



Research Journal of Pharmaceutical, Biological and Chemical Sciences

Pathvisio++: An Interactive Web Server for Analyzing Metabolic Pathways.

Udayakumar Mani*, Krishna Nishand M, Raghu Ram G, and Sai Mukund Ramakrishnan.

Department of Bioinformatics, School of Chemical and Biotechnology, SASTRA UNIVERSITY, Tanjore-613402, Tamilnadu, India.

ABSTRACT

Metabolic Pathways represents organized information about genes, proteins and metabolites. Wikipathways is a pathway database designed and managed by BigCat which contains both metabolic and proteomic pathways. The pathways in Wikipathways are of GPML format which is unique to itself. Wikipathways is an extension of GenMAPP pathway collection. Pathvisio is the only software designed to interpret and analyse GPML files. Pathways can be directly edited from a java applet editor where you can draw genes, proteins and their interactions like in any popular drawing tool. It is a static interface which shows only the pictorial representation of the pathways. Pathvisio++ is a highly interactive web based server which improves the functionality of Pathvisio, making it highly user friendly and efficient tool for analysing pathways. It is designed using XML, PHP, MySQL, JavaScript, HTML, Ajax, Perl, Bio Perl, HTML Maps and Apache server integrated with the WAMP server. It has many features like JavaScript viewer for visualization of pathways, ontologies for the genes involved in the pathways, modifying cellular location and export the pathways as annotated table. Pathvisio++ webserver can be browsed at <http://scbt.sastra.edu/pathvisio>.

Keywords : pathways, regulation, GenMAPP, GPML, Ontologies, small molecule, the PHP GD.

**Corresponding author*

INTRODUCTION

Pathways are represented as networks which provide intuitive views of the myriad of interactions. They can represent receptor-binding events, protein complexes, phosphorylation reactions, translocations and transcriptional regulation, with only a minimal set of symbols, lines and arrows. These pathways are stored in different types of file formats like GPML (GenMap Mark-up Language), SBML (1), XML (Extensible Mark-up Language) etc. There are many databases to store such formats like Wiki pathways, KEGG (2). Wikipathways (3) is a pathway database designed and managed by BigCat which contains both metabolic and proteomic pathways. Wiki Pathways thus presents a new model for pathway databases that enhances and complements on-going efforts, such as KEGG, Reactome (4) and Pathway Commons. The pathways in Wikipathways are of GPML format which is unique to itself. Wikipathways is an extension of GenMAPP (5) pathway collection. Building on the same MediaWiki software that powers Wikipedia, we added a custom graphical pathway editing tool and integrated databases covering major gene (6), protein, and small molecule systems (7). The familiar web-based format of Wiki Pathways greatly reduces the barrier to participate in pathway curation. More importantly, the open, public approach of Wiki Pathways allows for broader participation by the entire community, ranging from students to senior experts in each field. This approach also shifts the bulk of peer review, editorial curation, and maintenance to the community. PathVisio, a standalone pathway editor and pathway analysis tool. Besides drawing pathways, it can be used to visualize data. PathVisio is suitable for the creation and exploration of pathways, while relying on GenMAPP for visualization of experimental data and MAPPFinder for statistical analysis (8). PathVisio is written in the Java programming language as opposed to Visual Basic. Thus, PathVisio is cross-platform, easier to integrate with other scientific software (often written in Java), and works with web technologies such as Java applets and Java Web start. The existing GenMAPP pathway archive can be used in PathVisio and other GPML-compliant programs. GPML has already been extended with new shapes and the capability to define relationships between nodes, allowing a network view of the pathway. Because of the novel features of GPML, it is preferable to use this format for pathway storage even if the MAPP format is used in later analysis steps.

