Research Journal of Microbiology

ISSN 1816-4935

Academic Journals Inc.

www.academicjournals.com

Research Journal of Microbiology 4 (3): 132-137, 2009
ISSN 1816-4935

Graph Theoretic Approach on Metabolomic Networks of Mycobacterial Strains for Potential Drug Targets

[1]V. Baths, [2]V.V. Rohit Kumar, [3]G.V.R. Praneeth and [1]U. Roy
[1]Birla institute of Technology and Science Pilani, Goa Campus, India
[2]Ah3 102,
[3]Ah5 242, Birla Institute of Technology and Science Pilani, Goa Campus, India

Abstract: A special strain of *Mycobacterium tuberculosis*, H37Rv's Gluconeogenesis pathway is analysed for clusters in the pathway using the principles of spectral graph theory to find out a drug target for tuberculosis. The software named Visant was used and the data set was obtained from KEGG. The large-scale properties of chemical reaction systems, such as metabolism, can be studied with graph-based methods. To do this, one needs to reduce the information, lists of chemical reactions, available in databases. There are several ways by which this reduction can be done even for the simplest type of graph representation. Present study is aimed to apply the knowledge of graphs and graph theoretic concepts to compare the metabolic network in *Mycobacterium tuberculosis*. The study is done on the gluconeogensis pathway, a pathway that is important for the growth of *M. Tubercoulosis* H37Rv strain. Each metabolite of the pathway is taken as node of a network with the edge between the nodes representing the reaction. Spectrum and spectral radius of this network were obtained using spectral graph theory, manually. The spectral radius of this network is found out to be 0.9254.

Key words: Tuberculosis, spectral graph analysis, H37Rv, cluster formation

INTRODUCTION

Recent advances in high throughput technologies and network theory have made it possible to reconstruct and analyse large genome-scale networks of organisms in silico. Several types of networks reflecting different aspects of metabolism and regulation in organisms have been reconstructed. The transcriptional networks based on microarray data, protein-protein interaction networks based on high-throughput yeast two-hybrid type of experiments and metabolic networks based on reaction annotation of the individual proteins coded by the genome are some examples. Several of these studies have focused on elucidating the general principles underlying the structure and organisation of metabolic networks of a large number of organisms (Verkhedkar *et al*., 2007).

Here, we seek to exploit the reaction-based networks for comparative network analysis. We use concepts from spectral graph theory to systematically determine how differences in the basic metabolism of various organisms are reflected at the systems level. In the present study, we have constructed and characterised the metabolic networks of organism: *Mycobacterium tuberculosis* H37Rv strain which is obligate intracellular pathogens.

Various concepts from graph theory have previously been used to construct and analyse metabolic networks for several fully sequenced organisms (Wolfang *et al*., 2007). Representing metabolic networks as graphs makes them amenable to various analyses, such as the detection of shortest and alternate paths. Such analysis have also resulted in the identification of the highly

Corresponding Author: Veeky Baths, Birla institute of Technology and Science Pilani, Goa Campus, India
Tel: (+91)9823764987

connected giant strong components of the networks, as well as metabolites central to the network (Patra and Vishveshwara, 2000). Graph spectral analysis can be carried out to obtain information on central hubs.

MATERIALS AND METHODS

Bacterial Strain

Mycobacterium tuberculosis H37rv Strain.

Softwares Used

- Visant-for analysisng networks of the pathways
- Keggdraw-for viewing structures of compounds involved
- MATLAB-For calculating Eigen values

Database Used

We used the KEGG (Kyoto encyclopedia of genes and genomes) database to reconstruct the reaction networks of *M. tuberculosis* H37Rv. A list of metabolic pathways and their constituent biochemical reactions were downloaded as flat files. These files contain information about reactants, products, reversibility and study state stoichiometry of biochemical reactions.

