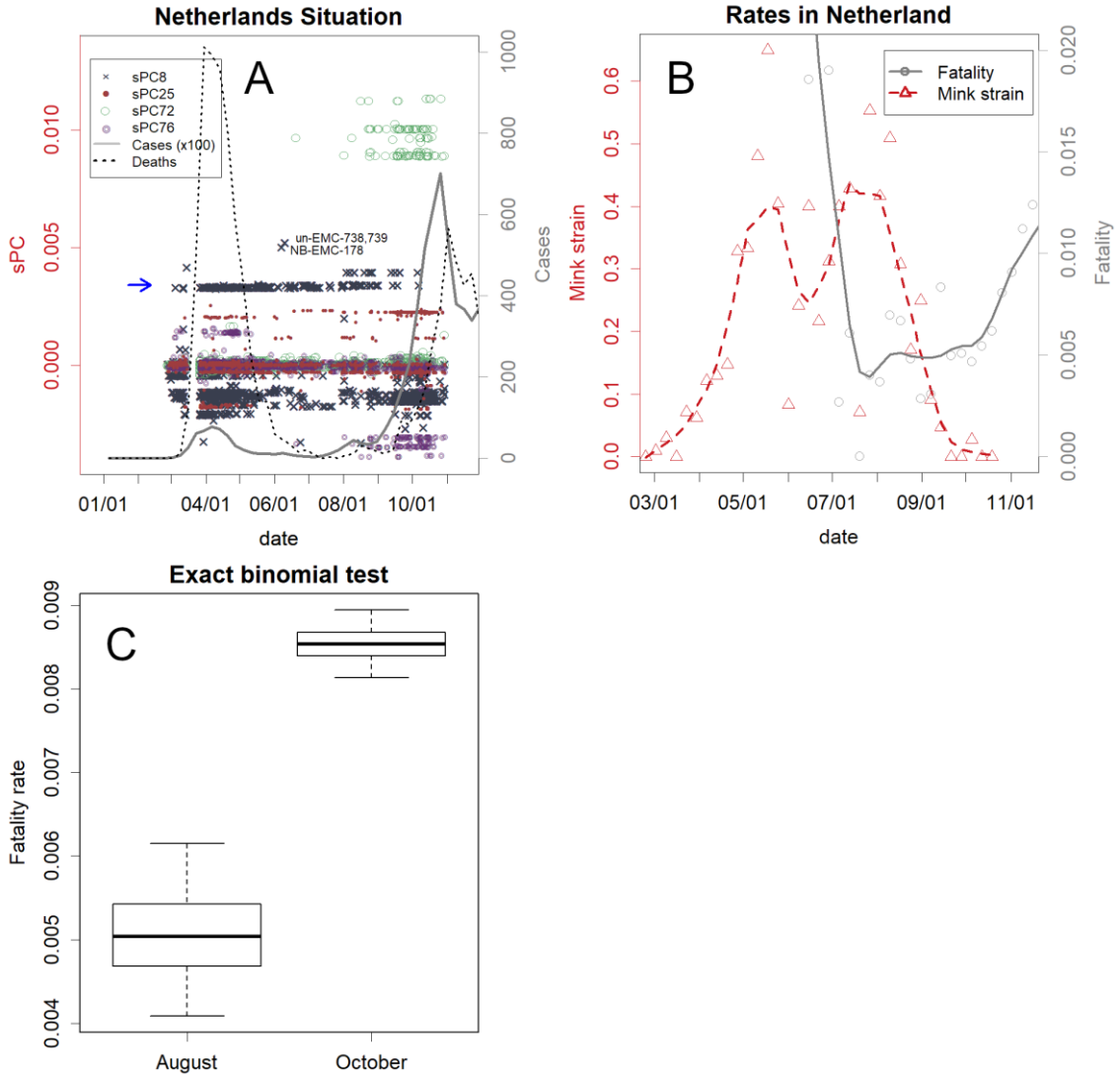


Figure 2. The SARS-CoV-2 outbreak in the Netherlands. **A**. Number of confirmed cases, deaths, and PC for human-virus. Before the second wave, variants had changed (coloured

