Intro Guide to Analyzing Metabolomic Data using mzMine 2.53 and GNPS

MichiganState 2020 iGEM | Bioinformatics

Last Updated: 10/24/2020

Preface

This short document is meant to share our feature-based metabolomic networking workflow/protocol (from ingesting raw .mzML mass spec files to exporting data for analysis in GNPS) and does not include any statistical processes to validate our results. This guide is primarily targeted towards new users of mzMine/GNPS; and as such, we've included our raw data on our wiki so that you can follow this protocol. Furthermore, we assume readers understand the basic concepts of molecular networking prior to opening this doc; but just in case, we've linked more resources below that offer more comprehensive explanations of FBMN methods. Additionally, the parameters used for this analysis may not be appropriate for all situations, and we cannot guarantee when the information contained in this document will be deprecated. In anycase, we hope the information enclosed in this document will be informative to future iGEM teams looking to incorporate metabolomics analysis to their projects.

MZmine/GNPS FBMN Outline/Table of Contents/Summary w/specific parameters used

- Graphical Overview
- Prerequisites
- Ingest Data to MZmine (“Raw Data Methods” → “Raw data import”)
- Mass Detection (“Raw data methods” → “Feature Detection” → “Mass Detection”)
  - MS1 noise level: 1e4
  - MS2 noise level: 1e3
- ADAP chromatogram builder (“Raw data methods” → “Feature detection” → “ADAP Chromatogram Builder”)
  - min group size in # scan: 5
  - group intensity threshold: 3e6
  - min highest intensity: 3e6
  - m/z tolerance: 0.02 or 10 ppm
- Chromatogram deconvolution (“Feature list methods” → “Feature Detection” → “Chromatogram deconvolution”)
  - min peak height: 1e5
  - peak duration: 0 to 3 minutes
  - baseline level: 1e4
- Isotope Grouping (“Feature list methods” → “Isotopes” → “Isotopic peak grouper”)
  - m/z tolerance: 0.02 or 10 ppm
  - RT tolerance: 0.25
  - max charge: 2
- Joining aligner (““feature list methods” → “Alignment” → “Join aligner”)
  - m/z tolerance: 0.02 or 10 ppm
  - weight for m/z: 75
  - RT tolerance: 0.25
  - weight for RT: 25
- Peak filtering (aka “feature list row filter") (“Feature list methods” → “Filtering” → “Feature list rows filter”)
  - min peak in row: 5
  - m/z range: 75-1500
  - keep peaks with MS2 for GNPS: TRUE
  - reset peak number ID: TRUE
- Gap-filling: (aka “Peak finder”) (“Feature list methods” → “Gap filling” → “Peak finder”)
  - Intensity tolerance: 10%
  - m/z tolerance: 0.02 or 10 ppm
  - RT tolerance: 0.25
- *Manual Validation
- Export and Upload Data to GNPS

Graphical Overview

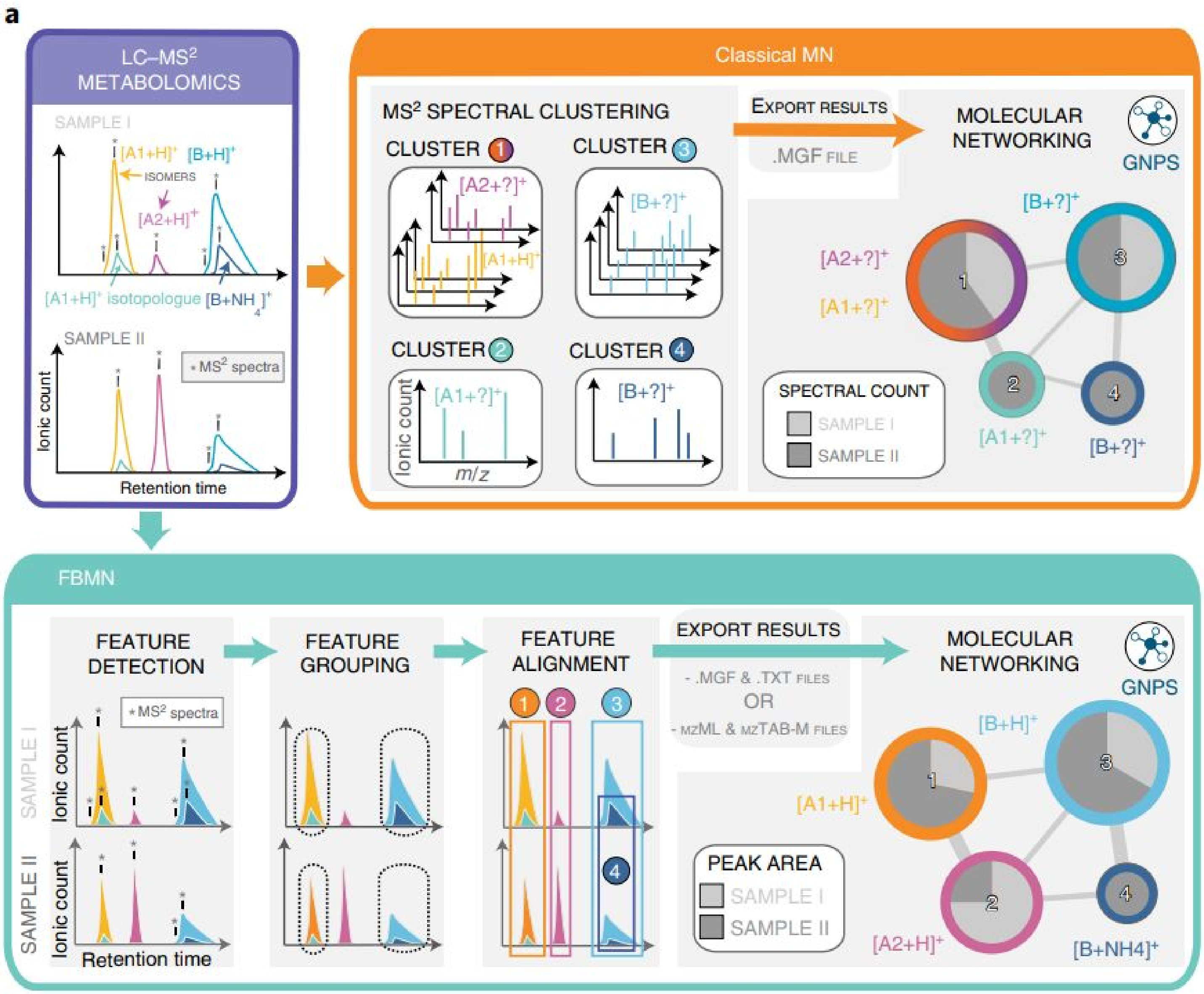


Figure 1A. Graphical overview of FBMN compared to Classical MN. Adapted [reprinted] from “Feature-based molecular networking in the GNPS analysis environment,” by Nothias et. al., 2020, *Nature Methods,17,* pp. 905-908.

Jump back to table of contents

Prerequisites

Learn about Feature-Based Molecular Networking

- This paper by Nothias et al. (2020) offers a great comparison between FBMN and classical networking and offers insight into how you can use FBMN within the GNPS Environment.
    - Nothias, L., Petras, D., Schmid, R. et al. Feature-based molecular networking in the GNPS analysis environment. Nat Methods 17, 905–908 (2020). https://doi.org/10.1038/s41592-020-0933-6
- This more extensive tutorial is a great follow-up example that goes more in-depth in more complex topics like manual validation of feature lists generated in the MZmine workflow.
    - **https://ccms-ucsd.github.io/GNPSDocumentation/featurebasedmolecularnetworking-with-mzmine2/**

Get mzMine

- Follow instruction here to download the latest version of MZmine: https://github.com/mzmine/mzmine2/releases
    - Installation isn't necessary as you initialize the application from start-up command/batch files for your respective operating system.
        - [Please note the images on this doc are generated on Windows, but the GUI does not change between operating systems]
    - More in depth documentation on starting and operating MZmine can be found within this manual:
        - https://docs.google.com/document/d/1JtdNwp1y6wz_Boz8qVGESH-kSrMhcGaJWR18iGmhstU/edit?usp=sharing

Set-up Account on GNPS

- https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp?redirect=auth
- More details about GNPS can be found in its Documentation page:
    - https://ccms-ucsd.github.io/GNPSDocumentation/

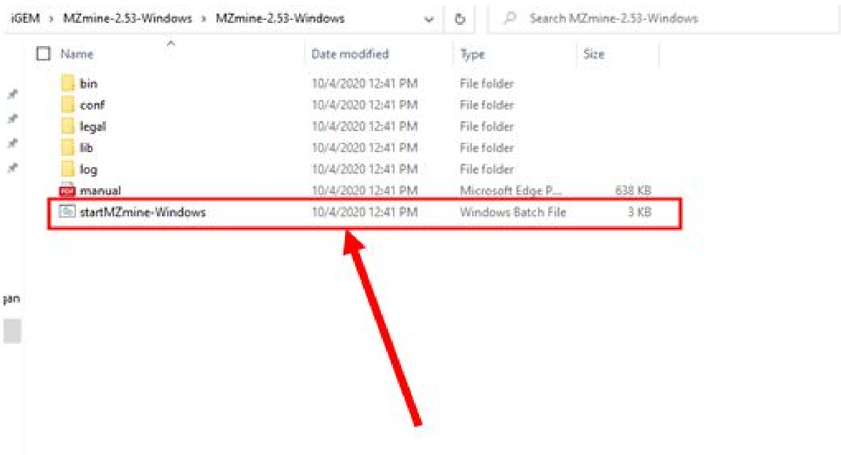
Setup file transfer client to upload data to GNPS

- Use these instructions to easily upload files to your GNPS account to analyze your samples.
    - https://ccms-ucsd.github.io/GNPSDocumentation/fileupload/
- FileSharing Clients make these file transfers easier:
    - Filezilla: https://filezilla-project.org/
    - WinSCP: https://winscp.net/eng/download.php

