Article

Integration of Epidemiological and Genomic Data to Investigate H5N1 HPAI Outbreaks in Northern Italy in 2021-2022

Diletta Fornasiero [1]*, Alice Fusaro [1], Bianca Zecchin [1], Matteo Mazzucato [1], Francesca Scolamacchia [1], Grazia Manca [1], Calogero Terregino [1], Tiziano Dorotea [1]*, Paolo Mulatti [1]

[1] Istituto Zooprofilattico Sperimentale delle Venezie, 35020 Legnaro, Padova, Italy
* Correspondence: D.F., dfornasiero@izsvenezie.it; T.D., tdorotea@izsvenezie.it

Abstract: Between October 2021 and April 2022, 317 outbreaks caused by highly pathogenic avian influenza (HPAI) H5N1 viruses were notified in poultry farms in the northeastern Italian Regions. The complete genomes of 214 strains were used to estimate the genetic network based on the virus similarity. An exponential random graph model (ERGM) was used to assess the effect of at-risk contacts, same owners, in-bound/out-bound risk windows overlap, genetic differences, geographic distances, same species and poultry company, on the probability of observing a link within the genetic network, which can be interpreted as the potential propagation of the epidemic via lateral spread or a common source of infection. The variables same poultry company (Est.=0.548, C.I.=[0.179;0.918]) and risk windows overlap (Est.=0.339, C.I.=[0.309;0.368]) were associated with a higher probability of link formation, while the genetic differences (Est.=-0.563, C.I.=[-0.640;-0.486]) and geographic distances (Est.=-0.058, C.I.=[-0.078;-0.038]) indicated a reduced probability. The integration of epidemiological data with genomic analyses allows monitoring the epidemic evolution and helps explain the dynamics of lateral spreads suggesting the potential diffusion routes. The 2021-2022 epidemic stresses the need to further strengthen the biosecurity measures, and to encourage the reorganisation of the poultry production sector to minimize the impact of future epidemics.

Keywords: HPAI; H5N1; Italy; genetic network; epidemiological investigation; contact tracing; ERGM

1. Introduction

The end of the 2021 summer season was characterized by several highly pathogenic avian influenza (HPAI) outbreaks of the H5 and H5N1 subtypes in wild and domestic birds in western Russia and close to the western and eastern borders with Kazakhstan and Mongolia countries. These areas are well-known autumn staging sites for wild water birds that are later ready to migrate to Europe for overwintering. The migration of infected wild birds towards the northern and eastern European countries has been considered one of the major risks for the introduction of avian influenza (AI), and its further spread to the southern and western areas of Europe, recurring almost every autumn-winter season [1]. In 2021–2022 Europe experienced the largest highly pathogenicity avian influenza (HPAI) epidemic ever observed, affecting a total of 37 European Countries, with approximately 2,500 outbreaks notified in poultry, 47.7 million birds culled, 187 outbreaks in captive birds, and around 3,600 in wild birds [1].

In the past three decades, Italy has been recurrently affected by the circulation of AI viruses, with epidemics of both HPAI and Low Pathogenicity Avian Influenza (LPAI). This is likely due to the high concentration of poultry farms in limited geographic areas (Densely Populated Poultry Areas - DPPAs), in close proximity to vast wetlands located along the Black Sea–Mediterranean and the East Atlantic migratory flyways, which may

 doi:10.20944/preprints202212.0404.v1

serve as wintering and nesting areas to many migratory wild waterfowl [2,3]. Despite the large epidemics of HPAI and LPAI that occurred in the late 1990s and early 2000s, only sporadic cases were observed until 2016, when an H5N8 HPAI virus was responsible for two consecutive epidemic waves, causing a total of 83 outbreaks in poultry farms and 14 in wild birds, mainly in the northern Regions of Italy [4].

On 18 October 2021, an H5N1 HPAI virus was detected in a fattening turkey farm in Verona Province, in the very core of the DPPAs. In the following weeks, the number of confirmed outbreaks rapidly increased and the epidemic spread to the neighbouring Provinces and Regions. By April 2022, a total of 317 outbreaks were notified in poultry farms and 23 were reported in wild birds, affecting seven different Italian Regions. The viral circulation mainly occurred within the DPPAs located between the Veneto and Lombardia Regions and in particular in the Provinces of Verona, Padova, Vicenza, Brescia, and Mantova. Similarly to the 2017-2018 HPAI H5N8 epidemic [4], the AI viruses appeared to have been directly introduced as highly pathogenic strains. The epidemic predominantly affected fattening turkey farms, as commonly observed in past events [4–6]; however, unlike the epidemics experienced since 1999, the 2021-2022 epidemic was characterized by a marked involvement of the broiler sector, although often with scarce or no symptoms reported. Of the 23 confirmed cases reported in wild birds, only a limited number involved birds belonging to the Anseriformes order (considered as target species for the passive surveillance activities for AI in wild populations), while several cases were reported in species usually considered marginally affected by HPAI viruses, such as magpies and owls [7,8]. The phylogenetic analyses suggested that at least 12 different introductions into the domestic sector and wild populations occurred during the epidemic, indicating that lateral transmission events most likely happened, determining a very high number of secondary cases in the poultry sector [9].

Hereby we present an approach to integrate genomic and epidemiological data collected during the 2021-2022 HPAI H5N1 epidemic in Italy, applying techniques of network analysis to genetic networks. Such an approach allowed to assess the effects of potential drivers of virus circulation on the emergence of secondary outbreaks via lateral spreads in domestic poultry premises. The results may provide insights on shortcomings in the management of an epidemic, and could be used to inform future prevention measures to limit the massive spread between farms in DPPAs.

2. Materials and Methods

Genomic analysis and network construction

The complete genome of 342 Italian viruses was generated as previously described [4]. Virus sequences were deposited in the GISAID EpiFlu Database (https://platform.gisaid.org/) with the accession numbers reported in Supplementary Table S1. Sequences were aligned in MAFFT v7 [10], and compared to the most related sequences available in GISAID. Maximum likelihood phylogenetic trees of each gene segment were constructed by using IQTREE v1.6.6 (https://github.com/iqtree/iqtree1) and an ultrafast bootstrap resampling analysis (1000 replications). Phylogenetic trees were visualized in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). The datasets used to generate the phylogenetic trees included all the Italian viruses sequenced from October 2021 to April 2022.

The phylogenetic network was generated using the Median Joining (MJ) method implemented in NETWORK 10.2.0.0 [11], for the eight concatenated gene segments of non-reassortant H5N1 viruses, allowing visualizing how viral genomes are connected on the basis of their genetic similarity. To this purpose, the phylogenetic network analysis was run on a dataset composed of 229 Italian non-reassortant H5N1 viruses (including viruses from domestic and wild birds, marked with a star in Supplementary Table S1), nine H5N1 viruses from Poland and eight H5N1 viruses from the Czech Republic belonging to the same genetic cluster. The MJ network uses a maximum parsimony ap-

 doi:10.20944/preprints202212.0404.v1

proach to reconstruct the relationships between highly similar sequences; it consists of nodes and links (also referred to as ties or edges) that connect the nodes. Two connected nodes within the network represent a dyad. The nodes can be either isolated viral strains or median vectors, which represent hypothesised sequences used to connect the existing viruses in the most parsimonious way.

From the hypotheses generated by the interpretation of the phylogenetic tree and genetic network described above, a second MJ network was constructed using only a subset of the available viral strains isolated from the Italian poultry compartment (referred to as cluster A), containing only a single isolate per outbreak. This network was then used for the analysis to assess the impact of epidemiological drivers on the potential viral spread between poultry farms.

Network covariates

The information collected through on-field epidemiological enquiries, contact tracing activities, and phylogenetic analyses were exploited to derive a series of factors to be further analysed as covariates in the genetic network. All data cleaning, variables calculation and statistical analysis operations were performed using the R [12] and RStudio [13] software. The raw data were processed to obtain three types of terms, with different information accordingly to the specific network feature to which they refer: i) network density, ii) nodal-level covariates and, iii) edge-level covariates. Network density refers to the total number of edges observed in the network, and it is used to evaluate whether the observed network exhibits an intrinsic non-random structure that could be related to a series of other potential drivers. The nodal-level covariates encompass information associated with the nodes within the network, and related to the outbreaks. Specifically, the nodal homophily (i.e. the tendency of two nodes to share the same attribute) was taken into account for species reared and poultry company, to include information on the similarity of the two end nodes (i.e. outbreaks) of a potential tie (i.e. highest genetic similarity) in rearing the same species and belonging to the same company. Therefore, two dichotomic matrices were built considering all the outbreaks pairwise comparisons, to indicate whether the compared outbreaks had the same condition (i.e. rearing the same species or belonging to the same poultry integrator). The edge-level covariates included a list of 5 pairwise matrices containing information characterising each outbreaks pair (i.e. dyadic relationship): (i) overlap between risk time windows; (ii) number of at-risk contacts; (iii) geographical distances; (iv) belonging to the same owner; and (v) genetic differences between the viruses identified at each infected farms.

