medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Increasing SARS-CoV-2 mutations against vaccination-acquired immunity

Tomokazu Konishi

Graduate School of Bioresource Sciences, Akita Prefectural University, Akita 010-0195, Japan

key words

COVID-19, vaccination, Omicron variant, variants of concern, animal virus

Summary

Monovalent vaccines using RNA or adenoviruses have successfully controlled the COVID-19 epidemic in many countries. However, viral mutations have hampered the efficacy of this approach. The Omicron variant, in particular, has caused a pandemic which has put pressure on the healthcare system worldwide. Therefore, administration of booster vaccinations has been initiated; however, there are concerns about their effectiveness, sustainability, and possible dangers. There is also the question of how a variant with such isolated mutations originated and whether this is likely to continue in the future. Here, we compare the mutations in the Omicron variant with others by direct PCA to consider questions pertaining to their evolution and characterisation. The Omicron variant, like the other variants, has mutated in its human vectors. The accumulated mutations exceeded the range of acquired immunity, causing a pandemic, and similar mutations are likely to occur in the future. We also compare Omicron with variants that have infected animals and discuss the possibility of a vaccine using a weaker variant of the virus.

Introduction

The COVID-19 epidemic continues, despite the efforts of many countries to bring it to closure[1]. It is postulated to have started in Wuhan, China. Then, by April 2020, it had spread to Europe and North America. During this progression, it rapidly mutated to form four major sub-groups, three of which are still prevalent today[2]. COVID-19 is also known to have spread among several susceptible animal species; a problem that, as in the case of humans, continues to manifest itself[3-6]. Subsequently, increased surveillance at national borders has slowed the spread of the disease across national lines. Further, more potent variants have emerged in each country independently[7]. The most infectious sub-types have spread across borders and have been designated as variants of concern (VOC) by the WHO[8].

Countries have taken measures to prevent the spread of the disease by surveying and isolating the infected people. Vaccines have been rapidly developed; particularly monovalent

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

vaccines. This has been made possible through the use of new technologies, such as RNA vaccines, that have become popular globally. These have proven very effective, and have led to a significant reduction in the number of people infected for a time; even in countries where detection and isolation did not work well[9-10].

However, as the virus continues to mutate, the vaccines' effectiveness is waning. The Delta variant has become emblematic of this situation. In turn, more recently, infections by the Omicron sub-type have exploded[11-14]. Even in Australia, where vaccination rates are high and effective control measures are in place, the latter variant has caused many cases[15]. Due to its virulence, despite rigorous counter-measures, the Omicron is thought to be the sub-type with a mutation in the Spike protein. This is because the said alteration almost abrogates the effect of the monovalent vaccine; thus it is likely to be infectious even after the third dose[16].

Here, we characterise the Omicron variant's genetic sequence using direct principal component analysis (PCA)[17] and discuss the mechanism informing its virulent manifestation. This is an objective method to evaluate the characteristics of a sample based on sequence differences; with each PCA axis presenting differences in the nucleotide sequences at specific positions. In this analysis, several axes are used to determine factors, such as the origin of each variant, how it has changed, and its basic characteristics. This data will be compared with that obtained from variants infecting animals to discuss the possibility of developing a vaccine using a weakened variant of the virus.

Results

When we evaluated PCA axes from the data up to April 2020 and compared it to the recent data on these axes, the variants appeared as several groups on three routes (Fig. 1A)[2]. At this stage, the virus's acclimation to human vectors seems to be complete. The changes made here are of great importance, and this is probably why the current variants retain the same characteristics. In Fig. 1A, the axis shows 27,000 random samples of variants registered up to 27 December 2021. In blue is the VOC of the WHO. All the Omicron variants belonged to group 1.

Incidentally, approximately 1500 sequences of the SARS-CoV-2 variants infecting animals have been registered as of 27 December 2021. Notably, they too belong to one of these mentioned groups (Fig. 1B). In particular, many of the sub-forms prevalent in minks, deer, dogs, cats, and zoo animals are thought to have been transmitted by humans; specifically, they likely originated from variants where many human cases have been observed.

The currently prevalent variants have many more mutations. Fig. 2A reflects the magnitude of mutations in the variants. To equalise the weights, two WHO-VOC each were selected to set the axes. The Omicron variants were observed to be distant from the others. The samples from African countries recorded changes in the variant. As the samples in the upper right corner increased, the number of reported cases increased, suggesting that they became more infectious. The upper rightmost variants have a three-amino acid insertion in the spike protein sequence. This is noteworthy because, while many of the newer variants have some deletions, insertions are rare. Thus, this is probably the variant that has mutated the most.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

However, the spread of mutations is not the process of change observed on a time-series basis. The first two reported cases in South Africa were already heavily mutated (10/12 and 10/24). The earliest Omicron variant, which was still less mutated, would have been located farther down to the left. It is likely that the disease spread elsewhere, matured, and then the most prevalent variant moved to the sequencing countries (Fig. S1).

The global data in Fig. 2A are shown over time (Fig. 2C and 2D). It can be seen that each epidemic was caused by a single variant, where the change in variants was discontinuous. The gap to the Omicron variant is emphasized by the absence of sufficient African records. This is distinctly different from the case of H1N1 influenza. If the mutations were to accumulate sequentially in one variant, PCs would show sine curves, as seen with H1N1 mutations (Fig. S2). There was one variant of H1N1 per year somewhere in the world, which moved annually while changing itself. After a few years, the variant would change by approximately 15–30/1000 bases and then return to the same location to cause another epidemic. This is likely because the flu infects many people who then gain acquired immunity.

Omicron was first reported in South Africa; however, Group 1 variants to which Omicron belongs were not prevalent in this country after August 2020 (Fig. 2B). The only Group 1 variant that appeared briefly in July 2021 was C.1.2, which is also quite far from the Omicron variant (Fig. 2A). A closer group 1 variant is B.1.1.519, which was reported by Botswana and Morocco. The relationship between this variant and Omicron and its origin remains unknown because of lack of records.

Omicron is a mutated human variant of COVID-19. However, this variant's mutations did not resemble any of the existing coronaviruses (Fig. 3A and 3 B)[18], nor did it have anything in common with SARS-CoV-2 that had infected animals (Fig. 3C). Thus, this eliminated the possibility that it was transmitted from animals[19]. In particular, the rodent data were completely unrelated to Omicron's mutations (Fig. S4). This animal vector hypothesis originally arose as a result of processing the phylogenetic results with PCA. However, phylogenetic trees are a form of one-dimensional data created based on the distances between sequences. Therefore, these sequences are not comparable. Further, given that PCA is a method for observing multidimensional data, processing one-dimensional results is not its original purpose. Artefacts caused by inappropriate data processing were apparently the source of this concern. As seen in Fig. 1A and 2A, this variant gradually changed from group 1.

