Genomic Analysis and Geographic Visualization of H5N1 and SARS-CoV

Andrew W. Hill, BS¹, Boyan Alexandrov², Robert P. Guralnick, PhD¹, Daniel Janies, PhD²

¹Department of Ecology and Evolutionary Biology and Museum of Natural History, University of Colorado, Boulder CO 80309;
²Department of Biomedical Informatics, Ohio State University Medical Center, Columbus OH 43210

Abstract

Emerging infectious diseases and organisms present critical issues of national security, public health, and economic welfare. We still understand little about the zoonotic potential of many viruses. To this end, we are developing novel database tools to manage large comparative genomic datasets. These tools add value because they allow us to summarize the direction, frequency, and order of genomic changes. We will perform numerous real-world tests with our tools with both Avian Influenza and Coronaviruses.

Introduction and Statement of Problem

As demonstrated by the coordinated response to SARS and avian influenza, diseases are being addressed via genomic sequencing. However, our ability to derive information from large comparative genomic datasets lags far behind acquisition. Knowledge about genomic as well as temporal and geographic correlates of past transmission and pandemic events will be valuable to predict how, where, and when various strains of influenza may become pandemic and design effective vaccines to novel strains.

Project Description

To meet this need we have developed a tool for analysis and visualization of disease evolution over time and space. By taking full genome phylogenetic analysis of infectious disease, our tool can project lineage movement over Google Earth. The movement can then be analyzed in not only its genomic context, but in its temporal and geographic context as well. The phylogenetic hypothesis can also be coupled with character state optimizations, such as host type, giving us the ability to view movements of not only genotypes but phenotypes also (Figure 1).

Figure 1. Geographic analysis of spread of H5N1 avian influenza based on the projection of a phylogeny into a virtual globe. Red and white colored branches indicate key genotypes that confer increased replication of the virus in mammals.

Our biomedical aims have been to study the molecular changes associated with zoonotic events from a whole genome perspective. Therefore, the first study involved full genome analysis of H5N1 avian influenza[1]. Through our integrated approach the we found that particular avian host taxa responsible for the spread H5N1 are largely context-specific temporally and geographically. Our results indicate that H5N1 has entered different regions of the world via different host pathways and diverse host types. We are extending the application of these tools to better understand the genomic, temporal, geographic and phenotypic characteristics that mediate the movements of the coronaviruses and specifically the unexpected outbreak and spread of SARS-CoV.

References