YETI: Yeast Exploration Tool Integrator

R. J. Orton[1], W. I. Sellers[2] and D. L. Gerloff[1],*

[1]Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh, EH9 3JR, UK and [2]Department of Human Sciences, University of Loughborough, Loughborough, LE11 3TU, UK

ABSTRACT

Summary: Yeast Exploration Tool Integrator (YETI) is a novel bioinformatics tool for the integrated visualization and analysis of functional genomic data sets from the budding yeast *Saccharomyces cerevisiae*.

Availability: YETI is freely available for use over the WWW, or download under license, at http://www.bru.ed.ac.uk/~orton/yeti.html

Contact: orton@bch.ed.ac.uk

The budding yeast *Saccharomyces cerevisiae* is one of the most widely studied eukaryotic organisms due to its value as a model organism in biological research. Its complete genome sequence is well annotated (Goffeau *et al*., 1996) and a variety of large functional genomic data sets are publicly available. The *Saccharomyces* Genome Database (SGD) (Cherry *et al*., 1998) contains a wealth of genomic and biological information, is regularly updated and makes a large amount of its data publicly available. There are also a number of publicly available yeast two-hybrid protein–protein interaction data sets available from online databases and individual studies (Mewes *et al*., 1997; Uetz *et al*., 2000; Ito *et al*., 2001).

YETI

The Yeast Exploration Tool Integrator (YETI) is a novel bioinformatics tool for the integrated visualization and analysis of *S.cerevisiae* functional genomic data sets. YETI consists of a MySQL relational database for the storage and management of data and a Java program for visualization and analysis. YETI uses publicly available data from the SGD and individual studies to populate its own database and scripts have been written to keep the database up-to-date. The first release of YETI consists of three sections:

(1) A *Genome Section* for the informative display of the *S.cerevisiae* chromosomes, known and predicted genes and associated annotations. This section allows users to view the location of groups of genes on a genomic schematic (Fig. 1A). These groups can be specified by the users themselves or selected from predefined

---

*To whom correspondence should be addressed.

lists based on Gene Ontology (GO) annotations (Ashburner *et al*., 2000). Users are able to scroll along chromosomes, zoom in or out, click on genes of interest to view datasheets or select regions of interest to investigate further (Fig. 1B).

(2) A *Proteome Section* for the effective visualization of protein–protein interactions as inferred from yeast two-hybrid screens (Fig. 1C). A protein interaction map is generated by inputting the name of a protein of interest; YETI then dynamically displays all relevant interactions on a graphical panel, as inspired by Mrowka (2001). Some advanced features include reorganizing the map, interaction direction, deselecting, expanding out, datasheets and protein interaction paths. Users can also select multiple proteins or interaction clusters to investigate further.

(3) An *Analysis Section* provides users with a graphical interface to the database (Fig. 1D). One of the limitations with online databases is that the primary and sometimes only way of searching is by entering a single name and then viewing the results one at a time. YETI offers a much more dynamic and flexible search system that allows for complex query searches and provides an interactive data table to collectively view results.

All sections are closely inter-linked enabling users to move swiftly between sections and investigate all aspects of any genes or proteins of interest. Furthermore access to textual information, including GO annotations, is provided at any point.