When SARS-CoV-2-infected animals, such as mink, deer, dogs, and cats, a ping-pong effect occurred, thereby increasing the number of infected animals. In these animals, acclimatisation occurred quickly. This is similar to the situation in which the initial SARS-CoV-2 variants were acclimatised to humans by April 2020. For example, mutations in PC21 and PC25 (Fig. 3D) on the animal sample axis suggest acclimation to minks and deer in some countries.

The concern about re-infection from these animals to humans is natural. However, variants that are sufficiently far from the human variants, as shown in Fig 3D, are not evident. This is why 27,000 human samples are clustered in the centre. If massive re-emergence should occur in the future, it would be easily confirmed by sequencing. In fact, the only variant that has ever been prevalent in humans is the one in the Netherlands, indicated by the blue arrow. This variant is far from human viruses, but it is even farther from the mink viruses. Thus, it

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

is probably the process of acclimation to the mink. During the epidemic phase dominated by this variant, the mortality rate in the Netherlands reduced by a factor commensurate with the variant titer[20].

The mutations occurred mainly in spike glycoprotein (S) and nucleocapsid phosphoprotein (N) (Fig. 4). This is very different from influenza, in which all ORFs change simultaneously at the same rate[21]. The mutations in Delta variant are larger than those of Alpha. Further, since these are opposite mutations across the initial variant (Fig. 2A), Delta would have been spared much of the immunity gained by Alpha. Lambda has more mutations than these, with Omicron having even more of them. The mutations are mainly in S and N, which the surface proteins of the virus; therefore, there must be strong selection pressure to avoid immunity[22]. In Omicron, there was a high density of S mutations suggesting that there was selection pressure to avoid the acquired immunity imparted by monovalent vaccines. In Omicron, the mutations are also in the smaller ORFs, which are relatively well preserved. The mutation in the envelope (E) is only one amino acid, but it is very rare. In addition, there are three amino acid mutations in M.

The animal viruses did not show the same concentration of S and N mutations as human viruses, for example, Alpha. Fig. 5 shows the number of mutations for the variants farthest from the human virus population (Fig. 3D). There were more missense mutations; therefore, some amino acid mutations may have been desirable for each host's specificity. However, many small ORFs were retained, and none of them caused major mutations, such as Alpha and Omicron. This does not necessarily mean that variants that are more acclimated to humans are less likely to infect animals, but the examples shown here were relatively early in the process of infecting animals incidentally (so they would have had more time to get away from humans). Newer variants, for example, delta and lambda, can also infect animals (Fig. S5).

Discussion

Omicron did not arise in South Africa. Specifically, the parent of this variant was not prevalent in South Africa. Rather, it probably originated in areas without sequence testing, matured sufficiently to overcome the vaccine-acquired immunity and then entered the sequencing countries. By the time the danger was recognised in the South African survey, the variant had probably already spread to other parts of the world. The current global epidemic may be the result of this delay. Omicron variant, like other variants, has mutated among humans to overcome vaccine-induced immunity. It is likely that mutations that overcome immunity provided by newer vaccines will occur again in the future.

In contrast to the mutations of the influenza H1N1 virus, SARS-CoV-2 mutations were discontinuous. This is because there were three groups in the early stages that evolved independently, in different regions, and after the borders were closed, and the evolved stronger infectious variants were successively released on a global scale. Even a variant as infectious as Delta, for example, does not infect everyone; this is because people are consciously protecting themselves. However, if a new, more infectious variant arises, it can break through these artificial defenses. It is also possible for a very different variant to overcome acquired immunity. With the widespread use of monovalent vaccines, many people

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

are now immune to certain variants. Omicron has been able to evade this immunity and has spread the disease due to high variations from previous variants.

Africa is home to 1.2 billion people, but there are few areas where sequencing is routinely performed. The number of sequences per population in Africa was only 1/150 of that in Europe (Fig. S3), and 40% were from South Africa. Hence, there is a relative dearth of records compared to other regions. This is also the case in many Asian and Latin American regions. Similar gaps in the records can be seen in the H1N1 influenza viral mutations[21]; which mutate continuously every year, but still sometimes reveal large gaps. Thus, the gaps in Lambda and Omicron (Fig. 2A) are likely due to this lack of records.

The USA and the UK are the most prolific sequencers. However, considering the COVID-19 situation in these countries, it seems that their huge amount of sequencing is not doing much to prevent the spread of infection. If these countries had been more generous lending some of their capacity for sample sequencing to developing countries, they would have been able to detect the new variants more quickly. If detection had occurred at an earlier stage, quarantine could have stopped the spread. Thus, there is a need for international cooperation to conduct such surveys.

Monovalent vaccines have been used to combat the COVID-19 epidemic. These targeted the S-protein and worked well, but the Omicron variants were more capable of evading this immunity. For this reason, many countries and regions are rushing to grant booster vaccinations. However, repeated vaccinations may not be sustainable[23-25]. In fact, in many areas, even the first round of vaccination has not been completed[26]. With regard to Israel, the effectiveness of boosters is said to be questionable[27]. In fact, there is a report that repeated boosters do not work[16]. There have also been concerns about the dangers of repeated booster doses[28]. Therefore, quarantine based monovalent vaccines must be revised.

I wish to point out the possibility of using animal-adapted variants to develop a multivalent SARS-CoV-2 vaccine, such as that for the vaccinia virus for smallpox. In fact, a half-adapted mink variant was barely able to spread among humans. It probably had low virulence and was quickly replaced by a more infectious variant. A more adapted variant would probably not be able to spread from humans to humans. Once a weakly toxic variant is selected, it can be maintained and propagated in its host and cultured cells. The efficacy can be expected from the fact that SARS-CoV-2 does not mutate, particularly small ORFs. Perhaps the virus does not have sufficient flexibility. However, in the body, all proteins are presented as antigens. This is why all the ORFs were altered in the influenza virus and this virus has been prevalent for decades[21]. Such viruses may be less effective in preventing infection than RNA vaccines targeting the S protein. However, they are more resistant to S protein mutations and may hold the potential for preventing severe symptoms.

