High pathogenicity avian influenza virus transmission is shaped by inter-specific social network structure.

*Jamie Dunning[1], Anna Gamża[2], Josh A. Firth[1,3], Adham Ashton-Butt[4,5], Rowland Kao[2,6], Ian A. Brown[7,8] and Alastair I. Ward[1].

[1]School of Biology, University of Leeds, Leeds, United Kingdom
[2]The Roslin Institute, University of Edinburgh, Edinburgh, United Kingdom
[3]Department of Biology, University of Oxford, Oxford, United Kingdom
[4]British Trust for Ornithology, Thetford, United Kingdom
[5]School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom
[6]School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
[7]International Reference Laboratory for Avian Influenza, Animal and Plant Health Agency (APHA-Weybridge), Addlestone, United Kingdom
[8]The Pirbright Institute, Pirbright, United Kingdom

*j.dunning@leeds.ac.uk

Summary: The continued emergence of zoonotic and epizootic diseases among wild animals is a recurrent and often intractable threat to human and animal health, threatening pandemic preparedness [1,2] . Predicting transmission pathways in multi-host systems, like wild birds, where ecological and behavioural diversity interact [3-5], is a particular challenge to surveillance [6-8]. As a result, study of disease in such multi-species systems is typically reactive to outbreaks, limiting our ability to predict spatial and temporal patterns of spread. Here, we integrated a fine-scale social network of wild bird co-occurrence with phylogenetic data from High Pathogenicity Avian Influenza Viruses (HPAIVs) to show that a zoonotic disease system can be tracked through a multi-species network, using social links as proxies for transmission probability [9]. Links in the network predicted genetic similarity between viral genomes, consistent with HPAIV transmission among species that co-occur in the wild. Species connection strength, independent of abundance, also predicted viral genetic similarity. Our study is the first empirical demonstration of social network structure predicting viral genetic similarity, a proxy for pathogen spread, through a multi-species system, with consequences for surveillance and prioritization of zoonotic and epizootic disease.

Main text: The repeated emergence of zoonotic and epizootic diseases of pandemic potential from wild animals poses a recurrent threat to human and animal health [1,2], with far-reaching consequences for global economic and food security [10]. For example Ebola, SARS-CoV-2 and avian influenza viruses all originate from communities of wild animals, often with little detection prior to outbreaks in humans or livestock [6–8]. However, in non-human animal communities, patterns of disease transmission are highly variable, and cannot be solely explained by pathogen virulence. Instead, just as human mobility and interaction shapes the trajectory of an epidemic [11,12], animal social network structure, shaped by the social [3,9,13-15] and physical environment [16,17] influences disease propagation through multi-species networks of wild animals.

Wild animal social systems are complex and dynamic, shaped by demography, behavioural preferences, and physical environment [4,13,18-22]. This complexity is magnified in multi-host systems where many diverse species with varied behavioural traits and spatial preferences interact [3,4]. It is this inherent complexity that drives network structure, providing the pathways over which disease may spread through these ecological systems [3,4,9]. For example, during the 2021/22 high pathogenicity avian influenza (HPAI H5N1 clade 2.3.4.4b) outbreak in European seabirds, disease presence in isolated colonies prompted birds to abandon infected breeding colonies, joining others and increasing between-colony connectivity and driving further spread [23]. This host expansion had devastating consequences for the conservation of wild birds [24,25], driving incursion of HPAIV into new continents [26] and novel mammalian hosts [8,27–29]. Perhaps due to the inherent complexity in understanding the drivers of disease through multi-species networks, research into animal disease systems has typically focused on specific traits (e.g., shared habitat, dietary preference, migratory phenology), represented by network nodes [18–21, 30,31], rather than between-species social structure, represented by the interconnecting edges between those nodes. However, the mechanism by which host traits, or environmental components [16,17], influence transmission is often by shaping social network structures - forcing potential hosts together or driving them apart. This spatial and temporal clustering either facilitates or constrains non-vector borne pathogen transmission through social links.

Understanding the patterns of disease spread through multi-host networks, and the traits shaping those networks, represents an important knowledge gap for informing zoonotic and epizootic disease surveillance and prevention [4,9,13-15,18]. Both aspects have been challenging to test empirically, leaving evaluation of multi-host network disease dynamics reliant on theoretical models [32–34]. Furthermore, surveillance and predictive modeling remain limited by reactive sample collection, processing constraints, and taxonomic sampling biases [35-36], and social network approaches tend to only be applied to simple and single-species, indirect transmission systems [18,37–40]. The use of social network models to measure transmission is yet to be validated empirically, and their

application to more complex, multi-species transmission systems remains a critical knowledge gap [4,9]. There is therefore a pressing need for predictive approaches to identify the species and pathways driving zoonotic and epizootic disease spread, in order to inform risk-based policy interventions.

