AutoAnnotate Documentation

Release 1.3

Mike Kucera

Aug 19, 2019

AutoAnnotate is a Cytoscape App that identifies clusters of nodes and automatically draws shape and label annotations for each cluster. The generated labels provide a concise semantic summary of the data attached to the nodes in each cluster.

AutoAnnotate maintains a relationship between the annotations and the nodes in a cluster, when the layout of the network changes then the annotations are automatically repositioned.

AutoAnnotate maintains multiple sets of annotations for a single network, which allows you to experiment with different clustering algorithms and label generation strategies. Additionally, AutoAnnotate allows clusters to be collapsed, which can simplify large networks by reducing potentially large sections of the network into single nodes.
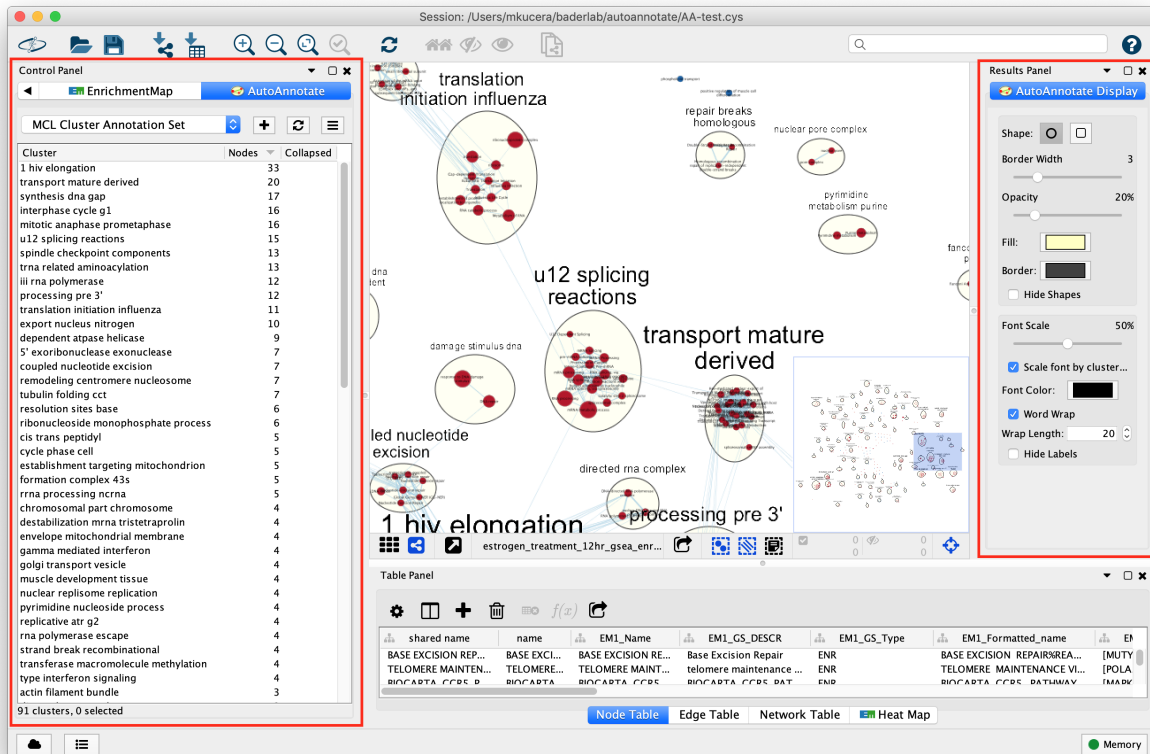
New in AutoAnnotate 1.3

- Requires Cytoscape 3.7

- Drawing annotations is much faster

- Word Wrap feature for cluster labels

- New Feature: copy annotations to another network

- Support for column namespaces

- "annotate" command now returns JSON

- Cleaner layout for "quick start" panel

- Cleaner layout for "display options" panel

- Cleaner menu layouts

- Several bug fixes

1.1 Installing

1.1.1 Requirements

Minimum versions (latest versions are recommended):

Package	Minimum Version
Cytoscape	3.7.0
WordClound	3.1.0
clusterMaker2	0.9.5

1.1.2 Installation

Install Cytoscape

- If you don't have Cytoscape please download and install the latest release from http://www.cytoscape.org/download.php.

Install AutoAnnotate 1.3

- Open Cytoscape

- In the main menu select **Apps > App Manager**

- In the App Manager select AutoAnnotate in the list of All Apps and click the Install button.

- WordCloud and clusterMaker2 will also be installed automatically along with AutoAnnotate

Alternatively the apps can be installed from the App Store.

- AutoAnnotate: http://apps.cytoscape.org/apps/autoannotate

- WordCloud: http://apps.cytoscape.org/apps/wordcloud

- clusterMaker2: http://apps.cytoscape.org/apps/clustermaker2

1.2 Creating an Annotation Set

- Start by selecting a network view.

- In the main menu select **Apps > AutoAnnotate > New Annotation Set…**

- The Create Annotation Set dialog will open.
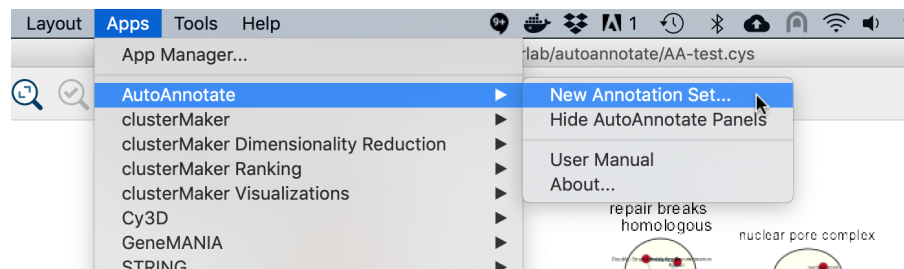
1.2.1 Create Annotation Set Dialog

This dialog is used to select the options for creating an Annotation Set.

Note: AutoAnnotate can detect if the network was created by EnrichmentMap, and if so sensible defaults for EnrichmentMap networks will be pre-selected.

Note: You may manually edit the contents and labels of the clusters after the Annotation Set is created, and you may create as many Annotation Sets as you like, so its not critical to get everything perfect at this step.
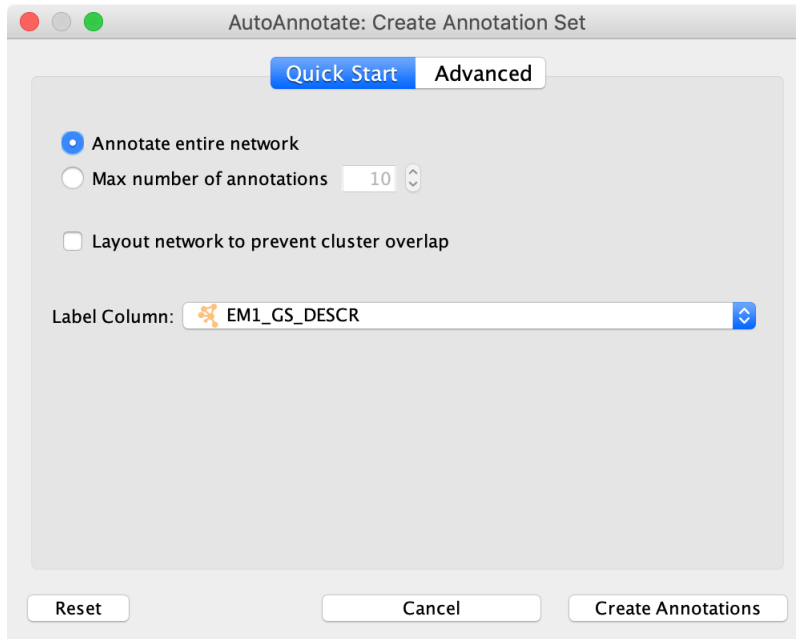

The dialog has two tabs, **Quick Start** and **Advanced**.