DISCUSSION

YETI is written in the Java programming language and should run on any operating system that has access to a Java 1.4 or higher virtual machine; it has been tested under Windows, Linux and Mac OS-X. YETI can be used over the WWW via a Java applet or it can be downloaded, under license, with the MySQL database and installed locally. One of YETI's strengths is its user-friendliness towards wet-laboratory users who will use YETI as a 'workbench' to explore the data. This is combined with a solid database structure enabling
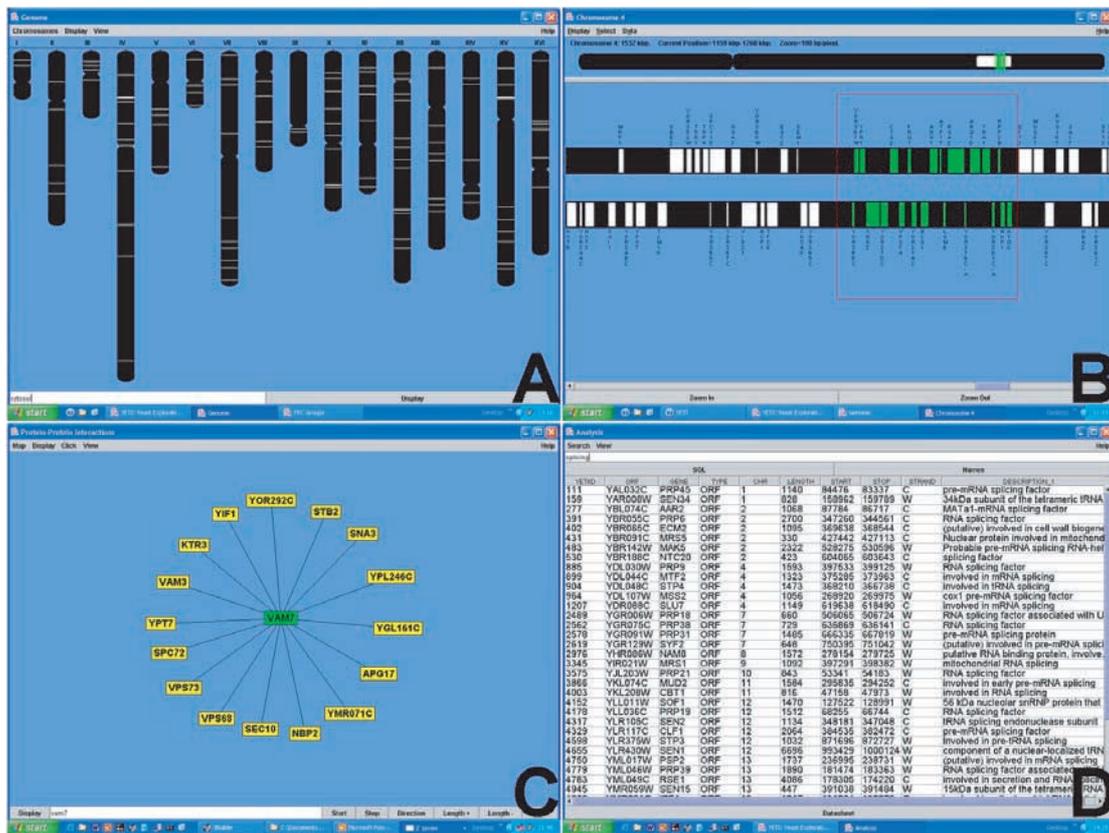

1

Fig. 1. YETI screenshots: (A) genome schematic; (B) chromosome scrolling; (C) protein interaction map; (D) interactive data table.

bioinformaticians to carry out collective analyses of the data sets. In addition, two new sections are currently under development. A Transcriptome Section will provide a means for effectively visualizing and analysing data from gene expression microarrays using both clustering algorithms and principal component analysis and a Cell Section will provide access to metabolic pathways. Although YETI's design is specifically aimed at *S.cerevisiae*, the framework of the program could be applied to other organisms with ease.

ACKNOWLEDGEMENTS

REFERENCES

Ashburner,M., Ball,C.A., Blake,J.A., Botstein,D., Butler,H., Cherry,J.M., Davis,P.A., Dolinski,K., Dwight,S.S., Eppig,J.T. *et al.* (2000) Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.

Cherry,J.M., Adler,C., Ball,C., Chervitz,S.A., Dwight,S.S., Hester,E.T., Jia,Y.K., Juvik,G., Roe,T., Schroeder,M. *et al.* (1998) SGD: *Saccharomyces* Genome Database. *Nucleic Acids Res.*, **26**, 73–79.

Goffeau,A., Barrell,B.G., Bussey,H., Davis,R.W., Dujon,B., Feldmann,H., Galibert,F., Hoheisel,J.D., Jacq,C., Johnston,M. *et al.* (1996) Life with 6000 genes. *Science*, **274**, 563–567.

Ito,T., Chiba,T., Ozawa,R., Yoshida,M., Hattori,M. and Sakaki,Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl Acad. Sci.*, **98**, 4569–4574.

Mewes,H.W., Albermann,K., Heumann,K., Liebl,S. and Pfeiffer,F. (1997) MIPS: a database for protein sequences, homology data and yeast genome information. *Nucleic Acid Res.*, **25**, 28–30.

Mrowka,R. (2001) A Java applet for visualizing protein–protein interaction. *Bioinformatics*, **17**, 669–670.

Uetz,P., Giot,L., Cagney,G., Mansfield,T.A., Judson,R.S., Knight,J.R., Lockshon,D., Narayan,V., Srinivasan,M., Pochart,P. *et al.* (2000) A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature*, **403**, 623–627.