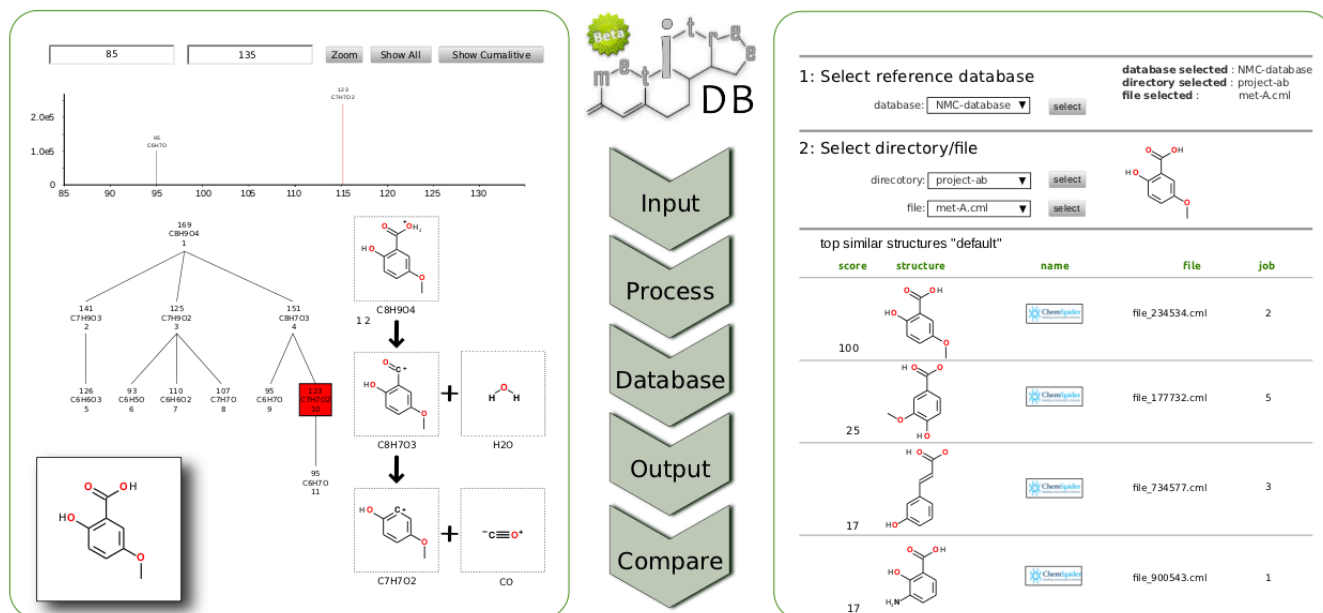


METITREE USER GUIDE

version 0.4
Fourth edition: June 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.....	3
Creating a new directory	3
Add or remove files	3
Extended view	4
Process and visualize.....	5
Storage and manage database(s) with MS ⁿ data	9
Querying an existing database(s) for similar MS ⁿ data.....	11
Visualize MS ⁿ data with chemical structure of the fragments	12
FAQs.....	13
Biography:	14

Upload MSⁿ data

Upload files with MSⁿ mass spectrometry data and store them into data directories.

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

Creating a new directory

1. Select “**MSⁿ data**” (tab), to open the “**MSⁿ data**” page.
2. First provide a name for the new directory then click '**create**' to save your directory.

create a directory


Use directories to group files. Directories can also be used to process multiple files at once.

name :

Add or remove files

1. Select one of the directories by clicking the magnifying glass.

directories

	name	file	created
	metabolites	19	2012-06-28

2. The page “**Files**” will be displayed. The files added to the directory will be listed.

files in: *metabolites* (change directory / add files)

				sort by: name / date			
ID	file	name		pubchem	chemspider	owner	delete
 19	dilu_c.mzxml	Glutathione				you	
 18	dilu_d.mzxml	Glutathione				you	
 17	hmdb_a.mzxml	6-hydroxy-m-Anisic acid				you	
 16	lab_c.mzxml	5-Hydroxylysine				you	
 15	hmdb_f.mzxml	N-Acetylserotonin				you	

