MultiAlign Tutorial 04 – Reviewing An Analysis

BRIAN LAMARCHE
About this tutorial

- This tutorial will describe the graphical user interface (GUI) after an analysis has been performed in great detail.

- You will learn:
  - Shortcuts for drilling through the data
  - How to interact with plots
  - Details about every view
Analysis overview

THIS SECTION PROVIDES A SET OF BACKGROUND INFORMATION TO HELP GUIDE YOU THROUGH THE REST OF THE TUTORIAL
Previous tutorials covered this material

1. Home Screen
2. Select Data
3. Set Parameters
4. Select Baseline and Mass Tag Database
5. Set Analysis Path and Name
6. Running Analysis
7. Analysis View Window
GUI Basics

UNDERSTANDING HOW TO INTERACT WITH PLOTS
This section describes how to interact with each plot.

It is intended to provide basic to advanced functionality of the user interface.
Zooming can be done by:
- Left Click on plot
- Hold
- Drag
- Release to complete zoom
Zooming can be done by
- Left Click on axis
- Hold
- Drag
- Release to complete zoom
Selecting a Series

- A series is a set of points
- Clicking on a dot will select the series of points it is a part of

Left Click on a dot in the series
Context Menu (1)

- Right Click on the plot
- A context menu will appear
- Show All Data
  - Auto-zooms the plot to show all data points
- Copy All Data
  - Copies data from the plot to system clipboard for pasting into other software applications, e.g. Excel
- View Port
  - Allows you to step back to a previous zoom range
Context Menu (2)

- **Copy**
  - Saves a WMF version of the plot, with the plot data still in raster form

- **Copy Image**
  - Will copy a rasterize image to the system clipboard

- **Save Image**
  - Saves the plot as an image to file

- **Select All Series**
  - Selects all series of points

- **Selected Series**
  - Details what series of points are selected
Some plots will allow you to measure the distance between points 
- e.g. ppm and NET differences

To Measure:
- Hold Shift
- Left Click
- Drag
- Release left mouse button

The distance in the left window shows:
- 1.86 ppm mass difference
- .04 NET difference
Reviewing An Analysis

GLOBAL VIEW OF ANALYSIS
To open an existing analysis, click the “Open Analysis” Link.
Or choose a recent analysis by clicking on the database link.
Review of output

- Remember that MultiAlign stores its results in a directory.

- Alignment related plots must exist in the Plots subfolder.

- The analysis database file should also exist.

- This is important later on because MultiAlign will search for these plots to create visualization elements in the GUI.
This is the main window that should appear.

If it does not, click on the **analysis button** at the bottom of the screen.
If a mass tag database is used in the analysis, this tab allows investigation of each mass tag.
Investigating clusters

THE CLUSTER TAB AND SELECTED CLUSTER VIEWS
The clusters scatter plot window displays all of the clusters found in the analysis.

Each blue dot represents a feature found across multiple datasets.

This plot is interactive.
The clusters data grid shows all detailed information about each cluster.

The columns are sort-able by clicking on the column header.

You can also reorder the column orders by clicking on a column header and dragging to its new location.

Left clicking on a row in the cluster data grid will display the details about the cluster below.
Selected Cluster View Basics

- Visualization of features within cluster, matching mass tags, and nearby clusters
- Detailed Information about each feature in the cluster
- Abundance profile plots of cluster
- Any MS/MS spectra plots associated with cluster
Data grid that shows any matching mass tags or tags that are nearby within 6 ppm or .03 NET

Scatterplot that shows a cluster (red) and the features that comprise the cluster (blue)

**NOTE:** Shape is based on charge state

**NOTE:** Drift time plot can be hidden in the analysis

Advanced Tab
Data grid that shows any nearby clusters within a 6 ppm and 0.03 NET
## Tags and Clusters (3) – Shapes and Colors – Cluster Charts

<table>
<thead>
<tr>
<th>Shape</th>
<th>Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>●</td>
<td>+1</td>
</tr>
<tr>
<td>■</td>
<td>+2</td>
</tr>
<tr>
<td>▲</td>
<td>+3</td>
</tr>
<tr>
<td>★</td>
<td>+4</td>
</tr>
<tr>
<td>✗</td>
<td>+5</td>
</tr>
</tbody>
</table>

### Color Based on Type of Feature

- **LC-MS Features from selected cluster**
- **LC-MS Features from nearby cluster (hollow)**
- **Selected Cluster Centroid**
- **Nearby Cluster Centroid**

### Features and Cluster Centroids

- **Matched mass tag**
- **Unmatched mass tag nearby**

### Mass Tags

- **Matched Mass Tag**
  - LC-MS Features from nearby cluster (hollow diamond) – nearby mass tag unmatched

### Table: Tags and Clusters

<table>
<thead>
<tr>
<th>Nearby Clusters</th>
<th>Matching Mass Tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Monoisotopic Mass</td>
</tr>
</tbody>
</table>

### Diagram

- Scatter plots showing mass tags and clusters with annotations for matched and unmatched mass tags.

