Biological Network Exploration with Cytoscape 3

Gang Su,[1] John H. Morris,[2] Barry Demchak,[3] and Gary D. Bader[4]

[1]Molecular Behavioral Neuroscience Institute, University of Michigan, Ann Arbor, Michigan
[2]Resource for Biocomputing, Visualization, and Informatics, University of California San Francisco, San Francisco, California
[3]Department of Medicine, University of California San Diego, La Jolla, California
[4]The Donnelly Centre, University of Toronto, Toronto, Canada

ABSTRACT

Cytoscape is one of the most popular open-source software tools for the visual exploration of biomedical networks composed of protein, gene, and other types of interactions. It offers researchers a versatile and interactive visualization interface for exploring complex biological interconnections supported by diverse annotation and experimental data, thereby facilitating research tasks such as predicting gene function and constructing pathways. Cytoscape provides core functionality to load, visualize, search, filter, and save networks, and hundreds of Apps extend this functionality to address specific research needs. The latest generation of Cytoscape (version 3.0 and later) has substantial improvements in function, user interface, and performance relative to previous versions. This protocol aims to jump-start new users with specific protocols for basic Cytoscape functions, such as installing Cytoscape and Cytoscape Apps, loading data, visualizing and navigating the networks, visualizing network associated data (attributes), and identifying clusters. It also highlights new features that benefit experienced users. *Curr. Protoc. Bioinform.* 47:8.13.1-8.13.24. © 2014 by John Wiley & Sons, Inc.

Keywords: Cytoscape • interactive network visualization • network analysis

INTRODUCTION

A network model (or *graph* in mathematics) represents associations between entities in a system. Such models are commonly used to study complex systems in many disciplines, including computer science, social science, and life sciences. The molecules in a biological system interact with each other and form molecular complexes, modules, or pathways that carry out various biological functions. In a biological network, nodes (or vertices) often represent proteins, genes, or metabolites, while edges often represent relationships, such as physical interactions or gene-expression regulation (Merico et al., 2009).

These networks can be generated from prior knowledge or deduced from experimental data. Many online repositories, such as KEGG (Kanehisa, 2002; *UNIT 1.12*), Reactome (Joshi-Tope and Gillespie, 2005; *UNIT 8.7*), BIND (Bader et al., 2001), HPRD (Goel et al., 2012), IntAct (Kerrien et al., 2012), iRefWeb (Turinsky, 2014), MiMI (Tarcea et al., 2009), STRING (Franceschini et al., 2013), and GeneMANIA (Zuberi et al., 2013), host a large amount of data that can be readily represented as a network and then analyzed. On the other hand, rapid technological advances in high-throughput technology have improved the feasibility of constructing networks automatically from tens of thousands of molecular profiles (Dutkowski et al., 2014; Dutkowski et al., 2012; Margolin et al., 2006). Cytoscape supports visualization, analysis, and interpretation of these networks, and helps better understand the biological systems they model.

Analyzing Molecular Interactions

Cytoscape was developed as a response to this network data explosion and the visualization and analytical challenges it poses, starting in 2001 (Shannon et al., 2003). While many tools can be used for general-purpose network visualization or analysis—e.g., Pajek (Batagelj and Mrvar, 1998), Gephi (Bastian et al., 2009), Jung (Fisher et al., 10), GraphStream (Dutot et al., 2007), and igraph (Csardi and Nepusz, 2006)—Cytoscape aims to satisfy the unique needs of biologists needing to interactively explore biological networks, such as metabolic pathways or gene regulatory networks, in the context of corresponding experimental data. For example, gene expression changes obtained from a transcriptomics experiment can be used to color nodes in a disease pathway so that researchers can study genes of interest (e.g., the most differentially expressed genes) in the context of existing pathway knowledge, and then infer potential novel gene functionality. A variety of text and numeric data, such as gene function annotations, pathways, and expression profiles, can be imported and projected onto Cytoscape networks. In addition to these core capabilities, Cytoscape differentiates itself from other network visualization tools by enabling and encouraging active third-party development of add-on visualization and analysis Apps. A large number of Cytoscape Apps (known previously as *plugins*) designed for biological networks are available from the Cytoscape App Store, providing functionality including data import from external repositories, functional annotation and discovery, module detection, literature search, network layouts, and network filtering (Lotia et al., 2013; Saito et al., 2012). This rapidly growing App ecosystem makes Cytoscape very attractive for users who need easy-to-use analysis tools for biological network data.

In the following text, we introduce several analysis protocols addressing frequently encountered Cytoscape usage scenarios. Following these protocols, you will learn how to set up Cytoscape, install and use Apps, import biological networks, use visual styles to map attribute data, find highly inter-connected clusters and generate network layouts to aid visualization.

ANALYZING GENE EXPRESSION DATA IN CYTOSCAPE

One common use of Cytoscape is to map attribute data (such as experimental data or text annotations) onto a biological network, such as a protein-protein interaction network or metabolic pathway. This helps visualize multiple types of data in the same plot to help identify patterns and relationships between data of diverse types. In this protocol, we will use a yeast protein-protein interaction network and a classic gene expression experiment to illustrate the process of integrative data visualization using 'attribute mapping'. We will also use a Cytoscape App to identify regions (sub-networks) that could be biologically important.

We begin by retrieving the gene expression dataset from a classic yeast experiment by Gasch et al. (2000) that explored how yeast gene expression changes in response to environmental stimuli. The data that reflect gene expression changes in response to temperature shock will be used for subsequent exploration. The dataset is stored in the NCBI Gene Expression Omnibus (GEO) data repository (Edgar et al., 2002; Barrett et al., 2007) with the accession number GDS112.

For installation of Cytoscape and necessary resources, see Support Protocol 1.

Fetch expression data from GEO

1. The full GEO data (`.soft`) format can be directly downloaded via a Web browser from this link: *ftp://ftp.ncbi.nlm.nih.gov/geo/datasets/GDSnnn/GDS112/soft/ GDS112_full.soft.gz*. You can also interactively explore these data using online tools via *http://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS112*.

2. The downloaded file is compressed in gzip format (with a `.gz` extension). You can decompress the file directly by using the 7-Zip utility (*http://7-zip.org/*) on Windows or the Archive utility on Macintosh.

Load the protein-protein interaction network into Cytoscape

3. Launch Cytoscape.

4. At the Welcome screen, click on the "*S. cerevisiae*" (yeast) button under From Organism Network. This will load a BioGRID (Chatr-Aryamontri et al., 2012) interaction network for *Saccharomyces cerevisiae* (baker's yeast). This network contains approximately 6,600 nodes and 340,000 edges, depending on the network version your Cytoscape loads. The imported network can be found listed in the Network tab.

   Note that, due to its large size, no network view is created automatically. This saves computer resources. Also, very large networks are often too dense to effectively visualize.

Import the gene expression data

5. Go to the File menu → Import → Table → File . . . , and select the unzipped SOFT dataset (`GDS112_full.soft`) that was downloaded above.

6. Note that there are a number of comment lines at the start of the file. We can skip over those comments by selecting Show Text File Import Options and setting "Start Import Row:" to 83. Then, click on Refresh Preview.

7. To associate the experimental gene expression data with the network, the same gene identifiers must be used in both datasets. The BioGRID yeast dataset uses Entrez Gene identifiers (IDs) as the node primary identifiers. Conveniently, we already have those data in our SOFT dataset labeled as Gene ID. To link these datasets, select Show Mapping Options and, under "Select the primary key column in table:" and select Gene ID (as shown in Fig. 8.13.1).

8. Select OK to import the data as attributes. The imported data include fold changes at five different time points after shifting the temperature from 30°C to 37°C, as listed in Table 8.13.1. These attributes include not only a number of numeric attributes, but also various symbols. We will use these in the next step.

Filter the network with the genes that have expression data

The BioGRID network contains too many nodes and edges to visually explore effectively. To explore a biologically relevant subset, we can obtain a sub-network using only genes in the experimental data:

9. Click the Select tab on the left side panel (Control Panel) in Cytoscape (see Fig. 8.13.2).

10. Click "+" to create a new filter and select Column Filter. A new selection box should be shown with the label "Choose column . . . ."

11. Click on "Choose column...," select "Node: Gene Symbol," click on the "contains" button, and change it to "matches regex." This allows us to define a regular expression (a way to define text patterns for searching; *http://en.wikipedia.org/wiki/Regular_expression*) as the way to match the genes we would like to select.

12. In the text box, type the term `[A-Z0-9]*`. This selects all entries that have upper-case letters or numbers. Note that all of the gene expression data have gene symbols that match that regular expression, but if the gene only exists in the BioGRID interaction network, that field will be blank (as in Fig. 8.13.2).
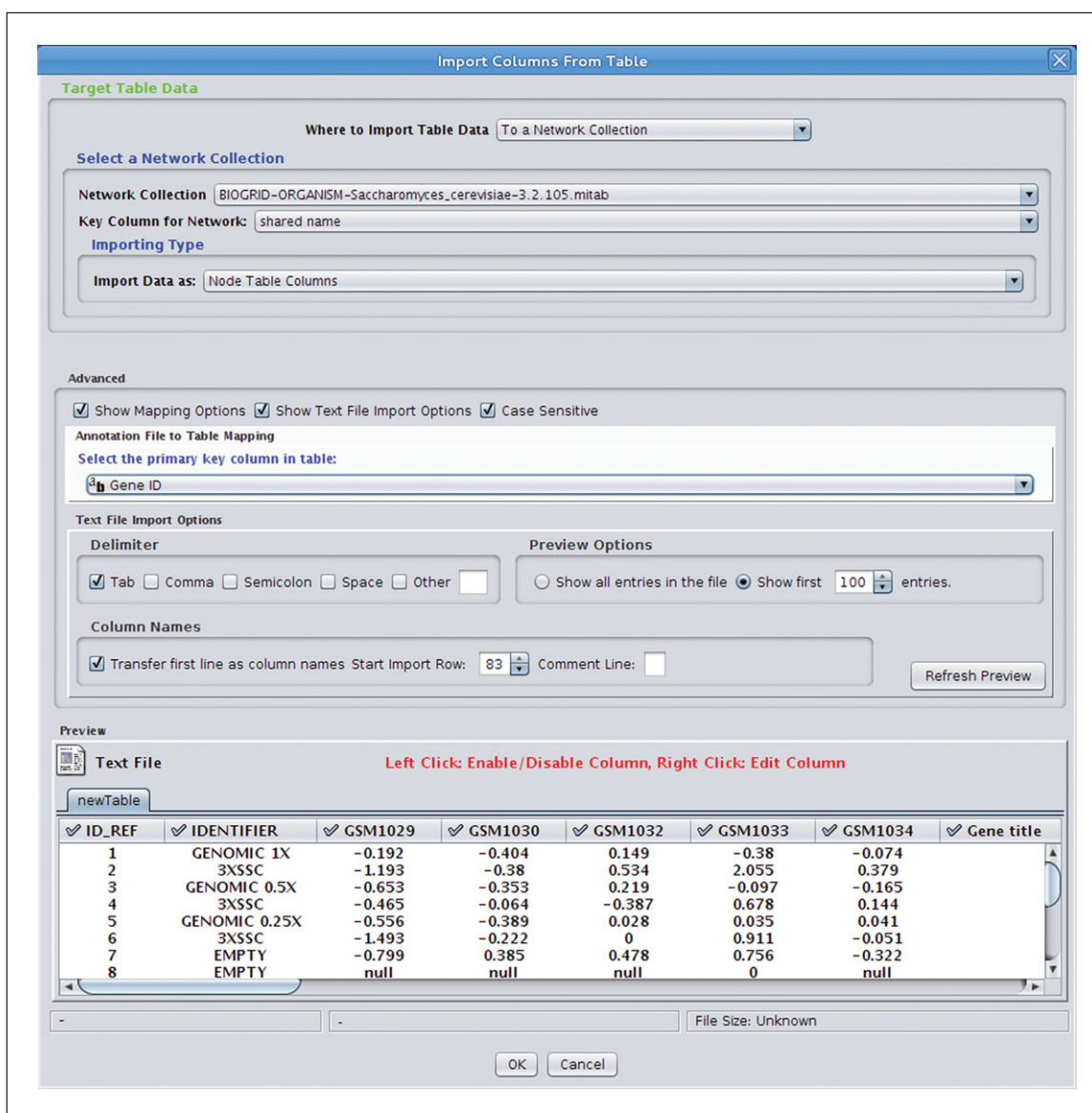
Figure 8.13.1 Prepare parameters to import data into Cytoscape.

