Protein Metabolite Interaction Network

Dissertation submitted in partial fulfillment of the requirement for the degree of

BACHELOR OF TECHNOLOGY

IN

BIOINFORMATICS

By

Kritika Sharma

(Roll No: 171513)

Under the Guidance of

Dr. Garlapati Vijay Kumar

JAYPEE UNIVERSITY OF INFORMATIONTECHNOLOGY, WAKNAGHAT

DEPT. OF BIOTECHNOLOGY AND BIOINFORMATICS

HP-173234, INDIA

May 2021

SUPERVISOR'S CERTIFICATE

This is to certify that the work reported in the B. Tech. thesis entitled **"Protein Metabolite Interaction Network"**, submitted by Kritika Sharma (171513) at Jaypee University of Information Technology, Waknaghat, India, is a bonafide record of his original work carried out under my supervision. This work has not been submitted elsewhere for any other degree or diploma.

(Signature of Supervisor)

Name of Supervisor -Dr. Garlapati Vijay Kumar

Designation -Associate Professor

Department of Biotechnology and Bioinformatics

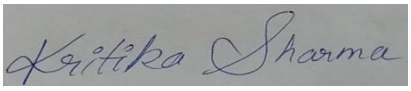
Jaypee University of Information Technology

Waknaghat, Distt-Solan, H.P. – 173 234

Date :18/06/2021

DECLARATION

I hereby declare that the work reported in the B. Tech. thesis entitled **"Protein Metabolite Interaction Network"** submitted at Jaypee University of Information Technology, Waknaghat, India, is an authentic record of my work carried out under the supervision of Dr. Garlapati Vijay Kumar, Dept. of Biotechnology and Bioinformatics, JUIT, Waknaghat, HP-173234, India. I have not submitted this work elsewhere for any other degree or diploma.



(**Kritika Sharma,171513**)

Department of Biotechnology and Bioinformatics,

Jaypee University of Information and Technology, Waknaghat, Solan, Himachal Pradesh, India.

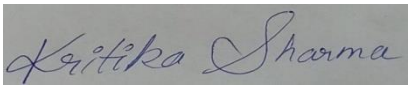
Date: 18/06/2021

ACKNOWLEDGMENTS

It’s a matter of great pleasure and privilege to be able to present this document. However, this would not have been possible without the consistent support and generous help of my mentor. I would like to extend our sincere thanks to him.

I am grateful to Dr. Garlapati Vijay Kumar for his guidance and constant supervision as well as for providing necessary information regarding the project & also for his support which helped us throughout this journey.

I will strive to use the gained skills and knowledge in the best possible way and I will continue to work on their improvement to attain desired career objectives.



(**Kritika Sharma,171513**)

TABLE OF CONTENTS

List of Tables

List of Figures

ABSTRACT

We have built up a metabolite data set by separating and coordinating data from a few public sources. By questioning this data set, Metscape permits clients to follow the associations among metabolites and proteins and show compound structures just as data for responses, chemicals, nucleic acids, and pathways. Applying this pathway filter, anyone can easily make sub-networks that consist of different compounds and reactions from any given pathway. It permits the users to transfer the test data and envision and investigate compound organizations over the long run or test conditions. Metabolites involve the most significant part of particles in cells. Yet, our insight into the protein-metabolite interactome severely lacks compared to the comprehension of protein-protein or protein-DNA interactomes. Our information shows instances of metabolite-initiated redesigning of protein buildings. Our report uncovers underlying standards of compound correspondence, reveals insight into the pre-dominance and components of protein indiscrimination, and empowers the extraction of boundaries on a quantitative level to bind metabolite.

Keywords: Protein-metabolite Interaction, Interaction Networks, Cytoscspe, Metscape, HMDB, KEGG, Biochemical pathways

Chapter 1

INTRODUCTION

==========

Large biomolecule, consisting of a long chain of amino acid units are termed as proteins. They have to carry out huge responsibilities inside living beings, which include the catalysis of metabolic responses, replication of DNA, reacting to touch, establishing structures of creatures and cells, and translocating particles from 1 region to the next.

A single protein molecule normally consists of one lengthy chain of amino acids. Short chain amino acids, with under twenty-thirty units, are rarely considered as protein molecules, are only termed as peptides, or in some cases oligo peptides.All the amino acid residues are bonded with each other through peptide bonds with the adjacent amino acid residues.

Metabolomics is an upcoming and developing stream of –omics research which is concerned with the efficient recognition, quantification, and classification of small metabolites which are present in the metabolome. [2]

The term metabolite usually refers to the small molecules present in the body. These molecules are responsible for different roles, which are to serve as fuel for basic activities, provide a structure to the cell, flagging, stimulate and inhibit consequences enzymatic reactions, and the catalysis reaction for themselves (for the most part as a cofactor to a protein).

Metabolites from chemical compounds, regardless of whether inalienable or drug, structure as a feature of the common biochemical cycle of debasing and wiping out the other compounds. The pace of degradation of a compound is a significant determinant of the span and force of its activity. Seeing how drug mixes are utilized and the expected symptoms of their metabolites is a significant piece of medication revelation.

There are several million protein molecules which are present within a single cell of bacteria [3] but the numbers of metabolites surpass proteins by almost a hundred times [4]. Past their functions as compounds at intermediate level during the transformation of metabolites, they

also create straightforward prompts or set off indirect adaptive feedbacks. Dietary states, stress, and natural conditions impact the intracellular degrees of countless metabolites, and the signs these particles intercede are sent through a progression of sub-atomic occasions, including the binding of metabolites to proteins. Several functionally important metabolite and protein interplay are depicted in a study by Chubukov et al., 2014. [5]

Protein structures are altered due to metabolite-protein interactions on a local or a global level [6-7]. The best-portrayed communications include the metabolites serving as a binding agents for the active site of proteins in the form of substrates, cofactors, or end product of an enzymatic response. They also work as controllers for allosteric sites, by tying themselves to the sites which are unique from active sites, and changing the movement of the protein molecule in quick and reversible manner. These allosteric connections with metabolites likewise impact proteins which lack all the enzymatic capacities for example, receptors present on the transmembrane and transcription factors. The function and assembly of protein complexes are also regulated by metabolite bindings [8] and is also responsible for high-molecular-weight protein assemblies [9].

