

RobinViz Manual

Ahmet Emre Aladağ, Cesim Erten, Melih Sözdinler

March 13, 2011

Contents

1	Manual	1
1.1	Overview	1
1.2	Installation	2
1.2.1	Linux Binary	2
1.2.2	Linux Source	2
1.2.3	Windows Binary	3
1.2.4	Windows Source	3
1.3	Quickstart	3
1.3.1	Starting the program	3
1.3.2	Running the Wizard	3
1.3.3	Preconfigured Settings	4
1.3.4	Last Settings	4
1.3.5	Manual Settings	4
1.4	Tutorial	4
1.4.1	Introduction	4
1.4.2	Precautions	4
1.4.3	Main Window	5
1.4.4	Menu	5
1.4.5	Execution Wizard	6
1.4.6	Co-Ontology Results	7
1.4.7	Co-Expression Results	15

1 Manual

1.1 Overview

We present our PPI network visualization system RobinViz (Reliability Oriented Bioinformatic Networks Visualization) which is designed to visually aid the prediction and verification processes of such networks. Embedding both the reliability (confirmation) values associated with the interactions and the verification data pertaining to them within a visualization model is a novel feature of the system. RobinViz is a free, open-source software protected under GPL. It is written in C++ and Python, and consists of almost 30,000 lines of code, excluding the employed libraries. You can find up-to-date version of this manual on <http://code.google.com/p/robinviz/wiki/Manual>

1.2 Installation

Here you can find instructions on how to install Robinviz.

Runtime Requirements: Python 2.7, PyQt 4.7

Library Requirements: LEDA 5.1+ Library from Algorithmic Solutions

Additional Windows Requirements: Windows XP SP3 / Vista (recommended), Visual Studio 2005 C++ or over for compilation from source.

1.2.1 Linux Binary

1. Use the following command to install PyQt4:

(a) `sudo apt-get install python-qt4`

2. Run the installer file: Robinviz-1.0.0-Linux-x86-Install. If it doesn't run by clicking, use the following command to give it executable rights and run it:

(a) `chmod +x Robinviz-1.0.0-Linux-x86-Install`

(b) `./Robinviz-1.0.0-Linux-x86-Install`

3. After the installation, open a new terminal/konsole to run the program with updated PATH variables. Move into the installation dir and run robinviz.exe

(a) `cd ~/robinviz`

(b) `./robinviz.exe`

1.2.2 Linux Source

1. Use the following command to install the required packages:

(a) `sudo apt-get install python-qt4 g++ libX11-dev`

2. Copy the distribution folder to anywhere you like (for example inside your home dir):

(a) `cp robinviz-1.0-source ~/robinviz`

3. Add the following lines to your `~/bashrc` (Make sure that incl folder is inside this LEDA_ROOT):

(a) `export LEDAROOT=/path/to/LEDA`

(b) `export PATH=$PATH:$LEDAROOT/Manual/cmd`

(c) `export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$LEDAROOT`

4. Give the following command on console:

(a) `source ~/.bashrc`

5. Move to the robinviz directory on console and run the following command:

- (a) `cd ~/robinviz`
- (b) `sh compile.sh`
- (c) `./robinviz.exe`

1.2.3 Windows Binary

1. Start the installation wizard to install the program under C:\Robinviz. Double click on the Robinviz icon on your desktop. Path should not include any spaces.

1.2.4 Windows Source

1. Follow the instructions at http://www.algorithmicsolutions.info/leda_manual/DLL_s_MS_Visual.html

or

1. Use sample Visual Studio 2005 Solution Template located at `src/cpp/Robinviz-Windows-Installer`.
2. Setup LEDA 5.1+ Library from Algorithmic Solutions
3. Follow the instructions at http://www.algorithmic-solutions.info/leda_manual/DLL_s_MS_Visual.html or if you use the template, you will need to add library path and include folder from menu bar Tools → Options → Project and Solutions → VC++ Directories.

1.3 Quickstart

1.3.1 Starting the program

When robinviz starts, you can follow debug information on the terminal (black window) that is opened aside and on the Log window that appears during the calculation. If you encounter any freeze or problems, you can figure out the reason for that from these sources. If you'd like to report any problems/bugs, please include debug information you see here.

