Spatiotemporal patterns, farm contacts and virus traits as determinants of SARS-CoV-2 transmission between mink farms, The Netherlands 2020

Gonzales JL[1], Hobbelen P[1], Kampfraath AA[1], Welkers F[2], Fisher E[2], Hakze-van der Honing R[1], Hagenaars T[1], van der Poel WHM[1].

[1] Wageningen Bioveterinary Research, Lelystad, the Netherlands
[2] Department Population Health Sciences, Faculty of Veterinary Medicine, Utrecht University, Utrecht, the Netherlands

ABSTRACT

Epidemic spread of COVID-19 between Dutch mink farms between April and November 2020 led to 69 farms becoming infected and being culled, and to the decision by the Dutch government to terminate mink farming by the end of 2020. To elucidate which factors determined the transmission dynamics, we here analyse the spatiotemporal outbreak pattern, virus genetic sequencing data, observed variation in clinical outcome and information on the between-farm contact structure. In particular, we use the spatiotemporal outbreak pattern to estimate the between-farm transmission kernel for different genetic clusters of outbreaks. This kernel is defined as the transmission hazard between an infected and a susceptible mink farm as a function of the distance between the two farms, and its distance dependence can be used as a signature to assess different potential transmission routes against. Subsequently, the analyses of data on clinical outcome and on the between-farm contact structure are used to elucidate the role of contact patterns versus virus strain properties as determinants of transmission that could explain between-cluster differences in the transmission kernel.

Phylogenetic analysis of the sequences found in samples from 66 outbreak farms suggests six distinct genetic clusters, with five clusters corresponding to separate introductions of COVID-19 into mink. Between four of these clusters, that together comprise the majority of outbreaks, we identify significant differences in the spatial transmission characteristics. One cluster in particular, referred to as A2, showed a higher transmission potential, with expected higher spatial range and higher number of farms expected to become infected in relation to the other virus clusters. Analysis of the between-farm contact structure shows a significant correlation between cluster and the frequency distribution of different types of contact. Different AA mutations in the spike protein were identified among these clusters with some specific mutations being dominant in Cluster A2. Infection characteristics such as shedding and clinical presentation were also assessed and compared among clusters. Combining genetic and epidemiological models led to assessing transmission in higher detail and identify potential differences in transmission between variants on the same pathogen.

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has likely emerged from spill over from wild animals to humans. Human-to-human transmission spread the virus worldwide and led to spill back transmission events from humans to domestic (Ref) and wild animals (Ref). Farmed minks (*Neovison vison*) are among the domestic animal species susceptible to infection and able to sustainably transmit infection between minks. As a consequence large epidemics in farmed minks have occurred in different countries (Boklund et al., 2021; Chaintoutis et al., 2021; Lu et al., 2021)

and bidirectional transmission from minks-to-humans and humans-to-minks could be observed (Hammer et al., 2021; Lu et al., 2021; Oude Munnink et al., 2021).

Outbreaks in farmed micks were first reported in April 2020 in the Netherlands (Oreshkova et al., 2020). Between April and November 2020 a total of 68 farms became infected and were culled as part of the implemented control measures (Lu et al., 2021). During this epidemic, genomic analysis of samples of infected micks from affected farms and people working in these farms allowed the confirmation of spillback events from minks to people (Lu et al., 2021; Oude Munnink et al., 2021). During this epidemic, five distinct mink farm sequence clusters were identified, which were labelled Clusters A-E. In addition samples collected from minks from the affected farms allowed the combination of genomic and epidemiological (field) data to identify factors that may have contributed to farm-to-farm spread of the virus (Lu et al., 2021). The analysis indicated differences in phylodynamic and transmission patterns among the different clusters, with amino acid changes in the spike protein, observed particularly in cluster A, indicating a possible association with increased transmissibility. In addition generalised linear model analysis, for data from cluster A, showed an association between distance and transmission, with transmission being most likely between farms at close proximity, therefore, as also seen for other diseases, transmission is dependent on the spatial distribution of farms and the distance-dependent probability of transmission. These dependencies can be characterized by quantifying the risk of transmission between farms as a function of time and distance by applying for example transmission kernels (Ref). Such analysis combined with the genetic and epidemiological data available can provide additional insides, to those already identified, into the factors contributing to the between farm spread of the different genetic clusters and contribute to confirm difference in transmissibility among these clusters. Although mink farming is no longer done in the Netherlands, thorough analysis of this epidemic to better understand the mechanism involved in between farm transmission is important to improve control measures in the fur production sector in other countries, protect the health of farmed minks and reduce the risk of the sector becoming a reservoir for future spill over events.