The risk time windows overlap represents the period during which each farm was exposed to the potential infection spreading from other infected premises. It is determined considering the overlap of two at-risk time windows: an 'out-bound risk window' (ORW), representing the period in which avian influenza virus (AIV) could have been spread from an infected farm, and an 'in-bound risk window' (IRW), which indicated the period in which an AIV could have been introduced into an infected farm. The ORW was estimated to range from four days prior to the onset of symptoms (to account for the period in which the birds can be infectious without showing symptoms [14]) until the extinction of the outbreak. The IRW was assumed to begin 15 days before the symptom onset, accounting for the viral incubation period [15], and ending on the day on which symptoms were observed. The risk time windows overlap was calculated as the number of days in which the ORW and the IRW of each pair of outbreaks overlapped. This measure could range from a maximum of 16 days (corresponding to the full overlap between the two risk windows) to the minimum value of 0 when the two periods did not overlay.

As for the at-risk contacts, all the available data on the movement of vehicles and personnel (i.e. technicians and veterinarians) were entered into a structured dataset, and analysed through the R package EpiContactTrace [16] considering the IRW and ORW previously described. This allowed identifying the total number of contacts occurred

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0404.v1

between farms that could have been considered productive in terms of infection spread. Geographic distances were included to account for farm density and proximity, globally considered among the most important drivers of disease spread [17,18]. For the geographic data manipulation and calculation, the R packages rgdal [19] and sp [20,21] were used.

Infected farms belonging to the same owner, or close family members, were identified during the on-field epidemiological enquiries. The potential inter-outbreaks contacts due to the possible unreported movements of the same owners were recorded into a symmetric dichotomic matrix.

The pairwise genetic distances between the complete genome of the viruses were calculated in MEGA 7.0.26 [22] and included in the analyses as the number of different nucleotides between the genotyped strains.

The list of all the network covariates included in the analysis is summarised in Table 1.

Table 1. List of network covariates included in the analysis.

Type	Name	Description
Density	Edges	The overall network density considering the total number of observed ties.
Nodal-level covariates	Species	Network homophily by species.
	Poultry company	Network homophily by poultry company.
Edge-level covariates	Risk windows overlap	Number of overlapping days considering the out-bound and the in-bound risk time windows.
	At-risk contacts	Number of potentially productive contacts occurred between farms.
	Geographic distances	Farms proximity calculated as Euclidean distance (m).
	Same owner	Presence/absence of an owner of multiple infected farms.
	Genetic differences	Number of different bases between the isolated strains.

Statistical analysis

The genomic and epidemiological information were combined to evaluate which drivers of infection could significantly explain the structure of the observed genetic network. Since a link between two nodes within the network represents the highest genetic similarity between viral strains, a potential lateral spread or common source of infection was assumed for the two infected farms the viruses were isolated from. An exponential random graph model (ERGM) framework was developed to assess which potential risk factor can influence the probability of observing a link between two nodes, and hence might have significant effects of inter-farm transmission. The ERGM is a statistical model for network data that takes a generalized exponential family form and specifies the probability of an entire network as a function of covariates hypothesised affecting the network structure [23], and evaluates alternative hypotheses on the processes that could lead to the observed outcome [24]. For more detailed information on the exponential random graph modelling framework, its definitions and mathematical notation refer to [23–26].

In this study, the ERGM was used to perform a regression-like analysis where the dependent variable is the presence (or absence) of a link between two nodes (i.e. two isolated viruses), while the explanatory variables are the characteristics of each outbreak and the dyadic relationships between pairs, as described above. The ERGM estimated coefficients can be interpreted in a similar fashion to those obtained from a logistic regression, and represent the change in the (log-odds) likelihood of a tie to occur, for a unit change in a predictor. The log-odds form of each predictor included in the analysis can be transformed into a corresponding probability (through the inverse-logit form), to ease the

 doi:10.20944/preprints202212.0404.v1

interpretation of how each variable affected the probability of observing ties between nodes. The model goodness of fit was assessed through the tie prediction statistic, which predicts the dyad states of the observed network by the dyad states of 1000 simulated networks. The tie prediction statistic includes the receiver-operating characteristics (ROC) and precision-recall (PR) curves [27].

All of the analyses were performed on the R software, using the following packages: ergm [28–30] for the model fit and analysis, btergm [31] for the calculus of the ties predicted probabilities, and ggplot2 [32], ggpubr [33], igraph [34] and ggraph [35] for the graphical representation of the genetic network and the results of the analysis.

3. Results

Evolution of the Italian epidemiological situation

A total of 317 HPAI H5N1 outbreaks in poultry farms and 23 in wild birds were confirmed between 18 October 2021 and 1 April 2022 in nine Italian Regions (Figure 1). The domestic outbreaks were located in seven Regions, with Veneto and Lombardia being the ones where most cases were identified (Table 2). Fattening turkey farms (n=150/317, 47.32%) were predominantly affected, as commonly observed in past events [4,5]; however, a strong involvement of the broiler sector was also registered, with about a quarter of the outbreaks notified in this production type (n=77/317, 24.29%). The outbreaks reported in wild birds were found in eight Italian Regions spanning from Piemonte to the southern Regions of Campania and Puglia (Figure 1, orange dots). Only six outbreaks out of a total of 23 involved birds belonging to the order Anseriformes (n=6/23, 26.09%), considered among the target species for passive surveillance for Avian Influenza in wild populations [7]. Most of the outbreaks in wild birds affected seagulls (n=6/23, 26.09%), birds of prey (n=4/23, 17.39%), owls (n=4/23, 17.39%), herons (n=2/23, 8.70%), and magpies (n=2/23, 8.70%). Some of the outbreaks in wild birds involved wildlife rescue centers, with cases in more than a single species.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0404.v1

Figure 1. Italian notified HPAI H5N1 outbreaks distribution in domestic poultry and wild birds during the 2021-2022 epidemic.

Table 2. Avian influenza outbreaks in poultry farms by Region.

Region	Fattening turkey	Broiler[1]	Laying hen[2]	Breeder[3]	Multi-species[4]	Other species[5]	Total
Emilia-Romagna					2		2
Friuli-Venezia Giulia		1					1
Lazio					1		1
Lombardia	20	9	20		5	6	60
Piemonte		1					1
Toscana					4		4
Veneto	130	65	32	5	7	9	248
Total	150	77	51	5	17	15	317

[1]Chicken broilers and cockerels; [2]Pullets and ready-to-lay hens; [3]Guinea fowl breeders, chicken breeders and turkey breeders; [4]Rural/backyard farms, captive birds and agritourisms; [5]Quails, pheasants and guinea fowls.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0404.v1

The epidemic showed a rapid evolution, with a steady increase in the number of weekly cases until the second week of December, peaking at 51 cases per week. During this period, most outbreaks were notified from Veneto and Lombardia, while rare incursions were detected in other Regions. From the third week of December, there was a sharp decline in the number of newly reported cases (Figure 2), and the latest were sporadic outbreaks confirmed in areas outside the area of major viral circulation, confirmed between mid-March and early April, in two rural farms in Toscana and Emilia-Romagna.

