A peer-reviewed version of this preprint was published in PeerJ on 30 April 2015.

View the peer-reviewed version (peerj.com/articles/934), which is the preferred citable publication unless you specifically need to cite this preprint.

Dalby AR, Iqbal M. 2015. The European and Japanese outbreaks of H5N8 derive from a single source population providing evidence for the dispersal along the long distance bird migratory flyways. PeerJ 3:e934 https://doi.org/10.7717/peerj.934

The European and Japanese outbreaks of H5N8 derive from a single source population that has most likely been dispersed along the long distance bird migratory flyways

Andrew Dalby, Munir Iqbal

The origin of recent parallel outbreaks of the high pathogenicity H5N8 avian flu virus in Europe and in Japan can be traced to a single source population, which has most likely been spread by migratory birds. By using Bayesian coalescence methods to analyze the DNA sequences of the virus to find the times for divergence and combining bird migration data we can show the most likely locations and migratory pathways involved in the origin of the current outbreak. This population was most likely located in the Siberian summer breeding grounds of long-range migratory birds. These breeding grounds provide a connection between different migratory flyways and explain the current outbreaks in remote locations. By combining genetic methods and epidemiological data we can rapidly identify the sources and the dispersion pathways for novel avian influenza outbreaks.

Andrew R. Dalby

Faculty of Science and Technology, University of Westminster, 115 New Cavendish Street, Westminster, W1W 6UW, UK.

Munir Iqbal

Avian Viral Diseases Programme, The Pirbright Institute, Compton Laboratory, Newbury, Berkshire. RG20 7NN, UK

A.Dalby@westminster.ac.uk

Introduction

The H5N8 subtype of influenza A is a comparatively rare influenza A subtype that was first isolated from a turkey in Ireland in 1983 (Murphy 1986). After that initial outbreak there were no more cases of H5N8 until 2001 when a case was identified during environmental monitoring in a wild bird in New Jersey. Since then there have been a few sporadic detections in the United States but the biggest single outbreak to date has been in Korea in January 2014 (Lee et al. 2014).

This outbreak was preceded by cases in Eastern China in 2010 that are distinct from the American virus. Although the Korean outbreak strains had the same subtype the Asian outbreak seems to have been the product of a re-assortment of viral segments from other H5 containing subtypes such as H5N1 or H5N5 and an N8 containing subtype, rather than from the evolution of the previous H5N8 lineages (Lee et al. 2014; Zhao et al. 2013). The Korean outbreak has been subdivided into two lineages one of which is closely related to the Chinese sequences and that has only been identified in two samples in Gochang and a second that contains all the other cases and that was originally identified in Buan (Fan et al. 2014; Jeong et al. 2014).

The H5N8 virus is an example of a highly pathogenic avian influenza A (HPAI). These HPAI viruses pose a significant threat to domestic poultry as mortality rates amongst chickens are particularly high and can reach 100%. Recent studies of the virus have shown that it has a pathogenicity index of 3 in chickens (Kim et al. 2014). This is significantly higher than that of the original H5N8 from Ireland although that is a distinct lineage (Alexander et al. 1986). The management of the outbreak in Korea in early 2014 resulted in the culling of over 10 million birds or 6% of the total Korean poultry flock

(Kang). Ducks and particularly wild ducks such as mallards are often asymptomatic but can still be carriers of the H5N8 virus (Bae et al. 2014; Kang ; Kim et al. 2014).

Currently unlike H5N1, H5N8 is not considered a threat to human health as there has not been a case of transmission to humans. However this might be a result of the low incidence of the subtype as studies have shown that it can be transmitted to ferrets and mice, and antibodies have been detected in domestic dogs (Kim et al. 2014). The results of genetic analysis of the H5N8 virus in infected ferrets have also shown that that mutations to a mammalian transmissible form occur rapidly.

In November 2014 H5N8 was detected in Europe with outbreaks in poultry farms in the Netherlands, Germany and the United Kingdom. At the same time the virus was also detected in farmed birds and wild birds in Japan. This study identifies the probable geographical source and pathway for dispersal of the November/December outbreaks of H5N8.

