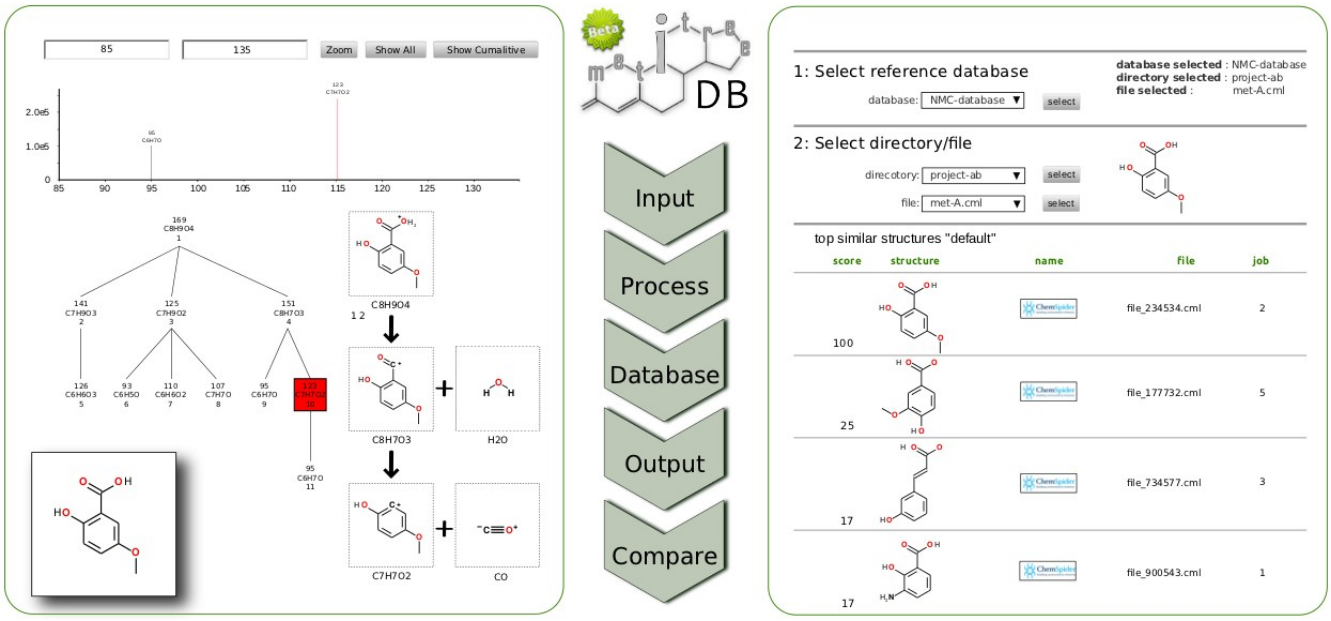
METITREE USER GUIDE

version 0.1
First edition: February 2012

This User Guide provides an overview of several functional features of the MetiTree web application. Use the content list to enter detailed information on how to use MetiTree.



MetiTree is a user-friendly, web application dedicated to organize, process, share, visualize, and compare high resolution multi-stage mass spectrometry (MS^n) data. It integrates dedicated features to export and visualize complex MS^n data, facilitating the exploration and interpretation of metabolomics experiments.

Table of Contents

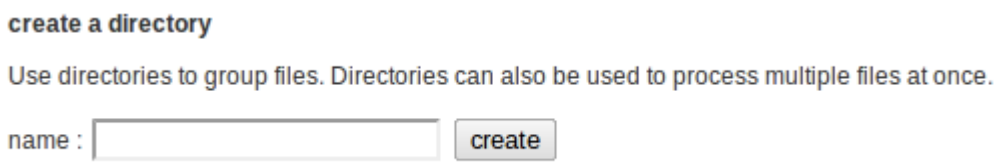
Upload MS^n data:

Upload files with MS^n mass spectrometry data and store them into data directories.