Table 8.13.1 Time Points for GDS112 Heat-Shock Accessions

Sample accession	Time point
GSM1029	0 min
GSM1030	5 min
GSM1032	15 min
GSM1033	30 min
GSM1034	60 min

13. Click on Apply to apply the selection. This should select approximately 5,500 nodes, as indicated at the bottom of the Select panel.

Create a new network with the selected subset

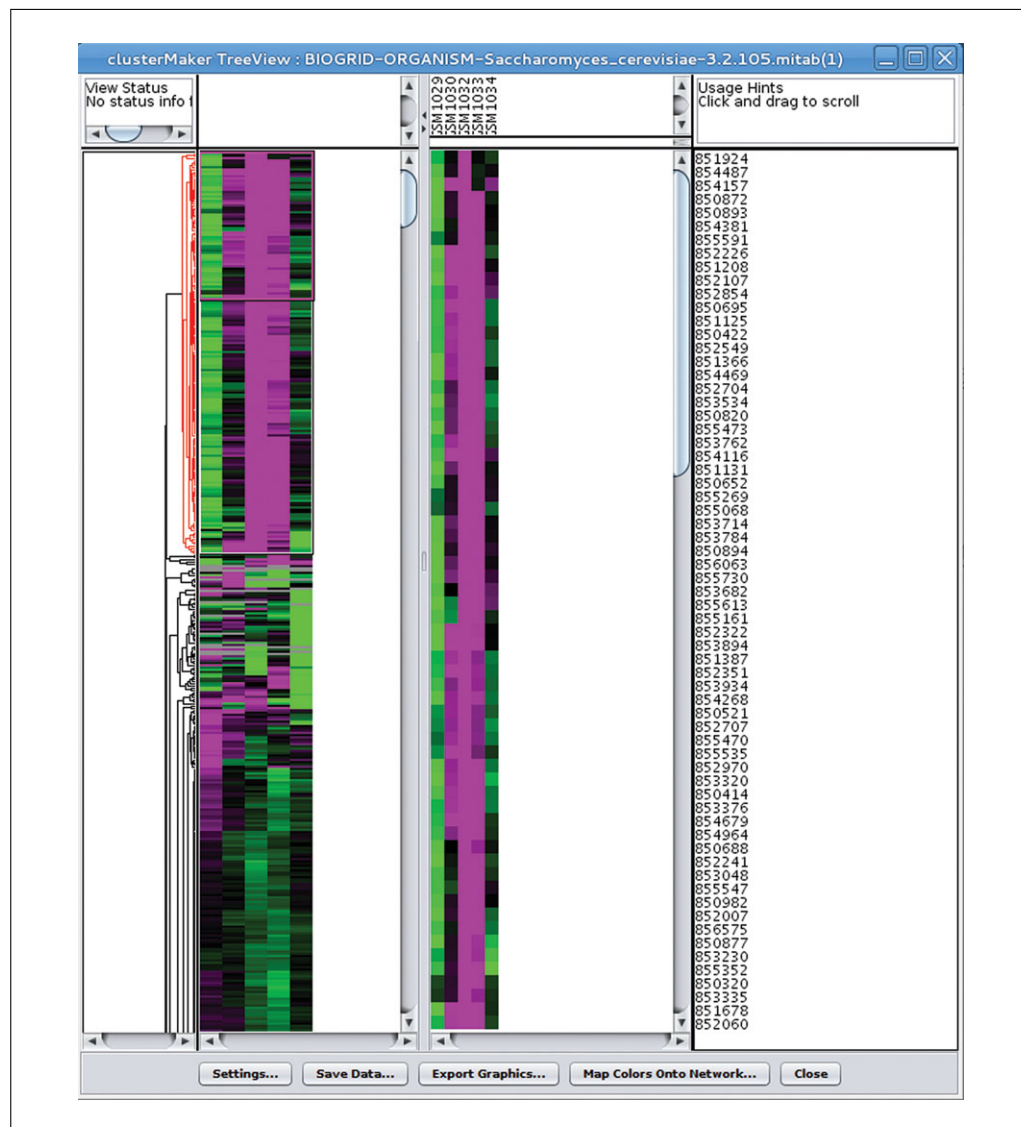Now, with the selected nodes, we would like to create a sub-network from the original BioGRID network:

Figure 8.13.2 Entering regular expressions for Cytoscape.

Figure 8.13.3 Initial display of loaded network.

Figure 8.13.4 Heatmap and dendrogram display.

14. Go to the File menu → New → Network → "From selected nodes all edges" to create a new network. Depending on the default settings, a network view may be created. This may take some time.

15. Click on the Network tab in the Control Panel. There should now be two network entries listed: `BIOGRID-ORGANISM-Saccharomyces_cerevisiae-3.2.105.mitab` and `BIOGRID-ORGANISM-Saccharomyces_cerevisiae-3.2.105.mitab(1)`. The first one is the original network and the second one is the sub-network filtered by the experiment. Right click on the first entry and select Destroy Network to save memory on your system.

16. If a view was not created for the new network, right-click on the new (and only remaining) network and select Create View. Otherwise, just click on that network to select it.

17. Lay out the network by going to Layout → Prefuse Force Directed Layout. This may take some time, after which you can add graphical details by going to View → Show Graphics Details. This may take some time, too. The result should look similar to Figure 8.13.3.
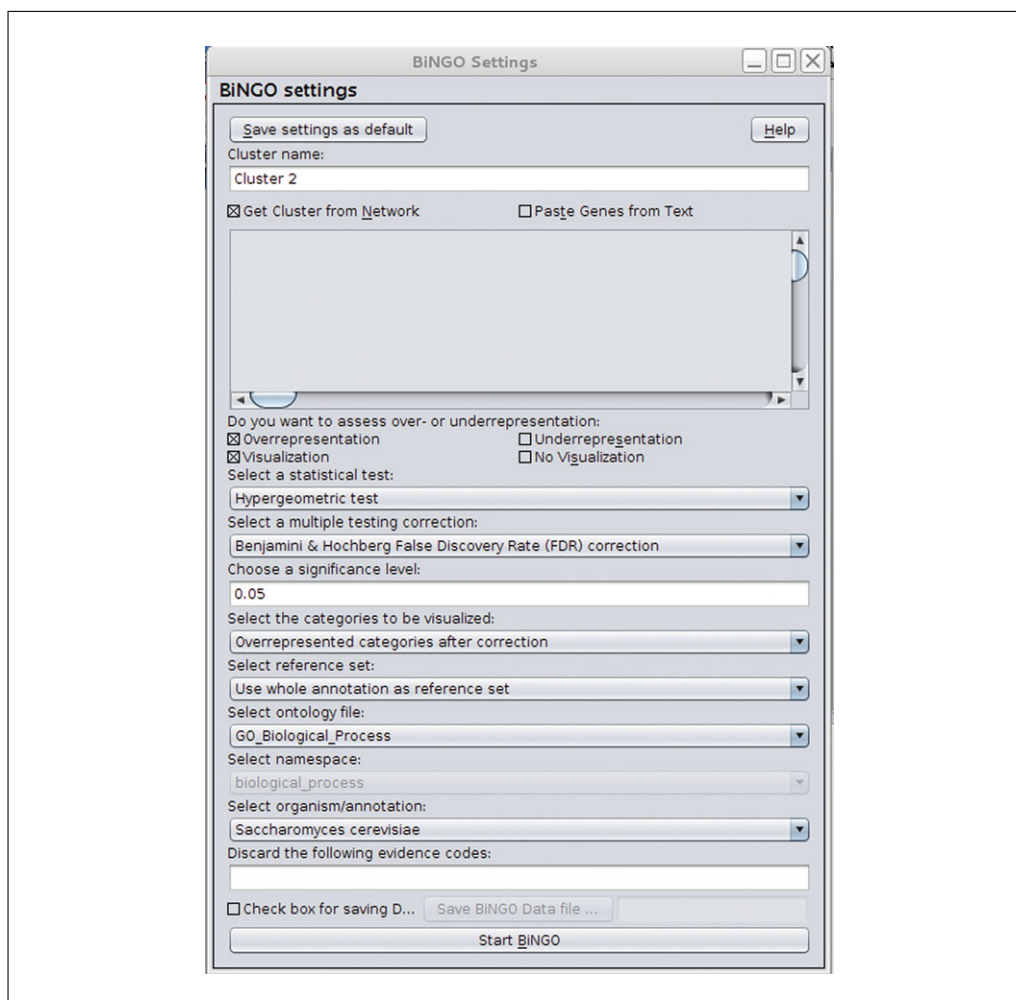
Figure 8.13.5 BiNGO parameter configurations.

18. Go to menu Layout → Bundle Edges → All Nodes and Edges. In the dialog, click OK. Edge bundling is a new Cytoscape feature that simplifies the view of a complex network by 'bundling' edges that are close to each other like ropes. This may take some time.

