

Systems biology

Newt: a comprehensive web-based tool for viewing, constructing and analyzing biological maps

Hasan Balci¹, Metin Can Siper^{1,2}, Nasim Saleh¹, Ilkin Safarli^{1,3}, Ludovic Roy⁴, Merve Kilicarslan¹, Rumeysa Ozaydin¹, Alexander Mazein ^{4,5}, Charles Auffray⁴, Özgün Babur^{2,6}, Emek Demir² and Ugur Dogrusoz ^{1,*}

¹i-Vis Research Lab, Computer Engineering Department, Bilkent University, Ankara 06800, Turkey, ²Molecular & Medical Genetics Department, School of Medicine, Oregon Health & Science University, Portland, OR 97239, USA, ³Visualization Design Lab, School of Computing, University of Utah, Salt Lake City, UT 84112, USA, ⁴European Institute for Systems Biology and Medicine, CIRI UMR5308, CNRS-ENS-UCBL-INSERM, Université de Lyon, 69007 Lyon, France, ⁵Luxembourg Centre for Systems Biomedicine, University of Luxembourg, L-4367 Belvaux, Luxembourg and ⁶Computer Science Department, University of Massachusetts Boston, Boston, MA 02125, USA

*To whom correspondence should be addressed.

Associate Editor: Wren Jonathan

Received on June 3, 2020; revised on August 25, 2020; editorial decision on September 16, 2020; accepted on September 18, 2020

Abstract

Motivation: Visualization of cellular processes and pathways is a key recurring requirement for effective biological data analysis. There is a considerable need for sophisticated web-based pathway viewers and editors operating with widely accepted standard formats, using the latest visualization techniques and libraries.

Results: We developed a web-based tool named Newt for viewing, constructing and analyzing biological maps in standard formats such as SBGN, SBML and SIF.

Availability and implementation: Newt's source code is publicly available on GitHub and freely distributed under the GNU LGPL. Ample documentation on Newt can be found on <http://newteditor.org> and on YouTube.

Contact: newteditor@gmail.com

1 Introduction

Lately, our knowledge about biological processes has been expanding at a rapid rate. With widely accepted standard formats such as BioPAX (Demir *et al.*, 2010), SBGN (Le Novère *et al.*, 2009) and SBML (Hucka *et al.*, 2003), pathway databases such as Pathway Commons (Rodchenkov *et al.*, 2020), Reactome (Croft *et al.*, 2011) and WikiPathways (Slenter *et al.*, 2018) are able to make this vast amount of biological process data available to the scientific community suitable for computation.

Visual communication via maps or diagrams is fundamental to the process of exploring concepts and disseminating information (Agrawal *et al.*, 2011). Biology is no exception. Although some standalone software applications have been developed to visually construct and analyze biological process data such as CellDesigner (Funahashi *et al.*, 2003), SBGN-ED (Czauderna *et al.*, 2010) and ChiBE (Babur *et al.*, 2014), currently there are no web applications offering a complete, end-to-end solution for effective analysis and editing of such data. With this motivation, Newt builds on SBGNViz (Sari *et al.*, 2015), an SBGN Process Description (PD)

viewer, constructing a full-blown pathway editor as well as adding many useful viewer-level features:

- In addition to viewing existing maps, Newt has advanced diagramming capabilities to build maps from scratch or edit existing ones.
- On top of SBGN PD notation, Newt has support for SBGN Activity Flow (AF) notation and Simple Interaction Format (SIF) (Rodchenkov *et al.*, 2020). SBGN persistency support was upgraded (SBGN-ML version 0.3) to include style and annotation of map objects. Furthermore, bi-directional conversion facilities with popular formats such as SBML (Hoksza *et al.*, 2019) and CellDesigner (Balaure *et al.*, 2020) are available.
- With Newt, possibly multiple sets of experimental data may be simultaneously shown in the context of pathways.
- Newt supports miscellaneous other new features such as semantic validation and guided fix for SBGN PD maps and color schemes to quickly adorn maps.

- Newt was designed in a modular fashion with a component-based approach (Newt architecture). Hence, its components can be used to build other visual analysis tools or can be embedded in other web pages.