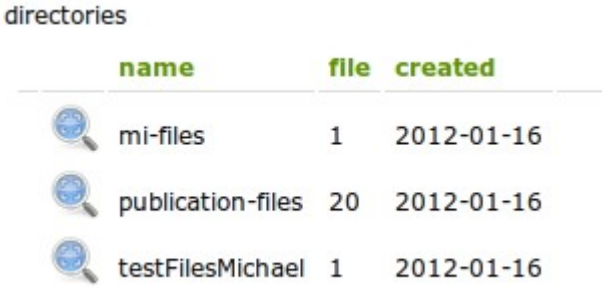
MetiTree accepts right now the uploading of mzXML files (version 3.1) containing MS^n mass spectrometry data. MetiTree allows the creation of directories for grouping mzXML files, assisting the organization of the data according projects or topics.

1. Select "**MS^n data**" (tab), to open the "**MS^n data**" page.
2. From the '**MS^n data**' you can perform the functions listed below:
   - Create a new directory
   - Modify an existing directory

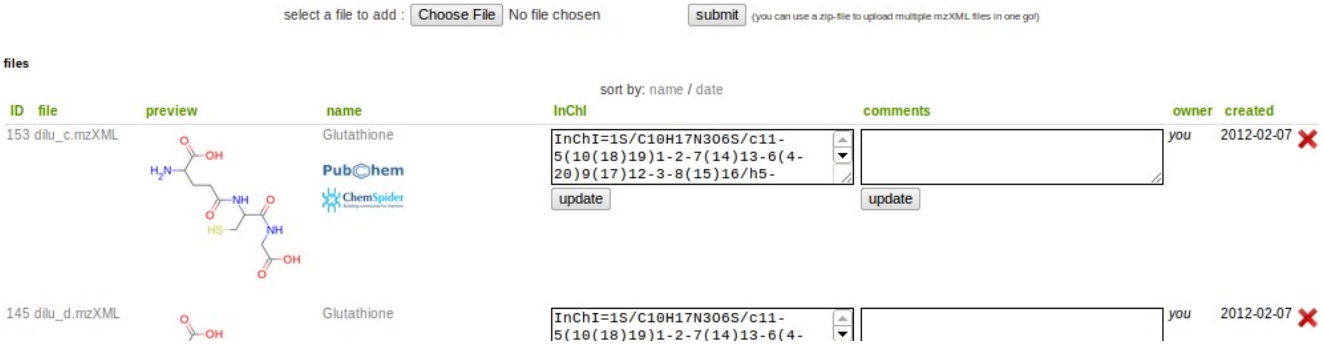
3. To create a new directory, first provide a name for the new directory. Click '**create**' to save your directory.



4. To modify an existing directory, select one of the directories by clicking the magnifying glass.



5. The page "**Files**" will be displayed. The files added to the directory will be listed.



6. Select "**Choose File**" to add new files from your desktop to this directory. Finally, select "**submit**" to save it.

You can upload a single mzXML file or a zip-file with multiple mzXML files.

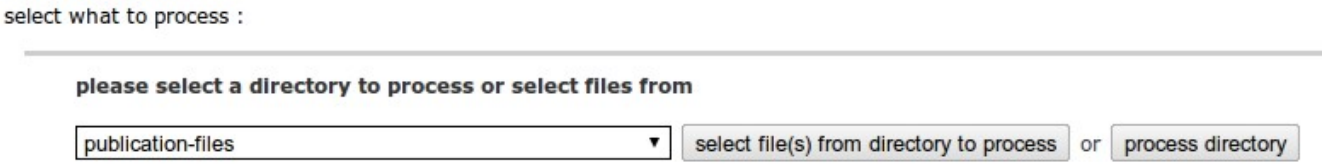
7. Once the file is uploaded you eventually can specify the InChI identifier of the neutral compound. To save the InChI identifier you have to select the "**update**" button. Automatically, the file will be cross-linked to the pubchem and chemspider databases.
8. The files can be sorted by name or data created by selecting "**name/date**".

