

Applications of large-scale phylogenetic analysis for research in emerging infectious disease

Speaker:

Daniel Janies Department of Biomedical Informatics, The Ohio State University
Daniel.Janies@osumc.edu

Co-authors:

Pablo Goloboff, Instituto Superior de Entomología, Consejo Nacional de Investigaciones Científicas y Técnicas; Argentina

Diego Pol Museo Paleontologico Egidio Feruglio, Consejo Nacional de Investigaciones Científicas y Técnicas; Argentina

Emerging infectious diseases and organisms present critical issues for public health. Evolution and geographic spread of emerging infectious diseases strains are now being addressed via the collection of large amounts of genomic data. However, our ability to derive information from large sequence datasets lags far behind their acquisition. Our work is directed at addressing this gap between data collection and analysis. In this talk we discuss application of phylogenetic analysis to study multiple co-circulating lineages, track host shifts, and measure surveillance quality among large datasets of influenza A genomes. We find that there are multiple co-circulating lineages within an antigenic subtype. Furthermore, the dominant strain of influenza is not easily predictable from the relative abundance and geographic distribution of viral strains in previous years. Using a longitudinal study of isolates ranging back to 1902 we find frequent host shifts between avian and human populations and many avenues of viral traffic among hosts. Using metrics designed to measure the quality of the fossil record we find that overall influenza surveillance is good - however the quality varies over time, geography, and in different strains.