



Article

A Mutation Network Method for Transmission Analysis of Human Influenza H3N2

Chi Zhang ¹, Yinghan Wang ¹, Cai Chen ¹, Haoyu Long ¹, Junbo Bai ¹, Jinfeng Zeng ¹, Zicheng Cao ¹, Bing Zhang ¹, Wei Shen ¹, Feng Tang ¹, Shiwen Liang ¹, Caijun Sun ¹ , Yuelong Shu ^{1,2} and Xiangjun Du ^{1,2,*} 

¹ School of Public Health (Shenzhen), Sun Yat-sen University, Guangzhou 510006, China; zhangch295@mail2.sysu.edu.cn (C.Z.); wangyh256@mail2.sysu.edu.cn (Y.W.); chenc57@mail2.sysu.edu.cn (C.C.); longhy7@mail2.sysu.edu.cn (H.L.); baijb@mail2.sysu.edu.cn (J.B.); zengjf7@mail2.sysu.edu.cn (J.Z.); caozch@mail2.sysu.edu.cn (Z.C.); zhangbing4502431@outlook.com (B.Z.); shenw3@mail.sysu.edu.cn (W.S.); tangf35@mail2.sysu.edu.cn (F.T.); liangshw5@mail2.sysu.edu.cn (S.L.); suncaijun@mail.sysu.edu.cn (C.S.); shuylong@mail.sysu.edu.cn (Y.S.)

² Key Laboratory of Tropical Disease Control, Ministry of Education, Sun Yat-sen University, Guangzhou 510006, China

* Correspondence: duxj9@mail.sysu.edu.cn

Received: 1 September 2020; Accepted: 1 October 2020; Published: 3 October 2020



Abstract: Characterizing the spatial transmission pattern is critical for better surveillance and control of human influenza. Here, we propose a mutation network framework that utilizes network theory to study the transmission of human influenza H3N2. On the basis of the mutation network, the transmission analysis captured the circulation pattern from a global simulation of human influenza H3N2. Furthermore, this method was applied to explore, in detail, the transmission patterns within Europe, the United States, and China, revealing the regional spread of human influenza H3N2. The mutation network framework proposed here could facilitate the understanding, surveillance, and control of other infectious diseases.

Keywords: influenza virus; mutation network; transmission; phylogenetic analysis

1. Introduction

Yearly epidemics of seasonal influenza cause an enormous disease burden around the globe, and they are responsible for approximately three to five million annual cases of severe illness and 290,000 to 650,000 respiratory deaths [1]. The annual spread of seasonal influenza virus is driven by the continuous evolution of the virus itself as a result of the immune system [2,3]. For the effective surveillance and control of influenza, the World Health Organization (WHO) has established an international network of collaborating laboratories to monitor the spread and antigenic variation of influenza viruses [4]. An understanding of the spatial spread of human influenza is of critical importance for targeted surveillance and precise vaccine recommendations, ultimately leading to better control of human influenza.

The global circulation pattern of human influenza virus, especially for the H3N2 subtype, was well characterized by many previous studies, where the East and Southeast Asia regions are central to the emergence of new variants, while the United States may play an important role in their global spread [3,5–7]. However, regional transmission patterns for migration hubs such as Europe, the United States, and China are less studied, despite their importance for both regional and global surveillance and the control of seasonal influenza.

Several methods were proposed to study the spatial transmission of seasonal influenza using phylogenetic theory [8–10]. With the rapid development of network theory and its wide application

in interdisciplinary fields [11–13], researchers are also exploring its potential in the transmission analysis of infectious diseases [5,14–16]. Compared with traditional phylogeographic methods, network presentation relies on fewer assumptions and aligns with the intrinsic nature of pathogen networks [14,17,18]. This field is still in its infancy, and the full potential of network theory in the study of infectious diseases needs to be explored.

In this study, we developed a mutation network framework to study the spatial transmission of human influenza H3N2 using sequence data. We further applied this framework to an in-depth transmission analysis in Europe, the United States, and China.

2. Materials and Methods

2.1. Simulations

The regional expansion and transmission of seasonal influenza H3N2 were simulated according to a realistic scenario. The expansion of human influenza H3N2 within a region was simulated using a specific transmission model with a periodic transmission rate involving a cycle of one year to simulate the seasonality of human influenza virus.

$$\Delta I(t+1) = \lambda(t)\Delta I(t) \quad (1)$$

$$\lambda(t) = \mu \left[1 + \sin\left(\frac{1}{26}\pi t\right) \right]$$

where ΔI is the number of new infected cases at time t , and $\lambda(t)$ is the transmission rate, formulated using a periodic function, with a scaling factor μ , which was fixed at 0.75. Infected people were expected to recover within 1 week. As a result, ΔI is also the number of infected cases at time t . The transmission events between regions (six regions: A, B, C, D, E, and F) was calculated on the basis of a predefined transmission intensity.