Process and visualize

MS^n data:

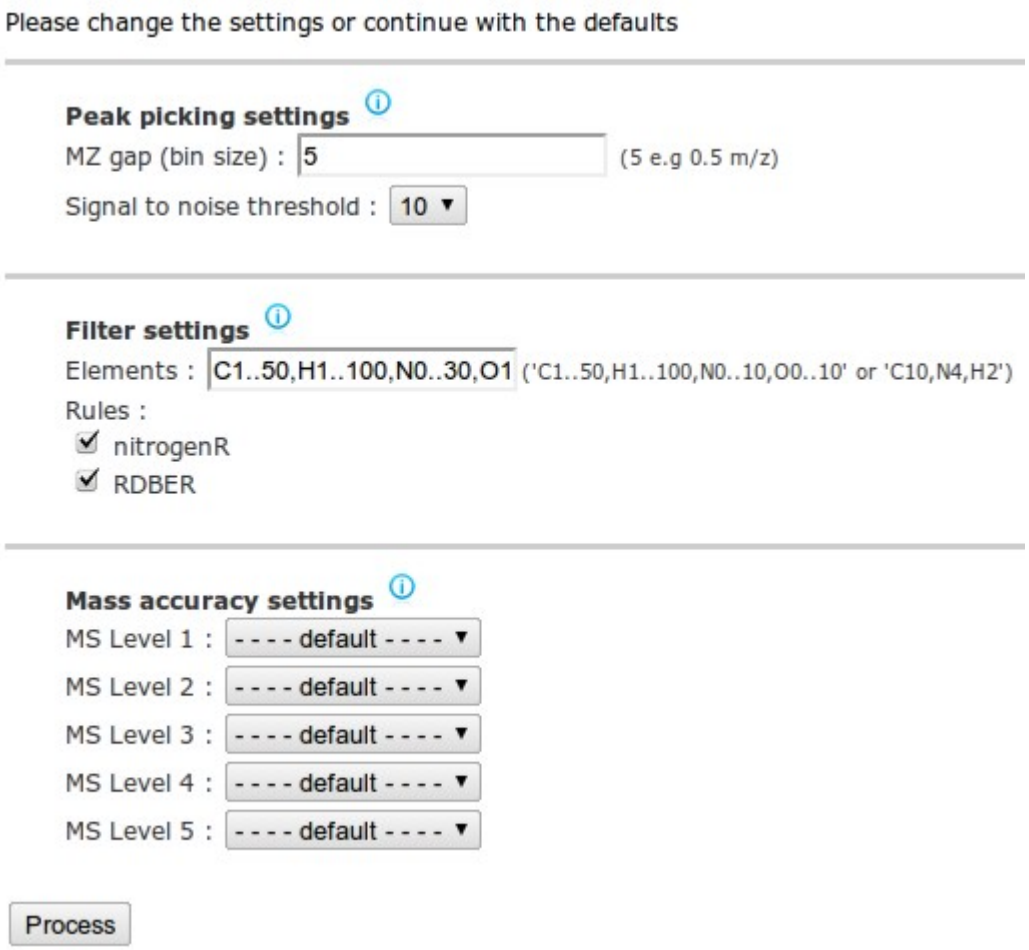
MetiTree integrates the MEF tool (Rojas-Chertó et al., 2011) and uses its functionality to process MS^n data. The required input to process MS^n data are mzXML files (http://en.wikipedia.org/wiki/MzXML) and the settings of the processing parameters. Processing parameters are grouped into those to extract the mass spectrometry information (m/z, intensity, and scan number) and those to enrich the MS data with chemical information (chemical elements and number of atoms). MetiTree allows the processing of individual or multiple files at the same time. Furthermore, the same mzXML file can be processed several times with different sets of processing parameters. Results and parameters information are stored to allow posterior revision.
Once the data is processed, it can be displayed using the spectral tree viewer. Here, the user can select a peak or a node, which represents a fragment. For each selection, the corresponding mass spectrum is displayed, together with the reactions that connect the parent ion with the selected fragment. The structures of the fragments are only visible when they have been assigned. The results generated by MetiTree can be exported to different file formats for further analysis. The CML format allows managing complex chemical content that other software can import to process it even further. The PDF format can be used to visualize the MS^n data, which makes it suitable to export images into reports or publications.

1. On the left navigation bar select "**Process**" (tab)
2. To process a file you must first select the directory where the file is found. To process all files of a determined directory select "**process directory**".



3. If you want to process specific files from the directory, select "**select file(s) from directory to process**". A list of files will be displayed. Check those files to process and select "**process files**". If you like to select all files at onces press the box with the sign plus.

4. The "**Define settings**" page will be displayed introducing the following processing settings:
    - MZ gap (bin size) = Minimal distance between adjacent peaks. Peaks with smaller intensity being excluded.
    - Signal to noise threshold = Signal to noise threshold allowed during the peak picking.
    - Elements = It defines upper and lower number of atoms admitted for each element in the elemental composition of the ion (e.g. C1..15, H1..9, O0..4).
    - Rules = It defines the constraint-rules that are applied to verify the elemental compositions generated. It distinguishes between the Nitrogen Rules and the RDB rules.
    - Mass accuracy settings = Mass accuracy error allowed per MS level.

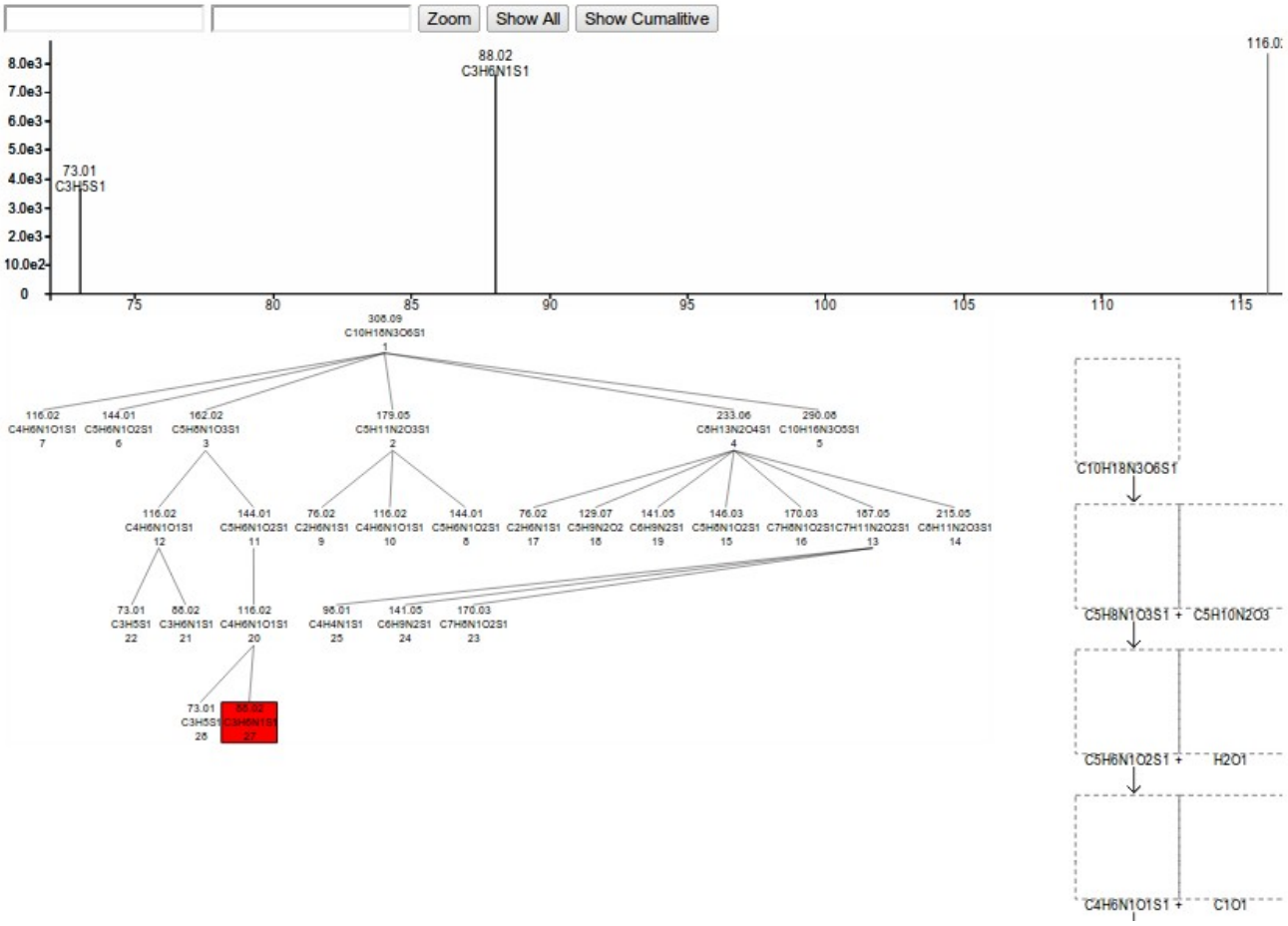


5. To submit the process select "**Process**".

6. The "**Process**" page will be displayed containing information about the progress of the different processing jobs.
7. It is possible to visualize the different results.
    - xcms = output generated by XCMS after the peak picking extraction.

- cml = chemical markup language file enriched with chemical information.
