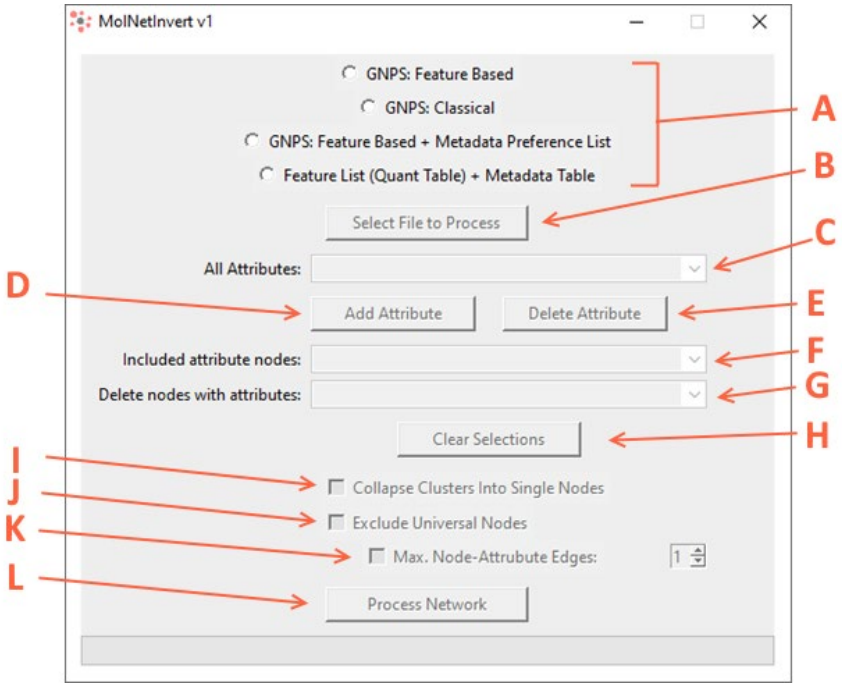
MolNetInvert User Guide

This is a user guide to using MolNetInvert, a graphical user interface for creating Metadata Based Molecular Networks (MBMN).



A: Four radio buttons for selecting the type of data that you wish to process into a MBMN.

B: Button for selecting the file/s you wish to process.

C: List of all metadata attributes that were found in your selected file/s.

D: An "Add Attribute" button for adding a metadata attribute selected in the list above. The added metadata attributes will be added in the resulting MBMN.

E: A "Delete Attribute" button for deleting a metadata attribute selected in the list above. The deleted metadata attributes will remove all features that have a relationship to that attribute in the resulting MBMN.

F: List of all metadata attributes that will be included in the resulting MBMN.

G: List of all metadata attributes that will be deleted in the resulting MBMN.

H: A "Clear Selections" button that will remove all selected attributes that have been included or deleted.

I: Option tick box to toggle the option of collapsing feature clusters into single nodes in the resulting MBMN.

J: Option tick box to toggle the option of excluding universal nodes in the resulting MBMN.

K: Option tick box to toggle the ability to set a maximum feature-to-metadata attributes allowed for individual features to be present in the resulting MBMN as well as a scroll box to set that maximum.

L: A "Process Network" button to process the selected data with the selected options into a MBMN.

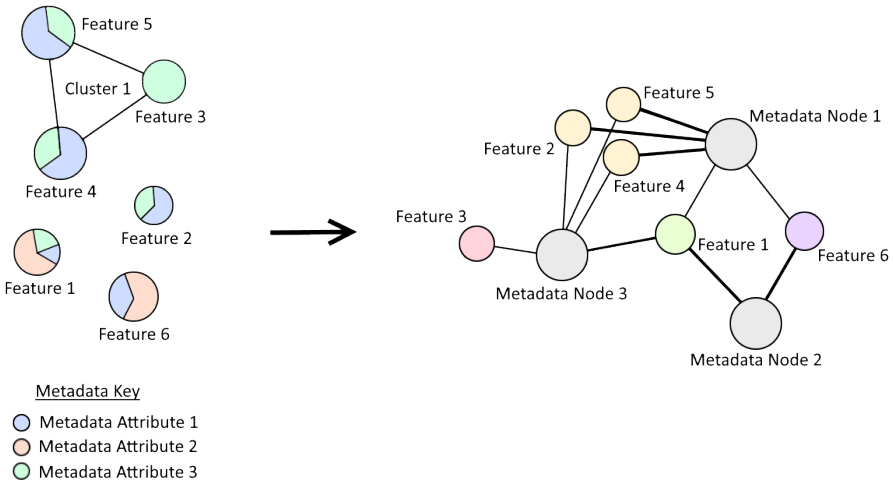


Figure 2: Example of the conversion of a Feature Based Molecular Network into a MBMN.

Choose Data Type

1. Using the radio buttons choose which type of data type you want to process into a Metadata Based Molecular Network (MBMN), select either:

    - **GNPS: Feature Based**

    Select this option for processing molecular networks that have been created via the GNPS Feature Based Molecular Network or Ion Identity Molecular Network workflow. The data file should be of type .graphml.

    - **GNPS: Classical**

    Select this option for processing molecular networks that have been created via the GNPS Classical Molecular Network workflow. The data file should be of type .graphml.

    - **GNPS: Feature Based + Metadata Preference List**

    Select this option for processing molecular networks that have been created via the GNPS Feature Based Molecular Network or Ion Identity Molecular Network workflow that have a large amount of metadata attributes. The Preference List allows you to ignore all metadata attributes that are not in the Preference List, making selecting attributes in MolNetInvert less tedious. The molecular network file should be of type .graphml, and the Preference List needs to be a .txt file with a list of metadata attributes. **NOTE:** the preferred metadata needs to have the exact same spelling as the metadata file used in the GNPS workflow.

    - **Feature List (Quant Table) + Metadata Table**

    Select this option for processing Feature Lists rather than GNPS based molecular networks. The Feature List file should be of type .csv similar to which is exported from MZMine3, and the Metadata Table should be a tab separated .txt file.

        - m/z values must be in a column with the heading “row m/z”
        - Retention time values must be in a column with the heading "row retention time"
        - Feature abundancies must be in a column with the following heading format: “SAMPLENAME Peak area”

- The metadata file must be a text file containing two columns on information, the first column containing the sample name (exactly the same as used in the above feature abundancies column), and the second column separated by a tab containing the metadata attribute

Select File to Process

2. Press the button "Select File To Process". This will open a dialogue box to select the file to process into a MBMN. Depending on the data type chosen in the previous step, particular file types should be selected.
    - For processing feature tables or molecular networks with metadata preference lists, you will first select the feature table or molecular network file, after which a new dialogue box will open for selecting the metadata list.

Select Attributes to Include and Exclude

3. Select which metadata attributes you would like to have present as metadata nodes in the resulting MBMN (metadata nodes will only be in the resulting network if they have accompanying features in the dataset). Also select which metadata attributes you would like to delete from the resulting MBMN.
   NOTE: Features that are present in the "Added Attributes" will not be present in the MBMN if they are also present in deleted attributes, however features that are in "Add Attributes" but also present in attributes that are not added (or deleted), will still be present in the resulting MBMN.

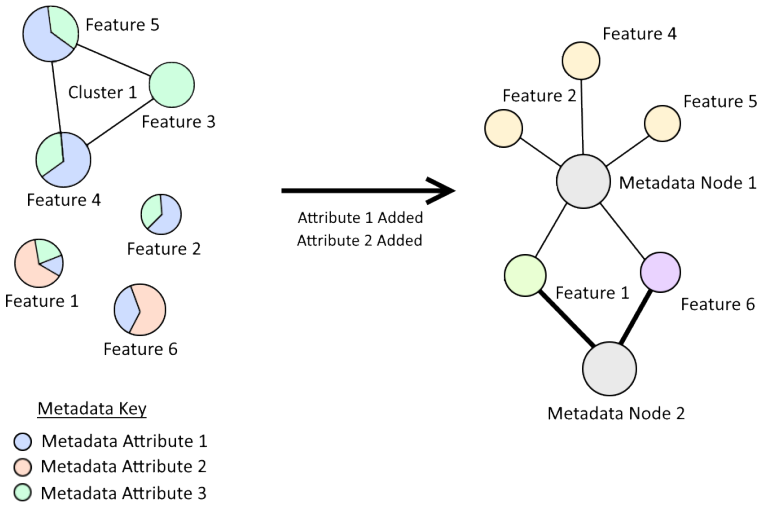


Figure X: Example of a resulting MBMN from the addition of metadata attribute 1 and metadata attribute 2. Features that only contain relationships to attributes that are not added (Feature 3) are excluded from the resulting MBMN.

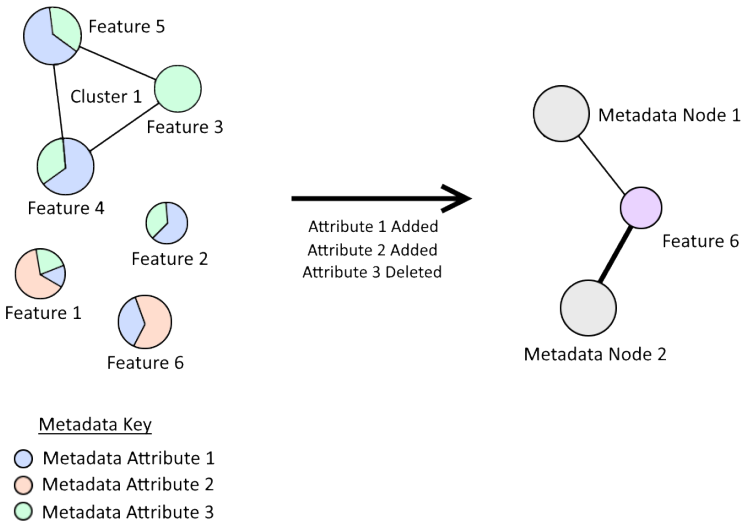


Figure X: Example of a resulting MBMN from the addition of metadata attribute 1, metadata attribute 2 and the deletion of attribute 3. Features that contain relationships to attributes that are deleted (Feature 3, Feature 4, Feature 5) are excluded from the resulting MBMN.

Select Processing Preferences

4. Select your preferences for the processing of the data into a MBMN. You have the following options:

    - **Collapse Clusters into Single Nodes** (not an option of Feature List datasets)

MBMN that are created from molecular networks via the GNPS workflow have the option to collapse molecular clusters into single nodes.

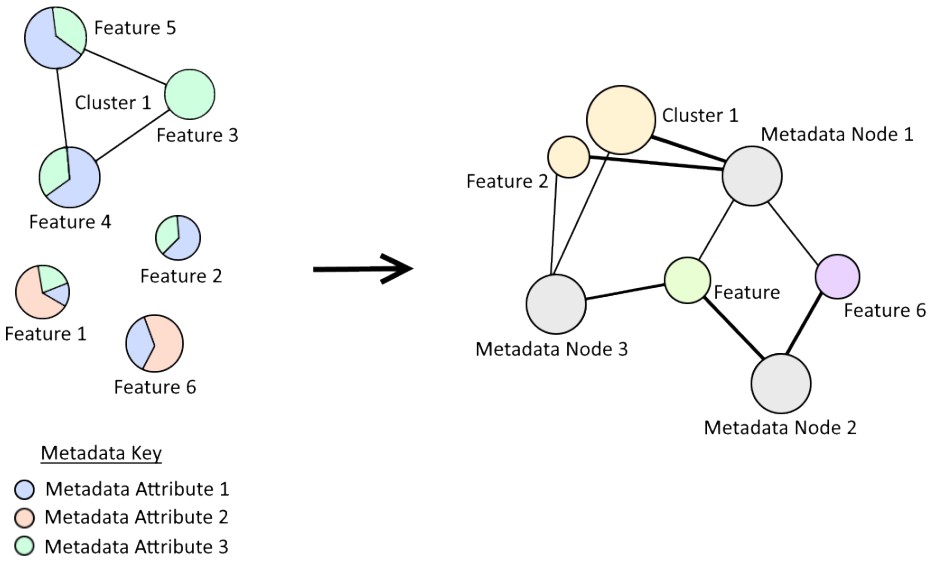


Figure X: The resulting MBMN that is created from a GNPS molecular network with the option of collapsing molecular features into node enabled. Feature 3, Feature 4, and Feature 5 are represented in the MBMN as a single, larger, node. The edge connections represent the clusters attributes as a whole, i.e. Feature 3 only has a relationship to metadata attribute 3, however it is represented in the MBMN Cluster 1 node, as the cluster as a whole has

relationships to attribute 3 and 1. The size of the Cluster 1 node is proportional the number of nodes that were in the cluster as created from.

- **Exclude Universal Nodes**

The option to exclude universal nodes will remove all features that have metadata attribute relationships to all included attributes.

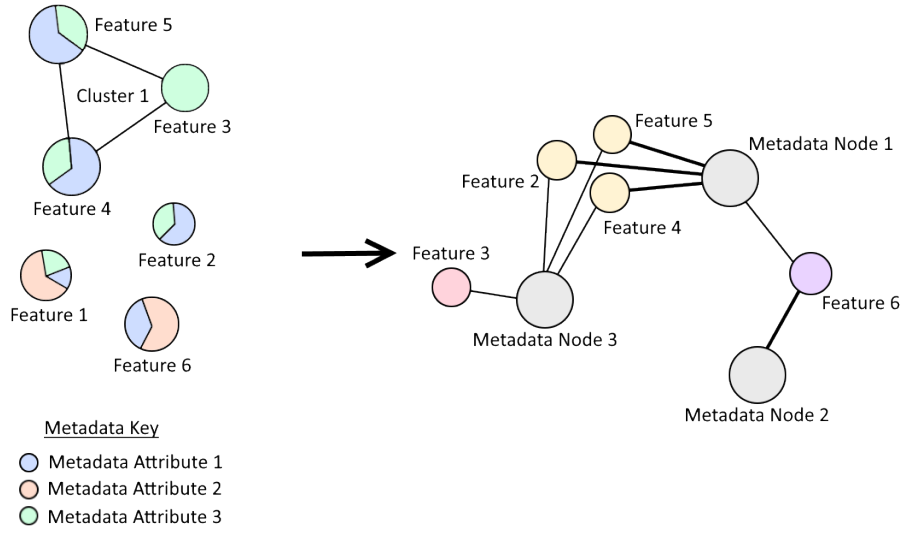


Figure X: Example of an MBMN that has used the option to remove all universal nodes (Feature 1) in the resulting network.

- **Max. Node-Attribute Edges**

This options allows for the restriction of features that have n+

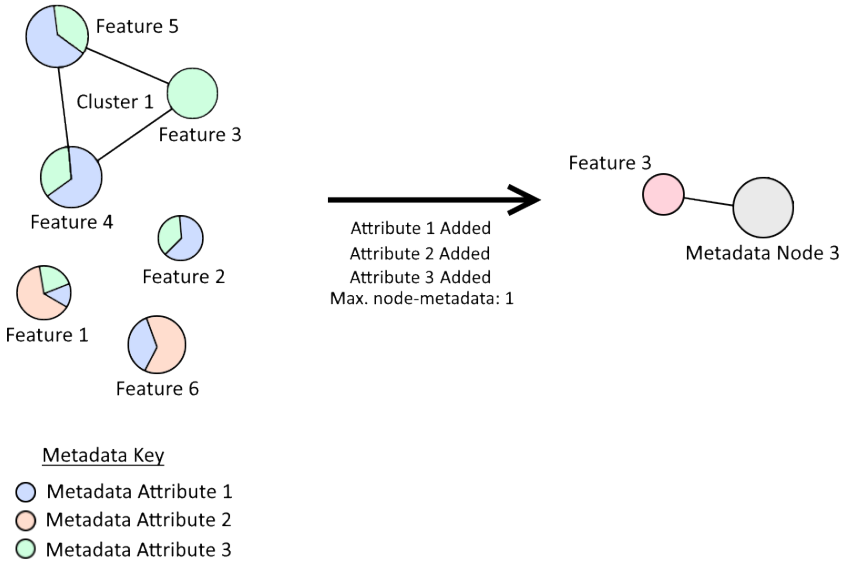


Figure X: Example of the resulting MBMN that has used the option to restrict the maximum number of allowed feature-to-metadata attribute relationships, in this case 1. Although Attribute 1 and Attribute 2 have been added to the MBMN, they have no nodes in the resulting network as there are no features that meet the requirement of a maximum of 1 feature-to-metadata attribute relationships (Features 1, 2, 4, 5 and 6 all have 2 or more feature-to-metadata attribute relationships).