$$N(t)_{j \rightarrow i} = \begin{cases} 0, & \Delta I(t)_j TI_{j \rightarrow i} < 1 \\ \Delta I(t)_j TI_{j \rightarrow i}, & \Delta I(t)_j TI_{j \rightarrow i} \geq 1 \end{cases} \quad (2)$$

where $N(t)_{j \rightarrow i}$ is the number of transmission events from region j to region i , $\Delta I(t)_j$ is the number of infected cases at time t for region j , and $TI_{j \rightarrow i}$ is the predefined transmission intensity from region j to region i (see Table S1, Supplementary Materials). The relative interregional transmission intensity was then calculated as the interregional transmission intensity divided by the total transmission intensity across all regions. The transmission events for a specific region were calculated as the sum of transmission events from all other regions. Each year, one specific region was selected as the source region with one initial infection as a function of the trunk proportion (0.35, 0.25, 0.15, 0.1, 0.08, and 0.07 for regions A, B, C, D, E, and F, respectively). The simulation was run weekly and lasted 20 years. The genetic sequences were simulated as a function of a predefined mutation probability [19] and a mutation rate of 6.7×10^{-3} per site and year [20]. The mutation probabilities between A and C, A and G, A and T, C and G, C and T, and G and T were set to 0.1, 0.5, 0.07, 0.04, 0.25, and 0.04, respectively [19]. The whole transmission process was recorded, including the sequences. For each region, 200 sequences were sampled every year, while 100 repeated simulations were carried out.

2.2. Bedford Data

A dataset from Bedford et al. [7] was used in this study for validation. In total, 4004 hemagglutinin (HA) sequences of the human influenza H3N2 virus, which were evenly sampled from nine regions in the period from January 2000 to April 2012, were included in the dataset. The number of passengers on interregional flights from their study was also used. For comparison, the interregional migration rate and trunk proportion for each region, obtained from phylogeographic analysis, were also used.

2.3. Nucleotide Sequence Data

Nucleotide HA sequences of human influenza H3N2 from Europe and the United States (US) between April 2008 and November 2019, and from China between September 2010 and April 2018 were downloaded from the Global Avian Influenza Data Sharing Initiative (GISAID) website. Overall, 22,867, 19,113, and 1802 sequences were used for Europe, the US, and China, respectively. Sequences were aligned with MUSCLE v3.7 using default settings [21]. Undetermined nucleotides or gaps were replaced with the highest-frequency nucleotides at the corresponding position on the basis of the top 10 closest sequences measured by Hamming distance. A phylogenetic tree was constructed using FastTree 2 with default settings [22]. Outliers on the phylogenetic tree and redundant sequences from the same place and season were removed from the dataset. The summer season was defined as the period from April to September, whereas the winter season was defined as the period from October to March of the next year. In total, 100 sampling datasets were constructed by randomly selecting five and 20 sequences in the summer and winter seasons, respectively, for each selected country in Europe and for each division from the US Department of Health and Human Services (HHS). Similarly, two and 10 sequences were randomly selected in the summer and winter seasons from each province or region in China. In situations involving a lack of sequences, they could all be used.

2.4. Mutation Network Construction

2.4.1. Initial Network

The initial network was constructed with nodes representing strains and links representing pairs of strains with a mutation probability no more than a predetermined threshold, which was defined as the 10th percentile of a collection of the highest mutation probabilities from each strain with their closest relative (Figure 1A, initial network). Baseline mutation probabilities among A, T, G, and C were extracted from pairs of sequences with single-nucleotide differences in the corresponding data. Mutation probabilities between pairs of strains with different numbers of mutations were calculated as the product of probabilities of single mutations.

2.4.2. Fully Connected Network

The initial network could not be fully connected if strains were too far apart. As a result, in the second step, an isolated subnetwork was connected to its closest subnetwork via the link with the highest probability. This process was iteratively repeated until the network was fully connected. Consequently, an undirected and weighted network was obtained, with the weight representing the mutation probability between two nodes (Figure 1A, connected network).

2.4.3. Directed Network

In order to define the direction of links, the earliest strain was selected as the root. The direction was then defined on the basis of the shortest paths from the root to all other strains. The shortest path was extracted using the Dijkstra algorithm, and the weights were transformed using the original probability with a negative logarithmic function. Then, a partial directed network was constructed, and directions for the remaining links were defined according to the following principles: (1) the direction went from the strain with a higher mutation probability from the root to the strain with a lower mutation probability from the root; (2) the direction between two strains with the same mutation probability from the root was defined as going from the strain with the earlier collection date to the strain with the later collection date; (3) a circular structure was not allowed, with the direction being flipped if one appeared. Finally, a directed and weighted network was constructed (Figure 1A, directed network).