Here, we used HPAIVs, a serious public health threat with trans-boundary transmission between hundreds of interacting hosts and multiple transmission pathways [5,8,] as a data-rich study system to test the relationship between social-linkage and indirect disease transmission. We combined social network analysis with viral phylogenetics and Bayesian multilevel modelling to test our prediction that social links would correlate with HPAIV viral genome similarity (as a proxy for transmission probability) in multi-species networks of wild birds. HPAI viral phylodynamics, a synthesis of phylogenetics and population dynamics, have been used to match HPAI H5N1 infection in poultry with an identical subtype present in sediment on nearby wetlands [41], implying landscape transmission, but without identified transmission pathways. We therefore used fine-scale species co-occurrence to build multi-species social networks to evaluate relationships between network structure and genomic relatedness of HPAI viruses sampled from hosts present in those networks. We demonstrated the relationship between bird species co-occurrence patterns and HPAI viral genome similarity at landscape scales. Our study provides the first empirical evidence that multi-species network structure drives pathogen spread in complex communities of wild animals.

Aligning social networks with HPAI viral phylogenetics

We constructed a multi-species social network from five years of Wetland Bird Survey (WeBS) data from the British Trust for Ornithology (BTO), a database describing the spatial spread of wetland birds at the landscape scale. We sampled communities of wild birds using multiple sites within a 50 km buffer of a farm cluster with recurrent HPAI outbreaks (nine outbreaks in the farm cluster, and 69 identified within the buffer since 2018) in the Great Britain, over five years, which we combined into a single global network (Supplementary Table 1a). From this global network we inferred 8,162 dyads (representing a species co-occurrence event) between 159 focal species nodes. Network structure (species composition and the relative number of interconnecting edges between species pairs was consistent between survey years (Supplementary Table 1b). We also accessed 715 full viral genome sequences of HPAIV H5N1 viruses collected in Great Britain, initially from 59 host species from the Global Initiative on Sharing All Influenza Data (GISAID), data repository. After filtering to only include species pairs which were also present in our viral phylogeny, our data included 763 instances of 172 unique combinations of 19 species (see full methods for details).

For each species pair we extracted two measures of edge centrality (a measure representing an individual species' contribution to global network structure) both of which we predicted were associated with transmission probability: 1) frequency of species co-occurrence in the landscape (dyadic edge-weight, mean = 166.81, SD = 285.87); and 2) species-level connection strength (Simple Ratio Index, SRI, mean = 0.005, SD = 0.005). SRI represents the connectedness of a species pair relative to opportunity for association, and regardless of abundance. A greater SRI is therefore a measure of dyadic clustering probability in the social network [42]. We calculated each measure from the global network, before filtering and scaling against our viral data. Each instance also comprised three measures from the viral phylogeny: conservative genetic similarity (maximum genetic distance between viral sequences, mean = 0.05; SD = 0.02), distance between samples in meters (mean geometric distance, mean = 279121.8, SD = 103676.6) and distance between samples in days (temporal distance, mean = 236.72, SD = 115.93).

To test the relationship between network centrality measures and viral genetic distance we ran two Bayesian Generalized Linear Mixed Models (GLMMs), controlling for spatial and temporal distances between sampling events. To account for repeated sampling species identities were included as multi-membership random effects. To sample from complex, high-dimensional probability distributions we used Hamiltonian Monte Carlo sampling. We built separate models to evaluate associations between genetic similarity as a proxy for transmission probability and two centrality measures, given as fixed effects, co-occurrence frequency and species connection strength. Because our study focused on transmission pathways, rather than host susceptibility to infection we did not constrain the phylogenetic relationship between hosts in our models. We accounted for species specific variance on viral genetic similarity independently of their dyad by linking a 'multi-membership' species variable within our random effects (see methods).

Network centrality predicts transmission pathways

We found that viral genetic similarity was negatively associated with greater frequency of co-occurrence in the landscape (estimate = -0.10; 95% CI = -0.10 to -0.05), consistent with species that frequently co-occurred in the landscape having more closely related viral sequences than those that co-occurred less frequently. The distance between sampling events in space (estimate = 0.08; 95% CI = 0.02 to 0.14) and in time (estimate = 0.40; 95% CI = 0.36 to 0.43) were both associated with less similar viral genomes. Using a multi-membership random effect structure, we found high variation in individual species contributions to the model, suggesting that species identity plays an important role in transmission probability (estimate = 1.03; 95% CI = 0.75 to 1.46)

Similarly to co-occurrence, we found that viral genetic similarity was also negatively associated with greater species connection strength (estimate = -0.14; 95% CI = -0.19 to -0.09). This suggests that not only are species that co-occur more frequently in the landscape associated with more similar viral genomes, but species pairs that cluster consistently, irrespective of abundance, are too. The distance between sampling events in space (estimate = 0.07; 95% CI = 0.01 to 0.12) and time (estimate = 0.41; 95% CI = 0.37 to 0.45) both were, again, associated with genetic similarity, confirming the assumption that sequences sampled closer in time and space were more similar to one another. Again the species identity explained significant variation in the model (estimate = 1.11; 95% CI = 0.80 to 1.58). Year, as a random effect, did not explain significant variance in either model, supporting the assumption that network structures remained broadly consistent across years.

