Bioinformatics, 36(8), 2020, 2616–2617
doi: 10.1093/bioinformatics/btz927
Advance Access Publication Date: 18 December 2019
Applications Note

Systems biology

SAMMI: a semi-automated tool for the visualization of metabolic networks

Andre Schultz and Rehan Akbani*

Department of Bioinformatics and Computational Biology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

* To whom correspondence should be addressed.

Associate Editor: Lenore Cowen

Received on August 7, 2019; revised on November 14, 2019; editorial decision on December 7, 2019; accepted on December 16, 2019

Abstract

Summary: Here we present a browser-based Semi-Automated Metabolic Map Illustrator (SAMMI) for the visualization of metabolic networks. While automated features allow for easy network partitioning, navigation, and node positioning, SAMMI also offers a wide array of manual map editing features. This combination allows for fast, context-specific visualization of metabolic networks as well as the development of standardized, large-scale, visually appealing maps. The implementation of SAMMI with popular constraint-based modeling toolboxes also allows for effortless visualization of simulation results of genome-scale metabolic models.

Availability and implementation: SAMMI has been implemented as a standalone web-based tool and as plug-ins for the COBRA and COBRApy toolboxes. SAMMI and its COBRA plugins are available under the GPL 3.0 license and are available along with documentation, tutorials, and source code at www.SammiTool.com.

Contact: rakbani@mdanderson.org

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

1 Introduction

The field of genome-scale metabolic models (GEMs) has rapidly expanded in recent years (O'Brien *et al.*, 2015), leading to the creation of a number of databases (Karp and Caspi, 2011) and algorithms (Lewis *et al.*, 2012) for the analysis of metabolic networks. In particular, the COBRA (Heirendt *et al.*, 2019) and COBRApy (Ebrahim *et al.*, 2013) toolboxes have become extremely popular tools for GEM simulation.

One important aspect of this field is the visualization of such networks and their associated data. Manually curated maps (Gawron *et al.*, 2016; King *et al.*, 2015) are often used for this purpose, but these are time-consuming to generate, are GEM specific, and can lead to discontinuity in data visualization due to node duplication. Automatically generated static maps (Jensen and Papin, 2014; Kelley *et al.*, 2017) account for some of these issues, but can still yield convoluted maps that can be difficult to interpret. In addition, many of the previously developed tools rely on standalone or independent software (Chazalviel *et al.*, 2018), making their integration with powerful, popular constraint-based modeling tools difficult. To account for these issues we present a Semi-Automated Metabolic Map Illustrator (SAMMI). SAMMI provides a wide array of automated map generation features as well as various manual curation and navigation tools. This combination allows powerful context-specific analysis as well as the generation of standardized, benchmark metabolic maps.

2 Application features

Using the standalone SAMMI interface, metabolic networks can be uploaded using Systems Biology Markup Language (SBML) (Hucka *et al.*, 2018) and Biological Pathway Exchange (BioPax) (Demir *et al.*, 2010) annotated models, which can be obtained from a number of different databases, or directly from the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa *et al.*, 2017) annotated networks. Alternatively, users can render SAMMI maps directly from the MATLAB and Python command lines using the COBRA and COBRApy SAMMI plug-ins respectively. SAMMI offers the option of partitioning models into multiple subgraphs using model annotation (e.g. reaction subsystems or metabolite compartments) or user defined groups of reactions. This parsing allows users not only to plot and navigate large models but to also render the visualization of desired reactions only, avoiding overcrowding by omitting undesirable reactions and ensuring connectivity since no nodes are initially duplicated. SAMMI networks are initially drawn as animated force-directed graphs, which allows for constant automated node positioning. Parameters of the force-directing layout, such as edge length and node repulsion, can be modified to tailor this process to the user's liking.

Once a model is loaded a wide array of manual curation functionalities are available for exploring the network and drawing visually appealing maps. These include, but are not limited to, fixing node position, arranging nodes in different shapes, duplicating

nodes, labeling nodes as secondary, curving edges, adding and deleting reactions and metabolites, defining network paths and components, re-defining and re-arranging labels, hiding reaction nodes, and annotating shapes and texts. In particular, nodes can be temporarily removed and later re-introduced into the network. This feature allows users to temporarily remove secondary metabolites for a less crowded visualization or while primary nodes are arranged.

Multiple datasets can be mapped onto SAMMI networks simultaneously as node and edge color, node size and edge thickness. Color and size scales can be defined for each independent dataset or globally for all datasets uploaded. Color scales can be customized by editing, adding or removing color breaks and re-defining colors. Color and size of nodes and edges with no associated data can also be defined separately, mapping data only to specified groups of reactions and metabolites.