If a new RNA vaccine becomes available for Omicron, a mutation may occur that overcomes the newly-developed immunity and causes the next pandemic. If this cycle repeats itself, SARS-CoV-2 may continue to change in a discontinuous fashion. This is a calamity that is difficult to control and will take many years to overcome. If, on the contrary, a multivalent vaccine is approved for practical use, the selective pressure would not be concentrated on the spike protein (S), even if SARS-CoV-2 continues to mutate, similar to influenza viruses. In this case, the epidemic will probably be small, making it possible to relax preventive measures.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

The H1N1 influenza haemagglutinin mutated and replaced most of the protein's surface between the 1970s and 2009[21]. If similar degrees of freedom exist in the S and N of SARS-CoV-2, then these should still have a high mutation potential. Omicron did not simply have many variations. Rather, they mutated just like the other variants which is identical to what we suspect will continue in the future. It is very important to stop this epidemic in each country so that we do not have another VOC. Hence, this effort must be coordinated on a global scale. The production and transportation of weaker variants are much more lower-tech than RNA vaccines, and is probably more sustainable.

Materials and methods

2.1 PCA

Sets of nucleotide sequences were downloaded from GISAID[29] on 27 December 2021. However, the set did not include samples from African countries other than South Africa. Thus, to increase the number of African samples, those with complete sequences from 1 July 2021 to 15 January 2022 were also downloaded. Only the complete sequences that contained less than 1,000 N were selected. The sequences were aligned using the DECIPHER[30]. Subsequently, they were converted to a Boolean vector and subjected to PCA[17]. Sample and sequence PCs were scaled based on the length of the sequence and the number of samples, respectively[31].

The PCA axis shows differences in a specific set of bases. The axis is determined using our designated search dataset. Therefore, depending on the set of samples used, the observed differences will vary. Depending on one's aim, there are several viable sets of axes available. One is the initial axes on human acclimatization, which was created using data up to April 2020, and spread radially across four groups, and was used to determine variant origin. The other axis was derived using two WHO-VOCs[8], Alpha to Omicron, to avoid weighting errors due to differences in the number of data. In this axis, the most highly mutated Omicron variant formed PC1. The remaining variants were divided in PC2. This was used to determine variation in the micron variants. In addition, to characterise the samples infecting animals, we used 1500 samples and two WHO-VOCs.

All calculations were performed using R[32]. The ID, acknowledgements, list of samples used for the WHO-VOC, PCA axes, and scaled PCs of samples and bases can be downloaded from Figshare[33]. The newest version of the R code is publicly available at GitHub[34].

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

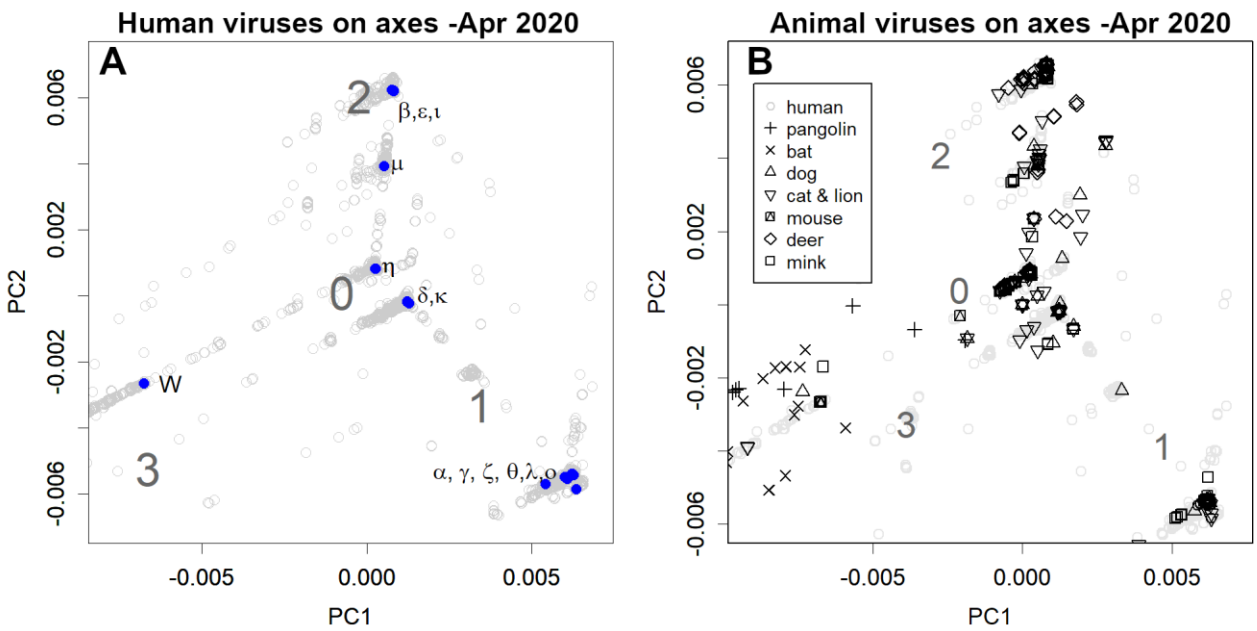

Fig. 1 Principal component analysis (PCA) with axes found in the data up to April 2020. The axes reflect the differences in the data at this point. **A.** These data, in addition to 27 000 randomly selected human samples from the data up to December 2021 have been shown. Blue is WHO-variant of concern (VOC); W is the first variant. **B.** Animal samples are shown on the same axis as in A.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