To elucidate which factors influenced the dynamics of farm-to-farm transmission, we analysed the spatiotemporal outbreak patterns, virus genetic sequencing data, clinical observations in the affected farms and information on the between-farm contact structure.

RESULTS

Identification of an additional genomic cluster

Previous phylogenetic analysis of the sequences found in samples from 66 outbreak farms identified five distinct genetic clusters each corresponding to separate introductions of COVID-19 into mink. These genetic clusters were categorized as A (n = 40 farms), B (n = 1), C (n = 14), D (n = 8) and E (n =1). We sequenced additional samples (Table S1) and confirmed these clusters. Additionally our phylogeny analyses suggested that cluster A, consisted of two separate subclusters which happen to be temporally separated (Figures 1 and S1). Hence, we subdivided these cluster in A1 (n = 11) and A2 (n =29).

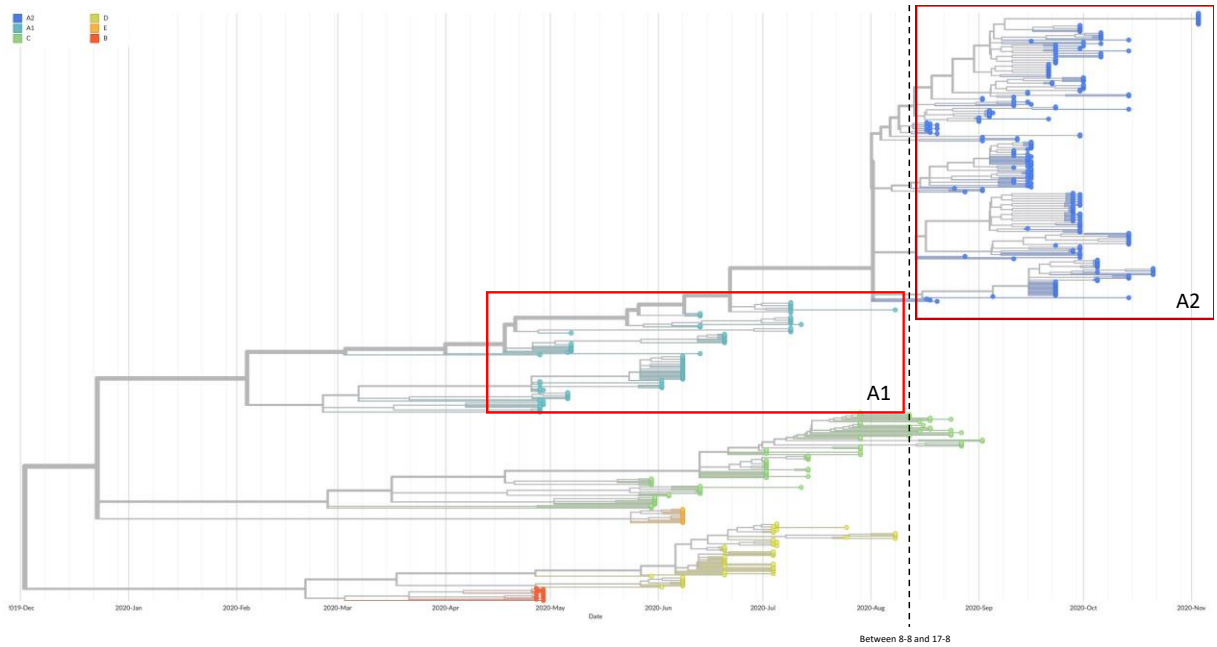

Figure 1. Time calibrated maximum likelihood phylogeny based on 485 SARS-CoV-2 sequences from infected mink in the Netherlands. Time calibration was done with TimeTree within the NextStrain pipeline. Tips are coloured based on the five clusters A-E previously identified (Lu et al., 2021). Red frames mark the separation made to the former cluster A into clusters A1 and A2 and the split in time between A1 and A2 is represented by the dashed line.

Spatiotemporal transmission of SARS-CoV2

Since virus sequences clustered as B or E were only found in one infected farm each, no transmission was assessed for these viruses. Clusters A1, A2, C and D, involved multiple farms each and a three parameter transmission kernel was used to quantify the spatial and distant-dependent probability of transmission for each of these clusters (Table 1).

Table 1 Maximun likelihood estimates of transmission kernel parameters estimated for clusters A1, A2, C and D. Values are the mean estimate and the 95% lower confidence limit (LCL) and upper confidence limit (UCL)

Cluster	Parameter	Mean	LCL	UCL
A1 - split 13 July	**Amplitude**	0.0004	0.0002	0.0059
	Power	13.6306	1.3630	46.3443
	Distance	18.7402	1.8740	33.7323
A2 - split 13 July	**Amplitude**	0.0239	0.0119	0.0525
	Power	4.5244	2.7147	9.0488
	Distance	7.0783	3.8931	10.2636
C	**Amplitude**	0.0334	0.0083	0.0954
	Power	1.8307	1.2815	3.0206
	Distance	0.8760	0.1752	2.7157
D	**Amplitude**	0.1561	0.0234	0.4838
	Power	1.9952	1.2969	2.9928
	Distance	0.3225	0.0484	1.2417

For these four virus clusters, significant differences in their spatial transmission characteristics were observed, with cluster A2 in particular showing higher probability of transmission at longer distances than the other clusters (Figure 2). To facilitate interpretation of the geographical farm-to-farm transmission risk expected for each of the clusters, we used each cluster's transmission kernels to estimate the expected R_h for each farm (Figure 2b) as a proxy measure of the farm "infectiousness" towards other farms given infection with one of these virus clusters (Figure 2c). The median (Q1 – Q3) R_h for cluster A2 (3.5 (0.6 – 8.8) was significantly higher ($p < 0.001$) than the median R_h of the other clusters. In the event of an epidemic with cluster A2 virus, more farms (n = 85) (Figure 2c) which are distributed in a wider geographical area, would successfully transmit infection ($R_h > 1$) to other farms (Figure 2b, 2c). In contrast, our analysis indicates that the risk of farm-to-farm transmission involving infection with Cluster A1, which was the first emerging cluster in the epidemic, was limited. Cluster C, had a significant higher R_h (median (Q1-Q3) 0.64 (1.2 – 1.8)) than cluster A1 (0.2 (0.1 – 0.7)) ($p < 0.001$) and possibly D (0.3 (0.1 – 1.3) ($p = 0.058$). No significant differences in the median R_h values between clusters A1 and C were observed ($p = 0.102$)

Number of farms grouped by their expected R_0

Cluster	R_0<=0.5	0.5<R_0<=1	R_0>1
A1	77	49	0
A2	29	12	85
C	56	19	51
D	73	14	39

Figure 2. Spatiotemporal transmission patterns of the A1, A2, C and D SARS-CoV-2 virus genetic clusters causing the epidemic in mink farms in the Netherlands. The figure shows the estimated transmission kernels (Top left), reflecting the rate of farm-to-farm transmission a function of between farm distance. The kernels' predicted risk of transmission maps (right) and the expected number of farms (bottom left) expected to have R0 > than 1 (high risk of between farm transmission). Note the higher spatial range and the higher number of farms expected to be able to transmit for cluster A2 in relation to the other virus clusters.

Factors associated with farm-to-farm spread.

The transmission kernel analysis indicated differences in transmissibility (far-to-farm) among the identified genetic clusters. To try to elucidate factors associated with this transmission differences we explored the association between the genomic clusters and the dominant amino acid changes identified in the affected farms, levels of virus shedding observed in sampled infected animals and the farm's contact structure.

Relationship between amino acid changes and virus clusters

Within the phylogeny, 22 subclusters of clonal spread between farms were found (Figures 1, 3), the three biggest subclusters all included samples with at least 5 farms. One subcluster was within cluster C and two within cluster A2, in which the relative number of smaller clonal clusters was also higher. The subcluster within C contains sequences from five farms (NB26, NB29, NB32, NB34 and NB39) and those within A2, one contained sequences from sixteen farms (NB33, NB35, NB36, NB43, NB46, NB47, NB48, NB50, NB52, NB53, NB54, NB56, NB59, NB61, NB63 and NB64) and the other subcluster contained sequences from five farms (NB52, NB57, NB58, NB59 and NB60). Only the samples of second subcluster of A2 share a link code (LC4). These three big clonal subclusters also share the F486L mutation in the spike protein, this mutation arose 6 times within minks and expanded twice, once in cluster C and once in cluster A2. Within C this expansion cooccurred with another mutation on the spike protein L452M that arose within NB06 at the root of cluster C. Nine of the 14 farms within C carried both mutations, which are all the farms infected within the cluster after 13-06-2020, except for NB24 (NB17, NB23, NB26, NB29, NB32, NB34, NB39, NB41 and NB45). Within cluster A2, two other spike mutations cooccur with F486L, Q314K that is linked to the exact same node in the tree and A262S that already occurred with the last part of cluster A1(NB21 and NB27).

