



HOKKAIDO UNIVERSITY

Title	A study on the population genetics of influenza A viruses [an abstract of dissertation and a summary of dissertation review]
Author(s)	Kim, Kiyeon
Degree Grantor	北海道大学
Degree Name	博士(獣医学)
Dissertation Number	甲第12395号
Issue Date	2016-09-26
Doc URL	https://hdl.handle.net/2115/63576
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	doctoral thesis
File Information	Kim_Kiyeon_abstract.pdf, 論文内容の要旨



学位論文内容の要旨
Abstract of the dissertation

博士の専攻分野の名称：博士（獣医学）

氏名：Kiyeon KIM
Name

学位論文題名
The title of the doctoral dissertation

A study on the population genetics of influenza A viruses
(A 型インフルエンザウイルスの集団遺伝学に関する研究)

Understanding the population genetics of an infectious disease pathogen is important for controlling the infectious disease caused by the pathogen. Genetic variation maintained in a microorganism population contains information about the past population and evolutionary dynamics of the microorganism. Tajima's D is a sequence statistics that can test whether or not a population of an organism undergo neutral selection with constant-sized population without subdivision. .

Firstly, I analyzed host-specific and segment-specific Tajima's D trends of influenza A virus through a systematic review using viral sequences registered in the National Center for Biotechnology Information. To avoid bias from viral population subdivision, viral sequences were stratified according to their sampling locations and years. As a result, I obtained a total of 580 datasets each of which consists of nucleotide sequences of influenza A viruses isolated from a single population of hosts at a single sampling site within a single year. By analyzing nucleotide sequences in the datasets, I found that Tajima's D values of viral sequences were different depending on hosts and gene segments. Tajima's D values of viruses isolated from chicken and human samples showed negative, suggesting purifying selection or a rapid population growth of the viruses. The negative Tajima's D values in rapidly growing viral population were also observed in computer simulations. Tajima's D values of PB2, PB1, PA, NP, and M genes of the viruses circulating in wild mallards were close to zero, suggesting that these genes have undergone neutral selection in constant-sized population. On the other hand, Tajima's D values of HA, NA, and NS genes of these viruses were positive, indicating HA, NA and NS have undergone balancing selection in wild mallards. Taken together, these results indicated the existence of unknown factors that maintain viral subtypes in wild mallards.

Secondly, I developed a method that estimates epidemiological parameters of an infectious disease outbreak from Tajima's D values observed in the nucleotide sequences of its causative agent. A total of 265 nucleotide sequences of NA gene of H1N1 influenza A viruses isolated during 2009 pandemic influenza in Buenos Aires were collected from the database

of National Center for Biotechnology Information. Then the time evolution of Tajima's D values was compared with those obtained from computer simulations of a mathematical epidemic model involving viral evolution. By using approximate Bayesian computation, the mode of the basic reproduction number of the 2009 pandemic in Buenos Aires was estimated to be 1.47 (95% credible interval: 1.19 – 4.93). The mode of recovery rate was estimated to be 0.12 (95% credible interval: 0.1 – 0.92). The mode of epidemic peak was estimated to be July 19, 2009 (95% credible interval: 7th July – 30th August). These estimations were consistent with precedent studies by others and reports from World Health Organization. Estimating epidemiological parameters at initial stage of epidemics is important to establish effective control strategies of infectious diseases. I anticipate that this method could be applicable to not only influenza but also other infectious diseases.