Metadata of the chapter that will be visualized online

Chapter Title	Practical Use of BiNoM: A Biological Network Manager Software	
Copyright Year	2013	
Copyright Holder	Springer Science+Business Media, LLC	
Author	Family Name	**Bonnet**
	Particle	
	Given Name	**Eric**
	Suffix	
	Organization	Institut Curie
	Address	26 rue d'Ulm, Paris, 75248, France
	Organization	INSERM, U900
	Address	Paris, 75248, France
	Organization	Mines ParisTech
	Address	Fontainebleau, 77300, France
Author	Family Name	**Calzone**
	Particle	
	Given Name	**Laurence**
	Suffix	
	Organization	Institut Curie
	Address	26 rue d'Ulm, Paris, 75248, France
	Organization	INSERM, U900
	Address	Paris, 75248, France
	Organization	Mines ParisTech
	Address	Fontainebleau, 77300, France
Author	Family Name	**Rovera**
	Particle	
	Given Name	**Daniel**
	Suffix	
	Organization	Institut Curie
	Address	26 rue d'Ulm, Paris, 75248, France
	Organization	INSERM, U900
	Address	Paris, 75248, France
	Organization	Mines ParisTech
	Address	Fontainebleau, 77300, France
Author	Family Name	**Stoll**
	Particle	
	Given Name	**Gautier**

	Suffix	
	Organization	Institut Curie
	Address	26 rue d'Ulm, Paris, 75248, France
	Organization	INSERM, U900
	Address	Paris, 75248, France
	Organization	Mines ParisTech
	Address	Fontainebleau, 77300, France
Author	Family Name	**Barillot**
	Particle	
	Given Name	**Emmanuel**
	Suffix	
	Organization	Institut Curie
	Address	26 rue d'Ulm, Paris, 75248, France
	Organization	INSERM, U900
	Address	Paris, 75248, France
	Organization	Mines ParisTech
	Address	Fontainebleau, 77300, France
Corresponding Author	Family Name	**Zinovyev**
	Particle	
	Given Name	**Andrei**
	Suffix	
	Organization	Institut Curie
	Address	26 rue d'Ulm, Paris, 75248, France
	Organization	INSERM, U900
	Address	Paris, 75248, France
	Organization	Mines ParisTech
	Address	Fontainebleau, 77300, France
	Email	andrei.zinovyev@curie.fr

Abstract	The Biological Network Manager (BiNoM) is a software tool for the manipulation and analysis of biological networks. It facilitates the import and conversion of a set of well-established systems biology file formats. It also provides a large set of graph-based algorithms that allow users to analyze and extract relevant subnetworks from large molecular maps. It has been successfully used in several projects related to the analysis of large and complex biological data, or networks from databases. In this tutorial, we present a detailed and practical case study of how to use BiNoM to analyze biological networks.

Key words (separated by "-")	Biological networks - Graph-based algorithms - Subnetworks - Molecular maps - BiNoM

Chapter 7

Practical Use of BiNoM: A Biological Network Manager Software

Eric Bonnet, Laurence Calzone, Daniel Rovera, Gautier Stoll, Emmanuel Barillot, and Andrei Zinovyev

Abstract

The Biological Network Manager (BiNoM) is a software tool for the manipulation and analysis of biological networks. It facilitates the import and conversion of a set of well-established systems biology file formats. It also provides a large set of graph-based algorithms that allow users to analyze and extract relevant subnetworks from large molecular maps. It has been successfully used in several projects related to the analysis of large and complex biological data, or networks from databases. In this tutorial, we present a detailed and practical case study of how to use BiNoM to analyze biological networks.

Key words Biological networks, Graph-based algorithms, Subnetworks, Molecular maps, BiNoM

1 Introduction

The last decade has seen unprecedented advances in the production of high-throughput experimental data in biology, fueled by drastic technological improvements in various ways of measuring biological entities. In return, those large amounts of biological information have stimulated the need of developing standards for an efficient representation and exchange of data. This is especially true for the field of systems biology, which aims at building models and quantitative or qualitative simulations of complex biological systems [1, 2]. To achieve this goal, it is obvious that a good communication and collaboration between modelers and experimentalists having various scientific backgrounds will be facilitated by the standardization of the representation of workflows, data formats, and mathematical models. Several complementary standards have already been created and are increasingly used in a large variety of projects. Most of them are based on an open-source and community-based organization, ensuring an easy access to the detailed specifications of the standard, flexibility, dynamic

Maria Victoria Schneider (ed.), *In Silico Systems Biology*, Methods in Molecular Biology: Methods and Protocols, vol. 1021,
DOI 10.1007/978-1-62703-450-0_7,

evolution, and wide acceptance. Examples of such community standards are the System Biology Markup Language (SBML) [3], a language focused on mathematical modeling, the Biological Pathway Exchange standard (BioPAX) [4], devoted to storing and exchange of pathway information and the Systems Biology Graphical Notation (SBGN), centered on the graphical notation for biological maps [5]. It is worth noting that there are now more than 40 databases and online resources supporting the BioPAX format, while more than 33 databases are using SBML (http://www.pathguide.org). Well-established examples are the Reactome database [6], BioModels [7], and MINT [8]. More and more systems biology software packages are also using standard formats to store and exchange data. For example, CellDesigner is a tool used to edit biological pathways diagrams [9] and is using a compatible SBML dialect to store the all the information related to a given diagram. Cytoscape is a widely spread program used for the visualization, modeling, and analysis of complex molecular and genetic interaction networks [10]. BiNoM was developed as a Cytoscape plugin, with the goal of facilitating the import and export of various systems biology formats, and also proposing a large set of graph-based algorithms for the extraction of relevant subnetworks from large molecular maps [11]. BiNoM was successfully used in several projects related to the analysis of complex biological networks [12]. Here, we present a set of detailed and concrete examples of how to extract relevant information from such maps using BiNoM.

2 Material

The Cytoscape [10] software should be installed on the computer (http://cytoscape.org). The BiNoM plugin can be installed in different ways. The first is to download BiNoM from our web page (http://bioinfo-out.curie.fr/projects/binom/) and copy the file under the directory "plugins" of the Cytoscape installation folder (administrator privileges might be necessary to perform this operation). The latest version of BiNoM (version 2.0) supports the latest versions of Cytoscape. The previous version (BiNoM 1.0) is also available on our website for older versions of Cytoscape.

Another possibility to install BiNoM is to use the plugin manager of Cytoscape. Starting in version 2.5, the plugin management has been added to allow users to search for, download, install, update, and delete plugins within Cytoscape.

1. Select the function "Plugins > Manage Plugins" from the menu.
2. Navigate in the tree view to the category "Analysis."

3. Select BiNoM v2.0 (or any more recent version available).
4. Click "Install."

All the files used throughout this tutorial can be downloaded from our website (http://bioinfo-out.curie.fr/projects/binom/).

3 Methods

3.1 Import and Export

A major function of the BiNoM software is to provide import and export functions for a given number of standard systems biology file formats. It is therefore possible to import data from SBML level 2 files, CellDesigner 3.X and 4.X files, BioPAX level 3, CSML files and also from simple text files formatted to the AIN (Annotated Influence Network) format. The aim of BiNoM import/export functions is not to be a universal converter but rather to favor a number of scenarios where the conversion is making sense (Fig. 1). It is worth mentioning that due to major changes in the specifications, the BioPAX level 3 format is incompatible with the previous level 2 format [4]. The previous version of BiNoM (version 1.0, still available from our website) was managing the BioPAX level 2, but the latest version of BiNoM (2.0) can only deal with BioPAX level 3 files. The current version of Cytoscape does not support a direct import of BioPAX level 3 files yet [10].

Let us take an example. The model of the yeast cell cycle by Novak and colleagues [13] was encoded as a graph using CellDesigner software [9]. We can easily import it in Cytoscape using the BiNoM functions. The file is available on the BiNoM website (file name: M-Phase.xml).

1. Select the function "Plugins > BiNoM 2.0 > BiNoM I/O > Import CellDesigner document from file" from the menu.
2. Select the file from the dialog window.
3. Click OK.

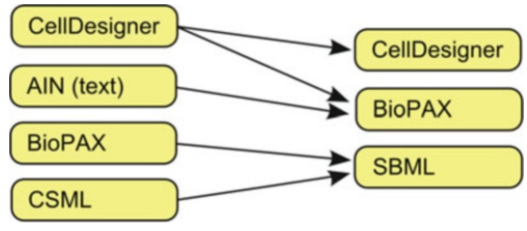

Fig. 1 The BiNoM import/export functions favor a number of scenarios that are illustrated on the figure (*left side*: import file formats, *right side*: export file formats). Note that for the CellDesigner to CellDesigner conversion, it is possible to split the network, change the layout, and change the color and the scale

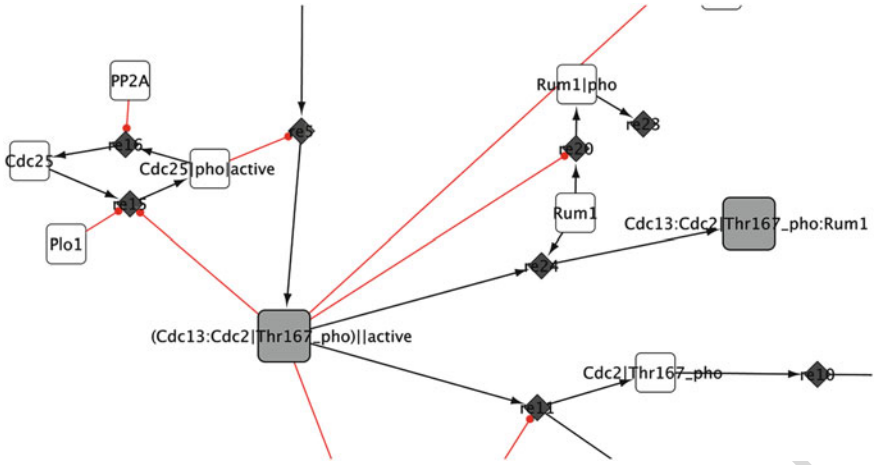

Fig. 2 Zoom on a simple cell cycle network imported from CellDesigner into Cytoscape using BiNoM

4. A new network is created as "M-Phase.xml" with 36 nodes and 42 edges (Fig. 2).

BiNoM uses its own visual mapping to represent the different molecules and their interactions, inspired by the SBGN standard [5], but presents a simplified version of it. For example, simple proteins are represented by white circles, while protein complexes are pictured as gray circles. Similarly, there is a specific mapping for the different relationships between molecules: for example, a catalysis relation will be represented as a red colored edge with a circular end. Figure 3 shows the BiNoM visual styles for BioPAX and CellDesigner. When importing pathway information, BiNoM generates meaningful names for every chemical species, following preestablished rules. Chemical species are defined as physical entities (e.g., a protein) with an optional cellular localization and posttranslational modifications. The name formatting rules are as follows:

(Entity1_name|Modification:Entity2_name|Modification)[_active|_hmN]@compartment

The colon symbol "**:**"delimitates the different components of a complex, the vertical bar "|" indicates the posttranslational modifications, while the "@" sign indicates the cellular compartment. The optional suffixes "active" or "hm" indicate the state of the chemical species and N-homodimer state, respectively. The parentheses delimitate the components concerned by the N-homodimer state and are useful to eliminate ambiguities (*see* Fig. 2 for examples).

We have recently developed a new import format-denominated AIN. The principle of this format is to encode an influence network, where edges represent either an inhibition or an activation,

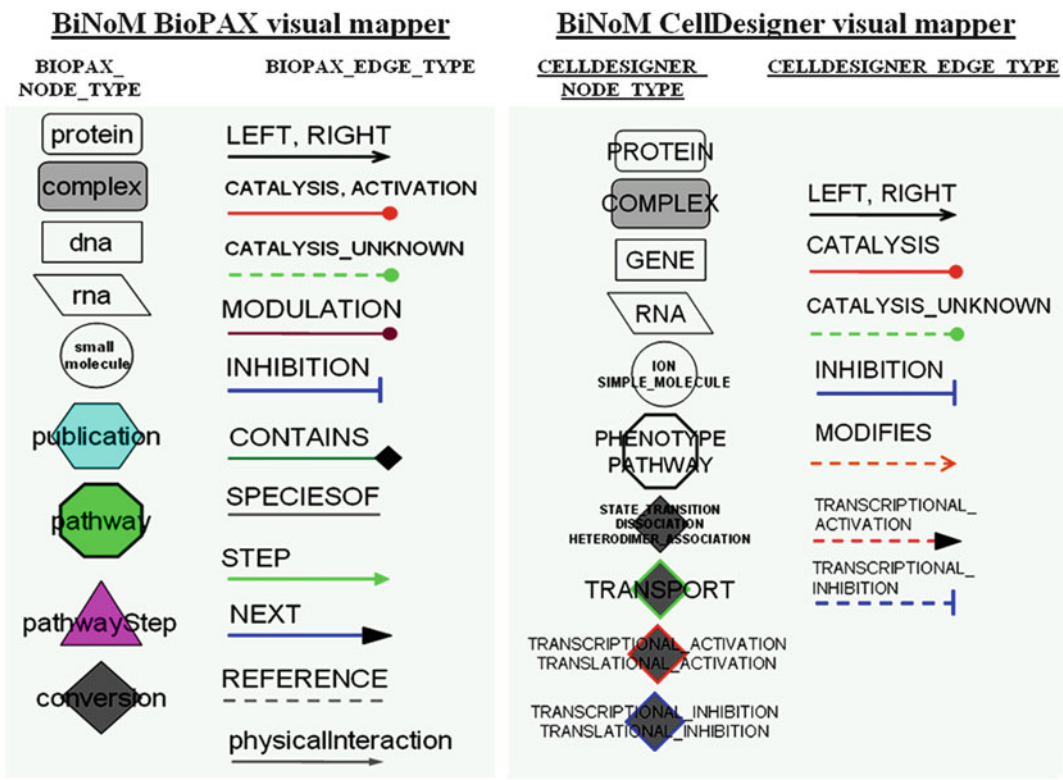

Fig. 3 Comprehensive visual representations followed by the BiNoM software for different entities and their relationships, for both the BioPAX and CellDesigner file formats

into a simple tab-delimited text file (*see* Table 1 for a detailed explanation of the AIN format). Using this format, it is rather straightforward for a biological expert to encode a network from his own expertise and/or from published results, using a spreadsheet program such as Excel, and then import it in Cytoscape using BiNoM, rather than using more sophisticated tools such as CellDesigner. All the information contained in the AIN file is automatically converted in the BioPAX format when the file is imported and can be subsequently retrieved with specific BiNoM functions. Let us now import a simple cell cycle model encoded as an AIN file (cell_cycle_AIN.txt, available from the BiNoM website).