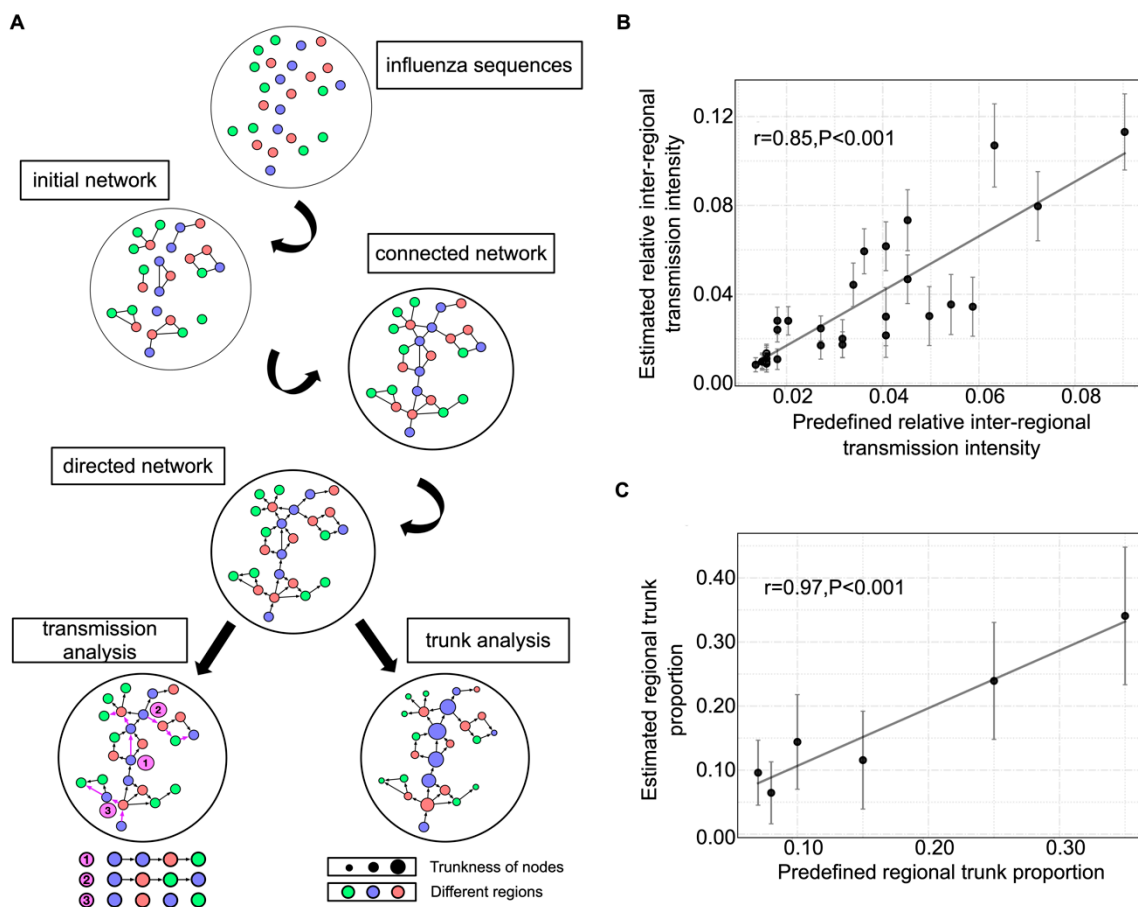


Figure 1. Mutation network framework and simulation validation. (A) Flowchart of the mutation network framework. Regions are differentiated using colors. Trunkness is the corrected betweenness centrality of nodes (see Section 2 for details). (B) Correlation between predefined and estimated relative interregional transmission intensity. (C) Correlation between predefined and estimated trunk proportion. The 95% confidence intervals are also given.

2.5. Transmission Analysis

In the mutation network, each node has a geographic attribute. In order to quantify the transmission pattern, random walks were undertaken in the mutation network. For a better representation of the starting point, a random node was selected on the basis of its Pagerank value [23]. Then, the next move was determined as a function of the weight, representing the mutation probability to its child nodes. This process was repeated until a terminal node (a node without children) was attained. A transmission event was recorded upon a change in location (Figure 1A, transmission analysis). The whole process was repeated 10,000 times to obtain a relative transmission intensity between two regions (i.e., the total number of transmission events between two regions normalized by the number of all transmission events).

2.6. Trunk Analysis

The betweenness centrality (BC) was used to quantify the “trunkness” of nodes in the network. Due to the single-trunk structure of the mutation network for human influenza H3N2, the BC for nodes located in the middle of the trunk was systematically amplified. To correct this bias, a new measure, named trunkness, was constructed by counting the number of the shortest paths passing through a node between a randomly chosen strain and each terminal strain in the network. This was

repeated 1000 times, and the relative proportion of trunkness for all nodes was used as an indicator of proportion in the trunk (Figure 1A, trunk analysis).

2.7. Statistical Analysis

Pearson correlation analysis was used in this study. All analysis results were carried out using a two-sided test with 0.05 as the significance level.

We used Python (version 3.7, Python Software Foundation, <https://www.python.org/>) and its extension libraries Pandas and Numpy for data cleaning and processing, Networkx for network analysis, Matplotlib and Basemap for mapping, and Gephi0.9.2 (<https://gephi.org>) for network visualization [24].

3. Result

3.1. Mutation Network Framework and Simulation Validation

The mutation network was constructed using large-scale influenza sequences through a series of processes and, eventually, a directed, weighted, and connected network was generated (Figure 1A). A range of algorithms, including random walk, shortest path, Pagerank, and BC, were used to study interregional transmission patterns and the source probability (probability of being the origin) of the human influenza H3N2 virus (see Section 2 for more details).

In order to test the effectiveness of the mutation network framework in inferring interregional transmission patterns of influenza and the source probability from genetic sequences, a well-calibrated simulation was carried out (see Section 2 for more details). The results show that the analysis using the mutation network could correctly retrieve the interregional transmission patterns and source probabilities. The correlations between estimated and predefined values were as high as 0.85 and 0.97 for relative transmission intensity and trunk proportion, respectively (Figure 1B,C).

