Bioinformation by Biomedical Informatics Publishing Group *open access*
www.bioinformation.net **Software**

PINAT1.0: Protein interaction network analysis tool

Sandeep K. Kushwaha*, Madhvi Shakya

Department of Bioinformatics, MANIT, Bhopal 462051, India;
Sandeep K. Kushwaha - E-mail: sandeepkushwaha09@gmail.com; Phone: +91 0755-2671950; Fax: +91 0755- 2670562; *Corresponding author

Received March 1, 2009; revised April 1, 2009; accepted April 8, 2009; published July 27, 2009

Abstract:
Cellular processes are regulated by interaction of various proteins i.e. multiprotein complexes and absences of these interactions are often the cause of disorder or disease. Such type of protein interactions are of great interest for drug designing. In host-parasite diseases like Tuberculosis, non-homologous proteins as drug target are first preference. Most potent drug target can be identifying among large number of non-homologous protein through protein interaction network analysis. Drug target should be those non-homologous protein which is associated with maximum number of functional proteins i.e. has highest number of interactants, so that maximum harm can be caused to pathogen only. In present work, Protein Interaction Network Analysis Tool (PINAT) has been developed to identification of potential protein interaction for drug target identification. PINAT is standalone, GUI application software made for protein-protein interaction (PPI) analysis and network building by using co-evolutionary profile. PINAT is very useful for large data PPI study with easiest handling among available softwares. PINAT provides excellent facilities for the assembly of data for network building with visual presentation of the results and interaction score. The software is written in JAVA and provides reliability through transparency with user.

Availability: PINAT is available at www.manit.ac.in/pinat

Keywords: PINAT, Co-evolution, Drug Target, PPI, Protein interaction network

Background:
Protein interaction network analysis is basic need for the understanding of functional association of proteins in cellular processes **[1]**. Protein interaction networks are being used in tasks such as assignment of function to uncharacterized proteins and searching for most and least association between proteins for various purposes **[2]**. Since almost all cellular processes are regulated by protein-protein interaction, so any interruption in interactions may be a cause of disorder or disease **[3]**. So, it is necessary to know the cellular impact of every association (inhibition or activation) on multiple signaling pathways. Various methods are existing for the study of protein-protein interaction i.e. Co-Evolution, Co-Occurrence, Co-Expression, Gene Fusion etc. **[4, 5]**. In co-evolution process, two or more species interact and influence genetic changes in one another. The process is also evident at the molecular level, where interacting proteins exhibit coordinated mutations to evolve at a similar rate **[6, 7]**. As a result, interacting proteins will seem to evolve at the same rate and have similar evolutionary histories. Observed co-evolution between interacting proteins has been used previously to predict protein interaction sites and to improve docking algorithms **[8]**. Cellular processes are regulated by interaction of various proteins i.e. multiprotein complexes and absences of these interactions are often the cause of disorder or disease. Such type of protein interactions are of great interest for drug designing **[9]**. Nowadays, protein-protein interaction is one of the recent approaches for drug target identification for various types of diseases. PINAT has been developed for drug target identification with the special relevance to host-parasite disease "tuberculosis".

Methodology:

Theoretical concept of PINAT

In the present work, well established Co-evolutionary profile model has been used for identification of interacting proteins **[6, 7, 10, 11]**. Identification of interacting proteins consists of the following sequential steps: (i) Retrieval of proteins and its orthologous sequences from KEGG databases (http://www.genome.jp/kegg) (ii) Construction of distance matrices among retrieved sequences for each proteins (iii) Upper triangular part of the distance matrix was transformed into a phylogenetic vector (iv) Normalization of the elements of each vector and (v) Interacting protein identification and network construction through correlation coefficient.

Software Input and Output

PINAT provides user defined (no. of protein sequences) browsing facility for loading of protein sequences files. After entering of desired number of proteins for protein-protein interaction analysis and network building, the browsing window will open. This window will browse the entered number of protein files. For large data, browsing can take a few seconds for uploading a browsed protein file. PINAT computes distance matrix (**Figure 1a**), Phylogenetic vector (**Figure 1b**), correlation coefficient matrix, interaction network (**Figure 2a**) and potential interactants (**Figure 2b**).

Network Construction and Potential Interaction Table

PINAT also provides the facility to visualize the computed association network among proteins as a non-directional graphical network. Network association is shown by potential interactions. Protein association interaction network has been developed considering correlation coefficient values (r-values).

