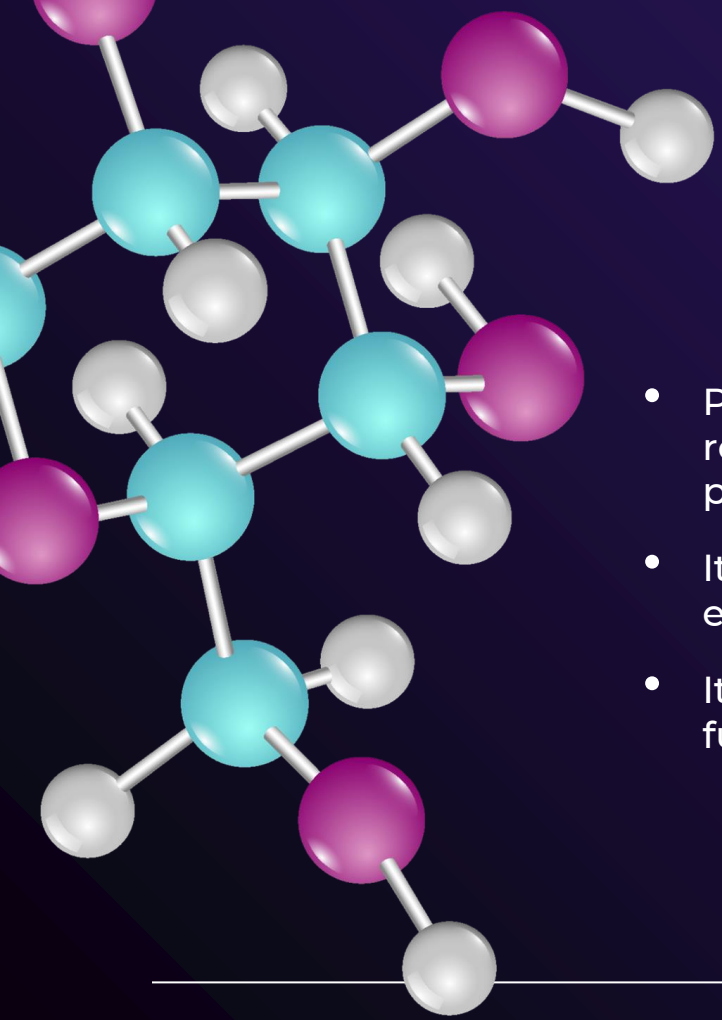
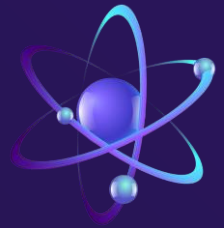
Protein-Protein Interactions Biological Network



INTRODUCTION

- Protein-protein interaction network plays a critical role in various biological processes, such as signaling pathways and networks.
- It describes the interplay between the biomolecules encoded by genes.
- It allows us to understand the complexities of cellular function and even predict potential therapeutics.

TABLE OF **CONTENTS**



01 Our Network Data

Biological Network directed interactome

02 Tools

Used packages for network analysis

03 Methods and Results

Our analysis steps and results

04 Conclusion

Overall findings of analysis

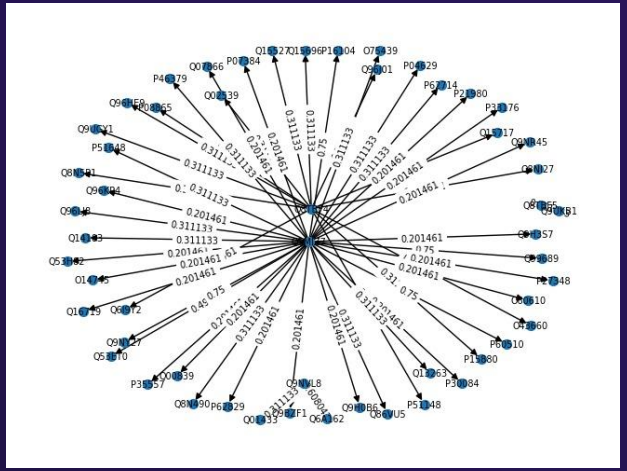
01

Network Data

You can enter a subtitle here if you need it

Network Data

PathLinker_2018_human-ppi-weighted-cap0_75
interactome txt file



Weighted Directed interactome data: each interaction starts from the tail node to the head node.

- Tail : start protein node.
- Head : end protein node.
- Edge weight : Interaction confidence.
- Edge Type: Method used to identify this interaction.

Each protein is represented by its UniProt ID.

