OrthoNets: simultaneous visual analysis of orthologs and their interaction neighborhoods across different organisms

Yanqi Hao[1], Anna Merkoulovitch[1], James Vlasblom[1,2,*], Shuye Pu[1], Andrei L. Turinsky[1], Denitza Roudeva[3,4], Brian Turner[1], Jack Greenblatt[3,4] and Shoshana J. Wodak[1,2,3]

[1]Molecular Structure & Function program, Hospital for Sick Children, 555 University Ave, Toronto, ON, Canada M5G 1X8, [2]Department of Biochemistry, [3]Department of Molecular Genetics and [4]Banting and Best Department of Medical Research, Donnelly Centre, University of Toronto, 1 King's College Circle, Toronto ON, Canada M5S 1A8

Associate Editor: Trey Ideker

ABSTRACT

Motivation: Protein interaction networks contain a wealth of biological information, but their large size often hinders cross-organism comparisons. We present OrthoNets, a Cytoscape plugin that displays protein–protein interaction (PPI) networks from two organisms simultaneously, highlighting orthology relationships and aggregating several types of biomedical annotations. OrthoNets also allows PPI networks derived from experiments to be overlaid on networks extracted from public databases, supporting the identification and verification of new interactors. Any newly identified PPIs can be validated by checking whether their orthologs interact in another organism.

Availability: OrthoNets is freely available at http://wodaklab.org/orthonets/.

Contact: jim.vlasblom@utoronto.ca

1 INTRODUCTION

Recent technological advances have contributed to a rapid growth in the amount of data on gene and protein interactions as well as on protein and gene function (Costanzo *et al.*, 2010; Gavin *et al.*, 2006; Krogan *et al.*, 2006; Stelzl *et al.*, 2005). Although organisms such as yeast are comparatively well characterized, others—including human—have data for only a small fraction of their predicted genomes or proteomes. In this case, orthology relationships can be used to transfer information on function and interactions from one organism to another, thus assisting the prediction and validation of new interactors (Yu *et al.*, 2004).

The OrthoNets plug-in for Cytoscape (Shannon *et al.*, 2003) enables simultaneous visual analysis of protein–protein interaction (PPI) networks in multiple organisms. By default, the plug-in provides direct access to updated networks from five model organisms—yeast, mouse, human, fly and worm—consolidated from 10 major PPI databases (Razick *et al.*, 2008) and available from the iRefWeb resource (Turner *et al.*, 2010) resource (http://wodaklab.org/iRefWeb). Interactions are annotated with supporting information, such as interaction type, experimental detection methods and supporting publications. Interacting proteins in OrthoNets are further annotated with domain architectures

*To whom correspondence should be addressed.

(Bateman *et al.*, 2004), Gene Ontology terms and (if applicable) disease information from the Online Mendelian Inheritance in Man (OMIM) database (McKusick *et al.*, 2006).

2 SIMULTANEOUS ANALYSIS OF INTERACTION NETWORKS IN TWO ORGANISMS

Once the user specifies genes or proteins of interest, OrthoNets displays their full interaction neighborhoods as well as the interaction neighborhoods of their orthologs (Fig. 1a). The organism views are synchronized, so that selecting a particular node in one network will automatically select the corresponding orthologs in the other network. Orthologs can also be identified by matching node colors (Fig. 1b). Orthology relationships are pre-computed using the InParanoid method (O'Brien *et al.*, 2005).

By default, proteins involved in chromatin modification (Kouzarides, 2007) are highlighted with a white border, which reflects the focus of our research team, but the user may accentuate any preferred group of proteins in this way. Pfam domain architectures as well as various protein and interaction annotations may be visualized (Fig. 1b–d). Of particular interest are disease annotations from the OMIM database (McKusick *et al.*, 2006), which can be mapped onto the visible network as customized node shapes. Additional graphing tools are also provided to help determine the most common diseases in the neighborhood of a protein. This last feature may be used to infer a disease association for the corresponding gene. OrthoNets also enables two PPI networks to be overlaid (Fig. 1a, left panel), with edge and node colors indicating PPIs that differ between or are shared by the two networks. This allows the user, for example, to compare newly detected interactions with those already in iRefWeb, to easily identify new PPIs. Furthermore, the plug-in highlights novel PPIs whose orthologs interact in the other selected organism (conserved PPIs). Conserved PPIs may be prioritized for follow-up experimental validation.

