
# MetDisease User Manual

## An App for Cytoscape

**Department of Computational Medicine & Bioinformatics, University of Michigan**

**3/1/2014**

# Contents



Please note that due to continuous software upgrades, the images in this handout may not exactly mimic what you see on the screen.

# Overview

MetDisease is an app for Cytoscape, the bioinformatics network visualization tool. The app is used to annotate a metabolic network with MeSH disease terms, explore related diseases within a network, and link to PubMed references corresponding to any network node and selection of MeSH terms. MeSH terms are controlled vocabulary terms used by the National Library of Medicine to describe the content of the articles indexed in PubMed.

Users can import and annotate any network where metabolites (compounds) are represented as nodes, referenced by KEGG or PubChem IDs. The edges can be arbitrarily defined by the users. MetDisease allows users to highlight and explore parts of metabolic networks related to one or more MeSH disease terms and provides links to relevant PubMed literature. Users have an option to import their own metabolic networks or to use MetDisease to annotate metabolic networks generated with the Cytoscape app MetScape (Gao, et al., 2010; Karnovsky, et al., 2012).

**Note:** This app requires Cytoscape 3.0 or higher to run correctly.

## Data Source

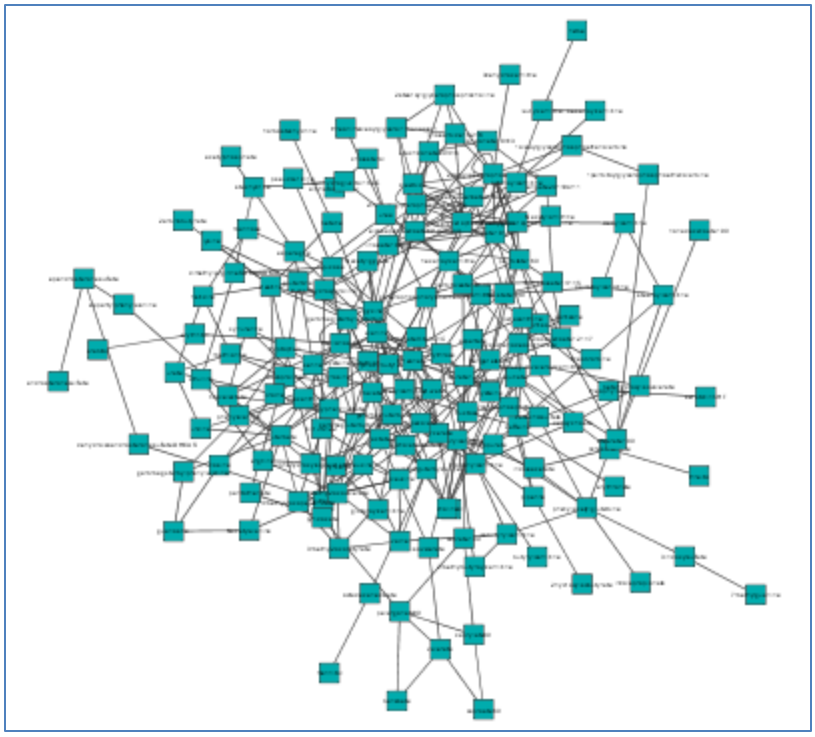
MetDisease uses the Metab2MeSH database (Sartor, et al., 2012). This data set is created twice a year by downloading the PubChem Compound and Substance databases and the NLM PubMed database, parsing them, and loading them into an in-house relational database. Associations between compounds and MeSH terms are calculated using two-sided Fisher's exact tests, and any results with p-value < 0.0001 are retained in the database. MetDisease then uses an internal service to access this database via SQL queries in order to determine relevant MeSH disease terms for the compounds in a given metabolic network.

# Workflow 1: Running MetDisease with a User-Imported Metabolic Network

A network must be built in Cytoscape before using the MetDisease app.

## Example Data

Users can import a metabolic network into Cytoscape and then use MetDisease. For the imported network example in this User Manual, a network was created in Cytoscape using a publicly-available metabolomics data set (Krumsiek, et al., 2012). A subset of known metabolites was downloaded from supplementary data. Adjusted partial correlation coefficient values less than 5e-4 were used to draw the edges in the resulting network. Metabolites are represented as nodes and partial correlation coefficients are represented as edges. You can download this sample input file at http://metdisease.ncibi.org/Fig_1_network.xgmml.



Go to the Opening MetDisease section to begin annotating this network with the MetDisease app. The database identifier for this network is PubChem Id.

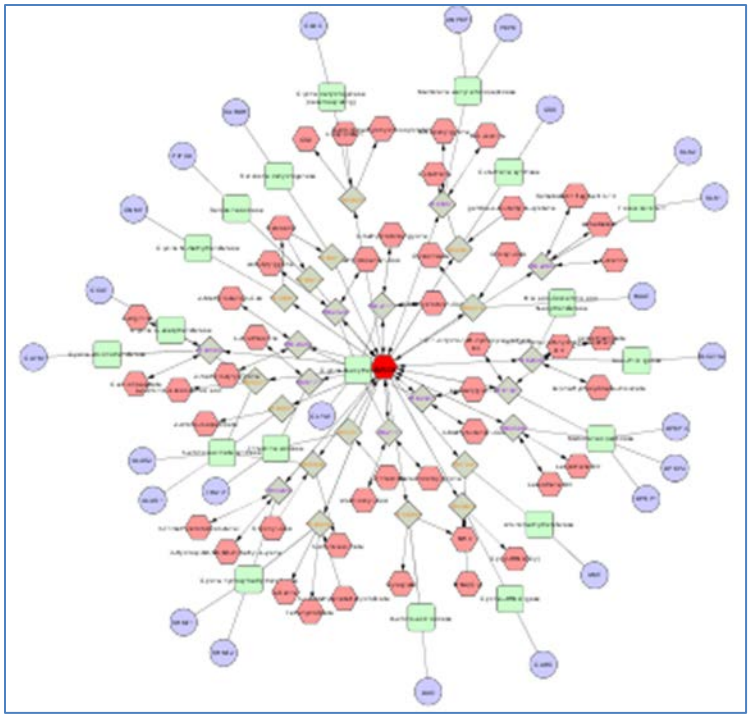
## Workflow 2: Running MetDisease with a Network Created Using the MetScape App

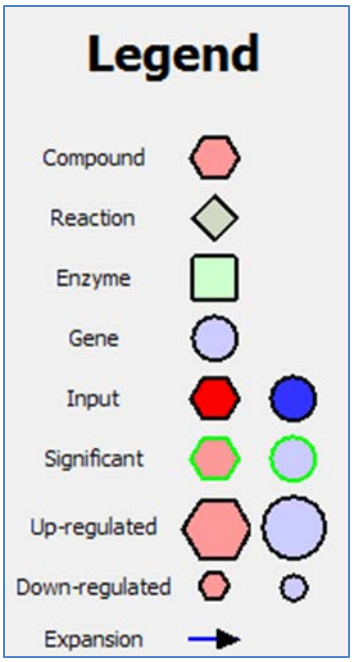
A network must be built in Cytoscape before using the MetDisease app.

### Example Data

For the MetScape-created network example in this User Manual, a network was created based on the compound Glycine using the MetScape app in Cytoscape. The MetScape app must be installed in Cytoscape when using this workflow.

1. Select **Apps -> MetScape -> Build Network.**
2. Select **Human** from the dropdown menu next to **Organism**.
3. Under the **Compounds** section, click **Add**
4. In the **Add Compounds** window, enter **Glycine** and click **OK**
5. In the **Select Compound Mappings** window, make sure Glycine is chosen under **Potential Matches** and click **OK**.
6. On the MetScape tab, under **Options** and **Network Type**, use the dropdown menu to select **Compound-Reaction-Enzyme-Gene**.
7. Click **Build Network**

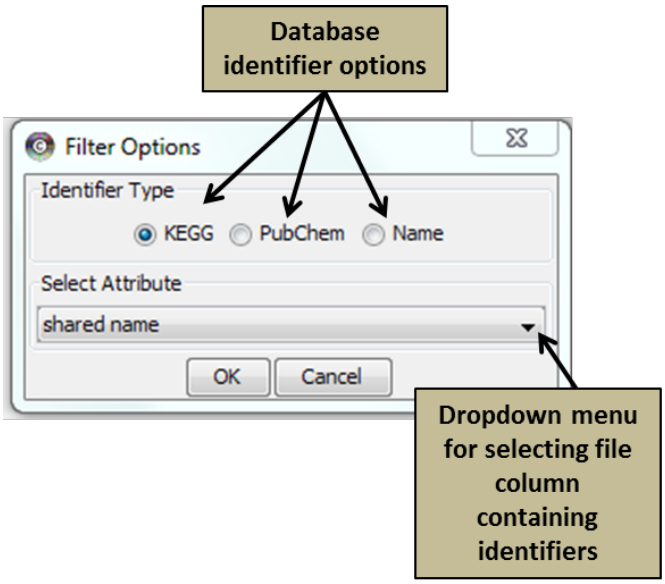





8. Go to the Opening MetDisease section to begin annotating this network with the MetDisease app. The database identifier for this network is KEGG Id.

## Opening MetDisease

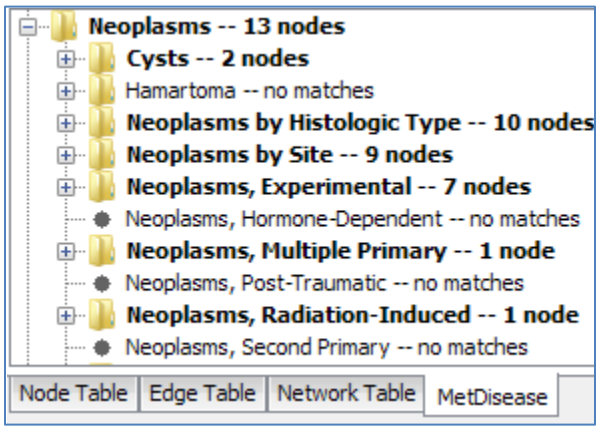
1. Select **Apps -> MetDisease -> Find MeSH Terms…**
2. The **Filter Options** dialog box appears.
3. There are 3 Database Identifiers to choose from:
    a. KEGG
    b. PubChem
    c. Name
4. Choose the **Database Identifier** used in the built network.
5. Using the dropdown menu under **Select Attribute**, choose the appropriate column from the input file that contains the identifier.
6. Click **OK**.



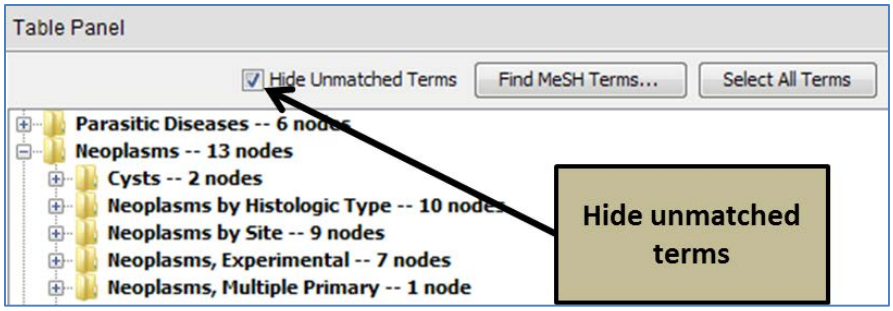

7. After the mapping completes, the disease branch of the MeSH tree is displayed. Go to the MeSH Tree section of this document for more information about the tree.

## MeSH Tree

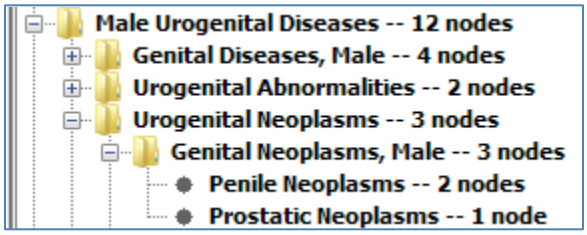
1. The disease branch of the **MeSH tree** is displayed in the **MetDisease tab** in the **Cytoscape Table Panel**.
2. Some MeSH terms are in bold while others are not:
    a. MeSH terms that are in bold have mapped compounds in the active network.
    b. MeSH terms that are not in bold have no matches in the active network.




3. To hide MeSH terms with no matches, click the box next to **Hide Unmatched Terms**, located above the MeSH tree. (**Note:** the unmatched MeSH terms can be made visible again at any time by unchecking the box next to Hide Unmatched MeSH Terms).

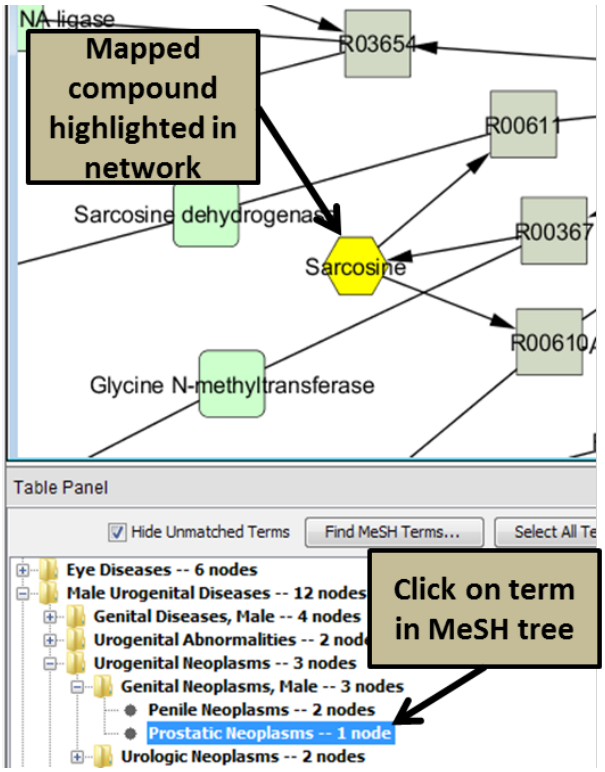



4. Only those MeSH terms that have mapped compounds in the active network are now visible.
5. Click the ⊞ sign next to any parent term to display its child term(s). If a term has a ⊟ sign next to it, the term has been fully expanded to show all child term(s).