- Windows: From Start Menu→Programs→Robinviz, click on Robinviz. A black window (terminal) and graphical user interface (GUI) of Robinviz will start.
- Linux: Start a new terminal/konsole. Change directory to robinviz installation dir and run the executable:

- `cd ~/robinviz`
- `./robinviz.exe`

1.3.2 Running the Wizard

Follow the File Menu→Execute path.

1.3.3 Preconfigured Settings

To have a quick run without messing with detailed configuration, you can run our preconfigured parameter settings.

1. In the Execution Wizard opened, check Use preconfigured settings radio button.
2. Select a configuration from the dropdown menu.
3. Click on Finish button. Please wait for a while, this may take a few minutes.

1.3.4 Last Settings

For any reason, if you'd like to re-run the latest configuration,

1. Check the Use the last settings radio button.
2. Click on Finish button.

1.3.5 Manual Settings

If you want to define your own parameters for custom execution,

1. Check Define your manual settings radio button.
2. Click on Next button to follow the next steps of the wizard.

1.4 Tutorial

1.4.1 Introduction

RobinViz (Reliability Oriented Bioinformatic Networks Visualization) is a protein-protein interaction (PPI) network visualization system designed to visually aid the prediction and verification processes of such networks. Embedding both the reliability (confirmation) values associated with the interactions and the verification data pertaining to them within a visualization model is a novel feature of the system. RobinViz is a free, open-source software protected under GPL. It is written in C++ and Python, and consists of almost 30,000 lines of code, excluding the employed libraries.

Executable binaries of the system can be accessed via the Downloads link from our website <http://code.google.com/p/robinviz>. These binaries can be downloaded and executed directly without any problems on most of the systems. Additionally we provide the source code implementations of the system. If the user wants to compile the source code and prepare her own executables we provide the necessary instructions under the Installation section of the manual. In this tutorial, we introduce features of Robinviz and how to make use of them and provide snapshots of a sample run.

1.4.2 Precautions

When switching between wizard steps, a progress bar is shown for user to understand some processing is being done on the data. Please be patient until these kind of processes ends. Just in case if no progress bar is shown, please follow the console messages to see the progress.

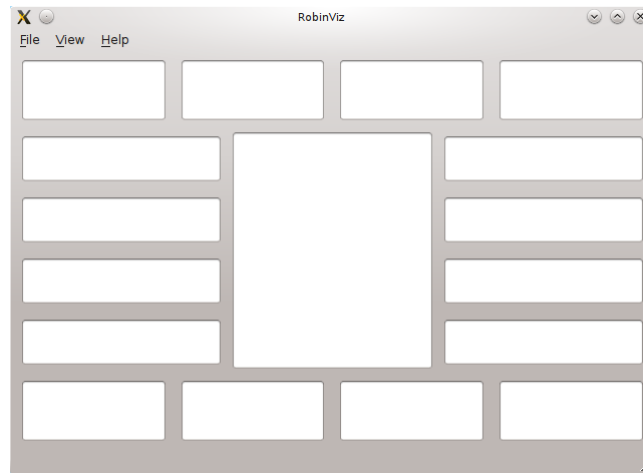


Figure 1: Empty Robinviz MainWindow

1.4.3 Main Window

When you run Robinviz, you will see an empty main window. In this window, the central part is the Central View and the smaller rectangles around the Central View are called the Peripheral Views. Window also has a menu bar and has menu options. (See Figure 1)

1.4.4 Menu

In the main window we have three main menus: File, View and Help.

- File
 - Execute (Ctrl+X): Displays the execution wizard and lets you select your inputs and perform the calculations.
 - Load Session (Ctrl+O): Loads previously saved session data to the program.
 - Save Session (Ctrl+S): Saves current calculation results for future use.
 - Display Last Session (Ctrl+D): Displays calculation results most recently obtained on the screen.
 - Update Local Data (Ctrl+U): Lets you download/update your database files.
 - Exit (Ctrl+Q): Quits the program.
- View
 - Color Legend (F6): Displays color legend for user to match the colors with the most popular high level categories.
 - Go To (Ctrl+G): Lets you quickly locate a bicluster by entering its number.