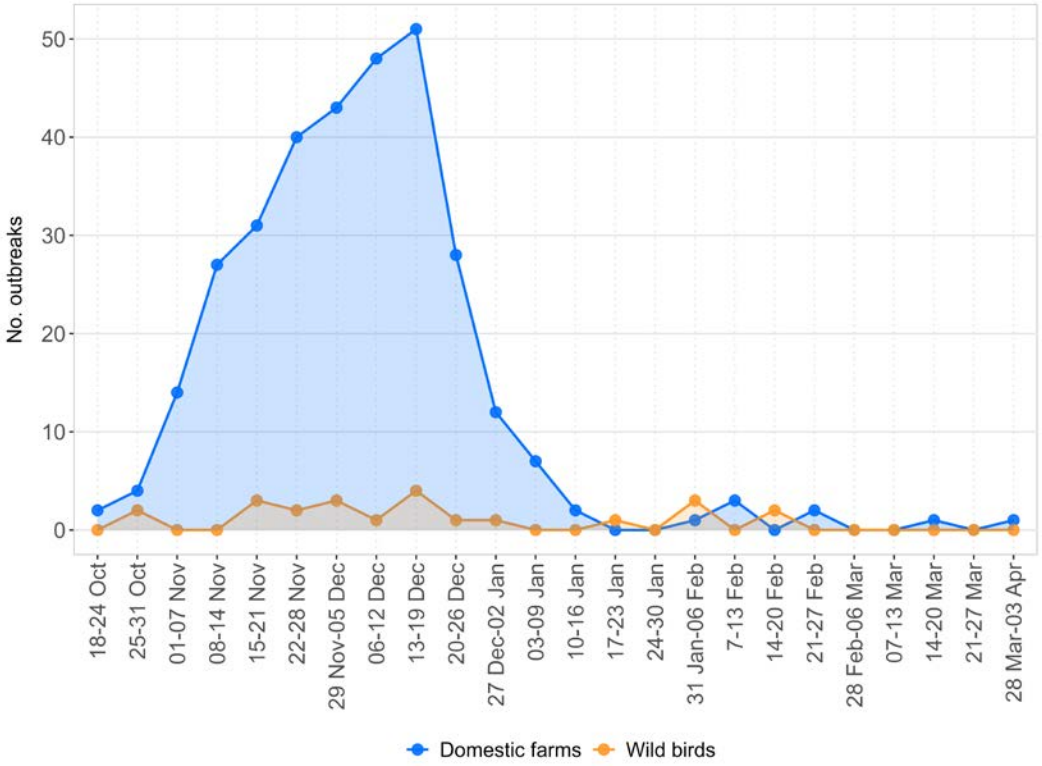

Figure 2. The epidemic curve showing the number of weekly cases in domestic and wild birds.

Genetic analyses

We characterized the complete genome of 342 HPAI H5N1 viruses collected from wild (n=21) and domestic (n=321) birds, for the outbreaks reported in Italy between October 2021 and April 2022; in some cases more than a single isolate per outbreak was sequenced. All the viruses belonged to the clade 2.3.4.4b [36], and clustered with viruses that had been circulating in Russia and Europe since the end of 2020. The phylogenetic analyses of the eight gene segments showed that the Italian HPAI H5N1 viruses belonged to six distinct genotypes originating from different reassortment events (Figure 3).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0404.v1

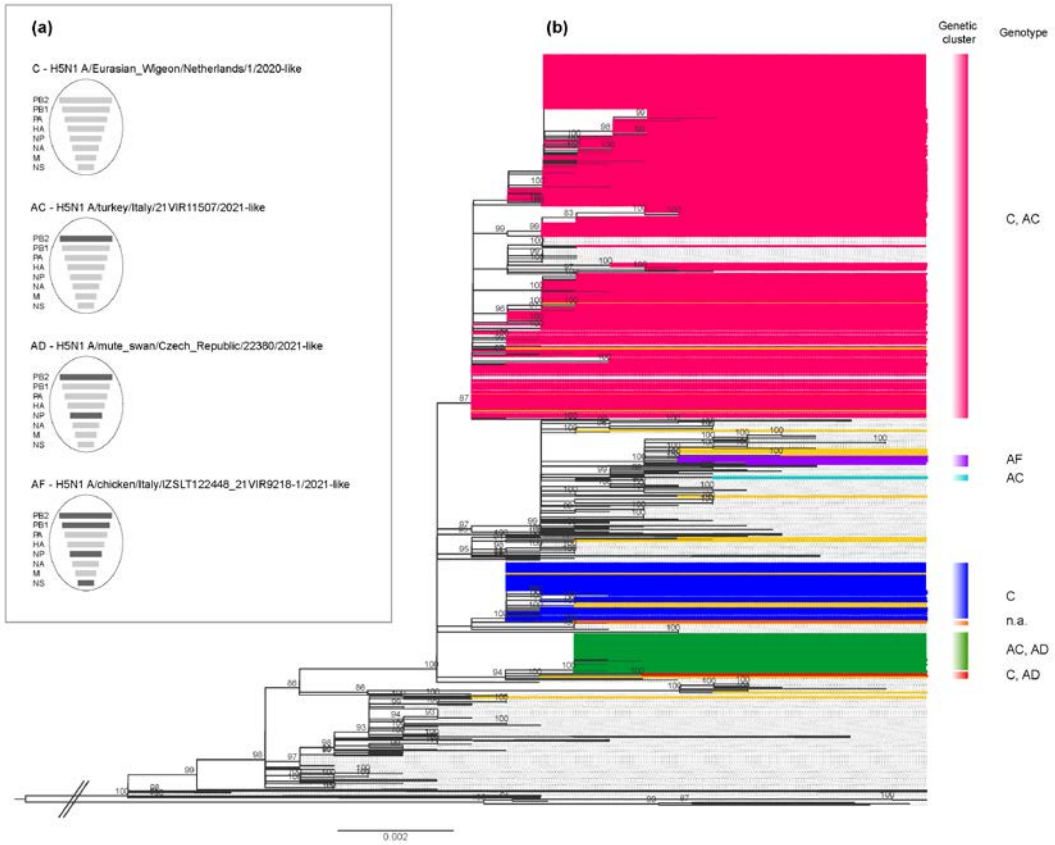

Figure 3. Phylogenetic tree of the hemagglutinin (HA) gene segment and genotypes identified among the Italian H5N1 viruses. (**a**) Schematic representation of the gene composition of each genotype. Light-grey bars: conserved segments; dark-grey bars: gene segments acquired by reassortment events. (**b**) Phylogenetic tree of the HA gene segment coloured according to seven different genetic clusters. C, AC, AD, AF: different genotypes characterizing the Italian viruses; n.a.: viruses not assigned to genotypes due to the absence of one or more gene segments; pink cluster: genetic group that includes most of the Italian viruses collected from October 2021 to April 2022; dark-yellow sequences: viruses collected from wild birds.

The genetic clustering suggested the occurrence of multiple viral introductions in the country, of which at least seven in domestic birds (marked with different colours in the HA phylogenetic tree, Figure 3b). Most of the viruses collected from poultry outbreaks fell within three major clusters (in pink, blue and green in Figure 3b), suggesting a sustained virus spread among poultry farms after the virus incursion into the domestic sector. Specifically, the genetic group which includes most of the Italian viruses (pink cluster in Figure 3b), was the most widespread in northern Italy and it was detected in several Provinces of the Veneto and Lombardia Regions (Figure 4), as well as in Poland and Czech Republic (Supplementary Figure S2). To elucidate the possible transmission dynamics among the affected farms, the viruses belonging to the pink cluster were subjected to a more in-depth study by using the phylogenetic network analysis based on the complete genome sequences of 229 Italian strains (Supplementary Figure S2). The viruses showed a clear clustering by Province, which might be indicative of the emergence of several secondary outbreaks caused by viral spread among the poultry farms. However, the occurrence of multiple introductions of genetically related viruses from wild birds into the domestic population cannot be completely excluded. The available genetic data make not possible to distinguish the new introductions from wild birds from secondary spreads. Therefore, the identification of viruses with similar genetic constellations in other countries, and the presence of long branches connecting poultry outbreaks with no evident epidemiological connection supported the hypothesis that this cluster could have

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0404.v1

been generated as a consequence of more than one virus incursion from wild birds into the domestic sector.

