

Find out how to access preview-only content

Look inside Get Access

Contemporary Computing
Communications in Computer and Information Science Volume 168, 2011, pp 548-550

Updated Version of PHYTODB, the First Data Warehousing and Mining Web Server for Phytoplasma Research

Abstract

PHYTODB contains a repository of phytoplasma genes and proteins. It provides a unified gateway to store, search, retrieve, update information about phytoplasma and computational resources for the analysis of nucleotide and aminoacid sequence data of phytoplasma. Server facilitates to differentiate and classify new phytoplasma for taxonomic purposes. PHYTODB database was updated by dividing the whole resources into two domains: *DataBanks* and *Tools*. *DataBanks* serve as the storage device of all information. Functional characterization of genes and protein are done. Updated Groupidentifier tool by rearrangement of RFLP classification scheme of phytoplasma and possibilities 6 new groups based on the new tool. PhytoDB can be obtained through <http://220.227.88.253/phytodb/>.

Page %P

Page 1

Updated Version of PHYTODB, the First Data Warehousing and Mining Web Server for Phytoplasma Research

R. Manimekalai, P. Anoop Raj, O.M. Roshna, Anil Paul, and George V. Thomas

Department of Biotechnology, Agribioinformatics center,
Central Plantation Crops Research Institute, Kasaragod – 67 1124, Kerala, India
rmanimekalai@rediffmail.com,
apposklr@gmail.com, agribioros@gmail.com,
paulanilpaul@gmail.com, georgevthomas@yahoo.com

1 Introduction

PHYTODB contains a repository of phytoplasma genes and proteins. It provides a unified gateway to store, search, retrieve, update information about phytoplasma and computational resources for the analysis of nucleotide and aminoacid sequence data of phytoplasma. Server facilitates to differentiate and classify new phytoplasma for taxonomic purposes. PHYTODB database was updated by dividing the whole resources into two domains: *DataBanks* and *Tools*. *DataBanks* serve as the storage device of all information. Functional characterization of genes and protein are done. Updated GroupIdentifier tool by rearrangement of RFLP classification scheme of phytoplasma and possibilities 6 new groups based on the new tool. PhytoDB can be obtained through <http://220.227.88.253/phytodb/>.

2 Design of Database

PHYTODB has 3-tier organization where Web 2.0 technologies like AJAX provide high quality dynamic interfaces. Apache web server provides full range of Web server features to hand over the client requests. DHTML and JavaScript are used for developing user interfaces with the help of Web 2.0 technology. GroupIdentifier tool for taxonomic classification of phytoplasma was developed using Java Servlets, Java Server Pages and BioJava 1.5. Other tools are developed using PHP for querying and response management and Perl CGI scripts for result generation. MySQL is the database server used for data and sequence storage. Apache-Tomcat server is used for GroupIdentifier tool and Apache web server for the remaining tools. NCBI (National Centre for Biotechnology Information) stands for the primary data source from where the phytoplasma nucleotide and protein sequences were retrieved.

3 Architecture

The web server resources are categorized into two domains: *DataBanks* and *Tools*. The *DataBanks* functions as repository of all phytoplasma information assembled in distinct sections facilitating easy data retrieval. *Tools* domain contains Hlogs,

S. Aluru et al. (Eds.): IC3 2011, CCIS 168, pp. 548–550, 2011.
© Springer-Verlag Berlin Heidelberg 2011

MSalign, PhyloCass and GroupIdentifier. GroupIdentifier achieves phytoplasma group determination based on similarity calculation.

3.1 Databanks

PHYTODB has 5 different databanks that comprises of DBgene, ProtB, 16Sr Groups, G-Nome and electronic literature service providing important research works done worldwide on phytoplasma (E-Lite). The functional categorization of all phytoplasma genes and proteins have performed 1600 phytoplasma genes and 2188 proteins entries are stored in 'DBgene' and 'ProtB' databanks respectively. 16S rRNA gene region of 775 various phytoplasma species are organized in taxonomic groups according to classification scheme recognized by Phytoplasma Taxonomy Group of the International Research Program on Comparative Mycoplasma [1]. in '16Sr Groups' cluster. *G-Nome* contains information regarding phytoplasma whole genome sequencing projects of various phytoplasma.