Materials and Methods

The complete set of available H5N8 nucleotide sequences were downloaded from the NCBI influenza virus resource and GISAID (Bao et al. 2008; Bogner et al. 2006). The search was restricted to complete sequences of H5N8 within the NCBI influenza virus resource.

All of the sequences were aligned with Muscle v3.8.31 (Edgar 2004). Manual inspection and editing of the sequences was carried out using Mega6.06 (Tamura et al. 2013). During manual editing the 5' end of the sequence was edited to remove the un-translated region. All sequences begin at the start codon. Sequences with missing nucleotides were removed. This included the German turkey sequence and 3 Chinese duck sequences. There was no editing at the 3' end of the nucleotide sequences as influenza uses a variety of stop codons that are sometimes repeated. Tip dates were assigned according to the year of collection.

A subset of the sequences was created for the detailed analysis of the hemagglutinin and neuraminidase sequences containing only the sequences from 2014. This provides a more detailed analysis for calculating the divergence dates from the Korean outbreak. For these calculations tip dates were given in months before December 2014. This solves the problem of missing data from earlier sequences where months might not be available.

Bayesian Coalescence trees were calculated for all the different segments and the subsets using Beast2.1.3 (Bouckaert et al. 2014). The model used assumed an exponential population growth and tip

dates were set from the sequence collection dates. The Hasegawa-Kishino-Yano nucleotide substitution model was used with an assumed rigid molecular clock (Hasegawa et al. 1985) as this was shown to give the highest posterior probabilities when compared to other substitution models. All simulations were performed as a single run with a minimum of 10 million iterations and 20 million for the NA, PB1, PB2, PA and NS segments, and a burn-in of 10%.

Analysis of the Bayesian coalescence output was carried out using Tracer1.6.0 (Rambaut & Drummond 2013a). All simulations were run until the effective sample sizes for all of the parameters in the model were over 200. The maximum clade credibility trees were calculated using Treeannotater 2.1.2 along with the median node heights and the final tree diagrams were generated using FigTree1.4.2 (Rambaut 2007; Rambaut & Drummond 2013b).

The locations of the H5N8 cases were taken from the EMPRES Global Animal Disease Information System (EMPRES-i) and information about the original reports were sourced from the Avian Flu Diary Blog (http://afludiary.blogspot.co.uk/). The map was created using Google maps and is available from: https://www.google.com/maps/d/edit?mid=zcvUWKLLjKsE.kvYJ1NxAer8k

Results and Discussion

The Bayesian coalescence analysis of the complete set of H5N8 sequences produces a consistent gene tree, where the same clade structure is produced for all eight segments (supplementary figures 1 to 8). These show that the European and Japanese sequences all form a single cluster closely related to, but distinct from those found in the Korean outbreak. This clustering suggests that the viruses are likely to come from a single source population.

Bayesian coalescent analysis of the 2014 sequences also permits the sequence divergence time to be calculated with greater accuracy (Lemey et al. 2009). The Bayesian coalescent trees for the 2014 hemaggluinin gene segments and the neuraminidase gene segments are shown in figures 1 and 2. The bars above the branch points represent the 95% highest posterior density for the distance between branches. The x-axis represents the date in months starting from January 2014. For the viral hemagglutinin gene segment the cluster of sequences responsible for the current European and Japanese outbreaks diverged between a median value of 1.58 and 5.53 months before December 2014 (95% highest posterior density), if the hemagglutninin sequence from the German sample was omitted. This sequence was excluded because of a nine base truncation at the 5' end. In a reconstructed tree where this sequence was included and the missing nucleotides were inferred to be the same as those from all of the other sequences the median time for divergence from the Korean sequences increases to between 3.22 and 7.33 months before December 2014 (95% highest posterior density). Only the hemagglutinin and neuraminidase segments are available from the infected German turkey and the median divergence time calculated from the neuraminidase tree is from 1.39 to 6.21 months before December 2014 (95% highest posterior density).