Quick Start

The quick start tab has a minimal set of basic options.

Annotate the entire network Every node in the graph will be part of a cluster.

Maximum number of annotations If this is selected then only the X largest clusters in the network will be given annotations.

Layout network to prevent cluster overlap Runs a layout algorithm that separates the clusters so that they do not overlap.

Label Column Column used to generate the labels.

Note: Quick start uses clusterMaker2 MCL algorithm, with the "Adjacent Words" label option.

Advanced

The Advanced tab provides full control over clustering and label options.

Cluster Options

Use clusterMaker App If this is selected then clusterMaker will be used to calculate the clusters.

Cluster algorithm Used to select the clustering algorithm provided by clusterMaker that will be used to calculate the clusters. See the table below for a list of clusterMaker algorithms currently available for use in AutoAnnotate.

Edge weight column Some clusterMaker algorithms use edge weights. This box is used to select the Edge Column to use for the edge weights. If the selected algorithm does not use edge weights then this box will be disabled (greyed out).

User-defined clusters You may provide your own clusters or use another clustering app to calculate the clusters. In this case the clusters must already be calculated and cluster identifiers must be stored in a Node Column. Any Node Column of may be used to identify the clusters. If you select a column of List type then the clusters may overlap.

Create Singleton Clusters Sometimes the cluster algorithm will leave some nodes un-clustered. If the 'create singleton clusters' option is selected then a cluster will be created for each un-clustered node.

Layout network to prevent cluster overlap Runs a layout algorithm that separates the clusters so that they do not overlap.

Label Options

Label Column Select a Node Column that will be used to calculate the cluster labels.

Label Algorithm Select an algorithm that is used to calculuate the label for each cluster. Two label algorithms are currently available:

1. WordCloud: Adjacent Words

   Uses the WordCloud app to calculate the labels.

   Based on the algorithm that is used by the "Cloud Style: Clustered Standard" option in Word-Cloud. Words in the label are chosen from the most frequent words in the cluster, and words

that are adjacent to the most frequent words. The higher the *"adjacent word bonus"* is, the more likely adjacent words will be in the label.

2. WordCloud: Largest Words

   Uses the WordCloud app to calculate the labels.

   Basically the same algorithm that is used by the "Cloud Style: No Clustering" option in Word-Cloud. Words in the label are chosen from the most frequent words in the cluster.

Click **Create Annotations** to create the new Annotation Set.

> **Warning:** AutoAnnotate will manage all groups (compound nodes) for the current network view. It is not recommended to manually create groups on a network view that has Annotation Sets. If you would like to manually create groups at any time please duplicate the network view.

1.2.2 ClusterMaker2 Algorithms

The following clusterMaker2 algorithms are made available through AutoAnnotate:

Algorithm	Weighted?
Affinity Propagation Cluster	Yes
Cluster Fuzzifier	Yes
Community cluster (GLay)	No
ConnectedComponents Cluster	Yes
Fuzzy C-Means Cluster	Yes
MCL Cluster	Yes
SCPS Cluster	Yes

Weighted algorithms require an edge attribute to be used as the edge weight.

1.2.3 Using an External Clustering Algorithm

AutoAnnotate allows clusters to be defined using any node attribute via the **User-defined clusters** option on the Advanced panel. This allows clustering algorithms provided by other Cytoscape Apps or by external scripts to be used with AutoAnnotate.

Using clusterMaker Manually

You may use a clusterMaker algorithm not provided through the **New Annotation Set** dialog.

Here is an example using the clusterMaker MCODE algorithm.

- Select **Apps > clusterMaker > MCODE Cluster**.

- A dialog with MCODE algorithm settings will be shown.

- Expand the **Cytoscape Advanced Settings** section and make note of the **Cluster Attribute Name**, in this case it is __**mcodeCluster**.

- Click **Ok** to run MCODE.

- A node attribute (column) called __**mcodeCluster** is created where each node is assigned a cluster identifier.

- Open the **New Annotation Set** dialog and choose the Advanced panel.

- Select **User-defined clusters** and then select the **__mcodeCluster** column.

- Click **Create Annotations**.

Using External Scripts

See the *Automating AutoAnnotate* for details on how to automate Cytoscape using external scripts.

A script must do the following to provide clusters to AutoAnnotate:

- Create a node attribute (column), any type may be used.

- Use the node attribute to assign a cluster identifier to each node. The cluster identifier may be blank, those nodes will not be included in clusters.

- Open the **New Annotation Set** dialog and choose the Advanced panel.

- Select **User-defined clusters** and then select the **__mcodeCluster** column.

- Click **Create Annotations**.
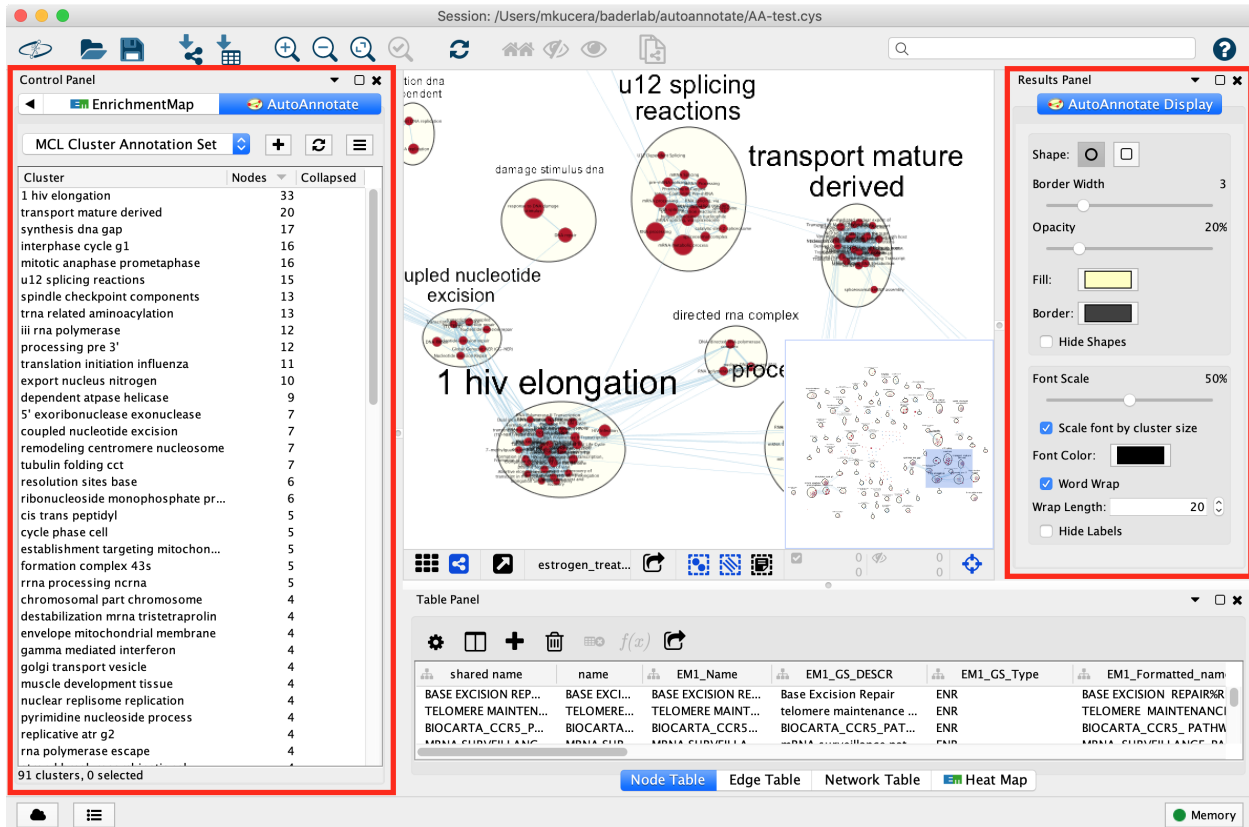
1.3 Working With AutoAnnotate

1.3.1 Panels

When an Annotation Set is created the AutoAnnotate panels will be shown.

Note: To show or hide the AutoAnnotate panels go to the main menu and select **Apps > AutoAnnotate > (Show/Hide) AutoAnnotate Panels**.


- In the **Control Panel** on the left side is the **Main AutoAnnotate Panel**.

  – This panel is used to manage Annotation Sets and Clusters.

- In the **Results Panel** on the right side is the **Display Options Panel**.

  – This panel is used to modify the look of the annotations.

1.3.2 Changing Display Options

This AutoAnnotate Display panel shows display options for the currently selected Annotation Set.
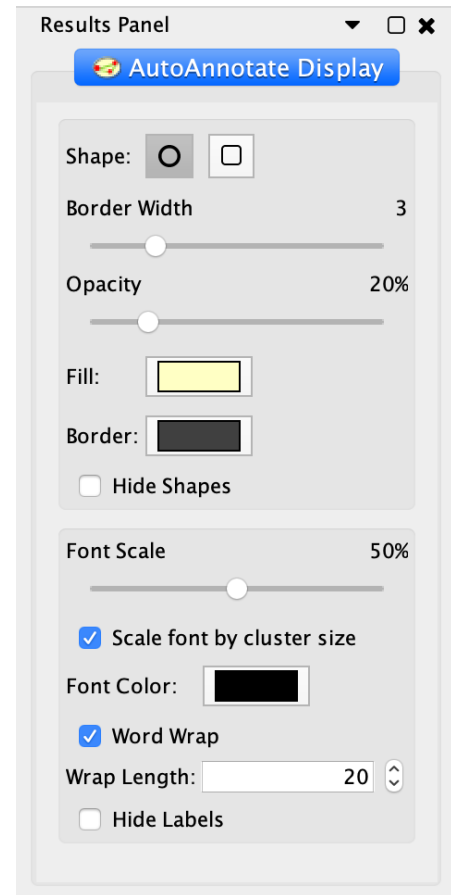
Shape Options

>   **Shape** There are two available shapes: Ellipse or Rectangle

>   **Border width** Slider that adjusts the width of the cluster annotation borders.

>   **Opacity** Slider that adjusts the opacity of the cluster annotation fill color. Slide to the left for completely transparent. Slide to the right for completely opaque.

>   **Fill Color** Opens a color picker for choosing the fill color.

>   **Border Color** Opens a color picker for choosing the border color.

>   **Hide Shapes** Select to hide all cluster shape annotations.

Label Options

>   **Font Size** Slider that adjusts the font size.

>>   **Scale font by cluster size** When selected the size of the label will be relative to the number of nodes in the cluster. Larger clusters will have larger labels. When deselected all the clusters will have the same size labels. Slider is used to adjust the font size or font scale.

>   **Font Color** Opens a color picker for choosing the font color.

>   **Word Wrap** Allows long labels to be broken and wrap onto the next line.

>>   **Wrap Length** Defines the maximum width of the label in number of characters. Individual words that are longer than the wrap length will not be broken. Has no effect if Word Wrap is disabled.

>   **Hide Labels** Select to hide all cluster labels.

1.3.3 Managing Annotation Sets

The main AutoAnnotate panel is used to manage Annotation Sets.
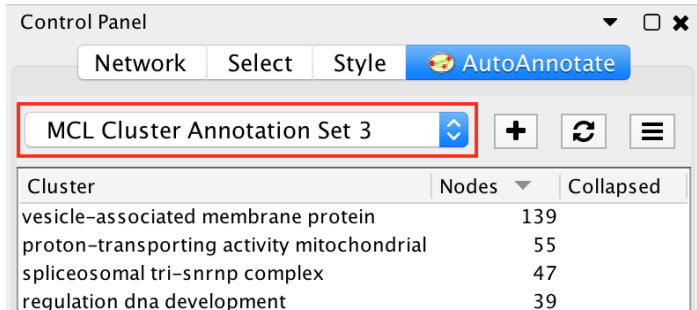
At the top of the panel is a dropdown box that allows you to switch between all of the Annotation Sets that have been created for the current network view.

- Select the (none) option to clear all annotations from the view.

- Select an Annotation Set to show all the annotations and labels in that set.

- When switching between Annotation Sets all the annotations in the view will be erased and redrawn.

    – Similarly when switching Annotation Sets any collapsed groups will be expanded.

1.3.4 Annotation Set Menu
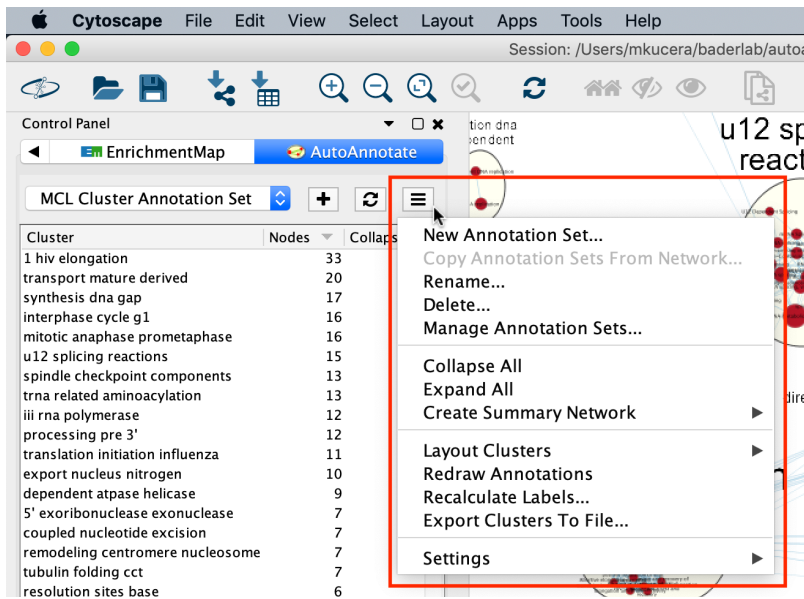
Click the ☰ button to show the Annotation Set Menu.

New Annotation Set. . . Opens the Create Annotation Set dialog. The dialog can also be opened by clicking the ✚ button.

Copy Annotation Sets from Network. . . Allows annotations to be copied from one network to another. See *Copying Annotations Between Networks* for more details.

Rename Renames the current annotation set.

Delete. . . Deletes the current annotation set.

Manage Annotation Sets. . . Opens a Dialog where you can change the order that Annotation Sets appear in the dropdown and delete multiple Annotation Sets at a time.

Collapse All Collapses all the clusters. **Warning**: Collapsing all clusters in a large network can be very slow. Consider using the **Create Summary Network** option instead.

Expand all Expands all the clusters.

Create Summary Network Generates a new network where each node represents a collapsed cluster in the original network. Very similar in functionality to "Collapse All", but runs significantly faster. See *Summary Network* for more details.

Layout Clusters The submenu provides access to cluster-aware layout algorithms. See *Cluster-Aware Layouts* for more details.

Redraw annotations Some versions of Cytoscape have a bug that cause annotations to occasionally be drawn in the wrong location. Redrawing the annotations can fix this problem. Annotations can also be redrawn by clicking the [↻] button.

Recalculate Labels Recalculates the labels for all of the clusters in the annotation set. Typically used when the WordCloud options or label algorithm changes, or if the contets of one or more clusters change. See *Label Options* for more details. (Note: to recalculate labels for a subset of clusters please select the clusters in the table and right-click to show the cluster context menu.)

Export Clusters to File. . . Creates a TXT file containing the labels of all the clusters.

Settings

Display Creation Parameters. . . Opens a dialog that shows the parameters that were chosen in the Create Annotation Set Dialog for the current Annotation Set.

Label Options. . . Opens a dialog where you can select the label algorithm that is used when recalculating labels. See *Label Options* for more details.

Settings. . . Opens a dialog where general AutoAnnotate settings can be changed.

1.3.5 Managing Clusters

On the main AutoAnnotate panel you will find the Cluster Table.

- This table shows each cluster, along with its label, the number of nodes in the cluster and a checkbox icon that indicates if the cluster has been collapsed.

- Click on a table header to sort the table by that column.

Most of the cluster options are available using the context menu of the cluster table.

- Start by selecting one or more clusters in the cluster table.

- To select multiple clusters hold the Shift key to select a range or hold the Ctrl/Command key to select individual clusters.
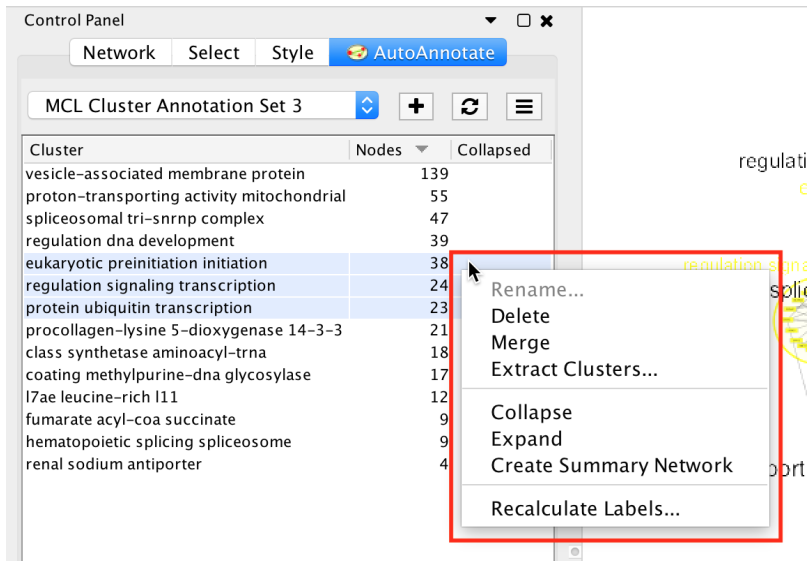
Right click with he mouse to show the cluster table context menu.

Rename… Enabled when 1 cluster is selected. Manually set the label for the cluster.

Delete Delete the cluster annotations. Note: This does not delete the nodes or edges in the cluster, just the annotations.

Merge Enabled when 2 or more clusters are selected. Merges the clusters into a single cluster. A new label for the merged cluster is automatically calculated.

Extract Clusters… Creates a new Annotation Set consisting of the selected clusters.

Collapse Collapses the selected clusters. Warning: can be very slow when many clusters are selected.

Expand Expands the selected clusters.

Generate Summary Network Generates a new network where each node represents a collapsed cluster that is selected in the original network. Very similar in functionality to "Collapse", but runs significantly faster.

Recalculate Labels Recalculates the labels for the selected clusters. See *Label Options* for more details.

1.3.6 Managing Clusters in the Network View

To manually create a cluster select some nodes in the network view, right click to show the context menu, and select **Apps > AutoAnnotate > Create Cluster**.
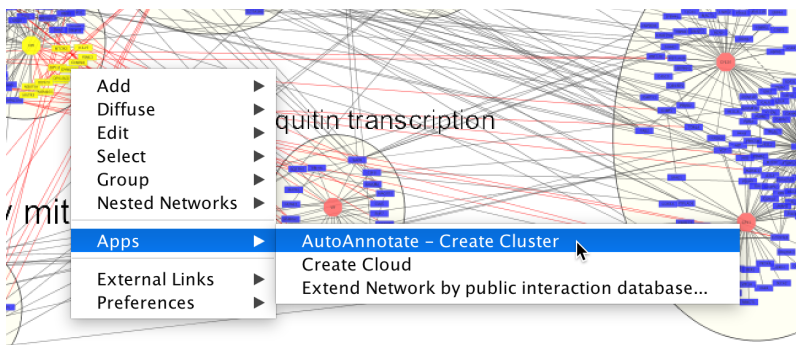

- A new cluster is created.

- A label for the cluster is automatically calculated.

- The cluster is selected in the cluster table.