Using Apps for additional analysis

A unique strength of Cytoscape is its rich collection of Apps that can perform various analyses. We will now use some of these Apps to analyze our data. Ensure that you have the specific App installed (as described in Support Protocol 2) before you attempt the protocol.

Identify network modules

One common task in biological network analysis is to identify clusters (or modules) of biological molecules that share similar properties. For instance, a cluster of genes whose expression changes similarly to external stimuli may have related function and participate in the same biological processes. The clusterMaker App (Morris et al., 2011) provides many frequently used network clustering algorithms. For instance, we can choose to find clusters in a gene network based on expression profiles using hierarchical or K-means clustering, or identify densely intra-connected sub-networks using Markov clustering or community clustering (Su et al., 2010).

19. Launch the clusterMaker App and the hierarchical cluster dialog: Apps → cluster-Maker → Hierarchical cluster.
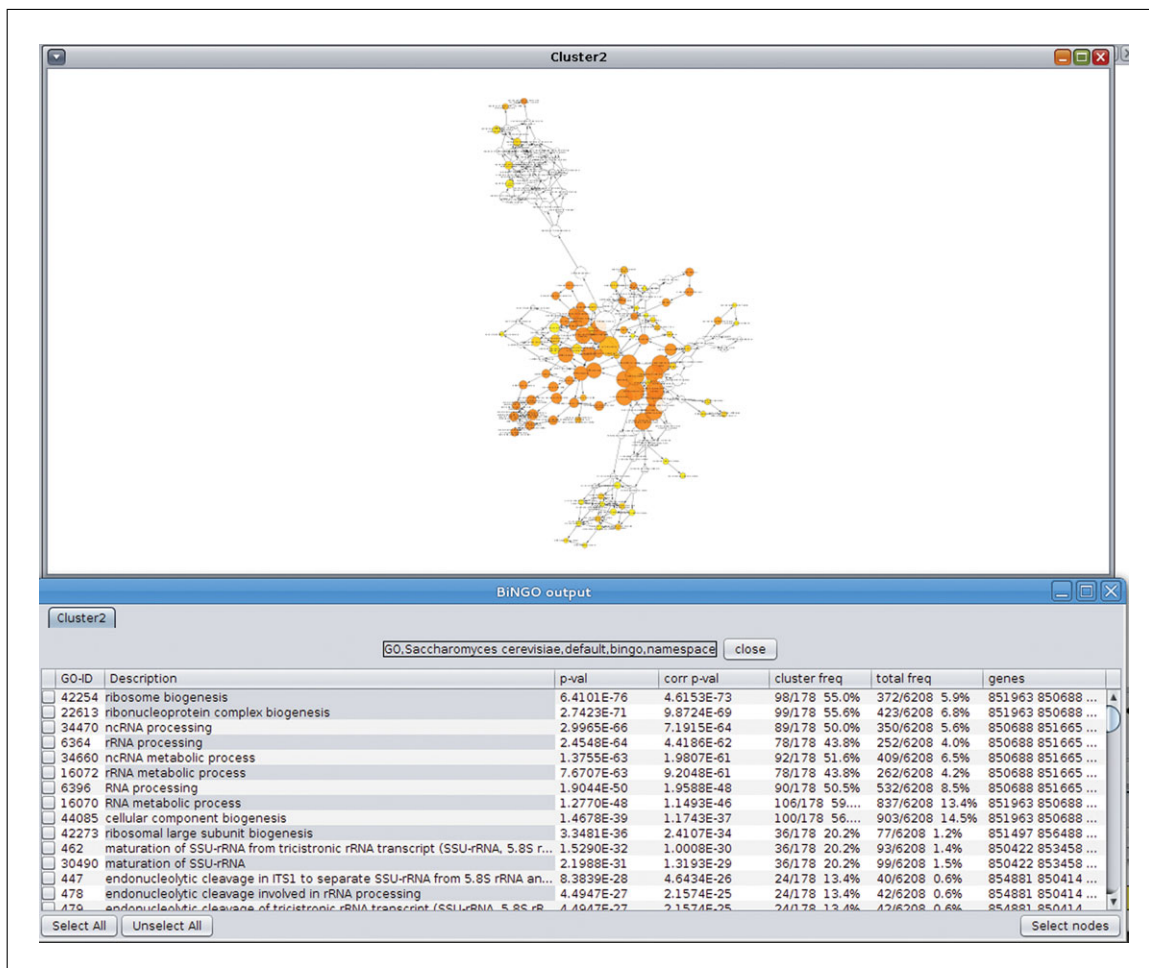
Cluster2

BiNGO output

GO,Saccharomyces cerevisiae,default,bingo,namespace [close]

GO-ID	Description	p-val	corr p-val	cluster freq	total freq	genes
42254	ribosome biogenesis	6.4101E-76	4.6153E-73	98/178 55.0%	372/6208 5.9%	851963 850688 ...
22613	ribonucleoprotein complex biogenesis	2.7423E-71	9.8724E-69	99/178 55.6%	423/6208 6.8%	851963 850688 ...
34470	ncRNA processing	2.9965E-66	7.1915E-64	89/178 50.0%	350/6208 5.6%	850688 851665 ...
6364	rRNA processing	2.4548E-64	4.4186E-62	78/178 43.8%	252/6208 4.0%	850688 851665 ...
34660	ncRNA metabolic process	1.3755E-63	1.9807E-61	92/178 51.6%	409/6208 6.5%	850688 851665 ...
16072	rRNA metabolic process	7.6707E-63	9.2048E-61	78/178 43.8%	262/6208 4.2%	850688 851665 ...
6396	RNA processing	1.9044E-50	1.9588E-48	90/178 50.5%	532/6208 8.5%	850688 851665 ...
16070	RNA metabolic process	1.2770E-48	1.1493E-46	106/178 59....	837/6208 13.4%	851963 850688 ...
44085	cellular component biogenesis	1.4678E-39	1.1743E-37	100/178 56...	903/6208 14.5%	851963 850688 ...
42273	ribosomal large subunit biogenesis	3.3481E-36	2.4107E-34	36/178 20.2%	77/6208 1.2%	851497 856488 ...
462	maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S r...	1.5290E-32	1.0008E-30	36/178 20.2%	93/6208 1.4%	850422 853458 ...
30490	maturation of SSU-rRNA	2.1988E-31	1.3193E-29	36/178 20.2%	99/6208 1.5%	850422 853458 ...
447	endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA an...	8.3839E-28	4.6434E-26	24/178 13.4%	40/6208 0.6%	854881 850414 ...
478	endonucleolytic cleavage involved in rRNA processing	4.4947E-27	2.1574E-25	24/178 13.4%	42/6208 0.6%	854881 850414 ...
479	endonucleolytic cleavage of tricistronic rRNA transcript (SSU-rRNA, 5.8S_rR...	4.4947E-27	2.1574E-25	24/178 13.4%	42/6208 0.6%	854881 850414 ...

[Select All] [Unselect All]   [Select nodes]

Figure 8.13.6 Force-based view of a Cytoscape network.

First, apply Support Protocol 2 to install the clusterMaker App if you have not already done so.

20. Select all of the expression data columns: GSM 1029, GSM 1030, GSM 1032, GSM 1033, and GSM 1034. Since these are time-series data, we probably do not want to cluster the attributes, so deselect "Cluster attributes as well as nodes" and select "Show TreeViewer when complete." Select 'OK'.

21. A clustered heatmap should now be shown, similar to Figure 8.13.4. Some well-defined clustering patterns can be identified. The first cluster has a single protein (HSP12), but the next cluster contains 178 genes. The corresponding nodes in a cluster can be selected by clicking on the horizontal lines in the dendrogram (as shown in Fig. 8.13.4). These 178 genes are all characterized by a tendency to have elevated expression at the time of the temperature change and significantly decreased expression after 15 min. Some of the other genes also show a tendency towards decreased expression after 30 min.

Perform an enrichment analysis using BiNGO

When we obtain gene clusters from a network, a natural follow-up question is: how do these clusters map to known gene function? BiNGO is a Cytoscape App that identifies statistically over-represented Gene Ontology (GO) gene function annotation terms in a gene set or sub-network (Maere et al., 2005). We will now use BiNGO to identify enriched functions in the previously identified clusters.

22. Launch the BiNGO App (Apps → BiNGO).

23. Select a name for the cluster (e.g., Cluster 2) and make sure the Get Cluster from Network option is selected (as in Fig. 8.13.5).

24. Click on Start BiNGO to run BiNGO. The results are displayed both as a table (ordered by p value of term enrichment) and a network of ontology terms where the node color represents the p value of the over-represented terms (as in Fig. 8.13.6). As might be expected, the top-scoring p values are all related to ribosome biogenesis and RNA processing. As the cell is shocked, the first step is to ramp up its ability to make proteins to respond to the new conditions. Later on in the cell's response, the cell no longer requires additional ribosomes and RNA-processing machinery, so it ramps down the expression of these genes.

EXPLORE A HUMAN DISEASE NETWORK

In this protocol, we import data from the human disease network constructed by Goh et al. (2007). This study constructed a global association network between diseases and genes using curated mutation data from the Online Mendelian Inheritance in Man (OMIM; Hamosh et al., 2005) database. A human disease network (HDN) was constructed by connecting diseases that share the same gene mutations, and a disease-gene network (DGN) was constructed similarly via associated diseases. Some of the resulting functional modules were interpreted quantitatively using microarray and protein-protein interaction networks. We will now explore the network from this paper.

For necessary resources, see Support Protocol 1.

Obtain human disease network dataset

1. Use a browser to load the *http://www.barabasilab.com/pubs/CCNR-ALB_ Publications/200705-14_PNAS-HumanDisease/Suppl/* Web page.

2. Download the following:

   Supporting Table S2: Network characteristics of diseases (*http://www.barabasilab.com/pubs/CCNR-ALB_Publications/200705-14_PNAS-HumanDisease/Suppl/supplementary_tableS2.txt*)
   Supporting Table S3: Network characteristics of disease genes (*http://www.barabasilab.com/pubs/CCNR-ALB_Publications/200705-14_PNAS-HumanDisease/Suppl/supplementary_tableS3.txt*)
   Supporting Table S4: List of human protein-protein interactions (*http://www.barabasilab.com/pubs/CCNR-ALB_Publications/200705-14_PNAS-HumanDisease/Suppl/supplementary_tableS4.txt*).

3. Download Network data: Human Disease Network (*http://www.barabasilab.com/pubs/CCNR-ALB_Publications/200705-14_PNAS-HumanDisease/Suppl/disease.net.w*) and Disease Gene Network (*http://www.barabasilab.com/pubs/CCNR-ALB_Publications/200705-14_PNAS-HumanDisease/Suppl/gene.net.w*). Rename the human disease network file from `disease.net.w` to `disease.net.txt`, and the human gene network file from `gene.net.w` to `gene.net.txt`.