3.2. Transmission Analysis for Bedford Data

We used the dataset published by Bedford et al. [7] to demonstrate the mutation network framework. Following the procedures outlined in Section 2, a fully connected, weighted, and directed network was constructed with 3946 nodes and 95,364 edges (see Figure 2A). As can be seen, the evolution process of human influenza H3N2 followed a single trunk structure with relatively short side branches. The out degree of nodes (number of child nodes) followed a power law distribution [25], suggesting the presence of a few hub strains with several children but many strains with a few children (see Figure S1, Supplementary Materials).

Nine geographical regions were defined for the dataset from Bedford et al. [7]: US/Canada, South America, Europe, India, North China, South China, Japan/Korea, Southeast Asia, and Oceania (Figure 2B). The results show that there was close communication between US/Canada and Europe, while regions from Asia were tightly connected (Figure 2B). We determined a high correlation between our proposed transmission matrix and that from Bedford et al. (Pearson correlation 0.81, $p < 0.001$; Figure 2C). According to current knowledge, flight traffic is the major factor driving global influenza transmission [6,26]. The correlation between travel flow and transmission intensity in our mutation network was also high (Pearson correlation 0.76, $p < 0.001$; Figure 2D).

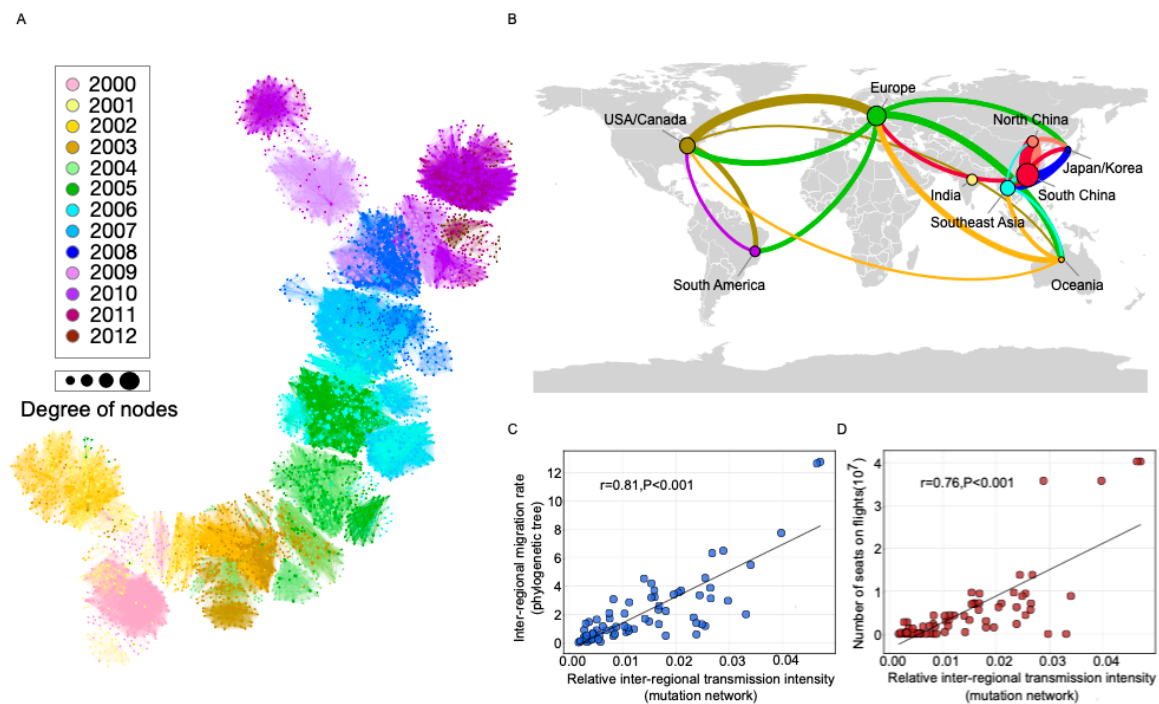


Figure 2. Global transmission and correlation analysis using the mutation network of human influenza H3N2. (A) Mutation network of human influenza H3N2. The color of the circle represents the time attribute of the node, while its size represents the degree of the node. (B) Global transmission network of human influenza H3N2 according to mutation network framework. The size of the circle represents the proportion of the region in the trunk of the mutation network. Transition lines between regions are colored on the basis of their origin. The thickness of the line represents the intensity of transmission and, for clarity, only lines with transmission intensity greater than 0.02 are shown. (C) Correlation between the relative transmission intensity from our mutation network and the results from the study by Bedford et al. [7]. (D) Correlation between the relative transmission intensity from our mutation network and the number of flight passengers between different regions.

3.3. Trunk Analysis for Bedford Data

The trunk in the mutation network represents the main spreading route during global migration of human influenza H3N2. A greater proportion of a node in the trunk denotes its greater importance in the global spread of human influenza H3N2. The results show that viruses from South and North China occupied the largest proportion of the trunk (18.5% and 9.9%, respectively), followed by Europe (15.9%), and the United States/Canada (13.2%) (Figure 3B). This result was highly correlated with the result from the study by Bedford et al. [7] ($0.69, p = 0.04$) (Figure 3C).