Metabololic Network Reconstruction

The metabolic networks of the strain taken were reconstructed. Each biochemical reaction in the metabolome of the organism is a node and nodes representing reactions that share a common metabolite as a substrate in one reaction and product in the other are connected by an edge. Thus only reactions which exhibit such a consecutive dependence on each other with respect to a metabolite is connected to each other. All edges have an equal weight of 1. In order to make the network amenable to network analysis, it is represented in the form of adjacency matrix or Reaction-Interaction Matrix (RIM), which is an nxn matrix; n being the number of nodes (biochemical reactions) in the graph. The elements of Aij of the RIM have values according to the following rules:

- $A_{ij} = 1$ if reactions i and j exhibit a consecutive dependence through a metabolite
- $A_{ij} = 0$ if the reactions donot exhibit such a dependence, or if i = j

To construct the RIM, the set of reactions in the flat file representing the metabolome was first represented as stoichiometric matrix S (m×n), eith every metabolite being represented by a row and every reaction by a column.

Graph Spectral Analysis

Graph spectral theory is a subfield of graph theory that deals with the analysis of the spectra (Eigen values and Eigen vector components) of nodes in the graph. Such an analysis provides information on the overall structure and topology of the graph. To obtain Eigen value spectra of the graph, the adjacency matrix is converted to a Laplacian matrix L, by the equation:

$$L = D-A$$

where, D is the degree matrix of the graph, A is the diagonal matrix in which the ith element of the diagonal is equal to the number of connections that the ith node makes in the graph diagnolisation of the Laplacian matrix which, yields the spectra of the graph comprising the Eigen values and corresponding Eigen vectors.

The analysis of vector components of the lower and higher Eigen values yields information about clusters present in the graph and connectivity of each node. Specifically, the vector components of second lowest Eigen value carry information about clusters present in the graph. A plot of sorted vector components and second lowest Eigen value versus node number clearly reveals the sub-cluster information with the nodes belonging to a sub cluster forming distinct plateaus on the curve and the nodes connecting the sub clusters having vector component values in between the plateaus.

RESULTS AND DISCUSSION

The Gluconeogeneis pathway of *M.tuberculosis* H37Rv strain is analysed in Visant (KEGG Database) (Fig. 1). Visant gives this pathway as a network of nodes and edges.The nodes representing the metabolites and the edges representing the reaction.