Jump back to table of contents

Ingest Data to MZmine

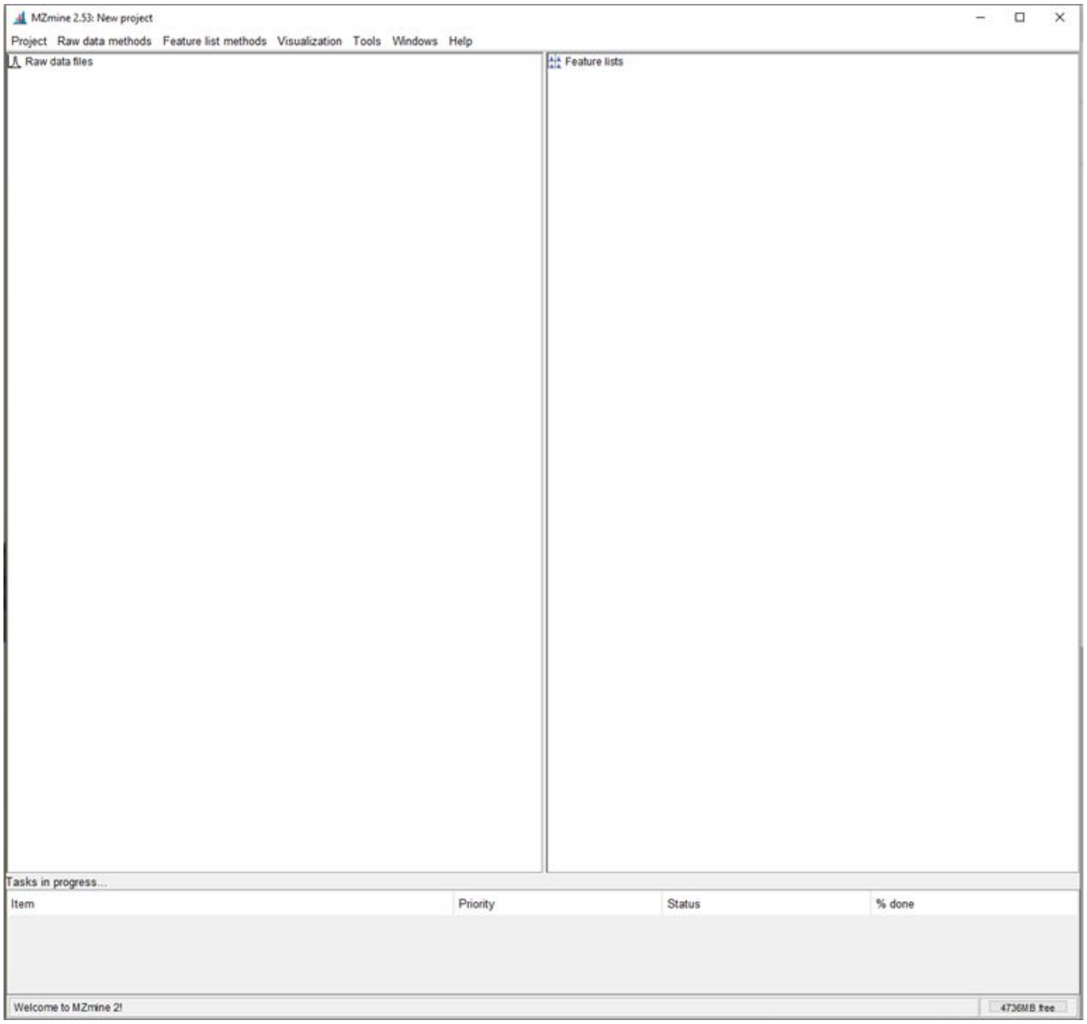
- After unzipping the mzMine program, initialize the program by clicking on the startMZmine script:



iGEM › MZmine-2.53-Windows › MZmine-2.53-Windows

Name	Date modified	Type	Size
bin	10/4/2020 12:41 PM	File folder	
conf	10/4/2020 12:41 PM	File folder	
legal	10/4/2020 12:41 PM	File folder	
lib	10/4/2020 12:41 PM	File folder	
log	10/4/2020 12:41 PM	File folder	
manual	10/4/2020 12:41 PM	Microsoft Edge P...	638 KB
startMZmine-Windows	10/4/2020 12:41 PM	Windows Batch File	3 KB

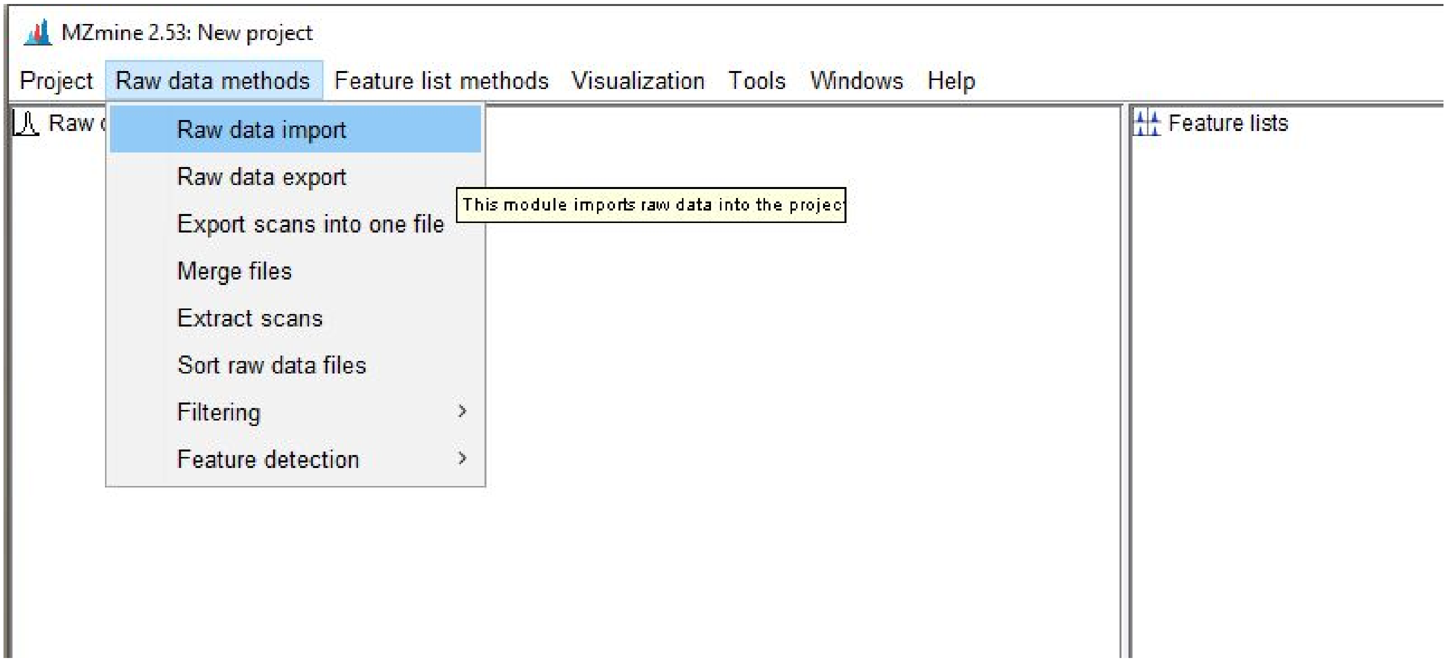
- Once initialized, you should see the following window:



Welcome to MZmine 2!

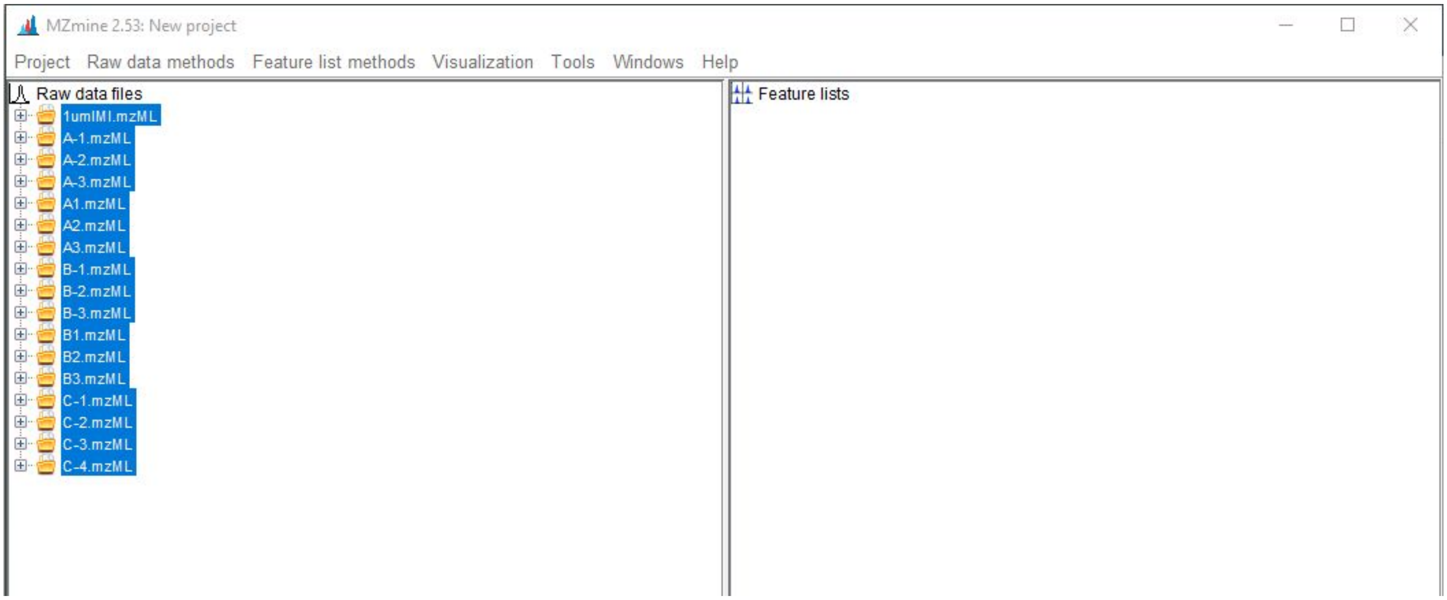
Import Raw Data Files

- Start importing your .mzXML files by clicking the “Raw data import” tool under “Raw Data methods” tab and selecting your files from your files:



This module imports raw data into the projec

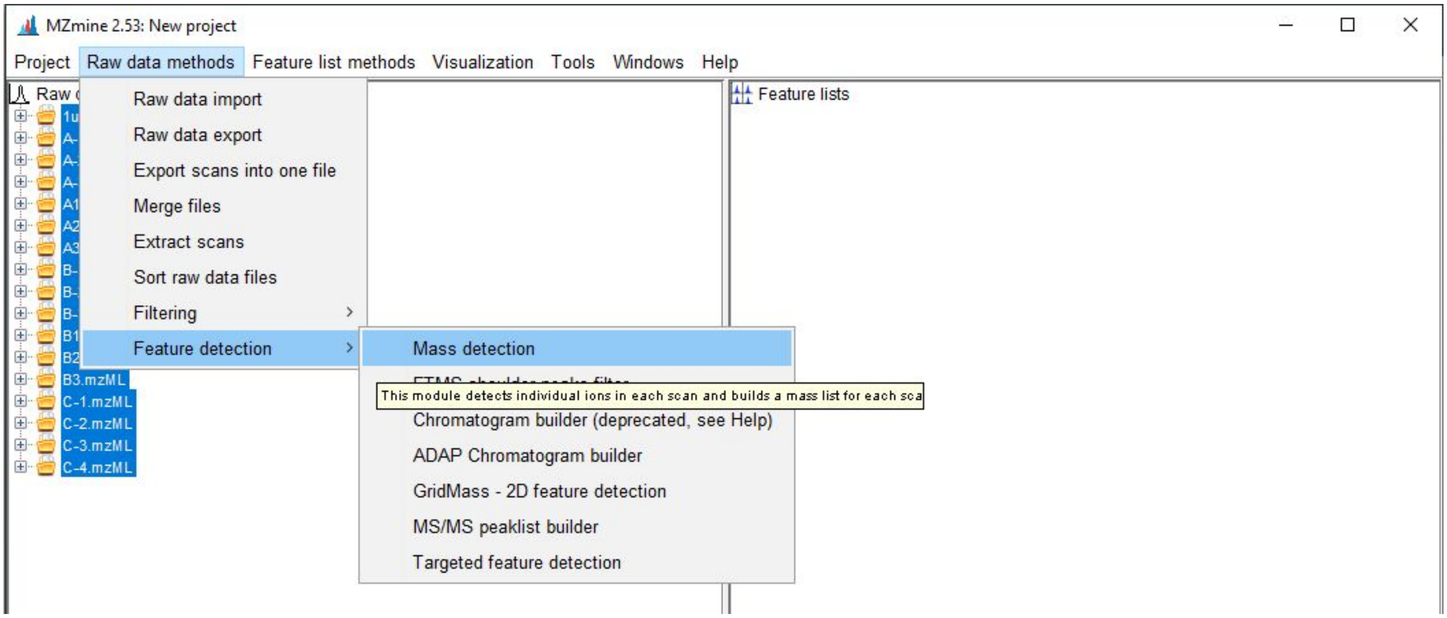
- Once, imported your raw data should show up in the left panel as individual folders that include the individual Mass-Spec files:



Jump back to table of contents

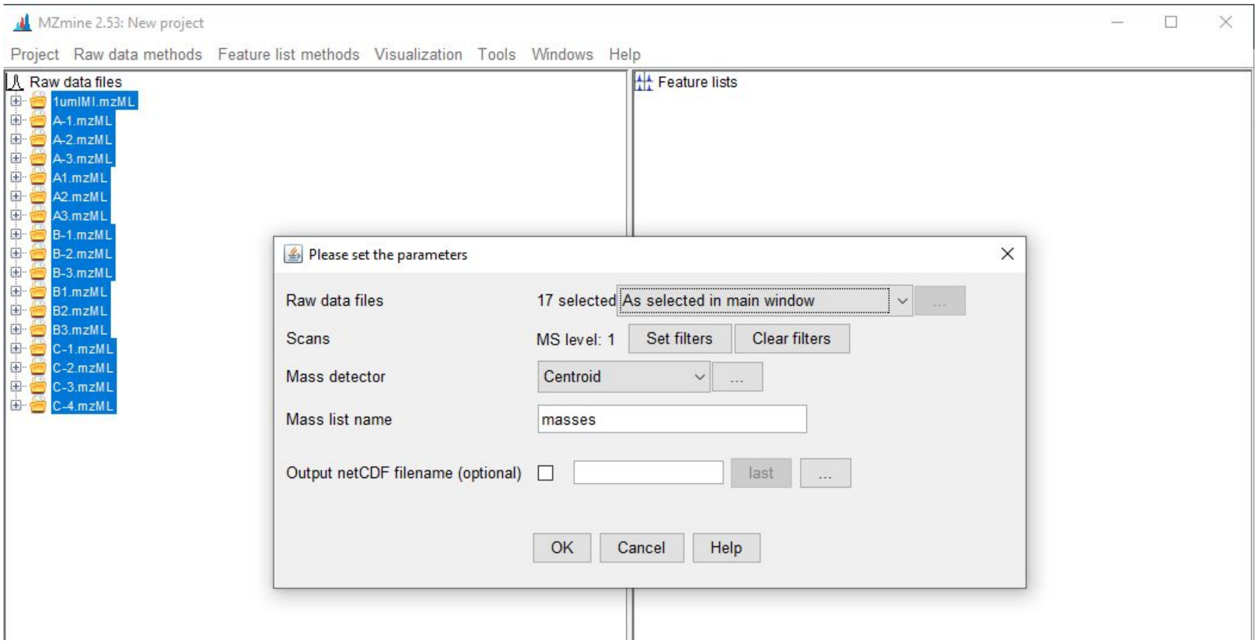
Mass Detection

- Select all of of the raw data files and click on the "Mass Detection" tool using the following path: "Raw data methods" → "Feature Detection" → "Mass Detection":



This module detects individual ions in each scan and builds a mass list for each sca

- The following pop-up box should appear:
    - Under Raw data files, select the files you want to analyze that were originally selected in the main window.
    - Mass Detector should be set to "Centroid"
    - Mass list name should be "masses"
        - This step filters the chromatograms based on a specific noise threshold and creates folders that are named after the original raw data files with "masses" appended to the end of the selected files



Please set the parameters

Raw data files: 17 selected — As selected in main window

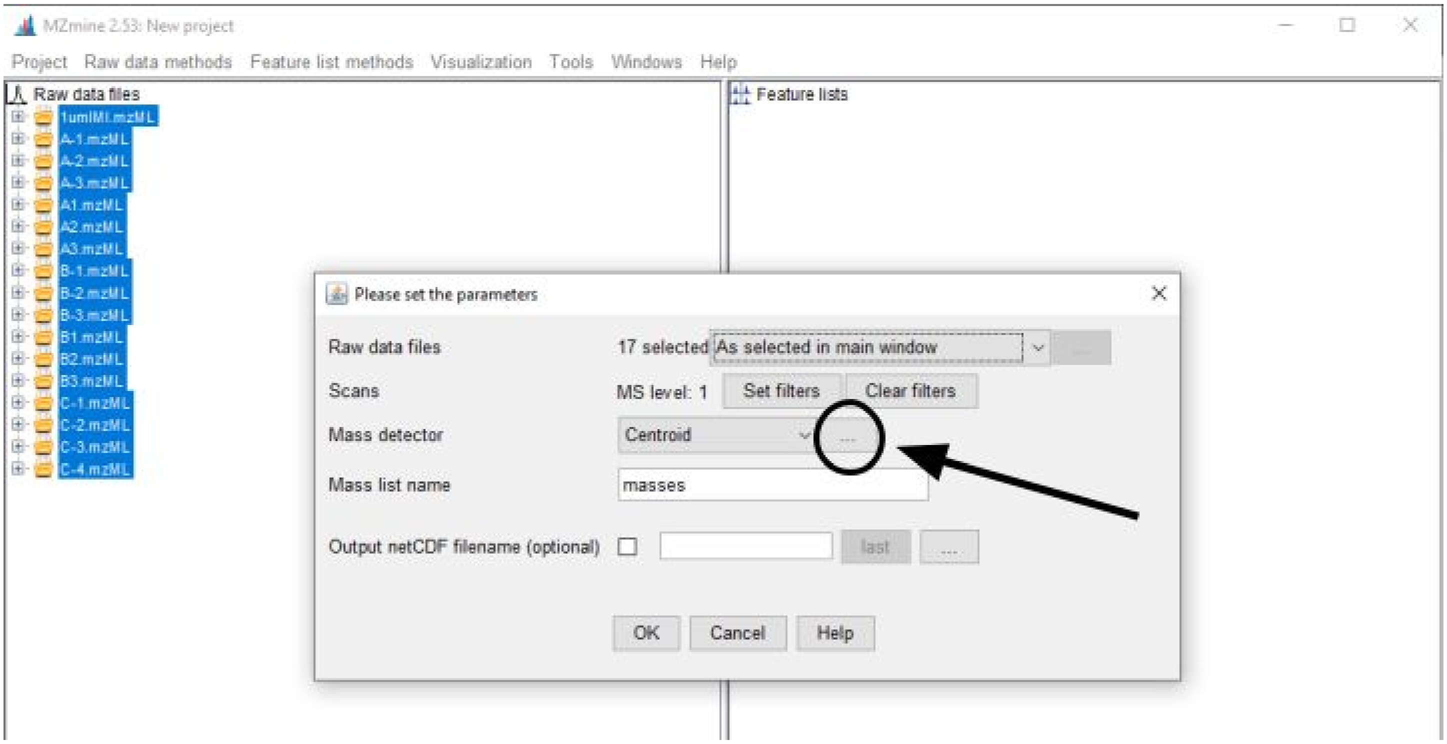
Scans: MS level: 1

Mass detector: Centroid

Mass list name: masses

Output netCDF filename (optional) □

- Set the MS1 noise levels by clicking on the three dots to the right of the dropbox menu of the Mass Detector field:



Raw data files: 17 selected As selected in main window

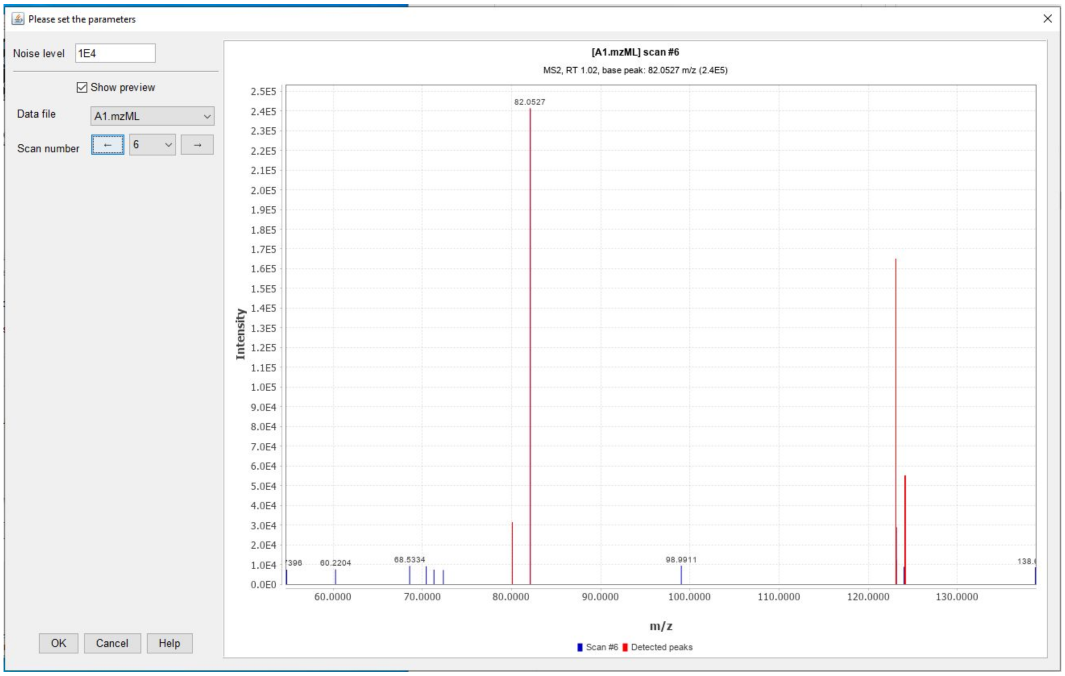
Scans: MS level: 1

Mass detector: Centroid

Mass list name: masses

Output netCDF filename (optional) □

- The following pop-up box should appear (note that the "Show preview" option is clicked):
  - Set the noise level appropriate for the run, in this case we set the intensity threshold to 1E4 (1 x 10^6)
  - In the preview box, you can see that the peaks that do not reach this threshold are not highlighted in red.



Noise level: 1E4

☑ Show preview

Data file: A1.mzML

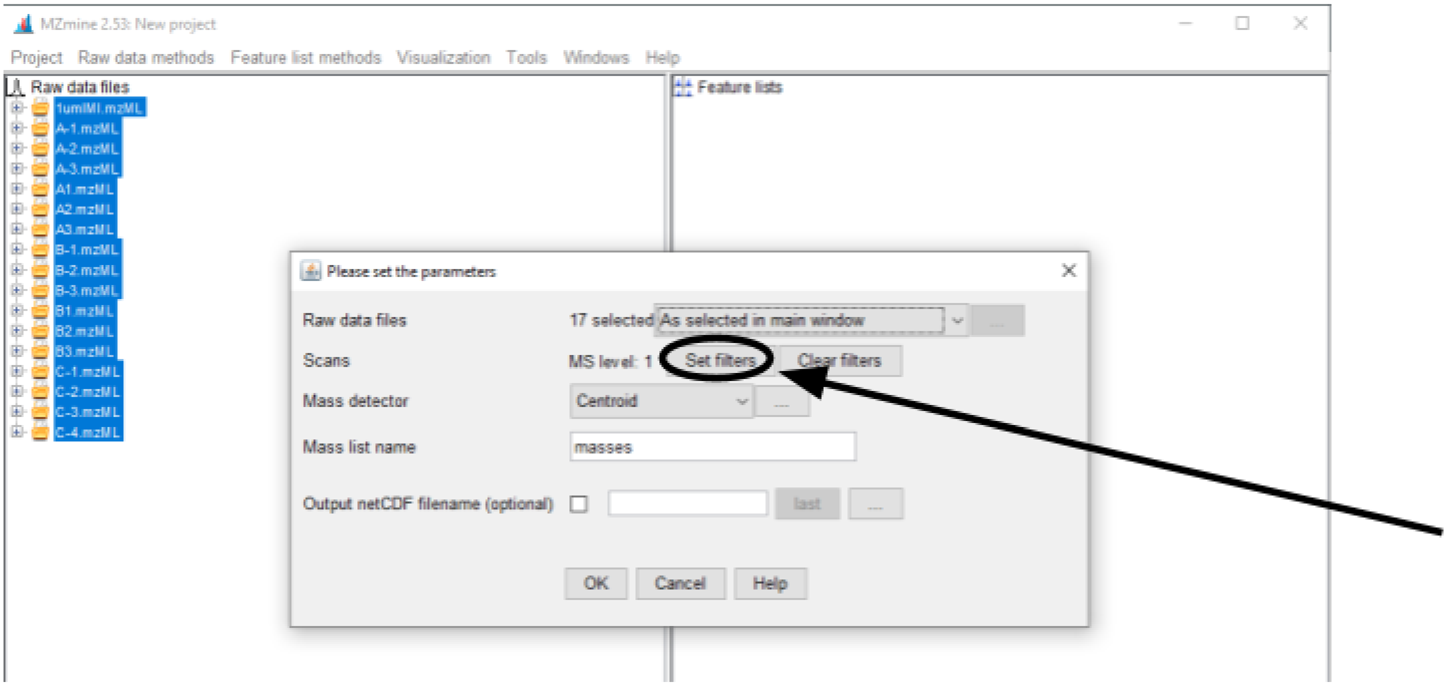
Scan number: 6

[A1.mzML] scan #6

MS2, RT 1.02, base peak: 82.0527 m/z (2.4E5)

Jump back to table of contents

- Repeat the last two steps, but now filtering the MS 2.
    - Change from setting noise levels from MS 1 to MS 2 by clicking the "Set Filter" box under the "Scans" field:



Raw data files: 17 selected As selected in main window

Scans: MS level: 1

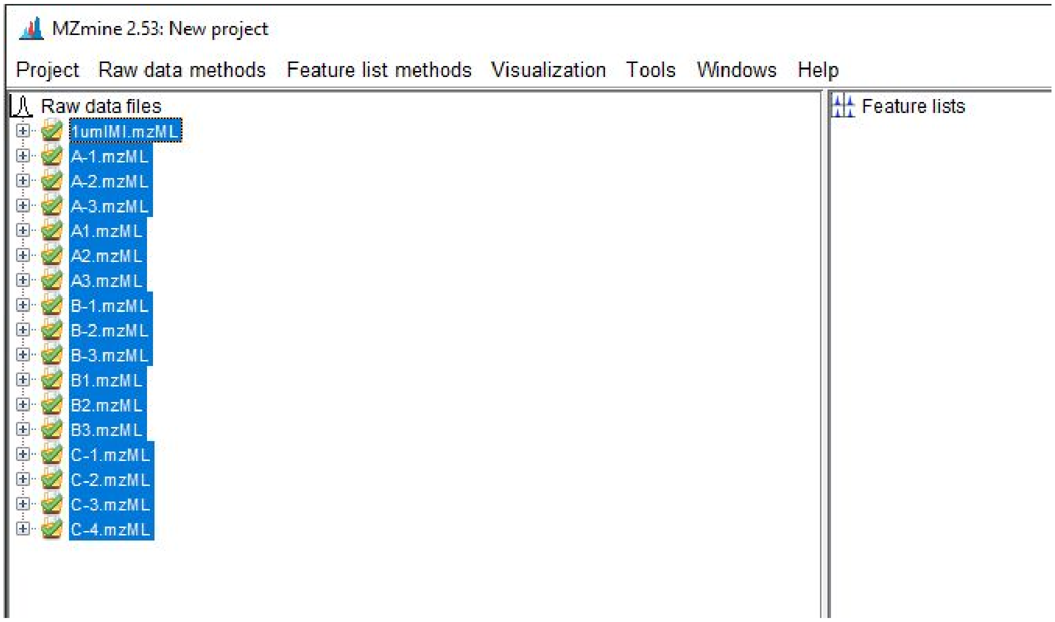
Mass detector: Centroid

Mass list name: masses

Output netCDF filename (optional) □

- The noise level should be set such that most of the noise is filtered out of your samples, but (in this example) it is not a quantitative step.
    - In general though, the MS2 scan level should be set less than the MS1 level.
    - For our files, we set noise levels to **1e4 and 1e3 for the MS1 and MS2** levels, respectively.