To test if our associations could be explained by social network structure alone, we also rewired the network and repeated our models over a series of permutations while maintaining the degree distribution for each species node**.** In both the co-occurrence and species connection strength models, the observed effect estimates fell outside the 95% confidence intervals of the permuted (null) distributions (co-occurrence: 95% CI = -0.03 to 0.03; and connection strength: 95% CI = -0.015 to 0.019) indicating that the observed associations were stronger than expected by chance alone, given the network structure. Spatial and temporal distance remained significant predictors of viral genetic similarity in both the observed and permuted models, consistent with expectations that sequences sampled closer in space and/or time were more similar (Fig. 1b).

Conclusion and Discussion:

By aligning a multi-species network of wild birds with a detailed viral phylogeny, 763 rows of 172 unique combinations between 19 wild bird species, we demonstrated that inter-specific social network structure significantly predicted genetic distance between HPAI samples, a proxy for disease transmission probability, through a multi-species community of wild animals. Our study represents the first empirical application of social network analysis to predict viral transmission pathways across multiple wild species in this way.

We used two large, open-data repositories underscoring their value to research into wildlife disease surveillance. Although we used WeBS data from the BTO, recent studies have also used similar open-data to model multi-species networks [30,43,44]. For example, Vickers et al. [30] used the eBird global repository of bird survey data [43] to link abundance of wild birds in the landscape with HPAI incursion into poultry, without specifying transmission pathways through a social network. In a social network context,

previous studies have modeled pathogen transmission within simple or single species transmission systems [37–40], or adopted a trait-based approach multi-species systems, requiring parameterization of network nodes [18]. Our results suggest, for the first time, that transmission of a disease of pandemic potential is also measurable through a multi-species network of wild animals, using social-links (edges) as a proxy for transmission. This result is also predicted in the context of host or environment traits, which drive pathogen transmission because they influence the shape of the social network, facilitating transmission by driving potential hosts into close proximity [23,30,31,38-40], facilitating social linkage. Our study demonstrates an empirical baseline for further theoretical predictive modelling of multi-species host-pathogen networks, by using the social links as a proxy for transmission, and could help to fill gaps in phylodynamic modelling, such as those resulting from lack of virus samples from species with infected individuals. Network stability over time, which we have also demonstrated here, and speed of survey data availability will be crucial to apply the concepts we have demonstrated here to predicting future outbreaks, using network edges as a proxy for transmission probability.

Although our study represents a complex, multi-species transmission system, we did not investigate the role of specific host (node based-)traits in shaping contagion, although much previous work has done so[23,30,31,38-40], without defining specific pathways of transmission. Host traits likely drive transmission through social links, by, for example, influencing immuno-competence [45], or behaviours triggered by the presence of disease in the community [23,] or indirectly, by shaping the social network through which disease is transferred, irrespective of the presence of disease [13-15]. It is also thertefore likely that host traits drive disease transmission at different scales, for example long-distance migration has been associated with intercontinental HPAI transmission [33-34], whereas local outbreaks in poultry have been associated with short-distance dispersal in sedentary birds [30-31,41]. Our study, at the landscape scale, demonstrates how these two scales are linked in the landscape - from intercontinental incursion to outbreaks in local farmyards. Evaluation of the role of traits in HPAI transmission may further support identification of putative bridge taxa; those critical intermediary hosts between waterfowl and poultry.

We note that a multi-species network of wild animals may include multiple mechanisms of transmission, and our study is appropriate only when applied to indirect transmission systems. For example, our HPAI viral genome database also included samples taken from predatory birds, raptors, that had close viral genetic distances to colonial seabirds [31]. However, those links were likely trophic, rather than social, probably representing ingestion of infected prey, rather than indirect transmission. We therefore suggest that future studies should define network edges by the mechanism of transmission, for

example social links, mobility, or predator-prey interactions, to evaluate patterns of transmission.

While we used 763 pairwise viral genome samples, and 172 unique dyads of 19 host species, we were limited by the availability of genome data for a large number of potential host species. The inferred social network, by comparison, comprised 8,162 edges between 159 species nodes. Although the sub-graph (blue, within the graph given in Fig.1a) offers proof of concept, we suggest proactive sampling of viruses from a sufficient number of a more comprehensive range of species would be required to map past transmission pathways of pathogen spread in order to enable greater confidence in prediction at multiple spatial and temporal scales. This may be particularly important for disease systems, such as HPAI, where species responsible for bridging the transmission gap between primary wild hosts (eg. anseriformes) and livestock are unidentified.

