

## jSquid: a Java applet for graphical on-line network exploration

Martin Klammer<sup>1,2,†</sup>, Sanjit Roopra<sup>1,2,†</sup> and Erik L. L. Sonnhammer<sup>1,\*</sup><sup>1</sup>Stockholm Bioinformatics Centre, Stockholm University, 10691 Stockholm, Sweden and <sup>2</sup>Department of Bioinformatics, Upper Austrian University of Applied Sciences, 4232 Hagenberg, Austria

Received on February 12, 2008; revised and accepted on April 28, 2008

Advance Access publication April 29, 2008

Associate Editor: Thomas Lengauer

### ABSTRACT

**Summary:** jSquid is a graph visualization tool for exploring graphs from protein–protein interaction or functional coupling networks. The tool was designed for the FunCoup web site, but can be used for any similar network exploring purpose. The program offers various visualization and graph manipulation techniques to increase the utility for the user.

**Availability:** jSquid is available for direct usage and download at <http://jSquid.sbc.su.se> including source code under the GPLv3 license, and input examples. It requires Java version 5 or higher to run properly.

**Contact:** erik.sonnhammer@sbcsu.se

**Supplementary information:** Supplementary data are available at *Bioinformatics* online.

## 1 INTRODUCTION

Protein networks can be visualized using a graph viewer like ProViz (Iragne *et al.*, 2005), Osprey (Breitkreutz *et al.*, 2003), SHARKview (Pinney *et al.*, 2005), Cytoscape (Shannon *et al.*, 2003) or Medusa (Hooper and Bork, 2005). Embedding these in a web site can be unsatisfactory, either because they cannot be run as an applet within a web page, or because the applet lacks functionality and usability.

Here we introduce jSquid, a Java applet providing various graph manipulation and exploration features. The program shares some core classes with Medusa, but most of the code is new and the functionality is greatly expanded. Although jSquid was developed for visualizing networks for the FunCoup database (<http://FunCoup.sbc.su.se>), it is a generic tool that can be used in other web sites for exploring graphs in order to investigate interactions of interest. A detailed comparison between jSquid and other online graph viewers can be found in Table S1.

## 2 OVERVIEW

### 2.1 User interface

The main panel displays the current graph, in which one can move single nodes or the whole graph and zoom in or out. On right-clicking a node or a selected group of nodes the user is offered a variety of options, such as changing the node's color

and shape, links to databases associated with this node or collapsing a group of nodes (see Section 2.3).

Via the top menu bar the user can change the appearance of the graph, such as redistributing the nodes, switching between summary and detailed edge views or toggling between different style switches. It also contains the grouping functionality (see Section 2.2), the preferences dialog, the export feature (see Section 2.4) and a detailed help function.

A set of tabs is placed to the right, where each tab lists a different kind of edge meta-type (currently Data, Species or Predicted class). Each meta-type is divided into subcategories, e.g. individual species that can be turned on or off independently. The subcategories contribute to a total confidence score for the interaction of two distinct nodes. Some edges, e.g. paralogs and orthologs, do not have a confidence score because they do not represent an interaction and are therefore listed below the tab box.

The bottom panel contains a zoom slider, an edge confidence slider and a text search box. The latter two elements were adapted from *prefuse* (Heer *et al.*, 2005).

### 2.2 Node grouping

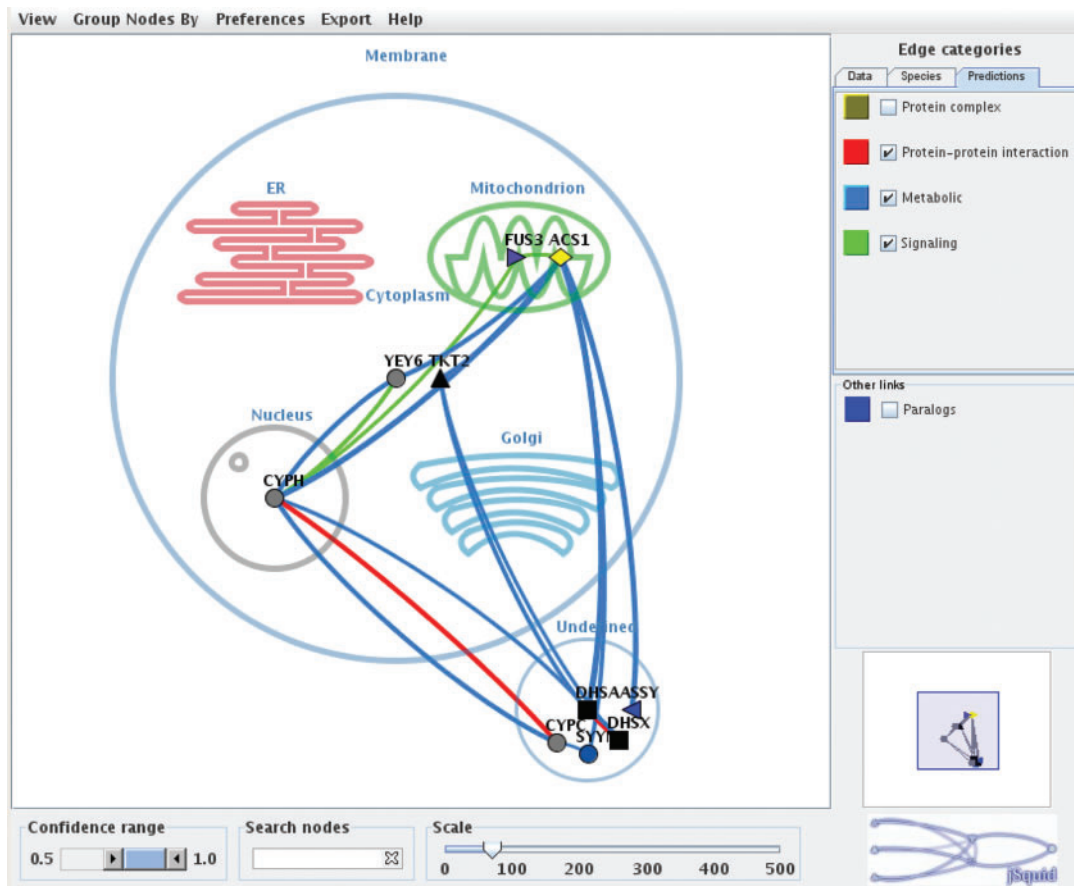
This feature separates jSquid from most of the other graph viewers currently available. It is a handy function, which clusters nodes (proteins in the case of FunCoup) into groups. These groups can be added through the XML data input, which allows an unlimited amount of different groups.