3. Select “**add files**” to add new files from your computer to this directory. Finally, select “**submit**” to save it.

add file(s)

select a file to add :

Browse...

submit

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

An mzXML file is an open data format for storage and exchange of mass spectroscopy data, developed at the SPC/Institute for Systems Biology. All known converters to convert vendor specific files to mzXML files are summarized on this page:

<http://tools.proteomecenter.org/wiki/index.php?title=Formats:mzXML>

Extended view

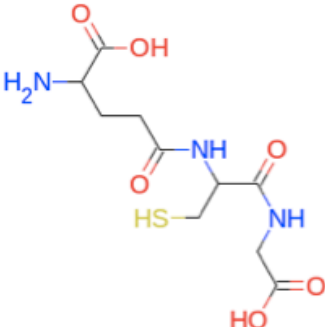
To see extended information of a file, click the magnifying glass.

1. Once the file is uploaded you eventually can specify the InChI (IUPAC International Chemical 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.
2. The files can be sorted by name or data created by selecting “**name/date**”.

sort by: name / date

ID	file	name	pubchem	chemspider	owner	delete
19	dilu_c.mzxml	Glutathione	PubChem	ChemSpider	you	X

2d image



InChI

```
InChI=1S/C10H17N3O6S/c11-5(10(18)19)1-2-7(14)13-6(4-20)9(17)12-3-8(15)16/h5-6,20H,1-4,11H2,(H,12,17)(H,13,14)(H,15,16)(H,18,19)/t5-,6-/m0/s1
```

update/save

comments

update/save

Process and visualize

MSⁿ data:

MetiTree integrates the MEF tool (Rojas-Chertó et al., 2011) and uses its functionality to process MSⁿ data. The required input to process MSⁿ data is an mzXML file (<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 MSⁿ data, which makes it suitable to export images into reports or publications.

1. On the 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**”.

select what to process:

please select a directory to process or select files from

metabolites

select file(s) from directory to process

or

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 once press the box with the sign plus.

files in directory *metabolites*

Process files



file

comments

created



dilu_a.mzxml

2012-06-28



dilu_b.mzxml

2012-06-28



dilu_c.mzxml

2012-06-28



dilu_d.mzxml

2012-06-28

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.

Please change the settings or continue with the defaults

Peak picking settings ⓘ

MZ gap (bin size as m/z) : m/z (5 e.g 0.5 m/z)

Signal to noise threshold : %

Filter settings ⓘ

Elements : ('C1..50,H1..100,N0..10,O0..10' or 'C10,N4,H2')

Rules :

☒ nitrogenR

☒ RDBER

Mass accuracy settings (ppm) ⓘ

MS Level 1 :

MS Level 2 :

MS Level 3 :

MS Level 4 :




MS Level 5 :

5. To submit the process select “**Process**”.

6. The “**Process**” page will be displayed containing information about the progress of the different processing jobs.

Status JOB#000000000010 (...finished!)

Output


name	compound	mzxml 	cml 	export (csv/tab) 	visuals 	created
hmdb_d.mzxml	3-Aminosalicylic acid 					2012-06-28

Settings

setting	value
accuracy-L1	15
accuracy-L2	15
accuracy-L3	15
accuracy-L4	15
accuracy-L5	15
elements	O3..3,N1..1,H8..8,C7..7
mzgap	0.5
rulenitrogenr	nitrogenR
rulerdber	RDBER
snthresh	1.0