At any rate, within a typical cell several million protein particles metabolize dwarf proteins by around 100-fold. Along these lines, a large number of practically important events of a metabolite binding to a protein could be present. Our insight into the domain of protein-metabolite interactions are still not complete, as they have low binding affinity and bind for a small period of time which has forestalled methodical examinations like those performed to recognize protein-protein or protein-nucleic acid interactions. Metabolite binding likewise controls the fabrication along with function related with the protein edifices.

Most portrayed connection between protein and metabolite have been found utilizing theory-driven tests that depend on in vitro activity assays. They are normally components for metabolites, which provide the formation of cofactors needed by an enzymatic reaction, supply the energy which is required to carry out a biochemical reaction, and regulates functioning which is related to the long polypeptide chains [10]. These tests are difficult, contingent upon the decision of the pertinent ligand, and block the recognizable proof of connections that don't change an in vitro quantifiable action. Strategies including synthetic adjustment of metabolites or protein labeling have been confined to investigations of lipid-protein cooperation and hydrophobic metabolites. We reasoned that recognizing ligand-

instigated primary adjustments and would give read out of all the interactions between the small molecules and proteins.

During the past decade, substantial efforts have been made for the analyses of biological networks, specifically for an interaction between protein molecules, binding of different transcription factors, and even the phosphorylation network for proteins [11]. These examinations have given an abundance of data for a better comprehension of how proteins work, how different segments come together to perform a function, and the essential standards of administrative arrangement of association.

In absolute quantity, these small metabolites are the most abundant in the cell segments, and similarly to proteins, these molecules are available in a wide scope in a cell while focusing and partaking in almost every biochemical assortment and administrative capacities. One illustration from the diverse pool of these small association between proteins and metabolites is galactose interacting and binding with a sensor protein [12], steroid hormones to transcription factors [13].

As responsible for controlling the functioning of proteins, a metabolite can govern numerous polypeptide molecules or explicitly focus on a set group of polypeptides. In resentment, their significance in interceding protein work and guidelines, precise methodologies for investigating their behavior in their natural surroundings are still not approached practically. This data is not only limited to clarification of the bio-chem exercises, but also provides guidelines of individual functioning of protein, yet for amassing and associations between natural paths. Besides, dietary admission of supplements can alter the level of metabolites, interpreting the guideline of cell measures by the level of metabolites shows a potential to remedy the deserts observed in the study of pathways.

Chapter 2

LITERATURE REVIEW

==========

2.1 Protein-Metabolite Interactions

The interactions between a protein and a metabolite is responsible for the maintenance of cellular hemostasis at several levels as they regulate many small interactions throughout the human body. Metabolites not only have a role as intermediate compounds in a metabolic conversion but are also involved in regulation of signal cascades to trigger different responses. Activity of a metabolite is directly dependent on the nutritional states and ecological stress over the body, Metabolites behave as allosteric regulators to a protein molecule that is binding to separate sites than active sites supporting the proteins activity (Gerosa and Sauer et al.,2011). The interactions between metabolites and proteins usually cause alterations to a protein structure at a local and a global level. The diagnosis of such interaction is done by detecting ligand-induced alterations in the structure of the target molecule. Metabolite and protein interactions occur at small scale and are very short spanned, so there are not many compulsive studies focused over these interactions. The study published by Xiyan et al.,2010 have demonstrated a mass spectroscopy assay for the identification of such interaction in yeast. [19]

2.2 Need for the Study of Protein Metabolite Interactions

Metabolite-protein interaction is critical for numerous cellular and physiological processes that occur throughout a biological system. Although, being crucial the protein-metabolite interactions are still understudied. [20] As metabolites are abundantly spread throughout the body and present in varying concentration all around the body and are just as involved as proteins in various biochemical reactions. With the study of interactions between protein and metabolite it creates an insight for understanding of biochemical processes related to signaling pathways, having a better grasp of protein metabolite interactions can prove to be very helpful in the understanding and manipulation of biochemical signaling pathways. [21] The recent discoveries and development in the mass spectroscopy techniques, highly sensitive metabolic profiling and high throughput proteomics a better horizon for studies of

protein metabolite interaction is developing which has the potential to completely revolutionize biochemical research of drug discovery. [22]

2.3 Available technique

Protein-metabolite interactions are known to help in better understanding of concepts involved in diagnostics of biochemical pathways and cellular activity which occur during a medical condition, for identification of such interaction is usually carried out by either directly studying the binding between a protein and metabolite or measuring the concentration of free molecules. The approaches for this characterization is usually classified in three categories: -

1. *In vitro*
2. *In vivo*
3. *In silico*

2.3.1 In vitro

One of the easiest and readily available method for characterization of protein-metabolite interactions, in vitro techniques utilizes well-controlled, standard conditions and chemical reagents to imitate biological systems. To identify different participants of an interaction in vitro methods relies on binding assays, in vitro techniques are also helpful in identifying the strength of an interaction, thermodynamics involved in it and the reaction kinetics behind the interaction. [25] It can indicate possible conformational changes that a molecule undergoes during an interaction. The structural configuration of a protein metabolite interaction can also be studied by in vitro techniques such as X-ray crystallography, NMR spectroscopy and mass spectroscopy. [21,22,23]

2.3.2 In vivo

In vitro methods have the capability to provide detailed insights about the protein metabolite interactions, but are not able to support the data in the case of proteins which undergo post-translational changes. In vivo techniques are better suited for such samples as the insights from these techniques directly displays how the protein and metabolite would interact inside a biological system, for proteins going through post-translational changes can be studied through these techniques. Methods utilized under the in vivo techniques usually share similarity with the methods used under in vitro techniques with just a simple difference which

is the use of more complex samples. Samples from patients can be studied using mass spectroscopy or NMR spectroscopy for structural studies. [24] In vivo studies provide a close and better understanding of protein metabolite interactions during a disease, and also shows the effects of diseases on the biochemical pathways inside the human body.

