Eric Bonnet
Institut Curie
Team Computational Systems Biology of Cancer
26 rue d'Ulm
75005 PARIS

Thursday, December 21, 2012
Biomed Central
BMC Systems Biology

Dear Dr Le Novère,

We thank you for the timely review of our manuscript entitled "BiNoM, a Cytoscape plugin for accessing and analyzing pathways using standard systems biology formats", reference MS1440652649801836. We have carefully analyzed and addressed each of the referee's comments and have incorporated the appropriate changes into our revised manuscript. Please find below the detailed answers to the referees comments.

Reviewer number 1: 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, otherformats, e.g., CellML, GraphML, SBGN-ML etc. are, however, not supported. Make more explicit that only these two formats are supported, not "such as".
We acknowledge that this sentence in the abstract was ambiguous. The text has now been reformulated to specify precisely what types of files can be imported and exported. Furthermore, we have added a new table to the manuscript detailing what are exactly the import/export possibilities of BiNoM (Table 3).

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.
In fact, BiNoM is designed to provide functions to import and export networks to the SBML level 2 format (the import is done through the "Import CellDesigner document from file" function). When a CellDesigner file is imported in BiNoM and then exported to the SBML level 2 format, there is a conversion from the CellDesigner layout representation to the

internal BiNoM format (which is detailed in the BiNoM manual), but without using the SBML layout extension. We agree that supporting this extension of SBML would be very useful for BiNoM users, and we have actually started to work on this function. We have added detailed explanations on those points in the main text.

Section Background:
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.
We fully agree with this remark and we have changed the text related to CellDesigner in the background section.

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?
We completely agree with the statements above on the development and usage of the community standards for Systems Biology. However, our focus on CellDesigner instead of SBML was motivated by pragmatic and priority considerations. CellDesigner is a very popular tool within the community of people constructing and using molecular maps. One of our priorities with BiNoM was to provide a robust set of functions to be able to manipulate and analyze the maps produced by CellDesigner. Obviously, SBML is the format of choice for mathematical modeling and is used as the standard format by most of the software tools in this field. Therefore, we have also implemented functions to export networks to SBML level 2 in BiNoM, precisely in the perspective where a user would like to perform mathematical modeling using more dedicated tools. As stated in the previous paragraphs, we completely agree with the idea of providing a more complete SBML support, and in fact we have started to work in that direction for the next release of the BiNoM plugin. We have changed the text in the background section to clarify all those points.

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.
We agree that this statement was a bit vague. As also requested by reviewer 3, we have now included a more detailed analysis and presentation of existing tools for the conversion of standard systems biology file formats, decomposition/reduction and analysis of large biological networks (Table 4).

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?

BiNoM manipulates the information contained in the standard systems biology files by mapping it onto a labeled graph, called index. The index does not try to map the totality of all details but rather serves as a connection map for the objects contained in other ontologies. The index contains the minimum information needed to graphically represent objects and connections between them. BiNoM index is a light-weight construction which can be easily regenerated, does not duplicate the information in existing files and serves only to facilitate the visualization and to access existing systems biology files. Currently, BiNoM index is mostly developed to map BioPAX ontology files and CellDesigner object schema. More specifically, we are using XmlBeans (xmlbeans.apache.org/) to create Java classes from the xml definition files in order to access all the elements contained in the CellDesigner and SBML files. For BioPAX files, we are using the Jastor library (jastor.sourceforge.net) to create Java classes from the BioPAX OWL specification file. We have added a paragraph in the "implementation" section of the manuscript to explain those details of the implementation.
Concerning the question of lossless transformation, we do not think that this problem is solvable without introducing some *ad hoc* extensions to the existing formats, and we did not have such a goal in BiNoM. BiNoM allows working in some practical scenarios: convert CellDesigner to BioPAX for sending pathway information to a database, convert BioPAX into SBML for further mathematical modeling, etc. All these scenarios are implemented to satisfy the needs met in concrete projects, rather than practicing a theoretical exercise on possibility of lossless transform. Since these functions were useful for us, we believe that they will be useful for other users. Nevertheless, some adequacy of transformations was systematically tested. For example, when a CellDesigner file is converted to BioPAX, we systematically compare the import from the resulting BioPAX file with the initially imported CellDesigner file and verify that it represents the same reaction network (in terms of species and reactions).

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.

We have now included a changelog in the supplementary methods.

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.
We agree with this remark and we are aware about possbile misleading conclusions which can be made by applying path analysis blindly to reaction graphs (bipartite graphs, or hypergraphs). BiNoM currently does not take this into account (a proper solution would be to use hypergraph traversals which can be indeed a subject for future improvements), providing just basic path finding algorithms. Short connections can be avoided by removing first a part of the hubs or very common "catalyzers" (such as ATP), but this remains for the user's decision.