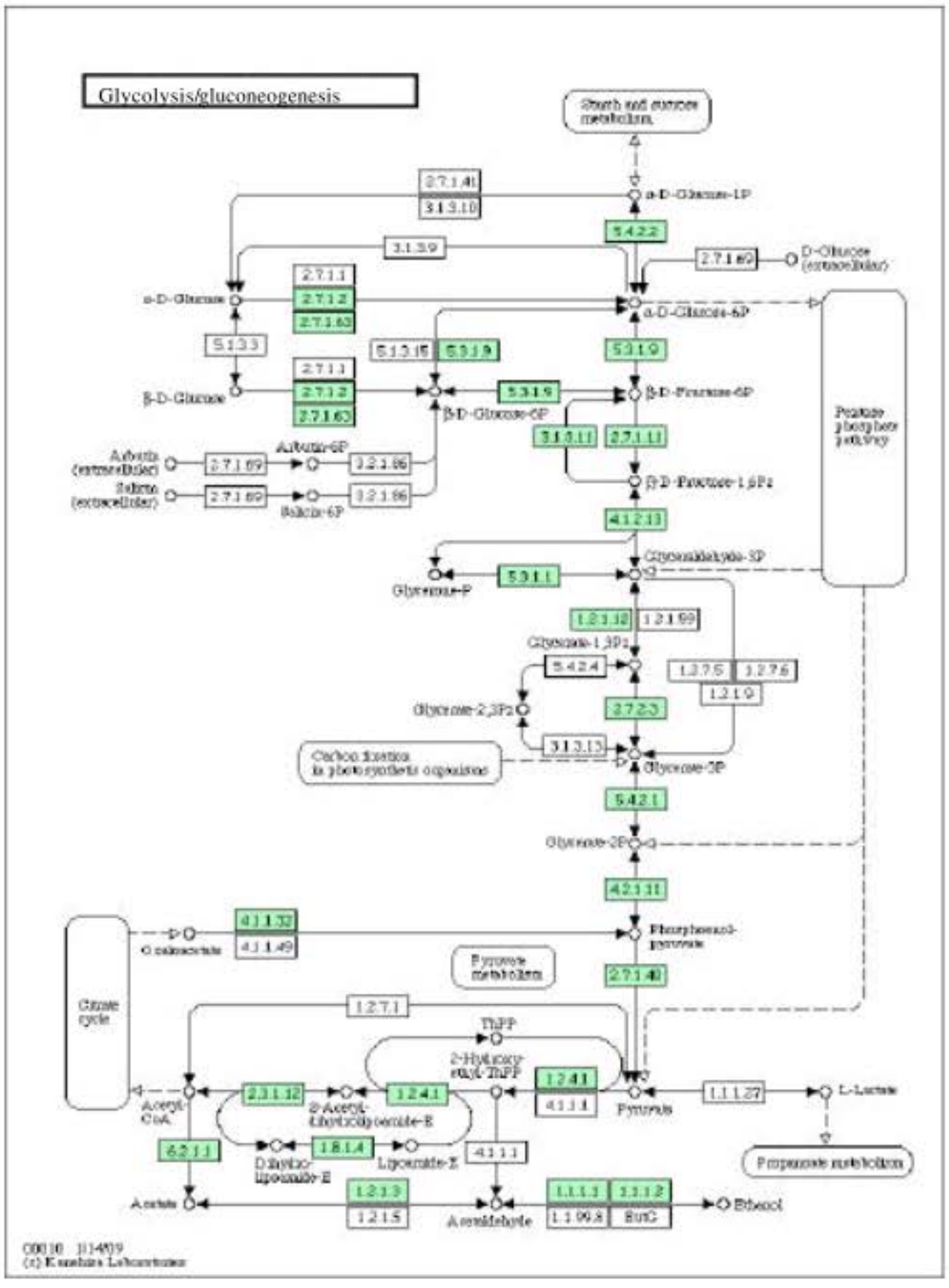

Fig. 1: Gluconeogensis pathway of *M. tuberculosis* H37Rv

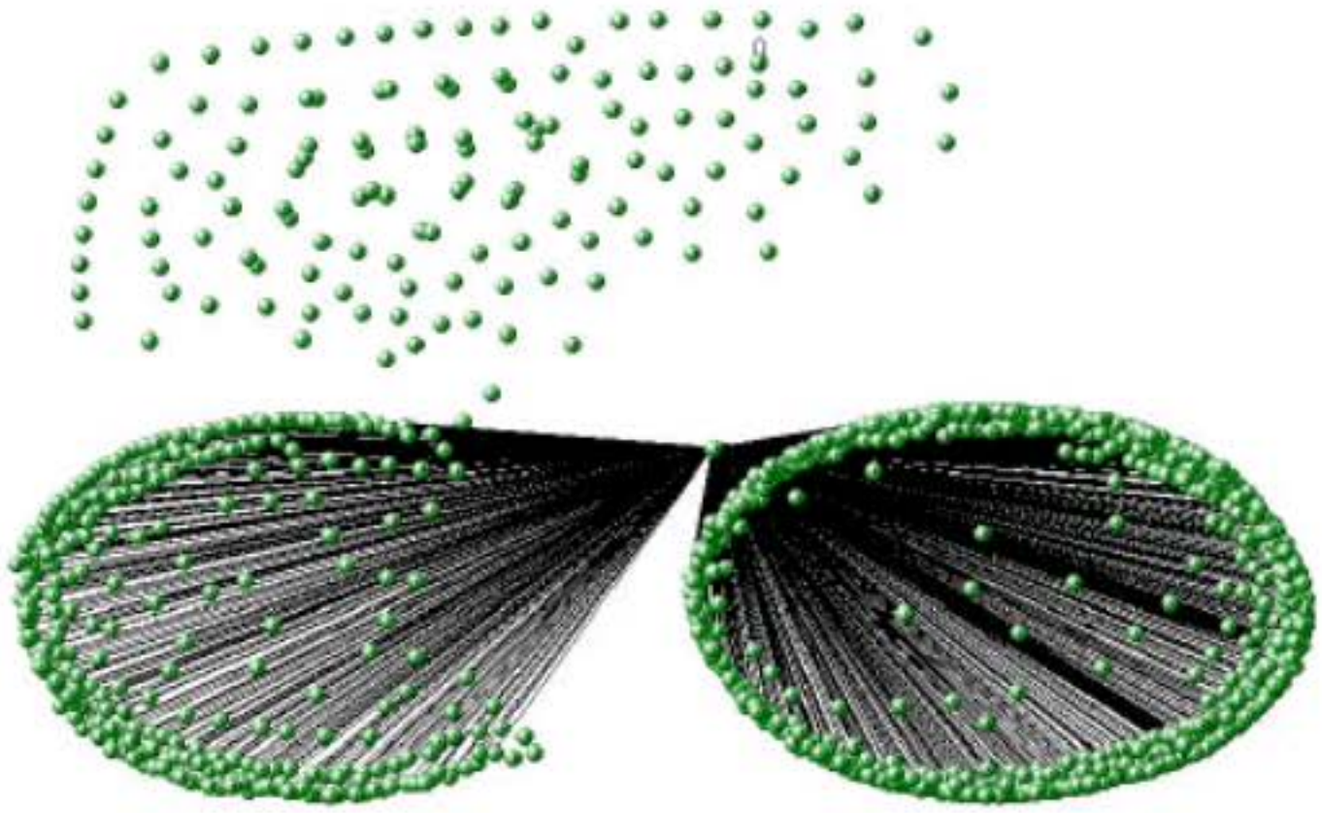

Fig. 2: Network representation of the pathway

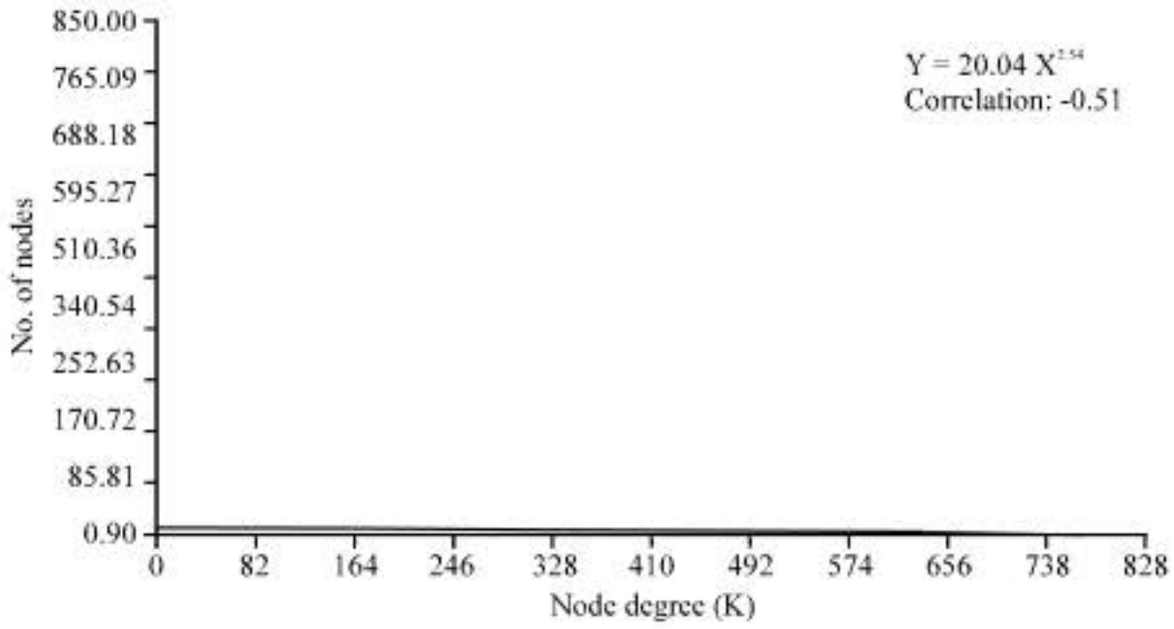

Fig. 3: Graph of number of nodes vs. the node degree

The gluconeogensis pathway is connected to various other pathways namely pentose phosphate pathway, starch and sucrose metabolism, citrate cycle and propanoate metabolism. The glycolysis pathway along with the above mentioned pathways is analysed for the network (Van Helden *et al*., 2002) in Visant (Fig. 2).