2.3.3 In silico

The use of computational data for the determination of configuration of a binding site of a metabolite and protein is referred to as in silico methods for the studying of protein metabolite interactions. In silico methods generally helps in the study of a protein metabolite complex by analysis of all the favorable conditions for the interaction by stimulating a virtual environment. These methods also provide an edge over other techniques as it can help in the prediction of binding site of metabolites and proteins through the homology study of amino acids present in the protein and how the metabolite will interact with it depending upon its chemical behavior [26]. In silico methods are used to create databases as it can process higher volumes of data at once. They can also compliment the structural characterizations of in vitro techniques [27].

OBJECTIVES

==========

The Major objectives of this project were as follows: -

- Exploring HMDB and KEGG.
- Exploring Cytoscape and its plug-in Metscape
- Identifying and understanding protein-metabolite interaction.

Chapter 3

METHODOLOGY

==========

3.1 HMDB

HMDB stands for Human Metabolome Database which is an uninhibitedly accessible computerized data set storing definite data related to little particle metabolites present inside the body of a human. The intended use for this data is directed towards the study of metabolites, support clinical science, identification of bio markers, and basic training.
The HMDB is available at https://hmdb.ca/

The distinguishing feature of HMDB which separates it from other metabolic assets is its broad help for more significant level information base looking and choosing capacities. HMDB also offers a search utility for chemical structures, a BLAST search as an additional feature to the already illustrated features which are classification and reckoning of data, this platform also provides a search utility for chemical structures with an added local BLAST search [14].

The basic goal behind the designing of this database is to accommodate or relate three kinds of data:

1. Information about the chemical nature of various metabolites
2. Clinical importance and effectiveness of metabolites
3. Data about the molecular orientation and biochemical behavior of a metabolite.

Formats in HMDB: -

1. FASTA format is used Protein/Gene Sequences
2. SDF format to support structural data
3. XML format for storing of metabolite and protein related data

Moreover, 5,702 protein successions are connected to these metabolite sections.

Steps

1. Visited the HMDB website homepage.
2. Downloaded all the metabolite data related to the TCA cycle in humans.

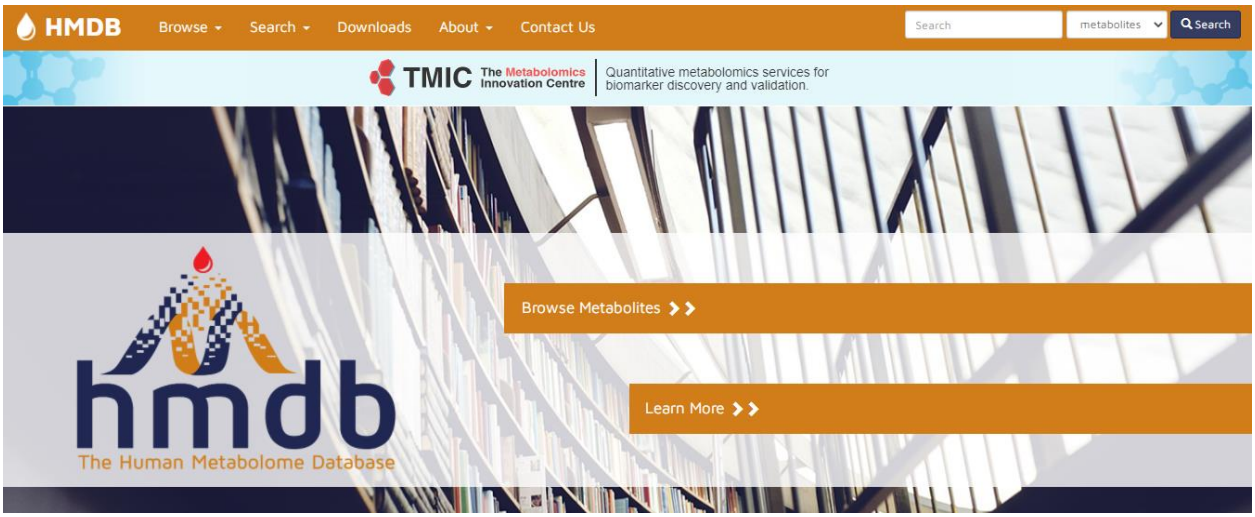


Fig 3.1 HMDB Homepage

3.2 KEGG

KEGG which is abbreviation for Kyoto Encyclopedia of Genes and Genomes is an approach to create link between the information from genomics and practical information of higher-order by uploading of insights based on the study of processes followed by a cell and creating a standard annotation for the genes [15].

- KEGG compiles reports of genomic data, chemical behavior and systematic functions in a single database.
- Principally, catalogs of genes which belong to a sequenced genome are connected with progression of the functions which are carried out by the cell, the organism, and the ecosystem.
- It is advanced by Kanehisa Laboratories.
- KEGG ENZYME is established upon the ExplorEnz database and is managed in the KEGG LIGAND.
- KEGG database was proposed in 1995.
- The KEGG is available at https://www.genome.jp/kegg/

A single gene or a molecule cannot account for the complete biological function of the living cell, instead it is a result of many molecules interacting together [16].

Steps

1. Visited the KEGG website homepage.
2. Downloaded all the proteins related to the TCA cycle in humans.

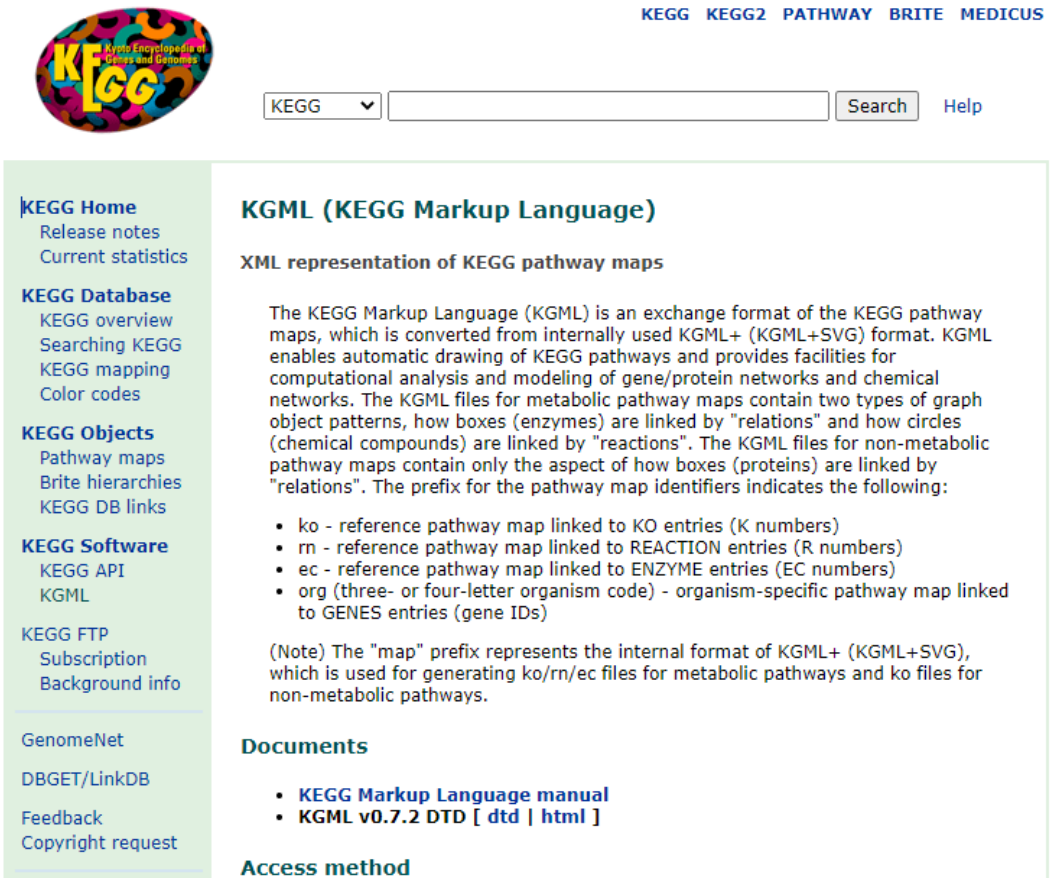


Fig 3.2 KEGG Homepage

3.3 CYTOSCAPE

Cytoscape provides a platform for the in the integration of bio-molecular interaction network with an efficient expression of data and other molecular states into a unified conceptual framework [17]. Cytoscape is one of the most famous open-source programming instruments for the visual investigation of biomedical organizations made out of protein, quality, and different kinds of communications. It offers scientists an adaptable and intelligent representation interface for investigating complex organic interconnections upheld by assorted explanation and exploratory information, accordingly encouraging examination errands, for example, anticipating quality capacity and pathway development.

Cytoscape gives center usefulness to stack, imagine, search, channel, and spare organizations, and several Apps stretch out this usefulness to address explicit exploration needs.

Cytoscape is generally utilized in natural organization investigation and it upholds many use cases in atomic and frameworks science, genomics, and proteomics:

1. It can import and load atomic and hereditary connection datasets in a few organizations.
2. It can utilize a few visual highlights that can viably feature key parts of the components of the organization.
3. This can be spared as visual styles, sent out and imported for re-use.