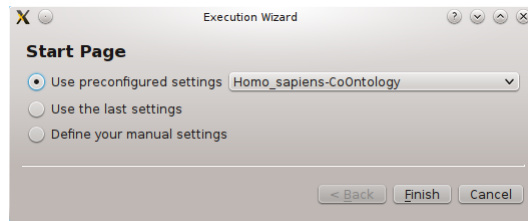


Figure 2: PreConfiguration Page

- Refresh(F5): Refreshes the drawings.
- Clear Views (Ctrl+L): Clears Central View and all the Peripheral Views.
- Fullscreen (F11): Toggles fullscreen mode.
- Help
 - Manual: Displays download links for the manual.
 - About: Displays the about dialog.

1.4.5 Execution Wizard

You can start this wizard following the File. There are three options you can select to move forward:

- Use preconfigured settings: This option provides you pre-selected data sources for execution so that you won't be lost in manual settings that are unfamiliar to you. It's helpful for the users running Robinviz for the first time. If you'd like to use this feature, please select this option and choose one pre-configured setting from the dropdown menu. Then click on the Finish button.
- Use the last settings: This option lets you re-perform calculations with the most recent settings without re-specifying them. Click on the Finish button after selecting this option. *Warning: If you just want to display last results without re-calculating, use File-Display Last Session.*
- Define your manual settings: This option lets you do custom configuration and input selection in the following wizard pages. Select this option and click on the Next button.

If you select one of the first two options, Robinviz will start performing calculations and display the results on the screen. But if you selected the Manual settings option, then you will have to provide some more preferences. Here are the following wizard pages:

1. PPI Network - Figure 3
2. Verification Concept - Figure 4
3. Color Assignment - Figure 5

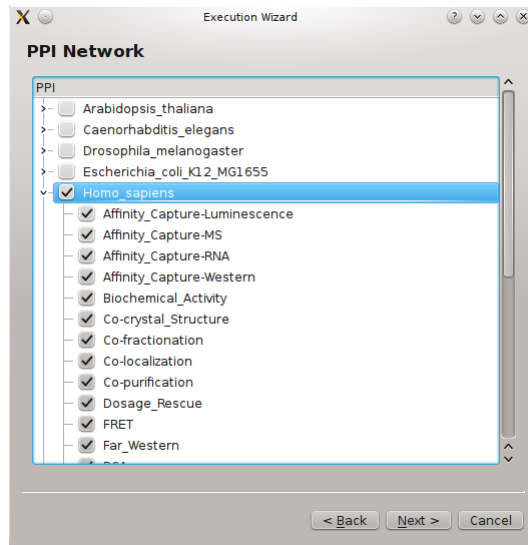


Figure 3: In this dialog, you will see a list of organisms and experiment types under each organism. Please select the PPI Network files you'd like to use here. If you select nothing, the selection in your last execution will be used.

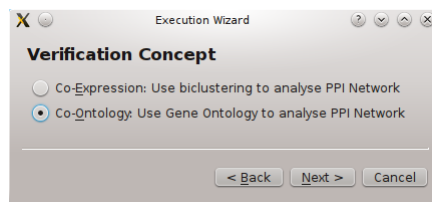


Figure 4: In this dialog, you are required to select the verification concept. Robinviz shall categorize genes according to their co-ontology or their co-expression information by looking at this option.

4. Central Nodes from GO Categories - Figure 6
5. GO Association Sources - Figure 7
6. GEO Expression Matrix - Figure 8
7. Biclustering Algorithm - Figure 9
8. Central Node Weights - Figure 10
9. Ready Page - Figure 11

Here are some screenshots from these pages:

1.4.6 Co-Ontology Results

Central View When you use Co-Ontology concept, Robinviz will display you a Central View with Central nodes each corresponding to the GO Categories

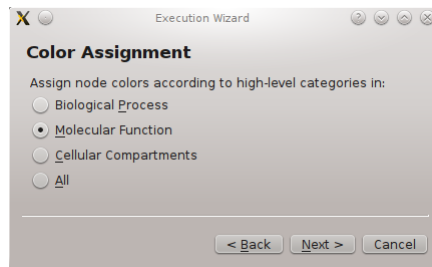


Figure 5: Nodes are colored to highlight their high level categories. You are asked to select which categories you'd like to use for this purpose. Top 10 categories that contain the most genes will be used for coloring.