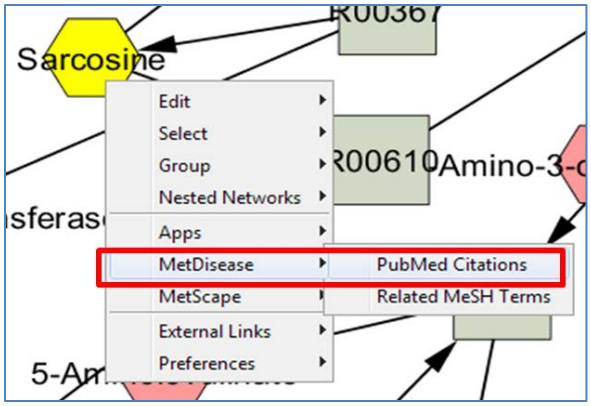



## Accessing Literature on Specific Metabolites and Diseases

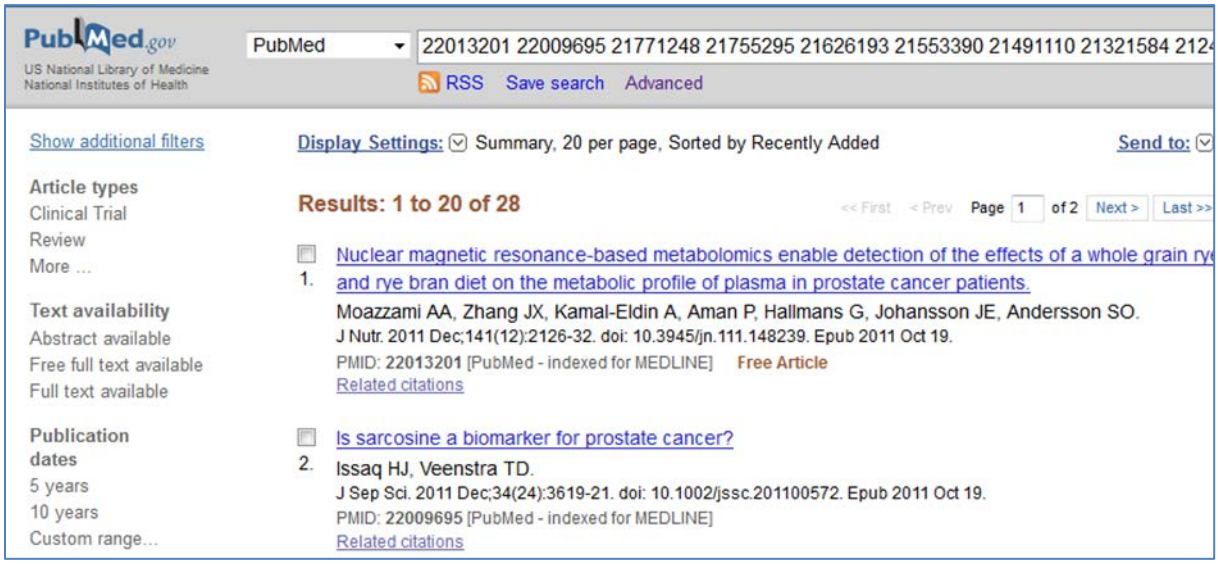
1. Click on a term in the MeSH term tree to highlight its mapped compound(s) in the drawn network.



