BACKGROUND

Alzheimer’s disease

Graph theory

Diseases and Networks

Essential proteins play critical roles in cell processes such as development and survival.

Biological network analysis

What are biological networks. What they signify. How they can be used.

Protein-Protein Interactions

What are proteins.

Why are proteins and their interactions important: book page 11

Little on how they are found

PPI databases, how they are built.annotated etc

Network Analysis

Important concepts:

Proteins.

Biological processes

Annotations, GO

BACKGROUND

1. Networks in biology

At every level of biology, biological elements interact with other biological and non-biological elements and form complex systems. In the study of biology, networks are a way to models such systems. They can be used to model biological systems at almost every level, from microscopic level with metabolic networks, gene regulatory networks, protein-protein interaction networks, to intercellular networks such as neural networks, and even higher organism level such as ecological networks.

//Little on how biological data is collected

The study of such biological networks is a vital part of understanding complex biological activities and mechanisms(3)(4). In recent decades the field of computational biology, which focuses more on holistic approach of biological systems, biological networks have been increasing studied as a means of modelling biological systems data and understand the functioning of the system. Network analysis has been used for several applications such as Protein function prediction(cite), disease gene prioritization(cite), identification of essential proteins/genes(cite) etc.

Protein-Protein Interaction Networks

Protein-Protein interaction networks(PPINs), are networks that model the interactions between proteins. They are modelled as undirected network graphs, where the nodes represent the proteins, and the edges represent the interactions between the proteins. The edges/interactions may or may not have weights, where the weights assign the interaction with a confidence of the interaction(6).

PPIN analysis has been used to study molecular evolution, assignment of functions to new proteins, identification of diseases(cite), identify disease-related subnetworks(cite), disease classification based on networks, network-based Genome wide association studies(cite) etc(5)(cite).

Networks analysis for diseases

One of the biggest applications of network theory for the analysis of biological networks is in medicine. It has been used for drug target identification, designing of effective strategies for the treatment of various diseases, identification of disease genes, development of precision therapeutics etc.

The traditional approach to drugs development, involved a linear causality module, where the genes responsible for a disease are targeted(1). Many diseases however, such as coronary heart disease, type 2 diabetes etc, do not originate from a significant mutation to a single gene, but rather from smaller defects in multiple genes. This makes the mutations different to identify and interpret. Usage of networks allows for a better representation of such diseases, and can lead to a better understanding of the disease causing genes and pathways, which in turn can offer targets for drug development(2)(7). This is done using a number of different analysis methods, such as topological, subnetwork analysis etc( discussed further in section about PPI network properties).

// add the examples of papers that used network analysis for diseases

2. Graph/Network Analysis for biological PPI networks

Network topology

The topology of a network describes the arrangement of nodes and edges in a networks. It has been shown as a method of finding influential proteins. Multiple studies have found that proteins influential proteins in the network are more essential for survival(cite), play critical roles in cell processes and may be linked with diseases. Network topological analysis such as centrality analysis, clustering are commonly used.

Centrality measures

Centrality measures are used to find how ‘central’ a node is in the network. They can be used to find influential nodes in the networks.

Degree

The degree of a node is the number of edges that are directly connected to/interact with the node. It a basic centrality measure, that measures how ‘popular’ a node is, but does not take into account the location of the node in the graph. //FORMULA

Betweenness

The betweenness of a edge/node is a measure of importance of a node in connecting different nodes in a network. It is the number of shortest paths between any two node, that pass through that edge/node. A node with high betweenness would play an important role in communication between other nodes(cite). //FORMULA

Closeness

Closeness of a node shows how close a node is from every other node in the networks. It is calculated by taking the inverse of the average length of the shortest distances of the node with every other node.

**Bridgeness**

Page rank

Modularity

Clustering and enrichment

clustering methods and algorithms

clustering validation

Enrichment