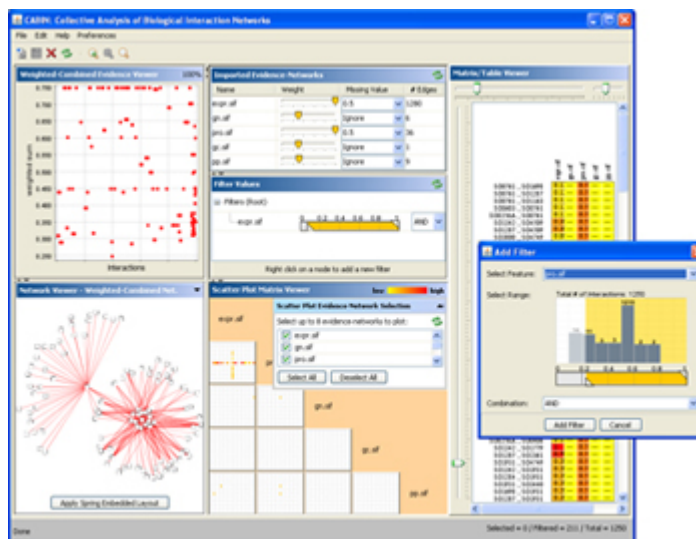


Collective Analysis of Biological Interaction Networks (CABIN) User Manual for CABIN Version 2.1

Kelly Domico
Mudita Singhal

June 2007



Prepared for the U.S. Department of Energy
Under Contract DE-AC05-76RL01830



Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

**Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov**

**Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161
ph: (800) 553-6847 fax: (703) 605-6900
email: orders@ntis.fedworld.gov
online ordering: <http://www.ntis.gov/ordering.htm>**

Collective Analysis of Biological Interaction
Networks (CABIN)
User Manual
for CABIN Version 2.1

Kelly Domico
Mudita Singhal

June 2007

**Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830**

**Pacific Northwest National Laboratory
Richland, Washington 99352**

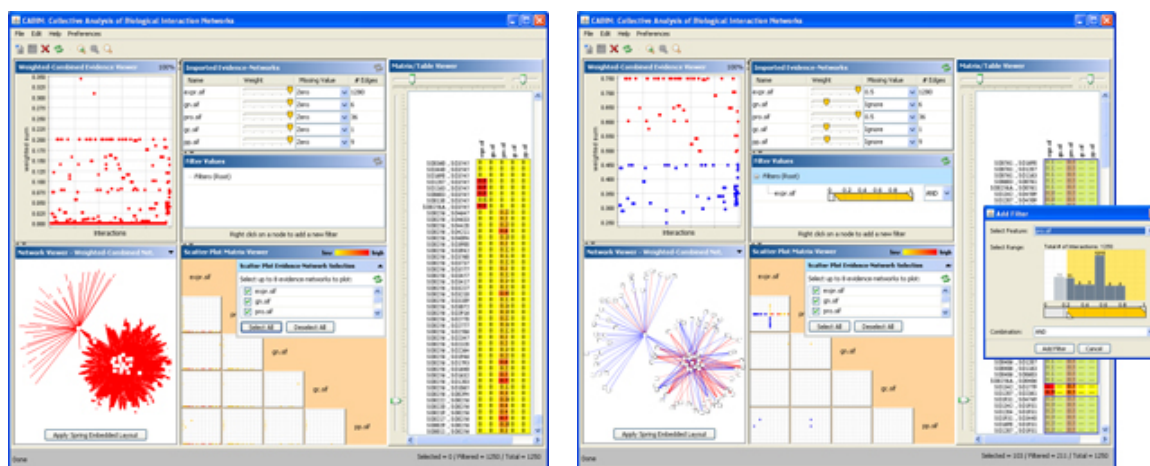
Contents

1. Introduction	1
1.1. License	1
1.1.1. CABIN Sample Data Release Policy	2
1.2. Technologies Used	2
2. Getting Started.....	3
2.1. Installing CABIN	3
2.2. Launching CABIN	3
2.3. Reporting Bugs.....	4
2.4. Functionality Requests	4
3. Features of CABIN.....	4
4. CABIN Browser.....	4
5. Importing Networks in CABIN.....	6
5.1. Importing an Evidence-Network in CABIN	6
5.1.1. Assigning Reliability Weights	8
5.1.2. Assigning Missing Values	8
5.1.3. Normalizing Values	9
5.2. Importing All Cytoscape Networks in CABIN	9
6. Getting Evidence from Prolinks.....	10
7. Multiple Coordinated Views.....	11
7.1. Scatter Plot Matrix Viewer.....	11
7.2. Weighted Combined Evidence Viewer	12
7.3. Matrix/Table Viewer	13
7.4. Network Viewer	14
8. Filters	15
9. Find/Search	17
10. Creating Sub-Networks.....	17
11. Saving Results	18
12. Acknowledgments	19
13. Citing CABIN.....	19
14. References.....	20
15. Appendix A.....	20

1. Introduction

The importance of understanding interaction networks has fueled the development of numerous interaction data generation techniques, databases as well as prediction tools facilitating the generation of more interaction data. However not all prediction tools and databases are a hundred percent accurate. Generation of networks with high confidence interactions formulates the first step towards deciphering unknown protein functions, determining protein complexes as well as drug discovery. The Collective Analysis of Biological Interaction Networks (CABIN) is an exploratory data analysis tool that enables fusion of interactions obtained from multiple sources of evidence, thereby increasing the confidence of computational predictions as well as validating experimental observations. CABIN has been written in JAVA™ and is available as a plug-in for Cytoscape [1] – an open source network visualization tool.

Figure 1.
Collective
Analysis
of
Biological
Interaction
Networks



a) CABIN Browser showing evidence-networks imported with default weight and missing values. Network Viewer shows the combined weighted view with spring-embedded layout applied. Scatter Plot Matrix Viewer shows plots between first five evidence-networks. The Matrix/Table Viewer shows a heat-map representation of the evidence in tabular format.

b) CABIN Browser showing evidence-networks with edited weight, missing values, and a filter already applied. Another filter is being created in the smaller window. Selection in Combined Weighted Evidence Viewer corresponds with selection in the Network Viewer, the Matrix/Table Viewer, and the Scatter Plot Matrix Viewer.

This manual describes the different functionalities in CABIN and demonstrates how it can be used to integrate multiple sources of evidence.

1.1. License

CABIN is designed to run on Windows XP/2000, Linux and Mac OS X operating systems. Please use them at your own risk. Neither PNNL nor the U.S. Department of Energy claim any responsibility for any problems associated with this software or any research done using it.

This computer software was prepared by Battelle Memorial Institute, hereinafter the Contractor, under Contract No. DE-AC05-76RL0 1830 with the Department of Energy

(DOE). All rights in the computer software are reserved by DOE on behalf of the United States Government and the Contractor as provided in the Contract. NEITHER THE GOVERNMENT NOR THE CONTRACTOR MAKES ANY WARRANTY, EXPRESS OR IMPLIED, OR ASSUMES ANY LIABILITY FOR THE USE OF THIS SOFTWARE. This notice including this sentence must appear on any copies of this computer software.

Refer to Section 13 *Citing CABIN* for citation information.