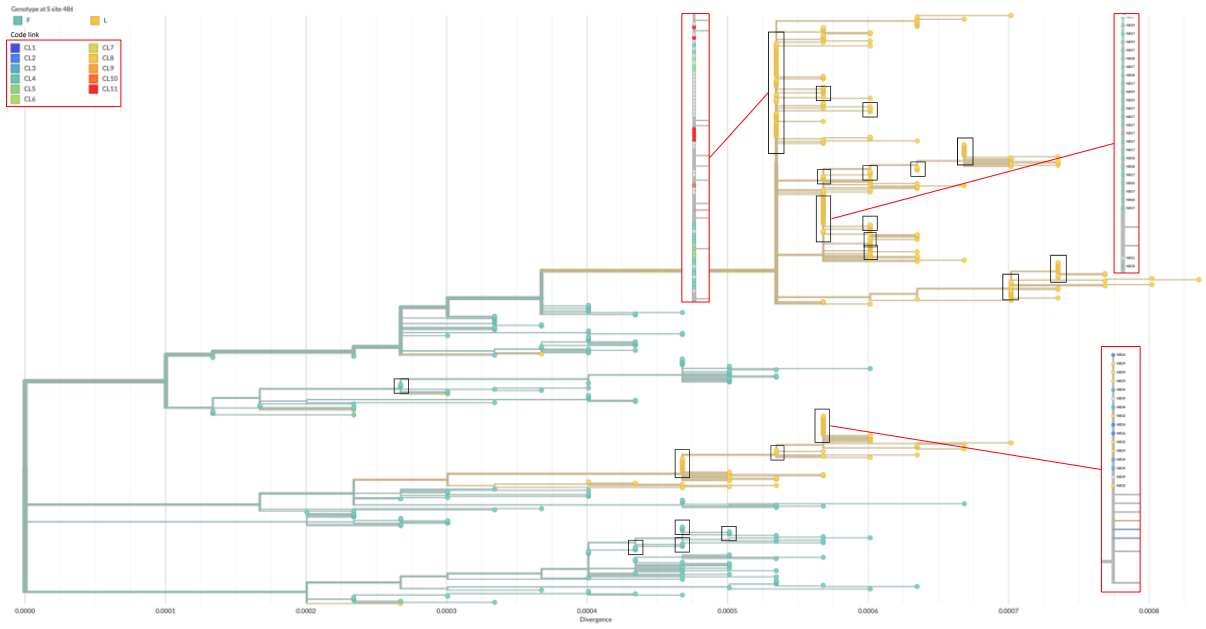

Figure 3. Maximum likelihood phylogeny in divergence view coloured by site 486 of the spike protein. Clonal clusters of sequences spread over multiple farms marked with black frames. Red frames show zoomed views of the three biggest clonal clusters coloured by link codes.

In summary we identified the following mutations among the affected farms: A262S, K314Q, L452M, Y453F F486L, D614G which appeared in high proportion of the samples sequenced within each farm. In Figure 4, the most frequent AA change identified at each farm are shown by cluster.

Two of these A262S and K314Q were not identified by Lu et al(ref) and these changes are clearly associated with cluster A2 (Figure 4).

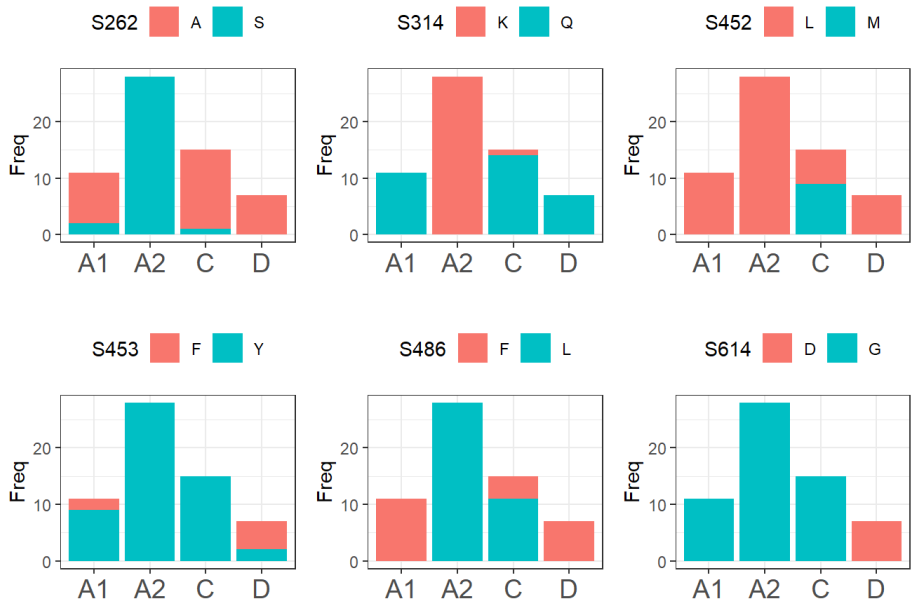

Figure 4. Frequency (number of farms) of observed dominant mutations in the Spike protein identified across the different genetic clusters.