The Cytoscape is available at https://cytoscape.org/

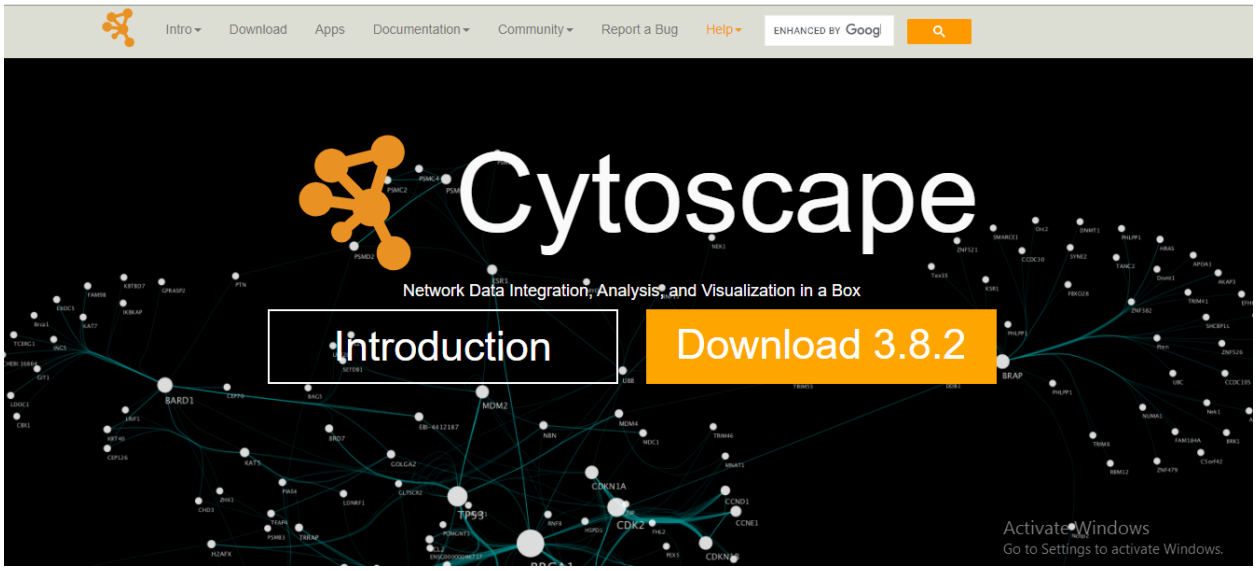


Fig 3.3 CYTOSCAPE Homepage

3.4 METSCAPE

The MetScape 3 App is an extension for Cytoscape which generates a framework that can visualize and interpret metabolomics data with respect to human metabolism [18].

It permits clients to assemble and dissect organizations of qualities and mixes, recognize advanced paths from articulation information, and imagine changes in metabolite information. MetScape utilizes an inner social information base that coordinates information from KEGG and EHMN.

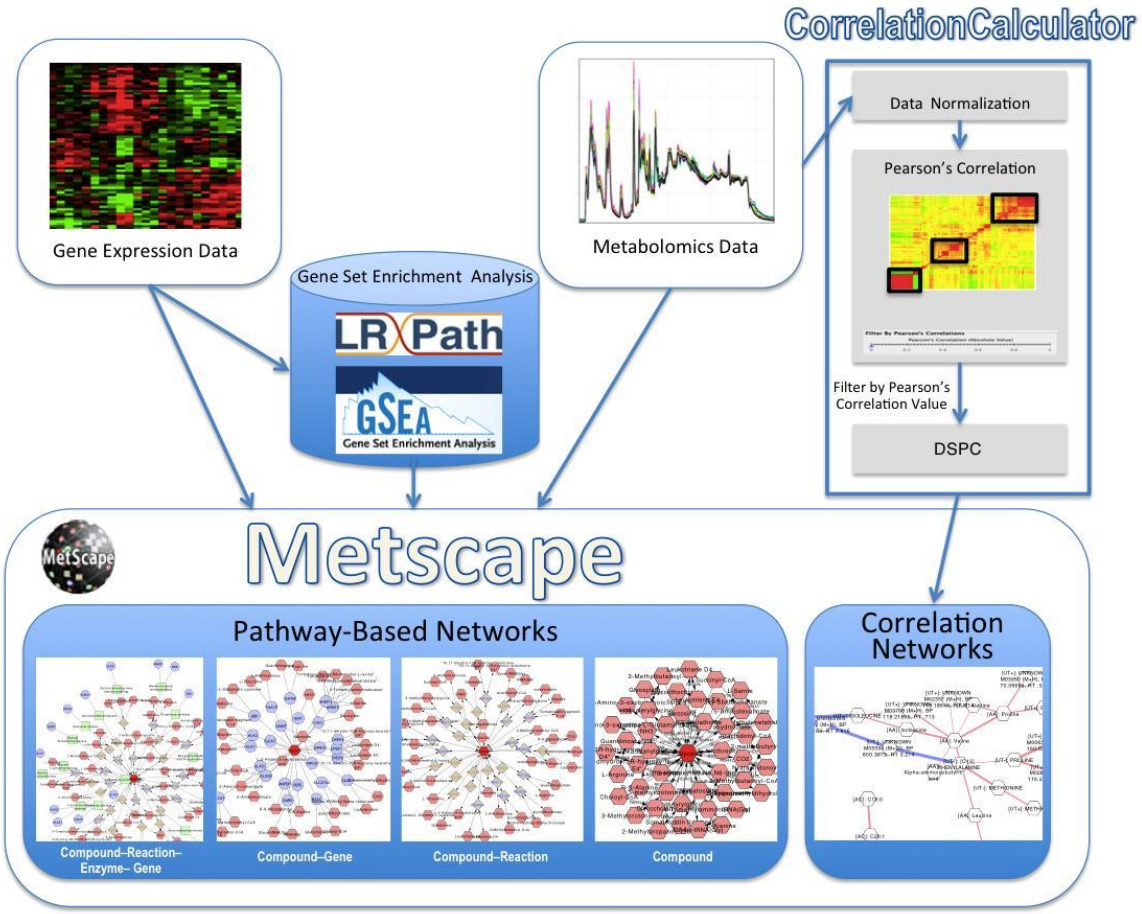


Fig 3.4 Metscape Workflow (http://metscape.ncibi.org/)

Steps

1. Open Cytoscape, go-to **apps** option.
2. Click on **app manager** and install metscape.
3. Installed METSCAPE Plug-in through CYTOSCAPE.
4. Selected **build network** option and then selected **pathway-based.**
5. Under the input section selected human as the organism.
6. Clicked on select under the heading of data file.
7. Imported the experimental data which was downloaded earlier.
8. Selected network type as the compound reaction.
9. Clicked on **build network.**