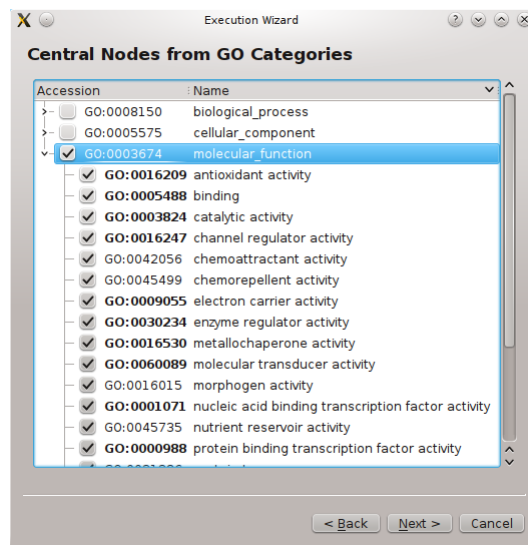


Figure 6: In this part, you are asked to define the what the central nodes (categories) will be in the central view. Genes will be categorized according to these categories you select. Note that some categories are in bold. You can double click on those categories to see its sub-category list. Selecting a highlevel category such as “binding” does not include its sub-categories automatically. This is because that the central node “binding” will cover all those sub-categories.

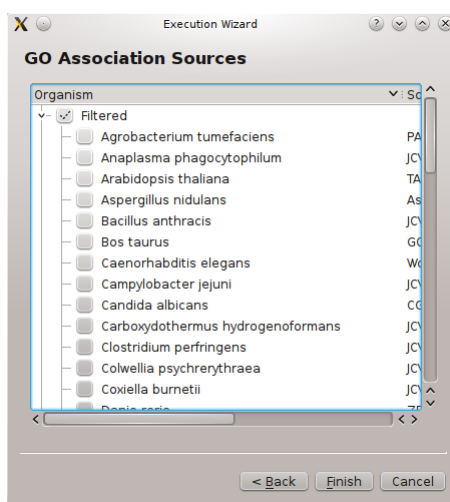


Figure 7: In this part, you are asked to select a GO Association source. These data tell us which gene is in which category. Multiple selection is doable but in this screenshot, only Filtered Homo Sapiens data is used.

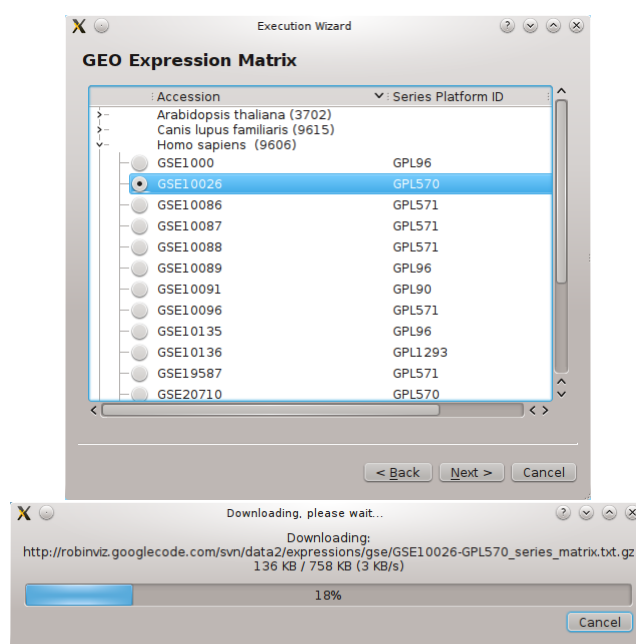


Figure 8: If you selected Co-Expression for verification method, you will see some more dialogs such as this one. Here, you are asked to select one Gene Expression Matrix data which will be downloaded from our servers. This data will be used for biclustering.

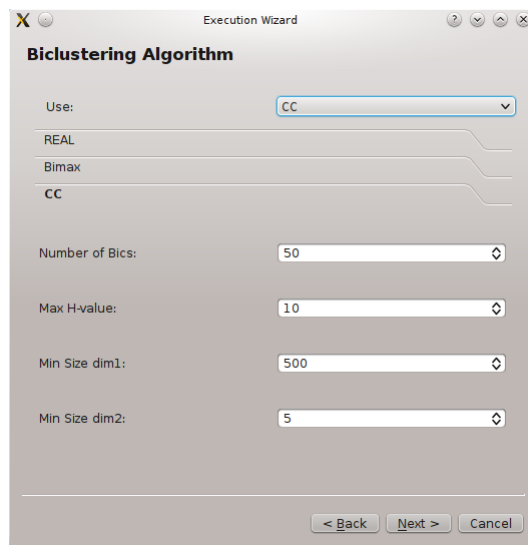


Figure 9: After GEO Expression Matrix selection, you are asked to select a biclustering algorithm and provide its parameters. Each GEO file might require different parameters so you might need to try different parameters to obtain to optimum biclustering.