Association between clusters and within farm levels of shedding and infection

Comparison of the shedding levels (adjusted for the animal clinical status) did not reveal any significant difference between the different virus genomic cluster (Figure 3). Overall, no differences in shedding between sick animals and those not showing clinical signs were observed and shedding in annal swabs was significantly lower. However, the apparent prevalence of infection (proportion PCR positives, adjusted for clinical status) was significantly higher (Odds = 6, p= 0.003) in farms affected with virus A2 than A1. No differences between A1 and the other genetic clusters were found (Figure 5).

Farm's contact structure

Analysis of the between-farm contact structure shows a significant association between genetic cluster and the frequency of contacts (Figure 6). With cluster A2 strongly associated with high frequency of contacts. No significant associations were found between veterinary practice assisting the farms and the genetic cluster. A significant association was found between genetic cluster and feed provider.

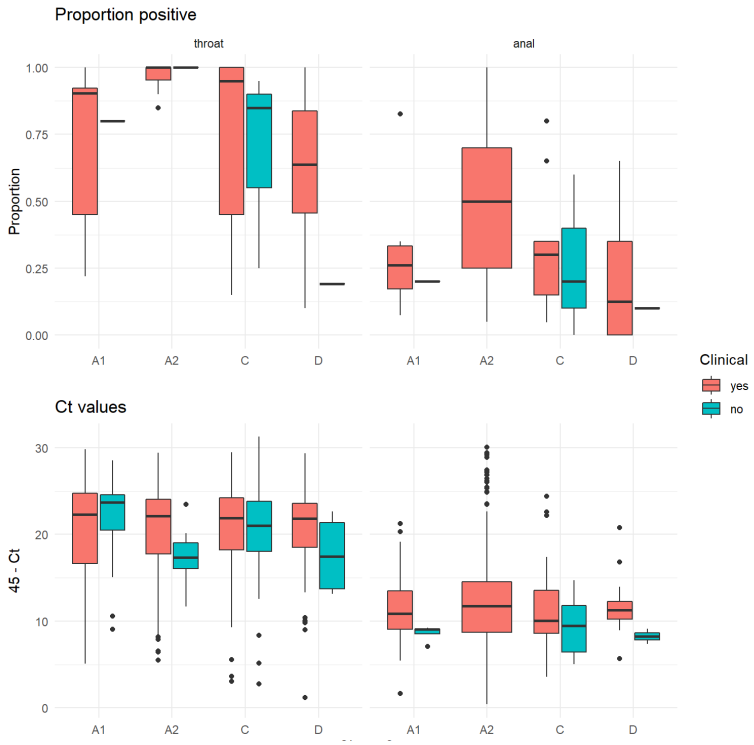

Figure 5. Distribution of within farm apparent prevalence (to panels) and shedding levels of infected animals (lower panels) observed for the virus genetic clusters. A significant higher within farm apparent prevalence was observed in farms infected with A2 viruses than farms infected with A1. No differences were observed between