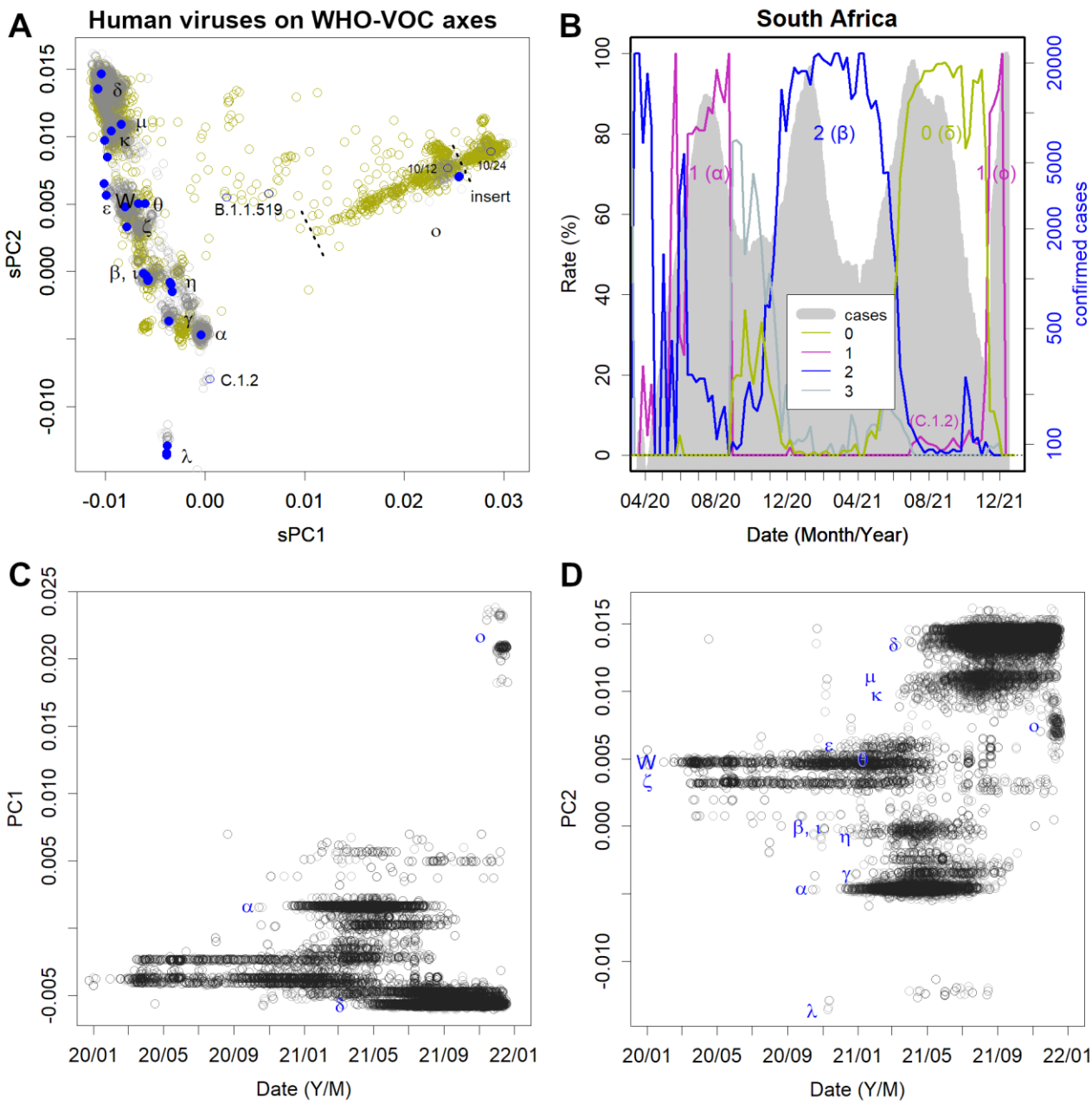

Fig. 2 Spread of Omicron variants. **A**. Human data is presented the axes determined by WHO-variant of concern (VOC). Each axis reflects the differences among these variants. Grey reflects 27 000 randomly selected data from the global data. Khaki reflects African data of 17 000 samples from July 2021 to January 2022. Omicron is depicted on the right-hand side of the figure; 10/12 and 10/24 were the first reported cases in South Africa (blue). **B.** Epidemics in South Africa (grey) and groups accounting for the percentage of cases at each time point (coloured lines) are specified. **C, D.** The time course of the global data is presented in A. Each variant has a range of mutations depending on the number of patients but does not change continuously. Rather, another, more potent variant creates the next epidemic.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

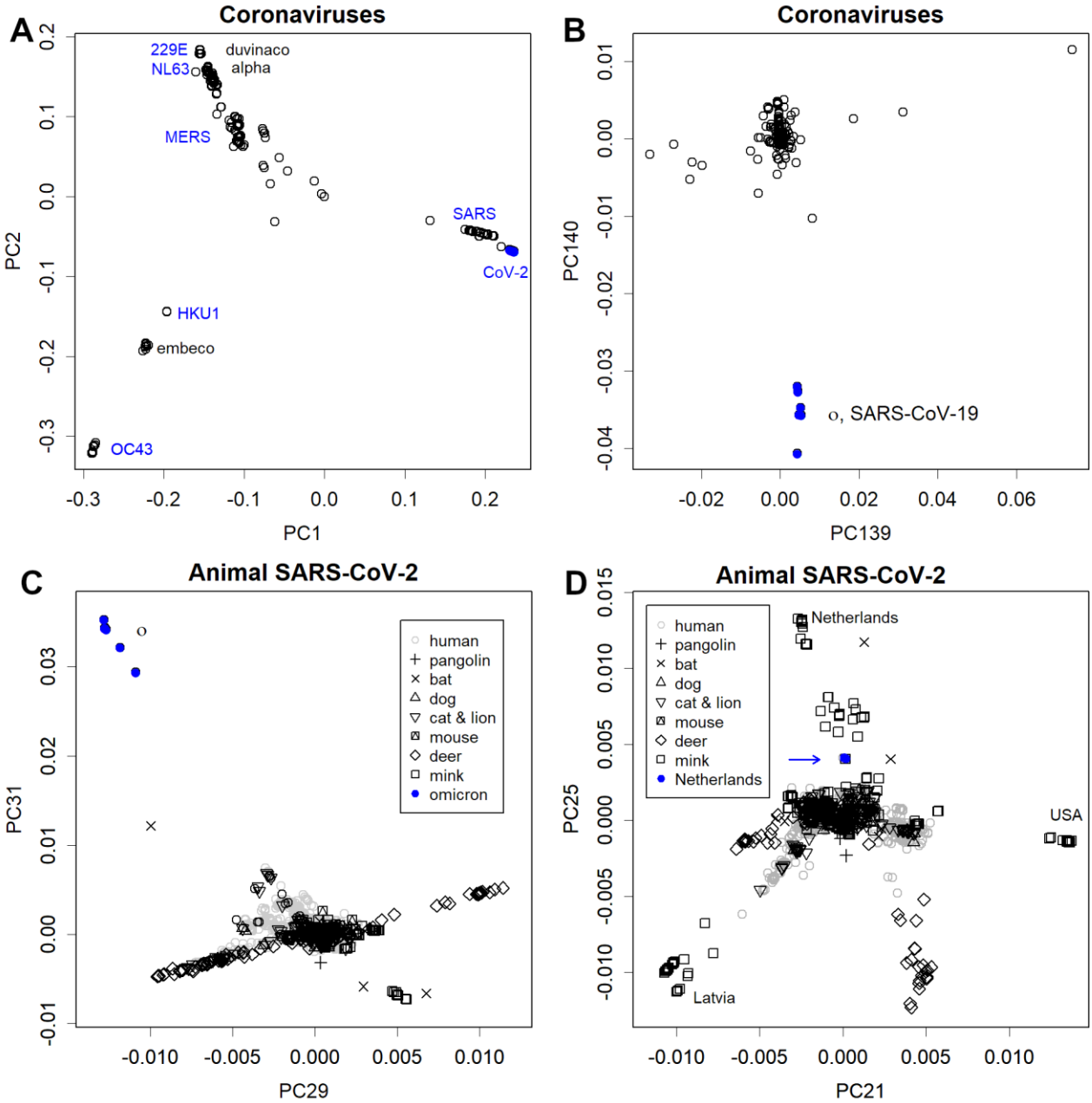

Fig. 3 Results for the axes found in the various samples. **A**. Axes found in coronaviruses. All SARS-CoV-2 are on the far right. **B**. The same series of axes of principal component (PC) 139 and 140, showing the characteristics of Omicron variant. The fact that all coronaviruses do not appear in this neighbourhood indicates that Omicron has features completely independent of other variants. **C**. Axes found in animal SARS-CoV-2 and WHO-variant of concern (VOC). There is no data in the vicinity of the omicron, indicating that the causal mutations were unique. **D**. PC 21 and PC 25 of the same set of axes as C, showing the characteristics of mink and deer variants in several countries. Blue arrow shows the mink variant prevalent in humans.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

