PhyloPro: a web-based tool for the generation and visualization of phylogenetic profiles across Eukarya

Xuejian Xiong1, Hongyan Song1, Tuan On1,2, Lucas Lochovsky1, Nicholas J. Provar3,4 and John Parkinson1,2,4,5,*

1Program in Molecular Structure and Function, Hospital for Sick Children, 2Department of Molecular Genetics, 3Department of Cell and Systems Biology, 4Centre for the Analysis of Genome Evolution and Function and 5Department of Biochemistry, University of Toronto, Toronto, Canada

Received on November 3, 2010; revised on December 21, 2010; accepted on January 6, 2011

In its simplest form, a phylogenetic profile is simply a pattern denoting the presence or absence of a homolog across a set of species. Previous studies have shown that genes sharing similar functions often have similar phylogenetic profiles, a phenomenon that has been exploited in various gene function prediction algorithms (Date and Marcotte, 2005; von Mering et al., 2007). Due to the lack of eukaryotic genomes, analyses involving phylogenetic profiles have typically been restricted to bacteria. However, the generation of increasing numbers of eukaryotic genome sequences, phylogenetic profiles of eukaryotic genes are becoming increasingly informative. Here, we introduce a new web-tool Phylopro (http://compsysbio.org/phylopro/), which uses the 120 available eukaryotic genome sequences to visualize the evolutionary trajectories of user-defined subsets of model organism genes. Applied to pathways or complexes, PhyloPro allows the user to rapidly identify core conserved elements of biological processes together with those that may represent lineage-specific innovations. PhyloPro thus provides a valuable resource for the evolutionary and comparative studies of biological systems.

Contact: jaarkin@sickkids.ca

© The Author 2011. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com
or tab-delimited text file. Users may also download the list of all orthologs shown in the heatmap. Figure 1 shows a typical application of PhyloPro, visualizing the phylogenetic profiles of 86 yeast genes associated with the cell cycle. From the heatmap, four groups of genes with specific patterns of orthology relationship are readily delineated: (A) highly conserved across all 120 Eukarya; (B) fungal-specific innovations; (C) duplications in vertebrates; and (D) duplications in yeast. PhyloPro has successfully been applied to explore the evolutionary landscape of the chromatin modification machinery (On et al., 2010), and the conservation of substrates of the Ned4 family of ubiquitin ligases (Persaud et al., 2009). There are several caveats associated with orthology detection (Ruano-Rubio et al., 2009). For example, in the absence of detailed phylogenetic analyses, domain gains, losses and shuffling events can significantly complicate orthology assignments. Hence, future development plans for PhyloPro include incorporating alternative orthology predictions from other graph or tree-based methods such as OrthoMCL (Li et al., 2003) and TreeFam (Ruan et al., 2008). Furthermore, the list of target species will be expanded as new genomes become available. Finally, we plan to extend the range of query species beyond the six model organisms to permit unbiased sequence-based searches. In summary, PhyloPro provides an intuitive web-based resource for performing rapid evolutionary and comparative studies of biological systems.

ACKNOWLEDGEMENTS
The authors would like to thank Dr James Wasmuth for advice during the creation of PhyloPro. Computing resources were provided by the Center for Computational Biology, Hospital for Sick Children and the SciNet HPC Consortium.

Funding: Canadian Institute for Health Research (CIHR - MOP82940); New Investigators award from CIHR (to J.P.); Early Researchers Award from Ontario Ministry for Research and Innovation (to J.P.).

Conflict of Interest: none declared.

REFERENCES