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

points of specified axes of PCs). The mink-derived human-virus appeared at and above the level of PC8 (x) indicated by the blue arrow. **B**. Rates of fatality (number of deaths in the following week/cases, grey) and the mink-derived variants (red). **C**. Comparison of fatality rates in August and October. The thick horizontal line indicates the estimation of the rate. The whiskers show 95% confidence intervals, and the boxes show the quartiles. The situation in Denmark is presented in Extended Data Fig. 1.

Materials and methods

PCA represents a sequence matrix, which is inherently multivariate data on multiple axes [14]. Each axis covers a certain set of base positions with specific weights. These are PCs for the bases. A sample is given a value on each axis, PC for samples. PCs for bases and samples are inextricably linked to each other. The high-level axes, such as PC1, represent differences associated with more samples and bases; conversely, the lower axes represent a minor difference, for example, a feature that appears only in a particular country or region.

Sequence data of 1,832 human-virus in the Netherlands, 6,980 in Denmark, and mink-virus of 188 Netherlands and 63 Denmark were downloaded from GISAID [15] on December 2nd, 2020. Sequences from 17,571 European human-virus downloaded previously were also used. The list of samples and acknowledgements are available in Figshare [13]. The sequences were aligned using DECHIPER [16] then analysed using PCA

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[14]. The axes were identified using 103 mink-virus and 6092 human-virus that were proportionally selected from each continent. These data may be comparable to 130 million animals; since the mink population is estimated to be 50 million [8], the axes provide two to three times the weight toward minks, which might have enabled the identification of unique mutations in mink-virus. Following this, the axes were applied to all data.

All calculations were performed using R [17]. All the codes used are presented in the author's page in Figshare [13]. The number of confirmed cases and deaths were obtained from the homepage of the WHO [18]. The rate of fatality was estimated as the rate between deaths in the following week and the number of cases during the week. The 95% and 50% confidence intervals of the rate were estimated using the binomial test [19], according to the estimation that deaths occur randomly among patients; if history repeats in the area with the same conditions, the observed rates will most likely vary as in this estimate. The mink-derived human-virus were found to have sPC8 > 0.003 and sPC9 > 0.001, and the rate was estimated within the countries. The weekly estimation of rates varied, as the numbers of infected and dead, and especially the number of sequences registered, were not very high. Therefore, the line representing the status changes was smoothened using the LOWESS function of the R [20].

1 Konishi, T. Principal component analysis of coronaviruses reveals their diversity and seasonal and pandemic potential. *PLoS ONE* **15**, e0242954,

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

doi:https://doi.org/10.1371/journal.pone.0242954 (2020).

2 Mallapaty, S. COVID mink analysis shows mutations are not dangerous — yet. *Nature* **587**, 340-341 (2020).

3 BBC. *Denmark to cull up to 17 million mink amid coronavirus fears*, <https://www.bbc.com/news/world-europe-54818615> (2020).

4 WHO. *Mink-strain of COVID-19 virus in Denmark*, <https://www.euro.who.int/en/countries/denmark/news/news/2020/11/mink-strain-of-covid-19-virus-in-denmark> (2020).

5 Mossburg, C. & Ries, B. *10,000 mink are dead in Covid-19 outbreaks at US fur farms after virus believed spread by humans*, <https://edition.cnn.com/2020/10/09/us/mink-covid-outbreak-trnd/index.html> (2020).

6 ipj/sri (dpa, R., AP, AFP). *Lithuanian mink also corona-infected, tearful Danish PM apologizes*, <https://www.dw.com/en/lithuanian-mink-also-corona-infected-tearful-danish-pm-apologizes/a-55743155> (2020).