In the above picture, the balls are the metabolites and the black edges are the reactions.

In the second phase of the study, the central gluconeogenesis pathway is taken for analysis, wherein the other connected pathays are taken as extra nodes for the calculation of degree (Vishweshwara *et al*., 2002) (Fig. 3). In this network (pathway), there are 31 metabolites which are the nodes of the network. The Adjacency matrix found out for this network (Ma and Zeng, 2003) is given in Fig. 4:

The Eigen values of this matrix are found out using MATLAB.

Eigen values							
0	-2.4659	-1.8894	1.4070	0.8078	0.4489	-0.7142	0
3.0477	-2.3201	-1.7333	-1.4879	1	0.2563	-0.2987	-1
2.8314	-2.0762	1.7288	-1.1832	1	0.1837	-0.4650	-1
2.7579	1.9591	1.5586	-1.1546	1.0055	-0.7651	-0.4393	

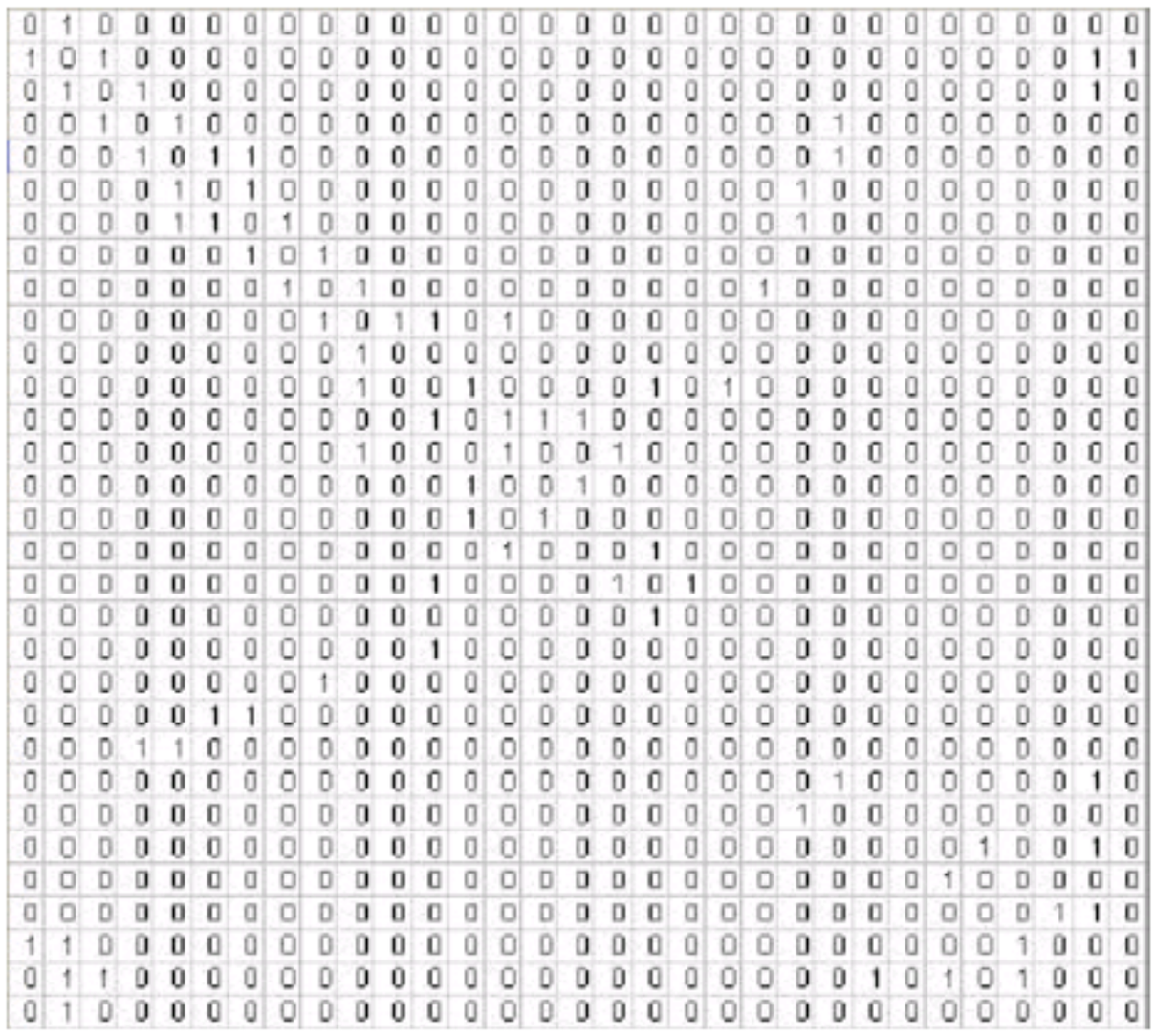

Fig. 4: Adjacency matrix drawn with its elements as zeroes and ones,one if there is a connecting adge between two nodes i and j and zero otherwise

In the next phase, the Laplacian Matrix (Walter, 2005) is found out. L = A-D. Here D is the diagonal degree matrix of our network.

D= diag [3 6 3 3 5 3 4 3 3 5 2 4 4 4 2 2 2 3 1 1 2 2 2 2 1 2 1 2 3 5 1]

L is normailsed using L' = [1/sqrt(D)]L[1/sqrt(D)]

This is normalized and the modulus of matrix is obtained.

The set of Eigen values for this matrix are obatined and the maximum of that set gives the spectral radius.

Final Eigen values

0	0.7771	-0.8300	0.2771	-0.6841	-0.2596 - 0.0225i	-0.5997	0.5
0.9254	0.6833	-0.8362	0.2159	-0.1670	-0.3133	-0.5804	0.0