PathLinker_2018_human-ppi-weighted-cap0_75.txt - Notepad

File Edit Format View Help

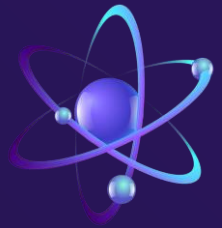
#tail	head	edge_weight	edge_type
Q8TBF5	Q9UKB1	3.11133e-01	MI:0004 (affinity chromatography technology)
Q8TBF4	Q15717	3.11133e-01	MI:0004 (affinity chromatography technology)
Q8TBF4	P08865	3.11133e-01	MI:0004 (affinity chromatography technology)
Q8TBF4	Q02539	3.11133e-01	MI:0004 (affinity chromatography technology)
Q8TBF4	Q96J01	2.01461e-01	MI:0401 (biochemical)

02

Tools

Used packages for network analysis

Python Packages



NetworkX

Main package used for creation, manipulation, and study of the structure, and functions of complex networks.

NumPy

Powerful library for scientific computing in Python.

Pandas

Python's widely used data manipulation and analysis library.

UniProt

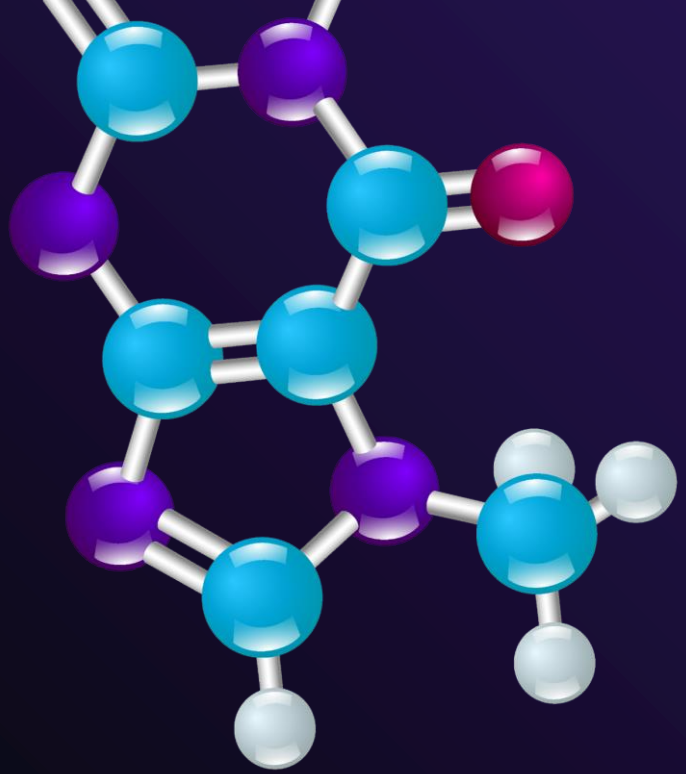
Freely accessible database of protein sequence and functional information

Matplotlib & Seaborn

Python data visualization libraries. Used it for visualizing our figures

StringIO

Used for reading UniProt gene information from a given protein.



03

Methods and Results

Network analysis steps and results

Methods and Results

- Reading Data from Text File .
- Extracting proteins' tails, heads, and edge weights.
- Data Exploration with Pandas.

- Result :

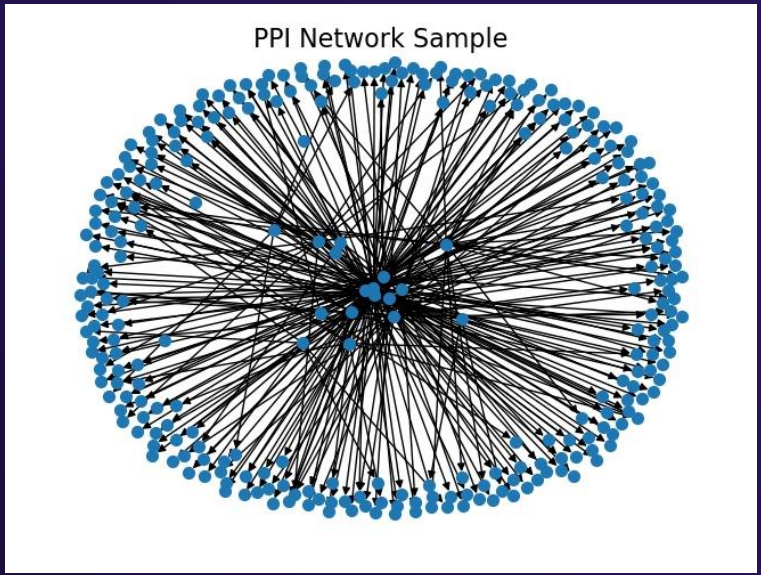
  Our network has : 17082 unique tails & 17095 unique heads.

