GENOME NETWORK ANALYSIS

Pinky Sindhu [1], Denee Lawrence [2], Hannah Lino [3], Dr. Catherine Putonti, Dr. Jason Shapiro

1. Department of Information Technology 2. School of Nursing 3. Department of Biology

Loyola University Chicago

Preparing people to lead extraordinary lives

INTRODUCTION

The genome network analysis project is about creating a tool that aids in analyzing complex genome networks.

The problems with analyzing such complex genome network is that in an absence of any computational tool it is always error prone and tedious. The homologous complex component in a cluster of genomes is always difficult to detect manually.

This research is intended to help visualize data presented as a similarity matrix so researchers would be able to efficiently compare and contrast genes shared among genomes. Specifically, for comparing large genome networks like viral communities.

THE VIRAL COMMUNITY IN GENOME NETWORK REPRESENTATION
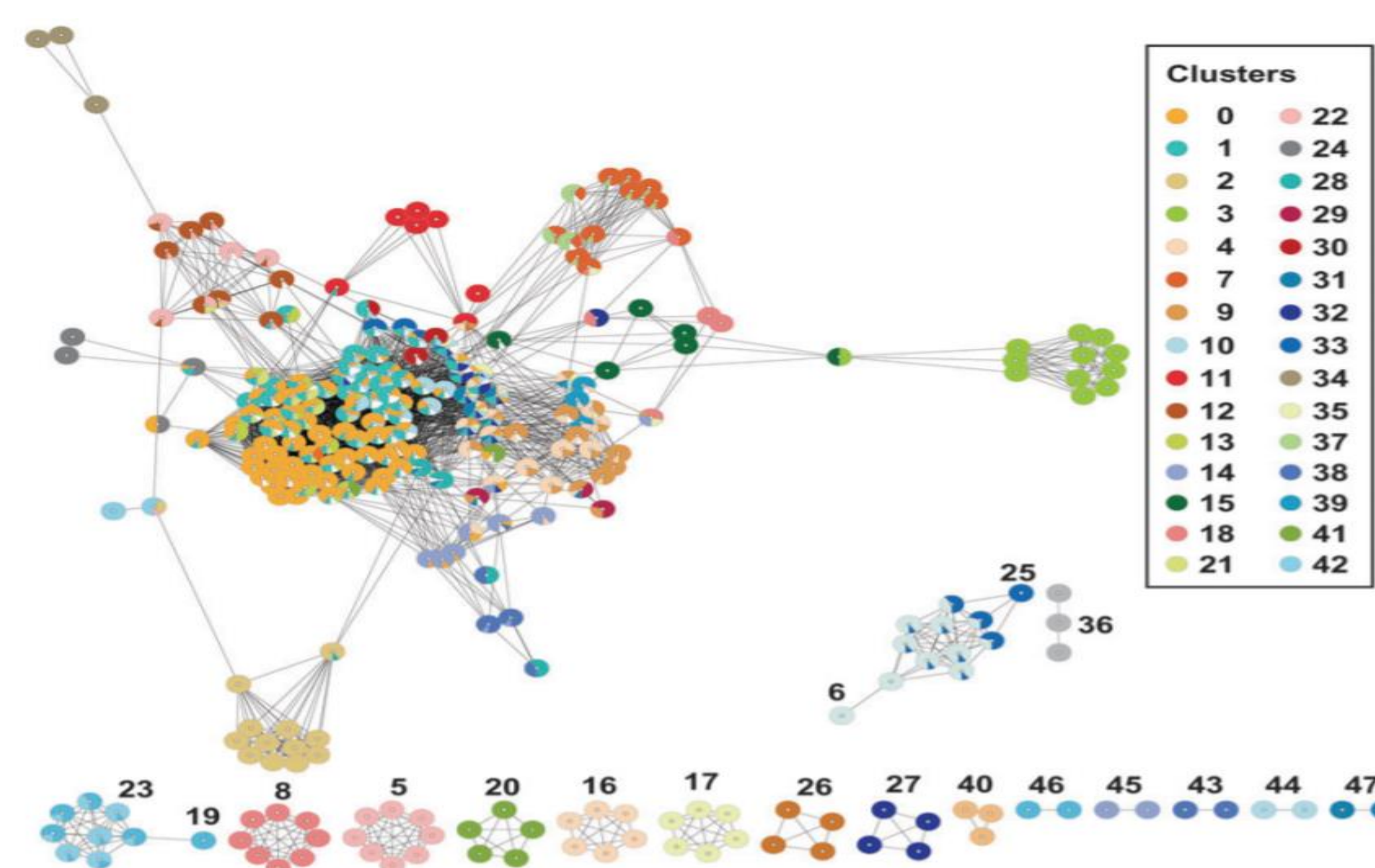
Each node in the representation denotes a virus genome in the cluster. This shows multiple genomes are combined to form several clusters called a network. Also, the multiple clusters are connected to each other through edges between nodes.