1. Select the function "Plugins > BiNoM 2.0 > BiNoM I/O > Import influence network from AIN file" from the menu.
2. Select the file "cell_cycle_AIN.txt" from the dialog window and click OK.
3. Click OK twice, for the windows "Defining families" and "Select constitutive reactions to add."
4. The network is imported as "cell_cycle_AIN" (Fig. 4).

t.1 **Table 1**
Description of the AIN format

t.2 Column number	Column caption	Description	Example(s)
t.3 1	ReviewRef	A reference (e.g., a PubMed ID) to an article	PMID:1234
t.4 2	ExperimentRef	A reference to an experiment	PMID:10783242
t.5 3	Link	Connection (activation or inhibition) between two entities. The name can represent a single protein, a protein complex "(C:D)," a phosphorylated protein "(C^p)," or a family. For the latter, the family can be given explicitly by the full list of the members "(C1, C2, C3)," or implicitly by using an undefined name where a dot will represent any character "(C.)"	"A->B," "A-|B," "((CCNE.):CDK2)->E2F5^p," "(E2F1, E2F2)->CDKN2A"
t.6 4	ChemType	Chemical type of the reaction	Binding
t.7 5	Delay	Delay of the reaction (numerical value and unit)	0.9 h
t.8 6	Confidence	Confidence level in the reaction (value between 0 and 1)	0.8
t.9 7	Tissue	Tissue where the reaction has been observed	Fibroblast
t.10 8	Comment	Comment about the reaction	"Specific phosphorylated site of E2F5"

t.11 Each line of the table represents a column of the AIN tab-delimited file. Columns are numbered from left to right. Missing values should be indicated by a single dot and text strings should be quoted. The only mandatory column is the Link (column number 3), representing the reaction

3.2 Manipulating Existing Networks

The cell cycle model of Novak et al. (Fig. 2) has 36 nodes and 42 edges in total, making it a rather small network. However, this is not very often the case. On the contrary, most networks publicly available from online databases such as Reactome [6] or large molecular maps built from the literature such as the RB/E2F map [12] have hundreds of nodes and edges, if not more. Such gigantic maps are barely readable and manageable when imported into visualization software such as Cytoscape. One of the main ideas of BiNoM plugin was to provide a set of network tools allowing users to extract meaningful subnetworks from large molecular maps and also to provide means to understand and read these maps [11]. We will see now through a set of examples how to extract such meaningful subnetworks.

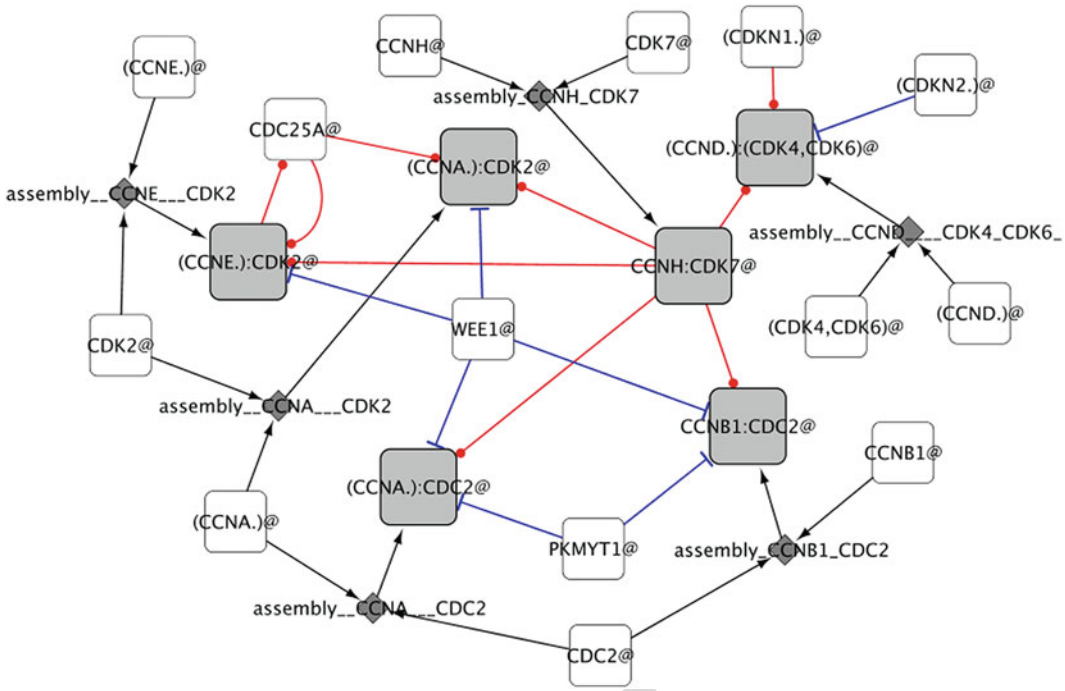

Fig. 4 A cell cycle network imported from an AIN text file (Annotated Influence Network)

As a first exercise, we create a simpler modular view of the M-phase example. Let us first decompose the cell cycle map we have imported in the previous paragraph by pruning the graph. In large networks, this step is important in order to simplify the network: we work on the connected graph rather than the whole graph.

1. Select the network "M-Phase.xml."
2. Choose the function "Plugins > BiNoM 2.0 > BiNoM Analysis > Prune Graph" in the menu.
3. Three networks are created: "M-Phase.xml_in," "M-Phase.xml_scc," and "M-Phase.xml_out."

The function is decomposing any network in three components corresponding to the nodes that are coming in (input), the nodes that go out (output), and the central cyclic part. The central part may sometimes be composed of several strongly connected components. In some situations they can be disconnected, forming several subnetworks. The decomposition of the strongly connected components part can be done in two ways: (1) by cycle decomposition and (2) by material components decomposition. Let us first see how to decompose a network into relevant directed cycles, which usually provides information about the life cycle of a gene or protein of the network.