- As the number of heads is very close to the number of tails, we may assume that the network is balanced, and this can lead to several insights.

  - Network stability
  - Symmetrical interactions
  - Functionality
  - Lack of dominant hub proteins.
  - Redundancy

Methods and Results

- Biological network graph construction of PPIs interactome data.
- By using NetworkX package
- Plot a graph of a sample of 300 edges out of 612,516 edges.



Methods and Results

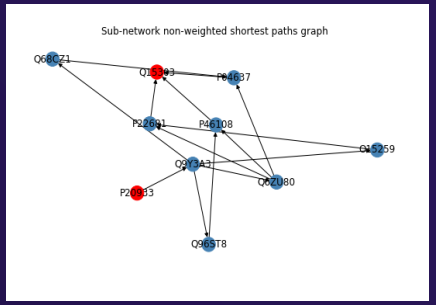
- Listing the acyclic shortest path(s) between two given proteins.

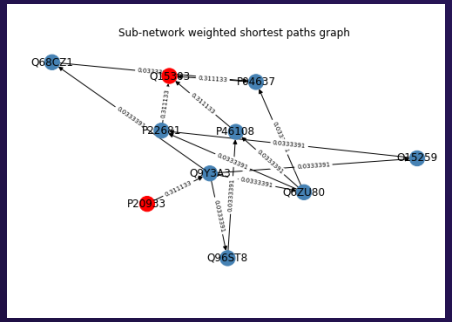
- The same 2 proteins may Have different paths with same weight

acyclic_shortest_path.txt - Notepad

File Edit Format View Help

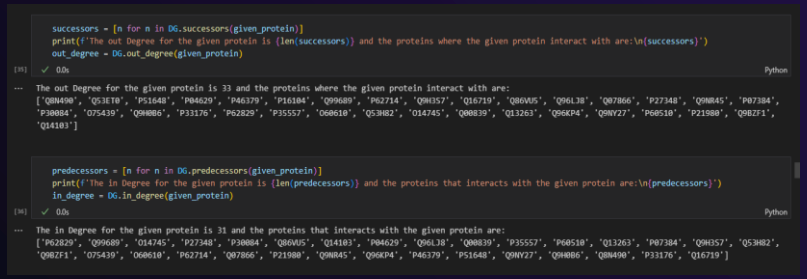
```
Shortest paths between P20933 and Q15303 Proteins

Number of shortest paths :6
Total path score : 0.688944

Path 0 :
P20933 > Q9Y3A3 > Q68CZ1 > P04637 > Q15303

        Tail    Head    Edge_weight
        P20933  Q9Y3A3  0.311133
        Q9Y3A3  Q68CZ1  0.0333391
        Q68CZ1  P04637  0.0333391
        P04637  Q15303  0.311133

Path 1 :
P20933 > Q9Y3A3 > Q6ZU80 > P04637 > Q15303

        Tail    Head    Edge_weight
        P20933  Q9Y3A3  0.311133
        Q9Y3A3  Q6ZU80  0.0333391
        Q6ZU80  P04637  0.0333391
        P04637  Q15303  0.311133
```





Methods and Results

- Listing all the directly connected proteins to given protein.
- Calculating the degree of this protein.
- When in-degree is close to out-degree that can result in stability of overall network topological.
- More stable network may decrease the probability of disease infection by providing more robust connections between proteins, reducing the risk of network failure and disease spread.



```
successors = [n for n in DG.successors(given_protein)]
print(f'The out Degree for the given protein is {len(successors)} and the proteins where the given protein interact with are:\n{successors}')
out_degree = DG.out_degree(given_protein)
```

[35] ✓ 0.0s Python

```
The out Degree for the given protein is 33 and the proteins where the given protein interact with are:
['Q8N490', 'Q53ET0', 'P51648', 'P04629', 'P46379', 'P16104', 'Q99689', 'P62714', 'Q9H3S7', 'Q16719', 'Q86VU5', 'Q96LJ8', 'Q07866', 'P27348', 'Q9NR45', 'P07384',
'P30084', 'O75439', 'Q9H0B6', 'P33176', 'P62829', 'P35557', 'O60610', 'Q53H82', 'O14745', 'Q00839', 'Q13263', 'Q96KP4', 'Q9NY27', 'P60510', 'P21980', 'Q9BZF1',
'Q14103']
```

```
predecessors = [n for n in DG.predecessors(given_protein)]
print(f'The in Degree for the given protein is {len(predecessors)} and the proteins that interacts with the given protein are:\n{predecessors}')
in_degree = DG.in_degree(given_protein)
```