1.1.1. CABIN Sample Data Release Policy

By using the sample protein-protein interaction datasets provided within CABIN, the user agrees to the following:

Pacific Northwest National Laboratory (PNNL) and Oak Ridge National Laboratory (ORNL) is not responsible for errors contained in the protein-protein interaction datasets, or for consequences arising from using the protein-protein interaction datasets. Forthcoming publications from us will detail methods and controls used in the acquisition of data underlying the sample protein-protein interaction datasets.

The following acknowledgement must be included in any publications, presentations, reports, databases, websites, or data analyses that have used the CABIN protein-protein interaction datasets: "Protein-protein interaction data were obtained from the Center for Molecular and Cellular Systems (mippi.ornl.gov) which is sponsored at Oak Ridge National Laboratory by the U.S. Department of Energy Office of Biological and Environmental Research." The following paper needs to be cited in any research facilitated by CABIN: Singhal, M., Domico, K., "CABIN: Collective Analysis of Biological Interaction Networks", *Computat. Biol. Chem.* (2007), doi:10.1016/j.compbiolchem.2007.03.006.

We reserve the right to publish a description of the overall protein-protein interaction datasets, and the methods used to obtain the contents of the datasets.

1.2. Technologies Used

This software has been written in Java. It is platform independent and should work on MAC, Windows and Linux operating systems, although it has not been tested on platforms other than Windows. CABIN uses a number of open source libraries listed below.

Cytoscape	http://www.cytoscape.org
JFreeChart	http://www.jfree.org/jfreechart
JMatrixView	http://jmatrixview.sourceforge.net
COLT	http://dtd.lbl.gov/~hoschek/colt
BiSlider	http://vernier.frederic.free.fr/BiSlider
SwingX	https://swingx.dev.java.net
JGoodies Forms	http://www.jgoodies.com/freeware/forms

2. Getting Started

2.1. Installing CABIN

A link for downloading the CABIN plug-in can be found on the following web page:

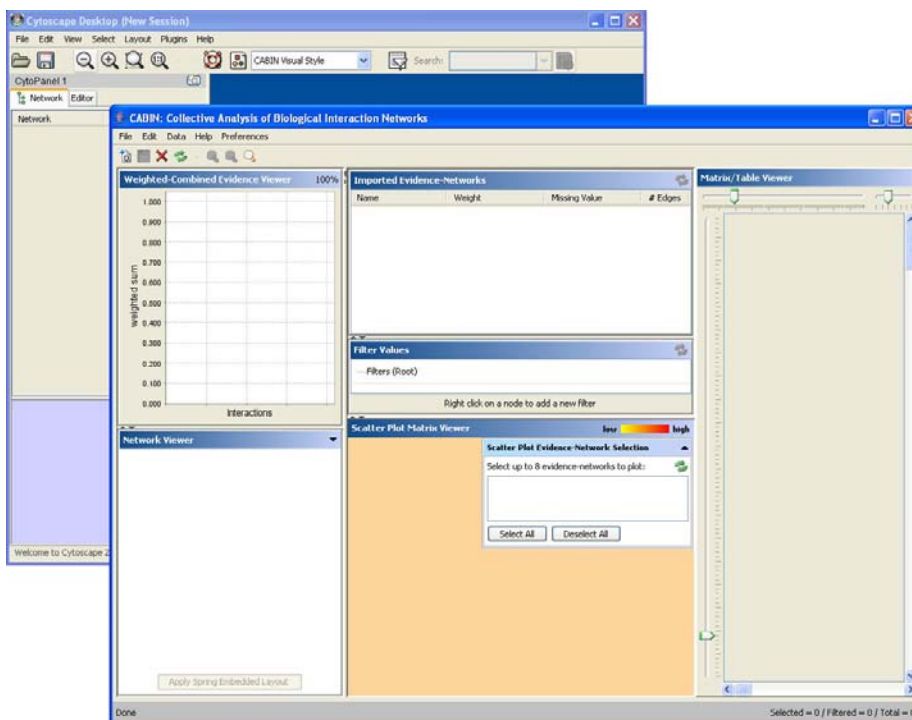
<http://www.sysbio.org/dataresources/cabin.stm>

As CABIN is a plug-in of Cytoscape, you will need to have Cytoscape installed on your local machine or have access to Cytoscape via WebStart. If you don't already have Cytoscape installed, you can download and install Cytoscape 2.4 (or earlier versions) from www.cytoscape.org. Download the CABIN jar file from the link above and place it in the 'plugins' directory of Cytoscape. Please contact cabin@pnl.gov with the subject "Installing CABIN problems", if you face any problems.

2.2. Launching CABIN

Launch the installed Cytoscape application either directly from your machine, or from a webpage using WebStart. If installed correctly, Cytoscape will automatically recognize CABIN as a plug-in and it will be listed under the "Plugins" menu in Cytoscape. To launch CABIN, choose "CABIN" from the "Plugins" menu. You will see the CABIN Browser window appear with empty views since there is no evidence-network imported yet. Consequently, the CABIN Browser will be cleared as well on the load of a new or saved Cytoscape session.

Figure 2.
Launching
CABIN in
Cytoscape



2.3. Reporting Bugs

Please send an email to cabin@pnl.gov to report bugs.

2.4. Functionality Requests

We always like to hear about what new functionality you would like to see in CABIN. Please email us a detailed description of the desired functionality at cabin@pnl.gov with the subject “New Functionality Request.”

3. Features of CABIN

CABIN provides multiple coordinated views that represent the problem at multiple levels of abstraction. A Scatter Plot Matrix Viewer shows the relation amongst the different evidence-networks for the interactions of interest serving to facilitate the weights assignment to each evidence-source. Matrix visualization within CABIN shows a heat map representation of the data values for all the evidence-sources in a tabular format with a functionality to sort by column. A Cytoscape network view shows the selected evidence-source in a node-edge format, providing the conventional visualization for protein-protein interaction networks. A filtering mechanism is provided to select subsets of interactions based on their evidence values shown in a histogram view. Finally a Weighted-Combined Evidence Viewer provides a view of the interactions with confidence based on the average weighted sum of its evidences. All four views have functionality to make selections that are reflected in the other views. Along with these basic views, CABIN also has functionalities to search for interactions, select interactions in a view, create a new network from all or selected interactions, save selected or all interactions to a tab delimited file on the local machine, and export all networks as SIF files to a specified folder on the local machine.

The following are the main tasks that can be performed in CABIN. These are discussed in the subsequent sections:

- Section 5 – **Importing Networks in CABIN** p. 6
- Section 6 – **Getting Evidence from Prolinks** p. 10
- Section 7 – **Multiple Coordinated Views** p. 11
- Section 8 – **Filters** p. 15
- Section 9 – **Find/Search** p. 17
- Section 10 – **Creating Sub-Networks** p. 17
- Section 11 – **Saving Results** p. 18

4. CABIN Browser

The main window you see after launching CABIN from the “Plugins” menu bar is called the CABIN Browser. This browser enables you to import networks from Cytoscape into CABIN as evidence-networks. Starting with a list of proteins of interest multiple evidence-networks can be created using computational and publicly available tools such as Prolinks[2], DIP[3], BIND[4], Literature Evidence, or the BioNetBuilder[5] plugin of

Cytoscape. Section 6 provides a brief description of how such networks can be created based on evidence from the Prolinks database and another Cytoscape plug-in called BioNetBuilder. It also provides a list of other resources that can be utilized to collect evidence for interactions.

The menu bar in the CABIN browser provides options for managing and/or using the displayed evidence-networks:

File menu:

- New Evidence-Network... – Opens a dialog box for adding a new evidence-network to CABIN (as described in Section 5.1 *Importing an Evidence-Network in CABIN*)
- Import All Evidence-Networks – Imports all networks within Cytoscape into CABIN (as described in Section 5.2 *Importing All Cytoscape Networks in CABIN*)
- Export All Evidence-Networks – Exports all networks within CABIN as SIF files into a specified directory
- Load Sample Networks – Imports a set of sample networks into CABIN (see Appendix A for details of the sample data)

Edit Menu:

- Find Interaction... – Provides a dialog facilitating searching for an interaction or all interactions involving a protein of interest (as described in Section 9 *Find/Search*)
- Delete Evidence-Network... – Opens a dialog box for deleting an imported evidence-network from CABIN
- Clear CABIN Browser – Removes all evidence-networks from CABIN

Data Menu:

- Get Evidence from Prolinks... - Opens a dialog box for retrieving data from the Prolinks database (as described in Section 6 *Getting Evidence From Prolinks*)



Currently, the *Getting Evidence from Prolinks* functionality is only available for internal users (please send an email to cabin@pnl.gov to obtain an internal version of the jar file. External users can utilize the BioNetBuilder plug-in of Cytoscape [<http://err.bio.nyu.edu/cytoscape/bionetbuilder>] which provides similar functionality to build interaction networks from the Prolinks evidence.

Help Menu:

- About CABIN... - Provides information about and the terms of use of CABIN
- Data Release Policy – Provides the sample data release policy of CABIN
- CABIN User Guide... - Opens the user guide for CABIN

Preferences Menu:

- Edit Preferences... - Opens a dialog box for editing the viewing preferences for the multiple viewers available in CABIN

5. Importing Networks in CABIN

A Cytoscape network is represented as an evidence-network in CABIN. There are two ways to add new evidence-networks to CABIN:

5.1. Importing an Evidence-Network in CABIN

For a network to be imported into CABIN, it should first be imported into Cytoscape by creating a SIF file for the network (refer to the Cytoscape documentation to learn more about the SIF format). Then select the “New Evidence-Network” menu item from the “File” menu. It will bring up a window (as shown in Figure 3) to fill in values for some basic parameters for the evidence-network. Provide a name to the network and select a Cytoscape network by clicking on the “Cyto” button. A Cytoscape networks browser window will appear, listing all networks available in Cytoscape for import. Choose the network you want from the list and click the “Select Network” button when finished. Alternatively, you can double-click on the network name to select and close. At any point, new networks can be imported into Cytoscape and they will be reflected in the import dialog of CABIN.

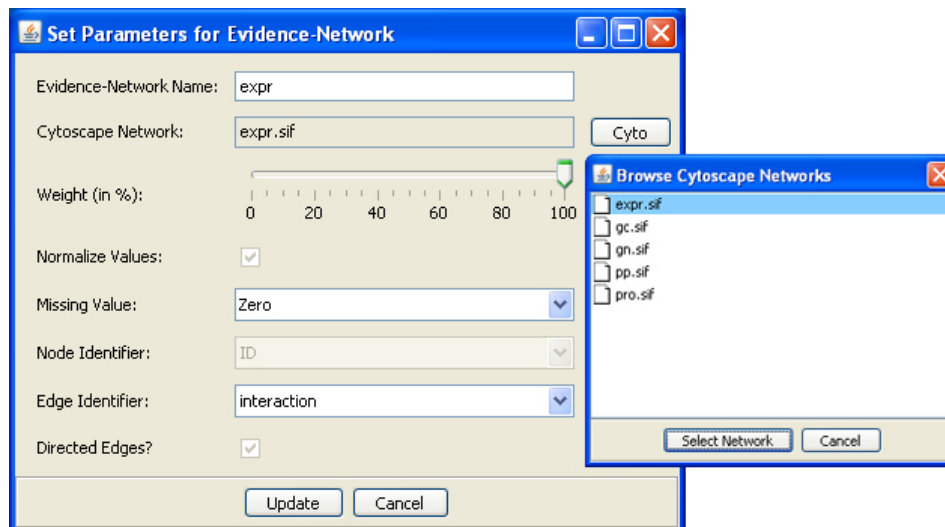
Assign a reliability score/weight to the evidence-source by selecting a value in the weights slider. The “Normalize Values” check box is disabled for now since all networks need to be normalized to the same score range (0-1 in this case) to be able to do a fair comparison. Assign a missing value to the network for interaction edges that either exist in some other evidence-network and not this one or do not have confidence values assigned to it. Section 5.1.2 discusses the effect of selecting different options for missing values in more detail.

You can also associate any edge attributes as the edge identifier (confidence value) for this evidence-network. Select the edge attribute you wish to assign using the “Edge Identifier” combobox. Click the “Update” button when you have finished assigning all the parameters.



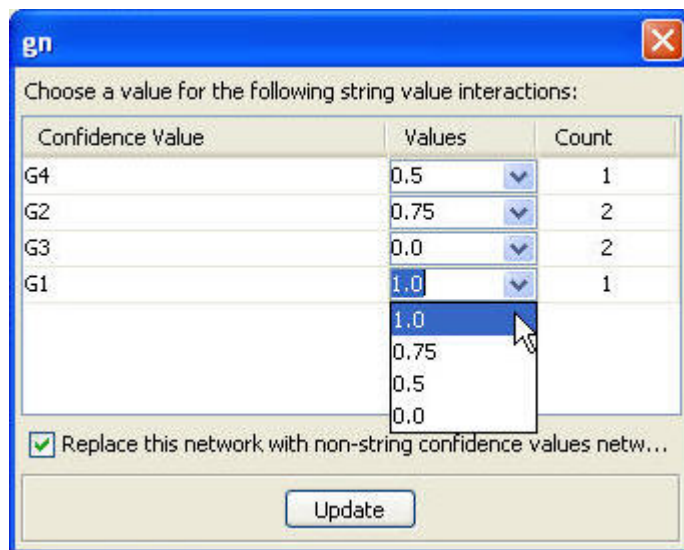
In a future release, options will be provided to suggest values for the reliability weights and the missing values calculated based on statistics of the interaction values of the imported network.

Figure 3.
Setting
Parameters
for an
evidence-
network



Any non-number confidence value (excluding missing values) assigned to an interaction will be noted and shown in a dialog at the completion of each network import, as shown in Figure 4. The first column displays the non-number confidence value, the second column allows you to assign a valid confidence value, and the third column displays a count of that confidence value. Select a valid confidence value for each non-number confidence value. Check “Replace this network with non-string confidence values network in Cytoscape” to replace with the chosen valid confidence values and click “Update” when finished.

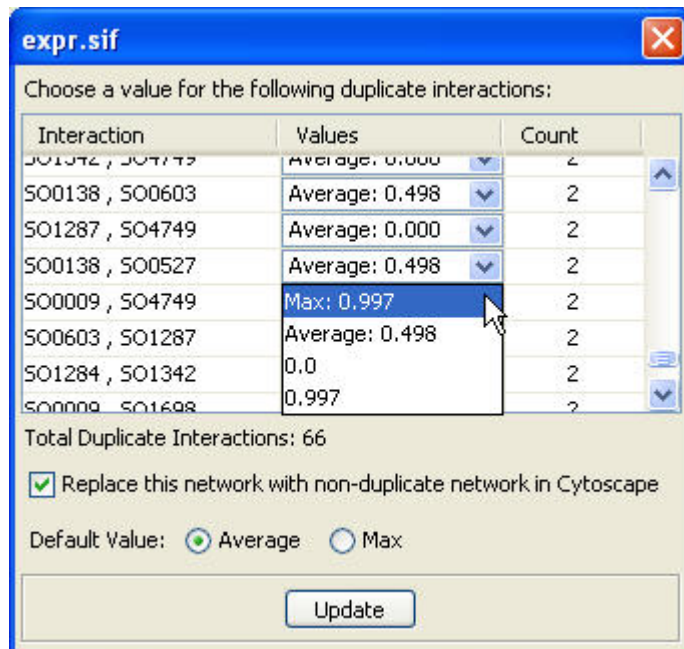
Figure 4.
Non-Number
Confidence
Value Dialog



Since multiple instances of the same interaction are not allowed in CABIN, any duplicate interactions in the imported network will also be collected and shown in a dialog (see Figure 5) at the completion of each network import. The first column shows the interaction containing duplicates, the second column allows you to choose the value to assign that interaction, and the third column displays the number of duplicates for that interaction. Select the “Replace this network with non-duplicate network in Cytoscape” checkbox if you wish to remove all duplicate interactions

from the network in Cytoscape and replace the confidence value with your chosen value for each interaction. For convenience, select the “Average” or “Max” radio button to select the average value or maximum value for all duplicate interactions, respectively. Click “Update” when you have finished assigning values to all interactions.

Figure 5.
Duplicate
Interaction
Dialog



5.1.1. Assigning Reliability Weights

The reliability weight is a value based on how much confidence you place in the prediction of interactions by the data source. The weights are represented as a percentage for each evidence-network. It is used in CABIN for the two combined weighted views, Weighted-Combined Evidence Viewer and Network Viewer, to weigh the values in each evidence-network appropriately prior to combining. You can edit the weight for each evidence-network in the *Imported Evidence-Networks Panel* by simply moving the slider position of the corresponding value. The “Update” (🔄) button needs to be clicked to update the weights through the CABIN Browser.

5.1.2. Assigning Missing Values


In case all evidence sources do not give confidence values for all interactions, we have the option to assign a default value as a missing value for each evidence-network. There are four options to choose from: Ignore, Median, Zero, or 0.5. Alternatively, you can enter a number value between 0 and 1 as a fifth option. These options are described below:

- Ignore – ignores that particular evidence for the interactions with missing values in the weighted combined views

- Median – replaces the missing value of interactions in that evidence-network with the median score of that particular evidence-network
- Zero – replaces the missing values of interactions in that evidence-network with a value of 0
- 0.5 – replaces the missing values of interactions in that evidence-network with the midpoint of the range (all values are normalized on a scale of 0-1)
- User entered value – replaces the missing values of interactions in that evidence-network with the user entered value. This value needs to be between 0-1 since all values are normalized to a scale of 0-1



Use this option with caution since it can bias the results.

To edit the missing value at anytime in the *Imported Evidence-Networks Panel*, simply select a new missing value from the list or enter a new valid value between 0 and 1. Click the “Update” () button to update the new missing value throughout the CABIN Browser.

5.1.3. Normalizing Values

If you choose to normalize the values, all interaction values are normalized to within the range of 0 and 1. Interaction values already within the correct range are not normalized. The following normalizing function is used:

$$(\text{current_value} - \text{min_value}) / (\text{max_value} / \text{min_value})$$

Currently, the option to NOT normalize the values is not provided. All values are normalized to a scale of 0-1 by default.



In a future release, advanced statistics will be used to normalize discrete and continuous data appropriately.

5.2. Importing All Cytoscape Networks in CABIN

Multiple networks available within Cytoscape can be imported into CABIN with default values for the reliability weight and missing value. Once the networks are imported, you can edit the values for the weights and missing value option for each of the imported evidence-networks.












To use this option, select the “Import All Evidence-Networks” menu item from the “File” menu and the “Set Evidence-Network Parameters” dialog will appear. Assign the default weight and missing value and click “Finished” when you have finished assigning all parameters. This will automatically add all networks from within Cytoscape into CABIN as evidence-networks with the following default parameters:

Evidence-Network Name – *the name of the Cytoscape network file*
 Weight (in %) – *the weight you chose*
 Missing Value – *the missing value you chose*

As in the case of importing networks individually, any duplicate interactions or interactions with non-number confidence values will be collected and displayed in the respective dialogs.

The *Imported Evidence-Networks Panel* (see Figure 6) shows a list of all evidence-networks currently added to the CABIN Browser. From here, you can edit each evidence-network's name, reliability weight, and missing value. To edit the evidence-network's name, simply double-click on the name, edit the old name, and press the "Enter" key on the keyboard when finished. The weight assignment for each evidence-network can be edited by simply moving the slider position of the corresponding evidence-network in the *Imported Evidence-Networks Panel*. The weights are not updated in all the views until you click the "Update" (↻) button. Similarly you can edit the missing value anytime by simply selecting a new missing value from the list or entering a new valid value between 0 and 1. Click the "Update" (↻) button to update the new missing value throughout the CABIN Browser.

Figure 6.
Imported
Evidence-
Networks
Panel

Imported Evidence-Networks 			
Name	Weight	Missing Value	# Edges
gc		Ignore 	1
gn		Ignore 	6
expr		0.5 	1193
pp		Ignore 	9
pro		0.5 	36

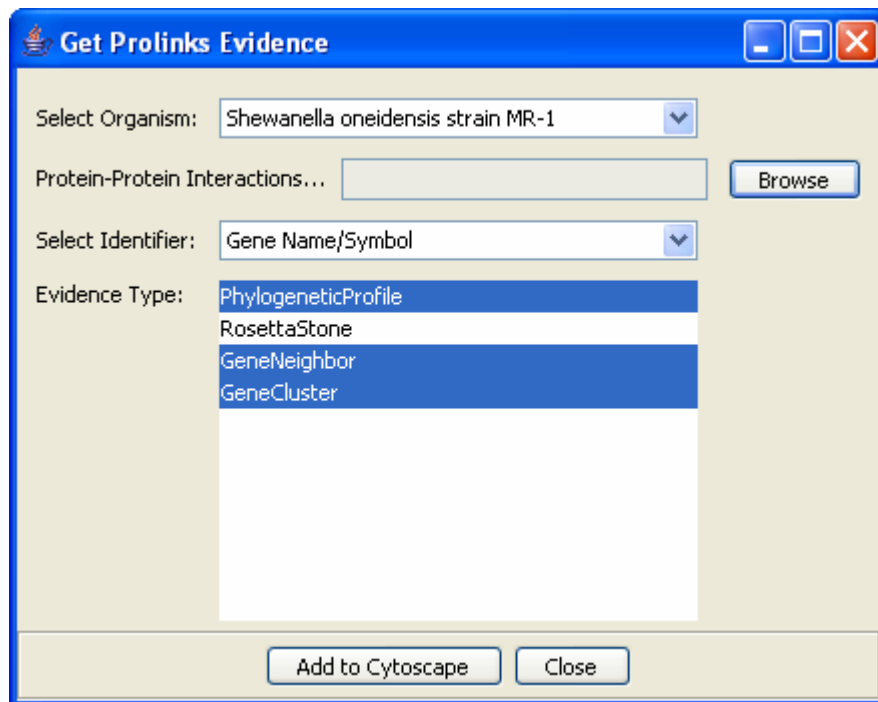


CABIN also provides the functionality to retrieve evidence from the Prolinks database from a list of protein-protein interactions.

6. Getting Evidence from Prolinks

Select the "Getting Evidence from Prolinks" menu item from the "Data" menu. Select an organism, protein-protein interactions file and the identifier type. Select one or more evidence types by holding the "Ctrl" key and selecting the evidence types from the list. Click "Add to Cytoscape" when finished; this will add all networks retrieved from Prolinks for the selected evidence types to Cytoscape. You can then import those networks into the CABIN Browser for further analysis.

Figure 7.
Get
Evidence
from
Prolinks



You can also use BioNetBuilder to acquire evidence from many different databases including DIP, BIND, Prolinks, KEGG, HPRD, The BioGrid, and GO. BioNetBuilder is a Cytoscape plug-in that provides you with several options to specify which interactions to retrieve.

7. Multiple Coordinated Views

7.1. Scatter Plot Matrix Viewer

The Scatter Plot Matrix view shows a matrix of scatter plots of the evidence-networks chosen in the *Scatter Plot Feature Selection Panel*. Each evidence-network is plotted against the other selected evidence-networks, showing a matrix with multiple scatter plots reflecting the relationship of all evidence-networks with all the others. A color gradient is used as an indicator of density with yellow corresponding to lower density and red corresponding to higher density. An example of a scatter plot matrix between five selected evidence-networks is shown in Figure 8.

Figure 8.
Scatter Plot
Matrix of
evidence-
networks



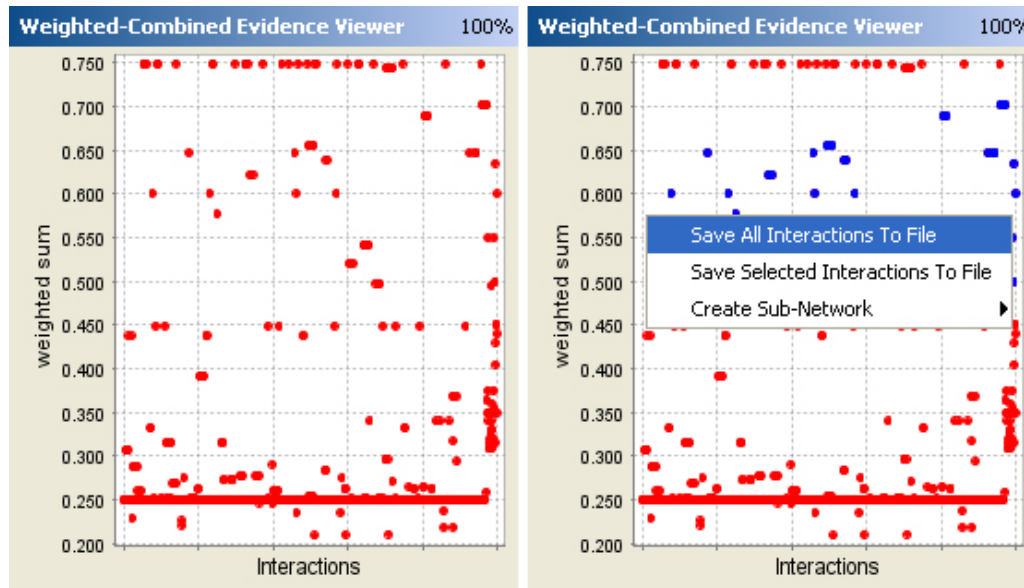
Select the evidence-networks you want plotted by clicking on the check boxes in the *Scatter Plot Feature Selection Panel* and click the “Update” (🔄) button when finished. Alternatively, you can click “Select All” or “Deselect All” to select or deselect all evidence-networks (or the first eight if there are more than eight evidence-networks) in the list, respectively. A minimum of two and a maximum of eight evidence-networks can be selected for plotting at a time. The status bar shows an estimated completion time and the progress of the operation. Click on any scatter plot for a more detailed view of the plot. Selecting points in the detailed view will update the selection in the Weighted Combined Evidence Viewer, Network Viewer, Matrix/Table Viewer and the Scatter Plot Matrix Viewer of evidence-networks.

By default, the values in the Scatter Plot Matrix View are filtered. To disable the filter, select “Edit Preferences” from the Preferences menu. Check the “Apply filter” checkbox and click the “Update” (🔄) button to refresh the view. You can also choose to display the R-squared value of each plot from this preferences menu. The R-Squared is calculated using the Pearson Product Moment Correlation.

7.2. Weighted Combined Evidence Viewer

Each point in the Weighted Combined Evidence Viewer represents an interaction. This view shows a scatter plot of the interactions with confidences based on the weighted sum of all its evidences. By placing your mouse pointer over a point, you can see the interaction and its weighted sum. Multiple interactions can be selected by clicking and dragging over a rectangular area in the view as shown in Figure 9. The selected area is depicted by the red, outlined rectangular box. On release of the mouse button, the selected points are represented in blue color. Selecting points in this view will update the selection in the Scatter Plot Matrix Viewer, Network Viewer and the Matrix/Table Viewer as well.

Figure 9.
Combined
Weighted
Evidence
View



Selected or all interactions in this view can be saved to a file on your local machine by right-clicking anywhere within the view (as described in Section 11 *Saving Results*). A new evidence-network of the selected points or all points can also be created from this menu. This evidence-network is imported automatically into CABIN.

Click on the “Zoom In” (🔍) or “Zoom Out” (🔍) buttons to get a closer look at the points in the view. To zoom only to a specified set of interactions, select the interactions in the view and click on the zoom in button. To return to the full view, click the “Zoom 100%” (🔍) button.



Future releases will provide options to choose advanced algorithms to combine these multiple sources of evidence utilizing the statistical reliability of the evidence-network, missing values and continuous or discrete nature of the values.

7.3. Matrix/Table Viewer

The Matrix/Table Viewer is a heat map representation of the data values for all the evidence sources in a tabular format. Additionally, the last column will show the weighted-combined confidence value for each interaction. To expand the size of the columns, use the horizontal slider at the top left of the view. To expand the size of the rows, use the vertical slider to the left of the view. If the labels are not visible in their entirety then the slider at the top right can be used to move the entire matrix diagonally from the top left corner to the bottom right corner of the view.

Figure 10.
Matrix/Table
Viewer

The screenshot shows the Matrix/Table Viewer window. The table contains interaction data with columns for different evidence types and a 'Weighted Combined' column. A context menu is open over the 'Weighted Combined' column, showing options for sorting and saving data.

	pro.sif	expr.sif	gc.sif	pp.sif	gn.sif	Weighted Combined
\$00256, \$04647	0.23	0.5	0.5	---	---	0.35
\$00256, \$04633	0.2	0.5	0.5	---	---	0.34
\$00256, \$04428	0.25	0.5	0.5	---	---	0.35
\$00256, \$04211	0.6	0.5	0.5	---	---	0.4
\$00256, \$04094	0.18	0.5	0.5	---	---	0.34
\$00256, \$03988	0.2	0.5	0.5	---	---	0.34
\$00256, \$03942	0.12	0.5	0.5	---	---	0.33
\$00256, \$03760	0.12	0.5	0.5	---	---	0.33
\$00256, \$03737	0.12	0.5	0.5	---	---	0.33
\$00256, \$03577	0.2	0.5	0.5	---	---	0.34
\$00256, \$03457	0.13	0.5	0.5	---	---	0.33
\$00256, \$03417	0					
\$00256, \$03227	0					
\$00256, \$03210	0					
\$00256, \$03189	0					
\$00256, \$03072	0					
\$00256, \$02916	0					
\$00256, \$02778	0					
\$00256, \$02777	0					
\$00256, \$02586	0					
\$00256, \$02347	0					
\$00256, \$02328	0.2	0.5	0.5	---	---	0.34

Context Menu Options:

- Sort Column (Ascending, Descending)
- Save All Interactions To File
- Save Selected Interactions To File
- Create Sub-Network
- Turn Filter On (checked)

You can sort the values of each column by right-clicking anywhere within the column to be sorted. Choose either “Ascending” or “Descending” from the popup menu. Once sorted, the order of the rows will be rearranged accordingly. The ‘---’ symbol in a cell represents a missing value of “Ignore”.

Click on the corresponding row to select it. To select a continuous block of rows, click on the first row that will represent the top of the selection block. Then, while holding down the “Shift” key on the keyboard, click the last row that will represent the bottom of the selection block. To select any number of discontinuous rows, hold down the “Ctrl” key on the keyboard while click on the desired rows. Selection in this view is also coordinated with selection in the other views. Similarly, right-clicking anywhere in the view provides options to save selected or all interactions to a file (as described in Section 11 *Saving Results*) and to create a new evidence-network from the selected interactions or all interactions.

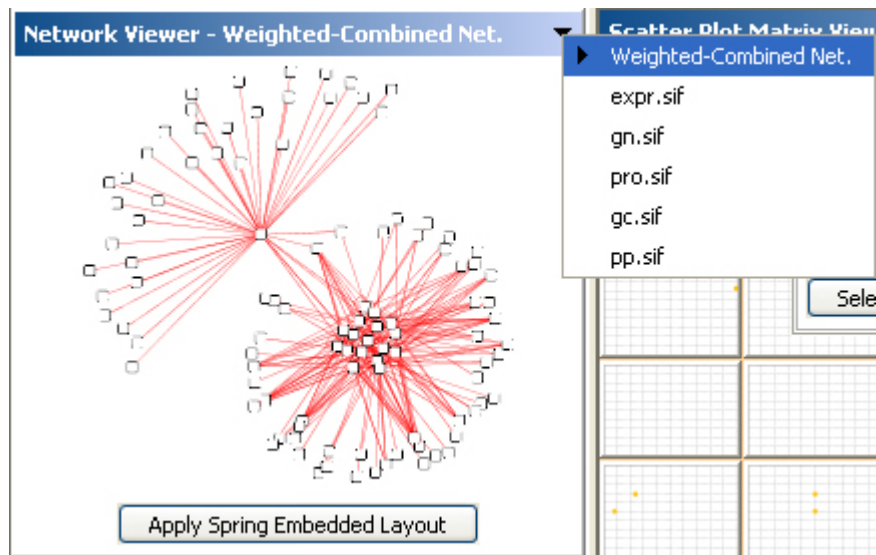
By default, the values in the Matrix/Table Viewer are filtered. To disable the filter, right-click anywhere in the view and select “Turn Filter On”. The filter is enabled if a checkmark is visible in the right-click menu.

7.4. Network Viewer

The Network Viewer shows the selected evidence-network in a node-edge format. A combined weighted view is also included; this view is essentially a graph/network representation of the Weighted Combined Evidence Viewer described in Section 7.2 *Weighted Combined Evidence Viewer*.

To switch between the views of each evidence-network, simply click on the down arrow icon (▼) to the right of the title. This will bring up a list of all evidence-network views as well as the combined weighted view. The right arrow icon (▶) in the list signifies which view is currently showing. Once an evidence-network view is selected from the drop-down list, a layout can be applied by clicking on the “Apply Spring Embedded Layout” button.

Figure 11.
Cytoscape
View

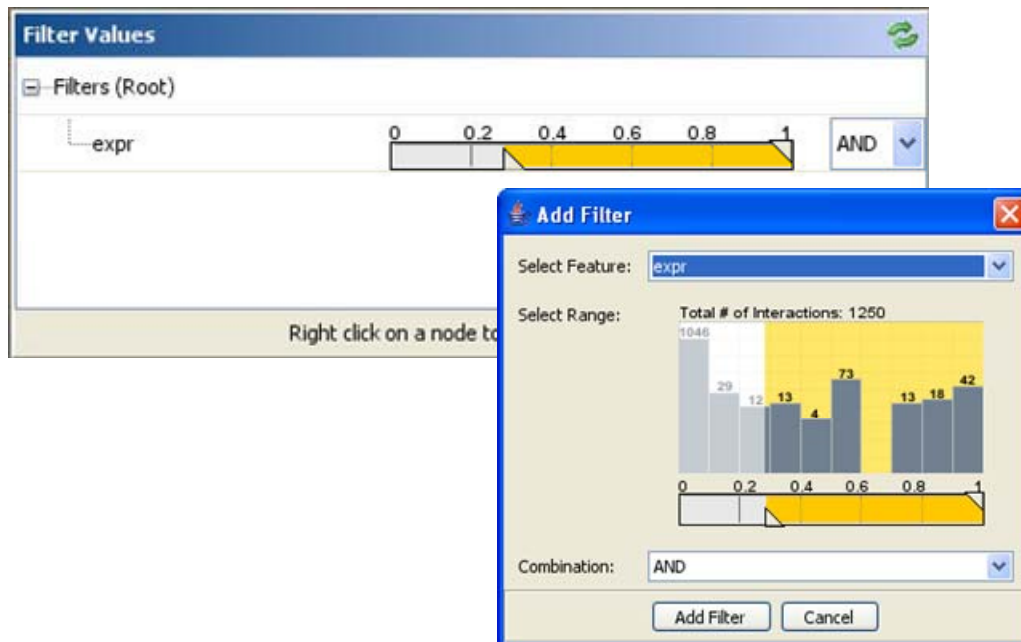


Selection in this view works in a similar manner to the selection in Weighted Combined Evidence Viewer discussed in section 7.2. It is coordinated with selection in the other views. Similarly, right-clicking on the view provides options to save selected or all interactions to a file (as described in Section 11 *Saving Results*) and to create a new evidence-network from the selected interactions or all interactions.

8. Filters

Often there's need to assign a cutoff value to the confidence values for interactions obtained from different sources of evidence. For this, CABIN provides the option to filter interactions in an evidence-network according to their confidence values. To create a filter for an evidence-network, right-click on the root node entitled “Filters (Root)” in the *Filter Values Panel*. The “Add Filter” window will appear, showing a histogram and a slider for each evidence-network selected from the drop-down list. Choose an evidence-network from the drop-down list. For that evidence-network, move the slider to a cutoff point, selecting the desired range and click ‘Add Filter’ when finished. For your reference, the histogram provides the number of interactions at each interval along with the total number of interactions in that evidence-network. There are ten intervals, separated equally between 0 and 1.

Figure 12.
Filter Values
Panel



To remove a filter, simply right-click on the filter you wish to remove in the *Filter Values Panel* and select “Remove Filter” from the popup menu. To apply the changes, click the “Update” (🔄) button. This will automatically update all the views in the CABIN Browser. Note that the “Update” (🔄) button will be disabled if there are no changes to be applied.

To view the histogram at any point, right-click and select “Show Histogram” from the popup menu. You can change the range settings from this option as well. Select and “AND” or “OR” option from the drop-down list to specify the relations between multiple filters. To view the entire formulated query, right-click on the “Filters (Root)” node and select the “Show Query” menu item. A window will appear outlining the filter query as shown in Figure 13.

Figure 13.
Evidence-
Network
Formulated
Filter Query

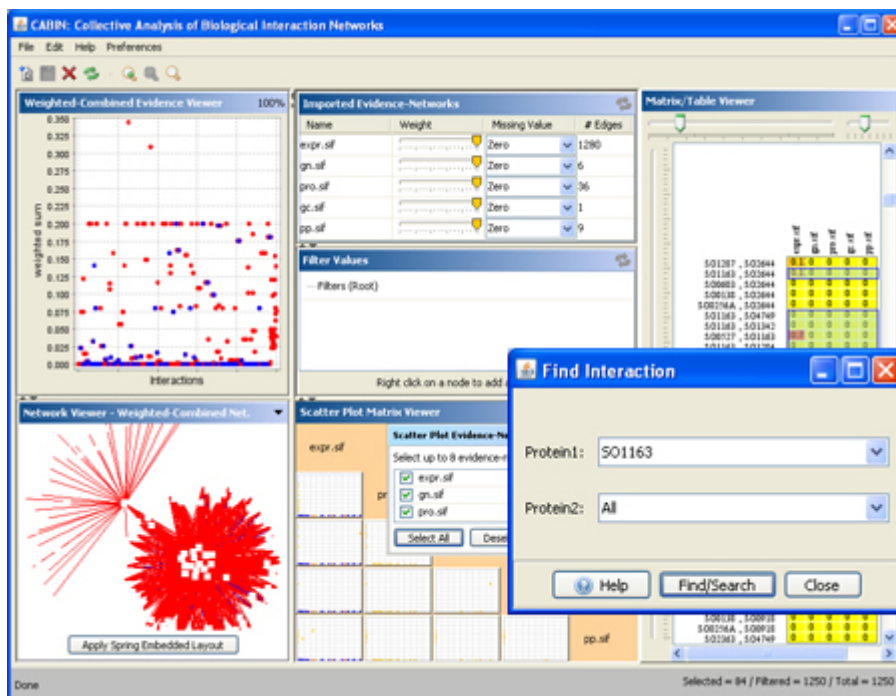


There should be evidence-networks uploaded to the CABIN Browser before you can create filters. Also, currently only one filter can be created per evidence-network.

9. Find/Search

To be able to look for an interaction of interest, CABIN provides the ability to search for an interaction between two proteins. From the “Edit” menu, select “Find Interaction”; this will bring up the “Find Interaction” window as shown in Figure 14. You can either enter the names of two proteins involved in the interaction, or enter a protein name for ‘Protein1’ and select “All” in either field. Selecting “All” for both fields will highlight all interactions. Selecting “All” for one field and a specific protein for the other will highlight all interactions with at least one protein being the specified protein.

Figure 14.
Find
Interaction

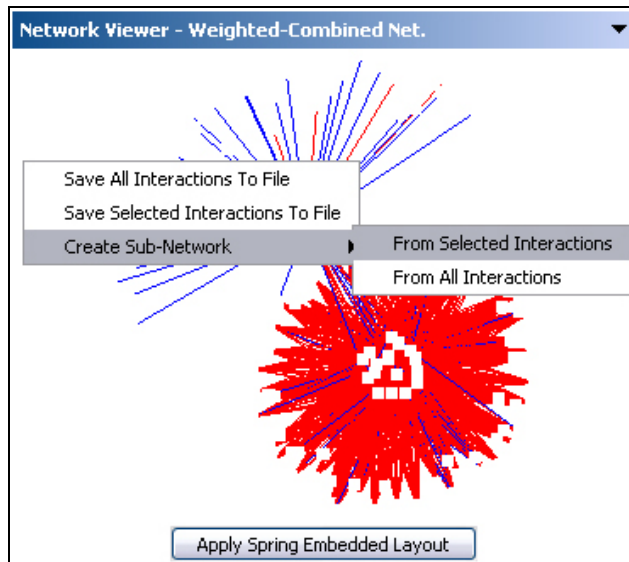


To create a more complex search expression, include character ranges ([A-z]) or number ranges ([0-100]) in the expression. For example, if looking for sequence of proteins from “SO0000” to “SO0100”, enter the expression `SO0*[0-100]`. You can also use the expression `“0*”` to indicate leading zeros. To search for multiple proteins in either field, you may enter a comma separated list of protein names. For more help in using these ranges or to see an example, click the “?” button on the “Find Interaction” window.

10. Creating Sub-Networks

CABIN provides the option to create a sub-network from all or selected interactions in the Weighted-Combined Evidence Viewer, Network Viewer, and Matrix/Table Viewer. This sub-network is automatically imported into CABIN. To create a sub-network, right-click on any of the three views and a popup menu will appear. From the “Create Sub-Network” submenu, select either all interactions or selected interactions as shown in Figure 15.