MOTIVATION

Viruses that infect bacteria are the most diverse and abundant organisms on this planet. Their genomes are composed in a very complex way, because they exchange genes continuously when they infect the same bacteria.

This is why genome networks are important in comparing and contrasting viral genomes. The study will be helpful for researchers to easily identify the homologous components between different viral communities.

Furthermore, it will allow researchers to identify the genes of the organisms at which these homologous components are exhibited.

NETWORK REPRESENTATIONS

Actual Representation and Alternate Representation

This diagram depicts a toy problem where the nodes represent genomes and the edges represent how closely related the nodes are to each other.
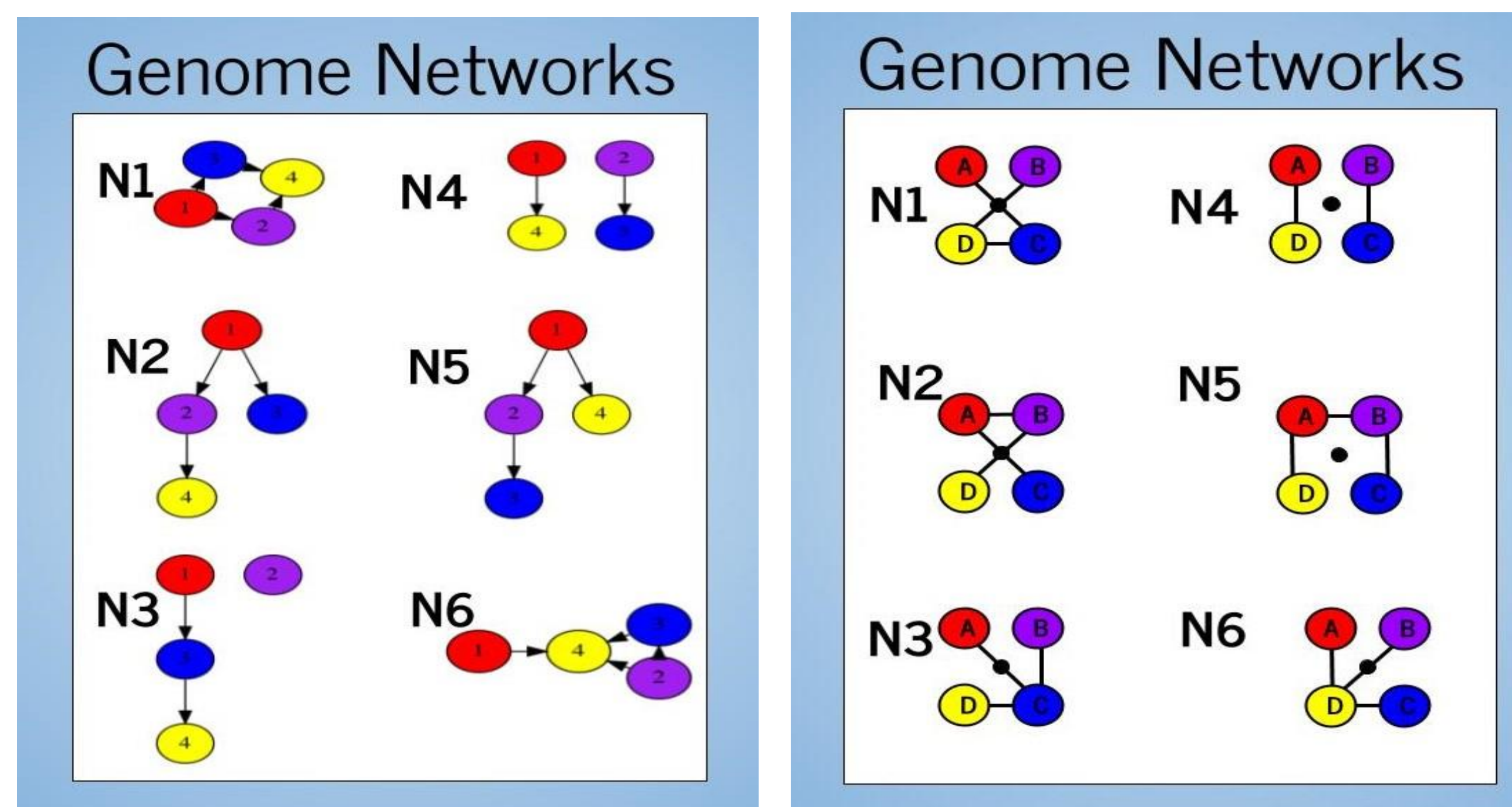
ANALYSIS APPROACH