During the Korean outbreak a large number of wild birds were also affected, particularly in the region around the Dong-Lim reservoir (Jeong et al. 2014). One of the bird species that was found to be infected was the Baikal Teal (*Anas formosa*) which is a migratory species that over- winters in Korea before returning to North Eastern Siberia to breed during the summer months (Allport et al. 1991). This migration coincided with the last Korean H5N8 sequences identified in wild birds during the initial outbreak. This migration also falls within the range of divergence dates from the Bayesian coalescence analysis for the current cluster of H5N8 cases in Europe. This result strongly suggests that the virus was carried to the Siberian breeding grounds as the Baikal teal migrated north and that the European and Japanese sequences evolved there.

The wide geographic dispersal of the current outbreaks gives further support to the contention that migratory birds are the source of the virus. Most of the recent cases occur close to the coastline and in areas where there lakes and known sites for waterfowl and migratory birds. In Holland the virus has been identified in widgeon and in Germany it was found in a common teal (*Anas crecca*) that had no apparent clinical symptoms. The Japanese have recently identified the wild bird species infected with the virus as tundra swans (*Cygnus columbianus),* white naped cranes (*Grus vipio*), pochards (*Aythya ferina*) and wild ducks (*Anas platyrhyncus*).

The breeding grounds and migratory staging grounds for Baikal teal overlap with those for many other migratory species including common species such as mallards, pochards, widgeon (*Anas penelope*), common teal, whooper swans (*Cygnus cygnus*) and tundra swans, as well as endangered species such

as white-naped cranes (Miyabayashi & Mundkur 1999). Mallard and teal have previously been identified as having a high prevalence (between 6 and 7%) for influenza A virus (Munster et al. 2007). The bird migrations flow from Siberia along the five different flyways that overlap in Central Siberia. They are the East Atlantic, East Asia Australian, East Africa West Asia, Central Asia and Black Sea Mediterranean flyways. So far H5N8 infections in birds have only been detected in the East Atlantic and East Asian Australian flyways (figure 9). The absence from other flyways can be explained either through transmission by a limited number of bird migratory species, or because or the lack of surveillance in these geographical regions. Although there have been a few cases recently reported in North America these are from the different lineage not related to the Korean outbreak.

This year the winter migration has been later than usual because of the warmer autumn weather. Ideally satellite-tracking data would be available for all of the migrating species from their summer breeding grounds. However tracking data is only available for species of interest that include Bewick swans (*Cygnus bewickii*), a sub-species of Tundra swans. These tracking data show that their migration was delayed until late October and early November, which coincided with the European outbreaks of the H5N8 virus (Slimbridge Wildlife Trust).

Gaidet and co-workers considering the spread of another HPAI, H5N1 suggested that the risk of transmission by migratory birds was only a low risk because of the need for asymptomatic infections and also taking into account the distances travelled, the time taken and the number of staging points along the journey (Gaidet et al. 2010). However low pathogenicity avian influenza have been shown to spread via migrating birds because the large majority of cases remain asymptomatic (Dusek et al. 2014;

Lam et al. 2012). In the case of high pathogenicity H5N8, the virus has been shown to be asymptomatic in mallards, and there would have been selection of virus variants that are asymptomatic amongst the Baikal teal if the disease has been carried by a migrant bird (Bae et al. 2014; Kang ; Kim et al. 2014). Dispersion of the virus through migratory flyways still requires that there is relay infection for the virus to spread over very long migratory distances.

Previous studies had shown that there was a spatio-temporal relationship between bird migration and the spread of the HPAI H5N1 subtype (Takekawa et al. 2010). However it was not possible to show that transmission by the migratory birds was the cause of this correlation. In this case the genetic data and the calculated divergence times show that the evolutionary events responsible for generating the European and Japanese cases occurred in the summer months in a single location.

The initial outbreak affected a large number of birds during the period close to the main spring migration this increased the likelihood of long-range transmission. The spread of the virus requires that there is relay infection so that it spreads amongst susceptible birds at the migratory staging points, in order to provide the next step in transmission. This is seen with the presence of an increased number of cases at staging points such as the Netherlands. This is supported by the current limited amount of data, although there are other staging points in Estonia and in Denmark where there have not been any reported cases (Beekman et al. 2002; Green et al. 2002).

Conclusions

The results presented here give strong support to the view that the H5N8 outbreaks that occurred in Europe and Japan in December 2014 originated from a single source population. Although there is no direct evidence of what this source population was it is likely that the virus was spread along long-range migratory routes as trade is a less likely source given the absence in disease infections during the summer months. This suggests that the summer breeding grounds for migratory species such as Baikal teal are the most likely geographical location for the source of the outbreaks.

Increased monitoring for HPAI is needed in areas where there is overlap between migrating species, especially if this zone links very disparate geographical regions. This could be achieved through environmental monitoring of faecal samples in areas where migratory birds congregate. In the case of H5N8 the main costs are economic as it is not currently a human pathogenic subtype, but it has had devastating consequences for the Korean poultry industry. However the longer the virus is in circulation in wild birds and poultry the more likely it is that a human case will occur, especially considering the close relationship to the H5N1 strains and existing evidence that shows the virus can reproduce in mammalian hosts.

It is important to involve local communities and experts as well as farmers so that we can significantly improve the monitoring network giving earlier warning of potential epidemics. This also means improved communication between international organizations and making the biological sequence data available in a timely manner.

Acknowledgments

194 Special thanks must go to Michael Coston whose Avian Flu Diary Blog has been an important source
195 of information for the current outbreak.

196

References

Alexander D, Parsons G, and Manvell R. 1986. Experimental assessment of the pathogenicity of eight avian influenza A viruses of H5 subtype for chickens, turkeys, ducks and quail. *Avian pathology* 15:647-662.

Allport G, Poole C, Park E, Jo S, and Eldridge M. 1991. The feeding ecology, requirements and distribution of Baikal Teal Anas formosa in the Republic of Korea. *Wildfowl* 42:98-107.

Bae Y, Lee S-B, Min KC, Mo J-S, Jeon EO, Koo B-S, Kwon H-I, Choi YK, Kim JJ, and Kim J-N. 2014. Pathological evaluation of natural cases of a highly pathogenic avian influenza virus, subtype H5N8, in broiler breeders and commercial layers in South Korea. *Avian Diseases*.

Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, Ostell J, and Lipman D. 2008. The influenza virus resource at the National Center for Biotechnology Information. *Journal of virology* 82:596-601.

Beekman JH, Nolet BA, and Klaassen M. 2002. Skipping swans: fuelling rates and wind conditions determine differential use of migratory stopover sites of Bewick's Swans Cygnus bewickii. *Ardea* 90:437-460.

Birdlife International. Baikal Teal Sibirionetta formosa. *Available at* http://www.birdlife.org/datazone/species/factsheet/22680317 (accessed 4th December 2014).

Bogner P, Capua I, Lipman DJ, and Cox NJ. 2006. A global initiative on sharing avian flu data. *Nature* 442:981-981.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, and Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology* 10:e1003537.

Dusek RJ, Hallgrimsson GT, Ip HS, Jónsson JE, Sreevatsan S, Nashold SW, TeSlaa JL, Enomoto S, Halpin RA, and Lin X. 2014. North Atlantic Migratory Bird Flyways Provide Routes for Intercontinental Movement of Avian Influenza Viruses. *PloS one* 9:e92075.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* 32:1792-1797.

Fan S, Zhou L, Wu D, Gao X, Pei E, Wang T, Gao Y, and Xia X. 2014. A novel highly pathogenic H5N8 avian influenza virus isolated from a wild duck in China. *Influenza Other Respir Viruses* 8:646-653.

Gaidet N, Cappelle J, Takekawa JY, Prosser DJ, Iverson SA, Douglas DC, Perry WM, Mundkur T, and Newman SH. 2010. Potential spread of highly pathogenic avian influenza H5N1 by wildfowl: dispersal ranges and rates determined from large-scale satellite telemetry. *Journal of Applied Ecology* 47:1147-1157.

Green M, Alerstam T, Clausen P, Drent R, and Ebbinge BS. 2002. Dark-bellied Brent Geese Branta bernicla bernicla, as recorded by satellite telemetry, do not minimize flight distance during spring migration. *Ibis* 144:106-121.

Hasegawa M, Kishino H, and Yano T-a. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of molecular evolution* 22:160-174.