Section Pathway influence quantification algorithm:
Illustrate the PIQuant score with a small figure if appropriate.
We would be happy to provide an extra figure illustrating the PIQuant score, but we are not exactly sure what kind of illustration is suggested by the reviewer (the score is already illustrated by a toy influence network model on figure 5 linked to the detailed explanations in the text).

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.
Indeed this work was done in our lab. The sentence has been changed in the text.

Is the qual extension supported for GINsim?
No, the qual extension is not supported yet 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.
We agree with this remark and we have now removed from the conclusion any reference to ongoing or future projects implementations for BiNoM that are not published yet.

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).

As requested by another reviewer, we have now included a much more complete comparison of BiNoM to other Cytoscape plugins and related software tools into the manuscript (see the background section and Table 4).

Supplementary Excel spreadsheet:

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

We have now included a graphical representation of the table 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.***

There is graphical support for compartments, but they are indicated in the name of the nodes (i.e. the name of the species), by following the internal BiNoM text ontology. For example, the entity p53 located in the nucleus will be indicated as "p53@nucleus". A detailed description of the BiNoM naming procedure is indicated in the main text (section "Implementation"), as well as in the manual.

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

The current version of BiNoM provides support for BioPAX level 3 only and therefore a check is done on the input file to verify the level. However, I have checked the Apoptosis file provided on our website (section "Documentation and Source Code", "BiNoM manual and the files mentioned in it"), and the file is at the BioPAX level 3 format, and can be opened in BiNoM. Unfortunately, our current staff does not allow us to support all versions of Systems Biology standards: for doing this, a user should rely on external converters or use older versions of BiNoM.

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

BiNoM implements its own graphical style to represent entities and their relationships and this style is inspired by the SBGN standard, that we are trying to follow as much as possible. However, in some cases, our representation might be slightly different. We have included a complete overview of the BiNoM graphical styles in the manual (see the section 8).). We think that having an arrow head to a reaction node is useful in any structural analysis of the reaction graph, showing its directionality.

**** 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.***

We are not sure to fully understand the statement above. When importing data from a CellDesigner or BioPAX file, BiNoM is using a naming service to create "meaningful" names for each entity (provided that they are present in the file, of course), specifying the type of the molecule (protein, RNA, small molecule, protein), the modifications (phosphorylation or any other type of modification) and the compartment. . This is not a "cryptic" abbreviation but rather a well-defined way for textual representation of the structure of the nodes in process diagrams, which makes possible many semantic-based analyses afterwards. Our experience shows that a user easily gets used to this representation. More specifically for a BioPAX file, BiNoM selects the name of the entity from the "displayName", when it is provided, and other "name" fields present in the file. Furthermore, BiNoM provides a powerful BioPAX querying mechanism (see the section BiNoM BioPAX utils & query) that allows the user to retrieve most of the data encoded in the BioPAX file such as database crosslinks, publications, etc.

**** 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.***

It is indeed a very valuable suggestion to include the MIRIAM annotations, and we will surely include this type of information in the next BiNoM release. We have included a warning in the main text about the MIRIAM annotations, that are indeed lost in the current version.

**** 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.***

Indeed the support of SBML in BiNoM is restricted to SBML level 2. This is so for the simple reason that CellDesigner currently follows SBML level 2, and for us CellDesigner is a main source of reaction graphs (also we did not find useful to re-implement import functions

already existing in Cytoscape or many other tools). We have made this very clear throughout the text. We have also introduced a new table in the manuscript, detailing exactly what are the type and version of standard System Biology file formats that BiNoM supports for import-export operations.

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

There was indeed an error in the code preventing the correct import of SBML level 2 files, but this error is now corrected in the current BiNoM version.

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).***

**** 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.***

We have corrected all those errors.

**** Implementation section: why is "Module manager" and "Utilities" written in upper cases?***
That is actually the way they are spelled in the BiNoM menu.

**** "gene regulatory networks" -> "gene-regulatory networks"***
We believe that "gene regulatory networks" is correct and is actually encountered frequently in the literature.

**** BiNoM's website does not provive JavaDoc, i.e., API usage difficult.***
The JavaDoc has been added to the website.

Reviewer number 2

Discretionary Revisions
1. In the background section, the authors mentioned three standard pathway exchange formats: BioPAX, SBML and SBGN. The manuscript describes the support of BiNoM to both BioPAX and SBML, but there is no mention about SBGN. The figures don't seem to be compliant with SBGN standard. It would be helpful if the authors can provide some description about their plan on the graphical representation standards.
We agree that the background section is not very explicit about the relationships between the SBGN standard and BiNoM. In fact, the visualization in BiNoM of the different types of entities and their relationships is inspired by the SBGN standards although we do not strictly follow all the SBGN rules. As a matter of fact there is a Cytoscape plugin under development that is dedicated to the import and visualization of SBGN files (cySBGN, http://www.ebi.ac.uk/saezrodriguez/cysbgn/). We have changed the text in the background section to clarify the point mentioned by the reviewer.