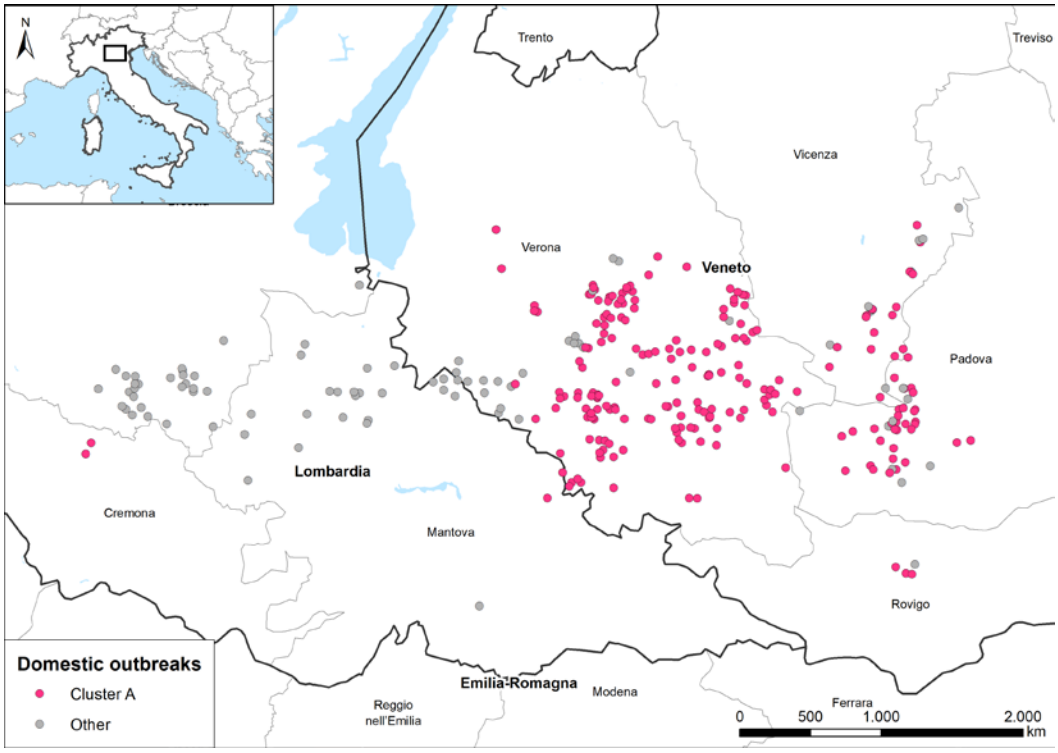

Figure 4. Geographic distribution of the domestic outbreaks in the highly densely populated poultry area. Pink dots: cluster A outbreaks; grey dots: non-clustered outbreaks or outbreaks belonging to other minor genetic groups.

To further investigate the possible lateral transmission dynamics among the affected farms, a subset of the viruses belonging to the phylogenetic network was considered to construct a new genetic network (cluster A, Figure 5). This subset of sequences was selected as they belong to the largest group in terms of domestic outbreaks isolates and available epidemiological data. Moreover, the various genetic groups have markedly different phylogenetic characteristics, suggesting that they might have different epidemiology and transmission dynamics.

The cluster A network contains 249 nodes (214 of which were individual viral strains isolated from single Italian domestic outbreaks belonging to the larger genetic group, while 35 are the estimated median vectors), and 345 edges. The 214 outbreaks were confirmed between October and the end of December 2021 and they were mainly distributed in the Veneto Provinces (3 in Rovigo, 32 in Padova, 17 in Vicenza and 158 in Verona), while only 4 were located in the Lombardia Region (2 in Cremona and 2 in Mantova Provinces), as shown in Figure 4.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0404.v1

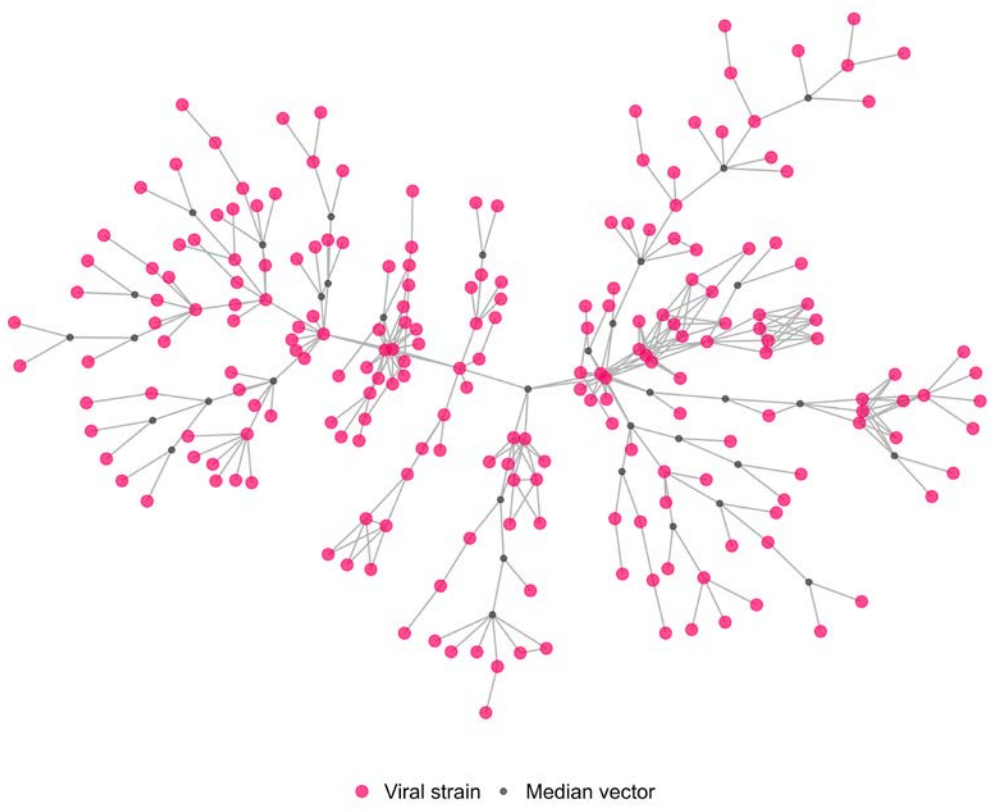

Figure 5. Graphical representation of the estimated genetic network (cluster A). Pink nodes: H5N1 HPAI virus isolated from individual domestic outbreaks belonging to the larger genetic group; grey nodes: estimated median vectors; links: highest genetic similarity between nodes.

Epidemiological analysis of the genetic network

The epidemiological data collected for the domestic outbreaks belonging to the genetic cluster A were summarized through an initial exploratory analysis. The contact tracing operations identified a total of 58 at-risk contacts through the movements of personnel and feed trucks, which involved 46 farms (up to a maximum of 4 at-risk contacts per farms dyad). Forty-five farms were found to belong to the same owners or relatives in the cluster included in the analysis, with up to six premises per single farmer. A total of 26 poultry companies, in addition to the rural sector, resulted affected by the end of the epidemic, and the one most affected had 108 outbreaks. All of the species and productive types were represented in cluster A: breeders (n=3/214, 1.40% of the total outbreaks in the cluster), broilers (n=55/214, 25.70%), laying hens (n=27/214, 12.62%), fattening turkeys (n=114/214, 53.27%), multispecies (n=6/214, 2.80%) and other species (n=9/214, 4.21%). The geographic distances between the infected farms ranged from a minimum of 130 meters to a maximum of approximately 126 km, with a median of 22 km (interquartile range - IQR: 14-35). Lastly, the viral genetic differences ranged from a minimum of 0 (identical viral strains) to a maximum of 163 different bases, with a median of 21 (IQR: 13-27) bases.

The ERGM output can be similarly interpreted as the results that would be obtained from a logistic model. The coefficients are the change in the (log-odds) likelihood of a tie for a unit change in a predictor. Each log-odds of a tie can be transformed, through the logistic function, into a corresponding probability, considering the additive effect of each coefficient included in the model. The coefficient estimate for the *edge* term, which represents the network density, is negative ($Est._{log\text{-}odds}$=-3.856, $C.I._{log\text{-}odds}$=[-4.009; -3.702]), indicating a much lower probability of observing a tie than expected in a network of the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0404.v1

same size (tie baseline probability of 0.021). This can be interpreted as the fact that the observed network owns an intrinsic structure, in which the presence of ties between nodes (i.e. highest similarity between different isolated viral strains) can be explained by the influence of other variables. Four out of the seven predictors in the ERGM were significantly associated with the presence of network ties (Table 3). The *poultry company* and *risk time windows* variables were positively associated with a +13.4% higher probability of observing a tie when two farms belonged to the same company ($Est._{log\text{-}odds}$=0.548, $C.I._{log\text{-}odds}$=[0.179; 0.918]), and +8.4% for each additional day of exposure ($Est._{log\text{-}odds}$=0.339, $C.I._{log\text{-}odds}$=[0.309; 0.368]), respectively. Conversely, the *genetic differences* and *geographic distance* variables were negatively associated with the probability of observing two nodes connected within the genetic network. Specifically, a 13.7% lower probability was calculated for each additional different genomic base in the sequenced strains ($Est._{log\text{-}odds}$=-0.563, $C.I._{log\text{-}odds}$=[-0.640; -0.486]), and -1.4% for each additional kilometre of distance between farms ($Est._{log\text{-}odds}$=-0.058, $C.I._{log\text{-}odds}$=[-0.078; -0.038]). No significant effects were associated with the *species*, *at-risk contacts* and *same owners* variables.

