

The Establishment of Influenza Virus Bioinformatics System

Abstract:

Core facilities are usually set up to increase both the analysis efficiency and the work quality when a large-scale research project is initiated in many countries around the world. The purpose of these core facilities is to serve the needs of other related projects. This influenza information core facility collects all the available influenza data and information, and provides them to the other team projects after integrating with the RSS technology. Another role of this core facility will be coordinating the interactions among different team projects, and getting the most out of the collaborations. The core will also seek for the capacity of processing large dataset and forming pipeline and quality control for data analysis. In the first year, we have accomplished the following works:

1. completed the collection of the following three influenza sequence data resources through the use of web agent and other tools:
 - Influenza Virus Resource (IRV) of NCBI (National Center for Biotechnology Information), USA
 - Influenza Sequence Database (ISD) of LANL (Los Alamos National Laboratory), USA
 - Influenza data from the Center of Disease Control (CDC) Taiwan
2. completed the integrated influenza sequence database and the graphical user interface (GUI) of its query and analysis functions:
 - BLAST sequence similarity search
 - CLUSTAL W multiple sequence alignment and PHYLIP phylogenetic Genotype analysis
 - Proteotype comparison and phylogenetic analysis
 - Protein structure information through PDB (Protein Data Bank)
3. completed the instant information and news system through the RSS Technology
4. already started to provide bioinformatics analysis service to team members of the other related influenza research projects.

Keywords: influenza information core facility, web agent, integrated information, phylogenetic analysis, molecular evolution, proteotype comparison, protein structure