Combined, our results support the hypothesis that species co-occurrence in a landscape is linked with pathogen genetic similarity, implying predictable transmission through a complex, multi-species wildlife social network. Our social network approach allied with increasing availability of open data on wild species occurrence and relative abundance offers significant potential for evaluating and predicting pathogen transmission pathways among wildlife communities, and hence opportunities to better inform surveillance and epizootic disease control policy.

Methods:

To test the hypothesis that host species' social linkages predict viral genetic distance across multi-species networks of wild birds, we combined social network analysis with phylodynamic and Bayesian mixed modelling approaches. We tested if our results differed from the expected network, without social structure, by running a series of permutations that mirrored our observed data models.

Survey effort and Networks: To build species co-occurrence networks, we accessed five years of data from the British Trust for Ornithology Wetland Bird Survey (WeBS) [46], 2018 - 2022 for two overlapping regions of East Yorkshire, Great Britain, each defined by a 50km circular buffer around two recurrent HPAI outbreak sites. Each year represented a number of sampling events (surveys) within the non-breeding season (October to March), combined into a global network representing co-occurrence in the landscape.

From these networks we extracted two established dyadic (pairwise - between two species) measures of association. To measure the extent to which two species co-occur

in the landscape, and therefore, may be expected to share virus more easily, we used pairwise edge weight, which is the absolute frequency of co-occurrence in the network. Then, to capture the extent to which a species pair is associated in space and time, accounting for absolute co-occurrence frequency relative to the opportunity for co-occurrence, we also used the Simple Ratio Index (SRI;[42]), which measures between 0 (no overlap between a species) and 1 (where one species is recorded so is the other, without exception). Although both measure interaction between species, pairwise weight and SRI represent different aspects of interaction in the landscape. Where a high pairwise weight denotes more frequent co-occurrence in the landscape, a higher SRI value represents stronger relative association, regardless of abundance. For example, in the WeBS data, the highest SRI values were between relatively scarce species whose habitat overlaps, for example Dipper *Cinclus cinclus* and Grey Wagtail *Motacilla cinerea*, which both favor fast flowing upland streams. Although they were both rare in the landscape, they were almost always recorded at the same sites. Finally, we restricted the network used in our models to only include the species also present in the viral phylogenetic data (see below), while weights and SRI measures were calculated and carried over from the global social network.

Network similarity: We tested the assumption that our networks are comparable between years, that is that the dyads that make up the general network structure were similar across our five years of data collection. We did this using two measures designed to compare non-uniform networks, in this case, with variation in the species recorded (the network actors, nodes) between years: 1) Jaccard similarity describes the proportion of shared dyads between two networks, regardless of weight; and, 2) a Spearman rank correlation test to measure similarity of dyadic edge weights between species pairs present in both networks. In both cases 1 denotes identical network structure, and 0 denotes complete disimialirty.

Viral genetic similarity: We accessed 715 full genome sequences of H5N1 viruses collected in Great Britain, initially from 59 host species from the Global Initiative on Sharing All Influenza Data (GISAID). To define genetic distance between viral genome sequences, we classified them by reassortment profile [28]. We aligned the sequences using MAFFT v7.511[47], trimmed them to coding regions, removed insertions present in fewer than 10% of sequences and rejected sequences with nucleotide coverage lower than 80% for all segments. We identified clusters for each of the eight gene segments using nearest-neighbor hierarchical clustering of nucleotide distances in the bioseq package in R [48,49]. We used a 0.5% distance threshold for the Polymerase Basic (PB)1, PB2, and Polymerase Acidic gene segments, and 1% for all others [31]. We aligned each sequence with host information, after filtering out mammals, domestic birds and taxonomically vague entries (none of which are recorded as part of WeBS surveys), as well as non- H5N1 strains (H5N5 and H5N8). We recorded the date associated with

each sampling event and a location, provided by the Animal and Plant Health Agency . We compared each sequence with all others in the database to quantify a genetic distance between pairwise viral genomes using the Tamura and Nei 1993 model [50] using the APE package in R [49, 51] that calculates genetic distance as a proportion that differ between each pair of sequence, accounting for difference in rates for both kinds of transition and transversion, as well as measures of geometric (space between sampling events) and temporal (time between sampling events) distance. For more details on how these data were generated see Gamża et al. [31]. Finally, we further filtered these data to only include species pairs also present in at least one of our five social networks, leaving 763 pairwise samples between 19 wild host species.