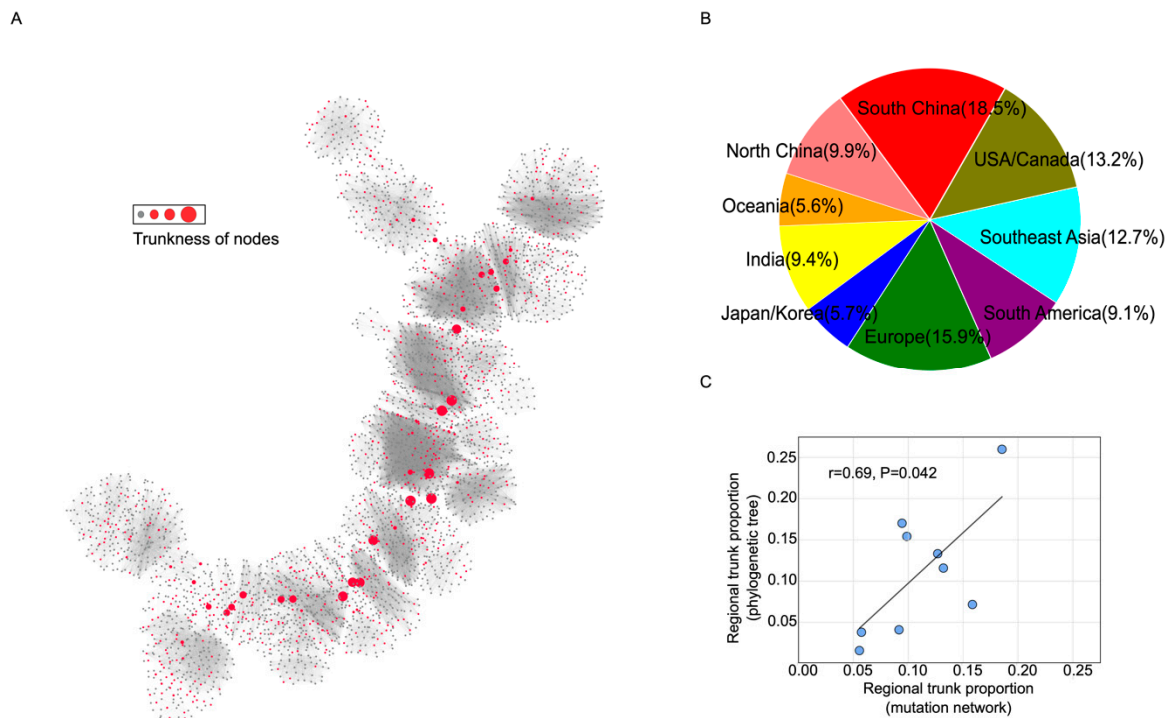


Figure 3. Trunk and correlation analysis using the mutation network of human influenza H3N2. (A) Trunkness of nodes in the mutation network. The trunkness of a node was defined by counting the number of shortest paths passing through a node between a randomly chosen strain and each terminal strain (node without children) in the network. Red circles represent the nodes contributing to the trunk, whereas their size represents the extent of the contribution. Gray circles represent the nodes not contributing to the trunk. (B) Proportion of each region in the trunk of the mutation network. (C) Correlation between the proportion in the trunk from the mutation network and that from the study by Bedford et al. [7].

3.4. Transmission Analysis for Europe, the United States, and China

Using the mutation network framework proposed herein, we determined more detailed transmission patterns for specific regions. In this study, we chose Europe, the United States, and China for further evaluation, as they play important roles in the global transmission of human influenza H3N2. To reduce the bias caused by sequence data from different regions, we carried out our analysis using mutation networks constructed with balance-sampled sequences from different countries or regions (see Section 2 for more details). We employed the 11 European countries (Sweden, Norway, the United Kingdom, the Netherlands, Germany, France, Spain, Austria, Italy, Russia, Switzerland) with sufficient data, and we used the 10 divisions of the US Department of Health and Human Services (HHS). Furthermore, in order to avoid sampling biases during the summer, when influenza activity is low in Europe and the United States, these strains were not counted. For China, we only focused on the southern provinces and northern regions with sufficient data in the summer season (Beijing, Tianjin, and Shandong), while we used the three regions proposed by Yu et al. for the winter season [27].

The results show that selected European countries contributed substantially to the trunk (Figure 4A). Regions from the northwest US occupied a larger proportion of the trunk, which indicates their important role in the spread of human influenza H3N2 (Figure 4B). In China, Guangdong, was the most important region in the spread of human influenza H3N2 during the summer (Figure 4C). During the winter, all three geographic regions played important roles in the spread with a comparable proportion in the trunk (Figure 4D).