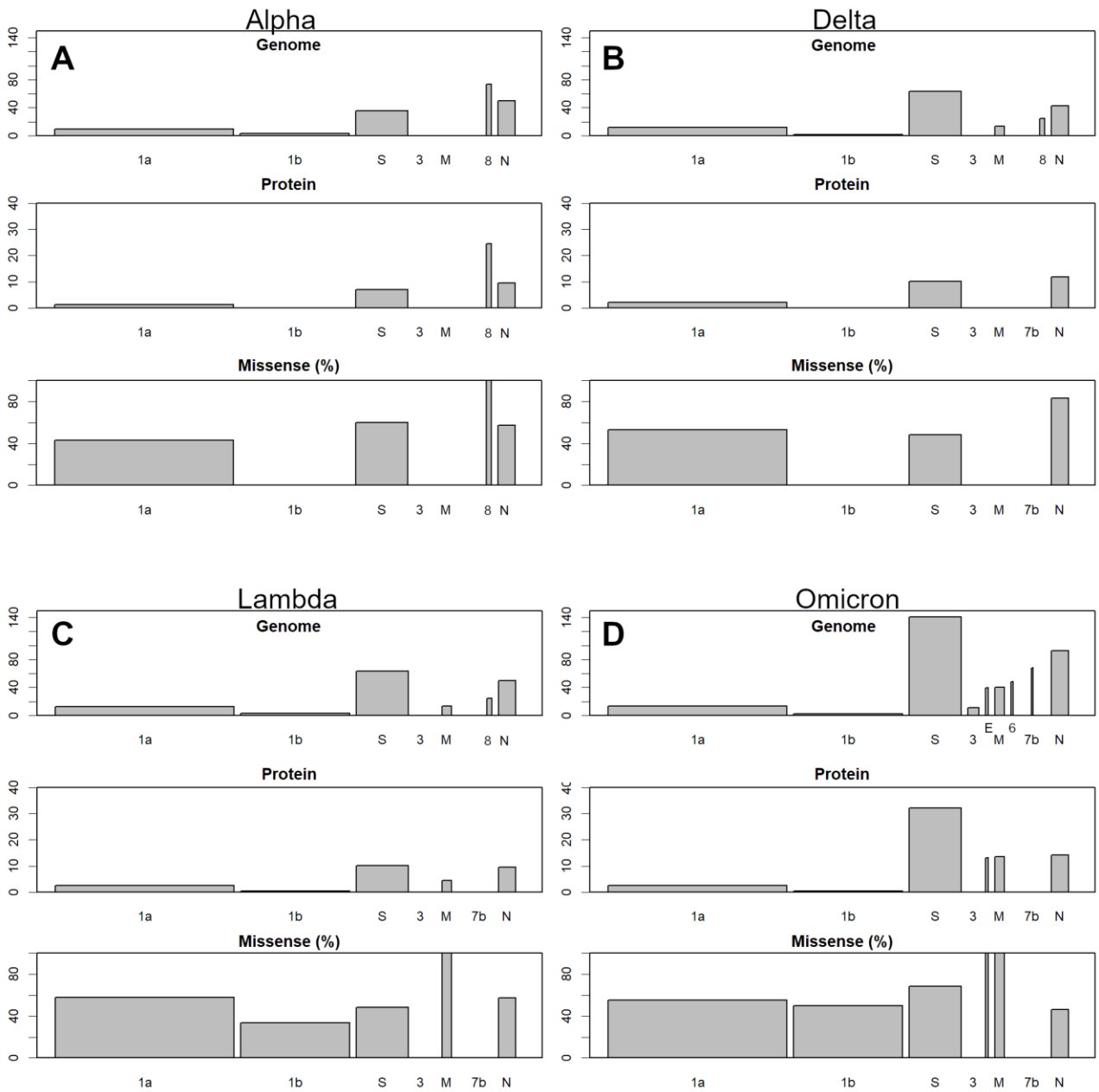

Fig. 4 Changes in each variant compared to average data as of April 2020. Each panel consists of three sub-panels; from top to bottom: genomic variation, protein variation, and the proportion of missense mutations. The top two panels show the number of mutations per 1000 bases or residues, with S and N being the most prominent: **A**, Alpha; **B**, Delta; **C**, Lambda; and **D**, Omicron. Omicron variant has a particularly high number of S protein mutations.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