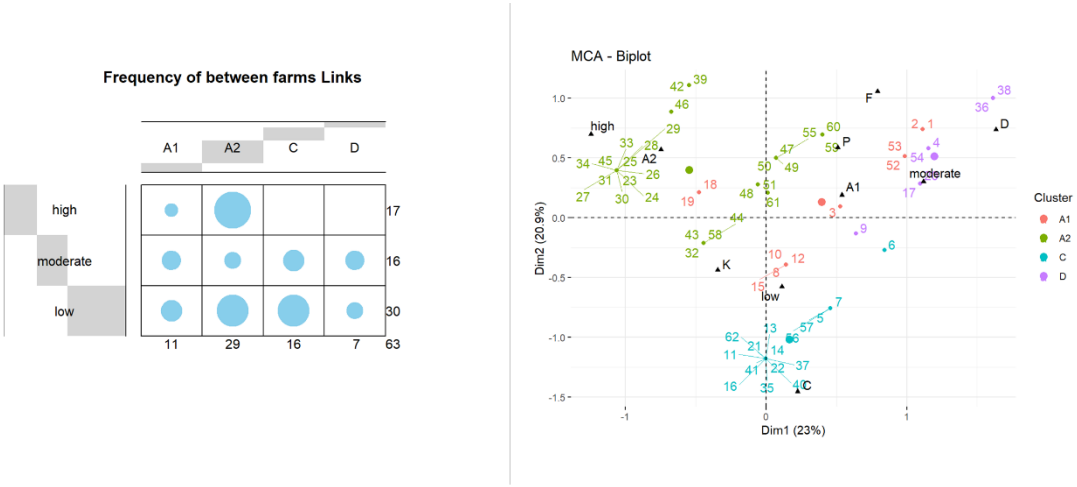

Figure 6. Frequency of contacts (left panel) and multivariable correspondence analysis (right) assessing the association between the frequency of contacts among infected farms in relation to the virus genetic cluster as well as the feed provider (significant variables identified in the multinomial regression). Notice the correlation between A2 and the high frequency of contacts (labels A2 and high, in the upper right quadrant). Labels P, K and F represent feed suppliers.

DISCUSSION

Here we combined genetic and epidemiological data and analysis methods to characterise the farm-to-farm transmission dynamics of SARS-COV-2 infections in mink farms in the Netherlands. Here we provide further confirmation that farm density, similar to other pathogens, has a positive influence of the risk of transmission, with farm dense areas showing a higher risk of between farms transmission than dispersed areas. The "epi-genetic" combined analysis confirmed the evolution of one more genetic cluster, here referred to as cluster A2, than those previously detected (A1, B, C, D and E). The transmission kernel analysis indicated differences in between farm transmission among the different genetic clusters; with Cluster A2 being the most transmissible cluster, as expressed in higher transmission intensity and larder spatial range than the other virus clusters. Cluster A1, the first introduced SARS-COV-2 virus into mink farms in the Netherlands, was the least transmissible cluster. In addition we explored the drivers that could explain differences in transmission between the different clusters. We looked at the presence and distribution of genetic mutations, the between farm contact structure and frequency and infection parameters at host level within the farms.

Looking at the distribution of AA mutations in the spike protein among the clusters, mutations A262S, Q314K and F486L are clearly dominant in Cluster A2, with mutations Q314 and F486L also present to lower a extend in cluster C. Interestingly Cluster C also showed a higher transmissibility than clusters A1 and D.

In addition to the 3 AA mutations dominant in Cluster 2, this cluster was also associated with a higher frequency of contacts between the infected farms. A particular type of link between the farms infected with A2 was that several of this farms had the same ownership and shared equipment and personnel. It is difficult to infer whether the higher transmission is due to presence of this mutations, higher contact frequency or both.

While similar levels of shedding were observed among the different clusters, an apparent higher within farm prevalence of infection were observed in the A2 affected farms than A1. However, we cannot confirm whether this observation reflects higher transmissibility of the virus among minks and consequently between farms or it is a surveillance result. Say something about the influence of time of detection and sampling for A2.

METHODS

Samples and phylogenetic analysis

Samples collected from infected farms (around 20 samples per farm) were deep sequenced and analysed phylogenetically to assess the presence of different viral genetic clusters which could indicate independent introductions (from humans) and potential between farm transmission. Samples were processed by two different labs one sequencing with short reads Illumina (N=315) and the other with Nanopore long reads (N=338). A quality filter was done to exclude all sequences with over 5% undermined bases. In addition, there was some overlap between the samples, thus this collection was filtered to remain with one sample per individual mink, index samples that were not linked to one specific mink were removed, throat samples were favoured over rectal samples and Illumina sequences were favoured over Nanopore as they have a lower error rate. This resulted in a set of 485 sample (Table S1) with max 20 and an average of 7 sequences per farm. For farms NB30, NB31, NB37, NB38 and NB66 no sequences of acceptable quality were available. A maximum likelihood phylogeny and a time calibrated version were created with the NextStrain

pipeline (v 3.0.5). This detailed tree of SARS-CoV-2 infected mink was then used to detect breaks within clusters and clonal spread between farms.

Spatiotemporal transmission patterns

We used the spatiotemporal outbreak pattern (farm location, distance between farms and the time period when a farm was infected and recovered/was removed) (Figure S1) to estimate the between-farm transmission kernel for different virus genetic clusters of outbreaks. This kernel is defined as the transmission hazard between an infected and a susceptible mink farm as a function of the distance between the two farms, and its distance dependence can be used as a signature to assess different potential transmission routes against.