- trees = Tree Viewer interconnects three MS^n items: the spectrum, which contains mass peaks, the fragmentation tree, which contains fragment nodes/elemental compositions, and the fragmentation reactions, which contain structures.
- master tree = A viewer that filters the tree file based on the occurrence of each fragment.
- csv = comma separated values file with enriched chemical information
- pdf = portable document format to visualize the spectral data.

8. Select "**master tree**" to visualize spectral data with the tree viewer



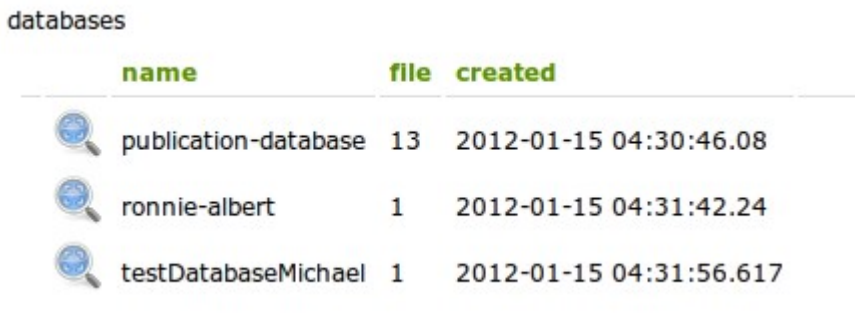
Storage and manage database(s) with MS^n data

Processed MS^n data can be stored in one or multiple internal databases. Because the users are organized in groups, they can share files and libraries with other group members. Only the administrator user has the faculty to create a new group and assign users to it.

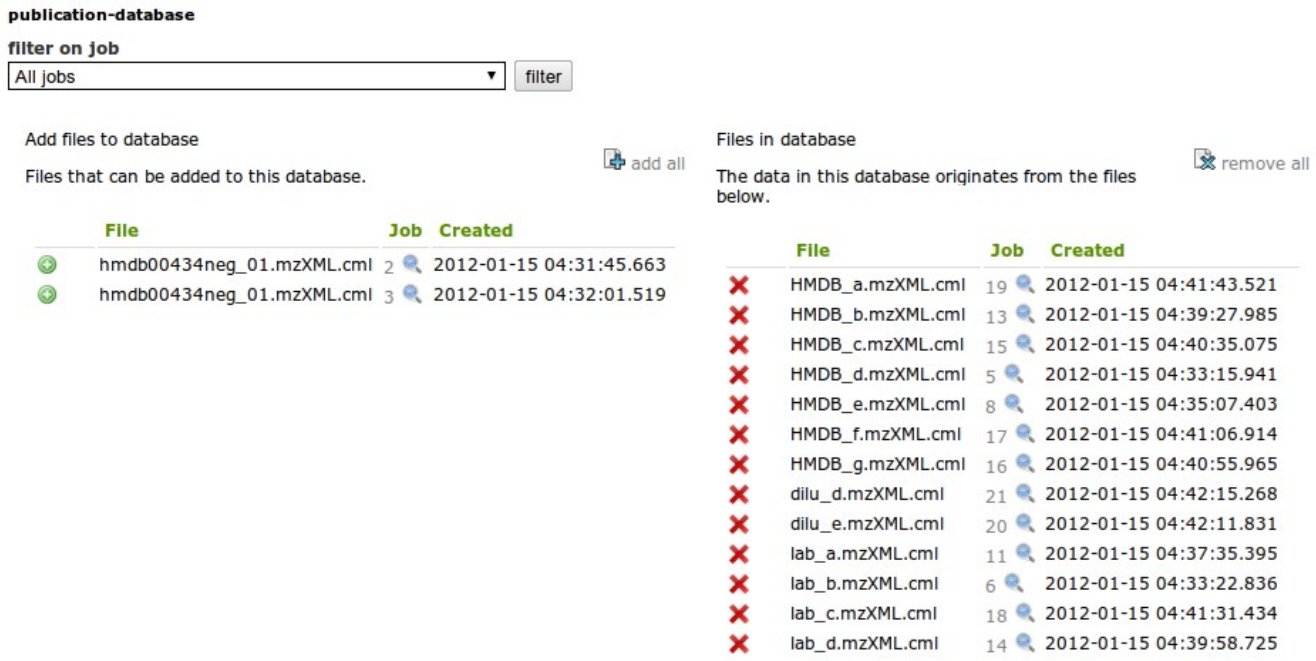
1. On the left navigation bar select the “**Databases**” (tab)
2. The '**Databases**' page will be displayed. From this page, you have access to the functions listed below:
    - Create a new database
    - Modify an existing database

