Reviewer's report

Title: BiNoM, a Cytoscape plugin for accessing and analyzing pathways using standard systems biology formats

Version: 1 **Date:** 19 November 2012

Reviewer number: 1

Reviewer's report:

The authors describe BiNoM, a Java-based tool that can be used as a stand-alone program as well as a plugin for Cytoscape to deal with systems biological networks given in BioPAX or SBML format. As already indicated in the article's title, the main usage of the tool is via Cytoscape. The support for BioPAX is limited to Level 3 and SBML functionality is limited to those networks created with the program CellDesigner (or to SBML networks with an identical annotation). BiNoM can import and export these formats, i.e., do some basic interconversion, display the network graphs and perform a large set of analysis for the graphs. The project has been developed since a longer period, the article refers to an earlier publication from 2008, in which the previous version of the program has been published. The program is well described, the article is well written, elaborated tutorials, web sites, and examples guide the user. The source code of the program is also available. However, several aspects should be improved before publication.

Major concerns

==========

In the abstract:

It sounds as if BiNoM supports multiple different input formats, of which two examples are listed ("such as SBML and BioPAX"), but there might be other supported formats not listed here, maybe due to space limitation. However, other formats, e.g., CellML, GraphML, SBGN-ML etc. are, however, not supported. Make more explicit that only these two formats are supported, not "such as".

When stating that "BiNoM can be used to import and analyze les created with the CellDesigner software" this raises the question: Is CellDesigner's layout representation converted to SBML Layout extension? Since the way how CellDesigner stores its layouts is not a community standard way, it would be nice to have also support for the (official) SBML layout extension (import/export). In the abstract, however, the reader doesn't know if this is the case.

Section Bacground:

Here, more details are provided about CellDesigner's layout encoding. However, it is not really "proprietary" because there is a more than 90 pages documentation about it, provided for download at CellDesigner's web site. I would rather say, the developers of CellDesigner use their own layout extension, in

contrast to the "official" SBML layout extension. One reason to use the community standard instead; better compatibility, better code base in libraries such as libSBML or JSBML. So Why only CellDesigner's layout specification instead of the official SBML layout extension?

The statement "Some tools are available for a speci fic conversion of one fi le format into another" sounds as if all kinds of systems biology formats would be transformed by the tools that are cited. Can you also cite a converter that transforms CellML, SBGN-ML or what ever to something else? At least be a bit more specific which file formats are meant here. Furthermore, it is not clear to me if these tools for interconversion between SBML and BioPAX are used within BiNoM to interconvert between these file formats. If there is a self-made mapping of the data structures for each input format to some internal data structure, describe how this mapping is performed in order to guarantee a lossless import and export. Is it possible to read SBML and write BioPAX or the other way arround? What would happen if the same file would be interconverted multiple times? Would the output be still the same as the original file?

If these libraries are not used, explain why. What is the internal data format in BiNoM?

The change log, i.e., the most important differences between BiNoM in reference [23] and the current publication, would be of great benefit: What has been improved since the previous publication? Whenever this paper is cited, I ask myself if the described feature was already part of the earlier publication. It is therefore really important to have some change log somehwere.

Section Path analysis algorithms:

Watch out! Finding shortest paths within metabolic networks without considering the atom balance can be misleading! You might end up in a short path, e.g., connecting two reactions via ATP as a common product/substrate, but the actual path must follow the flow of matter due to the fact that metabolic networks are actually hypergraphs, indicated by using a bipartite display with a second kind of node for reactions. Explain how following invalid short connections can be avoided by your analysis. Similarly, in Section "BiNoM Utilities" there are also statements that might lead to wrong short abbreviations within the pathway.

Section Pathway influence quanti cation algorithm:

Illustrate the PIQuant score with a small figure if appropriate.

Section Results and Discussion:

The passive sentence "The comprehensive map of the RB/E2F network was built using CellDesigner" raises the question if the authors created this map using CellDesigner or if it was downloaded somewhere. It remains unclear who the author of the work is.

Section Conclusions:

Which libraries are used to read/write BioPAX and SBML? Maybe paxtools, libSBML, or JSBML? Or did you use self-made parsers/data objects? If any

third-party library has been used, these should be mentioned and cited accordingly.

Is the qual extension supported for GINsim?