2. A supplement with more details on the installation of BiNoM would be helpful, especially to users who are novice to Cytoscape.
We have added a section devoted to the installation of BiNoM in the supplementary methods.

Reviewer number: 3
Major Revisions
(1) As a variety of methods and tools for the conversion of standard systems biology file formats and the analysis and decomposition/reduction of large biological networks exit already, a more detailed consideration of related or previous work should be given and the benefit/novelty of the tool in comparison to already existing tools should be stated more clearly. Some related work with respect to file conversion includes:

• Sybill: Rübenacker,O. et al. (2009) Integrating BioPAX pathway knowledge with SBML models. IET Syst. Biol., 3, 317–328.

• System Biology Format Converter (SBFC):

http://www.ebi.ac.uk/compneur-srv/sbml/converters/SBMLtoBioPax.html Some related work with respect to network decomposition/clustering/modularization and structural analysis within Cytoscape

include:

• ShortestPath Plugin

• GLay: Su.G. et al. (2010) GLay: community structure analysis of biological networks. Bioinformatics, 26, 3135#3137.

• MCode: Bader,G.D. and Houge,C.W.V (2003) An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics, 4, 2.

• ModuLand: Szalay-Beko M. et al. (2012) ModuLand plug-in for Cytoscape: determination of hierarchical layers of overlapping network modules and community centrality. Bioinformatics 28(16):2202-4

• ClusterMaker: http://www.cgl.ucsf.edu/

• NeMo: Rivera et al. (2010) NeMo: Network Module identification in Cytoscape. BMC Bioinformatics 18;11 Suppl 1:S61.

This point was also mentioned by another reviewer. We have now introduced a detailed review of the BiNoM functions compared to all the tools mentioned by the reviewers. More specifically, a comparison table (Table 4) has been introduced and the main text has been modified in the “Background” section.

(2) Please state more clearly the novelty and progress of the current work with respect to the work published earlier (Zinovyev et al., 2008).

This point was also raised by another reviewer. We have added a “changelog” section to the supplementary methods, detailing the changes that have been made between the original version of BiNoM and this one.

Minor Essential Revisions

(1) Tutorial:

There is no G1S.xml file provided on the website as mentioned in the tutorial. Instead a Cytoscape session file is provided, but the Material Component analysis of the given G1S network does not result into 36 subnetworks. Please correct and/or provide the xml-file mentioned.

Indeed the file and the associated numbers were not correct. We have now provided the correct file, and we have also checked and corrected the complete modularization procedure.

(2) Figures/Tables:

• Figure 2: Due to its low informative value, it might be reasonable to change figure 2 and provide a smaller example providing node labels

• Table 1: As there is no difference between the listed path analysis algorithms with respect to the listed options, it is unclear why the table is provided instead of just mentioning the information in the text

The Figure 2 has been modified and now depicts a zoom on the Apoptosis network, showing the node labels. For the Table 1, we preferred to include this information as a small separate table, in order to avoid complicated and heavy explanations in the text.

(3) References in the text:

As already done for some algorithms implemented in BiNoM, in the text for each implemented method a respective literature reference should be given (if available). Implemented methods without a reference are for example:

• Material component decomposition (page 5)

• Path analysis algorithms: suboptimal shortest path, non self-intersecting paths (page 6)

Actually the first algorithm was developed specifically for BiNoM and is not (as far as we know) described elsewhere in the literature. Path analysis algorithms are adapted from the some standard graph theory algorithms (such as breadth-first search), however, we made some optimizations for them so referring to, for example, a graph theory textbook would be misleading. More explanations are given on how they work in the BiNoM manual.

(4) Network clustering, subsection Path analysis algorithms, 2nd paragraph: Obviously, the result of some decomposition functions will result in subnetworks that share some components, as it is for example often the case with the decomposition in material components. Therefore, BiNoM also includes a function to cluster networks, based on common components such as protein or protein complexes. To determine the size of the clusters, the user can specify a percentage of intersection (ranging from 0 to 100%) that will be used as a threshold to create the clusters.

As the subject of network clustering rather belongs to network decomposition instead of path analysis, it might be useful to include this paragraph into the section of network decomposition.

We have moved the paragraph mentioned above in the section devoted to network decomposition.

(5) Recheck spelling and language, e.g.:

• Section Application of PIQuant on an influence network, 3rd paragraph: The PIQuant score is then automatically calculated for each association between aa annotated node and a target node.

• Section Results and Discussion, 4th paragraph: This map contains a lot of valuable information but rather difficult to extract

Those errors have been corrected.

(6) Recheck title of sections/subsection as they are sometimes written in capital letters and sometimes not

This has been checked.

(7) Recheck formatting of the Reference section, e.g.:

- ***Ref. 24: journal name written in capital letters, instead of small ones as done by the others***
- ***Ref. 30: journal name written in its abbreviated form, instead of its full name as done by the others***

We have checked all the references and made corrections when necessary.