3 IMPLEMENTATION AND CUSTOMIZABILITY

OrthoNets data are stored in standard Cytoscape node or edge attribute files, and tab-delimited text files. Files deposited into the appropriate directory are automatically loaded by the plugin, enabling the user to add their own set of gene, protein or interaction annotations. In particular, the user may upload custom PPI networks
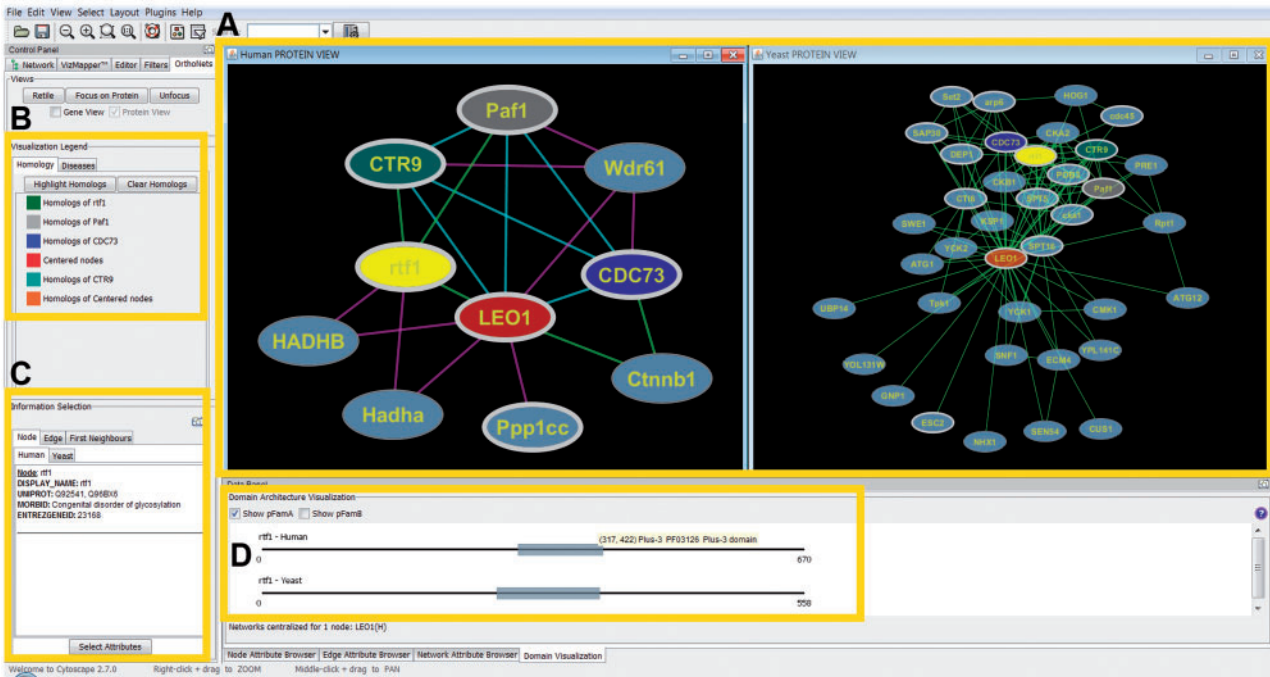
Fig. 1. (**A**) The network views (left: human, right: yeast): the user centers the human interaction network on the LEO1 component of the Paf1 complex, associated with RNA Polymerase II (left panel). OrthoNets simultaneously centers the yeast network (right panel) on the yeast LEO1 ortholog. Affinity purification data (Mak *et al.*, 2010) is overlaid on the human network: green edges denote those present in iRefWeb, purple edges indicate those found by experiment, but not in iRefWeb and cyan are those found in both. Although the Wdr61 (hSki8) interactions are not found in iRefWeb, Wdr61 has been reported as a subunit of the Paf1 complex (Zhu *et al.*, 2005). (**B**) Homology legend indicating the color coding of nodes displayed in (A). (**C**) Displays information available for the selected node or edge (in this case for human RTF1). (**D**) Pfam domain architecture for the selected node (here, RTF1), if available.

from iRefWeb filtered on the basis of various types of supporting PPI evidence, such as experimental method, number of publications and others. OrthoNets was developed in Java and is available at http://wodaklab.org/orthonets/.

REFERENCES

Bateman,A. *et al.* (2004) The Pfam protein families database. *Nucleic Acids Res.*, **32**, D138–D141.

Costanzo,M. *et al.* (2010) The genetic landscape of a cell. *Science*, **327**, 425–431.

Gavin,A.C. *et al.* (2006) Proteome survey reveals modularity of the yeast cell machinery. *Nature*, **440**, 631–636.

Kouzarides,T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.

Krogan,N.J. *et al.* (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*, **440**, 637–643.

Mak,A.B. *et al.* (2010) A lentiviral functional proteomics approach identifies chromatin remodeling complexes important for the induction of pluripotency. *MCP*, **9**, 811–823.

McKusick,V.A. *et al* (2006) *Online Mendelian Inheritance in Man, OMIM*. McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University, Baltimore, MD and National Center for Biotechnology Information, National Library of Medicine, Bethesda, MD.

O'Brien,K.P. *et al.* (2005) Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Res.*, **33**, D476–D480.

Razick,S. *et al.* (2008) iRefIndex: a consolidated protein interaction database with provenance. *BMC Bioinformatics*, **9**, 405.

Shannon,P. *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.

Stelzl,U. *et al.* (2005) A human protein-protein interaction network: a resource for annotating the proteome. *Cell*, **122**, 957–968.

Turner,B. *et al* (2010) iRefWeb: interactive analysis of consolidated protein interaction data and their supporting evidence. *Database*, **2010**, baq023.

Yu,H. *et al.* (2004) Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. *Genome Res.*, **14**, 1107–1118.

Zhu,B. *et al.* (2005) The human PAF complex coordinates transcription with events downstream of RNA synthesis. *Genes Dev.*, **19,** 1668–1673.