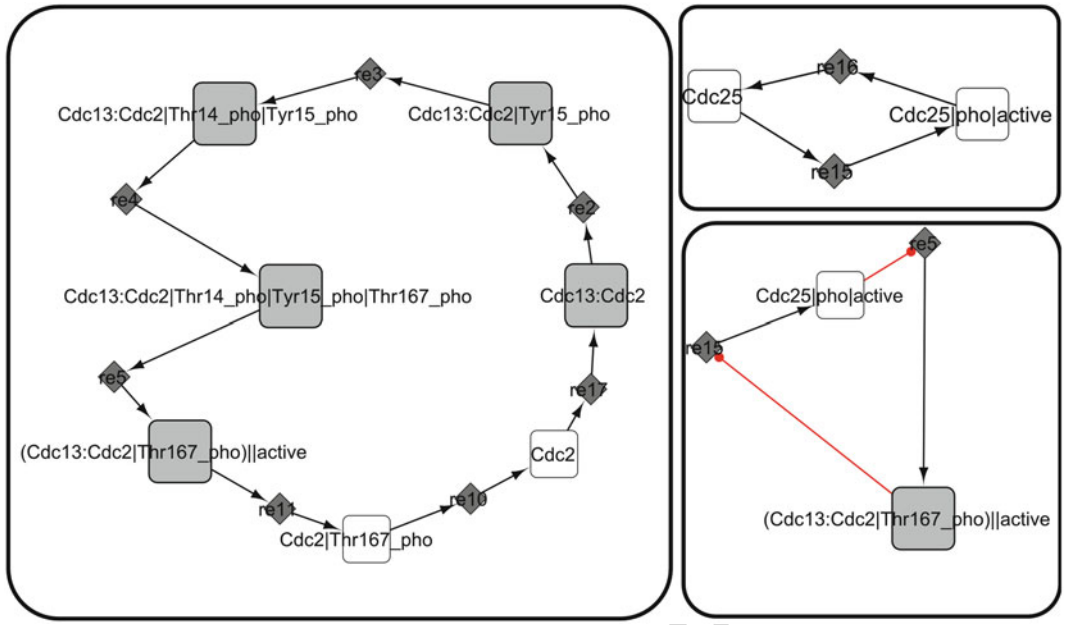

Fig. 5 Subnetworks (cycles) extracted from the M-Phase network using BiNoM functions

1. Select the network "M-Phase.xml_scc" (highlighted in the Cytoscape navigation panel).
2. Select the function "Plugins > BiNoM 2.0 > BiNoM Analysis > Get cycle decomposition" from the menu.
3. Three new networks are created: cycle1, cycle2, and cycle3 (Fig. 5).

Let us now merge in clusters networks that share a certain number of components.

1. Select the function "Plugins > BiNoM 2.0 > BiNoM Analysis > Cluster Networks" from the menu.
2. In the dialog window that appeared, select the networks cycle1, cycle2 and cycle3 (holding down the CTRL key for multiple selection).
3. Set the intersection threshold to 35 % using the sliding bar.
4. Click OK. Two networks are created: "cycle1" and "cycle2/cycle3."

In fact, only the networks cycle2 and cycle3 were clustered, because they share a component (Cdc25 phosphorylated and active; in a two-component network, they share more than 35 %). Now that the modules are created, we need to include the inputs and outputs that were put aside at the beginning of the analysis.

In order to merge networks, we can use a Cytoscape built-in function.

1. Select the function "Plugins > Advanced Network Merge" from the menu.
2. From the dialog window, select "Union" in the field "Operation."
3. In the list of networks, select "Cycle1," "M_Phase.xml_in," and "M_Phase.xml_out" and then click "Merge."
4. The resulting network is named "Union" and should have 30 nodes and 19 edges.
5. Rename the network to "Union1" by right-clicking on its name and selecting "Edit Network Title."
6. Using the same procedure as above, merge the networks "cycle2/3," "M_Phase.xml_in," and "M_Phase.xml_out." The resulting network should have 22 nodes and 12 edges.
7. Rename it to "Union2/3."

Some edges present in the original file have been lost during all these operations, and they need to be included again. For that, we will now update the networks.

1. Select the function "Plugins > BiNoM 2.0 > BiNoM Utilities > Update connections from other network" from the menu.
2. In the dialog window, select "M-Phase.xml" for the field "From Network" and select the networks "Union1" and "Union2/3" from the list "Networks to Update," and click OK.
3. The networks "Union1" and "Union2/3" are updated to 30 and 20 edges, respectively.

We can now remove unconnected and unnecessary components.

1. Select the network "Union1."
2. Change the layout by using the Cytoscape function "Layout > yFiles > Organic" from the menu (this step allows to visualize the unconnected components more easily).
3. Select all unconnected nodes and remove them.
4. The Wee1 and Rum1 genes should be in a network of their own, so we propose to remove them and all the edges connected to them (they are in fact important proteins that do not share a function with the two modules created).
5. The resulting networks should have 20 nodes and 20 edges for "Union1" (Fig. 6) and 8 nodes and 9 edges for "Union2/3" (Fig. 7).

Note that the analysis requires to make a certain number of choices, based on biological knowledge and related to the final goal of the user. Here, we want to create a modular view of the initial network to highlight the main mechanisms involved in the yeast cell cycle progression. Finally, we can now generate a modular view from the networks Union1 and Union2/3.

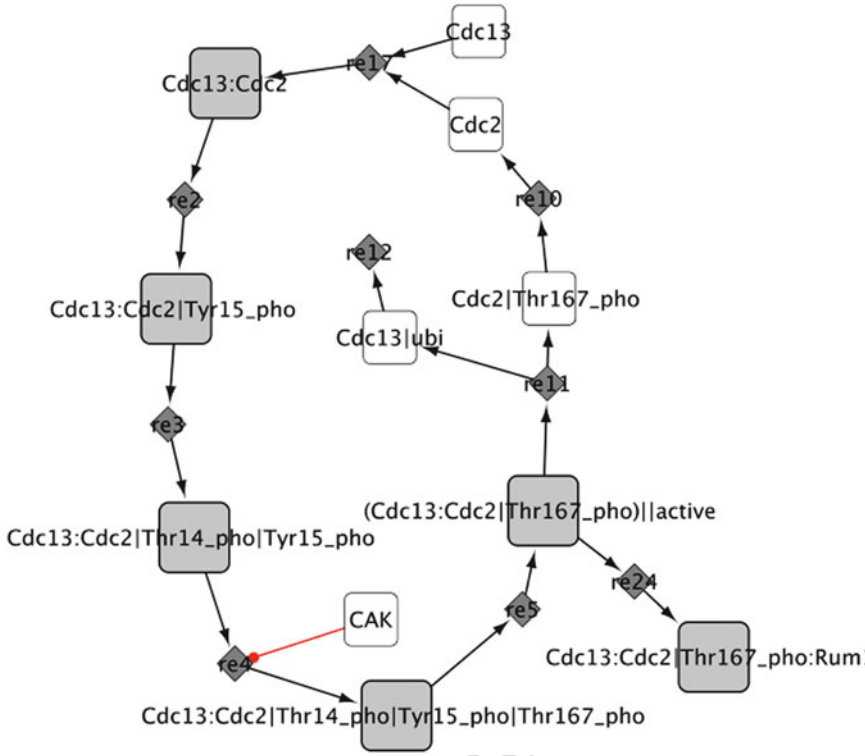

Fig. 6 A subnetwork resulting from the union of two subnetworks

This figure will be printed in b/w

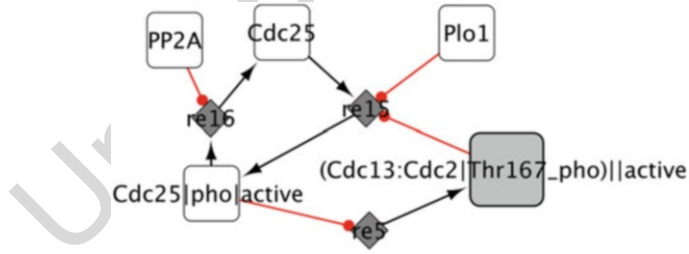

Fig. 7 A subnetwork resulting from the union of two subnetworks

This figure will be printed in b/w

AU2

1. Select the function "Plugins > BiNoM 2.0 > BiNoM module manager > Create Network of Modules" from the menu.
2. Select "Union1" and "Union2/3" from the list in the dialog window and click OK.
3. Select the function "Plugins > BiNoM 2.0 > BiNoM module manager > Create connections between modules" from the menu. Select the network "M-Phase.xml" from the list in the dialog window and click OK.

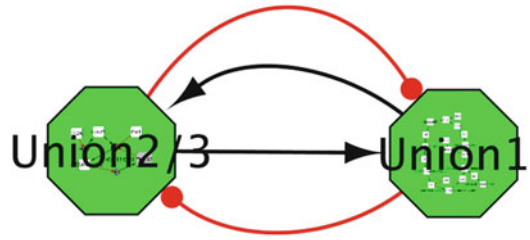

Fig. 8 A modular representation of two subnetworks, Union2/3 and Union1

4. Rename the network to "Module1" by right-clicking on it (Fig. 8).

The resulting map is a modular map of the initial network, in which modules participate in a specific process. For instance, "Union 1" shows all the events that lead to the activation of the maturation promoting factor, a heterodimer composed of the cyclin-dependent kinase Cdc2 and the cyclin B protein Cdc13. Note that in order to navigate from a module to the corresponding subnetwork, you have to perform the following operations:

1. Right-click on the module of interest. A contextual menu appears.
2. In the menu, choose "Nested Network" and then "Go to Nested Network." The corresponding subnetwork is now brought to the front window.

3.3 BiNoM and BioPAX Files

Biological Pathway Exchange (BioPAX) is a standard language that represents biological pathways at the molecular level and facilitates the exchange of pathway data [4]. The current BioPAX specification (level 3, released in July 2010; see http://www.biopax.org), supports representation of metabolic and signaling maps, molecular and genetic interactions, and gene regulation. Furthermore, there are several additional constructs available to store extra details such as database cross-references, chemical structures, sequence feature locations, and links to controlled vocabulary terms encoded in various ontologies (such as the Gene Ontology). BiNoM has a powerful set of functions to manage large BioPAX files, allowing the user to import, export, analyze, and extract the knowledge encoded using this specification [11]. We have recently updated the BiNoM plugin software to provide support for the latest BioPAX specification (level 3; see http://www.biopax.org/specification.php).

3.3.1 Import and Information Extraction from a BioPAX File

In the next example, we will be working with a relatively large molecular map representing the Apoptosis pathway in human, extracted from the Reactome database [6]. The file is available from our website (Apopotosis3.owl). Let us import the file in Cytoscape using BiNoM functions.

1. Select the function "Plugins > BiNoM 2.0 > BiNoM I/O > Import BioPAX 3 Document from file" from the menu.
2. Select the file "Apoptosis3.owl" from the dialog box.
3. A new dialog window appears. The three types of network should be imported. For that, check the boxes "Reaction Network," "Pathway Hierarchy," "Make Root Pathway Node," "Include Pathways," "Include Interactions," and "Interaction map."
4. Click OK. Three new networks are created, corresponding to the reaction network ("Apopotosis3 RN"), Apoptosis Pathway ("Apoptosis3 PS"), and Apoptosis protein–protein interactions ("Apoptosis3 PP").
5. Change the layout of each network for a better readability: choose "Organic" ("Layout > yFiles > Organic") for "Apoptosis3 RN" and "Apoptosis3 PP" and the type "Hierarchic" ("Layout > yFiles > Hierarchic") for "Apoptosis3 PS."

The three networks represent different types of knowledge extracted from the BioPAX file. We call them *network interfaces*, because they allow to access the different parts of the content of a BioPAX file. The Reaction Network (RN) is a graph which contains nodes of two types: "species" and "reactions." Proteins are represented as white rounded squares, complexes as gray rounded squares, and reactions as small gray diamonds (Fig. 9c). The Pathway Hierarchy (PS) contains pathway knowledge with two types of nodes: pathways, pictured as green hexagons, and pathway steps, pictured as pink triangles (Fig. 9b). The last interface contains an interaction map (IM) extracted from the proteins and protein complexes present in the BioPAX file, with edges of type "contains" (Fig. 9a).

The whole network being quite large, it is not always easy to find specific information. In the next example, we propose to query the graph by performing a simple analysis on a BioPAX imported file: the extraction of a path.

1. Select the network "Apoptosis3 RN" by clicking on the name in the navigation panel.
2. Select all nodes and edges by using the function "Select > Select All Nodes and Edges" from the menu.
3. Select the function "Plugins > BiNoM Analysis > Path Analysis" from the menu.
4. A dialog window appears; choose the node "TNF:TNFR1@plasma_membrane" in the "Sources" list and the node "GIG3:RIP:TRADD:TRAP3@cytosol" in the "Targets" list.
5. Take the default search options "Find shortest paths" (you can try the other options as an exercise).