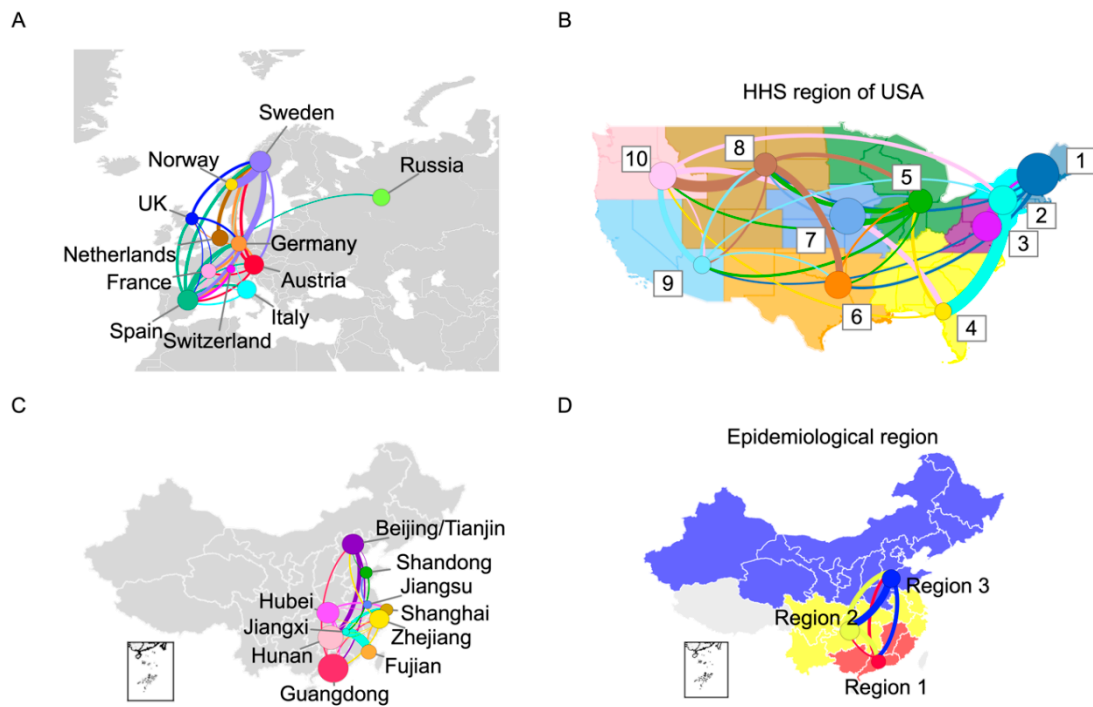


Figure 4. Transmission patterns in Europe, the United States, and China. **(A)** Regional transmission pattern in Europe. **(B)** Regional transmission pattern in the United States. Numbers 1-10 represent the 10 divisions of the US Department of Health and Human Services (HHS). **(C)** Transmission pattern in China for the summer season. **(D)** Transmission pattern in China for the winter season. The size of the circle represents the proportion in the trunk of the mutation network. Transition lines between regions are colored on the basis of their origin. The thickness of the line represents the transmission intensity (incomparable between subfigures) and, for clarity, only lines with interregional transmission intensity greater than 0.01 are shown.

4. Discussion

In this study, we proposed a sequence-based mutation network framework and applied it to study the transmission of human influenza H3N2. Our study indicates the feasibility of network theory in the study of infectious diseases, allowing for more detailed transmission information to be ascertained within the areas of Europe, the United States, and China, which will be helpful for the targeted surveillance and control of human influenza H3N2.

The mutation network framework was validated using a simulation, and results obtained using the mutation network were consistent with those obtained using a traditional phylogenetic method. Our results in the United States and China were also consistent with epidemiological data from the literature [28–30]. In Europe, previous studies showed that there is a phenomenon of spread from Western to Eastern Europe, whereas the trend between Northern and Southern Europe is not significant [31,32]. In this study, due to the availability of sequences, almost all countries in this study were from Western Europe (except Russia) and, indeed, no clear trend was observed within these countries with comparable trunk proportions. Regions from the northeast US play important roles in the spread of human influenza H3N2. Surveillance in these regions should be strengthened. South China plays an important role in the spread of human influenza H3N2 in the summer, whereas three regions contribute equally during the winter season, which indicates the rapid circulation of human influenza H3N2 during this period in China. North China does not typically experience seasonal influenza in the summer; however, our results indicate that Beijing and Tianjin also play important roles in the spread of human influenza H3N2 during this period. These two cities are within the capital economic region and are deemed major transport hubs in China. As municipalities, these two cities have strengthened

their surveillance system, and, as a result, they may better capture new variants of the human influenza H3N2 virus.

Our study also had several limitations. Recently, a median-joining network was used to study the evolutionary relationship of the novel coronavirus (SARS-CoV-2) [33]. However, this method was criticized due to its distance-based phenetics and lack of direction [34]. There were also other explorations applying network theory in the study of infectious diseases, but most were based on simple assumptions [5,35,36]. The mutation network proposed here was a directed network with mutation probabilities representing the weights of links; however, the network did not include intermediate or hidden nodes, which may represent unsampled strains. Furthermore, the transmission and trunk analyses were based on either Pagerank or BC with random walks; thus, other measures should be carefully evaluated [37]. Moreover, although we tried to balance the temporal and spatial distribution of the sequence data through our sampling approach, its effect was limited. This drawback may be exacerbated when fewer sequence data are available. With the development of sequencing technologies, especially portable equipment with lower cost, more sequences will be accumulated in the future; therefore, mutation networks are expected to play an increasingly important role in the study of infectious diseases.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1999-4915/12/10/1125/s1>.