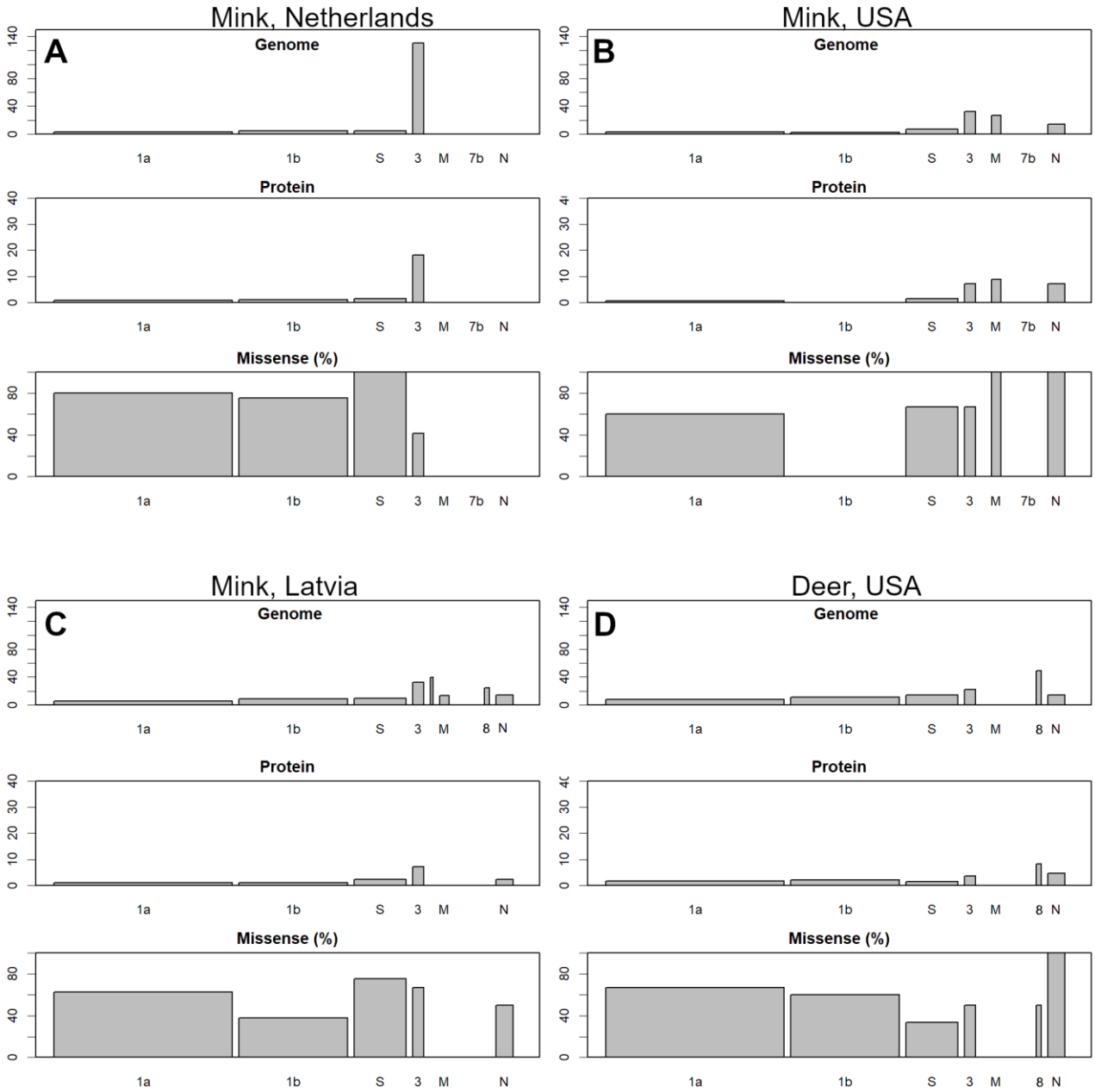

Fig. 5 Animal variants compared to the average human data up to April 2020. Those are the most distant variants found in Fig. 3D. The mutations are smaller than those of WHO-variant of concern (VOC) in Fig. 4 and are not concentrated in the S or N.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Supplement

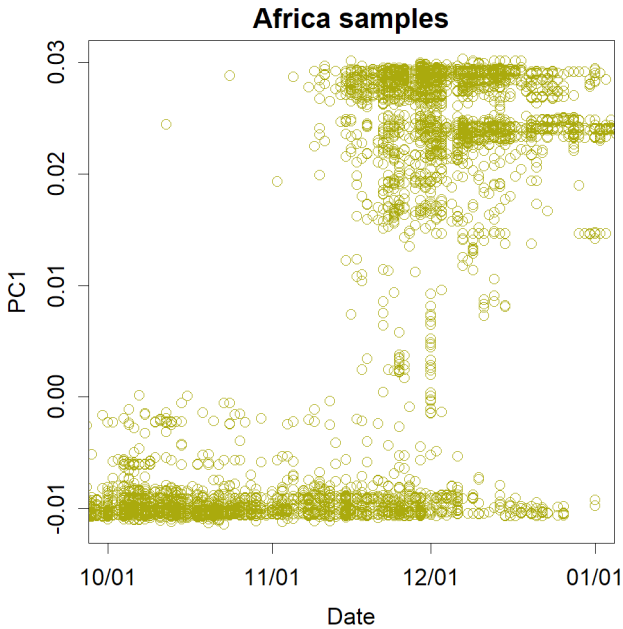

Fig. S1 Principal component 1 (PC1) of the African samples in Fig. 2A observed in chronological order. The upper third of the samples are cases with Omicron variant and the middle part is B.1.1.519. The earliest reports are from the uppermost part, which is the most mutated and infectious. The changes involved in the formation of the Omicron variant do not appear in this time series. Rather, it is more likely that a mature variant has entered the countries where these sequence testing methods are being carried out.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

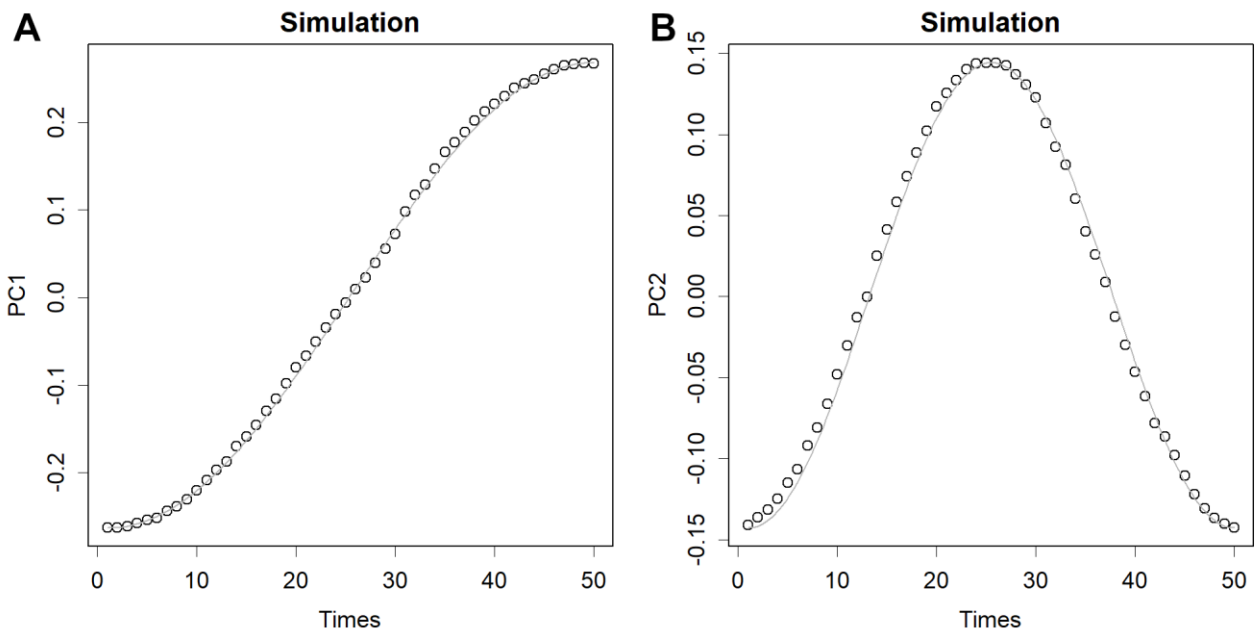

Fig. S2 Simulation of a random mutation of a single variant: A random walk (dots). The gray line is a sine curve, PC1 is half a cycle, and PC2 is one cycle. Influenza H1N1 showed a similar pattern. The length of the base was 1e4, and the number of trials was 1000. The average of all results was used for centering before PCA. The results are scaled, but they are still much larger than those for COVID-19, which has not yet produced as many mutations.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