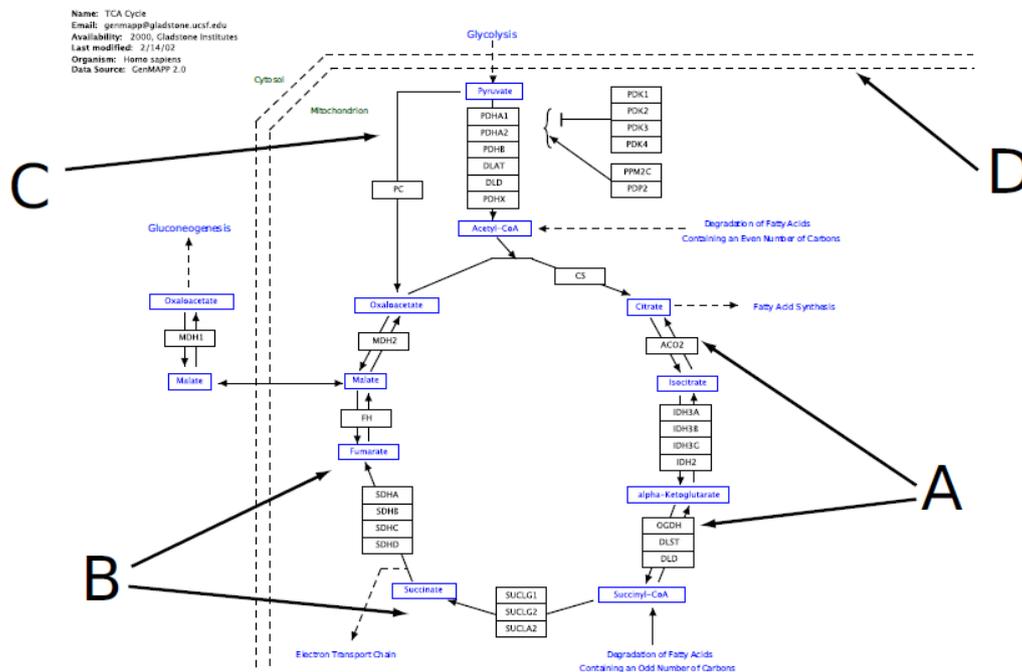


Figure 1: A biological pathway as represented by PathVisio has three main classes of objects:

Data Nodes, Lines and Shapes. (A) Data nodes can represent genes, proteins, Metabolites (B) Data Nodes can be linked to an online database. (C) Various shapes and text labels can be used to explain the pathway. (D) Pathways are stored in the GPML file format.

MATERIALS AND METHODS

A)PATHVISIO++ Implementation:

In Pathvisio++, the GPML file is read with the help of PHP and various data nodes are extracted like shapes, labels, data nodes and lines as separate entities. It uses the in-built library functions to extract all the XML tags present in the GPML file of our interest. Now all the entities that are extracted are stored as arrays which are used to construct the pathway diagram using PHP GD module. PHP(9) cannot be executed without a server and it needs a powerful server like Apache web server to be executed. Pathvisio++ makes use of PHP in almost all the scripts i.e. the pathway image creation and the linking among modules with their corresponding Perl scripts. MySQL(10) is the relational database management (RDBMS) that runs on the web server and acts as the backend for all the modules. It provides multi-user access to a number of databases Pathvisio++ is one such profile constructed using MySQL. The cellular location module makes use of MySQL which is connected through a PHP script to interact with the backend. Pathvisio++ receives the XML formatted data in the form of a GPML file from Wiki pathways and that acts as the input for the PHP GD module to convert it into a JPEG file as a pictorial representation. Pathvisio++ make use of AJAX(11) in retrieving information about which GPML file has been uploaded in the server for analysis from the browser cookie. The tool has used JavaScript(12) mainly for the visualisation module and hence displays the pictorial representation of the GPML file through the scripting. It brings an interactive interface for the visualisation and the analysis of pathways. Several modules in Pathvisio++ for the extraction work of information from NCBI database(13) have been extracted using Perl and combination of BioPerl(14) in it. This helps in extracting the contents with a string match terms effectively and in a simplistic manner.

B)Modules in PATHVISIO++:

i) Pathway Viewer for Wikipathways:

The objective is to create an interactive Pathway viewer enabling the users to view the pathway as a visualisation medium and simultaneously creating a JPEG image of it. The image is then split as different phases and just by clicking the gene in the image, all details can be fetched about the gene clicked. The underlying principle for this is HTML maps(15) which splits the image into several fragments thus each fragment having its on functionality. The module implementation is shown in **Figure2**.