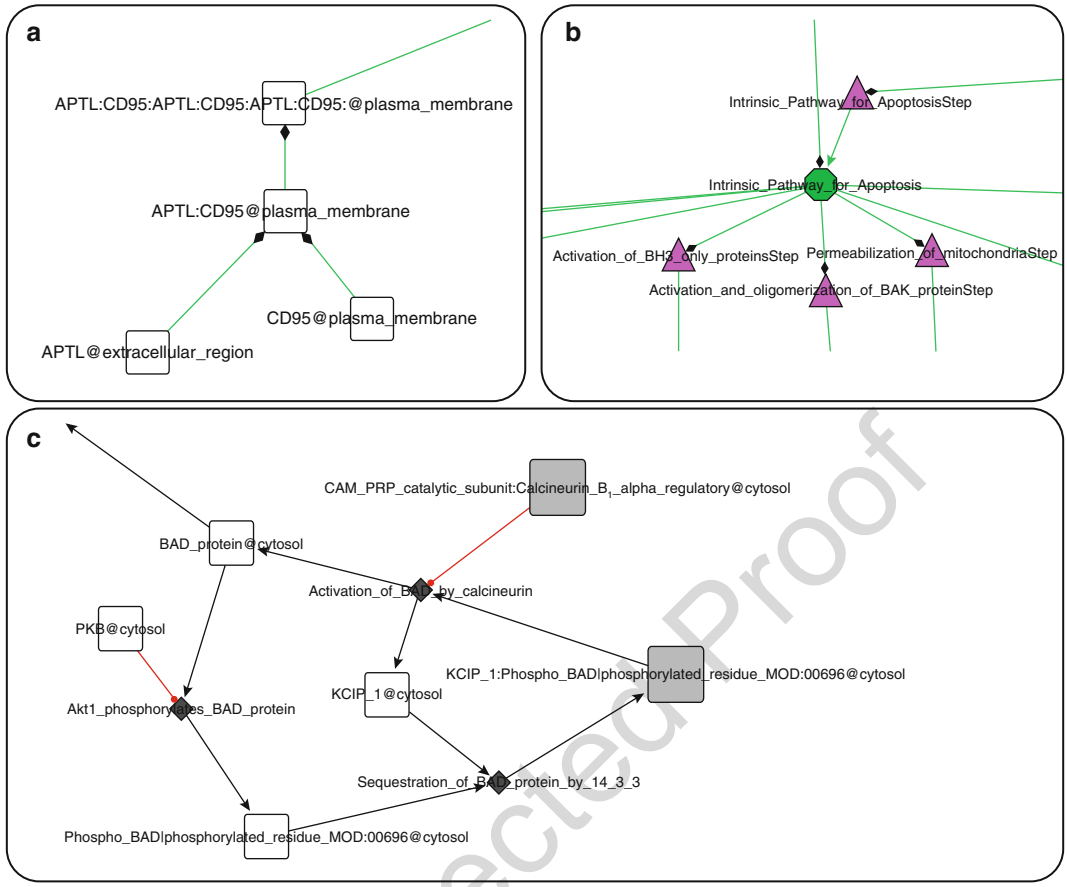

Fig. 9 Three types of networks resulting from the import of a BioPAX file. (**a**) Reaction network (RN). (**b**) Pathway hierarchical structure (PS). (**c**) Protein–protein interaction network (PP)

6. Click OK. The nodes of the shortest path between the two nodes are now highlighted in the network (note that it is not the case for the edges connecting them).
7. Extract the path as a new network by using the function "File > New > Network > From selected nodes, all edges" from the menu.
8. A new subnetwork is created with the name "Apoptosis3 RN—child."

3.3.2 Querying a BioPAX File

The BioPAX format is now used by an increasing number of databases and online repositories such as Reactome (http://www.reactome.org), Cancer Cell Map (http://cancer.cellmap.org), the Pathway Interaction Database (http://pid.nci.nih.gov/), or Pathway Commons (http://www.pathwaycommons.org). The amount of information contained in the files extracted from those databases can be very consequent, making it difficult for the average user to

efficiently query and retrieve relevant data. We have included in BiNoM an efficient BioPAX querying system. The BioPAX file is converted to an index, by mapping the BioPAX content on a labeled graph. This index can then be queried by the user for specific elements of interest. The result is returned as a graph directly in Cytoscape and can be further extended to include various elements such as all the complexes in which a protein of interest is involved, the reactions connected to those molecules, and the related publications. For example, let us extract the complexes related to a given protein from the Apoptosis BioPAX file.

1. First, we have to generate the index from the BioPAX file. Select the function "Plugins > BiNoM 2.0 > BiNoM BioPAX 3 Query > Generate Index" from the menu.
2. From the dialog window that appears, select the file "Apoptosis3.owl" for the field "BioPAX File." The second field named "Index File" will be filled automatically with the same file name, just changing the extension to ".xgmml." In this case, it will suggest the name "Apoptosis3.xgmml"; you can change that name if you wish or just accept the proposition. Click OK. The index is generated and saved.
3. Load the index with the function "Plugins > BiNoM 2.0 > BiNoM BioPAX 3 Query > Load Index" from the menu. Select the index file you have just created "Apoptosis3.xgmml" and click OK. The index is now loaded in memory (note that loading the index is an essential step to perform a query; the creation of the index is not enough).
4. Basic statistics related to the index file content can be obtained by the function "Plugins > BiNoM 2.0 > BiNoM BioPAX 3 Query > Load Index" from the menu. A window is displayed containing a table with counts for different elements of the index (proteins, complexes, publications, etc.).
5. Let us now do a basic query. Select the function "Plugins > BiNoM 2.0 > BiNoM BioPAX 3 Query > Select Entities" from the menu.
6. In the text field entitled "Input," type the name "SMAC," and click OK. A new network is created, having a single node named "SMAC@cytosol."
7. For a better visualization, you can set the visual style to "BiNoM BioPAX" on the tab "VizMapper" on the left-hand side of the Cytoscape interface.
8. Now we wish to expand this network by adding all the complexes in which this protein is involved. Select the function "Plugins > BiNoM 2.0 > BiNoM BioPAX 3 Query > Standard Query" from the menu. A window appears, named "BioPAX Standard Query from the index." Check the boxes for the

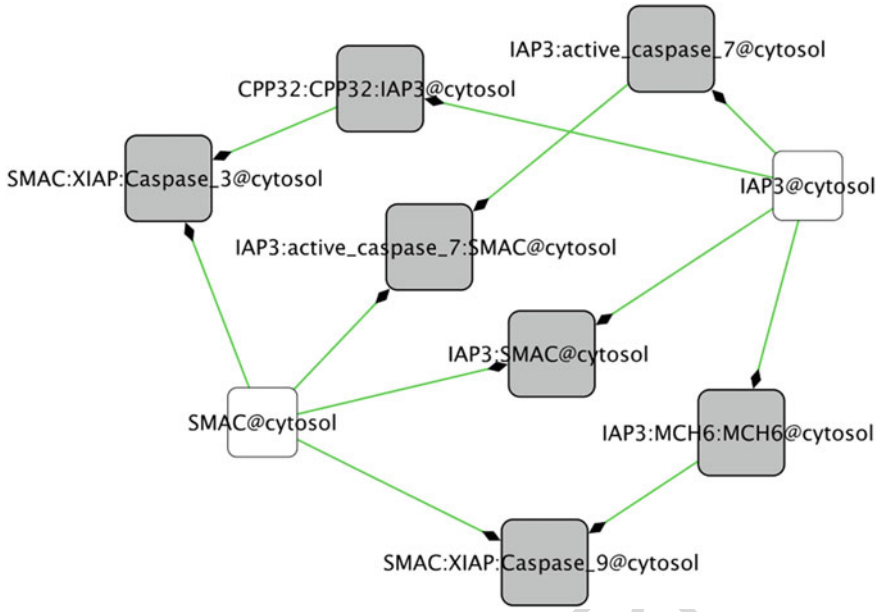

Fig. 10 A network constructed from a BioPAX query, centered on the SMAC protein, and including all protein complexes where this protein is involved. Data extracted from the Apoptosis data of the Reactome database

option "Add complexes" and "expand." Un-select the boxes "Add chemical species," "Add reactions," and "Add publications." Verify that the option "All nodes" is checked for the "Input" section and that "Output in the current network" is checked for the "Output" section (by checking those options, we make sure that all the nodes are by default selected as input and that the result of the query will be added to the current network). Click OK.