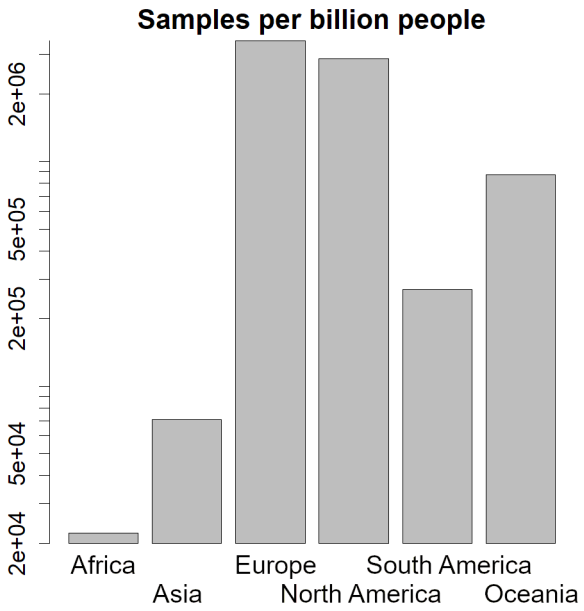

Fig. S3 Proportion of complete sequences registered in the GIS-AID per billion inhabitants in each region. The axes are logarithmic. Africa is two orders of magnitude lower than the other regions.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

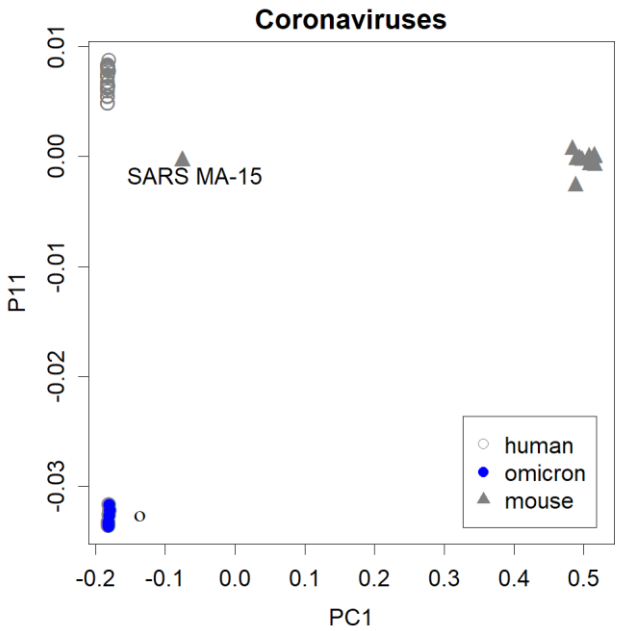

Fig. S4 Sequences of all Rodent coronaviruses from NCBI were compared with WHO-variant of concern (VOC). Five samples of Omicron variant were used and the axes were set from all said samples. The rodent variants were very far from the human SARS-CoV-2; a difference that appears in principal component 1 (PC1). Only one SARS variant infecting rodents appeared somewhat closer to humans. The characteristics of the Omicron variant appear in PC 11, whereas the rodent virus does not show these characteristics at all. These viruses are unrelated.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

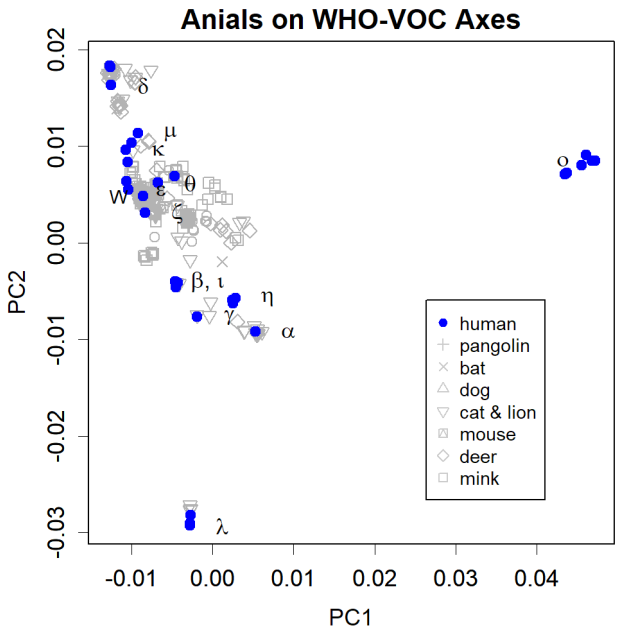

Fig. S5 WHO-variant of concern (VOC) axis with animal viruses. Blue is the WHO-VOC. Each variant infected animals to a degree; the most recent Omicron is the only exception.

[1] Callaway, E., Beyond Omicron: what's next for COVID's viral evolution. *Nature* 2021.
[2] Konishi, T., Continuous mutation of SARS-CoV-2 during migration via three routes. *PeerJ* **2021,** *in printing*.
[3] Sharun, K.; Dhama, K.; Pawde, A. M.; Gortázar, C.; Tiwari, R.; Bonilla-Aldana, D. K.; Rodriguez-Morales, A. J.; de la Fuente, J.; Michalak, I.; Attia, Y. A., SARS-CoV-2 in animals: potential for unknown reservoir hosts and public health implications. *Veterinary Quarterly* **2021,** *41* (1), 181-201.
[4] Shou, S.; Liu, M.; Yang, Y.; Kang, N.; Song, Y.; Tan, D.; Liu, N.; Wang, F.; Liu, J.; Xie, Y., Animal Models for COVID-19: Hamsters, Mouse, Ferret, Mink, Tree Shrew, and Non-human Primates. *Frontiers in Microbiology* **2021,** *12*.
[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.
[6] 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.
[7] Konishi, T., Progressing adaptation of SARS-CoV-2 to humans. *CBI Journal* **2022,** *22*, 1-12.
[8] WHO Tracking SARS-CoV-2 variants. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/.
[9] Europian Medicines Agency COVID-19 vaccines: authorised.

