

PDWSB: Public Domain Web Servers in Biology

G. P. S. Raghava

Institute of Microbial Technology, Chandigarh, India

Abstract

Web servers allow biologists to perform various tasks on-line over the Internet such as 1) access/retrieval of data from biological databases, 2) sequence analysis, 3) homology searching, 4) protein modelling, and 5) gene prediction. The number of web servers of biological interest in

the public domain is rapidly increasing. In order to keep track of these web servers, we created a database called PDWSB. This database collects and maintains information on web servers in biology. Presently, this database contains information about more than 700 web servers. All web servers are classified in categories based on their function. In order to provide a user-friendly interface, a search engine has been developed to retrieve information from PDWSB (www.imtech.res.in/jpdwsb). An email server was also developed to extract information from the database via e-mail

(freeware@imtech.chd.nic.in).

PDWSB provides a platform to the authors of a web server, from which they can distribute information about their server rapidly.

Introduction

Bioinformatics is the convergence of two technology revolutions: the explosive growth in biotechnology paralleled by the explosive growth in information technology (1-3). In the last 10 years, both the size of sequence databases (e.g., GenBank) and the power of computers have been doubling at about the same rate: every 18-24 months. The Internet has radically affected the way data are provided, handled, and analyzed. This powerful combination of data and tools, which allow easy access and analysis has changed and will continue to change our approach to the design and practice of biological research (4).

One of the major challenges for researchers or academicians working in the field of bioinformatics is to manage the growing biological resources available on the Internet. It is important for the bench scientist working in biology to have easy and efficient ways of wading through the data and finding what is important to his or her research. Broadly, the information on the Internet can be divided into two categories: static and dynamic. The static information is that found in the

form of web pages, where one can read the information on web pages as provided by its author. It is also called a one-way flow of information. On the other hand, dynamic web servers allow the user to retrieve desired information from databases. These dynamic web servers provide an interface between user and database. For example, almost all databases in molecular biology are accessible via a web browser (5,6).

Dynamic web servers are not limited to the access of information over the Internet, but also provide on-line computational tasks like 1) genome analysis, 2) protein structure prediction and, 3) protein sequence analysis (7,8). In order to utilize the full potential of this powerful resource, it is important to manage information on web servers. Recently, a comprehensive catalog of biological databases called DBcat was created, which contains information on roughly 500 databases in the form of a structured flat-file (5,6). There are a number of search engines that maintain information about web pages and web servers. It is difficult for biologists to extract useful information from these general-purpose search engines.

This paper describes a database that collects and maintains information on web servers of biological interest that are in the public domain and that are free for potential users. This is a human-edited database that allows a user

<p>to search the database by keyword or by directory. One of the unique features of this database is that it is accessible either via web browser (imtech.chd.nic.injpdwsb or www.lmtech.res.in) or via e-mail (freeware@imtech.chd.nic.in).</p> <p>Organization</p> <p>PDWSB stores the following information about a web server: 1) title of server, 2) location or URL, 3) category of web server, 4) description of server, 5) function, and 6) reference of server. All data is stored in PostgreSQL, which is a powerful Relational Database Management System (RDBMS).</p> <p>Accessing PDWSB</p> <p><i>Web Server</i></p> <p>A web server has been developed to access the information from the PDWSB database (imtech.chd.nic.injpdwsb). Users can access information from our server either by query search or by category browsing. Query search allows a user to query the database on any field such as title, function, description, keyword, etc. It provides a number of options such as: 1) case sensitive or insensitive search, 2) search only in a category, 3) select the fields to display, or 4) select the field on which the result is to be sorted. Category browsing allows a user to search for the desired site by browsing a relevant category which the user can select. This strategy is very useful when the user does not know which</p>	<p>server will be useful for their work. By visiting a desired category, users can easily find out which web servers will be useful for their work.</p> <p><i>E-mail Server</i></p> <p>We have also developed an e-mail server in order to provide off-web service. This e-mail server allows a user to extract information about web servers from the PDWSB server via e-mail. For a detailed description, send an e-mail to freeware@imtech.chd.nic.in with the word "help" in the body of the message.</p> <p><i>Submit A Web Server</i></p> <p>In order to submit a web server to the PDWSB database, a user needs to fill out a simple form available at the PDWSB web site. The following information about the web server needs to be entered: 1) title, 2) location, 3) description, 4) category, 5) function, 6) keyword, and 7) reference. Alternatively, an author can submit the information about the web server by e-mail (pdwsb@imtech.chd.nic.in).</p> <p>Discussion</p> <p>The PDWSB Server is intended to be a central repository of information on public domain web servers in the bioscience field. It will assist the scientific community in searching for a desired site from a single source. We are also assisting authors of web servers in the management and distribution of information about their servers.</p>	<p>We would greatly appreciate submission of new web servers directly from authors. For any query comment and suggestion, please contact the author (freeware@imtech.chd.nic.in).</p> <p>Acknowledgement</p> <p>The author is grateful to Dr. Amit Ghosh, Director, IMTECH Chandigarh, and Dr. J. R. Arora, adviser, DBT, Delhi, for their encouragement, support, and help in the creation of this database. The author is also grateful to Ms. Trantum Kaur for her assistance in the collection of software.</p> <p>References</p> <ol style="list-style-type: none"> Gentile, J.A. (1998). <i>Oncology</i>, 11, 356-9. Joubert, M, Volot, F., Fieschi, D., & Fieschi M. (1998). <i>Medinfo</i> 1, 161-5. Lugo-Vicente, H. (2000). <i>Semin. Pediatric Surg.</i> 1, 11-8. Bard J.B. & Davies JA (1995). <i>Bioassay</i> 11, 999-1001. Baxevanis, Andreas D. (2000). <i>Nucleic Acids Research</i>, 28, 1-7. Discala, C. Benigni, X. Barillot, & E. Vaysseix. (2000). <i>Nucleic Acids Research</i> 28, 8-9. Harper, R. (1994) <i>Curf. Orin. Biotechnoll</i>, 4-18. Eddy S. (1999). <i>Proteins</i> 37, 209-17. <p>Address for Correspondence</p> <p>Dr. G. P. S. Raghava Scientist, Bioinformatics Centre Institute of Microbial Technology Sector 39A, Chandigarh, India Phone: +91 172 690557 Fax: +91 172 690632 E-mail: raghava@imtech.res.in Web: imtech.res.injraghava</p>
--	---	--