7 Hikita, S. *Fur mink infected with corona, killing 90,000 animals Spain* (Asahi News Paper, Tokyo Japan, 2020). < https://www.asahi.com/articles/ASN7K2D9XN7KUHBI003.html >

8 International_Fur_Trade_Federation. The Socio-Economic Impact of *International Fur Farming*, <https://web.archive.org/web/20110713004402/http://www.iftf.com/publctns/4849Intls_eEng.pdf> (2009).

9 Jolliffe, I. T. *Principal Component Analysis*. (Springer-Verlag 2002).

10 Konishi, T. Changes in SARS-CoV-2 before the peak in each country, producing unique variants. *bioRxiv*, 413344, doi:https://doi.org/10.1101/2020.12.18.413344 (2020).

11 Konishi, T. Continuous mutation of SARS-CoV-2 during migration via three routes. *Figshare*, doi:10.6084/m9.figshare.12365777.v1 (2020).

12 Wang, T. *et al.* Probability of phenotypically detectable protein damage by ENU-induced mutations in the Mutagenetix database. *Nature Communications* **9**, 441, doi:10.1038/s41467-017-02806-4 (2018).

13 Konishi, T. SARS -CoV-2 has mutated among minks, reducing the lethality and infectivity to humans. Dataset. *Figshare*, doi:https://doi.org/10.6084/m9.figshare.13385192.v1 (2020).

14 Konishi, T. *et al.* Principal Component Analysis applied directly to Sequence

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Matrix. *Scientific Reports* **9**, 19297, doi:10.1038/s41598-019-55253-0 (2019).

15 Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Global challenges (Hoboken, NJ)* **1**, 33-46, doi:10.1002/gch2.1018 (2017).

16 Wright, E. S. DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics* **16**, 322, doi:10.1186/s12859-015-0749-z (2015).

17 R Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, 2020).

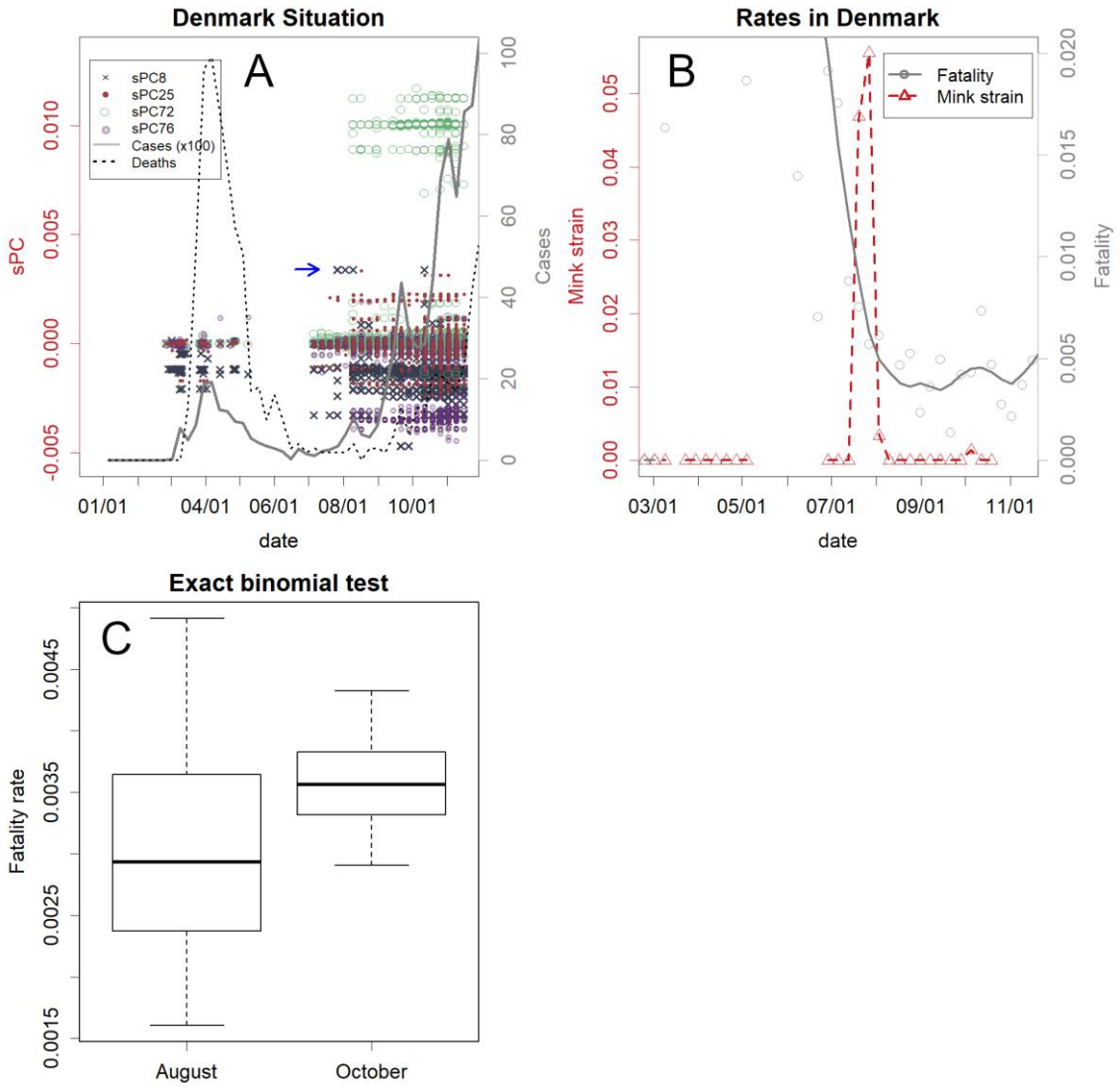
18 WHO. *WHO Coronavirus Disease (COVID-19) Dashboard*, <https://covid19.who.int/> (2020).