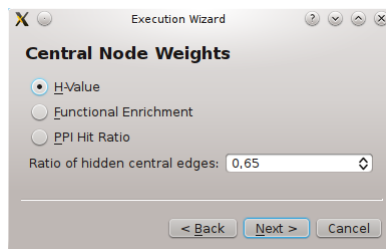


Figure 10: In this dialog, you are asked to define the method for calculating the central node weights and the ratio of hidden central edges ratio. If you keep this ratio close to 0, almost all the edges will be displayed which may result in cluttered graphs. If you increase this ratio, weaker edges will be eliminated so that only edges representing most reliable interactions will survive.

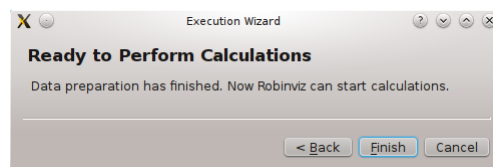


Figure 11: This dialog shows data preparation has been finished and you may now start Robinviz performing calculations.

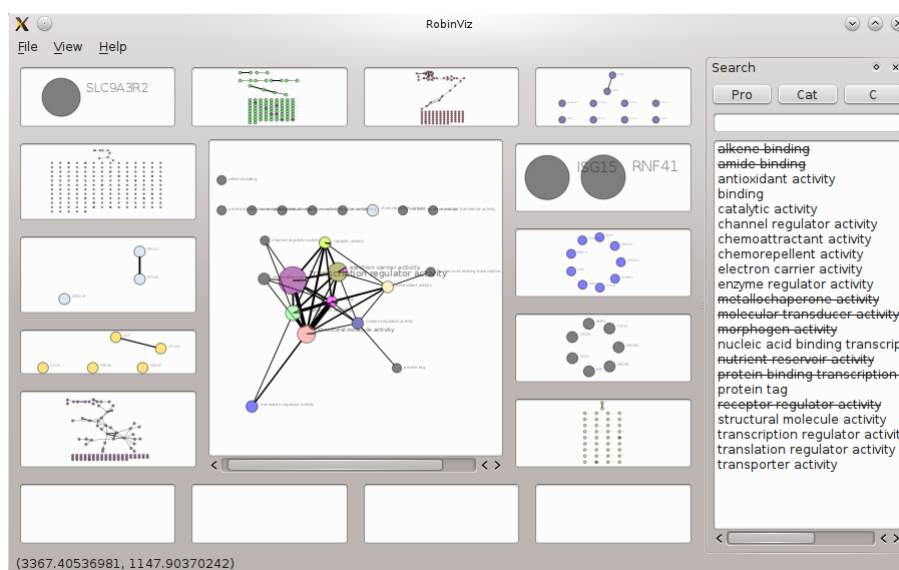


Figure 12: Co-Ontology results for Homo Sapiens PPI Network and Association data, categorized by molecular functions.

you selected in the wizard (See Figure 12). When you double click on a node, its contents (genes associated with this category) are displayed on an available peripheral view. When you re-double click, then the peripheral view will be cleared. These Central Nodes are piecharts and have some colors. These colors and their ratios represent the ratio of the highlevel categories of the inhabitants of this category. For example, peripheral view corresponding to this category has 10 genes and 9 of them belong to “binding” highlevel category and 1 to the catalytic activity, then node will have 90% binding slice and 10% catalytic activity slice.. You can move the central nodes. When you select a node, its edges are highlighted in red and its corresponding peripheral view is highlighted in yellow. You can zoom in and out by scrolling.

When you right click on a Central Node, you’ll see 2 options:

- Detailed Information (Online): Provides detailed information about this category from the AmiGO Browser. (See Figure 13)
- Enrichment Analysis: Provides Enrichment Analysis tables for this category and its contents. (See Figure 14)

When you right click on an empty space on the central view, you’ll encounter another menu with the following items:

- Open in new window: Opens the Central View in a separate, bigger window.
- Switch to Layout: Changes the current layout of the Central View.
- Save as Image: Saves the Central View as an image file.
- Save as GML: Saves the Central View as a graph file.