Table 3. Coefficient estimates, standard errors (S.E.), 95% coefficients intervals, P-values and corresponding probabilities for the explanatory variables included in the ERGM analysis.

Variable	Coefficient estimate (log-odds)	S.E.	C.I. [2.5%; 97.5%]	P-value[1]	Corresponding probability (inverse logit)	Probability vs randomness
Edges	-3.856	0.078	[-4.009; -3.702]	***	0.021	-47.9%
Species	0.218	0.179	[-0.132; 0.568]			
Poultry company	0.548	0.188	[0.179; 0.918]	**	0.634	+13.4%
At-risk contacts	-0.062	0.591	[-1.222; 1.096]			
Same owners	0.718	0.603	[-0.465; 1.900]			
Risk time windows	0.339	0.015	[0.309; 0.368]	***	0.584	+8.4%
Genetic differences	-0.563	0.039	[-0.640; -0.486]	***	0.363	-13.7%
Geographic distances	-0.058	0.010	[-0.078; -0.038]	***	0.486	-1.4%

[1] P-value: ***, P < 0.0001; **, P < 0.001.

The predicted probability distributions of the network ties are graphically presented in Figure 6. Specifically, the predicted tie probabilities are illustrated according to the three edge-level covariates that resulted significant in the model analysis (i.e. the geographic distances, the risk time windows and the genetic differences), and differentiated according to the actual presence/absence of the ties within the genetic network. The probabilities estimated for the linked nodes within the genetic network are overall higher compared to the ones obtained for the unlinked nodes, which is indicative of good model performance. The predicted probabilities are the highest when the between-farms distances are below 1 km (median 0.35, IQR: 0.13-0.71), and when the exposure is extended to the maximum period of 15-16 days (median 0.56, IQR: 0.33-0.69, and 0.56, IQR: 0.37-0.74, respectively). As expected, the estimated ties probabilities according to the genetic distances tend to decrease as the number of different genetic bases increases, reaching the lowest median value between the 6 and 8 bases (median 0.00, IQR: 0.00-0.09, and 0.00, IQR: 0.00-0.02, respectively).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0404.v1

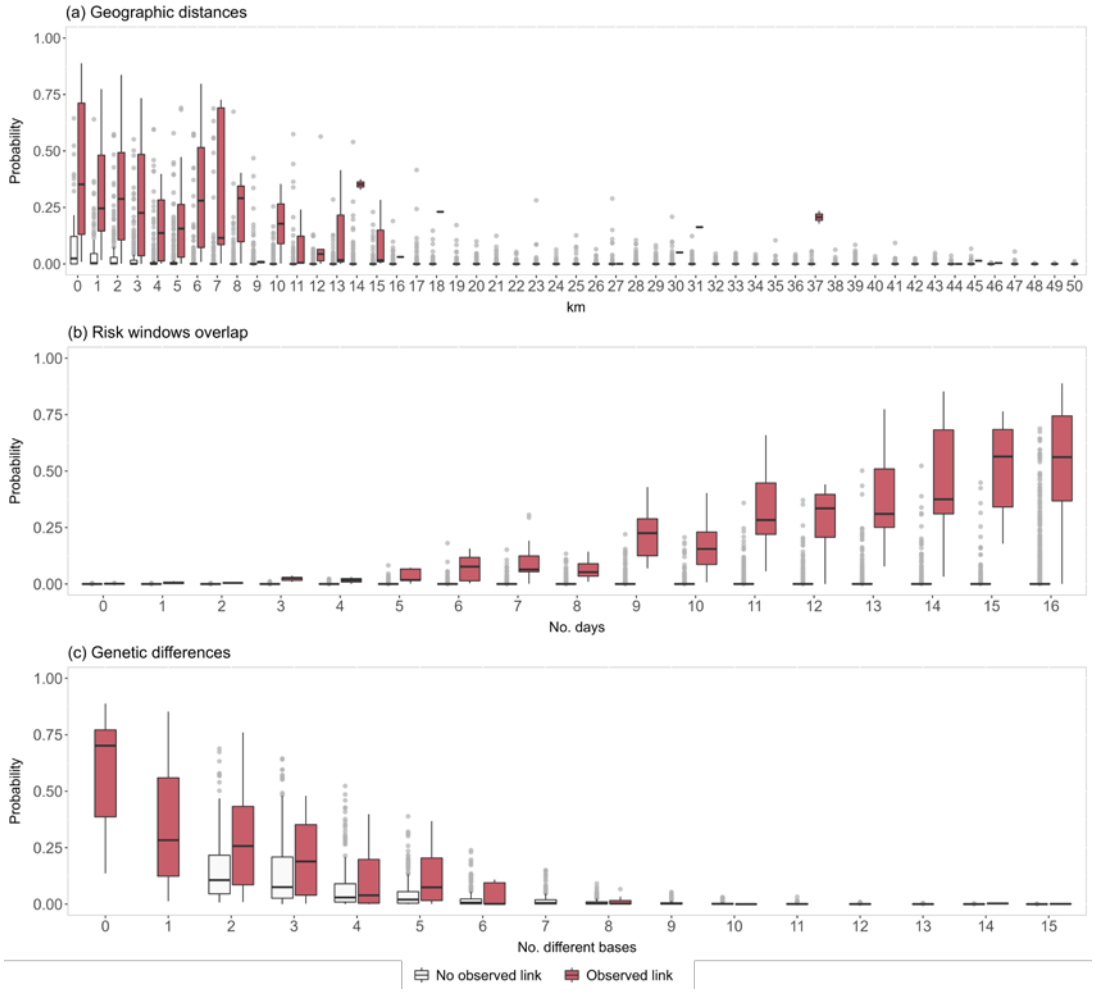

Figure 6. Graphical representation of the predicted probability distributions according to the (**a**) geographic distances, (**b**) risk time windows overlap, and (**c**) viral genetic differences. The x-axis scale for (**a**) and (**c**) was limited to 50 km and 15 bases, respectively, to improve the graph readability.

The model goodness of fit, intended as the tie prediction capability, was evaluated through the calculation of the ROC and PR curves and their related performance metrics (Figure 7). The area under the ROC estimated for the model was 0.957, indicating a very good performance compared to the estimated baseline of a random classifier. However, as the ROC curve can be affected by data imbalance (i.e. number of tie presences vs absences) and it is threshold-invariant (i.e. it does not take into account when the false positive and false negative errors are unequal), it is also important the evaluation of the PR curve, to avoid to overestimate the model performance. In this case, an area under the PR curve of 0.477 reflects a very good model performance, compared to the estimated baseline PR of 0.009 (which is indicative of a very imbalanced class distribution, i.e. the positive class prevalence is much lower compared to the negative one).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0404.v1

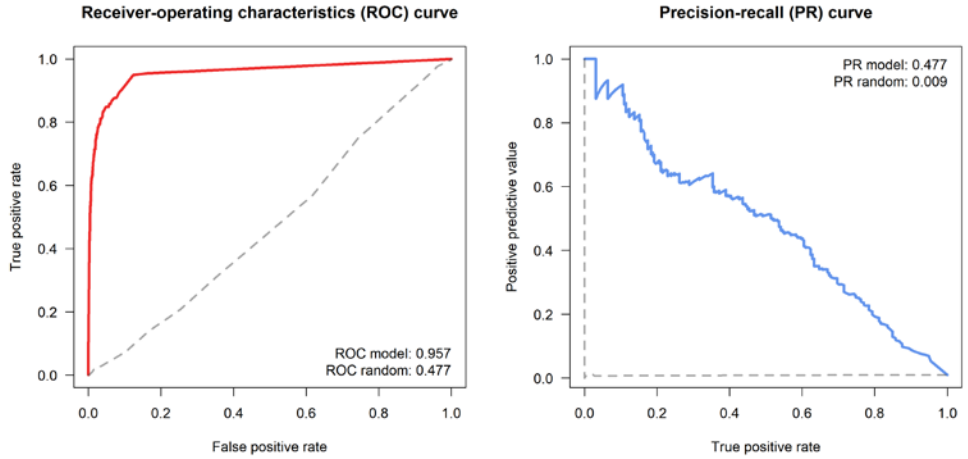

Figure 7. Tie prediction statistics including the ROC (left panel) and PR (right panel) curves. Solid lines: actual ROC/PR curves; dashed lines: baseline ROC/PR curves drawn for a random graph.