Deleting a node in an existing cluster will remove the node from the cluster. A new label is calculated.
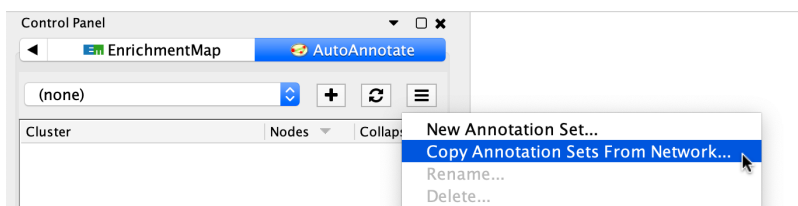
1.3.7 Copying Annotations Between Networks

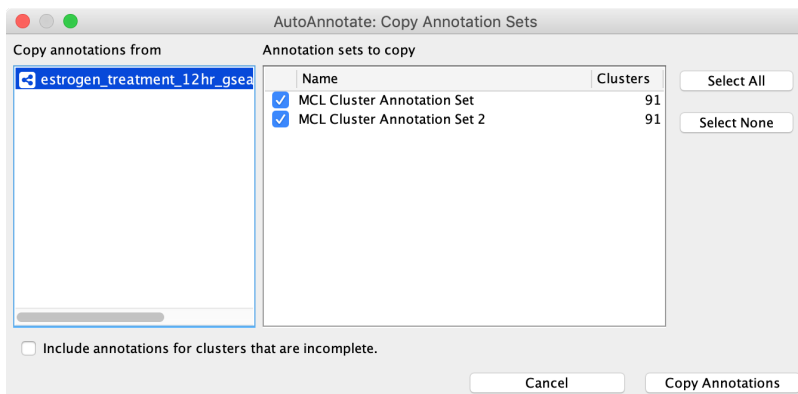Annotations can be copied between networks in the same network collection only.

Start by selecting the destination network in the **Networks** panel.
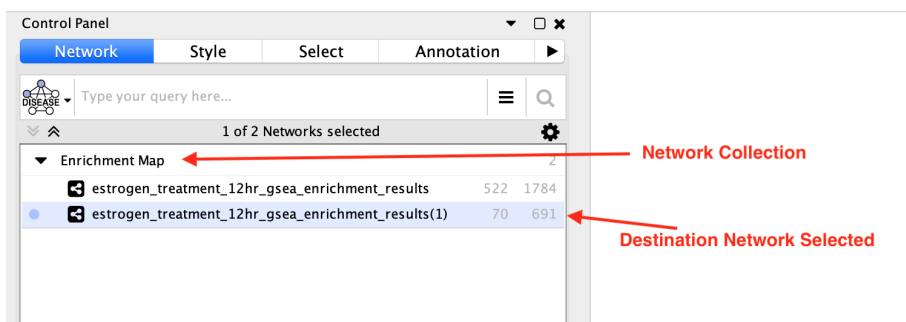
Then go to the **AutoAnnotate** panel and select **Copy Annotation Sets from Network…** in the panel menu.

The **Copy Annotation Sets** dialog will be shown.

Select the source network to copy annotations from. A list of annotation sets for that network is shown. Select the annotation sets to copy. Finally click the **Copy Annotations** button.

If the destination network contains fewer nodes than the source network then only the annotations for clusters that exist in the destination network will be copied. Select the **Include annotations for clusters that are incolmplete** option to copy annotations for clusters that do not have all their nodes present in the destination network.

1.3.8 Additional Options

AutoAnnotate has various warning dialogs that can be dismissed permanently by selecting "don't show this again". If you would like to see these warnings again they can be re-enabled. In the main menu select **Edit > Preferences > Properties**, then select **org.baderlab.autoannotate** in the Properties Dialog.
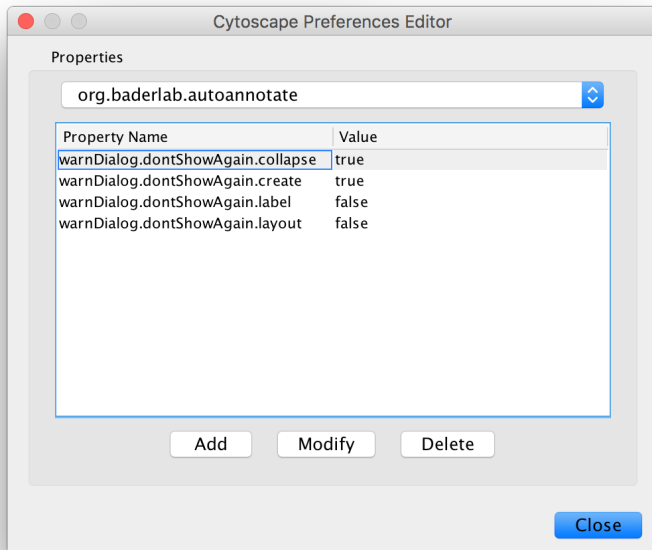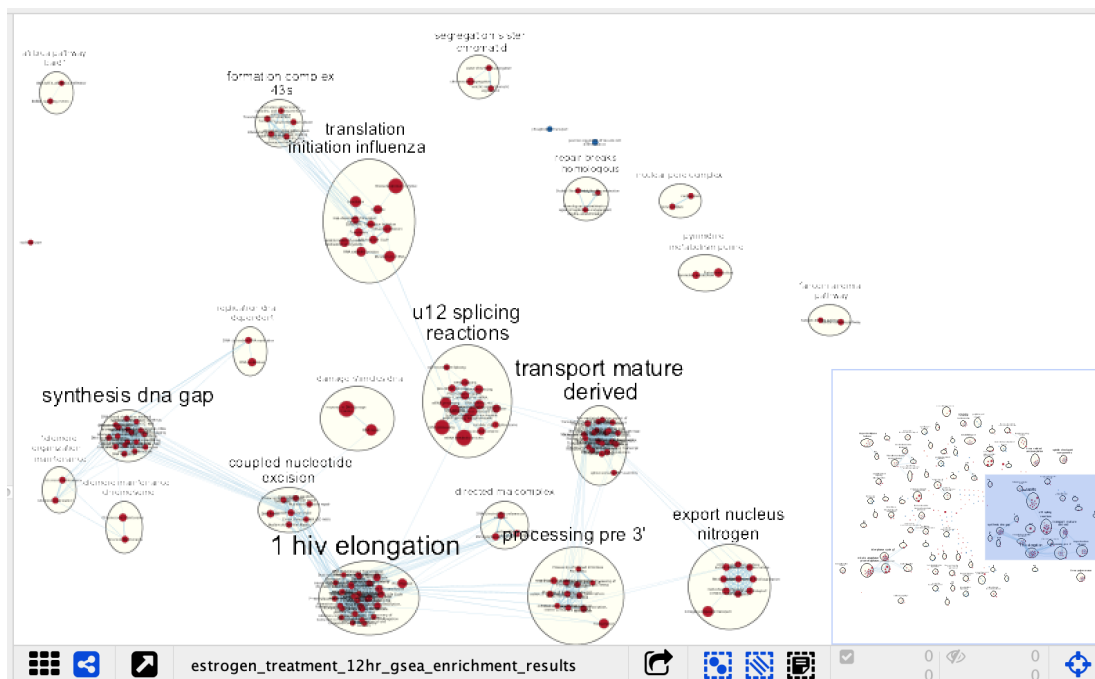
1.4 Grouping and Layout

1.4.1 Cluster-Aware Layouts

AutoAnnotate provides three cluster-aware layout algorithms. These are available from the panel menu.
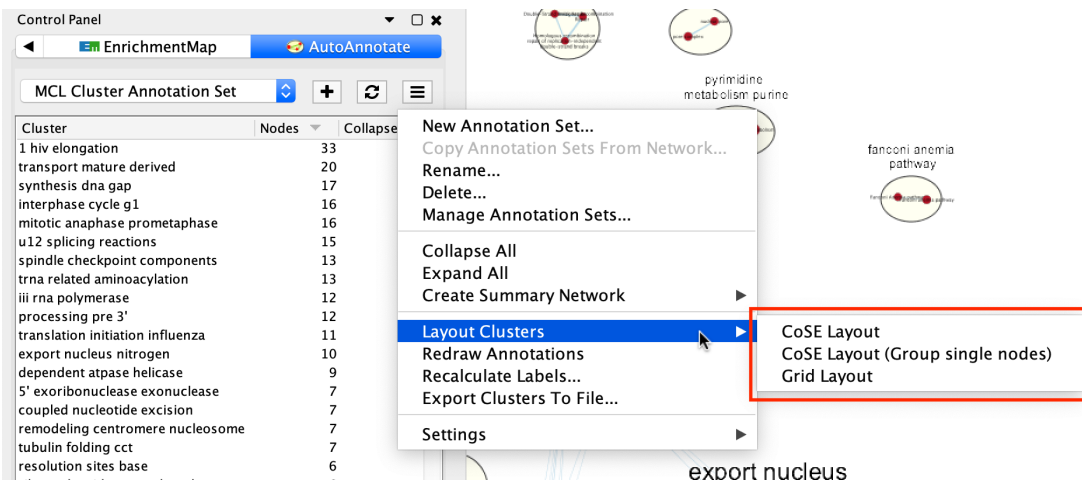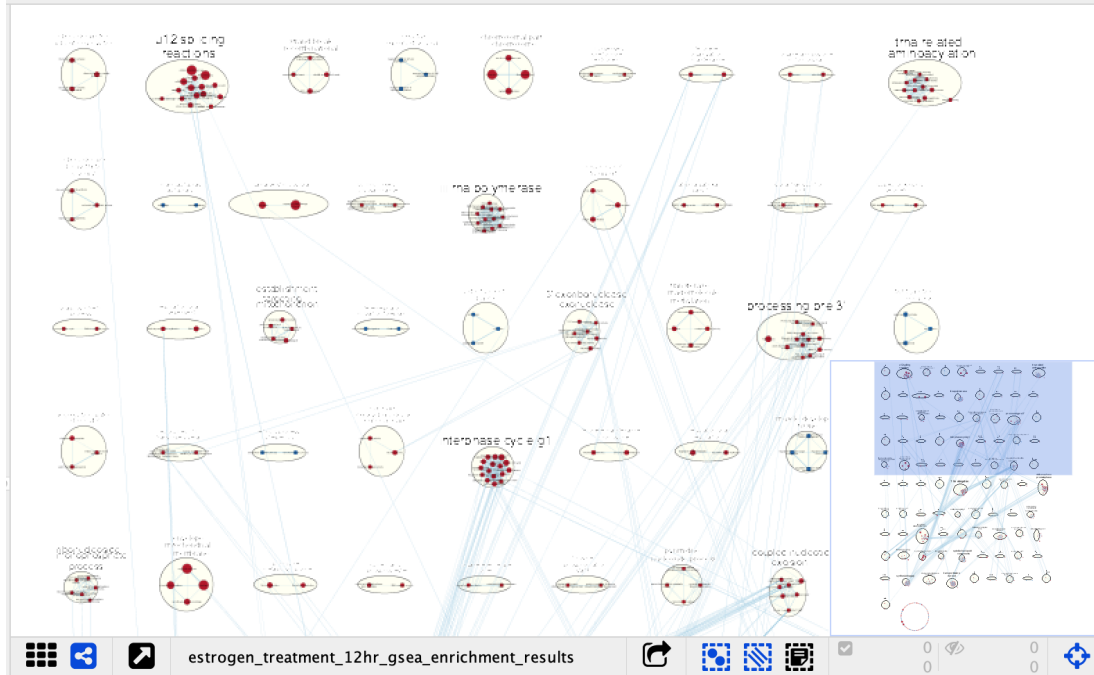
CoSE Layout

The CoSE (Compound Spring Embedded) layout is a force-directed layout that is cluster-aware. It will attempt to minimize cluster overlap while still placing connected clusters near each other.

Cluster annotations will still sometimes overlap. If this is happening try running the layout again or manually adjust the position of the clusters.

There are two versions of CoSE available. The standard version will allow unclustered nodes to be placed about the graph. The **CoSE Layout (Group Single Nodes)** version will group all unclustered nodes together rather than allowing them to be spread about the graph.

Grid Layout

The clusters are laid out in a grid pattern. The nodes within each cluster are laid out using a standard force-directed layout.

This layout is slower and less esthetically pleasing that the CoSE layout. This was the only layout available in older versions of AutoAnnotate and is still available mainly for backwards compatibility.

1.4.2 Collapsing and Expanding Clusters

AutoAnnotate can summarize a network by collapsing some or all of the clusters automatically.

> **Warning:** Depending on the number of clusters this function can be very slow. Consider using the **Generate Summary Network** option instead because it runs significantly faster. See the section below for more details.
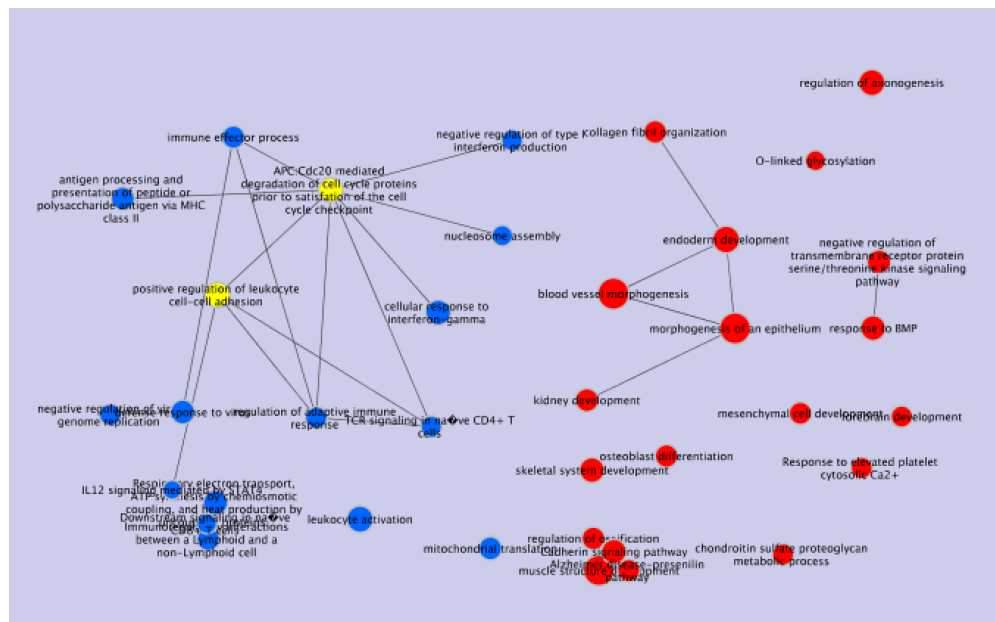
The Cytoscape "group nodes" feature is used to collapse and expand clusters. A group node is a compound node that contains other nodes.