In the FunCoup context it is possible to group nodes by membership in specific KEGG (Kanehisa *et al.*, 2008) pathways, by subcellular localization, or by connectivity and confidence. The data for KEGG pathway membership and localization grouping is located in the XML file. Grouping by connectivity is done by a simple greedy algorithm, similar to *n-club* (Mokken, 1979) but incorporating both connectivity and confidence. It searches for the edge with the highest confidence and forms a group with the two nodes joined by this edge. Then it tries to find other nodes that are connected to this group, and add them if their average connectivity and confidence lie above a user-specified threshold. If the algorithm cannot find any new nodes to add, it creates a new group from the highest confidence edge in the graph between nodes that are not in a group, and add more nodes as described above. This continues until no more new groups can be created.

Groups are displayed as circles with an optional headline, with the corresponding nodes inside. In case of subcellular

\*To whom correspondence should be addressed.

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.



**Fig. 1.** The jSquid applet displaying a subnetwork from FunCoup using *Saccharomyces cerevisiae* ACS1 as query. Nodes that are not assigned to a compartment are clustered in the group 'undefined' outside the cell.

localization grouping, a cartoon cell is drawn and the nodes are placed into compartments according to their GO (Harris *et al.*, 2004) annotation (Fig. 1).

### 2.3 Collapsing

jSquid permits collapsing of node sets, in order to get an overview of dense graphs. To do this, the user selects a node set with the mouse and collapses or expands them via the right-click menu.

### 2.4 Data

For robustness and expandability of the input data, jSquid reads XML defined by an XSD schema (jSquidXML). This is available on the jSquid homepage, together with a tutorial on how to create a valid input file. The applet also reads XGMML (Punin and Krishnamoorthy, 2001) input files. However, some data (i.e. data for edge meta-types) cannot be described in this format, and therefore not all of jSquid's features can be used.

Although jSquid offers many graph manipulation features, it cannot fully replace sophisticated tools like Cytoscape. Therefore, jSquid can export graphs to the XGMML format to make the data readable for other programs. Visible nodes along with all edges of the currently active meta-type are exported. Furthermore the graph can also be exported in

jSquidXML format to store the current node positions, and as a picture in GIF or EPS format.

*Conflict of Interest:* none declared.

## REFERENCES

- Breitkreutz, B.J. *et al.* (2003) Osprey: a network visualization system. *Genome Biol.*, **4**, 22.
- Harris, M.A. *et al.* (2004) The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.*, **32**, D258–D261.
- Heer, J. *et al.* (2005) prefuse: a toolkit for interactive information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Portland, Oregon, USA, pp. 421–430.
- Hooper, S.D. and Bork, P. (2005) Medusa: a simple tool for interaction graph analysis. *Bioinformatics*, **21**, 4432–4433.
- Iragne, F. *et al.* (2005) ProViz: protein interaction visualization and exploration. *Bioinformatics*, **21**, 272–274.
- Kanehisa, M. *et al.* (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res.*, **36**, D480–D484.
- Mokken, R.J. (1979) Cliques, clubs and clans. *Qual. Quant.*, **13**, 161–173.
- Pinney, J.W. *et al.* (2005) metaSHARK: software for automated metabolic network prediction from DNA sequence and its application to the genomes of *Plasmodium falciparum* and *Eimeria tenella*. *Nucleic Acids Res.*, **33**, 1399–1409.
- Punin, J. and Krishnamoorthy, M. (2001) XGMML (eXtensible Graph Markup and Modeling Language). Available at <http://www.cs.rpi.edu/~puninj/XGMML/draft-xgmml.html> (last accessed date May 14, 2008).
- Shannon, P. *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2540.