Jeong J, Kang H-M, Lee E-K, Song B-M, Kwon Y-K, Kim H-R, Choi K-S, Kim J-Y, Lee H-J, and Moon O-K. 2014. Highly pathogenic avian influenza virus (H5N8) in domestic poultry and its relationship with migratory birds in South Korea during 2014. *Veterinary microbiology* 173:249-257.

Kang H-M. Novel Reassortant Influenza A (H5N8) Viruses among Inoculated Domestic and Wild Ducks, South Korea, 2014.

Kim Y-I, Pascua PNQ, Kwon H-I, Lim G-J, Kim E-H, Yoon S-W, Park S-J, Kim SM, Choi E-J, and Si Y-J. 2014. Pathobiological features of a novel, highly pathogenic avian influenza A (H5N8) virus. *Emerging Microbes & Infections* 3:e75.

Lam TTY, Ip HS, Ghedin E, Wentworth DE, Halpin RA, Stockwell TB, Spiro DJ, Dusek RJ, Bortner JB, and Hoskins J. 2012. Migratory flyway and geographical distance are barriers to the gene flow of influenza virus among North American birds. *Ecology letters* 15:24-33.

Lee Y-J, Kang H-M, Lee E-K, Song B-M, Jeong J, Kwon Y-K, Kim H-R, Lee K-J, Hong M-S, and Jang I. 2014. Novel reassortant influenza A (H5N8) viruses, South Korea, 2014. *Emerging infectious diseases* 20:1087.

Lemey P, Rambaut A, Drummond AJ, and Suchard MA. 2009. Bayesian phylogeography finds its roots. *PLoS computational biology* 5:e1000520.

Miyabayashi Y, and Mundkur T. 1999. *Atlas of key sites for Anatidae in the East Asian Flyway*: Wetlands International.

Munster VJ, Baas C, Lexmond P, Waldenström J, Wallensten A, Fransson T, Rimmelzwaan GF, Beyer WE, Schutten M, and Olsen B. 2007. Spatial, temporal, and species variation in prevalence of influenza A viruses in wild migratory birds. *PLoS pathogens* 3:e61.

Murphy T. 1986. The control and epidemiology of an influenza A outbreak in Ireland. *Acute Virus Infections of Poultry*: Springer, 23-28.

Rambaut A. 2007. FigTree, a graphical viewer of phylogenetic trees. Available from http://tree.bio.ed.ac.uk/software/figtree.

Rambaut A, and Drummond A. 2013a. Tracer v1. 5 Available from http://beast.bio.ed.ac.uk/Tracer.

Rambaut A, and Drummond A. 2013b. TreeAnnotator v1. 7.0.

Slimbridge Wildlife Trust. Bewick's Swan Diary. *Available at* http://www.wwt.org.uk/news/all-news/2014/10/wwt-slimbridge-diaries/wwt-slimbridge-diaries-bewicks-swan-diary/bewicks-reach-britain/ (accessed 4th December 2014).

Takekawa JY, Newman SH, Xiao X, Prosser DJ, Spragens KA, Palm EC, Yan B, Li T, Lei F, and Zhao D. 2010. Migration of waterfowl in the East Asian flyway and spatial relationship to HPAI H5N1 outbreaks. *Avian diseases* 54:466-476.

