PATIKA*web*: A Web service for querying, visualizing and analyzing the graph-based PATIKA pathway database

O. Babur, E. Belviranli, D. Berktas, A. Cetintas, R. Colak, E. Demir[†], A. Dilek, U. Dogrusoz, E.Z. Erson, E. Giral, K. Karatepe, H. Kucuk, C. Kucukkececi, O. Kucuktunc, and S. Tekin

Bilkent Center for Bioinformatics, Bilkent University, Ankara 06800, Turkey

[†]Presenting author

Although there has been an enormous amount of effort for creating ontologies, standards and tools, current bioinformatics infrastructure is far from coping with the dramatically increasing amount of biological data. The PATIKA Project aims to provide the scientific community with an integrated environment for modeling, analyzing and integrating cellular processes.

Central to the PATIKA Project is an ontology, based on a qualitative two-level model. At entity-level, interactions and relations can be addressed in an abstract manner, where the exact state of the related parties is unknown, such as protein-protein interactions, inferred relations and literature-derived information. At state/transition-level, each entity is associated with a set of states interacting with each other via transitions. This level can capture more detailed information such as compartments, molecular complexes and different types of biological events (e.g. covalent modification, transportation and association). This two-level ontology elegantly covers most biological phenomena and is capable of integrating information present in the literature and molecular biology databases. PATIKA Ontology also uses the notion of compound graphs to represent abstractions, which are logical groupings used to handle complex and incomplete nature of the data.

The PATIKA Server acts as a central database and provides XML-based Web services for querying and integrating pathway models. Currently the server contains data integrated from several public databases including Reactome (around 5000 states and over 1600 transitions/reactions), and interfaces with major databases and ontologies such as Entrez-Gene, UniProt and GO. A CVS-like concurrent versions system is used for integration logic, allowing continuous and concurrent updates to the server via other PATIKA tools.

Perceiving the increasing demand on easily accessible tools, PATIKA*web* supplies a Web-based service with a user-friendly interface without requiring any registrations or installations. Since it is expected that majority of PATIKA users will only be interested in knowledge acquisition, PATIKA*web* focuses on providing a simple yet powerful interface for querying and visualizing the PATIKA database. PATIKA*web* fully supports the PATIKA ontology, including a multiple-view schema for bioentity and mechanistic levels, compartments and compound graphs for visualizing molecular complexes, pathways and black-box reactions. A specialized algorithm which can handle both compound graphs and compartments is used to automatically layout pathways. Graph editing functionalities such as scrolling and zooming, do/undo and delete as well as advanced features such as expand/collapse of nested abstractions are available to manage the current pathway model. PATIKA*web* implementation uses the JSP (JavaServer Pages technology) edition of the Tom Sawyer Visualization technology to handle this highly-dynamic and advanced visual content.

Querying component of PATIKA both supports SQL like queries and an array of graph-theoretic queries for finding feedback loops, positive/negative paths, common targets and regulators, or “interesting subgraphs” based on user’s genes of interest. Once retrieved from the database, the query results may be merged to the user’s current view and highlighted to provide an incremental user-friendly retrieval and analysis interface. Constructed models can be saved in XML, exported to standard formats such as BioPAX and SBML or converted to static images. Work for importing from BioPAX is underway.

Microarray data analysis component of PATIKA facilitates analysis of expression data, or comparison of two related experiments, on top of the pathway data, through visual techniques such as color-coding and labeling. Work for support of methods on pathway activity inference using expression data and methods for clustering, based on expression data is also underway.

We believe that PATIKA*web*’s unique visualization and querying features, coupled with its user-friendly Web-based interface fills an important gap in the pool of currently available tools and databases. In parallel to the dramatic expansion in its data and features, we expect PATIKA*web* to quickly become a popular tool among molecular biologists.

Avalilability: Freely available (http://web.patika.org) for non-profit use