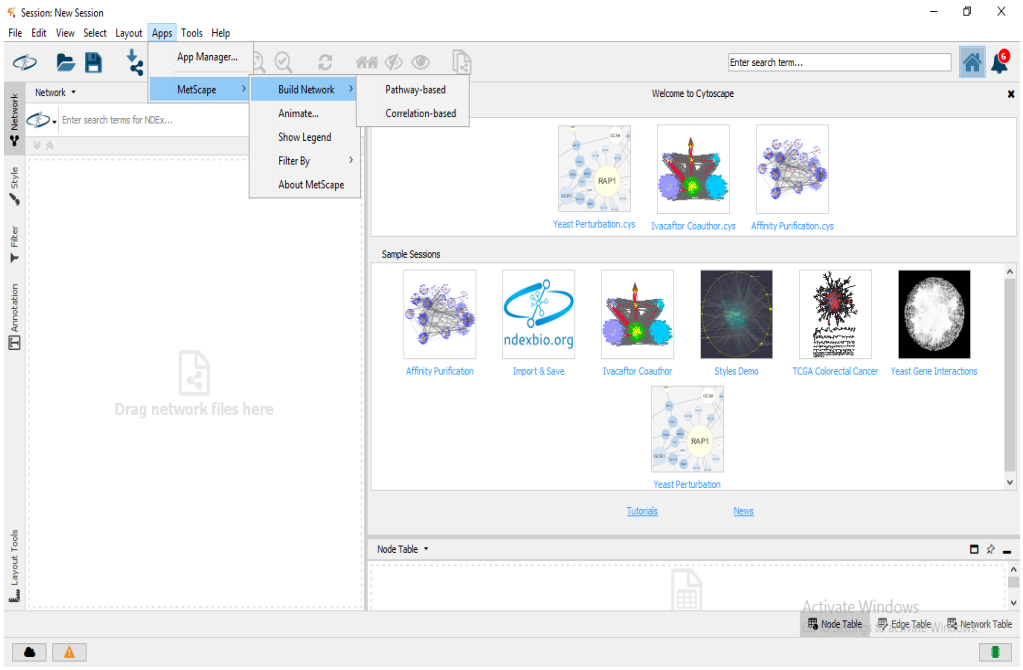


Fig 3.5 Procedure to run Metscape

Chapter 4

RESULTS & DISCUSSION

==========

4.1 Results

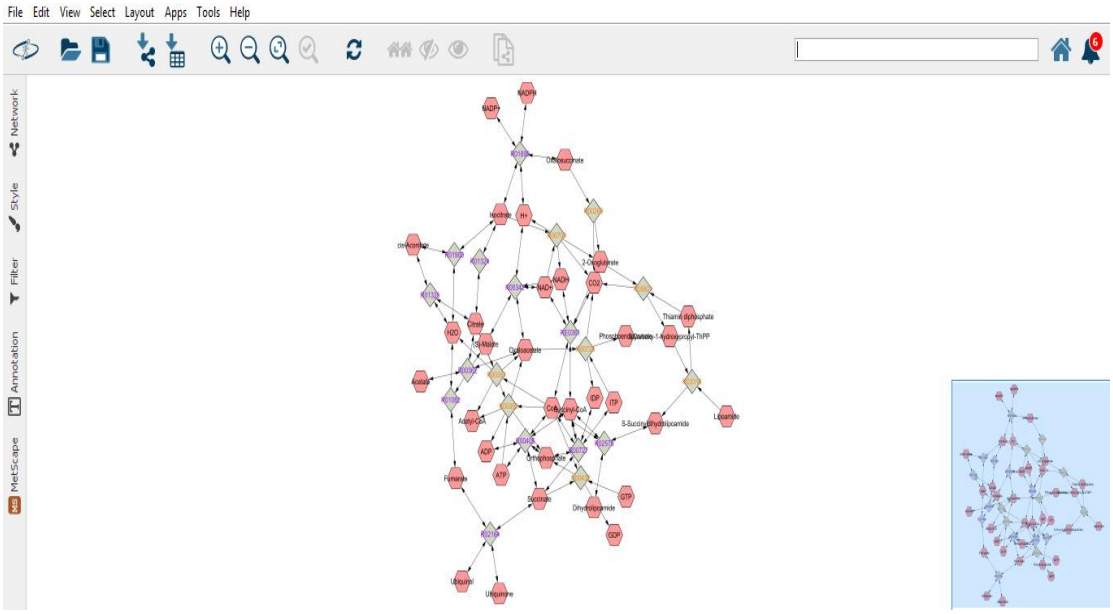


Fig 4.1 The output obtained from Metscape

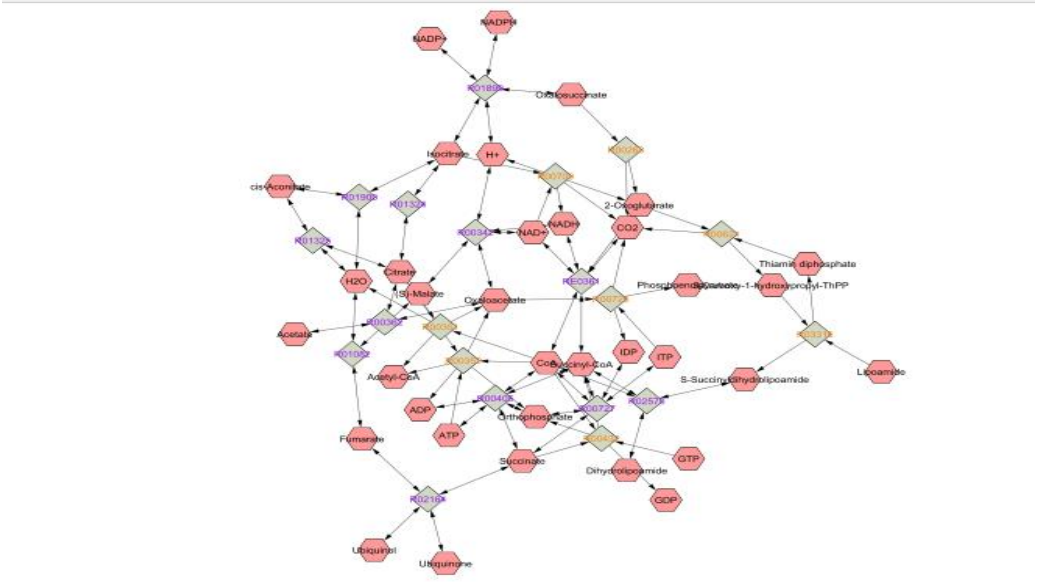


Fig 4.2. Protein-metabolite network with protein (hexagons) and metabolite (rectangles)