9. Several nodes and edges have been added to the current network. For a better visualization, adjust the layout with the function "Layout > yFiles > Organic" from the menu. A green arrow with a diamond ending represents the inclusion of one protein in a complex form. The resulting network should have 9 nodes and 11 edges (Fig. 10).

As we have seen from the standard query interface, it is possible to expand the network by including the chemical species, the reactions connecting all present species that have a common reaction, and the publications related to any of the components of the network (for more information on how to use those options, please consult the BiNoM manual available from our website). Note that the resulting network of interest can be exported as a SBML or BioPAX file as described in the previous paragraphs.

3.4 Other Useful BiNoM Functions

We have seen that BiNoM has several useful functions to extract relevant information from large-scale databases encoded with standards defined by the systems biology community. Very often, the results of those analyses will be one or more subnetworks of interest, possibly grouping a set of molecules involved in a particular biological function (cell death, cell cycle, apoptosis, etc.). An example of such an insightful extraction of a subnetwork is shown in Calzone et al. [12], where a compact modular view of the RB/E2F pathway composed of 16 protein modules and 8 E2F target gene modules (*see* Fig. 3 of this chapter) was extracted from a comprehensive network of hundreds of different molecules and interactions. Once the map is constructed, several options are possible to generate interesting and useful insights. These options include (non-exhaustive list) (1) the creation of a computational predictive model, making possible the analysis of the consequences of deletion or mutation of various elements of the network, and (2) superimposing external and experimental available data related to the function of the network, in order to visually appreciate the effect of different states/perturbations/disease effects. For example, bladder tumor expression data was superimposed on the RB/E2F pathway compact representation mentioned above, for both invasive and noninvasive cases (see http://bioinfo-out.curie.fr/projects/rbpathway/case_study.html). The nodes of the network are then colored according to the averaged expression levels of the different modules, indicating what parts are over- or under-expressed. Clear differences can be seen between the invasive and noninvasive state of the tumor samples, informing of the evolution of tumors at the expression level of genes of the network.

Let us now see an example of how to color a map using BiNoM functions, based on the M-phase network.

1. Import the CellDesigner file M-Phase.xml as described in the Subheading 3.1.
2. Now import values for each gene. They are stored in a simple text file having two columns "NODE_NAME" and "CONCENTRATION." Select the function "File > Import > Attribute from table (Text / MS Excel)." In this case, the values represent expression levels randomly generated, but they could be any type of scoring. Note that for experimental data, proteins with posttranslational modifications will not be colored.
3. Select the input file "M-Phase-Expression.txt." A preview of the file content should appear at the bottom of the dialog box.
4. In the "Advanced" box, click the box "Show text file import options."
5. A new box appears, entitled "Attribute names." Check the box "Transfer first line as attribute name." Now the column titles should read "NODE_NAME" and "CONCENTRATION."

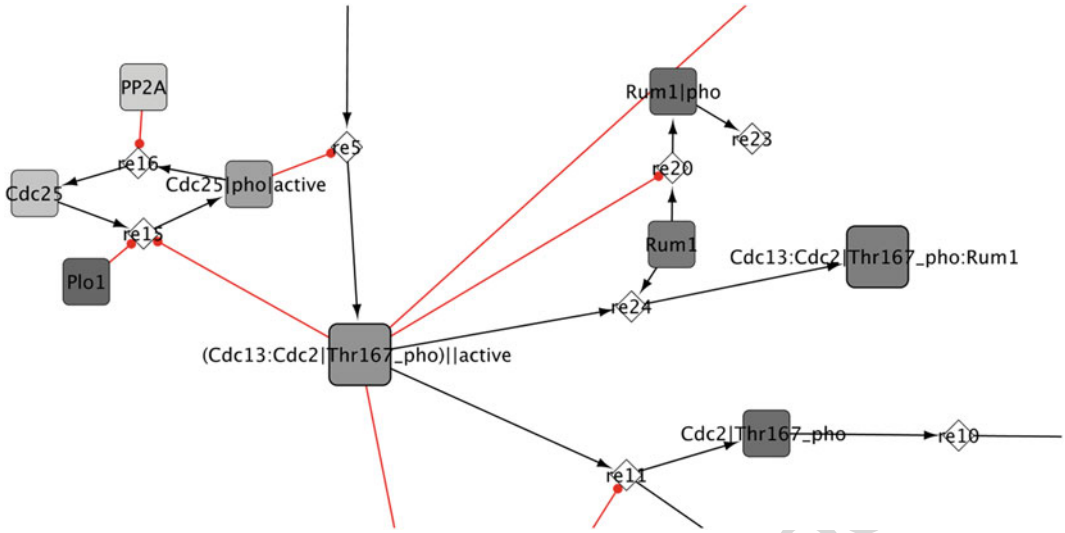

Fig. 11 M-Phase network (zoom) with nodes in shades of gray according to their expression values, ranging from low values (*light gray*) to high values (*dark gray*)

6. Click on "Import." The file is imported, and a new numerical attribute "CONCENTRATION" is created for all the genes.
7. Now click on the "VizMapper" tab, on the left-hand panel of Cytoscape.
8. In the "Visual Mapping Browser" box, click on the "Node Color" small triangle to display the properties.
9. Change the property "Mapping Type" to "Continuous Mapping."
10. Change the value of "Node Color" to "CONCENTRATION."
11. Click on "Graphical View," a new dialog box will appear. Set the minimal and maximal values according to the values of your dataset by clicking on the "Min/Max" button.
12. Set the colors by clicking on the small triangles located above the minimum and maximum values. Click OK.
13. The nodes of the network should be colored according to their expression value, as shown on the Fig. 11.

4 Conclusion

- Model building in systems (and mathematical) biology is a complex multistep process: from the definition of a suitable biological problem, knowledge is first collected and formalized into a network and then translated in mathematical terms.

BiNoM helps with intermediate steps of this process, in the construction of a network of biochemical or regulatory interactions, and in the analysis of the structural properties of this network. For this, BiNoM provides multiple ways: to access pathway databases through their BioPAX representations, manipulate (cut, decompose, reorganize) the network, apply algorithms from graph theory to the network, and map available quantitative data on it.