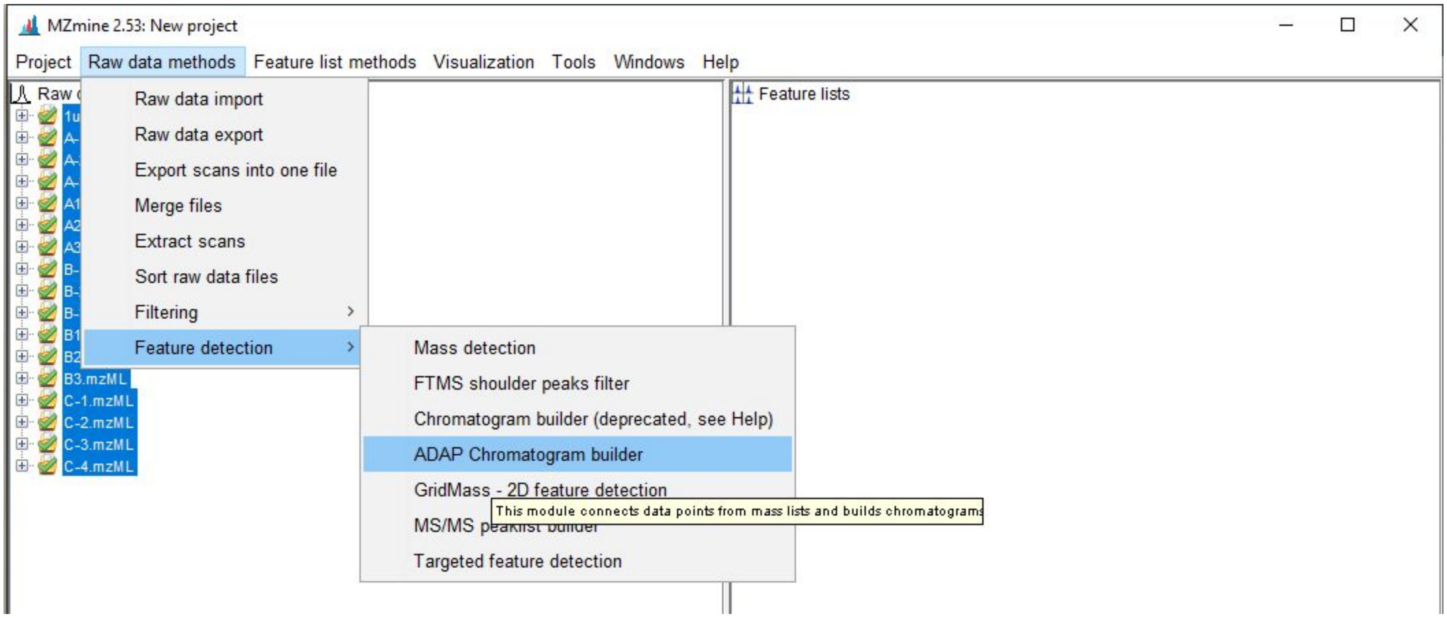
* Note how the file icons change once you apply the filters with a green check mark on each file icon.



Jump back to table of contents

ADAP Chromatogram Builder

- Next, we will build the chromatograms from our mass spec files:
    - Clicking on the ADAP Chromatogram builder from the following path: “Raw data methods” → “Feature detection” → “ADAP Chromatogram Builder”:

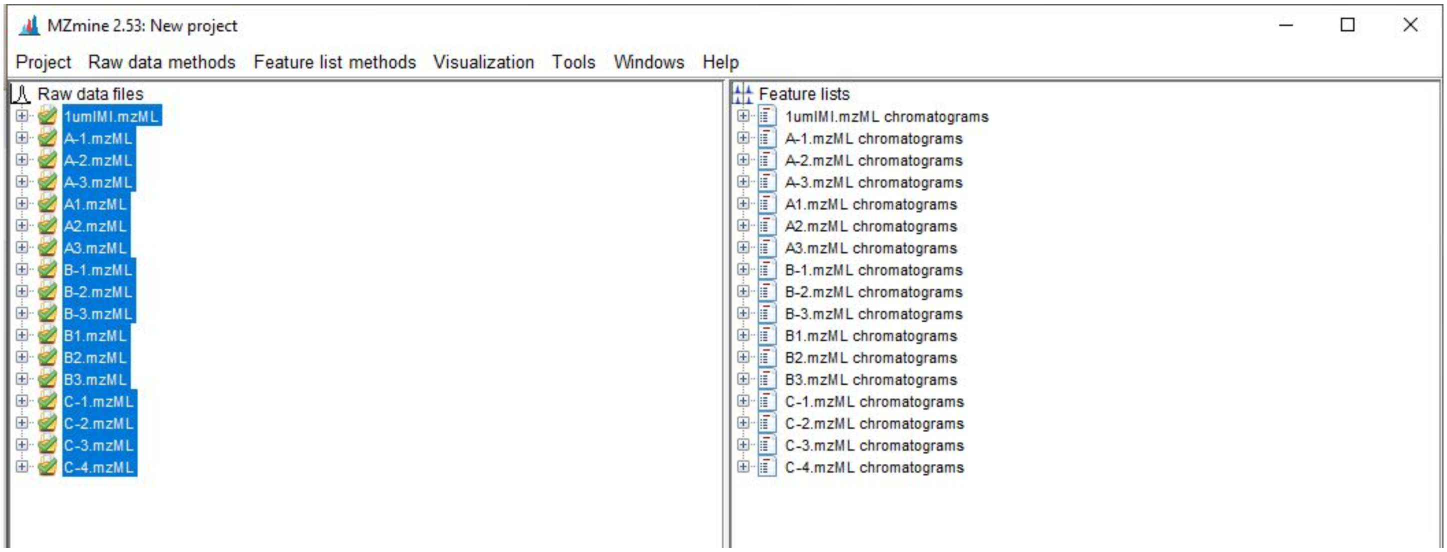


This module connects data points from mass lists and builds chromatogram

- For our experiment we set the following parameters to build our chromatograms:
    - Select the mass list by clicking the “Choose…” button to the right of the mass list option and picking the “masses” option in the dropdown menu.
- More details about these parameters can be found in this manual (http://mzmine.github.io/ADAP_user_manual.pdf), but the following explanations are excerpted briefly:
    - Min group size in # of Scans: 5
        - There must be this number of sequential scans having points above a set intensity threshold.
    - Group Intensity Threshold: 3e6
        - The threshold by which scans are compared against to count spectra as part of a group
    - Min highest intensity: 3e6
        - The group of spectra must have a peak of this intensity to be counted
    - m/z tolerance: 0.02 m/z or 10 ppm
        - Tolerance of the difference between data points in consecutive scans to be counted
    - Suffix: chromatograms
        - Appends the file name with “chromatograms”
- After entering these parameters, press ok to build your chromatograms.

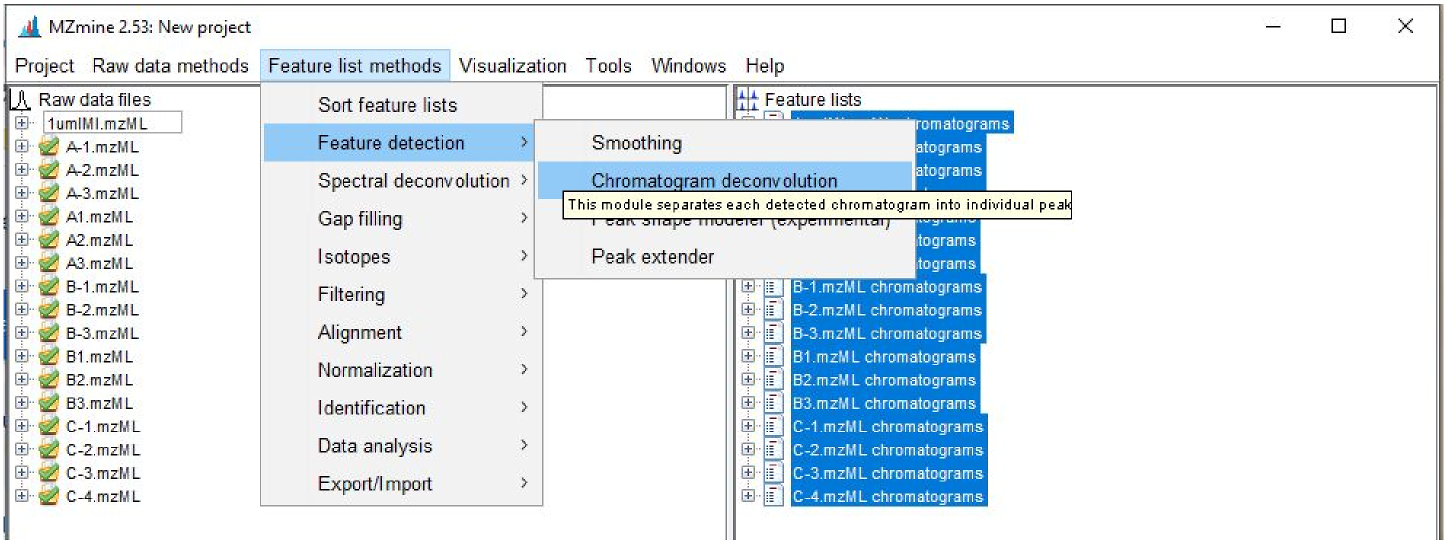
Jump back to table of contents

- The chromatograms will appear on the right in the “Feature lists” panel:



Chromatographic Deconvolution

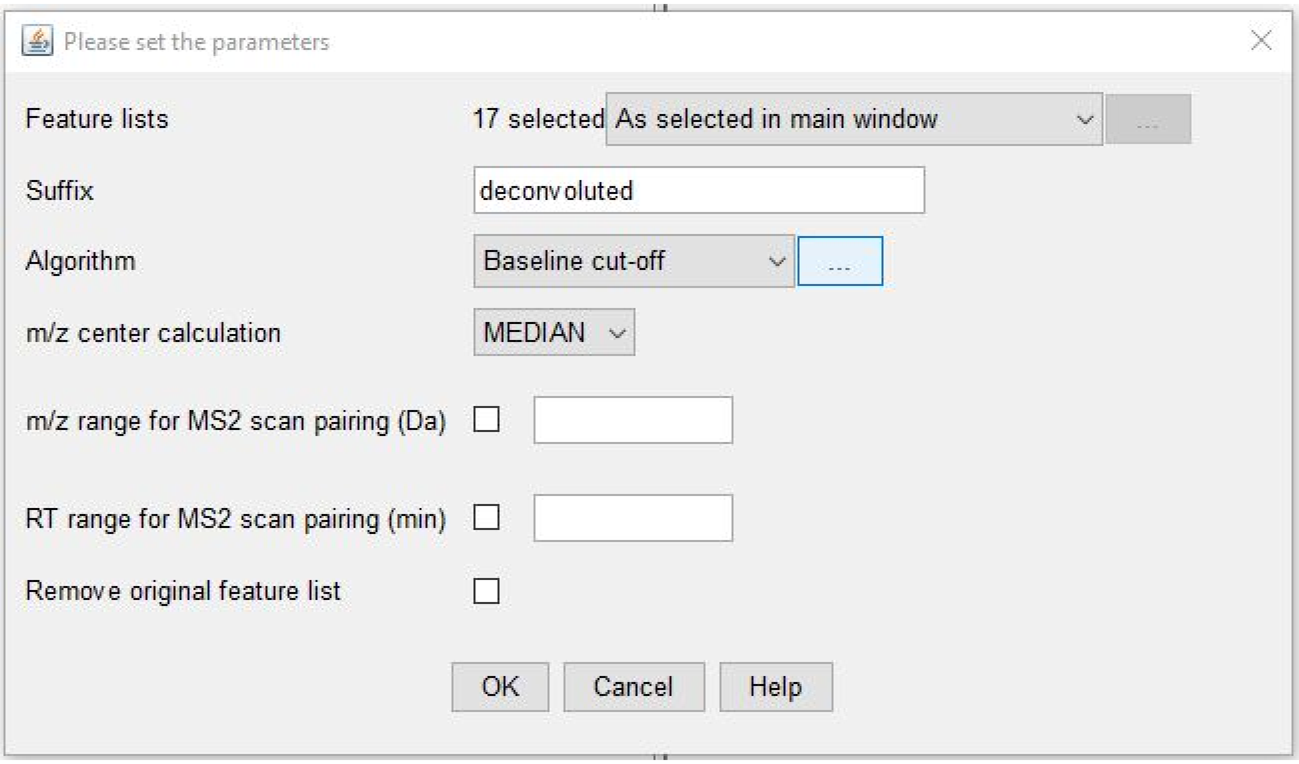
- Next, step is Chromatogram deconvolution
    - Chromatogram Deconvolution is the process of separating various mass spec data from the measured chromatographic information. Oftentimes in LC, there are samples that will co-elute in the same time period thereby “convoluting” the resulting mass spec data. Deconvolution reverses this process to separate the various masses that are co-eluted at the same time point.
        - Check out this page for a more in depth explanation of this process: https://www.spectralworks.com/products/analyzerpro-xd/learn-about-analyzerpro-xd/chromatographic-deconvolution/
- The Chromatogram Deconvolution Tool can be found under the following path: “Feature list methods” → “Feature Detection” → “Chromatogram deconvolution”
    - Be sure to select the chromatograms before opening this tool



This module separates each detected chromatogram into individual peak

Jump back to table of contents

- You should see the following pop-out



Please set the parameters

Feature lists: 17 selected As selected in main window

Suffix: deconvoluted

Algorithm: Baseline cut-off

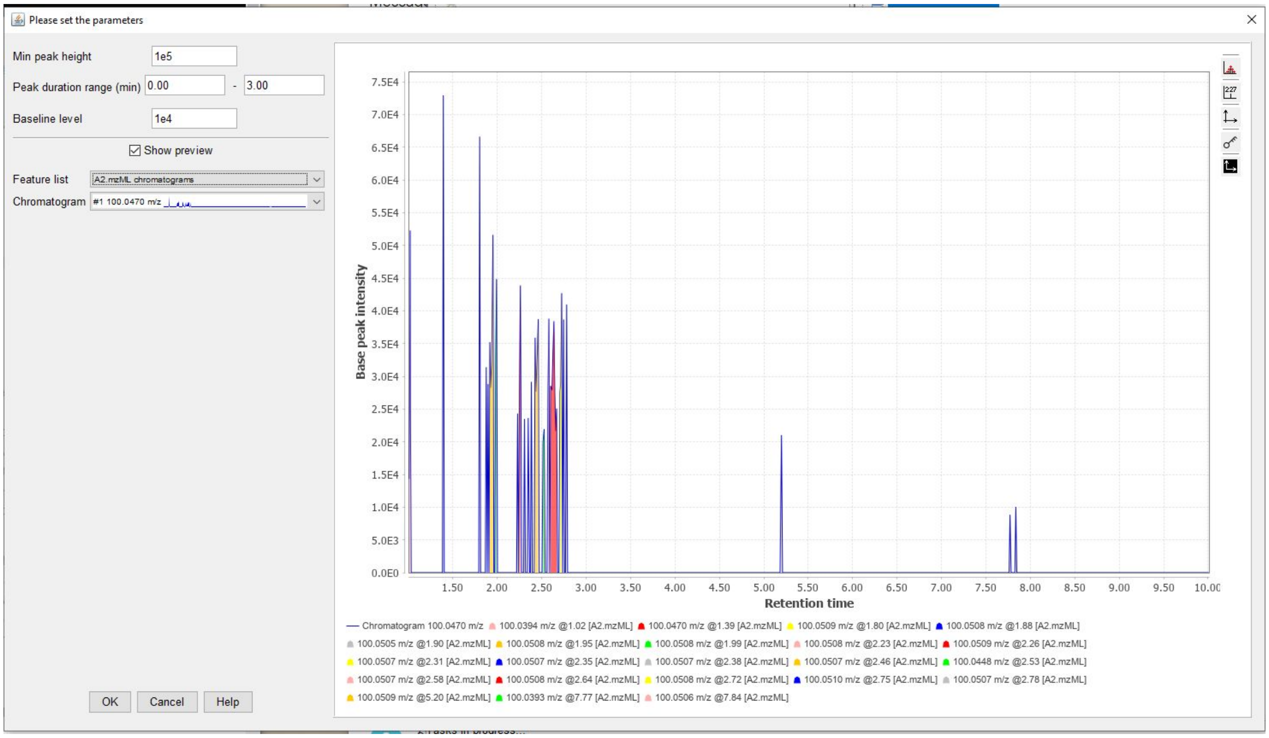
m/z center calculation: MEDIAN

m/z range for MS2 scan pairing (Da) □

RT range for MS2 scan pairing (min) □

Remove original feature list □

- We will use the following parameters:
  - Suffix: deconvoluted
  - Algorithm: Baseline cut-off (*Click the three dots next to the Baseline cut-off algorithm to modify the Algorithm parameters)
    - Min Peak height: 1e5
    - Peak Duration: 0 to 3 minutes
    - Baseline Level: 1e4
    - In the preview window, you can see individual peaks are color coded and you can adjust your parameters to ensure most of the features are identified and deconvoluted.
    - Individual peaks are identified by the m/z (mass to charge ratio) and retention time (in minutes)



Please set the parameters

Min peak height: 1e5

Peak duration range (min): 0.00 - 3.00

Baseline level: 1e4

☑ Show preview

Feature list: A2.mzML chromatograms

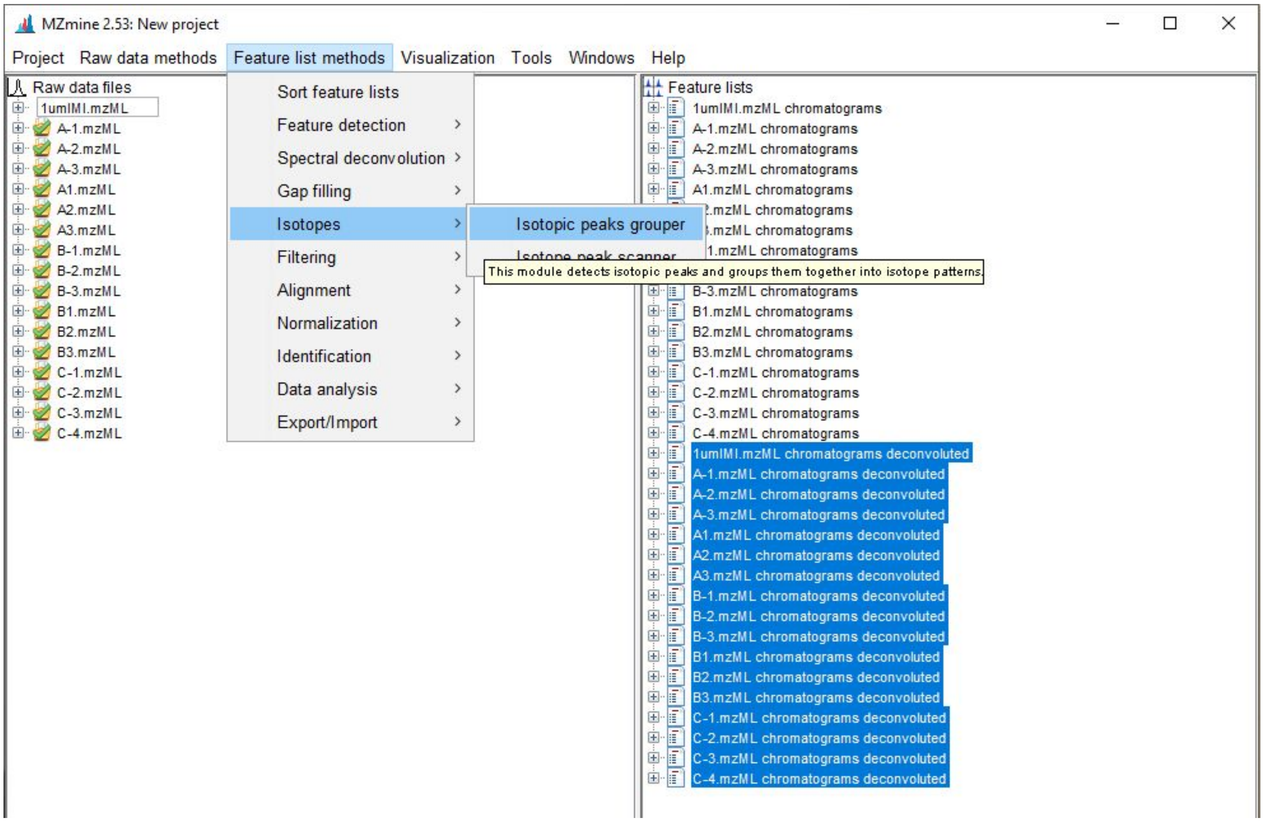
Chromatogram: #1 100.0470 m/z

OK
Cancel
Help
Base peak intensity
7.5E4
7.0E4
6.5E4
6.0E4
5.5E4
5.0E4
4.5E4
4.0E4
3.5E4
3.0E4
2.5E4
2.0E4
1.5E4
1.0E4
5.0E3
0.0E0
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
6.50
7.00
7.50
8.00
8.50
9.00
9.50
10.00
Retention time
Chromatogram 100.0470 m/z
100.0394 m/z @1.02 [A2.mzML]
100.0470 m/z @1.39 [A2.mzML]
100.0509 m/z @1.80 [A2.mzML]
100.0508 m/z @1.88 [A2.mzML]
100.0505 m/z @1.90 [A2.mzML]
100.0508 m/z @1.95 [A2.mzML]
100.0508 m/z @1.99 [A2.mzML]
100.0508 m/z @2.23 [A2.mzML]
100.0509 m/z @2.26 [A2.mzML]
100.0507 m/z @2.31 [A2.mzML]
100.0507 m/z @2.35 [A2.mzML]
100.0507 m/z @2.38 [A2.mzML]
100.0507 m/z @2.46 [A2.mzML]
100.0448 m/z @2.53 [A2.mzML]
100.0507 m/z @2.58 [A2.mzML]
100.0508 m/z @2.64 [A2.mzML]
100.0508 m/z @2.72 [A2.mzML]
100.0510 m/z @2.75 [A2.mzML]
100.0507 m/z @2.78 [A2.mzML]
100.0509 m/z @5.20 [A2.mzML]
100.0393 m/z @7.77 [A2.mzML]
100.0506 m/z @7.84 [A2.mzML]