2. Right Click on a compound in the network.
3. Select **MetDisease - > PubMed Citations**

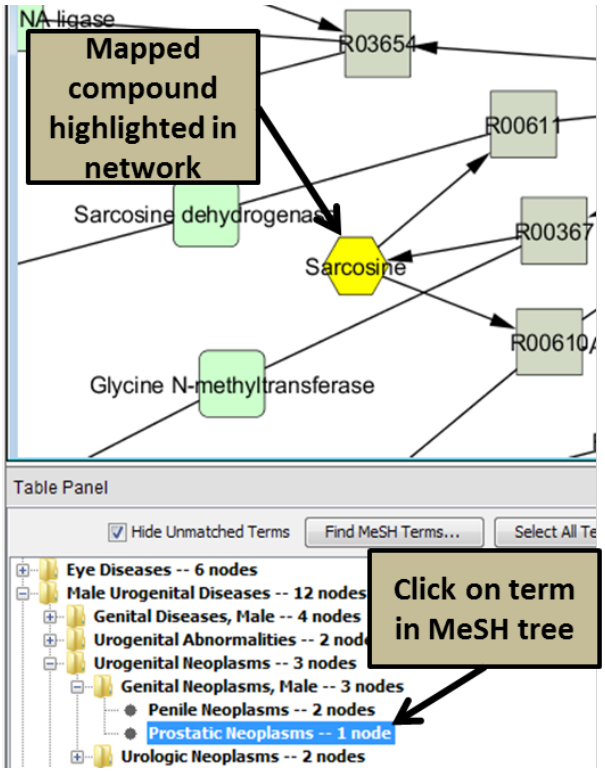



4. PubMed opens in a new browser window containing articles related to the designated compound and disease MeSH term.

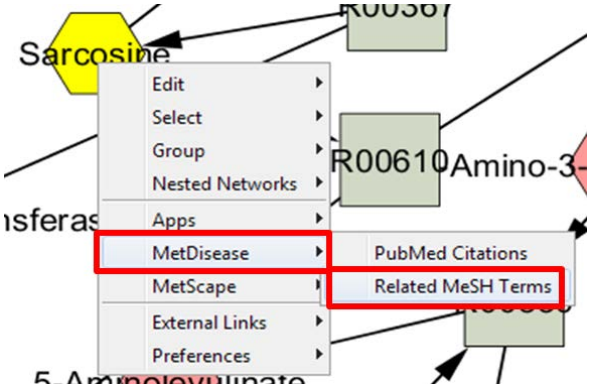



## Accessing Related MeSH Terms

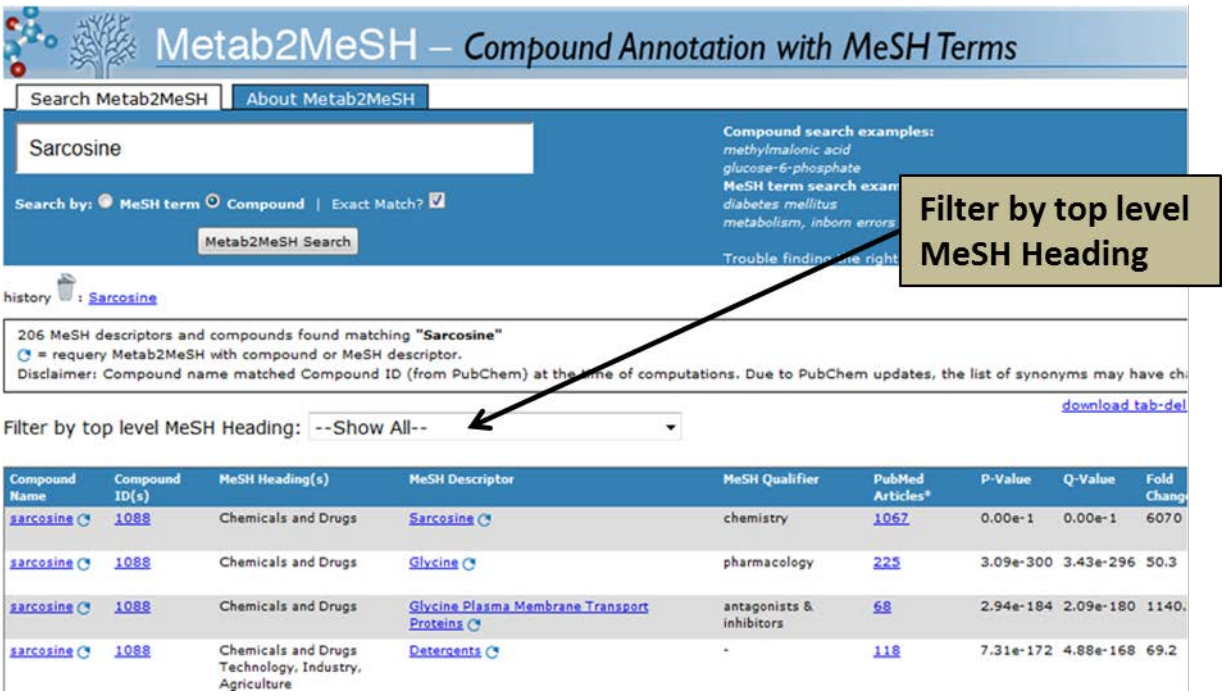
1. Click on a term in the MeSH term tree to highlight its mapped compound(s) in the drawn network.