4. Discussion

The evolution of the epidemiological situation in 2021-2022 in Italy was comparable only to the HPAI H7N1 epidemic of 1999-2000, in terms of spread and number of cases confirmed in the DPPAs [37]. Similarly to the 2017-2018 HPAI H5N8 outbreak [4], there was no initial phase of introduction of low-pathogenicity virus from wild to domestic, and subsequent adaptation to poultry followed by an increase in pathogenicity (as in the 1999-2000 epidemic). However, in 2017 the virus had initially circulated on the fringes of the DPPA with repeated and distinct introductions from the wild, and only in a second wave it involved areas with a higher density of poultry farms, with limited lateral spread [4]. On the contrary, the situation in 2021 immediately appeared to be more severe, as the HPAI virus was likely introduced directly in the core of the area with the highest poultry density of Italy, contributing to the rapid spread of the disease into the poultry sector, firstly involving the Verona Province and then the neighbouring Provinces. Moreover, the H5N1 virus behaved differently from what was expected, as it strongly affected the broiler sector (normally considered a low-risk species of AI infection [38]) with very limited or completely absent symptomatology [39].

The analysis presented in this study showed that there were significant effects of some characteristics, related to the domestic farms and the study area, on the probability that the viruses isolated from two different outbreaks were genetically connected. These associations could be explained by the multiple occurrences of lateral spreads, or the presence of common sources of infection, which plausibly happened massively during the course of the 2021-2022 epidemic and that would explain such a large number of poultry farms infected in such a short period. The drivers associated with a higher risk of infection spread are the longer exposure to active outbreaks, the geographical proximity of poultry farms, and belonging to the same poultry company. The period in which farms remained exposed to the risk of infection (hypothesized to occur via at-risk contacts through vehicles and personnel, or by proximity) due to the persistence of active outbreaks was found to be one of the most relevant determinants of the 2021-2022 epidemic evolution. The model showed that for each additional day that a farm is exposed to contact with an active outbreak, the risk of transmission increased by just over 8%. This finding is particularly relevant because it strongly prompts the need to extinguish the outbreaks as fast as possible to reduce the risk of further transmission in an area characterized by a very high density of poultry premises. Hence the importance of adequate preparedness to respond to epidemic emergencies, such as the development of *ad hoc* protocols for more rapid and efficient extinction operations, including culling, removal of carcasses, and disinfections, but also measures that aim at reducing the population ex-

 doi:10.20944/preprints202212.0404.v1

posed to active outbreaks (e.g. through targeted pre-emptive culling). The observed inverse relationship between the geographic distance and the risk of disease transmission between farms is consistent with the results found in the literature [40–42] and reflects the importance of reducing the farms density. Although the ban on restocking was applied within the restricted areas during the epidemic, the reduction of the number of operating farms could be included in a package of preventive measures following an increased risk of AIV introduction in the DPPAs. Such reduction could be reached, for instance, through the temporary ban on restocking farms in the DPPA, or anticipating the slaughtering of animals approaching the end of the production cycle, in response to the outcome of early warning systems. The organization of poultry companies may play a major role in the evolution of the epidemiological situation, as differences in biosecurity standards and operating procedures may reflect different management capabilities, which in turn results in different risk levels of disease introduction. However, in the 2021-2022 epidemic, even large poultry companies were severely affected, indicating a massive environmental viral pressure that could have hampered the biosecurity measures already in place, and prompting for further structural or managerial improvements or changes, to avert a recurrence of such a severe lateral transmission. A critical point identified during the last epidemic was related to the difficulties in collecting information on vehicles and personnel traceability, especially for smaller poultry companies. In fact, the variable *same poultry company* can be considered as a proxy for this missing information, indicating the possibility of more frequent contacts occurring between farms belonging to the same company than those documented. This variable could also include those at-risk contacts that were not properly reported in the visitor logbooks and that can be referred to each specific company (e.g., maintenance staff for water/gas/electrical systems, delivery of drugs for specific treatments, etc.). Finally, it is also possible that some procedures are typical to individual poultry companies and/or productive types (e.g., companies specialized in organic production, free-range eggs production, free-range farming, etc.). These procedures could be linked to particular types of biosecurity management and, thus, to a different risk of AIV transmission. Similar results were also discussed in the study of Yoo and colleagues [40], in which the variables 'shared poultry integrators' and 'closer geographical distance' were identified as significant contributing factors for potential transmission routes among premises. The variables related to the number of at-risk contacts, the same reared species, and belonging to the same owner were not significantly associated with the genetic similarity structure observed in the network and, thus, to the probability of lateral spreads. Several other studies evaluated the impact of potentially at-risk movements on the probability of AI spread among farms, and most of them report a significant association between the two. For example, some studies conducted on the 2016–2017 H5N8 epidemic in France found that live-duck movements and indirect contacts via the transit of transport vehicles might have played a crucial role in the AI diffusion. According to Guinat et al. [42], live-duck movements might have played a crucial role in the geographical spread of AIV; although, as reported by Chakraborty et al. [43], viral spread was mostly related to short-range movements, i.e. between neighbouring or within the same administrative regions. The analysis presented in the work of Bauzile et al. [44] also showed that transmission events among farms were likely to have occurred through live animal movements and the transit of transport vehicles. In Yoo et al. [40] an extensive list of 18 different types of movements was found associated with a higher risk of transmission. The reason for the discrepancy between the present study and those reported in the literature could be related to some differences in the traceability systems and logistics organization among the countries. For example, the Italian traceability system of the feed trucks and personnel movements relies, at the present time, on the collection of non-standardized data directly from the poultry integrators that assemble transport notes and personal diaries information for reconstructing the taken routes. The reliability and accuracy of this data collection system are not comparable, for example, to the one used in the Republic of Korea, which is based on satellite tracking using global positioning system technology [40], or

 doi:10.20944/preprints202212.0404.v1

the official registration system that is mandatory for livestock movements. In addition, in this study only contacts related to the movements of feed trucks and poultry company personnel could be retrieved for the analysis, which represent certainly only a part of all the movements that normally take place from and into farms. The number of the identified at-risk contacts potentially associable with lateral spread events (n=58) was, indeed, markedly low considering the total number of the retrieved movements among farms (n=1470), representing 3.94% of the overall movements only, potentially explaining the non-significance of this variable in our analyses. Nevertheless, it is important to take into account the severe risk associated with the presence of even a few contacts, which should not occur in the course of an HPAI epidemic. The presence of these potentially problematic movements might indicate the presence of breaches in the organization of transport logistics, prompting a more suitable organization of vehicle movements to prevent multiple contacts of the same lorries with different farms. At the same time, it is recommended to strengthen the biosecurity measures at the farm level, and evaluate the actual effectiveness of the disinfection procedures to and from the farms.

Rearing the same species or having the same production type, as well as belonging to the same owner or family members were also not significantly associated with the lateral spread events. Interestingly, in Yoo et al. [40] the variable accounting for the similarities in the reared poultry species resulted related to a higher risk of transmission. This difference might be due to the effect of the *geographical distance* variable. In the Italian DPPAs, the probability that farms in close proximity breed the same species or have the same owner is very small, indicating that the geographic distances are likely more strongly correlated with the risk of viral spread between outbreaks, resulting in hedging the effect of the other two variables. Moreover, from the information collected during the 2021-2022 epidemic, the same feed truck was reported as supplying up to five different poultry production types on the same day.

Lastly, although the variable related to the number of different bases between two viral strains does not have a typical epidemiological value, its inclusion in the model analysis was useful to estimate, on average, the magnitude of the viral genomic diversity necessary to consider two strains arising from two separate origins, even if belonging to the same cluster. The probability of observing two viruses connected within the genetic network (and thus, a higher chance of direct transmission between farms, or a common source of infection), decreases by 13.7% for each additional different base. Taking into account the characteristics of the H5N1 virus circulating in Italy in 2021-2022, the results indicate that at least 7-8 nucleotide substitutions might be sufficient to discriminate outbreaks related to each other from those that were not connected.