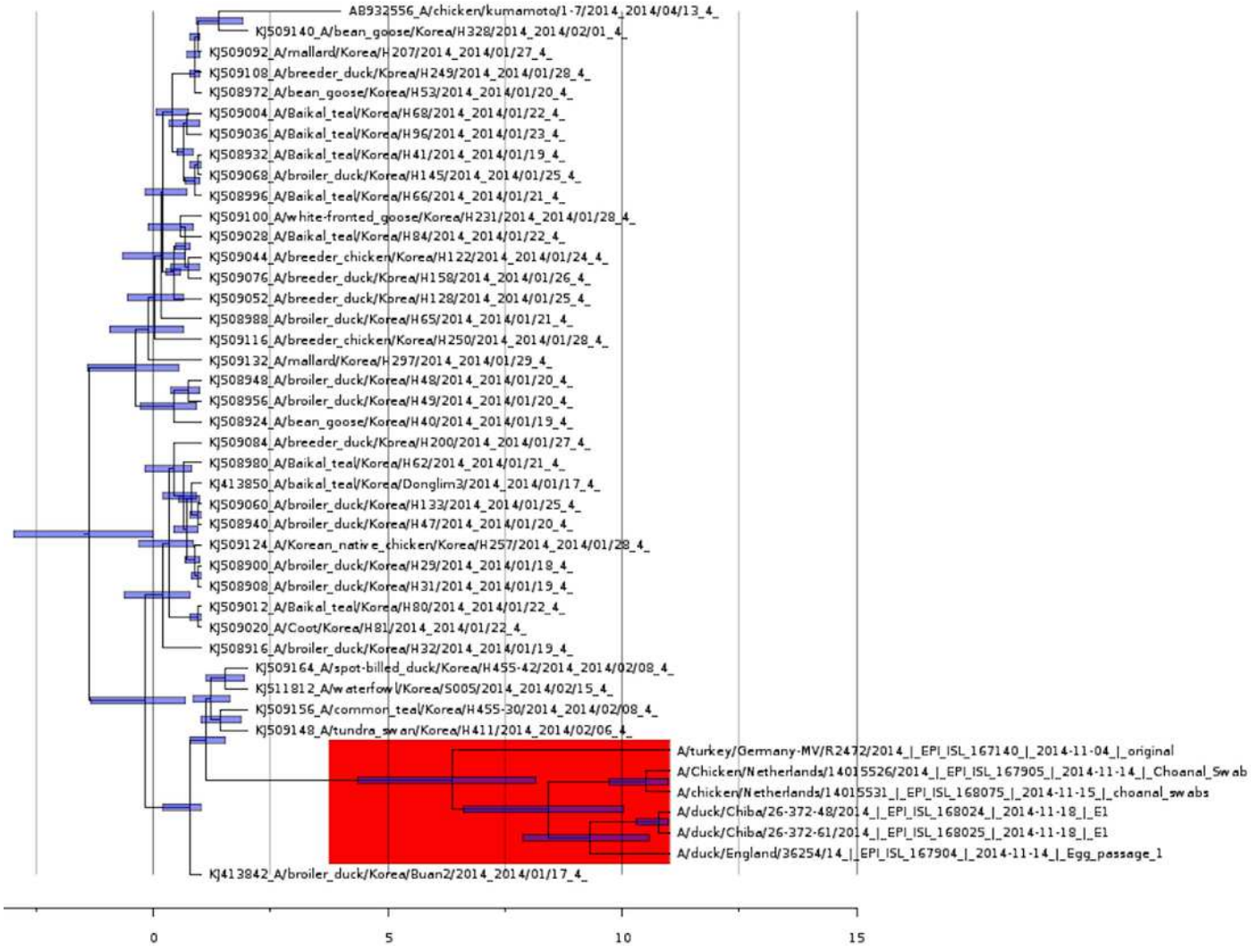
Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution* 30:2725-2729.

Zhao K, Gu M, Zhong L, Duan Z, Zhang Y, Zhu Y, Zhao G, Zhao M, Chen Z, and Hu S. 2013. Characterization of three H5N5 and one H5N8 highly pathogenic avian influenza viruses in China. *Veterinary microbiology* 163:351-357.

1

Bayesian coalescence gene tree for hemagglutinin

Bayesian coalescence gene tree for the 2014 H5N8 hemagglutinin sequences. The blue bars on the nodes represent the 95% highest posterior densities of the branch heights (this is the time for divergence in months). The European and Japanese clade is highlighted in red. 0 on the x-axis represents January 2014.



2

Bayesian coalescence tree for neuraminidase

Bayesian coalescence gene tree for the 2014 H5N8 neuraminidase sequences. The blue bars on the nodes represent the 95% highest posterior densities of the branch heights (this is the time for divergence in months). The European and Japanese clade is highlighted in red. 0 on the x-axis represents January 2014.