### Buttons
- **Save Feature Plots**
Features (1)

- Detailed information about each feature in cluster (consensus feature)
- Number of MS Features (i.e., number of parent scans) feature appeared in
- Number of MS/MS spectra associated with feature

Monoisotopic Mass vs. MS-Scan scatter plot for each charge state for the selected feature

LC-chromatogram of each charge state (depicted by color) for the selected feature
Features (2) – MS/MS

Red bar displays where on the LC-chromatogram the fragmentation occurred

Pre-cursor m/z and MS/MS scan
Abundance Profile

This slide currently shows the abundance (raw) from each dataset. Each dot represents the abundance for a feature from the dataset id below.

This plot is interactive

Future feature: organize each dataset by factor (e.g. age or time point)
Information about all MS/MS spectra related to selected cluster

Full MS/MS spectra
Investigating mass tag matches

USING A MASS TAG CENTRIC APPROACH, FINDING HOW MANY CLUSTERS MATCH
The mass tag plot shows all mass tags present in the AMT Tag database used for alignment and/or peptide identification via STAC.

Mass Tags are represented by orange diamonds.

This plot is interactive.
The mass tag histogram shows how many features (e.g. a peptide from an individual dataset) to the mass tag.

This plot is interactive.
The mass tag data grid shows detailed information about each mass tag.

Left clicking on a row in the mass tag data grid displays detailed information about the mass tag.
Detailed information about the mass tag
Selected Mass Tag Detail View

Synopsis about the mass tag

Visualization of mass tag and matching clusters (shown below)

Detailed information about the protein the mass tag is part of (proteomics only)
Quick information about the mass tag

Peptide sequence information, cleavage states (2 = fully tryptic, 1 = partially tryptic), and modification detail

Matched protein and cluster detail. A future feature is under development that allows users to drill back into cluster and protein detail.
Mass Tag - Tags and Clusters (1)

Clusters that match to this mass tag

And on other tab, nearby mass tags

Scatter plot that shows all matching clusters (red) and their features that comprise the cluster (blue)

Selected mass tag is displayed as orange diamonds. Nearby mass tags are displayed as hollow diamonds

NOTE: Feature and Cluster shapes are based on charge state

Drift time plot can be hidden in the analysis Advanced Tab
Mass Tag - Tags and Clusters (2)

Shape  Charge
●  +1
■  +2
▲  +3
▼  +4
×  +5

Color Based on Type of Feature
LC-MS Features of matched clusters
LC-MS Features from nearby cluster (hollow)
Matched Cluster Centroid
Nearby Unmatched Cluster Centroid

Features and Cluster Centroids

Selected mass tag
Nearby mass tags

Mass Tags

Selected Mass Tag
LC-MS Features from nearby cluster (hollow diamond) – nearby mass tag unmatched

Tags and Clusters
Features
Abundance Profile
MS/MS

Nearby Clusters  Matching Mass Tags
ID  Monoisotopic Mass  NET  Drift Time  Charge  Dif

Save Feature Plots
Details about the proteins that this mass tag rolls into. Shown here is multiple mass tags because the mass tag (peptide in this case) selected has multiple conformations.
Global Statistics Plots

UNDERSTANDING THE GLOBAL STATISTICS
Histogram of the number of datasets represented by each cluster

Histogram of the cluster sizes (number of features per cluster)

Histogram of charge states
Histogram of the number of datasets represented by each cluster

Histogram of the cluster sizes (number of features per cluster)

Histogram of charge states

Histogram of mass tags that match to individual features (sum of cluster sizes for clusters that match to a mass tag)

NOTE: The differences in the plots is from analyzing two different datasets and the total number of datasets (150 for this slide vs. 3 for the previous)
Datasets Plot

DATASET INFORMATION
Datasets Plot View

Allows the user to toggle between dataset plot view and data grid mode.
These plots are created from the high resolution images stored in the Plots directory where the database was created.

*Future feature will be to make these plots interactive*
Dataset Plot View

LC-MS Feature Scatter plot

Mass error histogram

Mass vs. Scan Residuals

NET vs. scan residuals

LCMSWarp alignment heatmap

Mass (m/z) vs. scan residuals

NET error histogram
Advanced Tab

APPLICATION CONFIGURATION
This view allows you to customize the application and visualization defaults.

This view currently has a minimal set of features displayed.
For more information see the MultiAlign website:

http://omics.pnl.gov/software/MultiAlign.php