The analysis of risk factors for potential lateral spread presented here emphasised the importance of prompt interventions following the outbreak confirmation, through its rapid extinction and possibly reducing the farms density at the local level. However, the high incidence of the disease, especially during the first weeks of the epidemic, led to a rapid saturation of the culling and disposal capacity. This resulted in numerous active outbreaks in close geographical proximity, a combination that led to a further massive spread of the virus. An ideal solution to address this critical issue would be the temporary reduction of the farms density, e.g. by restricting or banning the restocking of the most susceptible species in defined areas of the highest risk zones, in response to an adequately structured early warning system. Other hypotheses, regarding the effect of transports to the slaughterhouse of potentially infected animals in spreading the virus to the farms located in the proximity of the routes cannot be ruled out, and would require detailed data collection (e.g. information on the route taken by the trucks) and/or analytical methods based on complex mathematical simulations. The role played by the broiler species in the AI dynamics is a further crucial point to investigate, as they showed very little or no symptoms in course of infection [39], indicating that the time window to consider for the viral introduction should be carefully recalculated for the evaluation of the at-risk period of interest for the contact tracing activities.

 doi:10.20944/preprints202212.0404.v1

In addition to the assessment of the effect of implementing further control measures to counter the AI spread, especially in DPPAs, additional analyses are also necessary to gain knowledge on the rapidly changing epidemiology of AI over the last decade. These include, for example, adding temporal information to identify any significant fluctuation in the effect of the risk factors during the epidemic. Another interesting approach would include a Bayesian analysis to quantify and consider the effect of uncertainty associated with the potential epidemiological drivers, also reconstructing the missing information related to the median vectors nodes estimated during the generation of the genetic network. Mathematical models would be useful to investigate the AI spread, to calculate the actual reproductive number (Rt) during the course of the epidemic and to evaluate which control measures may have contributed the most to reducing Rt below the critical threshold necessary for spreading events. Finally, the results of the ERGM could be integrated with other mathematical models, e.g. based on an epidemic-tree approach, to reconstruct the most likely route of transmission occurred within the population and simulate how the modulation of specific measures (e.g. preventive culling within a defined radius, farms blocking, stamping out of at-risk contact farms, etc.) may affect the transmission dynamics.

 doi:10.20944/preprints202212.0404.v1

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Table S1: Accession numbers of the Italian viruses analyzed; Table S2: Information from the GISAID EpiFlu database on hemagglutinin segments of the European viruses; Figure S1: Phylogenetic network generated using the Median Joining method implemented in NETWORK 10.2.0., for the eight concatenated gene segments of non-reassortant H5N1 viruses (299 from Italy, 9 from Poland and 8 from Czech Republic); Circles: viral genotypes, the size is proportional to the number of viruses sharing the same genotype; the length of the branches is proportional to the number of nucleotide differences.

Author Contributions: Conceptualization, Diletta Fornasiero and Paolo Mulatti; Data curation, Diletta Fornasiero, Alice Fusaro and Bianca Zecchin; Formal analysis, Diletta Fornasiero, Alice Fusaro and Bianca Zecchin; Investigation, Diletta Fornasiero, Francesca Scolamacchia and Tiziano Dorotea; Methodology, Diletta Fornasiero; Supervision, Paolo Mulatti; Visualization, Diletta Fornasiero and Matteo Mazzucato; Writing – original draft, Diletta Fornasiero, Alice Fusaro and Bianca Zecchin; Writing – review & editing, Diletta Fornasiero, Alice Fusaro, Bianca Zecchin, Matteo Mazzucato, Francesca Scolamacchia, Grazia Manca, Calogero Terregino, Tiziano Dorotea and Paolo Mulatti. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data sharing not applicable due to privacy restrictions.

Acknowledgments: We are grateful to the Official Veterinarians who gave great support in the epidemiological investigation activities and data collection. We would like to thank the collaborators of the virology department for the intensive work of sample processing and laboratory analysis. We would also like to thank the poultry companies that collaborated in providing the contact tracing data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Adlhoch, C.; Fusaro, A.; Gonzales, J.L.; Kuiken, T.; Marangon, S.; Niqueux, É.; Staubach, C.; Terregino, C.; Guajardo, I.M.; Chuzhakina, K.; et al. Avian Influenza Overview June – September 2022. *EFSA J.* **2022**, *20*, doi:10.2903/j.efsa.2022.7597.
2. Mulatti, P.; Ferrè, N.; Marangon, S. Spatial Distribution of 2000-2007 Low Pathogenicity Avian Influenza Epidemics in Northern Italy. In *Pandemic Influenza Viruses: Science, Surveillance and Public Health*; Majumdar, S., Brenner, F., Huffman, J., McLean, R., Panah, A., Pietrobon, P., Keeler, S., Shive, S., Eds.; Pennsylvania Academy of Science: Easton, PA, USA., 2011 ISBN 978-0-945809-21-0.
3. Galbraith, C.A.; Jones, T.; Kirby, J.; Mundkur, T. *A Review of Migratory Bird Flyways and Priorities for Management*; UNEP/CMS Secretariat, Ed.; Secretariat of the Convention on the Conservation of Migratory Species of Wild Animals Recommended: Bonn (Germany), 2014; ISBN No. 27.
4. Mulatti, P.; Fusaro, A.; Scolamacchia, F.; Zecchin, B.; Azzolini, A.; Zamperin, G.; Terregino, C.; Cunial, G.; Monne, I.; Marangon, S. Integration of Genetic and Epidemiological Data to Infer H5N8 HPAI Virus Transmission Dynamics during the 2016-2017 Epidemic in Italy. *Sci. Rep.* **2018**, *8*, 18037, doi:10.1038/s41598-018-36892-1.
5. Mannelli, A.; Ferrè, N.; Marangon, S. Analysis of the 1999-2000 Highly Pathogenic Avian Influenza (H7N1) Epidemic in the Main Poultry-Production Area in Northern Italy. *Prev. Vet. Med.* **2006**, *73*, 273–285, doi:10.1016/j.prevetmed.2005.09.005.
6. Cecchinato, M.; Ceolin, C.; Busani, L.; Dalla Pozza, M.; Terregino, C.; Moreno, A.; Bonfanti, L.; Marangon, S. Low Pathogenicity Avian Influenza in Italy during 2007 and 2008: Epidemiology and Control. *Avian Dis.* **2010**, *54*, 323–328, doi:10.1637/8765-033109-Reg.1.
7. More, S.; Bicout, D.; Bøtner, A.; Butterworth, A.; Calistri, P.; Depner, K.; Edwards, S.; Garin-Bastuji, B.; Good, M.; Gortázar Schmidt, C.; et al. Avian Influenza. *EFSA J.* **2017**, *15*, doi:10.2903/j.efsa.2017.4991.
8. Lebarbenchon, C.; Chang, C.-M.; van der Werf, S.; Aubin, J.-T.; Kayser, Y.; Ballesteros, M.; Renaud, F.; Thomas, F.; Gauthier-Clerc, M. Influenza A Virus in Birds during Spring Migration in the Camargue, France. *J. Wildl. Dis.* **2007**, *43*, 789–793, doi:10.7589/0090-3558-43.4.789.
9. Zecchin, B.; Alice, F.; Barbierato, G.; Giussani, E.; Fornasiero, D.; Scolamacchia, F.; Mulatti, P.; Salviato, A.; Schivo, A.; Palumbo, E.; et al. Genetic Investigation of the HPAI H5N1 Viruses Responsible of HPAI Epidemic in Italy in 2021- 2022. In Proceedings of the ESVV 2022, 12th International Congress For Veterinary Virology; Ghent (Belgium), 2022; p. 125.
10. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780, doi:10.1093/molbev/mst010.
11. Bandelt, H.J.; Forster, P.; Rohl, A. Median-Joining Networks for Inferring Intraspecific Phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48, doi:10.1093/oxfordjournals.molbev.a026036.
12. R Core Team R: A Language and Environment for Statistical Computing 2022.
13. RStudio Team RStudio: Integrated Development Environment for R 2022.
14. Xu, W.; Berhane, Y.; Dubé, C.; Liang, B.; Pasick, J.; Vandomselaar, G.; Alexandersen, S. Epidemiological and Evolutionary Inference of the Transmission Network of the 2014 Highly Pathogenic Avian Influenza H5N2 Outbreak in British Columbia, Canada. *Sci. Rep.* **2016**, *6*, 1–15, doi:10.1038/srep30858.
15. Swayne, D.E. Overview of Avian Influenza. In *The Merck veterinary manual [online]*; Aiello, S., Moses, M., Eds.; Merck & Co. Inc: Kenilworth, NJ, 2016.
16. Nöremark, M.; Widgren, S. EpiContactTrace: An R-Package for Contact Tracing during Livestock Disease Outbreaks and for Risk-Based Surveillance. *BMC Vet. Res.* **2014**, *10*, 71, doi:10.1186/1746-6148-10-71.
17. Mulatti, P.; Kitron, U.; Jacquez, G.M.; Mannelli, A.; Marangon, S. Evaluation of the Risk of Neighbourhood Infection of H7N1 Highly Pathogenic Avian Influenza in Italy Using Q Statistic. *Prev. Vet. Med.* **2010**, *95*, 267–274,

doi:10.1016/j.prevetmed.2010.04.005.