Figure 13: Detailed Category Information from AmiGO Browser.

Enrichment Information respect to genes/proteins of Category(GO or Bicluster) inside PPI network				Gene Ontology Information of genes/proteins of Category(GO or Bicluster) inside PPI network		
Categories	Number of Genes	Ratio Respect to GO or Bicluster Gene Space	P-values	Gene(Protein)	Category Name	Gene Ontology Id
antioxidant activity	10	0.416667	0.000000	ALB	antioxidant activity	GO:0016209
binding	0	0.000000	0.056047	APOA4	antioxidant activity	GO:0016209
catalytic activity	0	0.000000	0.222255	APOE	antioxidant activity	GO:0016209
channel regulator activity	2	0.083333	0.082776	PRDX6	antioxidant activity	GO:0016209
chemoattractant activity	2	0.083333	0.066161	PRDX2	antioxidant activity	GO:0016209
chemorepellent activity	0	0.000000	0.635819	GO Id 8, can view html file at ouputs/go/		

Figure 14: Enrichment Analysis for antioxidant activity.

- **Print:** Prints the Central View with your printer.

Peripheral Views When you double click on a Central Node, a preview of its contents appear on a peripheral view. You can see its contents and even move its nodes. When you right click on an empty space in a peripheral view, you'll encounter such a menu:

- **Open a new window:** Opens the Peripheral View in a separate, bigger window.
- **Clear:** Disassociates this peripheral window and its category (i.e. clears it).
- **Enrichment Analysis:** Displays Enrichment Analysis table for this category.
- **Switch to layout:** Changes the current layout of the Peripheral View.
- **Save as Image:** Saves the Peripheral View as an image file.
- **Save as GML:** Saves the Peripheral View as a graph file.
- **Print:** Prints the Peripheral View with your printer.

When you look at the nodes in the peripheral view in detail, you will see that these nodes represent the proteins. They are in a piechart form and their piechart has colors of the corresponding highlevel categories' representation colors. For example, if a protein node has two color, it means it belongs to 2 high-level categories. When you right click on an empty space, you will encounter the same menu as mentioned above. But if you right click on a peripheral (protein) node, you'll see this menu:

- **Detailed Information (Online):** Displays more information about the protein from BioGRID website. (See Figure 15)
- **Display Neighbors in the whole PPI:** Displays the other proteins in the whole PPI (not just in this category) that interact with the protein in a new window. (See Figure 16)

When you hover on a protein node, you'll see what highlevel categories it belongs to. And when you display views in seperate windows, you will have File, View, Help menus.

Search Panel You can see a right panel in the main window and this panel lists the categories you had selected. The ones struck demonstrate the ones that do not have any genes associated. You can click on these categories to see which node they are correspond. The single nodes that are apart from the main drawing correspond to the struck categories. Clicking and double clicking on the panel items works as if you are clicking on the actual nodes. Moreover, you can list the genes/proteins by clicking on the Pro button. If you want to list the categories again, click on the Cat button. You can search for a gene by typing in the text edit box. Robinviz has an autocompletion feature that enables you to quickly locate your target. When you write a protein name and