Although the partitioning of large networks into pathways makes their visualization more appealing, this subdivision can hinder the visualization of network connectivity across pathways, an important feature in metabolic networks where many metabolites are present in a large number of pathways. To alleviate this issue, SAMMI offers advanced map navigation options across subgraphs. For instance, when searching for nodes in one graph, SAMMI also shows all other subgraphs that contain nodes matching that search, and indicates the number of nodes matching the search in each subgraph. Alternatively, users can choose to highlight a given metabolite and visualize all reactions in which that metabolite participates. Both options allow users to easily navigate to other significant subgraphs.

SAMMI maps can be easily shared as a SAMMI-specific JSON file. These files retain all nodes' state and position, edge curves and associated map data, allowing users to save the network in their current format for standardization. SAMMI maps can also be exported in formats compatible with Escher (King *et al.*, 2015) and MetExploreViz (Chazalviel *et al.*, 2018), as well as PNG and PDF figures. Supplementary Material gives a detailed comparison of several features between SAMMI and other popular graph visualization tools.

The SAMMI source code has been publicly shared on GitHub, where users are encouraged to submit additional feature suggestions and bug reports. A detailed overview of all SAMMI features, video tutorials, and examples of how to use the COBRA plugins directly from the command line are available in the SAMMI documentation, which can be accessed directly from the tool's interface.

3 Conclusion

Although automated node positioning is highly advantageous in order to draw large metabolic networks, a large degree of manual curation is also often desirable to provide context and finishing touches in visually appealing network representations. Furthermore, while static maps are often useful, their connectivity is often not suited for context-specific visualizations. SAMMI combines a series of automated and manual features that allow for both the development of highly curated, visually appealing maps, as well as the fast visualization of context-specific subnetworks and associated data. In addition, SAMMI also provides advanced navigation tools, easy file sharing, and cross-platform integration capabilities. We believe SAMMI will be a powerful tool for the visualization and analysis of metabolic networks.

Acknowledgements

We would like to acknowledge Chris Wakefield, Mary Rohrdanz and members of the MD Anderson Metabolomics Facility for helpful discussions and feedback on this project.

Funding

This work was supported by the grants National Institutes of Health/National Cancer Institute [U24CA210950, U01CA235510, P30 CA016672 and Department of Defense/Congressionally Directed Medical Research Programs W81XWH-16-1-0237].

Conflict of Interest: R.A. provides paid bioinformatics consultation services to the University of Houston.

References

Chazalviel,M. *et al.* (2018) MetExploreviz: web component for interactive metabolic network visualization. *Bioinformatics*, **34**, 312–313.

Demir,E. *et al.* (2010) The BioPAX community standard for pathway data sharing. *Nat. Biotechnol.*, **28**, 935–942.

Ebrahim,A. *et al.* (2013) Cobrapy: constraints-based reconstruction and analysis for python. *BMC Syst. Biol.*, **7**, 74.

Gawron,P. *et al.* (2016) MINERVA-a platform for visualization and curation of molecular interaction networks. *NPJ Syst. Biol. Appl.*, **2**, 16020.

Heirendt,L. *et al.* (2019) Creation and analysis of biochemical constraint-based models using the COBRA toolbox v.3.0. *Nat. Protoc.*, **14**, 639–702.

Hucka,M. *et al.* (2018) The systems biology markup language (SBML): language specification for level 3 version 2 core. *J. Integr. Bioinform.*, **15**,

Jensen,P.A. and Papin,J.A. (2014) MetDraw: automated visualization of genome-scale metabolic network reconstructions and high-throughput data. *Bioinformatics*, **30**, 1327–1328.

Kanehisa,M. *et al.* (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.*, **45**, D353–D361.

Karp,P.D. and Caspi,R. (2011) A survey of metabolic databases emphasizing the MetaCyc family. *Arch. Toxicol.*, **85**, 1015–1033.

Kelley,J.J. *et al.* (2017) MOST-visualization: software for producing automated textbook-style maps of genome-scale metabolic networks. *Bioinformatics*, **33**, 2596–2597.

King,Z.A. *et al.* (2015) Escher: a web application for building, sharing, and embedding data-rich visualizations of biological pathways. *PLoS Comput. Biol.*, **11**, e1004321.

Lewis,N.E. *et al.* (2012) Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nat. Rev. Microbiol.*, **10**, 291–305.

O'Brien,E.J. *et al.* (2015) Using genome-scale models to predict biological capabilities. *Cell*, **161**, 971–987.