About merging models: How is this planned? See semanticSBML for those approaches - will MIRIAM terms be used to identify shared components?

Personally, I am not a great fan of having multiple promises at the end of some article. It might happen that for some reason these nice ideas cannot be finished, or even that somebody else will do it more quickly. I would suggest to limit the article to what has been done so far, maybe give some outlook, but without too many promises.

Furthermore, there is already a web-based editor for SBML models: http://code.google.com/p/biographer/.

A comparison to another Cytoscape plugin for dealing with SBML should be included into the manuscript: http://www.charite.de/sysbio/people/koenig/software/cysbml/ (as related work).

Supplementary Excel spreadsheet:

* Plot the data contained therein and include the plots into the spreadsheet.

Tool testing:

* It seems there is no support for compartments? I imported a network from CellDesigner, but the compartments were not drawn.

* I could not open the example Apoptosis OWL file because it seems to be a BioPAX file prior to L3.

* CySBML

* In SBGN there is no arrow head at pointing towards a reaction node.

* When clicking at a node in a network there is not much details displayed to the user; only the name of some element. In case of imported SBML files, it seems that the "name" attribute is taken from the species node, but when working with a BioPAX file, the "displayName" tag seems not to be used; the name displayed to the user is actually the internal identifier of the element, which is most of the cases some cryptic abbreviation. Cross-links to databases etc. are not displayed. It would be very nice to display MIRIAM annotations to the user in some form. For testing purposes, any annotated model from BioModels database can be opened in CySBML, this gives in many cases even the molecule's structure. It would be nice to have a similar way to access these crosslinks in BiNoM as well.

* When exporting an BioPAX file containing lots of MIRIAM annotations to SBML, all the annotation is lost. However, the notes tag contains cryptic information about the lines between nodes. However, the notes element is intended to contain human-readable HTML-like information to be displayed, e.g., in a web browser. If information gets lost as it is the case here, this should be indicated to the user. It would, of course, be better, to keep MIRIAM annotation while doing such a conversion.

* The article should make more explicit, which Level/Version combinations of SBML are understood by the tool, because it seems that is restricted to L2 only, which is fine, but the most recent version is SBML Level 3. Hence, it should be indicated clearly as being limited to earlier versions. For BioPAX it is made explicit that only L3 is supported.

* When trying to open an SBML file without CellDesigner's extension, nothing is displayed in Cytoscape.

Minor issues

==========

* The term "large scale" is used very frequently and incorrectly. The authors do not talk about some "large scale", but rather about "large-scale networks" or "large-scale molecular maps". Please use a dash in all these cases. Avoid using this expression so often.

* Abstract: "... a Cytoscape plugin which" -> comma before "which"

* Use upper case for "Level" when talking about BioPAX or SBML (multiple occurrences).

* Do not use upper cases for "Systems Biology" -> "systems biology"

* Always insert a comma after "e.g." or "i.e."

* Mark the brand name "Java" with a TM symbol at its first occurence (but also not more).

* Use correct closing English quotation symbols (in the Implementation section).

* Implementation section: why is "Module manager" and "Utilities" written in upper cases?

* "gene regulatory networks" -> "gene-regulatory networks"

* Use upper case spelling when referring to specific figures or tables, e.g., Figure 1 shows... or Table 3 contains... instead of figure 1 and table 3 (same also with section etc.).

* Don't use upper cases within any headline, e.g. "BiNoM Strucutral Analysis" -> "BiNoM structural analysis"; similarly in many other head lines

* Always explain abbreviations at their first occurence, e.g., "PIQuant" (meaning of this abbreviation explained one page later; just spell it out at the first place).

* A BioPAX file is not "big", but "large".

* "BioPAX Query functions" -> query (multiple occurrences)

* Do not use the word "allow to" without an object (see section "BiNoM Module manager"), there it must be "allow users to"

* p. 11: "between aa annotated" -> an annotated

* Use serial commas, e.g., p. 12 between "E2F1" and the word "and"; as well as between "negative" and the word "or" on the same page.

* "human Apoptosis" -> apoptosis.

* Insert a blank after citation [18] at page 12.
* Spell the first "t" in "t-test" in italics (multiple occurrences).
* In the "User's manual" p. 52: check that words fit the margins.
* BiNoM's website does not provive JavaDoc, i.e., API usage difficult.

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician.