BIOINFORMATICS APPLICATIONS NOTE

Vol. 28 no. 6 2012, pages 891–892
doi:10.1093/bioinformatics/btr719

Systems biology

Advance Access publication December 30, 2011

CytoSEED: a Cytoscape plugin for viewing, manipulating and analyzing metabolic models created by the Model SEED

Matthew DeJongh*, Benjamin Bockstege, Paul Frybarger, Nicholas Hazekamp, Joshua Kammeraad, and Travis McGeehan

Computer Science Department, Hope College, Holland, Michigan, 49423, USA

Associate Editor: Trey Ideker

ABSTRACT

Summary: CytoSEED is a Cytoscape plugin for viewing, manipulating and analyzing metabolic models created using the Model SEED. The CytoSEED plugin enables users of the Model SEED to create informative visualizations of the reaction networks generated for their organisms of interest. These visualizations are useful for understanding organism-specific biochemistry and for highlighting the results of flux variability analysis experiments.

Availability and Implementation: Freely available for download on the web at http://sourceforge.net/projects/cytoseed/. Implemented in Java SE 6 and supported on all platforms that support Cytoscape.

Contact: dejongh@hope.edu

Supplementary information: Installation instructions, a tutorial, and full-size figures are available at http://www.cs.hope.edu/cytoseed/.

Received on October 31, 2011; revised on December 7, 2011; accepted on December 25, 2011

1 INTRODUCTION

The Model SEED (http://www.theseed.org/models) is a web-based resource for the automatic generation of metabolic models from prokaryotic genome sequences (Henry *et al.*, 2010). These models are automatically gap-filled by adding reactions that enable the model to produce all specified components of the organism's biomass. Although the Model SEED provides extensive capabilities for viewing and running analyses on these metabolic models, the graphical displays of the biochemical reaction networks are limited to static KEGG pathway maps (Kanehisa *et al.*, 2006). CytoSEED provides Model SEED users with a complementary means of visualizing their metabolic models using Cytoscape (Shannon *et al.*, 2003). Cytoscape is an open-source software platform for dynamic, graphical visualization, manipulation and analysis of networks. The CytoSEED plugin extends Cytoscape to enable users to load a metabolic model from the Model SEED, sort the reaction network into metabolic pathways, and modify the pathway visualizations to create the representations they find most informative. Through this process, users are able to identify and represent connections in the biochemical reaction network that are not readily apparent in the KEGG pathway maps. CytoSEED also enables users to display the results of flux variability analysis experiments for visualization in the context of a metabolic model, and to compare multiple models.

*To whom correspondence should be addressed.

2 SYSTEM OVERVIEW

CytoSEED provides a menu with options to load a model into a Cytoscape session, save a model session, and reopen a model session. To load a model, the user must enter the Model SEED identifier for the model; a Model SEED username and password are also required for loading private models. CytoSEED downloads the model data from the Model SEED web-server and stores it in a specified folder on the user's machine to speed up loading during subsequent sessions. When the download is complete, CytoSEED processes the model data, sorting the reactions into networks based on their association with KEGG pathways. A network is created for every KEGG pathway associated with the model reactions, so some reactions may be displayed in multiple networks; for example, the pyruvate kinase reaction is present in both the *Glycolysis/Gluconeogenesis* and the *Carbon Fixation in photosynthetic organisms* pathways, so models containing this reaction will result in the creation of separate networks for both pathways. Every network corresponding to a KEGG pathway is formatted to match the layout of the static KEGG pathway map using the *kgmlreader* plugin (freely available at http://code.google.com/p/kgmlreader/), if installed. In addition, CytoSEED creates several other networks to group common model reactions not generally associated with KEGG pathways, such as transport reactions. Finally, CytoSEED creates a network with the same name as the model identifier to group any leftover reactions (*e.g.*, the *Seed243273.1 Network* in Figure 1).

After the networks have been created, CytoSEED uses flux variability analysis results for any biomass production from the Model SEED to color the reactions based on the following categories: (1) *Essential* reactions, which must carry a non-zero flux to produce biomass, are colored gold; (2) *Active* reactions, which can carry a non-zero flux but are not essential, are colored blue; (3) *Inactive* reactions, which cannot carry a flux, are colored grey; and (4) *Gap-filled* reactions, which are automatically added by the Model SEED to complete the reaction network and are thus essential for any biomass production, are colored purple. Compounds are also colored based on several categories; see the legend in the lower right-hand corner of Figure 1 for details.

The Maps panel on the right-hand side of the Cytoscape desktop displays each network name and the number of total and unique reactions in the network. Using context menus associated with the Maps panel and each network, the user can create new networks, delete existing networks, and move or copy reactions between networks. Figure 1 displays the result of refining the *Mycoplasma genitalium* G-37 model visualization by deleting

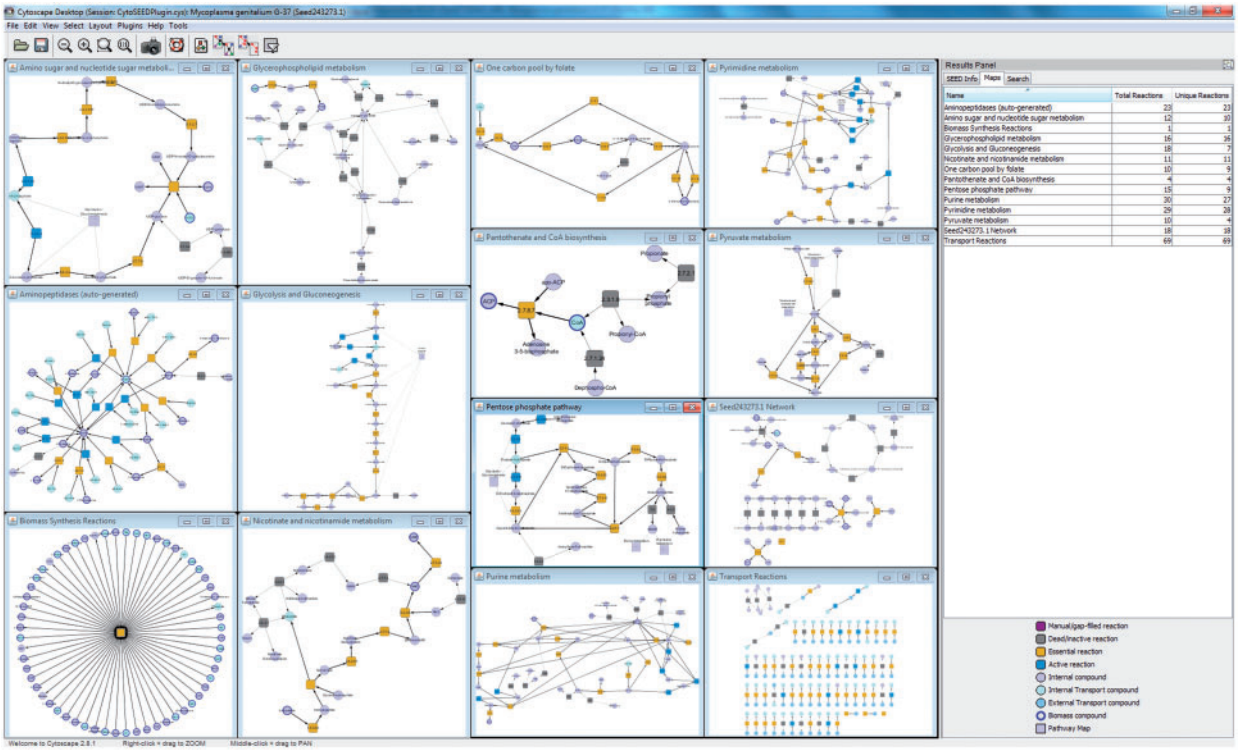

Fig. 1. CytoSEED interface displaying a refined visualization of the metabolic model for *Mycoplasma genitalium* G-37 (Model SEED identifier *Seed243273.1*). Reactions are colored based on flux variability analysis results for the model on Complete media.

extraneous networks such as the *Carbon fixation* pathway, and moving reactions to fill gaps in other pathways. A tutorial that steps the user through this process in more detail is available in the Supplementary information.

CytoSEED provides a dynamic visualization of a metabolic model that is useful for understanding the organism-specific biochemistry represented in the model. For example, *M. genitalium* is missing a gene for transaldolase, which converts the compound erythrose-4-phosphate into sedoheptulose-7-phosphate. This results in a gap in KEGG's static *Pentose phosphate* pathway map. However, two other reactions present in the model can bridge this gap, using sedoheptulose-1,7-bisphosphate as an intermediary compound. CytoSEED enables the user to move these reactions into the network representing the *Pentose phosphate* pathway, resulting in a visualization of the pathway that demonstrates it is complete and functioning in the model. CytoSEED also provides the capability of switching among the various flux variability analysis results available for a given model from the Model SEED web-server. For example, the *M. genitalium* model has results for both Complete (i.e., all transportable compounds are in the media) and SP4 media. Using the CytoSEED menu, the user can switch from the Complete media results to the SP4 media results, and the coloring of reactions is immediately updated, showing, e.g., that the glucose-phosphotransferase transport reaction switches from active to essential in the SP4 media.

CytoSEED also includes the capability of comparing multiple metabolic models in one Cytoscape session. See the tutorial for details.

ACKNOWLEDGEMENTS

The authors thank Aaron Best, Chris Henry, Scott Devoid and the members of the SEED and RAST teams for their support.

Funding: This work was supported by the National Science Foundation [grant numbers 0745100, 0850546 to MD]. BB, NH and JK were supported by the National Science Foundation REU program [grant number 0851293 to Hope College]. TM was supported by Exxon Mobil and Hope College through the Research Experiences Across Cultures at Hope program.

Conflict of Interest: none declared.

REFERENCES

Henry,C. *et al.* (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotech.*, **28**, 977–982.

Kanehisa,M. *et al.* (2006) From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res.*, **34**, D354–D357.

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