3. To create new database, provide a name for the new database and click '**add**' to save your database.

4. To modify an existing database, select one of  the magnifying glasses in front of the databases.



5. Now modify the files to be contained in the database as required.



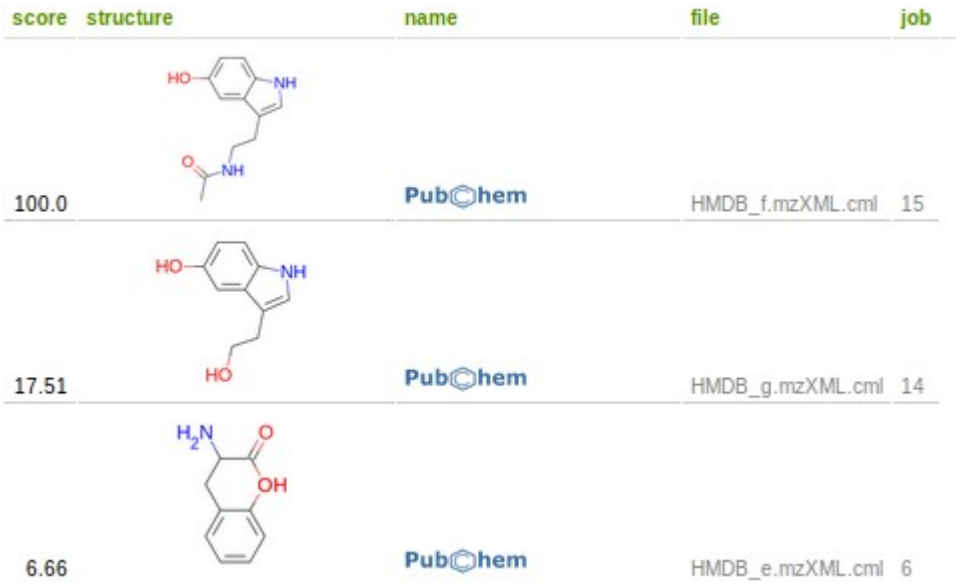
6. On the right side you have the files that actually are in the database. On the left side you have those files that you can upload to the database. Through "**filter on job**" you can filter the files by specific job number.

Querying an existing database(s) for similar MS^n data:.

MetiTree integrates the functionality to query for similar MS^n data stored in your own personal library. The results are presented in a list showing the structures of the compounds with most similar MS^n data together with the corresponding similarity value. The results are ranged between 0-100. A value near to 100 indicates that the MS^n data is highly similar; while a value close to 0 illustrates that it is very different (Rojas-chertó et al., 2012). We expect MetiTree to contribute two fold in the identification of metabolites. On one hand, we are interested to know if a fragmentation tree of the same compound is present in the library, which would be a partial identification. On the other hand, if similar fragmentation data is found, we can use this knowledge to give hints (e.g. substructure information) which structure the unknown compound could have.