Data structure: To test our predictions that high between-species co-occurrence, or greater association between a species pair, should result in a greater chance of viral transmission, we built and extracted edge-based centrality metrics from five years of BTO WeBS data and modeled these against similarity of host viral genomes (genetic distance). Virus samples were taken from relatively few wild bird host species (relative to the number in the social networks) so our models only contained species present in both data-sets.

Pairwise weight: We modeled pairwise weight, the mean of geometric distance between viral pairwise sample locations and the mean of pairwise temporal distance (time elapsed between sampling events) as fixed effects. We z-transformed all fixed effects to better align scales between other variables in the model. We used the mean (rather than the maximum) of geometric- and temporal distance to account for wide variation in sampling and to reduce marginal collinearity. Because most species appeared across many dyads, we accounted for species specific effects on genetic distance independently of their dyad (i.e. species A + species B, rather than speciesAB) by including species A and species B as a linked ‘multi-membership’ variable inside our random effects. This approach also avoided over-dispersion.

SRI: We repeated the first model, using the same model structure but with SRI instead of pairwise weight, as a fixed effect. We used the log-odds (using a logit link function) of SRI to better accommodate a linear relationship between SRI and other predictors in the model.

Model fit: To test the relationship between wild bird social network structure and maximum genetic distance between HPAI viruses (the response variable), and to accommodate multi-membership sampling in our species dyads, we ran a series of Generalized Linear Mixed Models (GLMMs) using Bayesian Markov Chain Monte Carlo (MCMC) simulations with the Hamiltonian Monte Carlo (HMC) and No-U-Turn Sampler (NUTS) algorithms in the brms package in R [49,52]. All models used the maximum genetic distance (least similar) between pairwise viral genomes, as a conservative proxy for

transmission between hosts, as the response variable. Otherwise, our models only differed only by the first predictor (fixed effect) variable, which was either pairwise weight or SRI. For all models, we used weakly informative priors, and ran each over 4 chains of 4,000 iterations, with a warm-up of 1,000 iterations and a thinning interval of 1. We used Effective Sample Sizes of >5000 in the posterior and tail distributions to maintain sampling efficiency and a scale reduction factor (Ř) values of 1 to ensure convergence. We also manually checked posterior trace plots to ensure model mixing. The fixed effects were considered statistically significant when the 95% credible interval (CI) excluded zero.

Permutations: In order to test that our results differed from the expected distribution without social structure in the network, we repeated our models with a series of network permutations. We did this by 're-wiring' the observed network, shuffling the edges between species pairs while maintaining the degree distribution for each species node [53]. We rewired networks 1000 times within years, before combining them into a global network within each permutation cycle. Edges were rewired between species nodes, breaking the structure of the observed network, but, because some species are much more frequent in our data -and therefore our networks have some strong ties- we sampled edges from the empirical distribution of edge weights [53]. This controlled for the chance that some effects are the result of some species having more interactions over all, rather than who those interactions were with. We maintained the structure of our original models and ran the permutation over 100 cycles. We inferred significance, that the effect of social structure on genetic distance in our observed models was stronger than expected under a system of random mixing, where our observed credible intervals fell outside of the 95% confidence intervals of the permuted estimate.