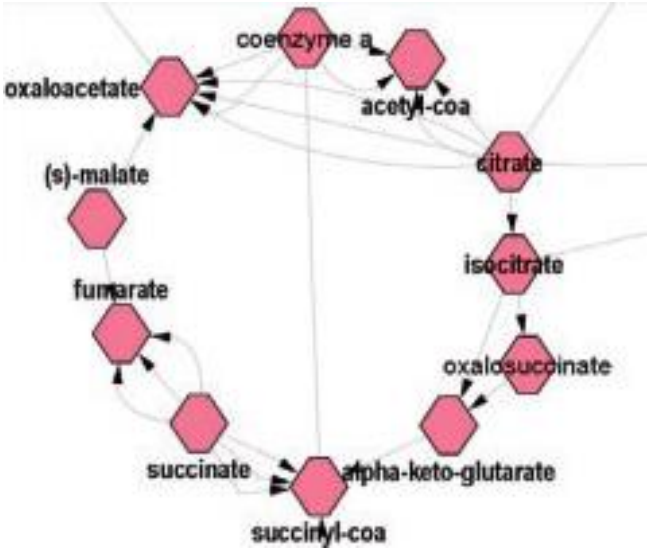


Fig 4.3 Part TCA cycle of the Protein network with Protein as nodes and reactions as edges

4.2 Discussion

This is a visual representation of the interaction network between proteins and metabolites involved in the TCA cycle. The compiled data here has 36 nodes and 20 edges, with metabolites representing nodes. By the study of these networks one can have a better understanding of molecular interactions as it provides several small reactions that occur at the molecular levels.

Table 4.1: List of Metabolites involved in TCA

Serial Number	**Metabolites**
1	Acetyl-CoA
2	Citrate
3	Isocitrate
4	α-ketoglutarate
5	Succinyl-CoA
6	Succinate
7	Fumarate
8	Malate
9	Oxaloacetate

Table 4.2: List of Proteins involved in TCA

Serial Number	**Proteins (enzymes)**
1	Citrate Synthase (CSY)
2	Aconitase (ACO)
3	Isocitrate Dehydrogenase (IDH)
4	α-ketoglutarate Dehydrogenase complex (αKGDHC)
5	Succinyl-CoA Synthetase (SCoAL)
6	Succinate Dehydrogenase (SDH)
7	Fumerase (FUM)
8	Malate Dehydrogenase (MDH)

Chapter 5

TECHNICAL CHALLENGES

==========

- The study of protein metabolite interaction is severely understudied compared to other biological interaction.
- As these interactions occur at a comparative small scale and are difficult to identify, but with the advancements in the detection techniques such high throughput proteomics and chemical profiling of different metabolites with exceptional sensitivity has created a wide scope for the study of interaction of different protein-metabolite Interaction.
- Visualization of such network has its own set of challenges as metabolomics data have an increased degree of connectivity. Which require several different layout algorithms for proper functioning and creating accurate visual outputs.
- The metscape plugin of cytoscape provides all these frameworks for visualization of metabolomics data, helping in creating the interaction network between proteins and metabolites. [28]

Chapter 6

CONCLUSIONS & FUTURE PROSPECTS

==========

Conclusions

- The systematic analysis of the metabolite protein network resulted in distinguishing a large number of recently known and revealed useful principles that control these interactions.

- It brings the result of high order underlying changes and gives a bioinformatics structure/framework for the analyzing and interpreting of data from the human metabolome.

- Metabolomics data from HMDB and KEGG are coordinated. Visualizing the protein grid provides an ease to follow associations among metabolites and proteins, and create a better awareness of the relation among the reactions, enzymes, genes, compounds and the followed pathways. Here, the number of nodes are 36, and the number of edges are 20.

- Visualizing metabolites with node color and size changes in an organization at a progression of time focuses can assist users with anticipating elements of metabolites.

- Connecting compounds and reactions to notable metabolomics data sources help in getting a better understanding of the metabolites.

Future Prospects

- The protein metabolite interactions are responsible for the regulation of energy flux and the flow of material in a biochemical process, the study of such molecular interaction provides a broad scope to better understand the reaction that occur in a biological system at a molecular level.
- Protein-metabolite interaction is a part of the field of metabolomics which has been growing recently and is related with the study of biological pathways and processes. The study of interactions between metabolites and proteins is still considered to be a fresh area in this field as most of the existing studies in this field is related with the structural analysis of different metabolites.
- The upcoming research in this area has shown that by studying the interaction network between metabolites and proteins one can better understand the molecular interactions involved in a biochemical pathways and also help in deciphering the effects of different diseases on these biochemical pathways.
- The study of protein metabolite interactions plays a key role in the fields of pharmaceutical sciences, personalized medicine, drug discovery and biomedical sciences.