Explore the protein-protein interaction network

4. Launch Cytoscape 3.

5. Go to menu File → Import → Network → File, choose `supplementary_tableS4.txt`.

6. The file import dialog should be shown.

Figure 8.13.7 Network import panel parameters.

7. In the Delimiter tab, check only Tab.

8. In the Column Names tab, check "Transfer first line as column names."

9. In the Start Import Row text box, click the spinner to change the number to 2.

    We want to skip the comment lines.

10. Click on Refresh Preview.

11. In the Interaction Definition tab, choose Column 2 as the Source Interaction, and choose Column 4 as the Target Interaction.

12. In the Interaction Type column, choose Column 5.

    There are three interaction sources: R and S indicate two literature sources, and L indicates literature curation.

13. Check the Show Text File Import Options checkbox.

14. If any of the columns are shown as blue, click on the column header to make them gray.

    We do not want to import the gene IDs as edge attributes.

15. If all settings are correct, the import dialog should look similar to Figure 8.13.7. Click on the OK button.

**Biological
Network
Exploration with
Cytoscape 3**

8.13.10

Supplement 47

Current Protocols in Bioinformatics
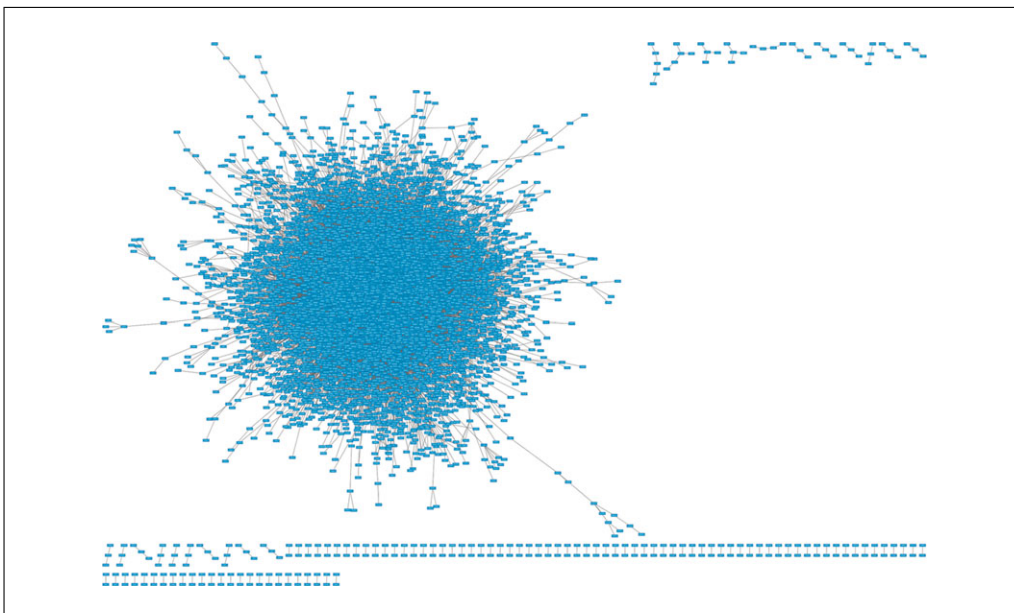
Figure 8.13.8 Force-based view on a very large network.

Layout and overlay information on the network

When a network is imported from a file, it has no layout information—by default, Cytoscape lays out the network as a grid. We will demonstrate different layouts.

16. Go to menu Layout → Apply Preferred Layout (or hit the corresponding button in the toolbar) to apply a force directed layout. Depending on your computer, this step may take a few minutes.

    Your view should look similar to Figure 8.13.8.

17. Click on the Styles tab to reveal the visual styles manager in the Control Panel.

18. Click the Edge tab at the bottom to reveal edge attribute mappings. Click on the Properties drop-down button to reveal the pop-up menu. Click on "Show all" to reveal all possible mappings.

19. Scroll down to reveal Stroke Color (Unselected). Click on the arrow on the right to reveal possible options.

20. In the Column field, click and select "interaction."

21. In the Mapping Type field, choose Discrete Mapping.

22. Choose a color for each of the interaction data sources for L, R, and S. For example, you can use Red, Green, and Blue for the three data sources and use the mixed colors Teal, Purple, Yellow, and Gray for LR, RS, and so on.

23. Now, let us make the nodes a little bit more transparent to reveal the distribution of data sources. Click on the Node tab. Click on the Properties drop-down menu and click on Show All. Scroll to the Fill Color tab and choose a shade of gray. Scroll to Transparency and click on 255, then change the value to 100.

    Your view should look similar to Figure 8.13.9.

24. Cytoscape sometimes hides labels, node graphics, and other information to improve visualization speed. You can force Cytoscape to draw all Graphics detail by going to menu View → Show Graphics Details.
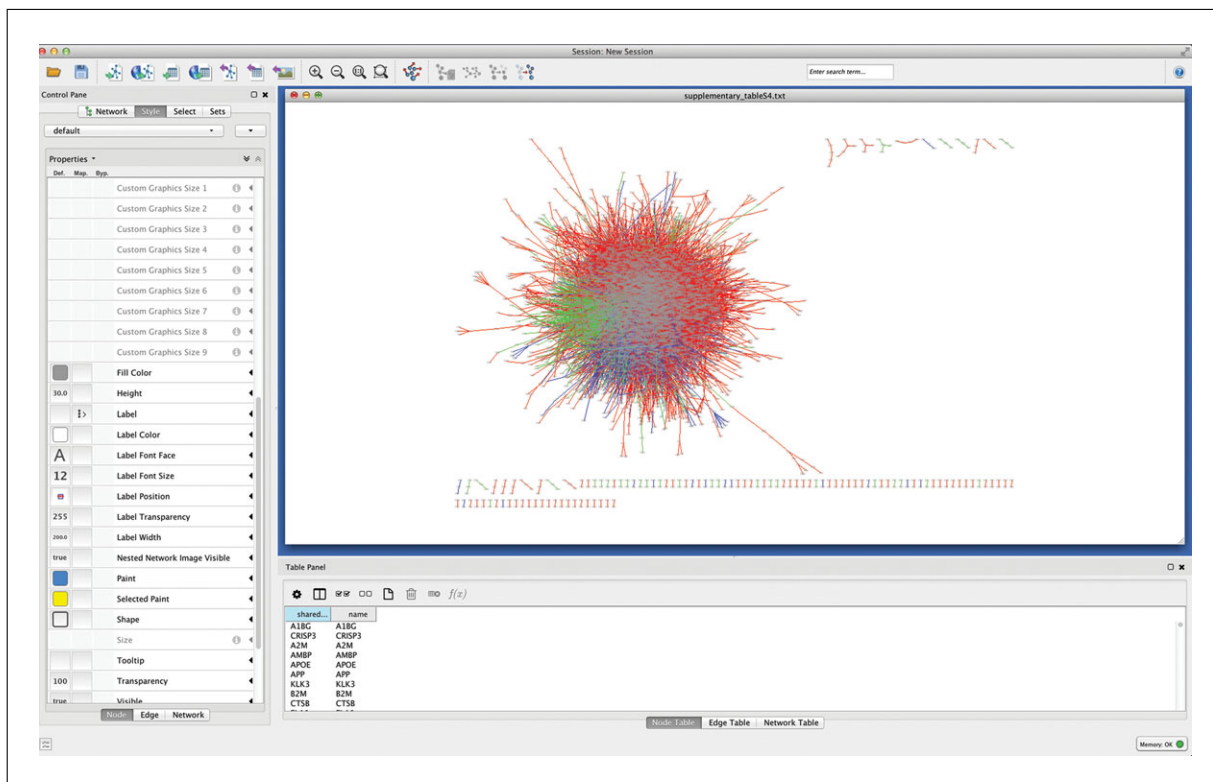
Analyzing Molecular Interactions

8.13.11

Figure 8.13.9 Using VizMapper to add annotation with attributes.

You can see that even though this is a very dense network, protein interactions from the same source tend to cluster with each other.

Discover local gene clusters using MCODE

Identifying densely connected nodes (e.g., genes) from a very densely connected network is useful for identifying biological modules, such as complexes, pathways, or other related sets of nodes. MCODE (Bader and Hogue, 2003) is one of many Cytoscape Apps that identifies local clusters that can be used to identify interesting modules.

25. Launch MCODE by going to menu Apps → MCODE → Open MCODE. The MCODE panel should now be visible in the control panel.

26. In the Find Cluster(s) tab, choose "in Whole Network," then click on "Analyze current network."

27. After some processing, the MCODE panel should open, and the identified clusters should be displayed.

28. Each of the clusters can be exported as a sub-network. Select the second cluster and click on the Create Sub-Network button. You should now see a network like Figure 8.13.10. There are mostly LSM genes, associated with small nuclear RNAs. In this case, MCODE identified a functionally homogenous cluster from a large interaction network based purely on how densely interconnected the nodes were.

29. You can grow or shrink the discovered local clusters by adjusting the Size Threshold slider. The clusters will change accordingly with regard to the 'seed' genes. Drag the slider several notches and the cluster will expand. You can create a new sub-network and check the functional term association using BiNGO following similar procedures in Basic Protocol 1.
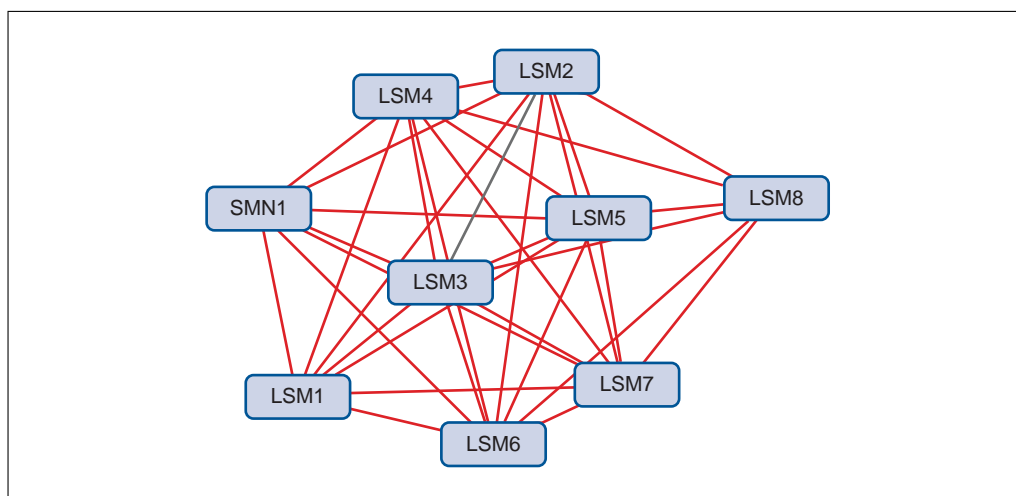
Biological Network Exploration with Cytoscape 3

8.13.12

Figure 8.13.10 A sub-cluster of LSM genes.