Author Contributions: X.D. designed the study; C.Z., Y.W., C.C., H.L., J.B., J.Z., Z.C., B.Z., W.S., F.T., and S.L. collected and analyzed the data; X.D., Y.S., and C.S. interpreted the data; X.D., C.Z., and Y.W. prepared the manuscript; X.D., Y.S., Y.W., and C.Z. edited the paper. All authors reviewed and approved the submitted manuscript.

Funding: This research was funded by the National Key R&D Program of China under Grant 2020YFC0840900, the National Natural Science Foundation of China under Grants 31970643 and 81961128002, the Shenzhen Science and Technology Program under Grant KQTD20180411143323605, and the Guangdong Frontier and Key Tech Innovation Program under Grants 2019B020228001 and 2019B111103001.

Acknowledgments: We gratefully acknowledge all authors from the laboratories responsible for obtaining the specimens, as well as the laboratories where genetic sequence data were generated and shared via the Global Avian Influenza Data Sharing Initiative (GISAID), on which this research was based.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

WHO	World Health Organization
HA	hemagglutinin
GISAID	Global Avian Influenza Data Sharing Initiative
HHS	US Department of Health and Human Services
BC	betweenness centrality
SARS-CoV-2	the 2019 novel coronavirus

References

1. World Health Organization (2018) Fact sheet Number 211. Available online: <http://www.who.int/mediacentre/factsheets/fs211/en/> (accessed on 6 November 2018).
2. Nelson, M.; Holmes, E. The evolution of epidemic influenza. *Nat. Rev. Genet.* **2007**, *8*, 196–205. [[CrossRef](#)] [[PubMed](#)]
3. Russell, C.A.; Jones, T.C.; Barr, I.G.; Cox, N.J.; Garten, R.J.; Gregory, V.; Gust, I.D.; Hampson, A.W.; Hay, A.J.; Hurt, A.C.; et al. The global circulation of seasonal influenza A (H3N2) viruses. *Science* **2008**, *320*, 340–346. [[CrossRef](#)] [[PubMed](#)]
4. Cox, N.J.; Brammer, T.L.; Regnery, H.L. Influenza: Global surveillance for epidemic and pandemic variants. *Eur. J. Epidemiol.* **1994**, *10*, 467–470. [[CrossRef](#)] [[PubMed](#)]
5. Chan, J.; Holmes, A.; Rabadan, R. Network Analysis of Global Influenza Spread. *PLoS Comput. Biol.* **2010**, *6*. [[CrossRef](#)] [[PubMed](#)]

