Title

Modeling and Comparative Analysis of SARS-CoV-2 Protein Interaction Networks Using Hierarchical Bayesian Models

Authors

Purvi Praful Kanfade[1,] Lisa Bramer[2], Moses Obiri[2], Erik VonKaenel[2]
1Bioinformatics Program, The University of Texas at El Paso, El Paso, TX
2Biological Sciences Division, Pacific Northwest National Laboratory (PNNL), Richland, WA

Abstract

SARS-CoV-2, the virus responsible for the COVID-19 pandemic, continues to evolve, resulting in distinct strains with varying pathogenic and epidemiological profiles. Understanding how variations in protein-protein interaction (PPI) networks across these strains influence viral behavior is critical for developing targeted therapies. This project compares the PPI networks of two SARS-CoV-2 strains, one from Washington (USA) and the other from Italy, each with its own unique set of protein interactions. By comparing these networks, we aim to reveal how differences in protein interactions contribute to the distinct characteristics of each strain.

The analysis began by preprocessing SARS-CoV-2 proteomics data to construct binary adjacency matrices that reflect the presence or absence of protein interactions for each strain. These matrices were transformed into network models, allowing us to identify differentially interacting proteins. Key network metrics, including degree centrality, betweenness centrality, and hubness scores, were calculated to evaluate the influence of key proteins within the networks. In addition, local and global network metrics were used to assess network stability and structure. Community detection methods, such as Louvain and Infomap, identified clusters of interacting proteins, highlighting essential hub proteins that play significant roles in viral replication and pathogenicity.

Our results suggest that proteins with high hubness and spreading scores in both strains may serve as potential antiviral targets, as disruptions in these proteins could alter network behavior. This comparative analysis highlights the importance of integrating machine learning techniques into biological network analysis to improve understanding of viral mechanisms and inform future therapeutic strategies.