0.9176	0.6230	0.4425	0.1257	-0.1541	-0.3690	-0.5573	-0.5
0.8529	0.5589	0.3574	0.0721	-0.2596 + 0.0225i	-0.4391	-0.5294	

Clearly, 0.9254 is the spectral radius.

CONCLUSION

The spectral radius of gluconeogenesis pathway network is obtained. The edges with an edge length less than this value form a spectrum /cluster. This small cluster, a subset of the full network can now be analysed for finding out a drug target.

ACKNOWLEDGMENT

The authors sincerely thank the Director, BITS Pilani, Goa Campus for providing all the facility to conduct the study.

REFERENCES

Ma, H.W. and A.P. Zeng, 2003. The connectivity structure, gaint strong component and centrality of metabolomic networks. Bioinformatics, 19: 1423-1430.

Patra, S.M. and S. Vishveshwara, 2000. Backbone cluster identification in proteins by a graph theoretical method. Biophys. Chem., 84: 13-25.

Van Helden, J., L. Wernisch Gilbert and S. Wodak, 2002. Grpah Based Analysis of Metabolomic Networks. In: Bioinformatics and Genome Analysis, Mewes, H.W., H. Seidel and B. Weiss (Eds.). Ernst Schering Research Foundation, Springer Verlag, pp: 245.

Verkhedkar, K.D., K. Raman, N.R. Chandra and S. Vishveshwara, 2007. Metabolome based reaction graphs of *M. tuberculosis* and *M. leprae*: A comparative network analysis. PLoS ONE, 2: e881-e881.

Vishweshwara, S., K.V. Bindra and N. Kanna, 2002. Protien structure: Insights from graph theory. J. Ther. Comput. Chem., 1: 187-211.

Wolfang, H., J. Vincent, L. Cagney Li, S. Falcon and G. Robert, 2007. Graphs in molecular biology. Biomed. Central, 8: 4-14.