Online Analytics
for Host-Pathogen Protein Interactions Analysis

Student: Ng Teng Ann   Supervisor: Dr Kwoh Chee Keong
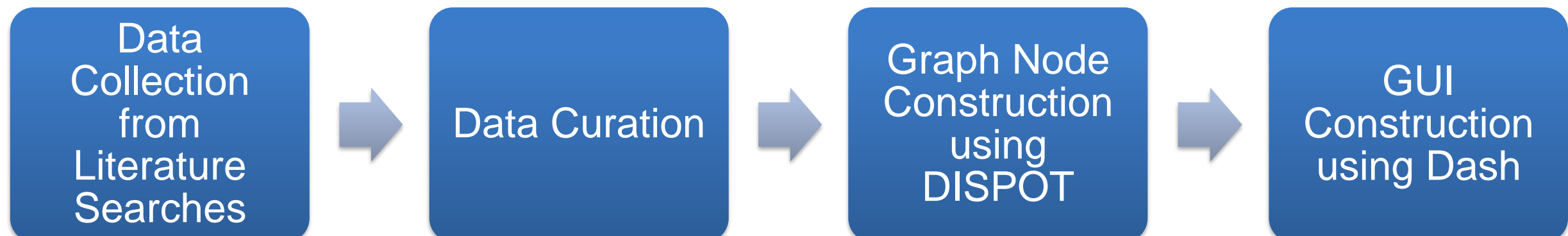
Introduction

Influenza A virus (IAV) is a significant danger to global human health and life. It is critical to have a deeper knowledge of the virulence factors responsible for IAV infections to counteract potential outbreaks.
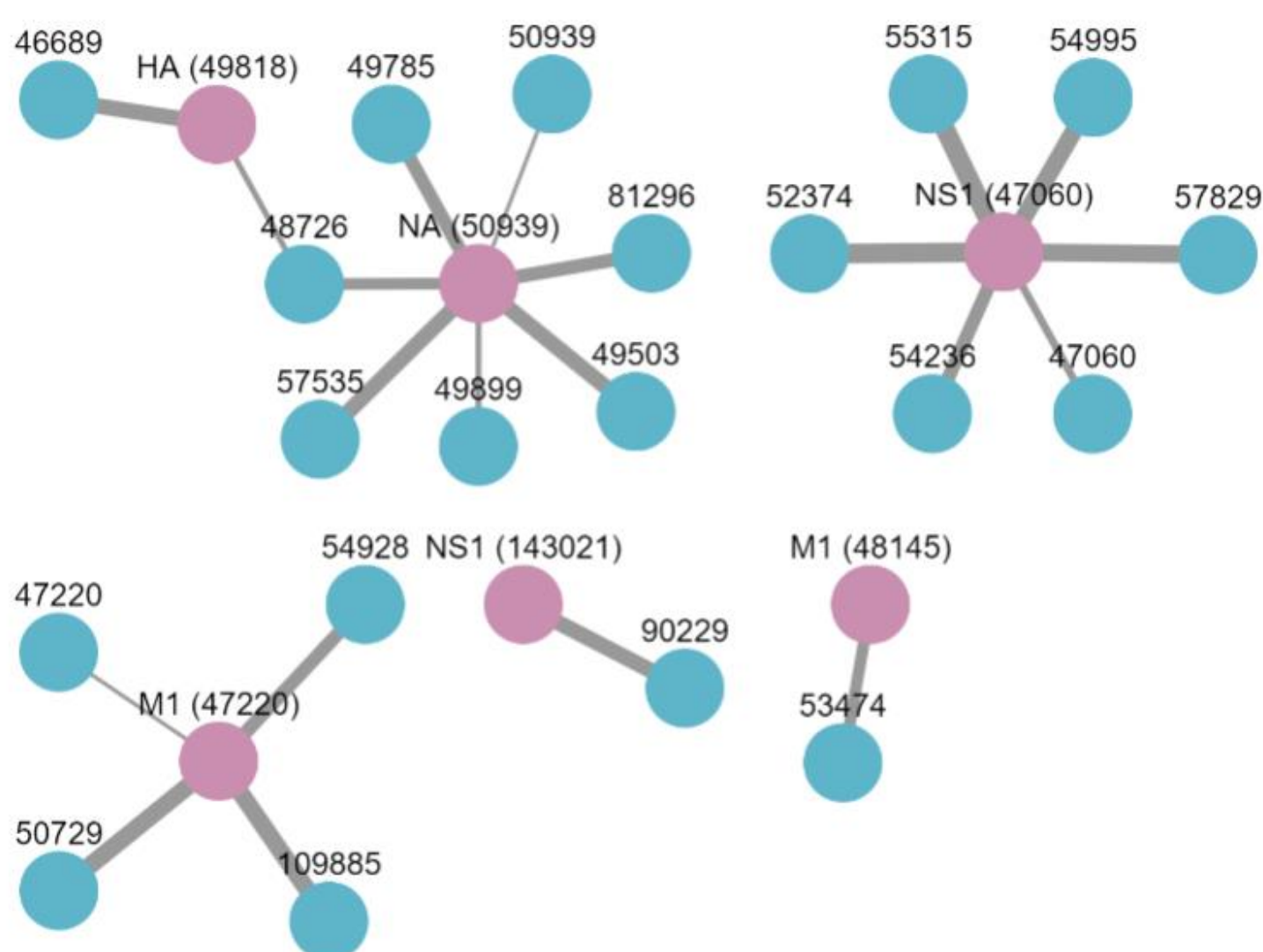
Objectives

- Conduct a meta-analysis, to deliver more accurate insights and offer adequate confidence on the viral factors accountable for the extreme harmfulness of IAV infections.
- Uncover specific domains from IAV and mouse proteins involved in host-pathogen protein interactions.
- Present interacting IAV-mouse protein domains as a network of nodes.

Methodology

Data Collection from Literature Searches → Data Curation → Graph Node Construction using DISPOT → GUI Construction using Dash

Network of Nodes

Pink nodes represents IAV protein segments,
Blue nodes represents mouse protein domains.

Summary

- Discovered that HA, NA, M1 and NS1 are the key segments involved in host-pathogen interaction
- Strongest interactions surround the S15/NS1 RNA-binding domain. (i.e., NS1 (47060) node)

Limitation

- Currently, it is not possible to differentiate IAV genome across different strains by using only the protein domains.

Future Work

- Machine learning techniques can be adopted to conduct detailed studies on protein embeddings and discover the co-relation between the various amino acid sequences.