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

https://www.ema.europa.eu/en/human-regulatory/overview/public-health-threats/coronavirus-disease-covid-19/treatments-vaccines/vaccines-covid-19/covid-19-vaccines-authorised.
[10] Higdon, M. M.; Wahl, B.; Jones, C. B.; Rosen, J. G.; Truelove, S. A.; Baidya, A.; Nande, A. A.; ShamaeiZadeh, P. A.; Walter, K. K.; Feikin, D. R.; Patel, M. K.; Knoll, M. D.; Hill, A. L., A systematic review of COVID-19 vaccine efficacy and effectiveness against SARS-CoV-2 infection and disease. *medRxiv* **2021**, 2021.09.17.21263549.
[11] WHO Classification of Omicron (B.1.1.529): SARS-CoV-2 Variant of Concern. https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern.
[12] Buchan, S. A.; Chung, H.; Brown, K. A.; Austin, P. C.; Fell, D. B.; Gubbay, J. B.; Nasreen, S.; Schwartz, K. L.; Sundaram, M. E.; Tadrous, M.; Wilson, K.; Wilson, S. E.; Kwong, J. C.; Investigators, o. b. o. t. C. I. R. N. P. C. N., Effectiveness of COVID-19 vaccines against Omicron or Delta infection. *medRxiv* **2022**, 2021.12.30.21268565.
[13] Willett, B. J.; Grove, J.; MacLean, O. A.; Wilkie, C.; Logan, N.; Lorenzo, G. D.; Furnon, W.; Scott, S.; Manali, M.; Szemiel, A.; Ashraf, S.; Vink, E.; Harvey, W.; Davis, C.; Orton, R.; Hughes, J.; Holland, P.; Silva, V.; Pascall, D.; Puxty, K.; da Silva Filipe, A.; Yebra, G.; Shaaban, S.; Holden, M. T. G.; Pinto, R. M.; Gunson, R.; Templeton, K.; Murcia, P.; Patel, A. H.; Haughney, J.; Robertson, D. L.; Palmarini, M.; Ray, S.; Thomson, E. C., The hyper-transmissible SARS-CoV-2 Omicron variant exhibits significant antigenic change, vaccine escape and a switch in cell entry mechanism. *medRxiv* **2022**, 2022.01.03.21268111.
[14] Gruell, H.; Vanshylla, K.; Tober-Lau, P.; Hillus, D.; Schommers, P.; Lehmann, C.; Kurth, F.; Sander, L. E.; Klein, F., mRNA booster immunization elicits potent neutralizing serum activity against the SARS-CoV-2 Omicron variant. *Nature Medicine* **2022**.
[15] Austrarian Governmant Dept Health COVID-19 Omicron variant. https://www.health.gov.au/health-alerts/covid-19/symptoms-and-variants/omicron.
[16] Kuhlmann, C.; Mayer, C. K.; Claassen, M.; Maponga, T.; Burgers, W. A.; Keeton, R.; Riou, C.; Sutherland, A. D.; Suliman, T.; Shaw, M. L.; Preiser, W., Breakthrough infections with SARS-CoV-2 omicron despite mRNA vaccine booster dose. *The Lancet* **2022**.
[17] Konishi, T.; Matsukuma, S.; Fuji, H.; Nakamura, D.; Satou, N.; Okano, K., Principal Component Analysis applied directly to Sequence Matrix. *Scientific Reports* **2019,** *9* (1), 19297.
[18] Konishi, T., Principal component analysis of coronaviruses reveals their diversity and seasonal and pandemic potential. *PLoS ONE* **2020,** *15* (12), e0242954.
[19] Wei, C.; Shan, K.-J.; Wang, W.; Zhang, S.; Huan, Q.; Qian, W., Evidence for a mouse origin of the SARS-CoV-2 Omicron variant. *J Genet Genomics* **2021**, S1673-8527(21)00373-8.
[20] Konishi, T., SARS-CoV-2 mutations among minks show reduced lethality and infectivity to humans. *PLOS ONE* **2021,** *May 26, 2021*, https://doi.org/10.1371/journal.pone.0247626.
[21] Konishi, T., Re-evaluation of the evolution of influenza H1 viruses using direct PCA. *Scientific Reports* **2019,** *9* (1), 19287.
[22] Harvey, W. T.; Carabelli, A. M.; Jackson, B.; Gupta, R. K.; Thomson, E. C.; Harrison, E. M.; Ludden, C.; Reeve, R.; Rambaut, A.; Peacock, S. J.; Robertson, D. L.; Consortium, C.-G. U., SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology* **2021**.
[23] WION. Not sustainable to vaccinate the planet every 6 months, says Oxford vaccine co-creator. https://www.wionews.com/world/not-sustainable-to-vaccinate-the-planet-every-6-months-says-oxford-vaccine-co-creator-442308.
[24] WHO, Interim Statement on COVID-19 vaccines in the context of the circulation of the Omicron SARS-CoV-2 Variant from the WHO Technical Advisory Group on COVID-19 Vaccine Composition (TAG-CO-VAC). **2022**.
[25] Dolgin, E., Omicron is supercharging the COVID vaccine booster debate. *Nature* **2021,**

medRxiv preprint doi: https://doi.org/10.1101/2022.01.30.22270133; this version posted February 14, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

02 December 2021.
[26] Our World in Data Coronavirus (COVID-19) Vaccinations. https://ourworldindata.org/covid-vaccinations.
[27] Staff, T. Israeli trial, world's first, finds 4th dose 'not good enough' against Omicron. https://www.timesofisrael.com/israeli-trial-worlds-first-finds-4th-dose-not-good-enough-against-omicron/.
[28] Anghel, I. Frequent Boosters Spur Warning on Immune Response. https://www.bloomberg.com/news/articles/2022-01-11/repeat-booster-shots-risk-overloading-immune-system-ema-says.
[29] Elbe, S.; Buckland-Merrett, G., Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob Chall* **2017,** *1* (1), 33-46.
[30] Wright, E. S., DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics* **2015,** *16*, 322.
[31] Konishi, T., Principal component analysis for designed experiments. *BMC Bioinformatics* **2015,** *16 Suppl 18*, S7.
[32] R Core Team, *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria, 2020.
[33] Konishi, T., Mutations in SARS-CoV-2 are on the increase against the acquired immunity. *Figshare* **2022**.
[34] Konishi, T. direct PCA for sequences. https://github.com/TomokazuKonishi/direct-PCA-for-sequences.