This study uses the SNAP library (Stanford Network Analysis Project) in Python in order to identify relationships between homologous genes and edges within two or more networks. As a team, we have devised a dimensionality reduction algorithm to visualize the relationships between networks on a 2D plot.
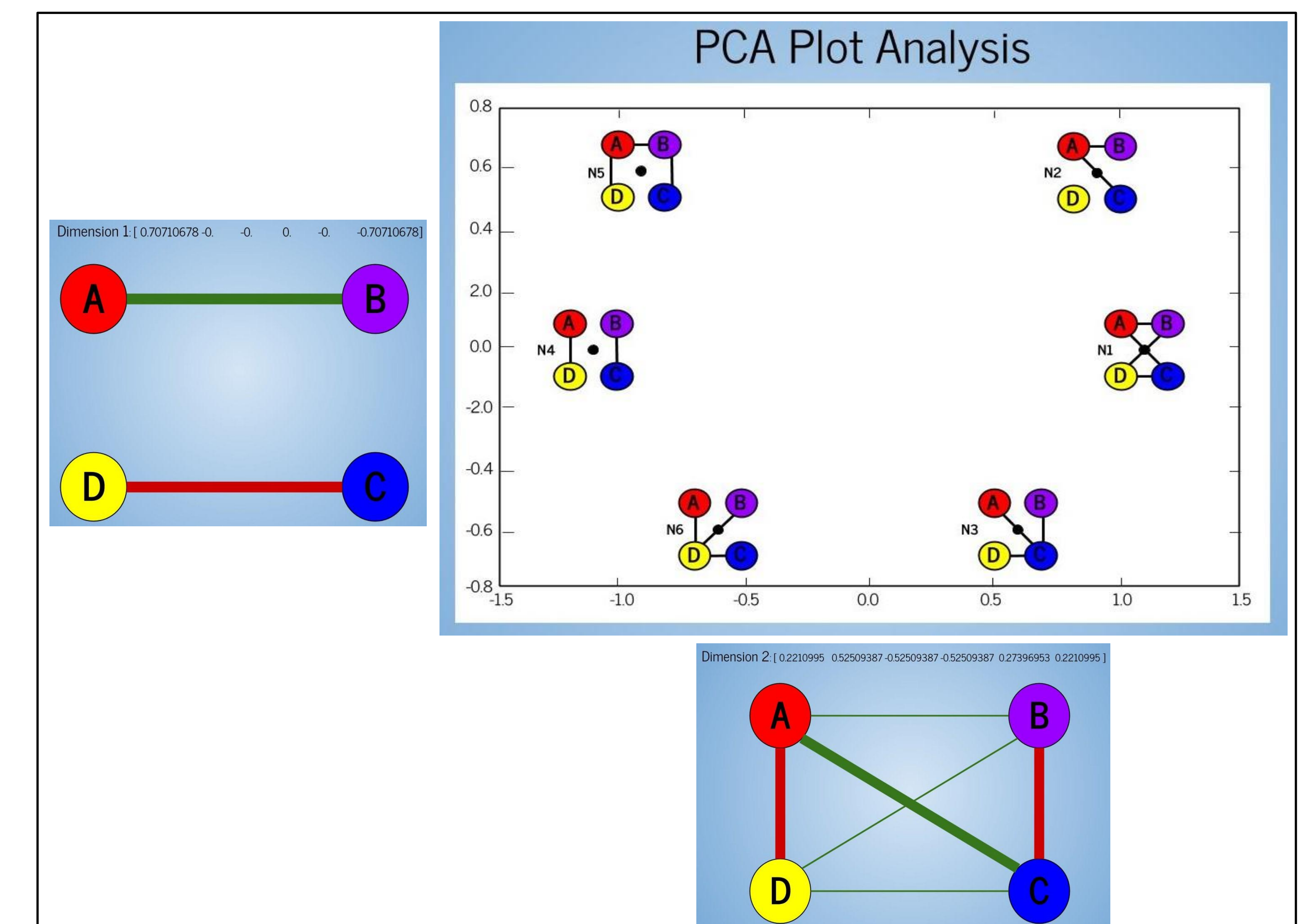
NETWORK ANALYSIS

Representation of Networks in a Matrix

	AB	AC	AD	BC	BD	DC
Network 1	1	1	0	0	1	1
Network 2	1	1	0	0	1	0
Network 3	0	1	0	0	0	1
Network 4	0	0	1	1	0	0
Network 5	1	0	1	1	0	0
Network 6	0	0	1	1	1	1

Comparison of Networks against each other in a Matrix

Networks	1	2	3	4	5	6
1	1	5/6	4/6	0	1/6	2/6
2		1	1/6	1/6	2/6	1/6
3			1	2/6	1/6	2/6
4				1	5/6	4/6
5					1	3/6
6						1

A Visual Plot for Network Relationships

ACKNOWLEDGEMENT

REFERENCE

- Who we are for ? Researcher- https://www.digital- science.com/researchers/
- Reticulate Representation of Evolutionary and Functional Relationships between Phage Genomes- https://academic.oup.com/mbe/article/25/4/762/1269810/Reticulate-Representation-of-Evolutionary-and