***

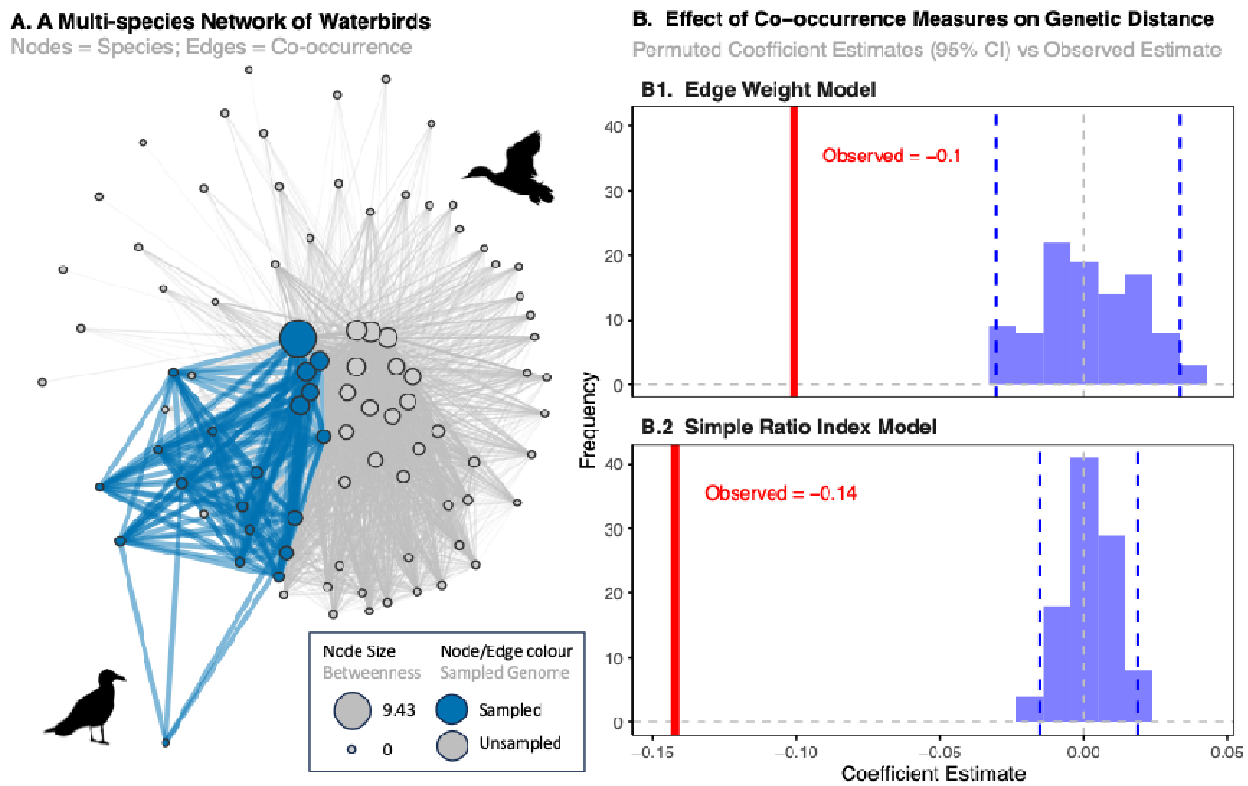
Author Contributions: JD was involved in all elements of this study - design, data acquisition, formal analysis, writing and editing; JAF was involved in all aspects of this study, but primarily social network and model specification, and project supervision; AG contributed viral genetic data, writing and editing; ABB contributed to study design and editing; RK and IAB contributed to study design, data acquisition and editing; AIW was involved in all elements of this study and project supervision.

Competing interests: The authors declare no competing interests

Acknowledgements: Data were provided by WeBS, a partnership jointly funded by the British Trust for Ornithology, Royal Society for the Protection of Birds and Joint Nature Conservation Committee with fieldwork conducted by volunteers and previous support from Wildfowl and Wetlands Trust. We thank Amanda Bretman, Bill Kunin, Talya Hackett, Elizabeth Duncan, Sam Lycett and Will Harvey for comments on this manuscript or the data therein. Our project was supported by Flu-Trial Map Avian: JD,

IAB and AIW under grant no. BB/Y007042/1; AG and RK were funded by the Strategic Program grant to Roslin Institute (grant no. BB/P013740/1), Flu-MAP project (grant no. BB/X006123/1), Flu-TrailMap (grant number BB/Y007271/1, BB/Y007298/1) and FluTrailMap- One Health (MR/Y03368X/1). JAF also acknowledges funding from NERC (NE/V013483/2) and WildAI (C-2023-00057).

****



Network centrality predicts transmission pathways. Figure 1a. Of the 159 species in the co-occurrence network, 19 were also associated with a genome for high pathogenicity avian influenza clade 2.3.4.4b, of which we had access to 763 samples. The colour of the nodes and edges denotes that a species pair had viral genome data available, blue, or not, grey. The node size denotes normalised betweenness value, where 9.43, sits on the greatest number of shortest paths between all of dyads in the network, in this case Mute Swan *Cygnus olor*, and 0, on no of the shortest paths (here including Northern Lapwing *Vanellus Vanellus*, Little egret *Egretta garzetta*, and Northern pintail *Anas acuta*). **1b.** The expected distribution of permuted estimates in blue, with 95% confidence intervals; blue dotted lines, plotted against the observed estimate, red vertical line, for two measures of network structure; edge weight and Simple Ratio Index. In both cases, the observed estimate falls outside the permuted distribution, suggesting that closer social proximity predicts greater similarity in HPAI

viral genome samples. This indicates that our results differ from what is expected when social structure is removed from the network.

Supplementary Table 1a. Descriptive statistics for within-year social networks showing the number of species involved in the network (network nodes); the number of edges between species pairs, representing co-occurrence within a recording event, and, the number of recording events. A recording event represents a single survey visit in time and space. **1b.** Network similarity over years: **Upper right matrix** showing Jaccard similarity, the number of shared dyads between years irrespective of edge weight; and, **Lower left matrix**, Spearman rank correlation coefficients between dyadic edge weight between shared dyads, between years. In both cases 1 = perfect correlation and 0 = no relationship.

a. Network structure					b. Network similarity between years				
	Species	Edges	Events			2018	2019	2020	2021	2022
2018	115	4132	354		2018	-	0.67	0.65	0.56	0.56
2019	131	5321	729		2019	0.89	-	0.71	0.64	0.67
2020	127	4938	717		2020	0.86	0.91	-	0.66	0.63
2021	134	6378	782		2021	0.83	0.87	0.86	-	0.73
2022	137	6154	748		2022	0.79	0.85	0.85	0.91	-

***

References

1. Kruse, H., Kirkemo, A.-M. & Handeland, K. Wildlife as source of zoonotic infections. *Emerg. Infect. Dis.* **10**, 2067 (2004).
2. Jones, K. E. et al. Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008).