The estimation of transmission kernels was made according to Boender et al., 2007a, Boender et al., 2007b, Boender et al., 2007c. We first calculated the probability of transmission $P(x_{i,j})$ from farm i to j, a distance $x_{i,j}$ apart according to the equation

$$P(x_{i,j}) = 1 - e^{-h(x_{i,j})T_i}$$

where T_i is the infectious period of flock i, and $h(x_{i,j})$ is a transmission kernel (Boender et al., 2007b). The latter describes how the transmission rate scales with distance. We fitted the epidemic data to the following kernel function

$$h(x_{i,j}) = \frac{h_0}{1 + (x_{i,j}/x_0)^{\alpha}}$$

where parameter h_0 determines the maximum value of the kernel, parameter x_0 sets the between-farm distance for which the kernel value is at half of its maximum and parameter α determines the steepness of the kernel.

We then used the kernels to estimate the corresponding R_h for each farm as proxi indicator of the potential infectiousness towards other farms for each of the genetic clusters. We then compared whether there were differences in infectiousness between the different clusters by comparing the median R_h of the different clusters using non-parametric methods.

Factors associated with infection and infectiousness (transmission)

Data on disease determinants (clinical reports, apparent within farm prevalence and virus shedding levels) and on the between-farm contact structure were used to elucidate the role of contact patterns versus virus strain (genetic cluster) properties as determinants of transmission that could explain between-cluster differences in the transmission kernel. For the analyses of contact patterns, a contact was assumed to be represented by a shared link (shared personnel, same ownership, shared equipment) between any two farms. We then used the number of contacts each infected farm had to categorize the frequency of contacts as low (≤ first quantile Q1), medium (≤ Q2) and high (> Q2). To assess whether there were differences in the frequency of contacts between farms infected with the different genetic clusters, we used multivariable corresponding analysis and multinomial regression. In addition to the frequency of contacts, the feed provider (multiple farms were provided by the same company), the veterinary practice visiting the farms, the province where a farms was located and potential interactions between these variables were included in the analysis. To assess whether virus strain (genetic cluster) infection properties (disease determinants) could have influenced transmission, we compared the within farm apparent prevalence at the time of sampling and the virus shedding levels observed in sampled (showing or not showing clinical signs) minks from each of the infected farms. These comparisons were done

using generalized mixed regression models (GLMM) with a binomial (compare apparent prevalence) or normal (shedding) error distribution.

ACKNOWLEDGMENT

This research was funded by the European Union's Horizon 2020 Research and Innovation programme, grant agreement No 773830: One Health European Joint Programme, projects COVRIN.

REFERENCES

Boklund, A., Hammer, A.S., Quaade, M.L., Rasmussen, T.B., Lohse, L., Strandbygaard, B., Jørgensen, C.S., Olesen, A.S., Hjerpe, F.B., Petersen, H.H., Jensen, T.K., Mortensen, S., Calvo-Artavia, F.F., Lefèvre, S.K., Nielsen, S.S., Halasa, T., Belsham, G.J., Bøtner, A., 2021. SARS-CoV-2 in Danish Mink Farms: Course of the Epidemic and a Descriptive Analysis of the Outbreaks in 2020. Animals (Basel) 11.

Chaintoutis, S.C., Thomou, Z., Mouchtaropoulou, E., Tsiolas, G., Chassalevris, T., Stylianaki, I., Lagou, M., Michailidou, S., Moutou, E., Koenen, J.J.H., Dijkshoorn, J.W., Paraskevis, D., Poutahidis, T., Siarkou, V.I., Sypsa, V., Argiriou, A., Fortomaris, P., Dovas, C.I., 2021. Outbreaks of SARS-CoV-2 in naturally infected mink farms: Impact, transmission dynamics, genetic patterns, and environmental contamination. PLoS Pathog 17, e1009883.

Hammer, A.S., Quaade, M.L., Rasmussen, T.B., Fonager, J., Rasmussen, M., Mundbjerg, K., Lohse, L., Strandbygaard, B., Jørgensen, C.S., Alfaro-Núñez, A., Rosenstierne, M.W., Boklund, A., Halasa, T., Fomsgaard, A., Belsham, G.J., Bøtner, A., 2021. SARS-CoV-2 Transmission between Mink (Neovison vison) and Humans, Denmark. Emerg Infect Dis 27, 547-551.