6. Lemey, P.; Rambaut, A.; Bedford, T.; Faria, N.; Bielejec, F.; Baele, G.; Russell, C.A.; Smith, D.J.; Pybus, O.G.; Brockmann, D.; et al. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Pathog* **2014**, *10*, e1003932. [[CrossRef](#)]
7. Bedford, T.; Riley, S.; Barr, I.G.; Broor, S.; Chadha, M.; Cox, N.J.; Daniels, R.S.; Gunasekaran, C.P.; Hurt, A.C.; Kelso, A.; et al. Global circulation patterns of seasonal influenza viruses vary with antigenic drift. *Nature* **2015**, *523*, 217–220. [[CrossRef](#)]
8. Venton, D. Phylogeography. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 4158. [[CrossRef](#)]
9. Bielejec, F.; Rambaut, A.; Suchard, M.A.; Lemey, P. SPREAD: Spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* **2011**, *27*, 2910–2912. [[CrossRef](#)]
10. Riddle, B.; Hafner, D. 7 Phylogeography in Historical Biogeography: Investigating the Biogeographic Histories of Populations, Species, and Young Biotas. *Biogeogr. Chang. World* **2007**, *1*, 161–176.
11. Kunin, V.; Goldovsky, L.; Darzentas, N.; Ouzounis, C.A. The net of life: Reconstructing the microbial phylogenetic network. *Genome Res.* **2005**, *15*, 954–959. [[CrossRef](#)]
12. Derrible, S.; Kennedy, C. Applications of Graph Theory and Network Science to Transit Network Design. *Transport. Rev.* **2011**, *31*, 495–519. [[CrossRef](#)]
13. Wedgwood, K.C.A.; Satin, L.S. Six degrees of depolarization: Comment on "Network science of biological systems at different scales: A review" by Marko Gosak et al. *Phys. Life Rev.* **2018**, *24*, 136–139. [[CrossRef](#)] [[PubMed](#)]
14. Bandelt, H.J.; Forster, P.; Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48. [[CrossRef](#)] [[PubMed](#)]
15. Campbell, E.M.; Jia, H.W.; Shankar, A.; Hanson, D.; Luo, W.; Masciotra, S.; Owen, S.M.; Oster, A.M.; Galang, R.R.; Spiller, M.W.; et al. Detailed Transmission Network Analysis of a Large Opiate-Driven Outbreak of HIV Infection In the United States. *J. Infect. Dis.* **2017**, *216*, 1053–1062. [[CrossRef](#)]
16. de Bernardi Schneider, A.; Ford, C.T.; Hostager, R.; Williams, J.; Cioce, M.; Çatalyürek, Ü.V.; Wertheim, J.O.; Janies, D. StrainHub: A phylogenetic tool to construct pathogen transmission networks. *Bioinformatics* **2020**, *36*, 945–947. [[CrossRef](#)]
17. Legendre, P.; Makarenkov, V. Reconstruction of biogeographic and evolutionary networks using reticulograms. *Syst. Biol.* **2002**, *51*, 199–216. [[CrossRef](#)]
18. Huson, D.H.; Scornavacca, C. A Survey of Combinatorial Methods for Phylogenetic Networks. *Genome Biol. Evol.* **2010**, *3*, 23–35. [[CrossRef](#)]
19. Yang, Z. Estimating the pattern of nucleotide substitution. *J. Mol. Evol.* **1994**, *39*, 105–111. [[CrossRef](#)]
20. Lofgren, E.; Fefferman, N.; Naumov, Y.; Gorski, J.; Naumova, E. Influenza Seasonality: Underlying Causes and Modeling Theories. *J. Virol.* **2007**, *81*, 5429–5436. [[CrossRef](#)]
21. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)]
22. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* **2010**, *5*, e9490. [[CrossRef](#)] [[PubMed](#)]
23. Brin, S.; Page, L. The anatomy of a large-scale hypertextual Web search engine. *Comput. Netw. ISDN Syst.* **1998**, *30*, 107–117. [[CrossRef](#)]
24. Bastian, M.; Heymann, S.; Jacomy, M. Gephi: An Open Source Software for Exploring and Manipulating Networks. *ICWSM* **2009**, *8*, 361–362.
25. Barabasi, A.-L. Scale-Free Networks: A Decade and Beyond. *Science* **2009**, *325*, 412–413. [[CrossRef](#)]
26. Thiemann, C.; Theis, F.; Grady, D.; Brune, R.; Brockmann, D. The structure of borders in a small world. *PLoS One* **2010**, *5*, e15422. [[CrossRef](#)]
27. Yu, H.J.; Alonso, W.J.; Feng, L.Z.; Tan, Y.; Shu, Y.L.; Yang, W.Z.; Viboud, C. Characterization of Regional Influenza Seasonality Patterns in China and Implications for Vaccination Strategies: Spatio-Temporal Modeling of Surveillance Data. *PLoS Med.* **2013**, *10*. [[CrossRef](#)]
28. Charu, V.; Zeger, S.; Gog, J.; Bjørnstad, O.N.; Kissler, S.; Simonsen, L.; Grenfell, B.T.; Viboud, C. Human mobility and the spatial transmission of influenza in the United States. *PLoS Comput. Biol.* **2017**, *13*, e1005382. [[CrossRef](#)]
29. Zinder, D.; Bedford, T.; Baskerville, E.B.; Woods, R.J.; Roy, M.; Pascual, M. Seasonality in the migration and establishment of H3N2 Influenza lineages with epidemic growth and decline. *Bmc Evol. Biol.* **2014**, *14*. [[CrossRef](#)]

30. Du, X.J.; Dong, L.B.; Lan, Y.; Peng, Y.S.; Wu, A.P.; Zhang, Y.; Huang, W.J.; Wang, D.Y.; Wang, M.; Guo, Y.J.; et al. Mapping of H3N2 influenza antigenic evolution in China reveals a strategy for vaccine strain recommendation. *Nat. Commun.* **2012**, *3*. [[CrossRef](#)]
31. Caini, S.; Alonso, W.J.; Séblain, C.E.-G.; Schellevis, F.; Paget, J. The spatiotemporal characteristics of influenza A and B in the WHO European Region: Can one define influenza transmission zones in Europe? *Eurosurveillance* **2017**, *22*. [[CrossRef](#)]
32. Paget, J.; Marquet, R.; Meijer, A.; van der Velden, K. Influenza activity in Europe during eight seasons (1999–2007): An evaluation of the indicators used to measure activity and an assessment of the timing, length and course of peak activity (spread) across Europe. *BMC Infect. Dis* **2007**, *7*, 141. [[CrossRef](#)] [[PubMed](#)]
33. Forster, P.; Forster, L.; Renfrew, C.; Forster, M. Phylogenetic network analysis of SARS-CoV-2 genomes. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 9241–9243. [[CrossRef](#)] [[PubMed](#)]
34. Kong, S.; Sánchez-Pacheco, S.J.; Murphy, R.W. On the use of median-joining networks in evolutionary biology. *Cladistics* **2016**, *32*, 691–699. [[CrossRef](#)]
35. Wagner, A. A genotype network reveals homoplastic cycles of convergent evolution in influenza A (H3N2) haemagglutinin. *Proc. R. Soc. B Biol. Sci.* **2014**, *281*. [[CrossRef](#)]
36. Wang, C.; Lyu, N.; Deng, L.; Wang, J.; Gu, W.; Ding, H.; Wu, Y.; Luo, J.; Wang, L.; Lyv, X.; et al. Evolutionary Pattern and Large-Scale Architecture of Mutation Networks of 2009 A (H1N1) Influenza A Virus. *Front. Genet.* **2018**, *9*, 204. [[CrossRef](#)]
37. Dolev, S.; Elovici, Y.; Puzis, R. Routing Betweenness Centrality. *J. ACM* **2010**, *57*. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).