3. Bansal, S., Grenfell, B. T. & Meyers, L. A. When individual behaviour matters: Homogeneous and network models in epidemiology. *J. R. Soc. Interface* **4**, 879–891 (2007).
4. Buhnerkempe, M. G. et al. Eight challenges in modelling disease ecology in multi-host, multi-agent systems. *Epidemics* **10**, 26–30 (2015).
5. Reid, S. M. et al. A multi-species, multi-pathogen avian viral disease outbreak event: Investigating potential for virus transmission at the wild bird–poultry interface. *Emerg. Microbes Infect.* **13**, 2348521 (2024).
6. Marí Saéz, A. et al. Investigating the zoonotic origin of the west african ebola epidemic. *EMBO Mol. Med.* **7**, 17–23 (2015).
7. Holmes, E. C. COVID-19—lessons for zoonotic disease. *Science* **375**, 1114–1115 (2022).
8. Xie, R. et al. The episodic resurgence of highly pathogenic avian influenza H5 virus. *Nature* **622**, 810–817 (2023).
9. Dunning, J., Firth, J. A. & Ward, A. I. The spread of HPAI is a social network problem. *PLoS Pathog.* (2024).
10. Bernstein, A. S. et al. The costs and benefits of primary prevention of zoonotic pandemics. *Sci. Adv.* **8**, eabl4183 (2022).
11. Firth, J. A. et al. Using a real-world network to model localized COVID-19 control strategies. *Nat. Med.* **26**, 1616–1622 (2020).
12. Kostandova, N. et al. A systematic review of using population-level human mobility data to understand SARS-CoV-2 transmission. *Nat. Commun.* **15**, 1–12 (2024).
13. Sah, P., Mann, J. & Bansal, S. Disease implications of animal social network structure: A synthesis across social systems. *J. Anim. Ecol.* **87**, 546–558 (2018).
14. Albery, G. F., Kirkpatrick, L., Firth, J. A. & Bansal, S. Unifying spatial and social network analysis in disease ecology. *J. Anim. Ecol.* **90**, 45–61 (2021).
15. Kulahci, I. G. & Quinn, J. L. Dynamic relationships between information transmission and social connections. *Trends Ecol. Evol.* **34**, 545–554 (2019).
16. King, K. C., Hall, M. D. & Wolinska, J. Infectious disease ecology and evolution in a changing world. *Philos. Trans. R. Soc. B* **378**, 20220002 (2023).
17. Carlson, C. J. et al. Climate change increases cross-species viral transmission risk. *Nature* **607**, 555–562 (2022).
18. Luis, A. D. et al. Network analysis of host–virus communities in bats and rodents reveals determinants of cross-species transmission. *Ecol. Lett.* **18**, 1153–1162 (2015).
19. Yin, S. et al. Functional traits explain waterbirds' host status, subtype richness, and community-level infection risk for avian influenza. *Ecol. Lett.* **26**, 1780–1791 (2023).