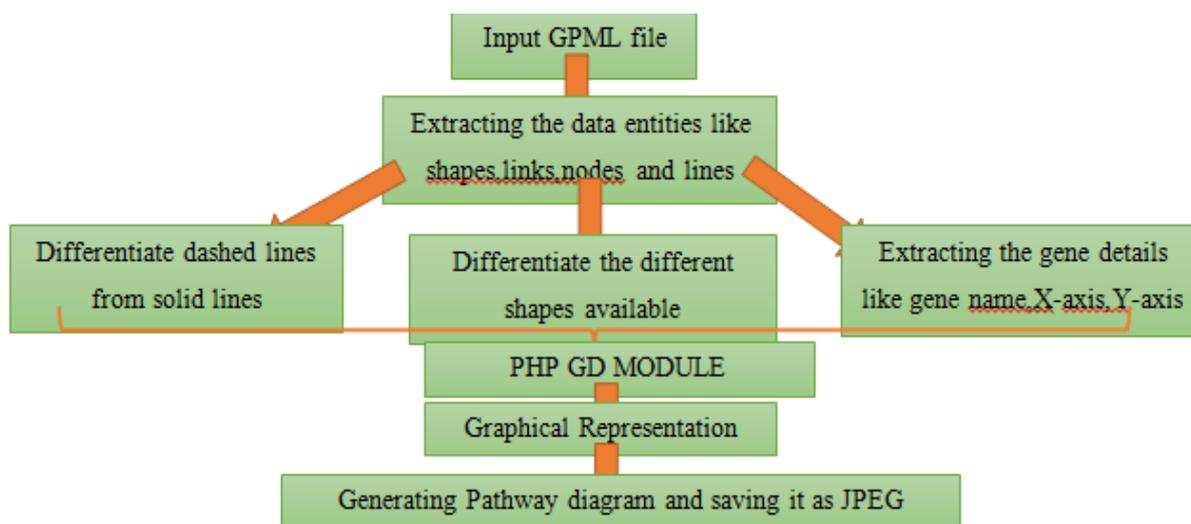


Figure2: Schema Diagram of Module 1 –Pathway Viewer for Wikipthways using GPML files in Pathvisio++

ii) Ontologies for Wikipathways:

Organizing and categorizing pathways is important to help finding relevant knowledge for a given topic. It can also help to identify related pathways or duplicate information that can be merged. Especially on a

Wikipathways, where users are free to create context-specific pathways (e.g. focused at a single tissue, cell-line or experiment); it is important to provide information about this context. Several biological ontologies exist that define a vocabulary of biological context that can be used for pathways. The module implementation is shown in **Figure3**.

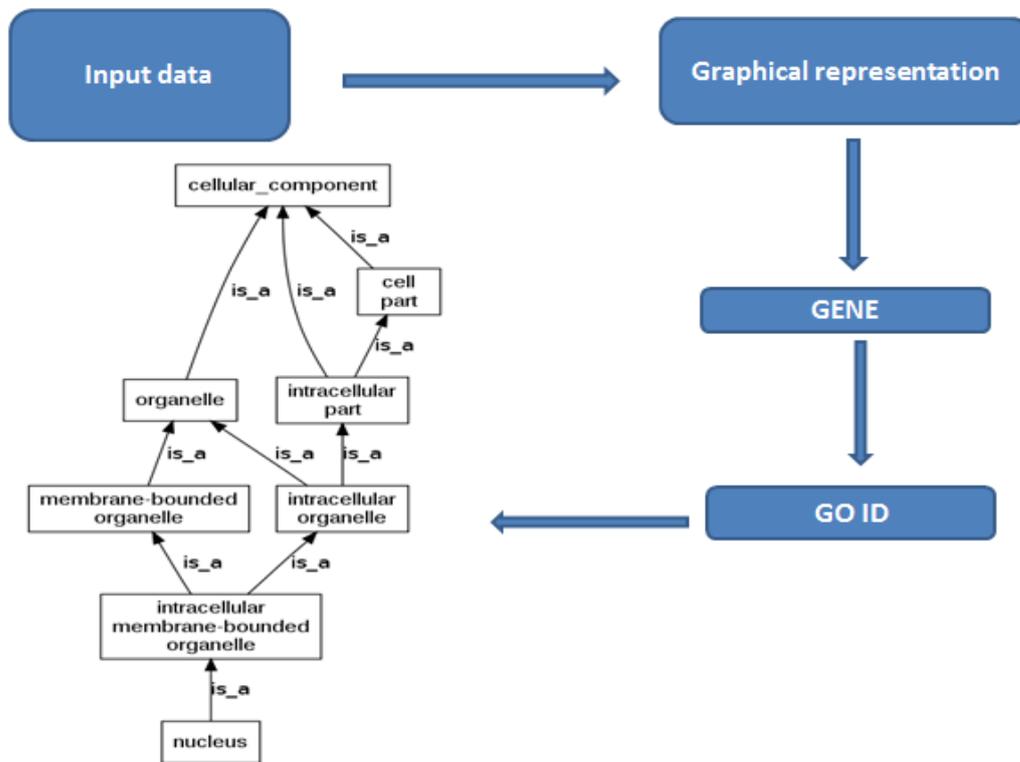


Figure3: Schema Diagram for finding Gene Ontologies in Pathvisio++

iii) Export Pathway as a Fully Annotated Table:

To export a pathway as a simple list of identifiers. To get, instead of a list, a table with not only the basic identifiers, but as much extra annotation as possible, such as gene name, description, and identifiers from all large databases such as Ensembl, Entrez, Unigene and Uniprot. The implementation of the module is shown in **Figure4**.