To the best of our knowledge, Newt is the only web-based tool for visual construction and analysis of biological pathways in widely accepted standard formats with a comprehensive set of features, using latest technologies and working on most modern browsers and touch-enabled devices without requiring any prior installation or set up. Other web-based tools such as Escher (King *et al.*, 2015) either use a more specialized, non-standard notation or lack a comprehensive feature set.

2 Materials and methods

2.1 Constructing pathways

Newt provides an easy to use Palette, from which users can drag and drop map objects onto the drawing canvas. A separate palette is provided for each supported map type (SBGN PD, SBGN AF and SIF), and hybrid maps are also allowed at the cost of giving up validation checks (e.g. a dissociation glyph can only be connected to one consumption glyph in an SBGN PD map) (Fig. 1).

A number of SBGN bricks (Junker *et al.*, 2012) are available under the Edit menu to quickly construct recurring biological network patterns such as complex associations. Created map objects can be inspected and edited via the Object tab. In addition, various map properties from map type, name and description to edge arrow sizes and label style are customizable through the Map tab. Furthermore, a number of color schemes, various coloring combinations and styles—solid, gradient or 3D—are available to quickly adorn maps.

Many advanced diagramming capabilities are available in Newt including pan/zoom operations, alignment tools, guidelines for aligning and uniformly distributing map objects, visual cues to properly resize nodes to content or manually as desired, and operations for interactive manipulation of edge bends.

Newt also allows one to query the Pathway Commons database using a number of different graph-based algorithms (Dogrusoz *et al.*, 2009a) to take a list of *genes of interest* and compute a *network of interest* consisting of these genes and their immediate set of relations as defined by the query type.

2.2 Analyzing pathways

A user will often like to upload and analyze an existing (or newly created) pathway. To aid such analysis, Newt provides some basic

inspection facilities (from object styling to annotations and links to external databases). In addition, one can search and highlight by label text or manually select and highlight parts of a map. Furthermore, the layout of the map can be manually or automatically adjusted by the provided operations (Dogrusoz *et al.*, 2009b).

One crucial set of operations available in Newt is for reducing the complexity of large maps through hide/show and expand/collapse operations, where the ability to nest maps is particularly useful (Dogrusoz *et al.*, 2018).

Newt also allows the user to check the semantic validity of their SBGN PD maps. Rules for semantic validity are defined by the SBGN standard and enforced by the libSBGN.js library. Furthermore, in cases where the map is not valid, Newt takes the user through each such error one by one using a wizard, focusing on map objects resulting in the error and possibly offering a fix for it whenever possible.

Finally, users may upload possibly multiple sets of data, such as experimental data or clustering data resulting from a computational method, to overlay and visually analyze such data in the context of their choice of a pathway, using a simple yet customizable data format.

2.3 Persistency and interoperability

Newt maps may be persisted on disk using a native SBGN-ML (van Iersel *et al.*, 2012) based format (.nwt file), which may be used to restore the map as is, including all properties of the map including style. The map may also be persisted as static (PNG or JPG) or vector (SVG) images. Furthermore, Newt maps may be exported to and imported from formats such as SBML, CellDesigner and SIF format.

Users may also produce uniformly styled maps or share their map and styling preferences such as label fit style and color scheme used through a preference file.

Another useful feature provided by Newt is the ability to launch Newt from a remote map file by URL (example). Remote launch can be invoked with a Pathway Commons URI as well.

3 Conclusion

Newt was developed to address the need for a customizable web-based component and an end-user application for constructing and visually analyzing biological pathways with state-of-the-art editing, diagramming, complexity management, and experimental data analysis facilities using standard notations. A user guide along with a rich set of tutorial videos, documentation on the tool architecture, an FAQ page and other miscellaneous documentation can be found on the tool website.

Acknowledgement

The authors acknowledge the support by Google Summer of Code to H.B., M.C.S. and I.S. during summers.

Funding

U.D. was supported by the Scientific and Technological Research Council of Turkey [118E131 and 5180088]. M.C.S., O.B. and E.D. were supported by Defense Advanced Research Projects Agency Communicating with Computers Program [W911NF-15-1-0544]. L.R., A.M. and C.A. were supported by the Innovative Medicines Initiative Joint Undertaking under grant agreement no. IMI 115446 (eTRIKS), resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (FP7/2007–2013) and EFPIA companies.