Lu, L., Sikkema, R.S., Velkers, F.C., Nieuwenhuijse, D.F., Fischer, E.A.J., Meijer, P.A., Bouwmeester-Vincken, N., Rietveld, A., Wegdam-Blans, M.C.A., Tolsma, P., Koppelman, M., Smit, L.A.M., Hakze-van der Honing, R.W., van der Poel, W.H.M., van der Spek, A.N., Spierenburg, M.A.H., Molenaar, R.J., Rond, J.d., Augustijn, M., Woolhouse, M., Stegeman, J.A., Lycett, S., Oude Munnink, B.B., Koopmans, M.P.G., 2021. Adaptation, spread and transmission of SARS-CoV-2 in farmed minks and associated humans in the Netherlands. Nature Communications 12, 6802.

Oreshkova, N., Molenaar, R.J., Vreman, S., Harders, F., Oude Munnink, B.B., Hakze-van der Honing, R.W., Gerhards, N., Tolsma, P., Bouwstra, R., Sikkema, R.S., Tacken, M.G., de Rooij, M.M., Weesendorp, E., Engelsma, M.Y., Bruschke, C.J., Smit, L.A., Koopmans, M., van der Poel, W.H., Stegeman, A., 2020. SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. Euro Surveill 25.

Oude Munnink, B.B., Sikkema, R.S., Nieuwenhuijse, D.F., Molenaar, R.J., Munger, E., Molenkamp, R., van der Spek, A., Tolsma, P., Rietveld, A., Brouwer, M., Bouwmeester-Vincken, N., Harders, F., Hakze-van der Honing, R., Wegdam-Blans, M.C.A., Bouwstra, R.J., GeurtsvanKessel, C., van der Eijk, A.A., Velkers, F.C., Smit, L.A.M., Stegeman, A., van der Poel, W.H.M., Koopmans, M.P.G., 2021. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. Science 371, 172-177.

Table S1: Number of sequences per cluster

clade	number of sequences	farms
A1	91	11
A2	238	29
B	11	1
C	79	14
D	53	8
E	12	1
NA	1	1
total	485	65

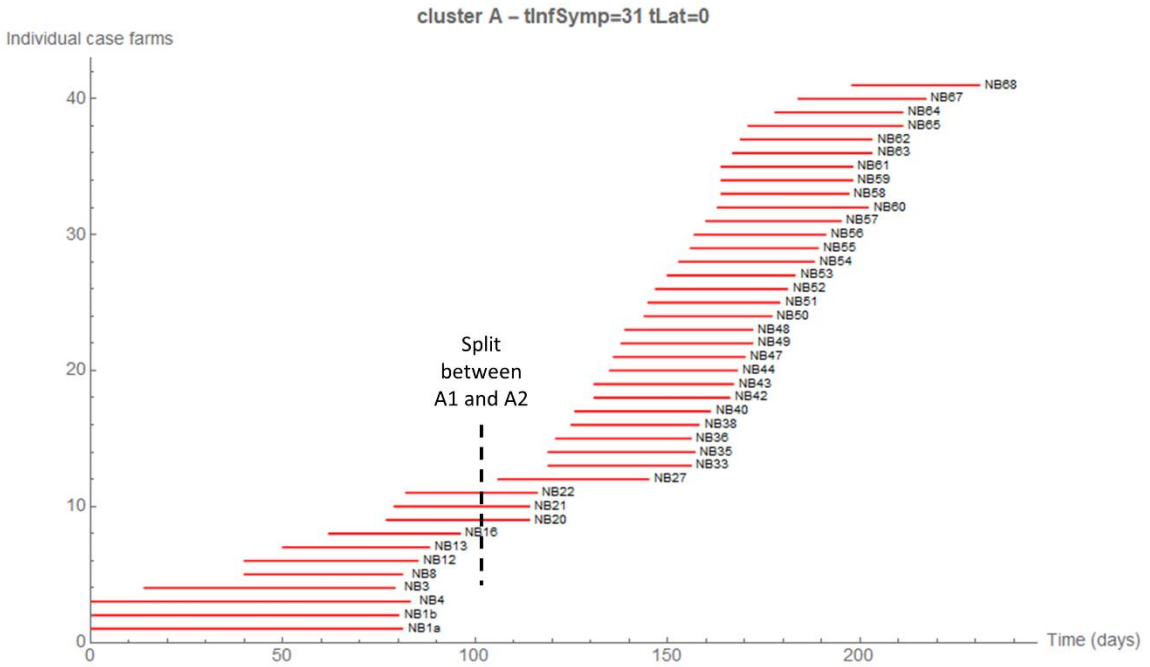

Figure S1. Temporal display of time of assumed infection and duration of infectiousness of affected farms within Clusters A1 and A2. The time when the expected divergence of cluster A2 from Cluster A1 is shown with a horizontal dashed line.