  - m/z center calculation: MEDIAN
  - All others: leave blank

Isotope Grouping

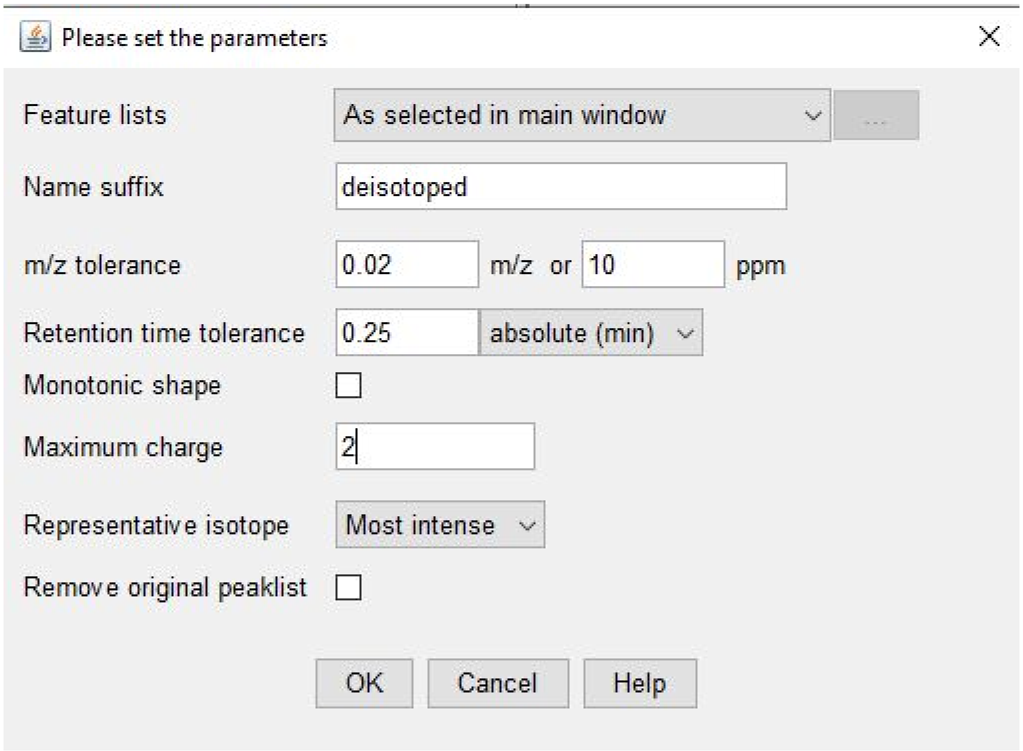
- Next, we will identify groups of isotopes with the "Isotopic peak grouper" tool which can be found under the following path: "Feature list methods" → "Isotopes" → "Isotopic peak grouper".



MZmine 2.53: New project

This module detects isotopic peaks and groups them together into isotope patterns.

- The following pop-up will appear, and we'll use the following parameters to group isotopes:
    - Name suffix: deisotoped
    - m/z tolerance: 0.02 or 10 ppm
    - RT tolerance: 0.25 (absolute (min))
    - max charge: 2
    - Representative isotope: Most intense



Please set the parameters

Feature lists: As selected in main window

Name suffix: deisotoped

m/z tolerance: 0.02 m/z or 10 ppm

Retention time tolerance: 0.25 absolute (min)

Monotonic shape □

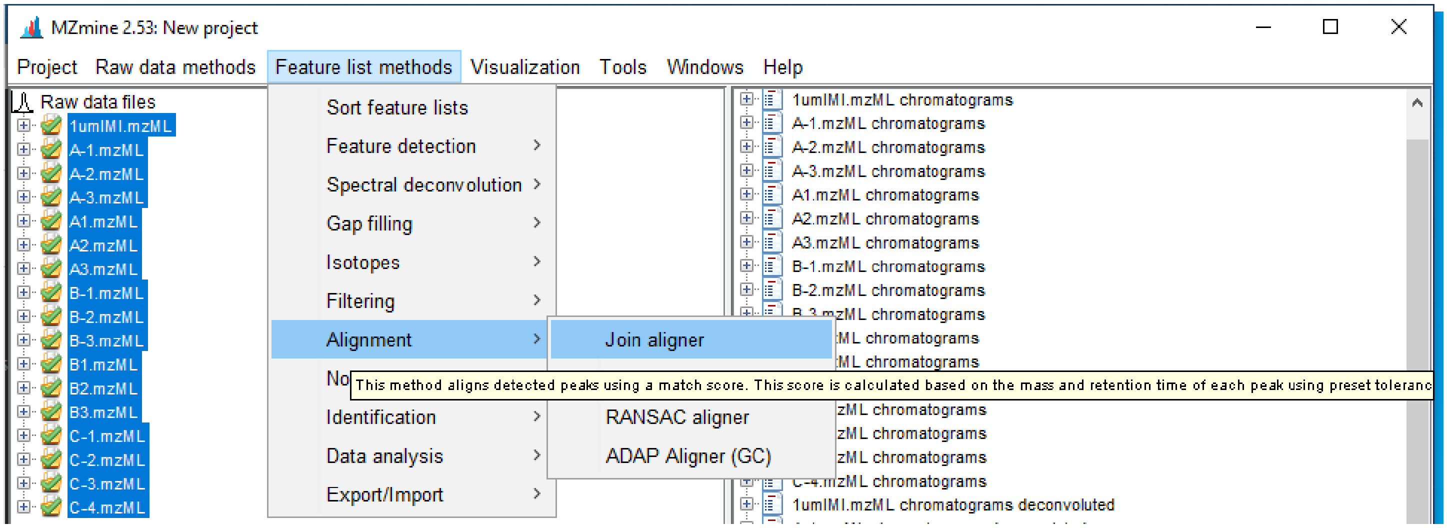
Maximum charge: 2

Representative isotope: Most intense

Remove original peaklist □

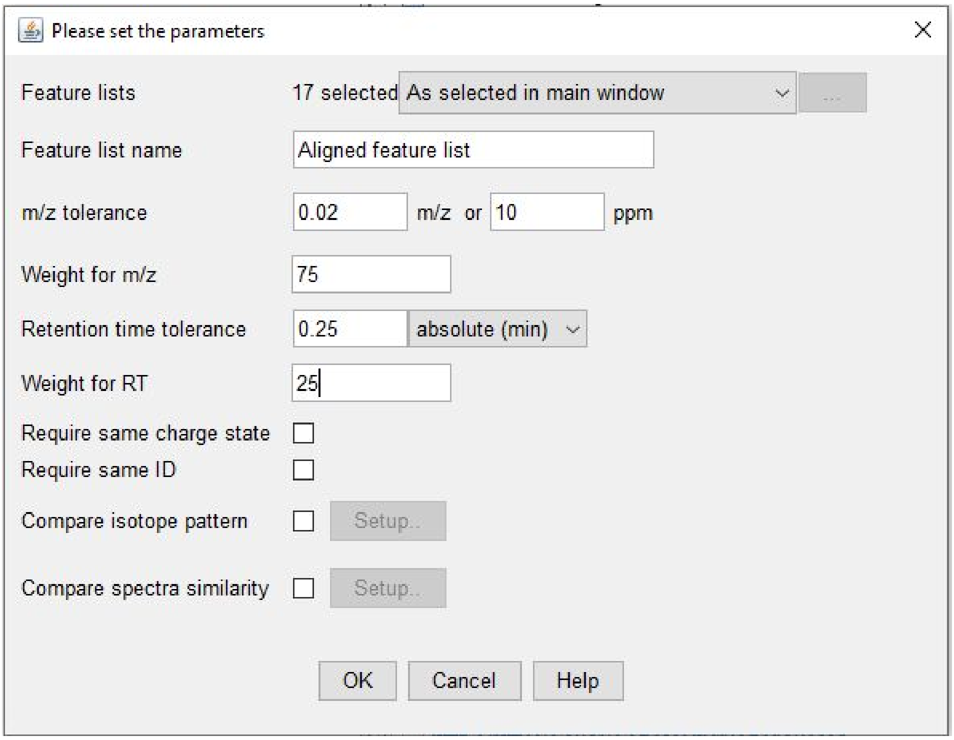
Join Aligner

- Next, we will align these groups of isotopes and other detected features according to their mass spec peaks and retention time for each sample.
  - Classical molecular networking primarily relies on the alignment of mass spec peaks to develop relationships between metabolites (as seen in the graphical overview).
- The alignment tools can be found with the following path: “feature list methods” → “Alignment”
  - For our samples we will use the simple Join aligner



This method aligns detected peaks using a match score. This score is calculated based on the mass and retention time of each peak using preset toleranc

- The following Pop-up should appear:



Feature lists: 17 selected As selected in main window

Feature list name: Aligned feature list

m/z tolerance: 0.02 m/z or 10 ppm

Weight for m/z: 75

Retention time tolerance: 0.25 absolute (min)

Weight for RT: 25

Require same charge state ☐

Require same ID ☐

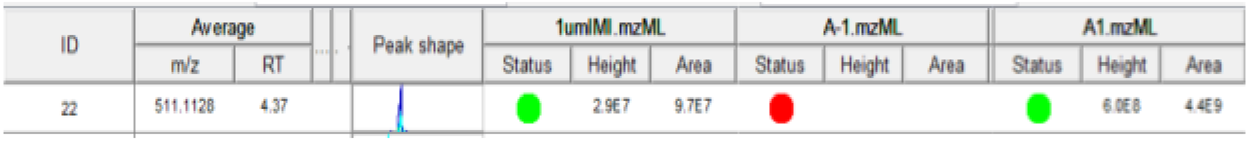
Compare isotope pattern ☐ Setup..