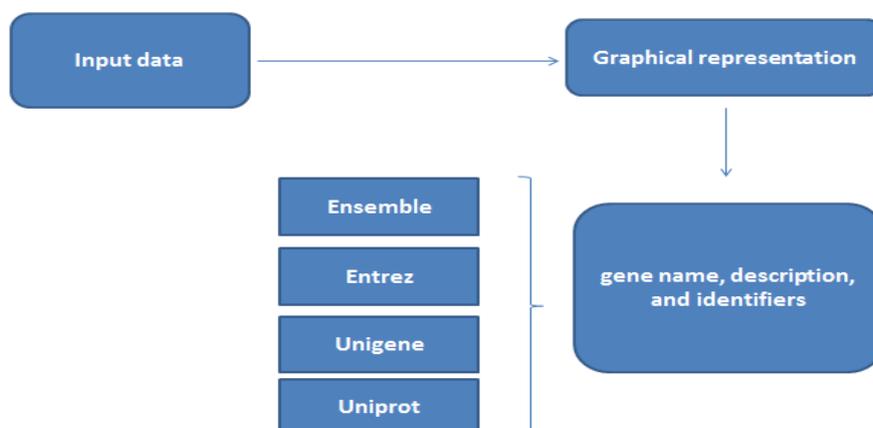


Figure4: Schema Diagram for Pathway as Annotated table

iv) Cellular Location of Genes:

In this module, the users can change the cellular location of the genes of their interest. A user friendly interface will be provided to the user to enter the co-ordinates of their interest and hence once on entering the location of their interest, they can easily change the cellular location of the genes. The image newly created will show the gene in the newly assigned location. This image will also be of JPEG type. **Figure5** shows the implementation of Cellular location of gens in Pathvisio++ and **Figure6** depicts the various modules implementation in the web server application .