[36] ✓ 0.0s Python

```
The in Degree for the given protein is 31 and the proteins that interacts with the given protein are:
['P62829', 'Q99689', 'O14745', 'P27348', 'P30084', 'Q86VU5', 'Q14103', 'P04629', 'Q96LJ8', 'Q00839', 'P35557', 'P60510', 'Q13263', 'P07384', 'Q9H3S7', 'Q53H82',
'Q9BZF1', 'O75439', 'O60610', 'P62714', 'Q07866', 'P21980', 'Q9NR45', 'Q96KP4', 'P46379', 'P51648', 'Q9NY27', 'Q9H0B6', 'Q8N490', 'P33176', 'Q16719']
```

Q5MIZ7_predecessors.txt - Notepad

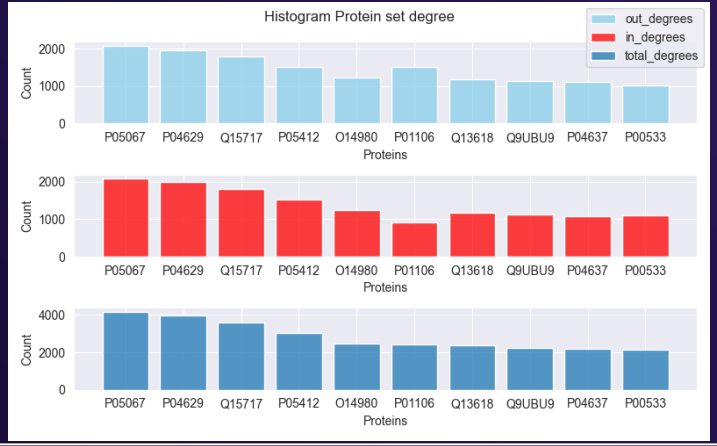
File Edit Format View Help

Q5MIZ7 In Degree = 31

Tail	Head	Edge_weight
P62829	Q5MIZ7	0.201461
Q99689	Q5MIZ7	0.75
O14745	Q5MIZ7	0.201461
P27348	Q5MIZ7	0.201461
P30084	Q5MIZ7	0.201461
Q86VU5	Q5MIZ7	0.311133
Q14103	Q5MIZ7	0.311133
P04629	Q5MIZ7	0.311133
Q96LJ8	Q5MIZ7	0.311133
Q00839	Q5MIZ7	0.201461

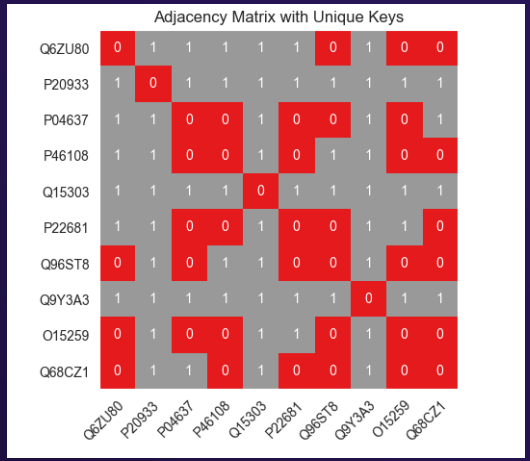
Methods and Results

- Listing degrees of given set of proteins and drawing Histogram
- We can conclude that both `P05067` and `P04629` are hub proteins and may be key scaffold proteins.
- The importance of hub proteins can be attributed to several factors, including:
  - Modularity
  - Robustness
  - Functionality
  - Evolution
  - Disease



Methods and Results

- Convert graph as an unweighted graph and save it using the adjacency matrix method
- We constructed this adjacency matrix from the last network we created of shortest paths where “P20933” is the tail, “Q15303” is the head, and “0.6889442” is the weight cost.
- Q9Y3A3 is the hub protein of this sub-network