Figure 8.13.11 Attribute import configurations.

Explore the human disease network

Although Cytoscape is often used to visualize gene, protein, and metabolic networks, it can be used to visualize other biomedical networks as well. In the following example, we will illustrate how to visualize the human disease network.

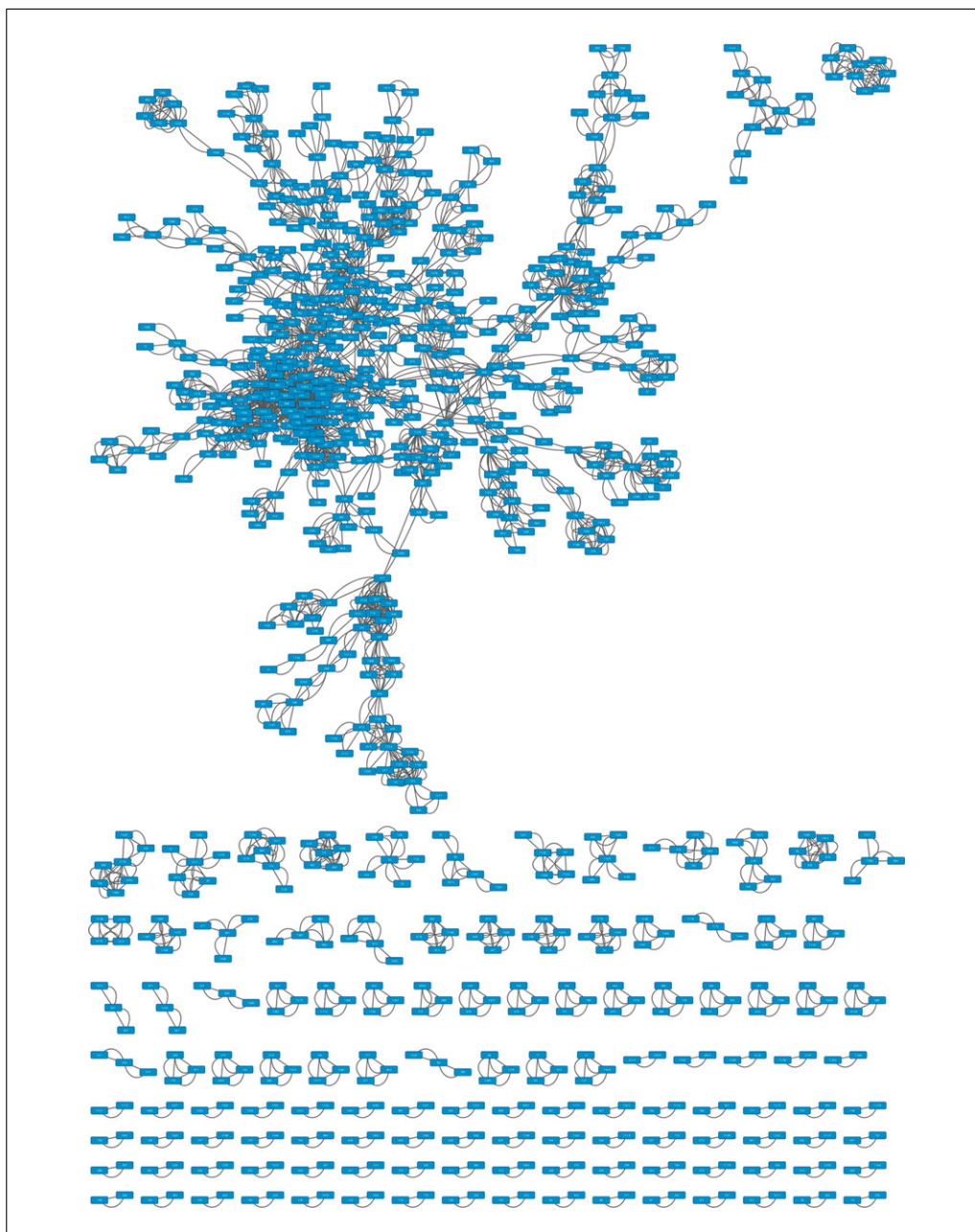30. Start a new session in Cytoscape by going to File menu → New → Session.

Figure 8.13.12 Force-based view of the human disease network (HDN).

31. Go to File menu → Import → Network → File, choose the previously downloaded `disease.net.txt`.

32. In the Interaction Definition tab within the Import Network From Table dialog, choose Column 1 as the Source Interaction and Column 2 as the Target Interaction.

33. Check Show Text File Import Options in the Advanced tab, and check "Transfer first line as column names."

34. Click on the Weight column header in the newTable preview to use Weight as an edge attribute. Your import dialog should look like Figure 8.13.11.

35. Click on the OK button.

36. Go to Layout → Apply Preferred Layout to apply a layout to the network. Your network view should be similar to Figure 8.13.12.
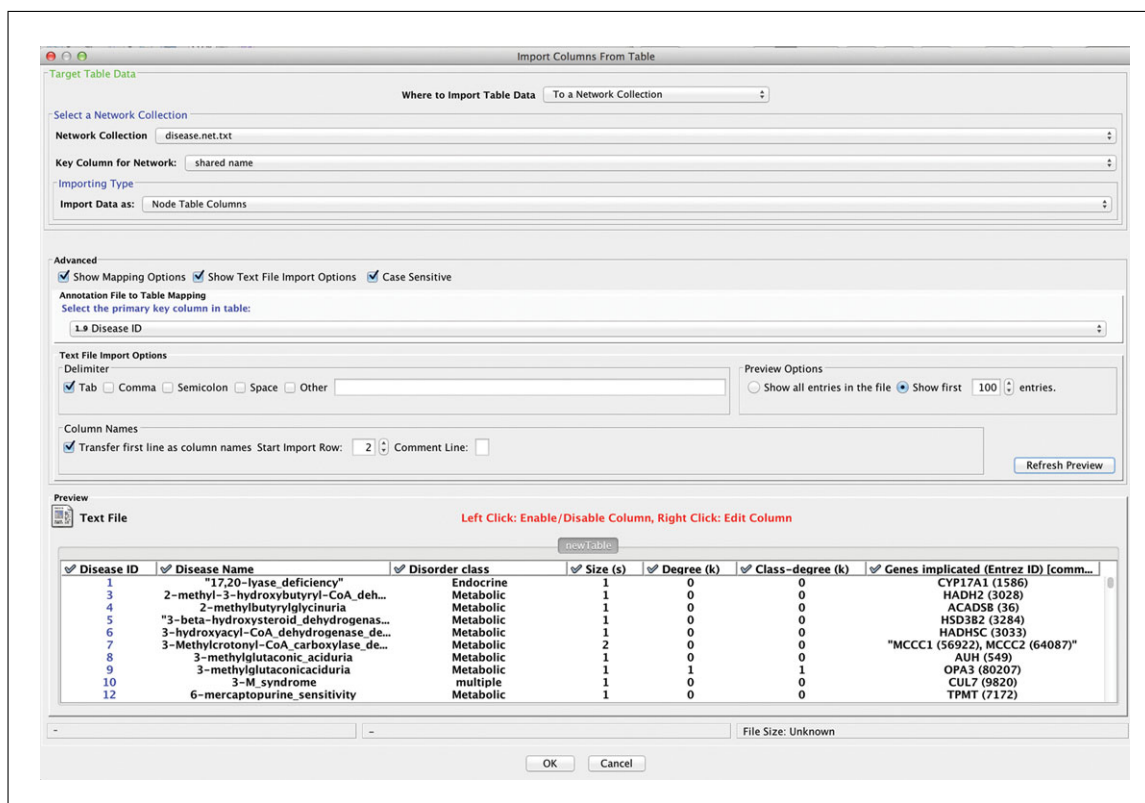
Figure 8.13.13 Node attribute import parameters.

Figure 8.13.14 Using attribute table to verify data were imported correctly.

37. Note that there are many duplicate edges. The diseases are linked by shared gene mutations and each link is documented twice in this file. We can remove the duplicated edges by going to menu Edit → Remove Duplicated Edges. Click on the network where you would like to remove duplicated edges, and then check "Ignore edge direction." Click on the OK button.

If the view is not automatically refreshed, pan or zoom the network view and the duplicate edges will disappear.

We can import additional attribute data to overlay on the disease network, as described in the following steps

38. Open `supplementary_tableS2.txt` with a text editor. On the second row, change `Name` to `Disease Name`.

Cytoscape attribute tables come preconfigured with a `Name` column when importing a network—importing this attribute again may cause a conflict.
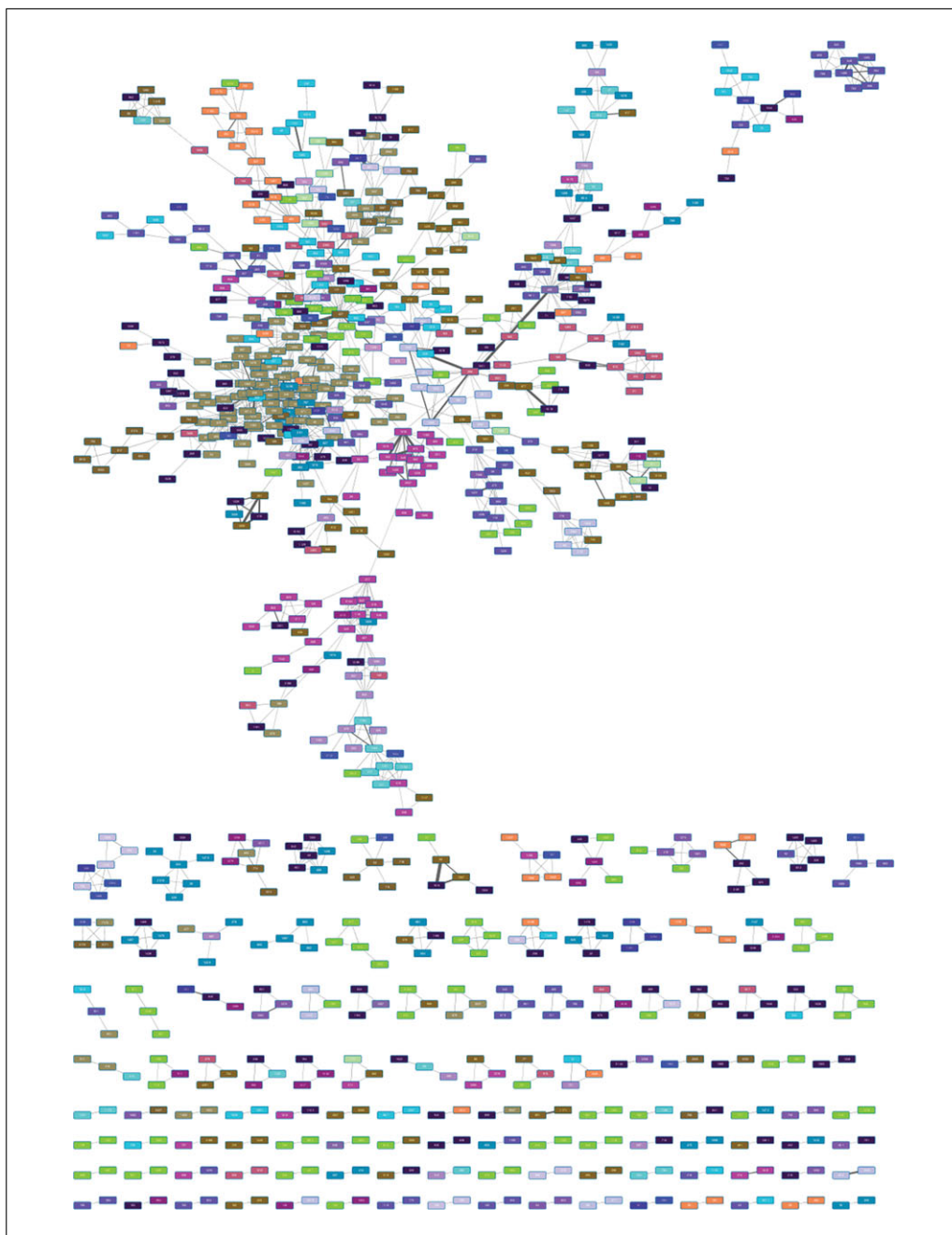
Figure 8.13.15 Visualize closely associated human diseases using attributes and VizMapper.