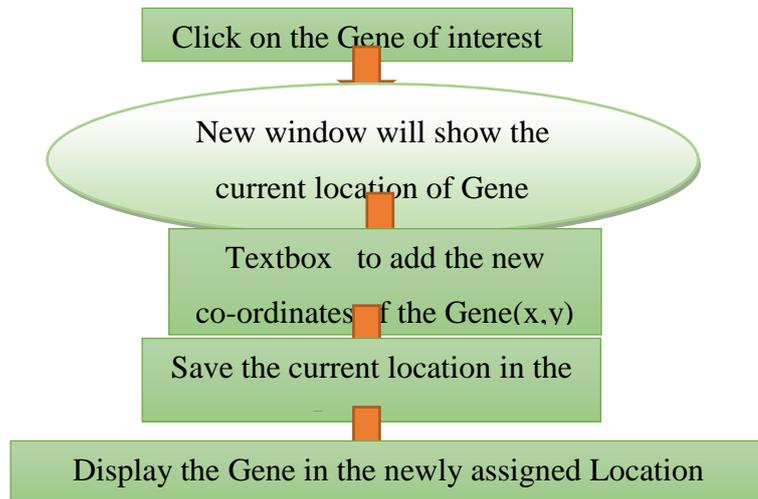


Figure5:Schema Diagram for change of Cellular Location of Genes in Pathvisio++

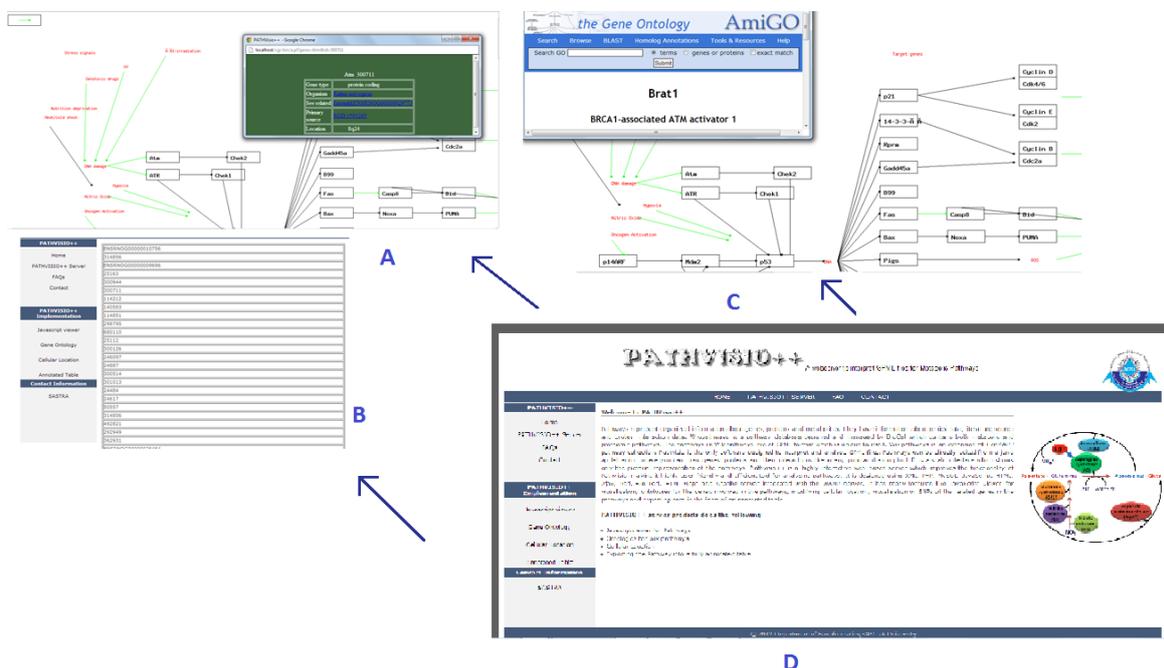


Figure6:PATHVISIO++ WEB SERVER (A)JavascriptViewer generated from GPML file using Pathvisio++.Box showing the Gene of interest clicked on the generated image from (A).Ontologies for Pathways .(C)Pathway as a Fully Annotated table .(D)Pathvisio++ Webserver Homepage.

CONCLUSION

Pathvisio++, an online server is an upgraded version of Pathvisio. It has taken into consideration the limitations of Pathvisio and also has several unique features to itself. So to enable all the

features, the GPML file has been first interpreted and all the useful information for drawing the image has been identified and with a powerful PHP algorithm, the pathway image has been drawn in JPEG format. It is later disintegrated into several small fragments where the user can click on the gene of interest and all information can be retrieved about the gene that has been clicked. This is done using HTML maps. The ontology module shows the list of ontologies related to the gene and on clicking the ontology; it shows the ontology flow for the gene with the functions. This information can be used in knowing the function of the gene in various ontologies, i.e. the maximum possible functions of the gene. Thus user can also change the cellular location of the gene. Pathvisio++ provides all the necessary information required for a gene by interpreting a GPML file that is reserved to Wikipathways alone. Thus a tool that can be interactive with a GPML file format is an advantage to the user.

REFERENCES

- [1] Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novère N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J; SBML Forum. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. 2003 Mar 1; 19(4):524-31.
- [2] Kanehisa, M. and S. Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*. 2000; 28, 27–30.
- [3] Pico AR, et al . WikiPathways: pathway editing for the people. *PLoS Biol*. 2008; 6:e184.
- [4] Joshi-Tope G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de Bono B, Jassal B, Gopinath G.R, Wu G.R, Matthews L, Lewis S, Birney E, Stein L. Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res*: 2005, January: 33(Database issue), D619–622
- [5] Dahlquist KD, Salomonis N, Vranizan K, Lawlor SC, Conklin BR: GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways. *Nat Genet* 2002, 31(1):19-20.
- [6] Udayakumar Mani, Krishna Nishand M , Raghu Ram G, Pharm J. Tool for Visualisation the Gene Loci of Multiple Genes: *Sci. & Res*. 2014, Vol. 6(11), 347-350
- [7] Udayakumar M, Prem Chandar D, Arun N, Mathangi J, Hemavathi K, Seenivasagam R: PMDB: Plant Metabolome Database—A Metabolomic Approach *Medicinal Chemistry Research*, 21(1): 47-52. 10.1007/s00044-010-9506-z
- [8] Alexander R Pico, Thomas Kelder, Martijn P van Iersel, Kristina Hanspers, Bruce R Conklin, and Chris Evelo : WikiPathways: Pathway Editing for the People 2008 July; 6(7): e184. PMID: PMC2475545
- [9] Julie C. Meloni, 'PHP essentials', II edition, Premier Press
- [10] <http://www.mysql.com>.
- [11] Lee Babin, 'Beginning AJAX with PHP', A press
- [12] Scott Duffy, 'How to do everything with JavaScript', Mc Graw-Hill.
- [13] NCBI website: <http://www.ncbi.nlm.nih.gov/>. Accessed on November 12 2014
- [14] <http://www.bioperl.org>
- [15] <http://www.w3schools.com/googleapi/>.