When AutoAnnotate collapses a cluster it first creates a group node that contains all the nodes in the cluster and then the group node is collapsed. When the cluster is collapsed Cytoscape will create "meta-edges" between the group node any other nodes it is connected to. The collapsed group nodes and the meta-edges provide a summary of the network.

When a cluster is expanded the group node is expanded and then deleted. The annotations are shown again.

Note: The annotation for a cluster is not shown when the cluster is collapsed. The label shown for a collapsed cluster is generated by Cytoscape and depends on the Group Preferences (explained below).


Warning: It is not recommend to manually collapse or expand clusters using the Groups menu as this may confuse AutoAnnotate. When switching between Annotation Sets all collapsed clusters are automatically expanded.

Warning: There are currently some limitations in Cytoscape with regards to saving and restoring annotations and groups in a session file. For this reason when you save your session all the clusters must be expanded first.

Clusters can be collapsed or expanded from the Annotation Set Menu or the Cluster Table Context Menu. Before collapsing clusters make sure that node aggregation is enabled. This will tell Cytoscape to generate a label for the group node that aggregates the labels of the nodes in the cluster.

- In the main menu select **Edit > Preferences > Group Preferences**.
- In the Group Preferences dialog select the option for Enable Attribute Aggregation.
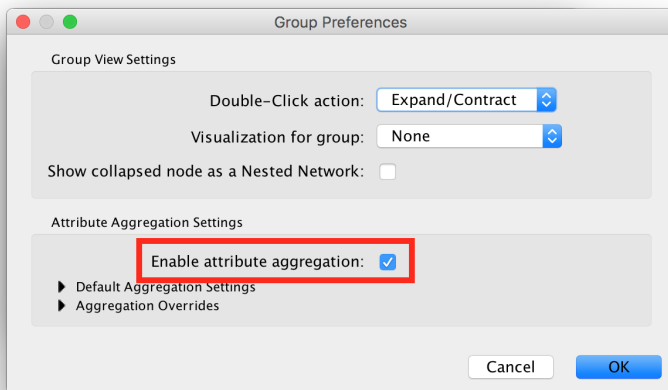
1.4.3 Summary Network

Using group nodes to collapse clusters can be very slow when there are many clusters. As an alternative a new option called **Generate Summary Network** has been added in AutoAnnotate 1.1. A new network is generated where each node represents a collapsed cluster in the original network, and each edge represents a meta-edge between clusters. The resulting network is very similar to the results you get from collapsing the clusters.

The main differences are:

- Nodes that are not contained in any cluster are ignored.

- There is no relationship maintained back to the original network. If the clusters in the original network change the summary network will not be updated. Create a new summary network at any time to reflect the current state of the clusters.

- A new visual style for the summary network is created based on the original visual style.

A summary network can be created from the Annotation Set Menu or the Cluster Table Context Menu.

Summary network uses the same attribute aggregation settings as collapsed group nodes. Available from **Edit > Preferences > Group Preferences** then expand the Default Aggregation Settings section.

1.5 Label Options

AutoAnnotate uses the WordCould app to generate cluster labels. Some basic label options are available from the AutoAnnotate panel menu, but more advanced options are available through WordCloud directly.

1.5.1 WordCloud App

The WordCloud app generates a "word cloud" for selected nodes based on one or more attributes (columns) of choice.
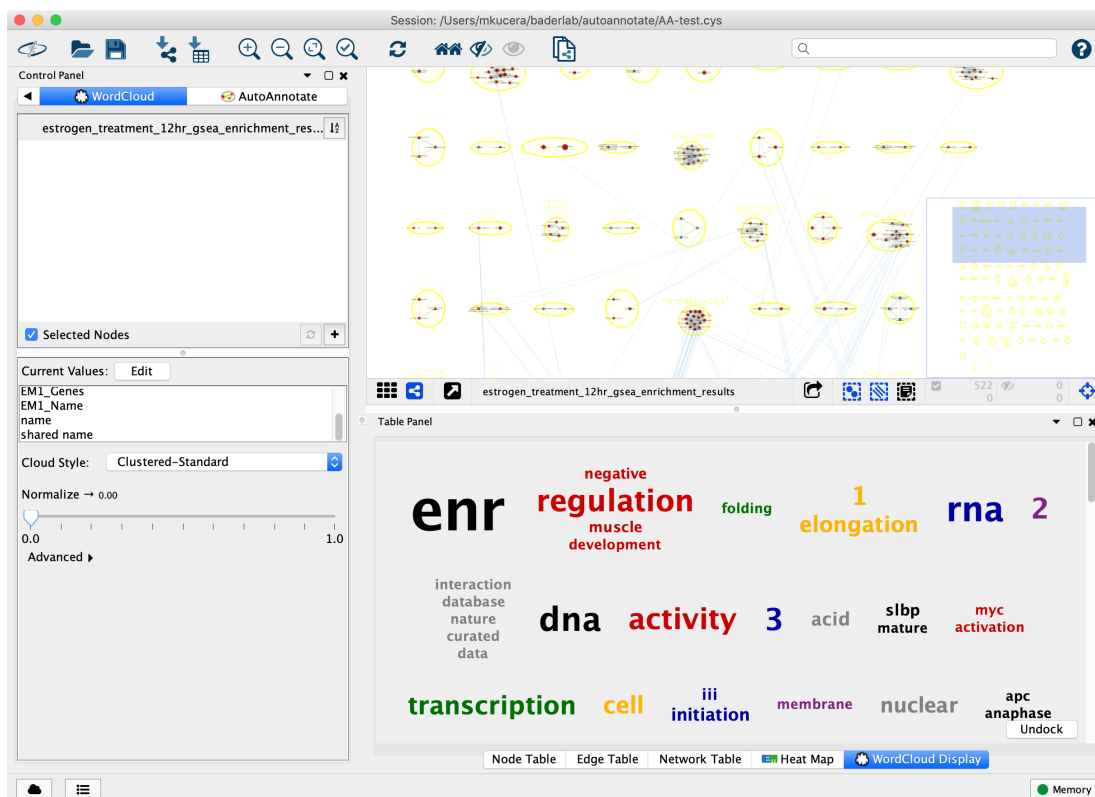
AutoAnnotate invokes WordCloud to generate a word tag cloud for the node data within each cluster, which is used to derive the text for the label annotations.