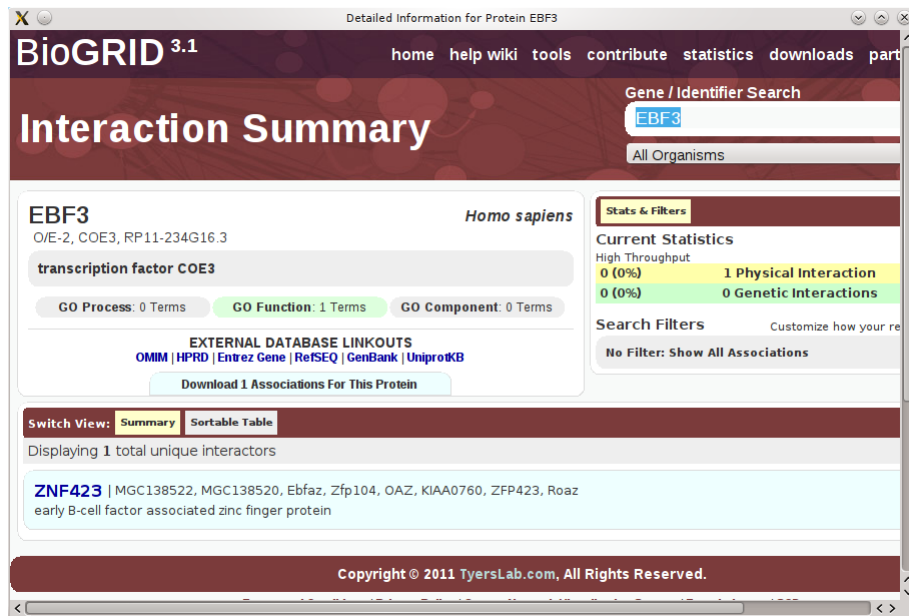


Figure 15: Detailed information for protein EBF3 on BioGRID website.

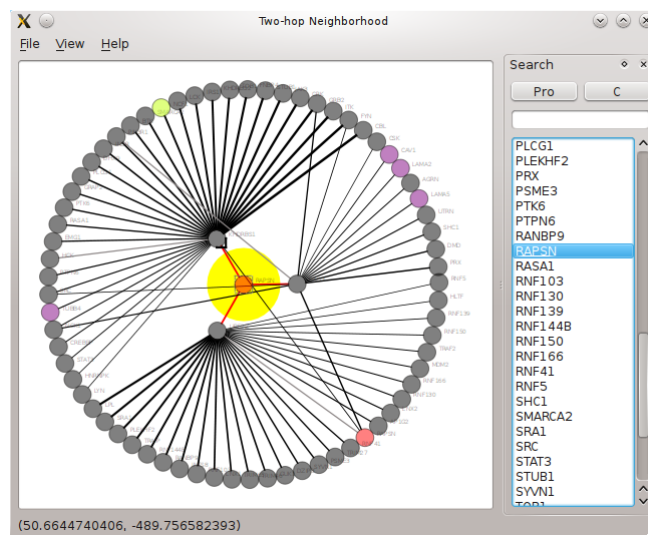


Figure 16: Two-hop neighborhood is displayed for protein RAPSIN. It has three one-hop neighbors and many other two-hop neighbors.

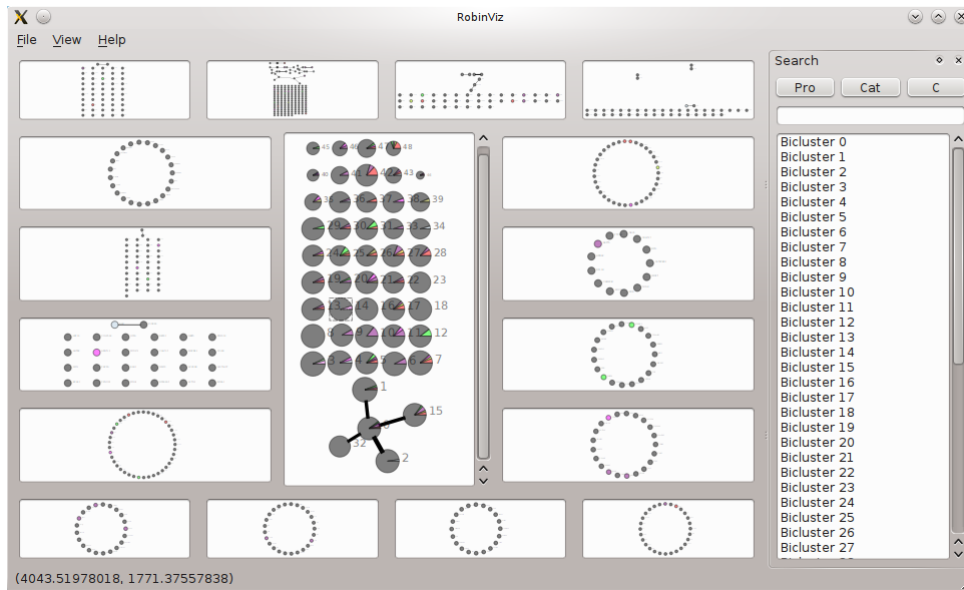


Figure 17: Co-Expression results for Homo Sapiens PPI/Association data with BiMAX Algorithm applied on Homo Sapiens GSE1000 GEO data.

press enter, the list will show you in which categories that protein resides in. You can click/double click on those categories to dive in.