39. Go to File menu → Import → Table → File, and choose the previous downloaded `supplementary_tableS2.txt`.

40. In the Import Column From Table dialog:

   a. Make sure the disease network is selected in the Network Collection combo.
   b. Check Show Mapping Options and Show Text File Import Options in the Advanced tab.
   c. In the Text File Import Options tab, check "Transfer first line as column names," set Start Import Row to 2, then click on the Refresh Preview button.
   d. Select "1..9 Disease ID" for "Select the primary key column in table."
   e. Your import dialog should look like Figure 8.13.13. Click the OK button.
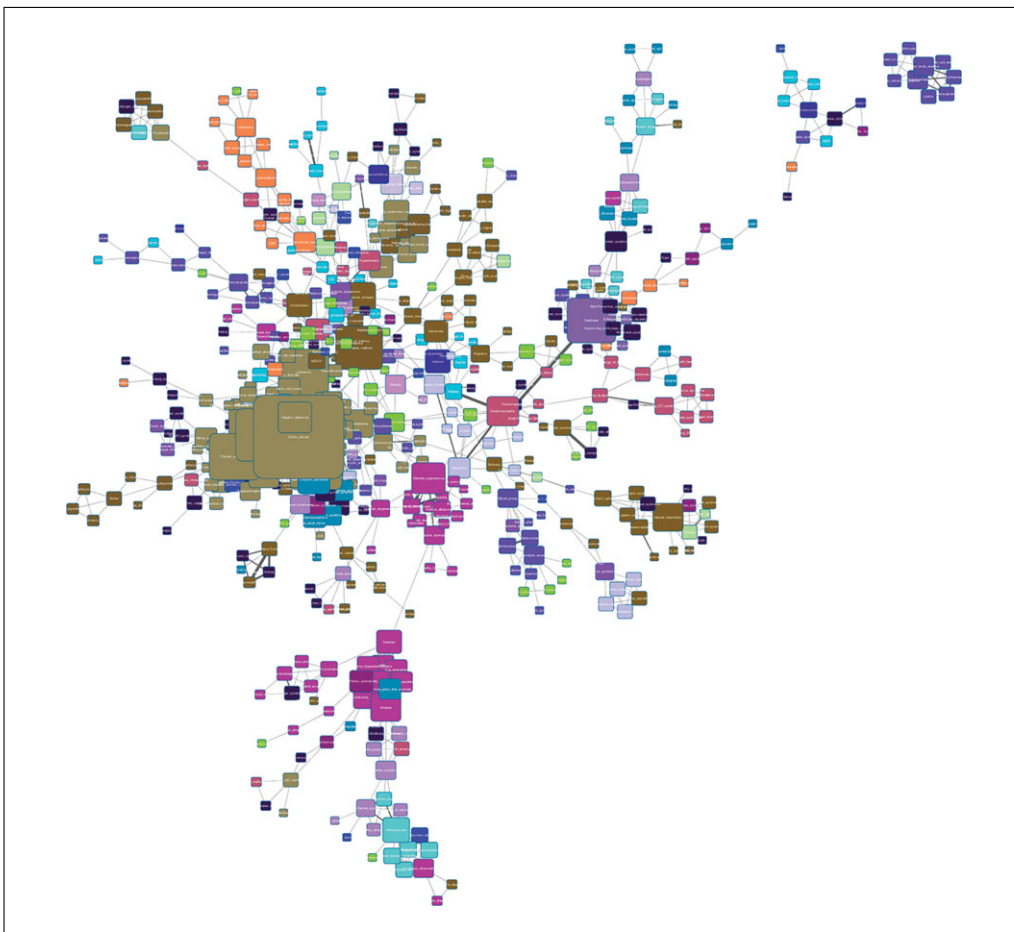
Figure 8.13.16 Fully annotated Human Disease Network (HDN).

41. After import, in the Table Panel, click on the first button, Change Table Mode, and check "Show all" in the pop-up menu. You can check the imported disease attributes (as in Fig. 8.13.14). Click on the Change Table Mode button again to switch back to show only attributes associated with selected nodes.

42. To search for diseases with keywords, you can type the search terms in the top right corner of the toolbar. Type `Neuro*` to select all diseases associated with Neurological disorders—they are quite close to each other in the network.

We can overlay the network with the attributes using visual styles, as in the following steps

43. Click on the Style tab to edit your visual styles.

44. Click on the Node tab to reveal Node properties.

45. Click on the Properties drop-down menu, then click on Show All.

46. Scroll to Fill Color. In the Column field, choose "Disorder class," and then choose Discrete Mapping.

47. Since the network contains many disease classes, we can right click and choose Mapping Value Generators → Random Color in the pop-up menu. Each disease class is then automatically assigned a random color as node fill.

48. Click on the Edge tab. Click on the Properties button, then click on Show All. Find the Width field, and choose Weight for the Column field. This value indicates how many gene mutations different diseases share. Choose Discrete Mapping, right click to

activate the pop-up menu, choose Mapping Value Generators, then choose Number Series. Set the start width to 1, and set the increment to 3. The network should look like Figure 8.13.15. Disorders that are connected with thick edges indicate more shared genetic mutations. For example, the "Leigh_syndrome" is closely connected to "mitochondrial_complex_deficiency."

49. Click on the Node tab. Find Label, and choose Disease_Name.

    Now the disease name will be displayed in each disorder instead of disease IDs.

50. Double-click on Height, and choose "Degree (k)." Choose Continuous Mapping. Double-click on the mapping bar and drag vertical ticks to make the node size range from 30 to 300 (mapped to degree 1 to 50). Repeat this step for Width.

51. We could now reproduce the visual style used in the poster at *http://www.barabasilab.com/pubs/CCNR-ALB_Publications/200705-14_PNAS-HumanDisease/Suppl/Goh_etal_poster.pdf*. Your network view should look similar to Figure 8.13.16. The two selected nodes are Leukemia and Deafness classes, respectively.

    Diseases in the same class tend to be placed near each other and form clusters that share similar gene mutations.

SETTING UP CYTOSCAPE

Cytoscape 3 was developed on Oracle Java and can be run on most operating systems (OS).

Necessary Resouces

Hardware

Minimum requirements: 1-GHz CPU or higher, dedicated graphics card, 500 MB hard drive space, 1 GB free physical memory, and a display that supports 1024 × 768 or higher resolution. Certain Apps such as literature search or pathway retrieval require an Internet connection.

Recommended requirements: dual-core or quad-core CPU at 2 GHz or higher, dedicated graphics card with 512 MB or more video memory, 1 GB or more available hard-drive space, 4 GB or more physical memory, two high-definition (HD) displays (1920×1080 or 1366×768), and high-speed Internet connection. For best performance, we recommend using a solid-state disk (SSD) instead of a hard disk.

Software

Operating system (OS):
Windows: Windows 8, 7, XP, or Vista; 64-bit OS is recommended for large networks
Macintosh: OS X 10.7 or later (with Intel CPUs); 10.8 or later recommended
Linux: Ubuntu 13.x, 12.x, 11.x, or Fedora
Java Runtime Environment (JRE):
64-bit Java Virtual Machine (JVM) is recommended—the latest Oracle distribution can be found at *http://www.java.com/en/*. As of this writing, Cytoscape has been tested with Java 6 and 7, but not later versions. For Windows systems, the 32-bit JVM is supplied at *java.com* by default—the 64-bit version is available at *http://java.com/en/download/manual.jsp* and is recommended. For Linux systems, Oracle Java 7 has caused Cytoscape crashes on some platforms, and OpenJDK7 (*http://openjdk.java.net*) is a good alternative. For Mac systems, the Cytoscape installer automatically triggers JVM6 to install if no other Java version is present.

For additional information, select the Release Notes button on the Cytoscape Web site (*http://cytoscape.org*).

NOTE: If your computer does not meet the above requirements, earlier versions of Cytoscape are available from the Cytoscape Web site, and they may run on older computer hardware.

1. Use a Web browser to load the Cytoscape Web page (*http://cytoscape.org*) and select the Download Cytoscape button.

2. Fill in your name, organization and e-mail address to register as a user (user registration helps with Cytoscape grant renewals). Check to agree to the terms of use (Lesser GNU Public License (LGPL) and optionally join the e-mail list. Click on the Proceed to Download button.

3. Click on the name of the distribution that matches your operating system (OS).

    Choose Windows 32-bit only if you have 32-bit Windows (e.g., Windows XP).

4. Execute the downloaded Cytoscape bundle to install

    The instructions for this depend on your Web browser.

5. Launch Cytoscape (the instructions for this depend on your operating system). For Mac or Linux, double-click on the Cytoscape icon in the installation folder. For Windows, open the Cytoscape folder through the Start Button and All Programs list, then click on the Cytoscape icon.