1.5.2 Basic Label Options

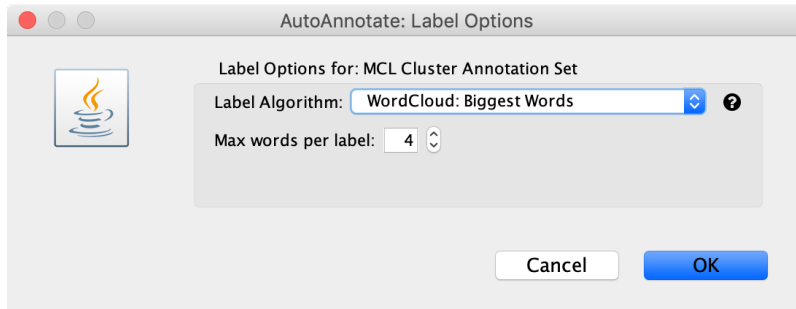Basic label options are available from the AutoAnnotate panel menu at **Settings > Label Options…**.
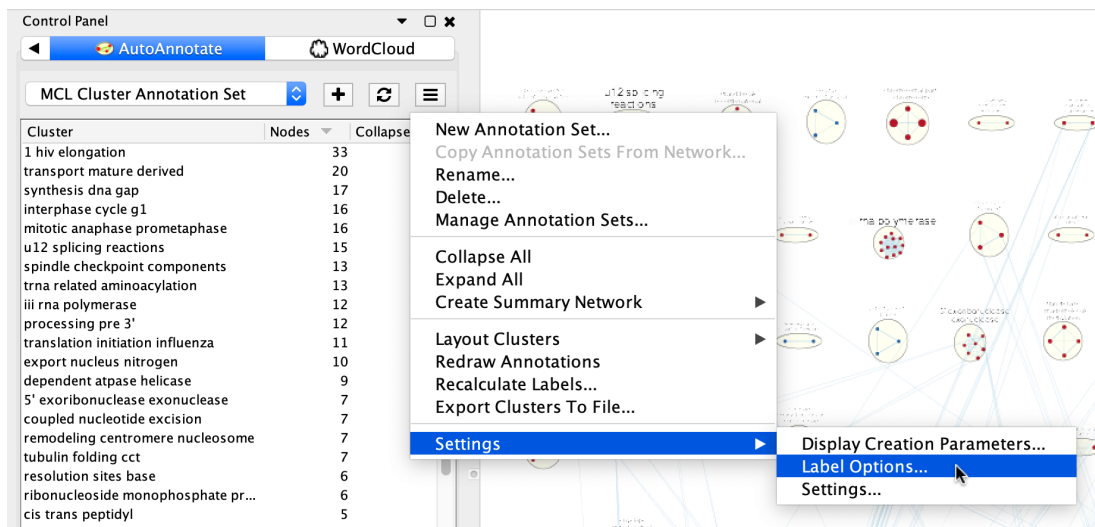
This will open the **Label Options** dialog. (These are the same options available in the **Create Annotation Set Dialog**.)
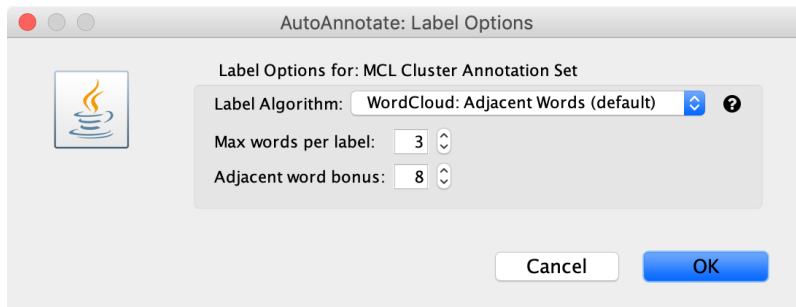
There are two label algorithms available…

Biggest Words

The "Biggest Words" option generates a word cloud for the nodes in each cluster. The words are sorted by size then the N largest words are chosen.

The only option available is to choose the maximum number of words in each label.

Adjacent Words

The "Adjacent Words" option is a heuristic that attempts to balance word size with word adjacency. Words that are adjacent to the largest words are given a size bonus making them more likely to be chosen for the label.

A word cloud is generated for the nodes in each cluster. The words are sorted by size, then a size bonus is added to every word that is adjacennt to the N largest words. This causes words that are adjacent to the N largest words to be more likely to be chosen. Then the list is sorted again by size and the N largest words (with bonus) are selected.

The **adjacent word bonus** option is used to decide how much of a size bonus to give to words that are adjacent to the largest workds. The default value is 8.

1.5.3 WordCloud Options

In the main menu select **Apps > WordCloud > Show WordCloud**.

In the WordCloud panel you can change options that effect the contents of each cluster's word cloud, for example using the Normalize slider.

Back in the AutoAnnotate panel select **Recalculate Labels** to regenerate labels using the updated WordCloud options.

1.5.4 Removing Words from Labels

If you do not want certain words to appear in labels they can be removed using WordCloud.

- Expand the **Advanced Options** section in the WordCloud panel.

- Click the **Excluded Words** button.

- A dialog will show.

- Enter the word that should be excluded, then click **Add**.

- Click **Ok**.

- Go back to the AutoAnnotate panel and select **Recalculate Labels** from the menu. The labels will be regenerated and the excluded words will no longer be chosen.

1.6 Automating AutoAnnotate

AutoAnnotate provides several commands which allow basic features to be automated via scripts, the command line or REST.

1.6.1 CyREST App

To call commands via REST the CyREST App is required. CyREST is normally installed by default.

CyRest is updated often. There are two ways to install the latest version:

1. App Manager

   Open the **App Manager** dialog from the main menu at **Apps > App Manager**. Then go to the **Check For Updates** tab. If there is a newer version available it will be listed.

2. App Store

   CyRest can also be installed or updated from the App Store Website

1.6.2 Command Documentation

CyREST Documentation

On-line documentation for AutoAnnotate commands can be accessed from the main menu at **Help > Automation > CyREST Command API**. This will open a web browser with documentation for all the commands that are available through CyREST. Navigate to the **autoannotate** entry and expand it for detailed documentation on each command.

Note: For more details on Cytoscape Automation see the Cytoscape Automation Manual Page. For more details on using CyRest see the CyREST Documentation

Command Line Documentation

Open the Automation Panel from the main menu at **View > Show Automation Panel**.

Type `help autoannotate` to list the available commands provided by AutoAnnotate. To get help on a particular command type, for example the `redraw` command, type `help autoannotate redraw`.

Note: For more details on the Automation Panel see the Cytoscape User Docs

1.7 Bug Reports and Enhancement Requests

Thank you for using AutoAnnotate. Your feedback is important and helps us to improve the App for everyone. If you have any comments, requests for new features or enhancements, or bugs to report, please enter them on the AutoAnnotate github issue tracker.

https://github.com/BaderLab/AutoAnnotateApp/issues