3.2 Tools

The domain has embedded with 3 sequence analysis tools (Hlogs, Msalign and PhyloClass) and a phytoplasma taxonomic group identification tool. *Hlogs* determine homologues, based on BLASTN algorithm [2]. The output of the similarity search has taxonomic group tag in sequence description line. *MSalign* produces biologically meaningful multiple sequence alignments of divergent sequences. It calculates the best match for the selected sequences. User can perform multiple sequence alignment with our 16Sr group database by entering a single 16S sequence as input. *PhyloClass* constructs phylogenetic tree from molecular sequence data. Algorithm behind the tool performs multiple sequence alignment and drawn a phylogenetic tree. The output tree file can be viewed through any tree viewer program. *GroupIdentifier* performs the group identification of phytoplasma. It represents an attempt of phylogenetic classification based on the most conserved 16S rRNA gene sequence of phytoplasma genome.

In our classification scheme 36 distinct phytoplasma groups are exist based on similarity calculation between sequences. The group system is developed based on similarity cutoff value >98.64%, which is verified through MSA and phylogeny analysis of nearly full length (~1246nts) 16S rDNA sequences of ~200 phytoplasma species. At this cutoff each phytoplasma groups can be differentiated clearly.

4 Conclusions

PHYTODB, publicly accessible resource of phytoplasma, developed by various computer languages and software such as Web 2.0 technologies, DHTML, JavaScript and Apache web servers. Various bioinformatics tools are incorporated using the Perl CGI scripts for sequence analysis. The database embraces fundamental information on phytoplasma. Tools embedded are useful for homologues search, multiple sequence alignment, phylogenetic analysis and 16S rRNA based group identification of new phytoplasma.

Srinivas Aluru · Sanjhamitra Bandyopadhyay
Umit V. Catalyurek · Deydatt P. Dubhashi
Phillip M. Jones · Miroslav Pajarski
Berthold Schmidt (Eds.)

Communications in Computer and Information Science

Contemporary Computing

4th International Conference, ICC 2011
Mumbai, India, August 2011
Proceedings

 Springer

Chapter Metrics

Citations



1. IRPCM Phytoplasma/ Spiroplasma Working Team – Phytoplasma Taxonomy Group, - ‘Candidatus Phytoplasma’, a taxon for the wall-less, non-helical prokaryotes that colonize plant phloem and insects. *Int. J. Syst. Evol. Microbiol.* 54, 1243–1255 (2004)
2. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J.: Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410 (1990)

About this Chapter

Title

Updated Version of PHYTODB, the First Data Warehousing and Mining Web Server for Phytoplasma Research

Book Title

Contemporary Computing

Book Subtitle

4th International Conference, IC3 2011, Noida, India, August 8-10, 2011. Proceedings

Pages

pp 548-550

Copyright

2011

DOI

10.1007/978-3-642-22606-9_58

Print ISBN

978-3-642-22605-2

Online ISBN

978-3-642-22606-9

Series Title

Communications in Computer and Information Science

Series Volume

168

Series ISSN

1865-0929

Publisher

Springer Berlin Heidelberg

Copyright Holder

Springer-Verlag GmbH Berlin Heidelberg

Additional Links

- [About this Book](#)

Topics

- [Computer Communication Networks](#)
- [Artificial Intelligence \(incl. Robotics\)](#)
- [Algorithm Analysis and Problem Complexity](#)
- [Information Systems Applications \(incl. Internet\)](#)
- [Information Storage and Retrieval](#)
- [Computation by Abstract Devices](#)








Industry Sectors

- [Electronics](#)
- [Telecommunications](#)
- [IT & Software](#)

eBook Packages

- [eBook Package english Computer Science](#)
- [eBook Package english full Collection](#)

Editors

- Srinivas Aluru  ⁽¹⁾
- Sanghamitra Bandyopadhyay  ⁽²⁾
- Umit V. Catalyurek  ⁽³⁾
- Devdatt P. Dubhashi  ⁽⁴⁾
- Phillip H. Jones  ⁽⁵⁾
- Manish Parashar  ⁽⁶⁾
- Bertil Schmidt  ⁽⁷⁾

Editor Affiliations

- 1. Iowa State University and IIT Bombay, India
- 2. Indian Statistical Institute
- 3. The Ohio State University
- 4. Department of Computing Science, Chalmers University
- 5. Dept. of Electrical and Computer Engineering, Iowa State University
- 6. TASSL, Dept. of Electrical & Computer Engineering, Rutgers, The State University of New Jersey
- 7. School of Computer Engineering, Nanyang Technological University

Authors

- R. Manimekalai ⁽⁸⁾
- P. Anoop Raj ⁽⁸⁾
- O. M. Roshna ⁽⁸⁾
- Anil Paul ⁽⁸⁾
- George V. Thomas ⁽⁸⁾

Author Affiliations

- 8. Department of Biotechnology, Agribioinformatics center, Central Plantation Crops Research Institute, Kasaragod – 67, 1124, Kerala, India

Continue reading...

To view the rest of this content please follow the download PDF link above.

