

## Project 5

### Prediction of Host Specificity based on Polymerase PB2

#### Background

Influenza A viruses (IAVs) are zoonotic pathogens that can infect a broad range of animals from birds to pigs and humans. Most of these species have wild birds as their natural hosts. IAVs are usually adapted and relatively restricted to a single host but occasionally the virus can jump and adapt to a new host species. The interspecies transmission requires that IAVs adapt to the new host and the whole process is facilitated by their high mutation rates. This can result in epidemics and pandemics with severe consequences for both human and animal life. Given our poor understanding of the host-adaptation process of the virus, which can be a major factor for such epidemics and pandemics, it is very hard to predict the type of the virus that will cause the coming outbreaks. The hemagglutinin protein plays a crucial part in defining the adaptation of the virus to different hosts since it binds to the receptor providing the entry into host cells. The avian strains of the IAVs are known to prefer a receptor with  $\alpha 2,3$ -sialic acid linkages while the human strains prefer a receptor with  $\alpha 2,6$ -sialic acid linkages. However, other proteins subunits have also previously been shown to play a role in the adaptation of IAVs to different hosts.

#### Aim of the study

The goal of this study is to find markers associated with human and avian hosts in Polymerase PB2 proteins.

#### Description of the dataset

The data is decision tables made from the multiple sequence alignments of the proteins. The decision classes are either Avian or Human indicating the hosts in which the viruses were found.

**Ref:** Khaliq, Zeeshan, et al. "A complete map of potential pathogenicity markers of avian influenza virus subtype H5 predicted from 11 expressed proteins." BMC microbiology 15.1 (2015): 128.

<https://bmcmicrobiol.biomedcentral.com/track/pdf/10.1186/s12866-015-0465-x>