20. Olival, K. J. et al. Host and viral traits predict zoonotic spillover from mammals. *Nature* **546**, 646–650 (2017).
21. Hill, N. J. et al. Ecological divergence of wild birds drives avian influenza spillover and global spread. *PLoS Pathog.* **18**, (2022).
22. Firth, J. A., Sheldon, B. C. & Brent, L. J. Indirectly connected: Simple social differences can explain the causes and apparent consequences of complex social network positions. *Proc. R. Soc. B* **284**, 20171939 (2017).
23. Jeglinski, J. W. E. et al. HPAIV outbreak triggers short-term colony connectivity in a seabird metapopulation. *Sci. Rep.* **14**, (2024).
24. Falchieri, M. et al. Shift in HPAI infection dynamics causes significant losses in seabird populations across great britain. *Vet. Rec.* **191**, 294–296 (2022).
25. Tremlett, C. J., Cleasby, I. R., Bolton, M. & Wilson, L. J. Declines in UK breeding populations of seabird species of conservation concern following the outbreak of high pathogenicity avian influenza (HPAI) in 2021–2022. *Bird Study* 1–18 (2025).
26. Kandeil, A. M. et al. Rapid evolution of A(H5N1) influenza viruses after intercontinental spread to north america. *Nat. Commun.* **14**, (2023).
27. Sidik, S. M. Bird flu outbreak in mink sparks concern about spread in people. *Nature* **614**, 17 (2023).
28. Plaza, P. I., Gamarra-Toledo, V., Euguí, J. R. & Lambertucci, S. A. Recent changes in patterns of mammal infection with highly pathogenic avian influenza A (H5N1) virus worldwide. *Emerg. Infect. Dis.* **30**, 444 (2024).
29. Sah, R. et al. Concerns on H5N1 avian influenza given the outbreak in US dairy cattle. *Lancet Reg. Health Am.* **35**, (2024).
30. Vickers, S. H., Raghwani, J., Banyard, A. C. & Brown, I. H. Utilising citizen science data to rapidly assess potential wild bridging hosts and reservoirs of infection: Avian influenza outbreaks in great britain. ***bioRxiv*** (2024).
31. Gamża, A., Lycett, S., Sanchez, A. & Kao, R. Using sequence data to study spatial scales of interactions driving spread of highly pathogenic avian influenza in great britain. ***arXiv*** preprint arXiv:2411.10424 (2024).
32. Albery, G. F., Eskew, E. A., Ross, N. & Olival, K. J. Predicting the global mammalian viral sharing network using phylogeography. *Nat. Commun.* **11**, 2260 (2020).
33. McDuie, F. et al. Mitigating risk: Predicting H5N1 avian influenza spread with an empirical model of bird movement. *Transbound. Emerg. Dis.* **2024**, 5525298 (2024).
34. Tian, H. et al. Avian influenza H5N1 viral and bird migration networks in asia. *Proc. Natl. Acad. Sci. USA* **112**, 172–177 (2015).
35. Wade, D. et al. High pathogenicity avian influenza: Targeted active surveillance of wild birds to enable early detection of emerging disease threats. *Epidemiol. Infect.* **151**, (2022).

36. Atkinson, P. W. et al. Evaluating the use of carcass and testing data to assess the high pathogenicity avian influenza (HPAI) related mortality in wild birds in the united kingdom and crown dependencies between 2021–2023. *Bird Study* **0**, 1–19 (2025).
37. Gómez, J. M., Nunn, C. L. & Verdú, M. Centrality in primate–parasite networks reveals the potential for the transmission of emerging infectious diseases to humans. *Proc. Natl. Acad. Sci. USA* **110**, 7738–7741 (2013).
38. Hamede, R. K., Bashford, J., McCallum, H. & Jones, M. Contact networks in a wild tasmanian devil (Sarcophilus harrisii) population: Using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease. *Ecol. Lett.* **12**, 1147–1157 (2009).
39. Wilson-Aggarwal, J. K. et al. High-resolution contact networks of free-ranging domestic dogs Canis familiaris and implications for transmission of infection. *PLoS Negl. Trop. Dis.* **13**, e0007565 (2019).
40. Proboste, T. et al. Quantifying feral pig interactions to inform disease transmission networks. *bioRxiv* 2024–08 (2024).
41. Giacinti, J. A. et al. Transmission dynamics of highly pathogenic avian influenza virus at the wildlife-poultry-environmental interface: A case study. *One Health* **19**, 100932 (2024).
42. Farine, D. R. & Whitehead, H. Constructing, conducting and interpreting animal social network analysis. *J. Anim. Ecol.* **84**, 1144–1163 (2015).
43. Sullivan, B. L. et al. eBird: A citizen-based bird observation network in the biological sciences. *Biol. Conserv.* **142**, 2282–2292 (2009).
44. Van Doren, B. M. et al. Social associations across species during nocturnal bird migration. *Curr. Biol.* (2025).
45. Hill, S. C. et al. Antibody responses to avian influenza viruses in wild birds broaden with age. *Proc. R. Soc. B* **283**, 20162159 (2016).
46. Woodward, I. D. et al. 2022/23: The Wetland Bird Survey and Goose & Swan Monitoring Programme. (2024).
47. Katoh, K. & Standley, D. M. MAFFT: Iterative refinement and additional methods. in *Multiple sequence alignment methods* 131–146 (Springer, 2013).
48. Keck, F. & Mahon, A. Handling biological sequences in r with the bioseq package. *Methods Ecol. Evol.* **11**, (2020).
49. R Core Team. R: A Language and Environment for Statistical Computing. (R Foundation for Statistical Computing, Vienna, Austria, 2023).
50. Tamura, K. & Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512–526 (1993).
51. Paradis, E. et al. Package 'ape'. Analyses of phylogenetics and evolution, version 2, 47 (2019).

52. Bürkner, P.-C. Brms: An r package for bayesian multilevel models using stan. *J. Stat. Softw.* **80**, 1–28 (2017).
53. Farine, D. R. A guide to null models for animal social network analysis. *Methods Ecol. Evol.* **8**, 1309–1320 (2017).