6. The Cytoscape desktop and the welcome screen should now appear.

SEARCH AND INSTALL APPS

Cytoscape Apps are optional extensions to the Cytoscape software that provide specific additional features (in Cytoscape 2.x, Apps were called *plugins*; Cytoscape 2.x plugins are not compatible with Cytoscape 3.x.). Many of these Apps are available for public browsing and download at the Cytoscape App Store (*http://apps.cytoscape.org/*). Apps can be searched by keyword or browsed by function tags. Each App has its own App page, containing a feature description, details on usage, download statistics, links to available tutorials, and user rating information. Apps can be installed directly from the App Store or from within Cytoscape (using the Cytoscape App Manager).

To install an App (e.g., the BiNGO gene set enrichment analysis App) directly from the App Store website:

1. Launch Cytoscape and keep it running.

2. Use a browser to load the BiNGO Web page: *http://apps.cytoscape.org/apps/bingo*.

3. Click on the Install button.

4. A dialog should pop out showing the progress.

5. When installed, the button on the BiNGO Web page will change to Installed.

To install an App (e.g., the MCODE network module detection App) from within Cytoscape using the Cytoscape App Manager

6. Launch Cytoscape. Close the Welcome screen if it is still visible.

7. Go to the Apps menu → App Manager.

8. In the "Search:" text box, type MCODE.

9. Select the MCODE App.

10. Click on the Install button.

11. When installed, the Installed button will become gray.

> *Multiple Apps can be installed one after the next. For instance, you can follow either of the above protocols to install the clusterMaker network clustering App and the enhancedGraphics App that provides additional rich information-visualization features for Cytoscape network nodes.*

> IMPORTANT NOTE: *Apps should be tested after being installed—before installing another App—to enable tracing any issue that may arise from a particular App.*

12. When installation is complete, click on the Currently Installed tab. All installed Apps should be shown.

GUIDELINES FOR UNDERSTANDING RESULTS

The protocols provided here can stand alone as methods for analyzing biological networks and also serve as a starting point for more in-depth analysis using various Cytoscape analysis and visualization Apps. The two basic protocols have focused on protein-protein interaction networks, but Cytoscape has been used to explore structural networks (Morris et al., 2007; Doncheva et al., 2011), protein-protein similarity networks (Atkinson et al., 2009), and biological pathways (see the Wikipathways App, the CyKEGG Parser App, and the MetScape App; Gao et al., 2010). If the starting point is a list of genes and no source network is available, Apps like the AgilentLiteratureSearch App (Vailaya et al., 2005) and GeneMANIA (Montojo et al., 2010) can provide useful starting points, as well as sources for information to augment existing networks. Additional Apps are available to extend the analysis and visualization we present here and delve further into the biological meaning of the interactions—they are available with documentation at *http://apps.cytoscape.org*.

Basic Protocol 1 demonstrates the annotation of a protein-protein interaction network with expression data, although the approach could be used to annotate networks with a wide variety of additional data. The hierarchical clusters derived from the expression data provide just one approach to exploring this data set. For instance, critical genes and proteins tend to be hubs (nodes connected to many other nodes) or part of the shortest path through the network between two other nodes (Yu, 2007). Various network parameters may be calculated using Cytoscape's built-in NetworkAnalyzer (via Tools → NetworkAnalyzer; Doncheva et al., 2012) or through Cytoscape Apps such as CentiScaPe (Scardoni et al., 2009), which calculates an even larger number of network parameters. In addition to the traditional hierarchical clustering and heat-map visualization, clusterMaker2 may be used to create a co-expression network where the edges between nodes represent expression profile similarities. The resulting similarity network can be further clustered to partition the network into "modules" of genes with similar expression patterns.

To explore the concept of modules in more detail, Cytoscape Apps such as jActiveModules (Ideker et al., 2002) can be used to find subnetworks where nodes show significant changes in expression levels. Unlike the co-expression network approach mentioned above, jActiveModules takes the network topology into account. This can highlight subnetworks where the genes in that subnetwork experience similar expression patterns. A plausible biological explanation for co-expression of genes or proteins is functional relatedness. This is especially true in prokaryotes, where functionally related genes may be organized into the same operons in the genome. Genes involved in a complex can exhibit just-in-time assembly, where one highly regulated critical gene controls the overall activity of the entire complex (de Lichtenberg et al., 2005) Comparing different expression

**Biological
Network
Exploration with
Cytoscape 3**

8.13.20

Supplement 47

Current Protocols in Bioinformatics

patterns across experimental conditions can also reveal different mechanisms that cause the same end result. As we saw, the BiNGO App finds significantly over-represented Gene Ontology terms annotated to the genes of interest. This helps identify functions enriched in a set of genes, including sets of genes that are co-expressed.

Basic Protocol 2 demonstrates the workflow for using Cytoscape to visualize and annotate large biomedical networks. One important and useful Cytoscape feature is its Style Manager (formerly called VizMapper), which allows researchers to translate a variety of attribute data, such as gene expression profiles, functional gene groups and pathways, and protein-protein interaction types, into intuitive graphic representations that facilitate exploratory knowledge discovery. In our examples, we used BiNGO, clusterMaker, and MCODE to identify closely associated gene and protein clusters. These clusters can be immediately visualized in the Network view, which is especially helpful for visualizing and understanding the local topologies and functional features in very large networks such as the Human Disease Network having thousands of nodes and edges. The Cytoscape App store contains many other examples (*http://apps.cytoscape.org/apps/with_tag/datavisualization*) tailored for visualizing data from various biological sources.

In our protocols, we import disease network and additional annotation data containing disease categories, and utilize such information to aid network visualization. Such data can also be imported directly from many external sources. As mentioned above, there are many other Cytoscape 3 Apps that enable data import from data repositories such as Reactome (Joshi-Tope and Gillespie, 2005), KEGG (Kanehisa, 2002), WikiPathways (Kelder et al., 2012), MetScape (Karnovsky et al., 2012), and Agilent Literature Search (Vailaya et al., 2005). Using Cytoscape, users can also integrate their own experimental data with existing network data and functional annotations.

COMMENTARY

Background Information

The term *biological network* usually refers to two types of data: those that are human-curated from the literature and those that are experimentally derived. The former are built on curated and verified knowledge such as those stored in pathway and protein interaction databases. The latter are derived from experiments, such as protein-interaction screens or gene-expression correlations. Combining these two data sources and other functional annotations enables researchers to support their experiments and identify new patterns from the data. Visual exploration tools are required for this, especially if the data are extensive.

The omics era has brought many opportunities and challenges for network analysis. The sharp decline in the cost of high-throughput technology has made it possible to efficiently measure tens of thousands of molecular profiles at once, for hundreds of different sample groups and experimental conditions. Such rich repositories of experimental data, along with the human-curated annotations from the literature, enable researchers to quickly identify novel connections between their observations and existing knowledge, thereby enabling testing of new hypotheses. In addition to the traditional genomics, transcriptomics, and proteomics, accurate metabolomics, phenomics, and lipidomics are also becoming more accessible. Together, these data offer different snapshots of a target organism. Even though robust and scalable computational and statistical methods have been developed to mine new signals, it is often difficult for the researcher to explore such data without high-performance, versatile, and interactive visualization software.

Cytoscape was originally designed as a simple tool to visualize networks with hundreds, or maybe a few thousands of nodes. Thanks to the continuous community support, it has expanded its capabilities and scope to handle bigger, more complex data and evolved into a sophisticated platform that can be used for many network analysis purposes. New features include:

• Improved performance of network layouts and visualizations of very large networks (using edge curving and bundling to reduce edge clutter).

• Flexible and fast search functions that enable the user to quickly find sets of nodes or

edges with custom criteria (such as a molecular function GO term).

- Connectivity to many external data repositories (e.g., Pathway Commons, Reactome).
- A more user-friendly online/local App management mechanism.
- A new network property table that supports complex annotations from a variety of sources.
- Comprehensive support for many network formats for import/export.
- A metanode (subnetwork) mechanism that enables the user to build hierarchy in a network (a network can be a node in another network), enabling a 'high-level', 'functional', or 'structural' view of big networks (instead of a 'flat' hairball of tens of thousands of nodes or edges).

With these new capabilities, users can:

- Explore a network obtained from experiments or annotations in Cytoscape.
- Overlay the nodes or edges with experimental values (using Styles).
- Obtain additional functional annotation from external sources such as Gene Ontology (GO).
- Partition the network for potential functional modules (e.g., using MCODE or clusterMaker Apps).
- Merge a network with existing curated networks (e.g., using Michigan Molecular Interaction, MiMI or GeneMANIA databases).

Our protocols above demonstrate common workflows, though many other workflows are possible. New protocols are regularly posted at *http://tutorials.cytoscape.org*. Also, new Apps are regularly posted to the Cytoscape App store, many enabling new workflows. A good way to find popular Apps is to rank all Apps by popularity (number of downloads) on the App store Web site. From the App store homepage, click All Apps at the top left, then click the "downloads" button at the top to sort Apps by number of downloads.

Critical Parameters and Troubleshooting

Out-of-memory errors

Symptoms: When loading a network from either a database source or from a Cytoscape session file, a "loading" message box is displayed and the progress animation continues, but no network is displayed either in the Networks tab or a network window. The Memory button in the lower right of the Cytoscape window contains Low instead of OK.

Possible causes: Cytoscape has run out of memory to load the network, or the operating system is swapping RAM to the hard disk because your `Cytoscape.vmoptions` file allocates more RAM to Java than your workstation has.

Remedies: Add more memory, then register the new memory with Cytoscape per the "Note on Memory Consumption" section of the Cytoscape user manual. If you have installed 32-bit Java and 32-bit Cytoscape, and already have 4 GB of RAM, consider using 64-bit Java and 64-bit Cytoscape, adding more RAM.

Mismatch between Java and Cytoscape

Symptoms: When starting Cytoscape on Windows, you receive messages indicating that the JMV could not be found, is defective, or the maximum heap size is too large.

Possible causes: You may have inadvertently installed 32-bit Java, which is the default download from *java.com* on all Windows systems. If you have installed a 64-bit version of Cytoscape, the 32-bit Java is inappropriate.

Remedies: Uninstall 32-bit Java and install 64-bit Java instead. Oracle maintains the latest Java version at *http://java.com/en/download/manual.jsp*. When you restart Cytoscape, you should see its splash screen.

Cytoscape freezes during startup

Symptoms: When starting Cytoscape, the splash screen appears and nothing more happens, or it shows the names of Cytoscape modules being loaded, but then freezes before showing a Cytoscape window.

Possible causes: Cytoscape and its code cache may have become unsynchronized, possibly as a result of installing a newer or older Cytoscape. Alternatively, a new Cytoscape installation could be taking extra time (up to 3 min) to build its code cache.

