Electronic supplementary materials
https://doi.org/10.1631/jzus.B1700374

Avian influenza viruses (AIVs) H9N2 are in the course of reassorting into novel AIVs

Hui-ping CHANG[§1], Li PENG[§2], Liang CHEN[3,4], Lu-fang JIANG[3,4], Zhi-jie ZHANG[4,5], Cheng-long XIONG[†‡3,4], Gen-ming ZHAO[4,5], Yue CHEN[6], Qing-wu JIANG[4,5]

Data S1 Materials and methods

Sequence collection and data pre-treatment

An isolate labeled A/Zhejiang/DTID-ZJU01/2013(H7N9) was taken as the representative of AIVs H7N9 that caused the first epidemic in China in 2013, and the one named Jiangxi Donghu was taken as the first human infected novel AIV H10N8. A/Yunnan/0127/2015(H5N6) was served as the representative one of AIVs H5N6 whenever needed. Their genomes were downloaded from the Global Initiative on Sharing Avian Influenza Data (GISAID) database (http://platform.gisaid.org/epi3/frontend). The six internal gene sequences of Asian influenza A viruses which were collected from 1990 to 2013 were derived from the NCBI Influenza Virus Sequence Database (http://www.ncbi.nlm.nih.gov/genomes/FLU/aboutdatabase.html).

MEGA6 (http://megasoftware.net) was used for the pre-treatment of genome data. The sequences were kept only one or more for further analysis if 1) they shared the same HA and NA subtypes, and were isolated from the same epidemic area within 2 years, and 2) they showed high homogeneity in the phylogenetic tree, since they might be the same strain but established from different individual host.

Cluster analysis

MEGA6 software, as well as p-distance trees, was used again to do the cluster analysis. After Clustal W alignment, trees of six large-scale gene sequence matrices were constructed. Significance testing was bootstrapped with 1,000 replicates.

Sequences shared close genetic relationships with human infected AIVs H7N9 and H10N8

and clustered within the same branches in each tree were searched by eyeball, and the distributions of them were also recorded down.

Computation of the mean evolutionary rate and the tMRCA

jModeltest 2.1.3 program (http://darwin.uvigo.es) was used to estimate the likelihood value of the model and to determine the best maximum likelihood tree. Three nucleotide substitution schemes and 24 candidate models were evaluated. Both the mean nucleotide institution rate and the tMRCA (the most recent common ancestor) were calculated using BEAST v1.6.1 (http://beast.bio.ed.ac.uk). For the tMRCA analysis, the collection dates of the H10N8 and H7N9 AIVs were calculated using the following formula:

[(month-1)×30＋date]/360＋2013.

In the tMRCA analysis, an uncorrelated lognormal relaxed clock model was engaged, and the nucleotide substitution and clock models were set as GTR＋I＋G and strict, respectively. The log files were imported into Tracer v1.5 (http://beast.bio.ed.ac.uk) to excavate necessary data.