Figure 15.
Create
Sub-
Network
from
Interactions



The “Create Sub-Network” dialog will appear as shown in Figure 16. Enter a name and choose a default weight for the new network. Choose which values to use for assigning the confidence values of the new network from the “Value to Keep” combo box. Finally, select the level to extend the network to. There are 3 level options:

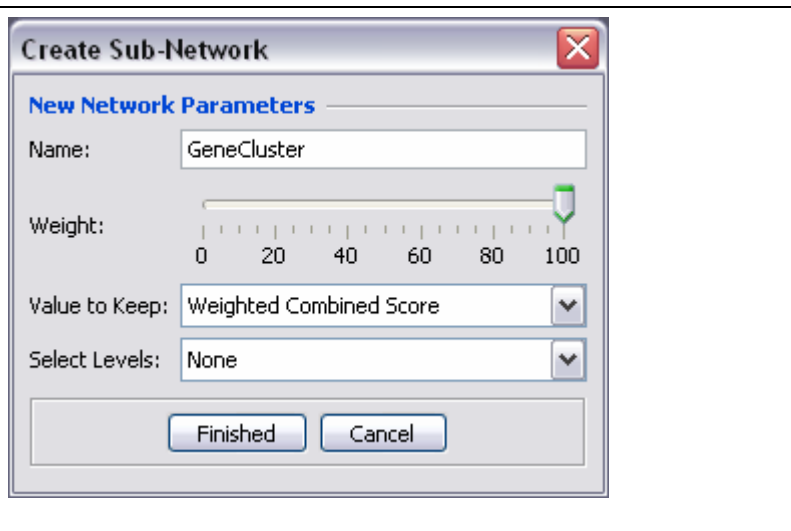
None - *only selected interactions*

Level 1 – *selected interactions and neighboring interactions one-hop away*

Level 2 – *selected interactions and neighboring interactions up to 2 hops away*

When you have finished assigning all parameters, click “Finished” and the new network will be automatically imported into CABIN.

Figure 16.
Create
Sub-
Network
Parameters

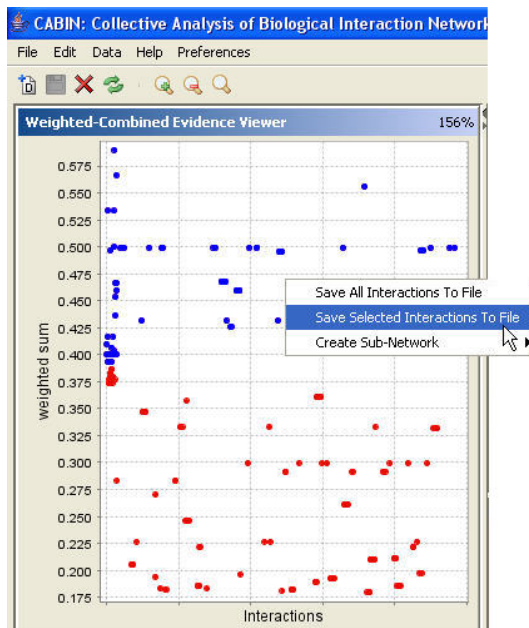


11. Saving Results

CABIN provides the option to save all or selected interactions in a view to a file on your local machine. To save interactions to a file, right-click inside the view to bring up a

popup menu with two options: save all interactions and save selection interactions. Choose one option from the menu depending on what you want to save. Choose the appropriate location from the file browser window that appears; specify a file name and click “Save” when finished.

Figure 17.
Save
Interactions
to File



To export all evidence-networks as SIF files to a selected directory, select “Export All Evidence-Networks” from the “File” menu.

12. Acknowledgments

The research described in this paper was conducted under the LDRD Program at the Pacific Northwest National Laboratory, a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy under Contract DE-AC06-76RL01830. The authors acknowledge contributions from Elie Naulleau [semiosys@semiophore.net] in obtaining the updated version of the jMatrixView open source library and making it available for public use.

Mac OS is a trademark of Apple Computer, Inc., registered in the U.S. and other countries.

Windows XP and Windows 2000 are registered trademarks of Microsoft Corporation in the United States and other countries.

Linux® is the registered trademark of Linus Torvalds in the U.S. and other countries.

13. Citing CABIN

Use the following text to cite CABIN:

Singhal, M., Domico, K., "CABIN: Collective Analysis of Biological Interaction Networks", *Computat. Biol. Chem.* (2007), doi:10.1016/j.compbiolchem.2007.03.006

14. References

1. Shannon, P. T., A. Markiel, et al. (2003). "Cytoscape: a software environment for integrated models of biomolecular interaction networks." *Genome Research* 13(11): 2498-2504.
2. [Prolinks : a database of protein functional linkages derived from coevolution](#)
Bowers PM, Pellegrini M, Thompson MJ, Fierro J, Yeates TO, Eisenberg D
Genome Biology 2004, 5(5):R35
3. Xenarios I, Rice DW, Salwinski L, Baron MK, Marcotte EM, Eisenberg D (2000) "DIP: The Database of Interacting Proteins". *NAR* 28:289-91
4. Bader, G. D. and C. W. Hogue; (2000). "BIND: A Data Specification for Storing and Describing Biomolecular Interactions, Molecular Complexes and Pathways." *Bioinformatics* 16(5): 465-477.
5. BioNetBuilder- <http://err.bio.nyu.edu/cytoscape/bionetbuilder/>
6. KEGG – <http://www.genome.jp/kegg>
7. HPRD – <http://www.hprd.org>
8. GO – <http://geneontology.org>

15. Appendix A

rpal_pulldown_specific.sif

The pulldown specific network was created using the Rpal854 dataset which includes 854 samples for 291 bait proteins, processed using the BEPRO3 algorithm in SEBINI. Interactions were included if BEPRO3 indicated that they were specific binding interactions based on homogeneity of the prey protein hits.

rpal_pulldown_nonspecific.sif

The pulldown nonspecific network was created using the Rpal854 dataset which includes 854 samples for 291 bait proteins, processed using the BEPRO3 algorithm in SEBINI. Interactions were included if BEPRO3 indicated that they might be false positive interactions based on homogeneity of the prey protein hits.

rpal_phylogenetic.sif (http://128.97.39.94/cgi_files/functionator/about.html)

The Phylogenetic Profile network is created by using the phylogenetic profile method from the Prolinks database which uses the presence and absence of proteins across multiple genomes to detect functional linkages.

rpal_genecluster.sif (http://128.97.39.94/cgi_files/functionator/about.html)

The Gene Cluster method from the Prolinks database is used to create the gene cluster network which uses genome proximity to predict functional linkage.

rpal_rosettastone.sif (http://128.97.39.94/cgi_files/functionator/about.html)

The Rosetta Stone network is created by using the Rosetta stone scores from the Prolinks database which uses a gene fusion event in a second organism to infer functional relatedness.

rpal_geneneighbor.sif (http://128.97.39.94/cgi_files/functionator/about.html)

The Gene Neighbor method in the Prolinks database uses both gene proximity and phylogenetic distribution to infer linkage.

rpal_interolog.sif (<http://bioverse.compbio.washington.edu>)

The interolog method predicts an interaction between two proteins if they are both homologs of two proteins known to interact. Known protein interactions are gathered from databases of experimentally-determined protein-protein interactions (e.g. BIND, DIP) and PSI-BLAST is used to determine similarity between this set and all proteins in a target organism.

rpal_regulog.sif (<http://bioverse.compbio.washington.edu>)

Regulogs are regulatory interactions inferred by homology. A regulog is predicted by determining similarity to a known transcription factor (TF) and the TF's target protein. Finally the nucleotide similarity in the upstream transcriptional promoter regions is determined and used to filter the regulog predictions: if there are similar promoter sequences then a regulog is predicted.

rpal_location.sif

The location network is created by looking at the subcellular compartment information of the two proteins in all possible interactions amongst the proteins in the pull-down experiments. The edge was assigned a value of 1 if both the proteins were localized in the same compartment and a value of 0 otherwise. If the compartment information for at least one of the proteins was unavailable or the protein was localized in multiple compartments, then a value of 0.5 was given to the interactions involving those proteins.