19 Clopper, C. J. & Pearson, E. S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**, 404–413 (1934).

20 Cleveland, W. S. Robust locally weighted regression and smoothing scatterplots. . *J. American Statistical Association* **74**, 829–836 (1979).

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Extended Data



Extended Data Figure 1 Situation in Denmark **A**. Number of confirmed cases, deaths, and PCs for samples. The blue arrow indicates the mink-derived human-virus (sPC8), which are far fewer than those in the Netherlands. **B**. Fatality rate (the number of deaths in the following week/the number of cases, grey) and percentage of mink-derived human-virus (red). The fatality rate remained fairly constant after the first wave subsided. **C**. The estimated confidence intervals confirmed the constancy of fatality.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424267; this version posted December 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	/1000	Missense
NB-EMC-6-13|EPI_ISL_523098	0.61	0.60
NB-EMC-26-1|EPI_ISL_523086	4.10	0.36
NB-EMC-7-2|EPI_ISL_523110	4.69	0.35
NB-EMC-45-4|EPI_ISL_577803	4.96	0.35
Humans*	0.47	0.57

Extended Data Table 1 Rate of differences between mink-derived human-virus in 1000 amino acid residues; the differences between each variant (positions are shown in Fig. 1B and 1C by the IDs) and NB-EMC-35-3 | EPI_ISL_577774 (a mink-virus considered to be the same as that of humans). The rates of missense mutations are also shown. *Humans: differences between the mink-derived variants and similar human-virus in group 2, the Netherlands. The mink-derived human-virus were NB-EMC-45-4, NB-EMC-39-5, NB-EMC-7-2, NB-EMC-26-1, NB-EMC-45-3, NB-EMC-39-3, and NB-EMC-41-4. The related variants were NB-EMC-312, ZH-EMC-379, ZH-EMC-844, and ZH-EMC-845. Differences between averages were used.