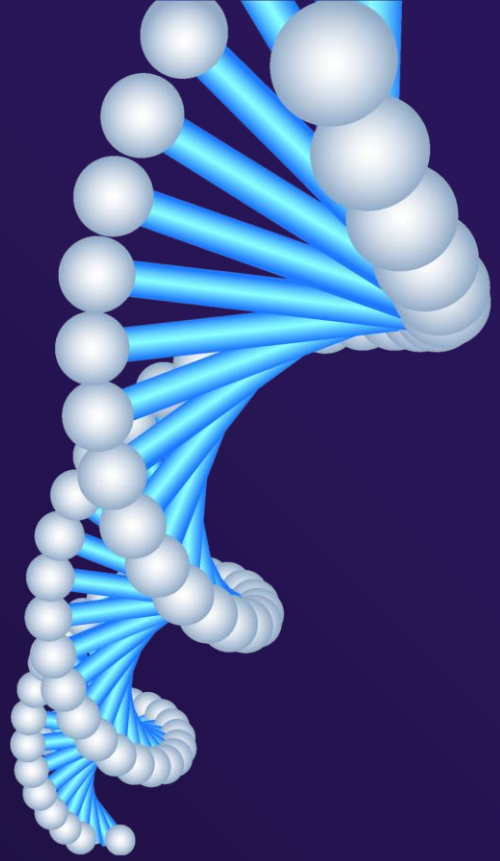
Methods and Results

- Conversion map between the protein UniProt ID and its gene name.
- As we concluded before that we have two hub proteins, in this section we retrieved their corresponding genes.

Entry	Entry Name	Reviewed	Protein names	Gene Names	Organism	Length
P05067	A4_HUMAN	reviewed	Amyloid-beta precursor protein (APP) (ABPP) (A...	APP A4 AD1	Homo sapiens (Human)	770
P04629	NTRK1_HUMAN	reviewed	High affinity nerve growth factor receptor (EC...	NTRK1 MTC TRK TRKA	Homo sapiens (Human)	796

04

Conclusion



Conclusion

- The analysis of protein-protein interactions is a critical step toward gaining a comprehensive understanding of biological systems and the formation of complex biological networks provided valuable insights into the organization and regulation of it as the power of computational analysis in understanding the complexities of cellular function and predicting potential therapeutics

OUR **TEAM**

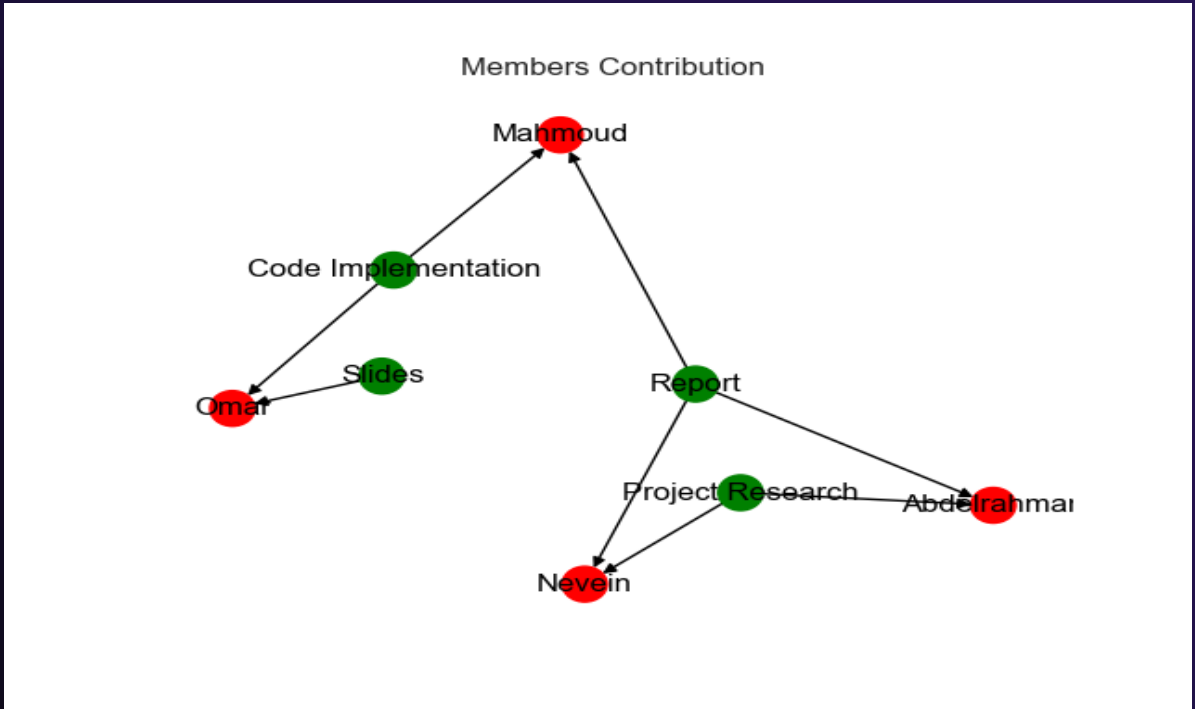
Abdelrahman Ali Farouk

Nevein Mohamed Ayman

Mahmoud Yaser Salman

Omar Saad EL-Gharbawy

Members Contribution



THANKS