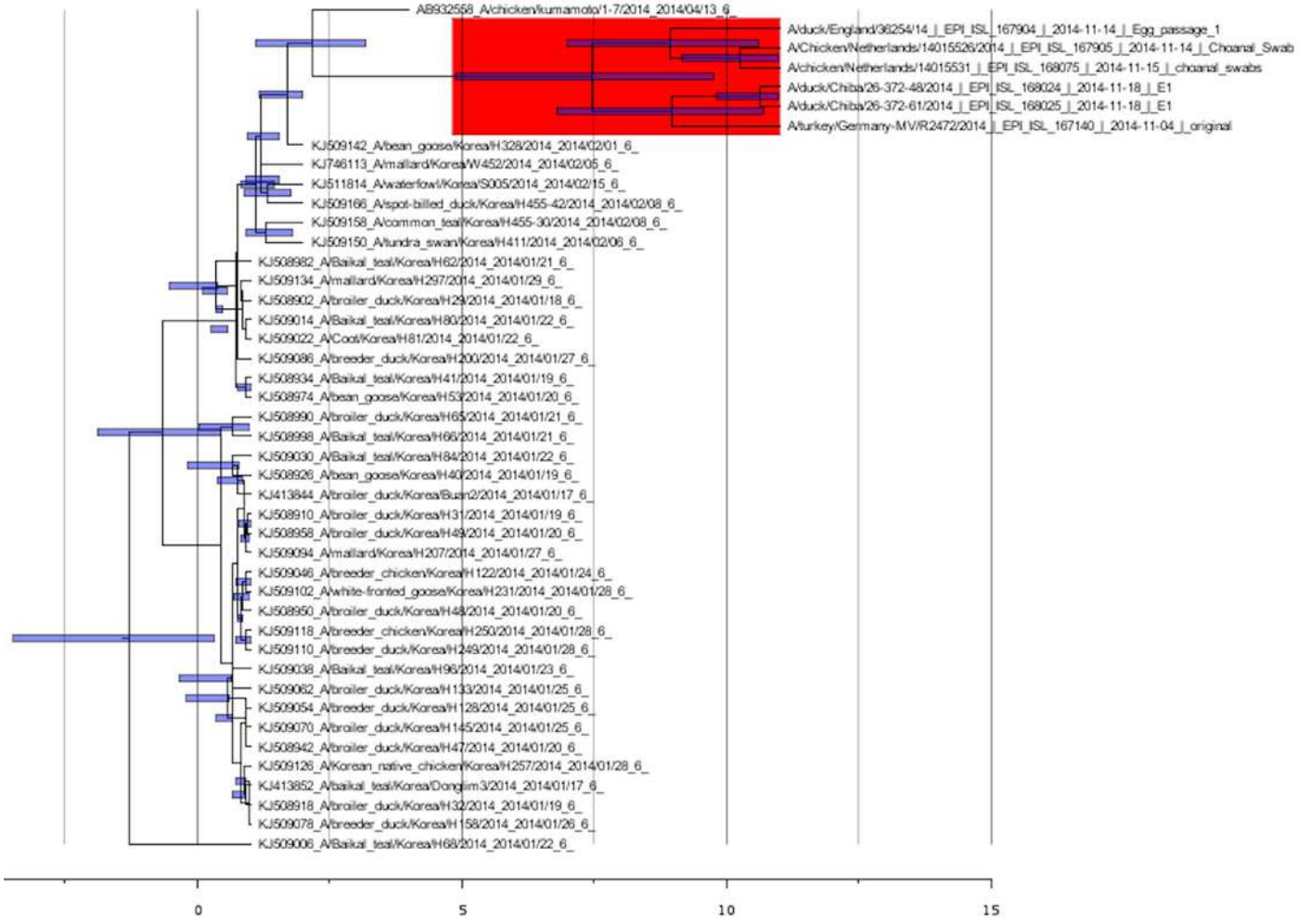


Figure 3(on next page)

Migratory flyways and cases of H5N8.

Bird migratory flyways and the December 2014 cases of H5N8. The Eastern Asian Australian flyway is in red. The East Atlantic flyway is in dark blue. An expandable version of this map is available from: https://www.google.com/maps/d/viewer?mid=zcvUWKLLjKsE.kvYJ1NxAer8k - Map Data (C) 2015 Google, INEGI

H5N8

H5N8 cases

All items

Flyways

East Atlantic Flyway

East Asian Australian Flyway

Map of H5N8 cases December 2014