Conflict of Interest: none declared.

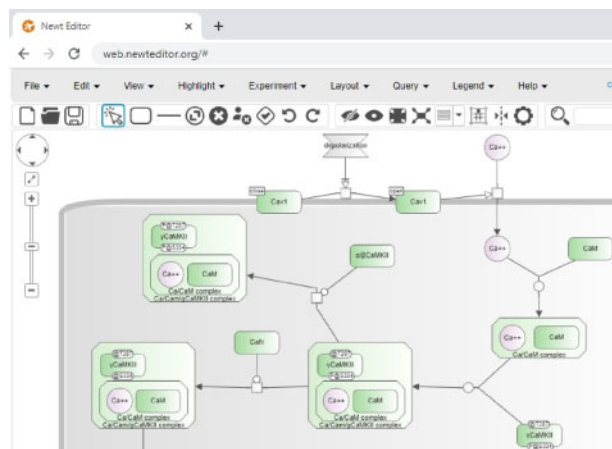


Fig. 1. A partial sample screenshot from Newt showing a pathway, where γ CaMKII shuttles $\text{Ca}^{2+}/\text{CaM}$ to the nucleus to trigger CREB phosphorylation and gene expression (Ma *et al.*, 2014)

References

- Agrawala, M. *et al.* (2011) Design principles for visual communication. *Commun. ACM*, **54**, 60–69.
- Babur, O. *et al.* (2014) Integrating biological pathways and genomic profiles with chibe 2. *BMC Genomics*, **15**, 642.
- Balaur, I. *et al.* (2020) cd2sbgnml: bidirectional conversion between CellDesigner and SBGN formats. *Bioinformatics*, **36**, 2620–2622.
- Croft, D. *et al.* (2011) Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.*, **39**, D691–D697.
- Czauderna, T. *et al.* (2010) Editing, validating and translating of SBGN maps. *Bioinformatics*, **26**, 2340–2341.
- Demir, E. *et al.* (2010) Biopax – a community standard for pathway data sharing. *Nat. Biotechnol.*, **28**, 935–942.
- Dogrusoz, U. *et al.* (2009a) Algorithms for effective querying of compound graph-based pathway databases. *BMC Bioinformatics*, **10**, 376.
- Dogrusoz, U. *et al.* (2009b) A layout algorithm for undirected compound graphs. *Inf. Sci.*, **179**, 980–994.
- Dogrusoz, U. *et al.* (2018) Efficient methods and readily customizable libraries for managing complexity of large networks. *PLoS One*, **13**, e0197238.
- Funahashi, A. *et al.* (2003) CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *Biosilico*, **1**, 159–162.
- Hoksza, D. *et al.* (2019) MINERVA API and plugins: opening molecular network analysis and visualization to the community. *Bioinformatics*, **35**, 4496–4498.
- Hucka, M. *et al.* (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, **19**, 524–531.
- Junker, A. *et al.* (2012) Wiring diagrams in biology: towards the standardized representation of biological information. *Trends Biotechnol.*, **30**, 555–557.
- 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.
- Le Novère, N. *et al.* (2009) The systems biology graphical notation. *Nat. Biotechnol.*, **27**, 735–741.
- Ma, H. *et al.* (2014) yCaMKII shuttles Ca²⁺/CaM to the nucleus to trigger CREB phosphorylation and gene expression. *Cell*, **159**, 281–294.
- Rodchenkov, I. *et al.* (2020) Pathway Commons 2019 Update: integration, analysis and exploration of pathway data. *Nucleic Acids Res.*, **48**, D489–D497.
- Sari, M. *et al.* (2015) SBGNViz: a tool for visualization and complexity management of SBGN process description maps. *PLoS One*, **10**, e0128985.
- Slenter, D.N. *et al.* (2018) WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res.*, **46**, D661–D667.
- van Iersel, M. *et al.* (2012) Software support for SBGN maps: SBGN-ML and LibSBGN. *Bioinformatics*, **28**, 2016–2021.