5. Right Click on a compound in the network.
6. Select **MetDisease - > Related MeSH Terms**



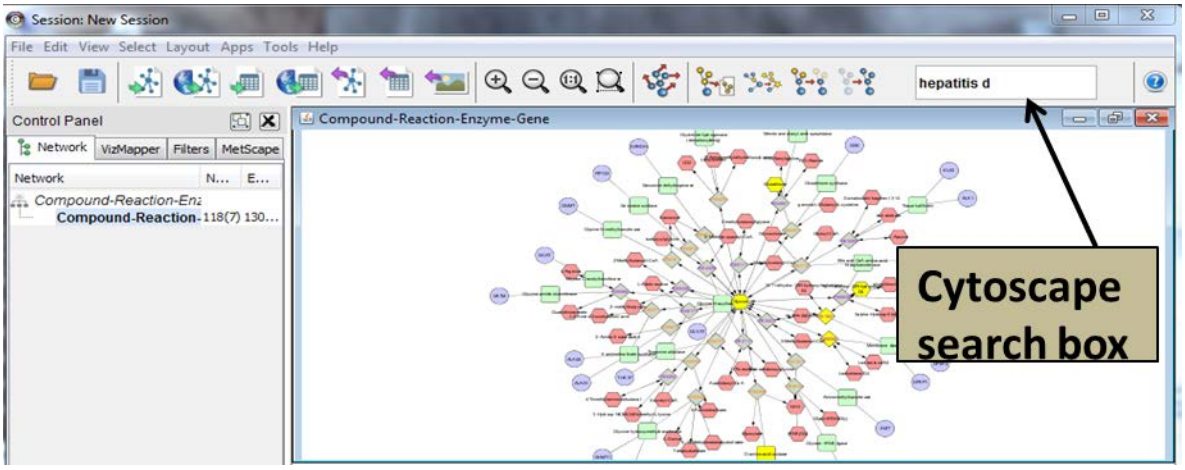

7. Metab2MeSH (Sartor, et al., 2012) opens in a new browser window showing other MeSH terms that are closely associated with the compound of interest. Metab2MeSH annotates compounds with the concepts defined in MeSH. The compound/MeSH terms displayed are those that are significantly associated in PubMed abstracts and are ordered highest to lowest by significance score. Users can filter by Diseases as a top level MeSH heading, producing results similar to MetDisease.




## Search by MeSH Terms

Use the Cytoscape search box to find nodes that are associated with a specific MeSH term.

1. Enter a MeSH term in the **Cytoscape search box** at the top right of the window. Hit the **Enter** key on your keyboard.
2. The nodes associated with this MeSH term are now highlighted in the network.




**Note:** When doing a search using Cytoscape's search box, it searches across all available attribute information, not just the MeSH terms. As a result, additional nodes may become highlighted. To

narrow the search, use quotation marks around phrases (ex. “hepatitis d” or “nephritis, hereditary”) to ensure that Cytoscape searches for the terms as a phrase.

## References

Gao J., et al. 2010. Metscape: a Cytoscape plug-in for visualizing and interpreting metabolomic data in the context of human metabolic networks. *Bioinformatics* 26(7): 971-973. PMID:20139469, PMCID:PMC2844990.

Karnovsky A., et al. 2012. Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data. *Bioinformatics* 28(3): 373-380. PMID:22135418, PMCID:PMC3268237.

Krumsiek J., et al. 2012. Mining the unknown: a systems approach to metabolite identification combining genetic and metabolic information. *PLoS Genetics* 8(10): e1003005. PMID:23093944, PMCID:PMC3475673.

Sartor M.A., et al. 2012. Metab2MeSH: annotating compounds with medical subject headings. *Bioinformatics* 28(10): 1408-1410. PMID:22492643, PMCID:PMC3348562.