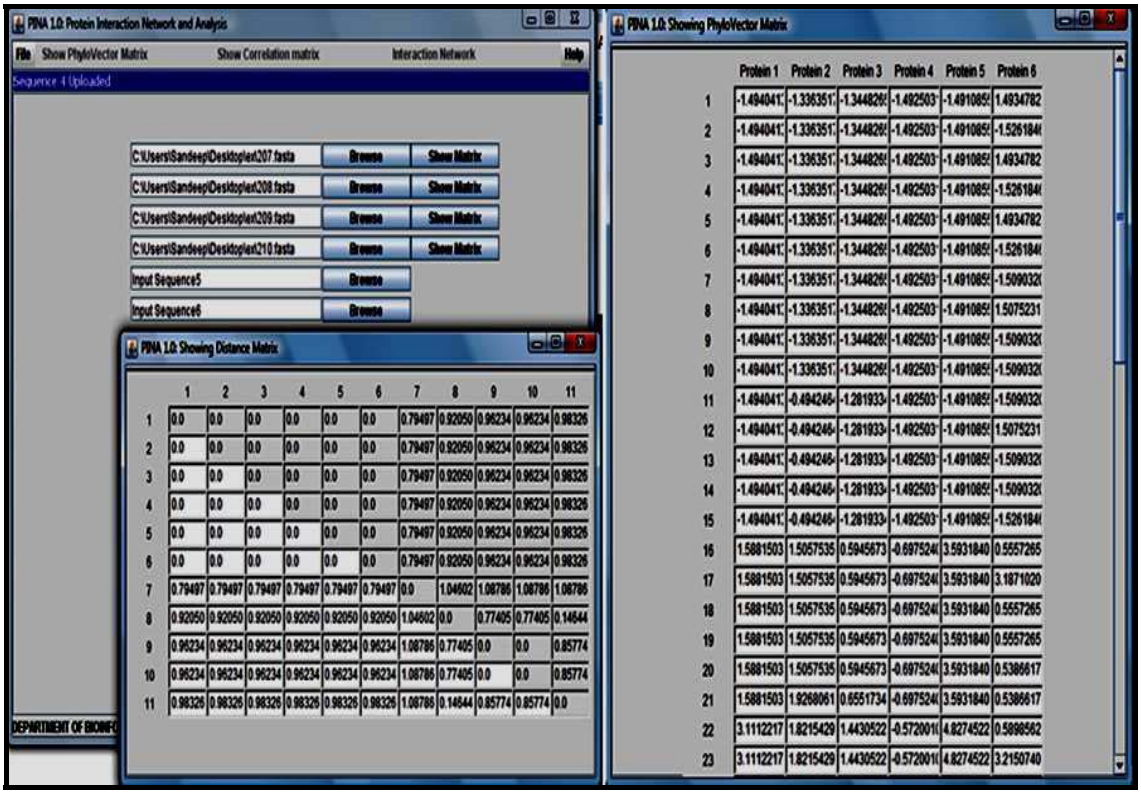

Figure 1: (a) Representative screen of PINAT with browse protein sequence along with distance matrix (b) Representative Screen of PINAT After phylovector generation.

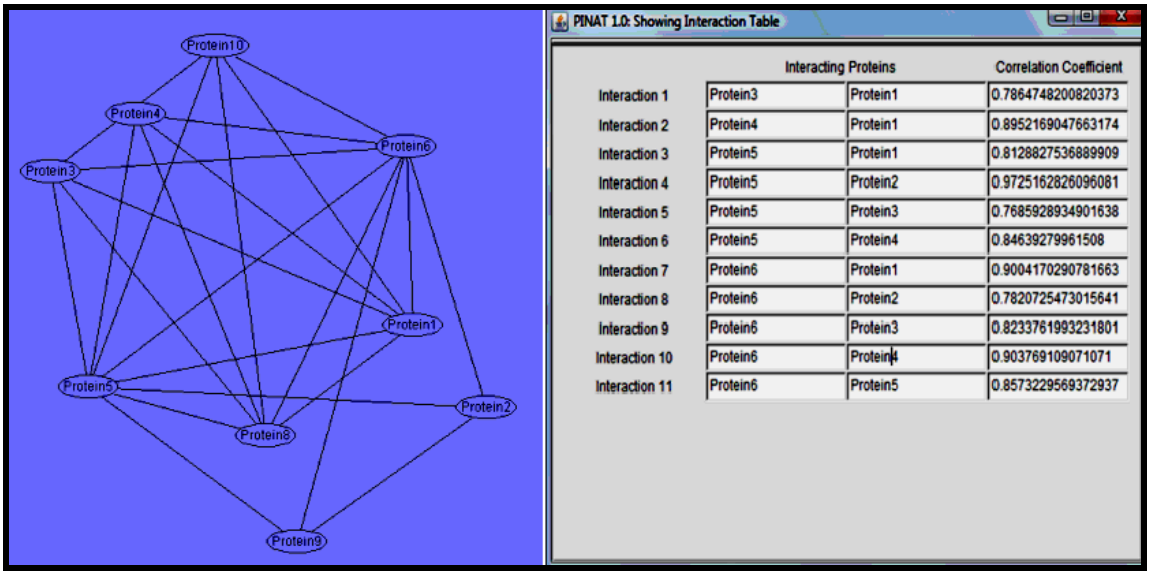

Figure 2: (a) Representative screen of PINAT with interaction network (b) Interaction list with potential interactants.

Acknowledgement:
We are grateful to Department of Bioinformatics, MANIT, Bhopal, M. P. India for support and cooperation and we are also thankful to Mr. Mohit Jha for many useful discussions and technical support.

References:

[1] L. Salwinski, *Curr. Opin. Struct. Biol.* 13: 377–382(2003) [PMID: 12831890].
[2] M. A. Huynen, *Curr. Opin. Cell Biol.* 15: 191–198(2003) [PMID: 12648675].

[3] A.L. Barabasi, *Nature Reviews Genetics* 5: 101-113(2004) [PMID: 14735121].
[4] R. A. Craig, *BMC Bioinformatics* 8:6 (2007) [PMID: 17212819].
[5] T. Sato, *Bioinformatics* 22: 2488-2496 (2006) [PMID: 16882650].
[6] T. Sato, *Bioinformatics* 21(17): 3482-9 (2005) [PMID: 15994190].
[7] H. B. Fraser, *Science* 296:750–752 (2002) [PMID: 11976460].
[8] M. Jucovic, *Proc. Natl. Acad. Sci.* 93:2343–2347(1996) [PMID: 8637875].
[9] A. I. Archakov, *Proteomics.* 3(4):380-91 (2003) [PMID: 12687606].
[10] S. H. Tan, *Nucleic Acids Res*. 32: W69–W72 (2004) [PMID: 15215353].
[11] F. Pazos, *Protein Eng.* 14(9): 609-14 (2001) [PMID: 11707606].

Edited by P. Kangueane
Citation: Kushwala *et al***,** Bioinformation 3(10): 419-421 (2009)