1.
2. Select the "**search/compare**" (tab), to open the "**database search**" page.
3. Choose a database to query "**database**", and the spectral tree file to be compared in "**directory**" > "**file**".

4. The '**database search**' page will display the results into a list of compounds ranked



from more similar to a less similar. On the left side it will shows the similarity value when the fragmentation trees are compared to each other.

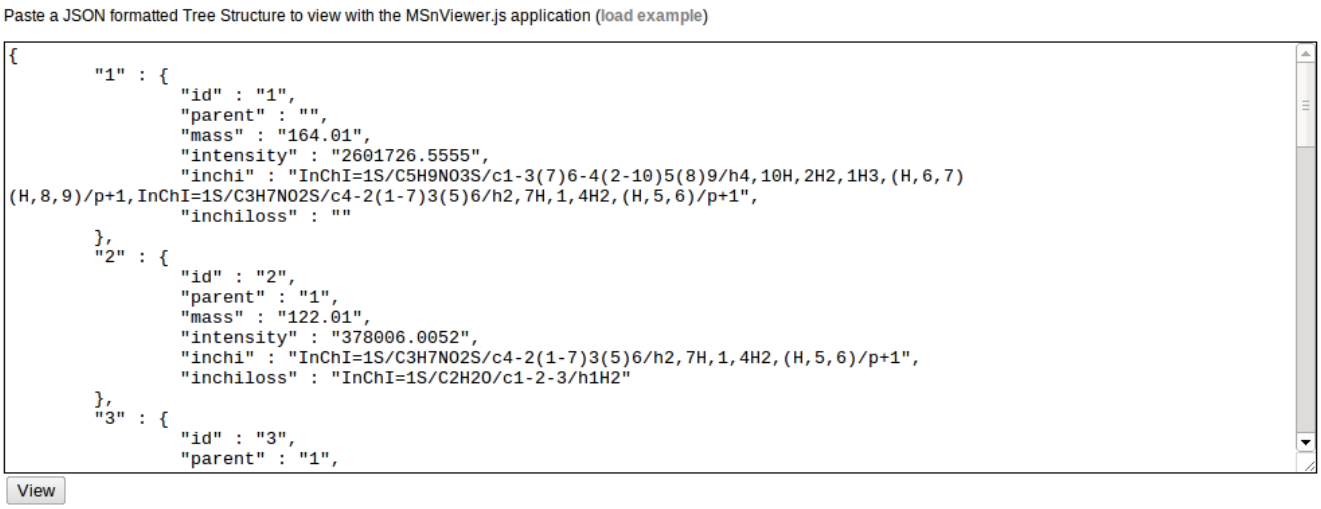
Visualize MS^n data with chemical structure of the fragments:

The spectral tree viewer is capable to show the chemical structure of the fragments in the MS^n data. At the moment it is only possible if you supply it with JSON file describing the MS^n data.

1. On the bottom of any page, select the "**MsnViewer Demo**".

2. You can upload an example by selecting "**Load example**" > "**view**".

You can modify the json message and add your own values.



3. After the selected view run, the MS^n viewer screen will be shown.

FAQs:

Where can I address for more help?
For any questions that could not be answered through this guide, you can send us an email directly to info@metitree.nl. We will answer you as soon as possible.

Can I maintain my data permanently in MetiTree?
Due to MetiTree is in a beta version every weekend all uploaded data or processed data is removed from the server. We expect that in a new release we will provide a persistent application.

How many files might I upload to process?
At the moment you can only upload X number of files. We must limit the uploading process with a small number to avoid collapsing of our server.

Rojas-Chertó, M., Kasper, P. T., Willighagen, E. L., Vreeken, R., Hankemeier, T., & Reijmers, T. (2011). Elemental Composition determination based on MSn. *Bioinformatics*, *27*(17), 2376–2383. Oxford Univ Press. doi:10.1093/bioinformatics/btr409

Rojas-chertó, M., Peironcely, J. E., Kasper, P. T., Van, J. J. J., Hooft, D., Vos, R. C. H. D., Vreeken, R., et al. (2012). Metabolite identification using automated comparison of high resolution MSn spectral trees. *Anal. Chem*.