REFERENCES

==========

1. https://en.wikipedia.org/wiki/Protein
2. German, J. Bruce, Bruce D. Hammock, and Steven M. Watkins. "Metabolomics: building on a century of biochemistry to guide human health." *Metabolomics* 1, no. 1 (2005): 3-9.
3. Milo, Ron. "What is the total number of protein molecules per cell volume? A call to rethink some published values." *Bioessays* 35, no. 12 (2013): 1050-1055.
4. Bennett, Bryson D., Elizabeth H. Kimball, Melissa Gao, Robin Osterhout, Stephen J. Van Dien, and Joshua D. Rabinowitz. "Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli." *Nature chemical biology* 5, no. 8 (2009): 593-599
5. Chubukov, Victor, Luca Gerosa, Karl Kochanowski, and Uwe Sauer. "Coordination of microbial metabolism." Nature Reviews Microbiology 12, no. 5 (2014): 327-340.
6. Boehr, David D., Ruth Nussinov, and Peter E. Wright. "The role of dynamic conformational ensembles in biomolecular recognition." *Nature chemical biology* 5, no. 11 (2009): 789-796.
7. Nussinov, Ruth, and Chung-Jung Tsai. "Allostery in disease and drug discovery." *Cell* 153, no. 2 (2013): 293-305.
8. Milroy, Lech-Gustav, Tom N. Grossmann, Sven Hennig, Luc Brunsveld, and Christian Ottmann. "Modulators of protein-protein interactions." *Chemical reviews* 114, no. 9 (2014): 4695-4748.
9. O'Connell, Jeremy D., Alice Zhao, Andrew D. Ellington, and Edward M. Marcotte. "Dynamic reorganization of metabolic enzymes into intracellular bodies." *Annual review of cell and developmental biology* 28 (2012): 89-111.
10. Förster, Jochen, Iman Famili, Patrick Fu, Bernhard Ø. Palsson, and Jens Nielsen. "Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network." *Genome Research* 13, no. 2 (2003): 244-253.
11. Snyder, Michael, and Jennifer EG Gallagher. "Systems biology from a yeast omics perspective." *FEBS Letters* 583, no. 24 (2009): 3895-3899.
12. Yano, Ken-ichi, and Toshio Fukasawa. "Galactose-dependent reversible interaction of Gal3p with Gal80p in the induction pathway of Gal4p-activated genes of

Saccharomyces cerevisiae." *Proceedings of the National Academy of Sciences* 94, no. 5 (1997): 1721-1726.

13. Evans, Ronald M. "The steroid and thyroid hormone receptor superfamily." *Science* 240, no. 4854 (1988): 889-895.
14. Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman. "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." *Nucleic acids research* 25, no. 17 (1997): 3389-3402.
15. Kanehisa, Minoru, and Susumu Goto. "KEGG: Kyoto encyclopedia of genes and genomes." *Nucleic acids research* 28, no. 1 (2000): 27-30.
16. Kanehisa, Minoru. *Post-genome informatics*. OUP Oxford, 2000.
17. Shannon, Paul, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. "Cytoscape: a software environment for integrated models of biomolecular interaction networks." *Genome Research* 13, no. 11 (2003): 2498-2504.
18. http://metscape.ncibi.org/
19. Li, Xiyan, Tara A. Gianoulis, Kevin Y. Yip, Mark Gerstein, and Michael Snyder. "Extensive in vivo metabolite-protein interactions revealed by large-scale systematic analyses." *Cell* 143, no. 4 (2010): 639-650.
20. Qin, Qian, Bohong Wang, Jiayue Wang, Mengmeng Chang, Tian Xia, Xianzhe Shi, and Guowang Xu. "A comprehensive strategy for studying protein-metabolite interactions by metabolomics and native mass spectrometry." *Talanta* 194 (2019): 63-72.
21. Yang, Grace Xiaolu, Xiyan Li, and Michael Snyder. "Investigating metabolite–protein interactions: an overview of available techniques." Methods 57, no. 4 (2012): 459-466.
22. Guo, Hongbo, Hui Peng, and Andrew Emili. "Mass spectrometry methods to study protein-metabolite interactions." Expert opinion on drug discovery 12, no. 12 (2017): 1271-1280.
23. Palmer, Rex A., and Hideaki Niwa. "X-ray crystallographic studies of protein–ligand interactions." *Biochemical Society Transactions* 31, no. 5 (2003): 973-979.
24. Bertini, Ivano, Marco Fragai, Andrea Giachetti, Claudio Luchinat, Massimiliano Maletta, Giacomo Parigi, and Kwon Joo Yeo. "Combining in silico tools and NMR

data to validate protein− ligand structural models: application to matrix metalloproteinases." *Journal of medicinal chemistry* 48, no. 24 (2005): 7544-7559.

25. Gestwicki, Jason E., Helen V. Hsieh, and J. Bruce Pitner. "Using receptor conformational change to detect low molecular weight analytes by surface plasmon resonance." *Analytical chemistry* 73, no. 23 (2001): 5732-5737.
26. Patti, Gary J., Oscar Yanes, and Gary Siuzdak. "Metabolomics: the apogee of the omics trilogy." *Nature reviews Molecular cell biology* 13, no. 4 (2012): 263-269.
27. Sun, Hao, and Dennis O. Scott. "Structure-based drug metabolism predictions for drug design." Chemical biology & drug design 75, no. 1 (2010): 3-17.
28. Gao, Jing, V. Glenn Tarcea, Alla Karnovsky, Barbara R. Mirel, Terry E. Weymouth, Christopher W. Beecher, James D. Cavalcoli et al. "Metscape: a Cytoscape plug-in for visualizing and interpreting metabolomic data in the context of human metabolic networks." *Bioinformatics* 26, no. 7 (2010): 971-973.
29. Liesener, André, and Uwe Karst. "Monitoring enzymatic conversions by mass spectrometry: a critical review." *Analytical and bioanalytical chemistry* 382, no. 7 (2005): 1451-1464.
30. Clarke, William, A. Roy Chowdhuri, and David S. Hage. "Analysis of free drug fractions by ultrafast immunoaffinity chromatography." *Analytical chemistry* 73, no. 10 (2001): 2157-2164.