18. Nickbakhsh, S.; Hall, M.D.; Dorigatti, I.; Lycett, S.J.; Mulatti, P.; Monne, I.; Fusaro, A.; Woolhouse, M.E.J.; Rambaut, A.; Kao, R.R. Modelling the Impact of Co-Circulating Low Pathogenic Avian Influenza Viruses on Epidemics of Highly Pathogenic Avian Influenza in Poultry. *Epidemics* **2016**, *17*, 27–34, doi:10.1016/j.epidem.2016.10.005.
19. Bivand, R, Keitt, T, Rowlingson, B. Rgdal: Bindings for the "Geospatial" Data Abstraction Library 2022.
20. Pebesma, E.J., R.S., B. Classes and Methods for Spatial Data in R 2005.
21. Bivand, Roger S., Pebesma, Edzer, Gomez-Rubio, V. *Applied Spatial Data Analysis with R*; Springer, N., Ed.; Second edi.; 2013;
22. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874, doi:10.1093/molbev/msw054.
23. Morris, M.; Handcock, M.S.; Hunter, D.R. Specification of Exponential-Family Random Graph Models: Terms and Computational Aspects. *J. Stat. Softw.* **2008**, *24*, doi:10.18637/jss.v024.i04.
24. Handcock, M.S.; Hunter, D.R.; Butts, C.T.; Goodreau, S.M.; Morris, M. Statnet : Software Tools for the Representation, Visualization, Analysis and Simulation of Network Data. *J. Stat. Softw.* **2008**, *24*, doi:10.18637/jss.v024.i01.
25. Hunter, D.R.; Handcock, M.S.; Butts, C.T.; Goodreau, S.M.; Morris, M. Ergm : A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. *J. Stat. Softw.* **2008**, *24*, doi:10.18637/jss.v024.i03.
26. Goodreau, S.M.; Handcock, M.S.; Hunter, D.R.; Butts, C.T.; Morris, M. A Statnet Tutorial. *J. Stat. Softw.* **2008**, *24*, doi:10.18637/jss.v024.i09.
27. Davis, J.; Goadrich, M. The Relationship between Precision-Recall and ROC Curves. In Proceedings of the 23rd international conference on Machine learning (ICML '06). Association for Computing Machinery.; New York, NY, USA, 2006.
28. Handcock, M.; Hunter, D.; Butts, C.; Goodreau, S.; Krivitsky, P.; Morris, M. Ergm: Fit, Simulate and Diagnose Exponential-Family Models for Networks. 2022.
29. Hunter, D.; Handcock, M.; Butts, C.; Goodreau, S.; Morris, M. Ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. *J. Stat. Softw.* **2008**, *24*, 1–29.
30. Krivitsky, P.N.; Hunter, D.R.; Morris, M.; Klumb, C. Ergm 4: New Features. **2021**.
31. Leifeld, P.; Cranmer, S.J.; Desmarais, B.A. Temporal Exponential Random Graph Models with Btergm : Estimation and Bootstrap Confidence Intervals. *J. Stat. Softw.* **2018**, *83*, doi:10.18637/jss.v083.i06.
32. Wickham, H. Ggplot2: Elegant Graphics for Data Analysis 2016.
33. Kassambara, A _ggpubr: "ggplot2" Based Publication Ready Plots_.
34. Csardi, G, Nepusz, T. The Igraph Software Package for Complex Network Research 2006.
35. Pedersen, T. _ggraph: An Implementation of Grammar of Graphics for Graphs and Networks_ 2021.
36. WHO (World Health Organization = Organisation mondiale de la Santé) Antigenic and Genetic Characteristics of Zoonotic Influenza A Viruses and Development of Candidate Vaccine Viruses for Pandemic Preparedness – Caractéristiques Antigéniques et Génétiques Des Virus Grippaux A Zoonotiques et Mise Au Point de Virus Vaccinaux. *Wkly. Epidemiol. Rec. = Relev. épidémiologique Hebd.* *95*, 525–539.
37. Capua, I.; Marangon, S. The Avian Influenza Epidemic in Italy, 1999-2000: A Review. *Avian Pathol.* **2000**, *29*, 289–294, doi:10.1080/03079450050118403.
38. Busani, L.; Valsecchi, M.G.; Rossi, E.; Toson, M.; Ferrè, N.; Pozza, M.D.; Marangon, S. Risk Factors for Highly Pathogenic H7N1 Avian Influenza Virus Infection in Poultry during the 1999-2000 Epidemic in Italy. *Vet. J.* **2009**, *181*, 171–177, doi:10.1016/j.tvjl.2008.02.013.
39. Gobbo, F.; Zanardello, C.; Bottinelli, M.; Budai, J.; Bruno, F.; De Nardi, R.; Patregnani, T.; Catania, S.; Terregino, C. Silent Infection of Highly Pathogenic Avian Influenza Virus (H5N1) Clade 2.3.4.4b in a Commercial Chicken Broiler Flock in Italy.

Viruses **2022**, *14*, 1600, doi:10.3390/v14081600.

40. Yoo, D.-S.; Chun, B. chul; Kim, Y.; Lee, K.-N.; Moon, O.-K. Dynamics of Inter-Farm Transmission of Highly Pathogenic Avian Influenza H5N6 Integrating Vehicle Movements and Phylogenetic Information. *Sci. Rep.* **2021**, *11*, 24163, doi:10.1038/s41598-021-03284-x.
41. Lee, H.; Suh, K.; Jung, N.; Lee, I.; Seo, I.; Moon, O.; Lee, J. Prediction of the Spread of Highly Pathogenic Avian Influenza Using a Multifactor Network: Part 2 – Comprehensive Network Analysis with Direct/Indirect Infection Route. *Biosyst. Eng.* **2014**, *118*, 115–127, doi:10.1016/j.biosystemseng.2013.11.009.
42. Guinat, C.; Durand, B.; Vergne, T.; Corre, T.; Rautureau, S.; Scoizec, A.; Lebouquin-Leneveu, S.; Guérin, J.-L.; Paul, M.C. Role of Live-Duck Movement Networks in Transmission of Avian Influenza, France, 2016–2017. *Emerg. Infect. Dis.* **2020**, *26*, 472–480, doi:10.3201/eid2603.190412.
43. Chakraborty, D.; Guinat, C.; Müller, N.F.; Briand, F.; Andraud, M.; Scoizec, A.; Lebouquin, S.; Niqueux, E.; Schmitz, A.; Grasland, B.; et al. Phylodynamic Analysis of the Highly Pathogenic Avian Influenza H5N8 Epidemic in France, 2016–2017. *Transbound. Emerg. Dis.* **2022**, *69*, doi:10.1111/tbed.14490.
44. Bauzile, B.; Sicard, G.; Guinat, C.; Andraud, M.; Rose, N.; Hammami, P.; Durand, B.; Paul, M.C.; Vergne, T. Unravelling Direct and Indirect Contact Patterns between Duck Farms in France and Their Association with the 2016–2017 Epidemic of Highly Pathogenic Avian Influenza (H5N8). *Prev. Vet. Med.* **2022**, *198*, 105548, doi:10.1016/j.prevetmed.2021.105548.