When you display peripheral views in a separate window, another search panel will help you locate the proteins in that view quickly. Just press Pro button and see the protein list in that view. If you look for a specific protein, type the protein name and press enter or select it from the list. Protein node will be highlighted in yellow.

1.4.7 Co-Expression Results

Co-Expression results are similar to Co-Ontology results. We will discuss about the differences in this part.

Central View You can see that Central Nodes this time have integer labels representing the bicluster number. When you right click on an empty space, you can see an extra option called Enrichment Table. Here you can see an extensive comparative table on Enrichment Analysis. Each cell in a row tells us the number of genes/proteins belonging to the category in the corresponding column (See Figure 18)

When you right click on a Central Node, you'll see an additional menu item called Visualization. Here we provide two visualization: Heatmap and Parallel Plot.

When you hover on a node, you will see the corresponding biclustering (such as H-value) score for that bicluster.

Peripheral View In a peripheral view right click menu on an empty space has an additional Visualization item will appear.

Category(Biclust Id	antioxidant activity	binding	catalytic activity	channel regulator activity	chemoattractant activity	chemorepellent activity	electron carrier activity	enzym regul activi
Category 1	38	36	22	6	7	5	11	13
Category 2	11	6	7	0	0	1	2	5
Category 3	3	4	1	1	0	1	0	0
Category 4	3	3	5	1	2	0	1	1
Category 5	5	3	1	0	1	0	0	1
Category 6	4	5	3	0	1	1	1	4
Category 7	1	2	2	0	2	0	1	0

Figure 18: Enrichment Table. Biclusters are on the left, highlevel categories are listed on the top.

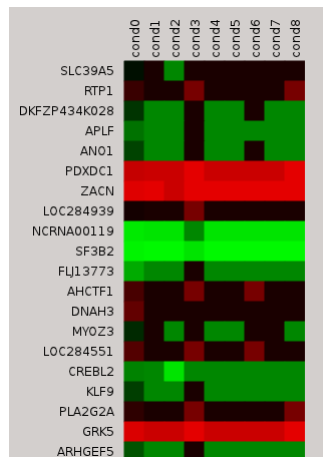


Figure 19: Heatmap representation of a bicluster. Black represents the median, lightest green represents the lowest gene expression value, lightest red represents the highest gene expression value. Dark colors represent values closer to the median.

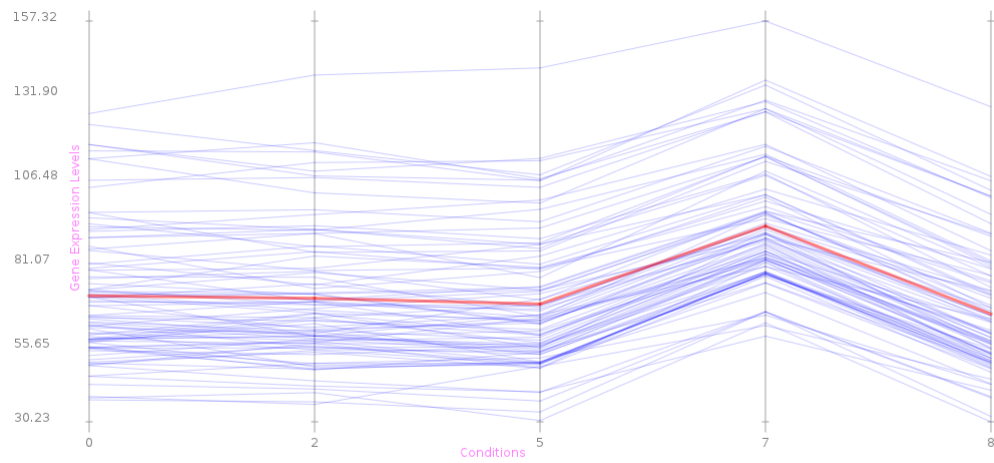


Figure 20: Parallel Plot diagram for a bicluster. y-axis represents the expresison levels whereas x-axis represents the conditions. Each blue line represent a gene's expression levels. Red line represent the average value for each condition.

Search Panel In the search panel, instead of categories, Bicluster numbers are displayed.