- The future developments of BiNoM will include functions such as merging several independent networks, finding minimal intervention sets to disrupt or modify the signaling flow from a set of source nodes to a set of target nodes, and the ability to generate a code for web-based representations of biological networks using the Google Map API and semantic zoom.
- BiNoM is not supposed to be a modeling software per se; it does not aim at implementing any engines for numerical simulations, but it has interfaces with external simulators through exporting networks to SBML and GINsim file formats (with use of GINsim Cytoscape plugin [14]). The main application of BiNoM is to facilitate the preparation phase of constructing, annotating, and structuring a biological network for further mathematical modeling and simulation, and this will determine its future development.

5 Notes

- Cycle decomposition can result in a huge number of cycles. It is advised to use it on small to moderate size networks.
- When trying to divide a large network into subnetworks, an alternative to the cycle decomposition described in the Subheading 3.2 is the function "Get Material Components" from the menu "Plugins > BiNoM 2.0 > BiNoM Analysis." This function is using node name semantics to isolate subnetworks in which each protein is involved.

Acknowledgements

EB, LC, DR, GS, EmB, and AZ are members of the team "Computational Systems Biology of Cancer," Equipe labellisée par la Ligue Nationale Contre le Cancer.

References

1. Brazma A, Krestyaninova M, Sarkans U (2006) Standards for systems biology. Nat Rev Genet 7(8):593–605. doi:nrg1922[pii]10.1038/nrg1922
2. Klipp E, Liebermeister W, Helbig A, Kowald A, Schaber J (2007) Systems biology standards–the community speaks. Nat Biotechnol 25(4):390–391. doi:nbt0407-390[pii]10.1038/nbt0407-390
3. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novere N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 19(4):524–531
4. Demir E, Cary MP, Paley S, Fukuda K, Lemer C, Vastrik I, Wu G, D'Eustachio P, Schaefer C, Luciano J, Schacherer F, Martinez-Flores I, Hu Z, Jimenez-Jacinto V, Joshi-Tope G, Kandasamy K, Lopez-Fuentes AC, Mi H, Pichler E, Rodchenkov I, Splendiani A, Tkachev S, Zucker J, Gopinath G, Rajasimha H, Ramakrishnan R, Shah I, Syed M, Anwar N, Babur O, Blinov M, Brauner E, Corwin D, Donaldson S, Gibbons F, Goldberg R, Hornbeck P, Luna A, Murray-Rust P, Neumann E, Reubenacker O, Samwald M, van Iersel M, Wimalaratne S, Allen K, Braun B, Whirl-Carrillo M, Cheung KH, Dahlquist K, Finney A, Gillespie M, Glass E, Gong L, Haw R, Honig M, Hubaut O, Kane D, Krupa S, Kutmon M, Leonard J, Marks D, Merberg D, Petri V, Pico A, Ravenscroft D, Ren L, Shah N, Sunshine M, Tang R, Whaley R, Letovksy S, Buetow KH, Rzhetsky A, Schachter V, Sobral BS, Dogrusoz U, McWeeney S, Aladjem M, Birney E, Collado-Vides J, Goto S, Hucka M, Le Novere N, Maltsev N, Pandey A, Thomas P, Wingender E, Karp PD, Sander C, Bader GD (2010) The BioPAX community standard for pathway data sharing. Nat Biotechnol 28(9):935–942. doi:nbt.1666[pii]10.1038/nbt.1666
5. Le Novere N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, Demir E, Wegner K, Aladjem MI, Wimalaratne SM, Bergman FT, Gauges R, Ghazal P, Kawaji H, Li L, Matsuoka Y, Villeger A, Boyd SE, Calzone L, Courtot M, Dogrusoz U, Freeman TC, Funahashi A, Ghosh S, Jouraku A, Kim S, Kolpakov F, Luna A, Sahle S, Schmidt E, Watterson S, Wu G, Goryanin I, Kell DB, Sander C, Sauro H, Snoep JL, Kohn K, Kitano H (2009) The systems biology graphical notation. Nat Biotechnol 27(8):735–741. doi:nbt.1558[pii]10.1038/nbt.1558
6. Joshi-Tope G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de Bono B, Jassal B, Gopinath GR, Wu GR, Matthews L, Lewis S, Birney E, Stein L (2005) Reactome: a knowledgebase of biological pathways. Nucleic Acids Res 33 (Database issue):D428–D432. doi:33/suppl_1/D428[pii]10.1093/nar/gki072
7. Le Novere N, Bornstein B, Broicher A, Courtot M, Donizelli M, Dharuri H, Li L, Sauro H, Schilstra M, Shapiro B, Snoep JL, Hucka M (2006) BioModels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. Nucleic Acids Res 34(Database issue): D689–D691. doi:34/suppl_1/D689[pii]10.1093/nar/gkj092
8. Licata L, Briganti L, Peluso D, Perfetto L, Iannuccelli M, Galeota E, Sacco F, Palma A, Nardozza AP, Santonico E, Castagnoli L, Cesareni G (2012) MINT, the molecular interaction database: 2012 update. Nucleic Acids Res 40(Database issue):D857–D861. doi:gkr930[pii]10.1093/nar/gkr930
9. Funahashi A, Morohashi M, Kitano H (2003) Cell Designer: a process diagram editor for gene-regulatory and biochemical networks. Biosilico 1(5):159–162
10. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B, Hanspers K, Isserlin R, Kelley R, Killcoyne S, Lotia S, Maere S, Morris J, Ono K, Pavlovic V, Pico AR, Vailaya A, Wang PL, Adler A, Conklin BR, Hood L, Kuiper M, Sander C, Schmulevich I, Schwikowski B, Warner GJ, Ideker T, Bader GD (2007) Integration of biological networks and gene expression data using Cytoscape. Nat Protoc 2(10):2366–2382. doi:nprot.2007.324[pii]10.1038/nprot.2007.324

11. Zinovyev A, Viara E, Calzone L, Barillot E (2008) BiNoM: a Cytoscape plugin for manipulating and analyzing biological networks. Bioinformatics 24(6):876–877. doi:btm553[pii]10.1093/bioinformatics/btm553

12. Calzone L, Gelay A, Zinovyev A, Radvanyi F, Barillot E (2008) A comprehensive modular map of molecular interactions in RB/E2F pathway. Mol Syst Biol 4:173. doi:msb20087[pii]10.1038/msb.2008.7

13. Novak B, Csikasz-Nagy A, Gyorffy B, Nasmyth K, Tyson JJ (1998) Model scenarios for evolution of the eukaryotic cell cycle. Philos Trans R Soc Lond B Biol Sci 353(1378):2063–2076. doi:10.1098/rstb.1998.0352

14. Gonzalez AG, Naldi A, Sanchez L, Thieffry D, Chaouiya C (2006) GINsim: a software suite for the qualitative modelling, simulation and analysis of regulatory networks. Biosystems 84 (2):91–100. doi:S0303-2647(05)00169-3[pii]10.1016/j.biosystems.2005.10.003

Author Queries

Chapter No.: 7 271122_1_En

Query Refs.	Details Required	Author's response
AU1	Please check if "his" should be changed to "his or her" for gender neutrality.	
AU2	Figs. 6 and 7 have same legends. Please check.	