Remedies: If a Cytoscape windows hasn't appeared after 3 min, use your OS to terminate the executing Cytoscape, then delete the Cytoscape cache directory maintained in your user directory at `<userdir>/CytoscapeConfiguration/3` (for a user named Bob using Windows 7, the directory would be `C:\Users\Bob\Cytoscape Configuration\3`). If restarting Cytoscape fails in the same way, delete the Apps, too, by removing `<userdir>/CytoscapeConfiguration` and restarting Cytoscape again.

No view window after large network load

Symptoms: After loading a large network, the network name appears in the Network tab, but there is no window showing the network.

Possible causes: For large networks, Cytoscape shortens the overall load time by not drawing the network view window.

Remedies: Right click on the network in the Network tab, and choose the Create View menu item. The network window will appear within a few seconds. You can create a more manageable subnetwork by using the procedure in step 16 of Basic Protocol 1.

Data integration errors

Symptoms: Expression or attribute data files are not properly integrated with the loaded network.

Possible causes: The gene identifier columns that synchronize the two files do not match exactly, or the files may not be in the correct format.

Remedies: Use the Node Table or Edge Table tabs in the Table Panel to check that the network identifiers match the identifiers in the expression or attribute data file per the tutorial at *http://opentutorials.cgl.ucsf.edu/index.php/ Tutorial:Network_Loading_And_ID_Mapping*.

Acknowledgements

Literature Cited

Atkinson, H.J., Morris, J.H., Ferrin, T.E., and Babbit, P.C. 2009. Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS One* 4:e4345.

Bader, G.D. and Hogue, C.W.V. 2003. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4:2.

Bader, G.D., Betel, D., and Hogue, C.W.V. 2001. BIND–The Biomolecular Interaction Network Database. *Nucleic Acids Res.* 29:242-245.

Barrett, T., Troup, D.B., Wilhite, S.E., Ledoux, P., Rudnev, D., Evangelista, C., Kim, I.F., Soboleva, A., Tomashevsky, M., and Edgar, R. 2007. NCBI GEO: Mining tens of millions of expression profiles—database and tools update. *Nucleic Acids Res.* 35:D760-D765.

Bastian M., Heymann S., and Jacomy M. 2009. Gephi: An open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media. Association for the Advancement of Artificial Intelligence, Palo Alto, Calif.

Batagelj, V. and Mrvar, A. 1998. Pajek-program for large network analysis. *Connections* 21:47-57.

Chatr-Aryamontri, A. Breitkreutz, B.J., Heinicke, S., Boucher, L., Winter, A., Stark, C., Nixon, J., Ramage, L., Kolas, N., O'Donnell, L., Reguly, T., Breitkreutz, A., Sellam, A., Chen, D., Chang, C., Rust, J., Livstone, M., Oughtred, R., Dolinski, K., and Tyers, M. 2012. The BioGRID interaction database: 2013 update. *Nucleic Acids Res.* 41:D816-D823.

Csardi, G. and Nepusz, T. 2006. The igraph software package for complex network research. *Inter-Journal Complex Systems* 1695.

de Lichtenberg, U., Jensen, L.J., Fausbøll, A., Jensen, T.S., Bork, P., and Brunak, S. 2005. Comparison of computational methods for the identification of cell cycle-regulated genes. *Bioinformatics* 21:1164-1171.

Doncheva, N.T., Klein, K., Domingues, F.S., and Albrecht, M. 2011. Analyzing and visualizing residue networks of protein structures. *Trends Biochem. Sci.* 36:179-182.

Doncheva, N.T., Assenov, Y., Dominiques, F.S., and Albrecht, M. 2012. Topological analysis and interactive visualization of biological networks and protein structures. *Nat. Protoc.* 7:670-685.

Dutkowski, J., Kramer, M., Surma, M.A., Balakrishnan, R., Cherry, J.M., Krogan, N.J., and Ideker, T. 2012. A gene ontology inferred from molecular networks. *Nat. Biotechnol.* 31:38-45.

Dutkowski, J., Ono, K., Kramer, M., Yu, M., Pratt, D., Demchak, B., and Ideker, T. 2014. NeXO Web: The NeXO ontology database and visualization platform. *Nucleic Acids Res.* 42:D1269-D1274.

Dutot, A., Pigne, Yoann, Dutot, Antoine, Guinand, Frederic, and Olivier, Damien. 2007. GraphStream: A tool for bridging the gap between Complex Systems and Dynamic Graphs. *In* Emergent Properties in Natural and Artificial Complex Systems. Satellite Conference within the 4th European Conference on Complex Systems (ECCS'2007).

Edgar, R., Domrachev, M., and Lash, A.E. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30:207-210.

Fisher, D., O'Madadhain, J., Smyth, P., White, S., and Boey, Y-B. 2005. Analysis and visualization of network data using JUNG. *J. Stat. Softw.* 10:1-25.

Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C., and Jensen, L.J. 2013. STRING v9.1: Protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 41:D808-D815.

Gao, J., Tarcea, V.G., Karnovsky, A., Mirel, B.R., Weymouth, T.E., Beecher, C.W., Cavalcoli, J.D., Athey, B.D., Omenn, G.S., Burant, C.F., and Jagadish, H.V. 2010. Metscape: A Cytoscape plug-in for visualizing and interpreting metabolomic data in the context of human metabolic networks. *Bioinformatics* 26:971-973.

Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell.* 11:4241-4257.

Goel, R., Harsha, H.C., Pandey, A., and Prasad, T.S. 2012. Human protein reference database and human proteinpedia as resources for phosphoproteome analysis. *Mol. Biosyst.* 8:453-463.

Goh, K.-I., Cusick, M.E., Valle, D., Childs, B., Vidal, M., and Barabási, AL. 2007. The human disease network. *Proc. Natl. Acad. Sci. U.S.A.* 104:8685-8690.

Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A., and McKusick, V.A. 2005. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.* 33:D514-D517.

Ideker, T., Ozier, O., Schwikowski, B., and Siegel, A.F. 2002. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* 18:S233-S240.

Joshi-Tope, G., Gillespie, M., Vastrik, I., D'Eustachio, P., Schmidt, E., de Bono, B., Jassal, B., Gopinath, G.R., Wu, G.R., Matthews, L., Lewis, S., Birney, E., and Stein, L. 2005. Reactome: A knowledgebase of biological pathways. *Nucleic Acids Res.* 33:D428-D432.

Kanehisa, M. 2002 The KEGG database. *Novartis Found. Symp.* 247:91-101.

Karnovsky, A., Weymouth, T., Hull, T., Tarcea, V.G., Scardoni, G., Laudanna, C., Sartor, M.A., Stringer, K.A., Jagadish, H.V., Burant, C., Athey, B., and Omenn, G.S. 2012. Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data. *Bioinformatics* 28:373-380.

Kelder, T., van Iersel, M.P., Hanspers, K., Kutmon, M., Conklin, B.R., Evelo, C.T., and Pico, A.R. 2012. WikiPathways: Building research communities on biological pathways. *Nucleic Acids Res.* 40:D1301-D1307.

Kerrien, S., Aranda, B., Breuza, L., Bridge, A., Broackes-Carter, F., Chen, C., Duesbury, M., Dumousseau, M., Feuermann, M., Hinz, U., Jandrasits, C., Jimenez, R.C., Khadake, J., Mahadevan, U., Masson, P., Pedruzzi, I., Pfeiffenberger, E., Porras, P., Raghunath, A., Roechert, B., Orchard, S., and Hermjakob, H. 2012. The IntAct molecular interaction database in 2012. *Nucleic Acids Res.* 40:D841-D846.

Lotia, S., Montojo, J., Dong, Y., Bader, G.D., and Pico, AR. 2013. Cytoscape app store. *Bioinformatics* 29:1350-1351.

Maere, S., Heymans, K., and Kuiper, M. 2005. BiNGO: A Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21:3448-3449.

Margolin, A.A., Wang, K., Lim, W.K., Kustagi, M., Nemenman, I., and Califano, A. 2006. Reverse engineering cellular networks. *Nat. Protoc.* 1:662-671.

Merico, D., Gfeller, D., and Bader, G.D. 2009. How to visually interpret biological data using networks. *Nat. Biotechnol.* 27:921-924.

Montojo, J., Zuberi, K., Rodriguez, H., Kazi, F., Wright, G., Donaldson, S.L., Morris, Q., and Bader, G.D. 2010. GeneMANIA Cytoscape plugin: Fast gene function predictions on the desktop. *Bioinformatics* 26:2927-2928.

Morris, J.H., Huang, C.C., Babbitt, P.C., and Ferrin, T.E. 2007. structureViz: Linking Cytoscape and UCSF Chimera. *Bioinformatics* 23:2345-2347.

Morris, J.H., Apeltsin, L., Newman, A.M., Baumbach, J., Wittkop, T., Su, G., Bader, G.D., and Ferrin, T.E. 2011. clusterMaker: A multialgorithm clustering plugin for Cytoscape. *BMC Bioinformatics* 12:436.

Saito, R., Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.L., Lotia, S., Pico, A.R., Bader, G.D., and Ideker, T. 2012. A travel guide to Cytoscape plugins. *Nat. Methods* 9:1069-1076.

Scardoni, G., Petterlini, M., and Laudanna, C. 2009. Analyzing biological network parameters with CentiScaPe. *Bioinformatics* 25:2857-2859.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. 2003. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13:2498-2504.

Su, G., Kuchinsky, A., Morris, J.H., States, D.J., and Meng, F. 2010. GLay: Community structure analysis of biological networks. *Bioinformatics* 26:3135-3137.

Tarcea, V.G., Weymouth, T., Ade, A., Bookvich, A., Gao, J., Mahavisno, V., Wright, Z., Chapman, A., Jayapandian, M., Ozgür, A., Tian, Y., Cavalcoli, J., Mirel, B., Patel, J., Radev, D., Athey, B., States, D., and Jagadish, H.V. 2009. Michigan molecular interactions r2: From interacting proteins to pathways. *Nucleic Acids Res.* 37:D642-D646.

Turinsky, A.L. 2014. Structural Genomics. Humana Press, Totowa, New Jersey.

Vailaya, A., Bluvas, P., Kincaid, R., Kuchinsky, A., Creech, M., and Adler, A. 2005. An architecture for biological information extraction and representation. *Bioinformatics* 21:430-438.

Yu, H. 2007. The importance of bottlenecks in protein networks: Correlation with gene essentiality and expression dynamics. *PLoS Comput. Biol.* 3:e59.

Zuberi, K., Franz, M., Rodriguez, H., Montojo, J., Lopes, C.T., Bader, G.D., and Morris, Q. 2013. GeneMANIA prediction server 2013 update. *Nucleic Acids Res.* 41:W115-W22.

Biological Network Exploration with Cytoscape 3

8.13.24