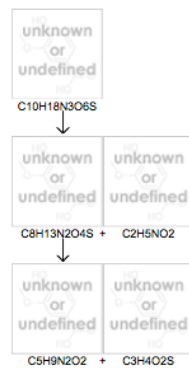
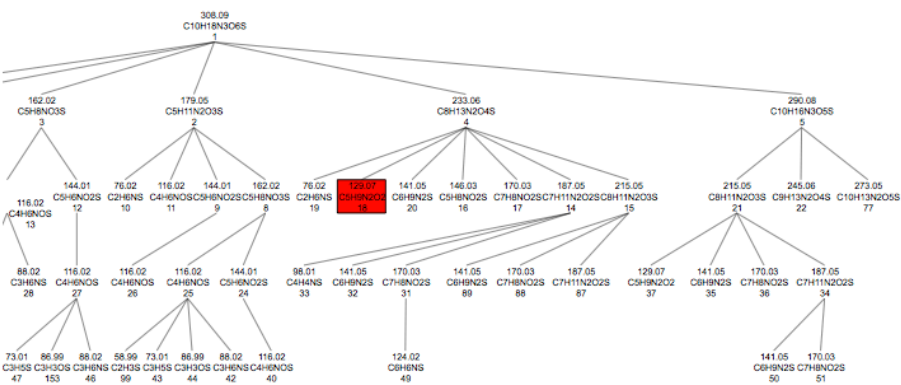
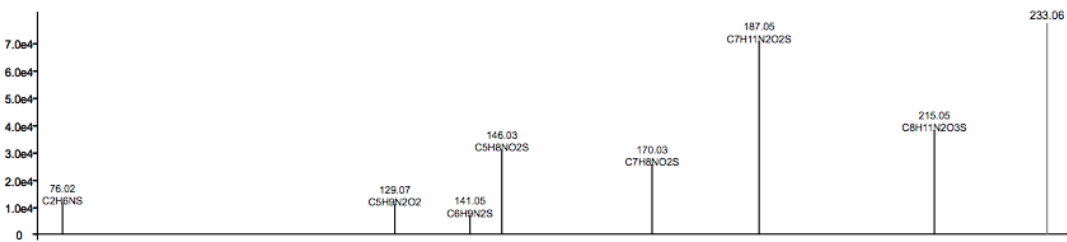
Job Log

```
#import job
```

7. It is possible to visualize the different results.
- mzXML = original mzXML file used to process the spectral data
 - cml = chemical markup language file enriched with chemical information.
 - csv/tab = plain text files with enriched chemical information
 - viewer = The Spectral Tree Viewer (<https://trac.nbic.nl/brsp201017>)
 - pdf = portable document format to visualize the spectral data
8. Select  to visualize spectral data with the tree viewer. The Spectral Tree Viewer interconnects three MSⁿ items: the spectrum, which contains mass peaks, the fragmentation tree, which contains fragment nodes/elemental compositions, and the fragmentation reactions, which contain structures.

cumulative spectra

Spectrum (x-axis: m/z , y-axis: absolute intensity)




Storage and manage database(s) with MSⁿ data

Processed MSⁿ 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 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.

add a new database

name :

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

databases

	name	created
	publicDb	2012-06-28

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

publication-database

filter on job

All jobs

Add files to database

Files that can be added to this database.



File	Job	Created
hmdb00434neg_01.mzXML.cml	2	2012-01-15 04:31:45.663
hmdb00434neg_01.mzXML.cml	3	2012-01-15 04:32:01.519

Files in database

The data in this database originates from the files below.



File	Job	Created
HMDB_a.mzXML.cml	19	2012-01-15 04:41:43.521
HMDB_b.mzXML.cml	13	2012-01-15 04:39:27.985
HMDB_c.mzXML.cml	15	2012-01-15 04:40:35.075
HMDB_d.mzXML.cml	5	2012-01-15 04:33:15.941
HMDB_e.mzXML.cml	8	2012-01-15 04:35:07.403
HMDB_f.mzXML.cml	17	2012-01-15 04:41:06.914
HMDB_g.mzXML.cml	16	2012-01-15 04:40:55.965
dilu_d.mzXML.cml	21	2012-01-15 04:42:15.268
dilu_e.mzXML.cml	20	2012-01-15 04:42:11.831
lab_a.mzXML.cml	11	2012-01-15 04:37:35.395
lab_b.mzXML.cml	6	2012-01-15 04:33:22.836
lab_c.mzXML.cml	18	2012-01-15 04:41:31.434
lab_d.mzXML.cml	14	2012-01-15 04:39:58.725

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 MSⁿ data

MetiTree integrates the functionality to query for similar MSⁿ data stored in your own personal library. The results are presented in a list showing the structures of the compounds with most similar MSⁿ data together with the corresponding similarity value. The results are ranged between 0-100. A value near to 100 indicates that the MSⁿ data is highly similar which would be a complete identification; 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. Select the “**search/compare**” (tab), to open the “**database search**” page.
2. Choose a database to query “**database**”, and the spectral tree file to be compared in “**directory**” > “**file**”. To see the database search result push “**search/refresh**” button.
3. 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.

1: select reference database:

database : publicDb

2: select query directory/file:

directory : metabolites

file : hmdb_f.cml - JOB#000C

database selected : publicDb

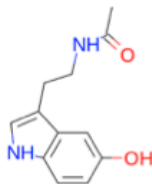
directory selected : metabolites

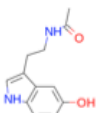
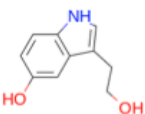
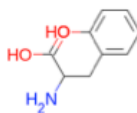
file selected : hmdb_f.cml

compound : N-Acetylserotonin

InChIkey : MVAWJSIDNICKHF-UHFFFAOYNA-N

InChI : InChI=1/C12H14N2O2/c1-8(15)13-5-4-9-7-14-12-3-2-10(16)6-11(9)12/h2-3,6-7,14,16H,4-5H2,1H3,(H,13,15)



score	structure	name	file	job
100.0		N-Acetylserotonin PubChem ChemSpider	hmdb_f.cml	13
16.66		3-(2-Hydroxyethyl)-1H-indol-5-ol PubChem ChemSpider	hmdb_g.cml	4
7.17		2-Hydroxyphenylalanine PubChem ChemSpider	hmdb_e.cml	1

Visualize MSⁿ data with chemical structure of the fragments

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

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



The Netherlands Metabolomics Centre in collaboration with The Netherlands Bioinformatics Centre (NBIC). A collaborative institute of the bioinformatics groups in the Netherlands.

Wiki - Roadmap - MsnViewer Demo
All rights reserved. Copyright © 2008 - 2012

2. You can upload an example by selecting “**Load example**” > “**view**”.
You can modify the json message and add your own values.

Paste a JSON formatted Tree Structure to view with the MsnViewer.js application (load example)

```
{
  "1" : {
    "id" : "1",
    "parent" : "",
    "mass" : "164.01",
    "intensity" : "2601726.5555",
    "inchi" : "InChI=1S/C5H9N03S/c1-3(7)6-4(2-10)5(8)9/h4,10H,2H2,1H3,(H,6,7)
(H,8,9)/p+1,InChI=1S/C3H7N02S/c4-2(1-7)3(5)6/h2,7H,1,4H2,(H,5,6)/p+1",
    "inchiloss" : ""
  },
  "2" : {
    "id" : "2",
    "parent" : "1",
    "mass" : "122.01",
    "intensity" : "378006.0052",
    "inchi" : "InChI=1S/C3H7N02S/c4-2(1-7)3(5)6/h2,7H,1,4H2,(H,5,6)/p+1",
    "inchiloss" : "InChI=1S/C2H20/c1-2-3/h1H2"
  },
  "3" : {
    "id" : "3",
    "parent" : "1",
```

3. After the selected view run, the MSⁿ 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.

How many files might I upload to process?

At the moment there is not a limit during the uploading process. If the number is big may be you can collapse the server.

Biography:

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., Hankemeier, T., & Reijmers, T. (2012). Metabolite identification using automated comparison of high resolution MSn spectral trees. *Anal. Chem.*
doi:10.1021/ac2034216