Compare spectra similarity ☐ Setup..

- And we’ll use the following parameters:
  - Feature list name: Aligned feature list
  - m/z tolerance: 0.02 m/z or 10 ppm
  - Weight for m/z: 75

Jump back to table of contents

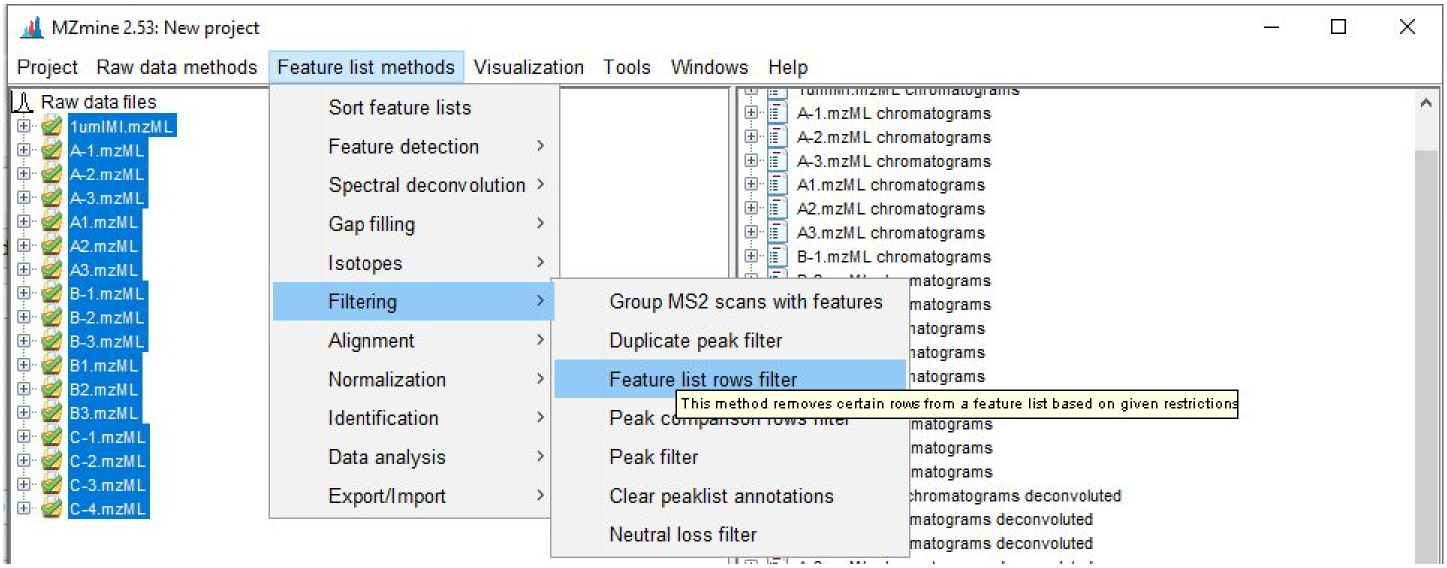
        - The is the weight of mass spec information on the scoring function fo the join aligner.
    - Retention time tolerance: 0.25 [absolute (min)]
        - This is the maximum allowed difference in LC-peak retention time between matched species.
    - Weight for RT: 25
        - Like above, this refers to the weight of RT on the scoring function.
- The product of this step result in a single list "Aligned feature list" which can be viewed in a separate window:
    - Your various samples will be displayed along the top, and the individual identified peaks will be arranged according to an "ID" number assigned during our Feature list method steps.
    - Some interesting quantities you should watch out for:
        - m/z ratio, retention time (RT), Height, and Area
    - For example in the following peak ID (to save space, only one sample from triplicates of the negative control and treated samples are shown below):
        - The features grouped in this node corresponding to a unique species have an average m/z of ~511and a retention time of ~4 minutes.
        - The height, or intensity of chromatographic peak, is 2.9E7 for our imidacloprid (IMI) standard and 6.0E8 for our *Pseudomonas* sample exposed to imidacloprid.
        - And the area of the peak under the chromatographic peak is similarly larger for our sample exposed to imidacloprid where our concentration of Imidacloprid is higher (1.76 mM vs 1 uM IMI)
        - Using this relationship, we designed future experiments to semi-quantitatively characterize the degradation of Imidacloprid.



ID	Average		Peak shape	1umIMI.mzML			A-1.mzML			A1.mzML		
	m/z	RT		Status	Height	Area	Status	Height	Area	Status	Height	Area
22	511.1128	4.37		🟢	2.9E7	9.7E7	🔴			🟢	6.0E8	4.4E9

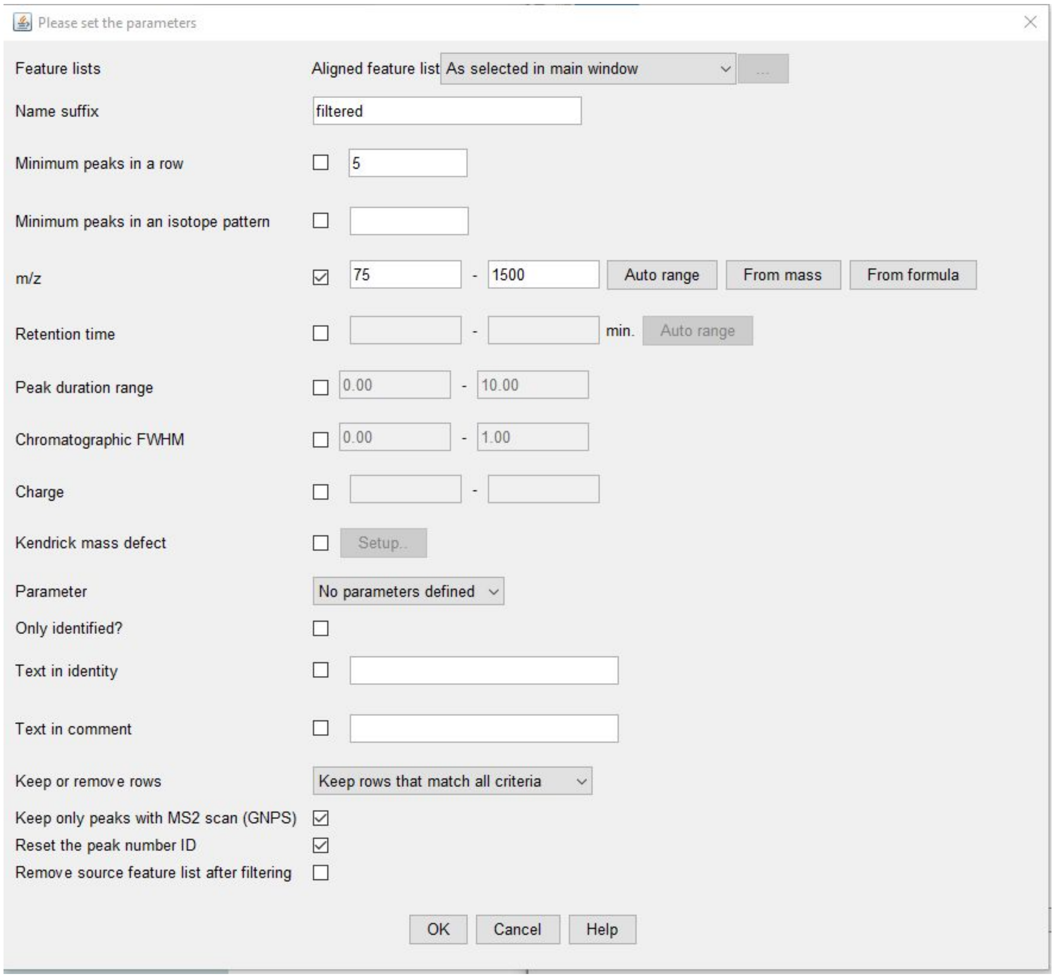
Peak Filtering

- The previous aligned feature list step identified over 22,000 nodes/species, of which many are not necessary for downstream analysis. This step allows users to selectively choose which species to analyze in later analysis steps. For our purposes, we will be filtering species based on the number of peaks for each feature, and the m/z of the species. For our project, we have relied on visualizing networks within the GNPS environment, and this step also allows us to introduce several processing steps to make it easier to upload data to GNPS.
- The peak filtering method for this example, "Feature list rows filter", can be found through the following path: "Feature list methods" → "Filtering" → "Feature list rows filter"



This method removes certain rows from a feature list based on given restrictions

- The following pop-up will appear, and we'll use the following parameters (leaving all other default settings unchanged):
    - min peak in row: 5
    - m/z range: 75-1500
    - keep peaks with MS2 for GNPS: TRUE
    - reset peak number ID: TRUE



Please set the parameters

Feature lists — Aligned feature list: As selected in main window

Name suffix: filtered

Minimum peaks in a row: □ 5

Minimum peaks in an isotope pattern: □

m/z: ☑ 75 - 1500 Auto range From mass From formula

Retention time: □ - min. Auto range

Peak duration range: □ 0.00 - 10.00

Chromatographic FWHM: □ 0.00 - 1.00

Charge: □ -

Kendrick mass defect: □ Setup..

Parameter: No parameters defined

Only identified?: □

Text in identity: □

Text in comment: □

Keep or remove rows: Keep rows that match all criteria

Keep only peaks with MS2 scan (GNPS): ☑

Reset the peak number ID: ☑

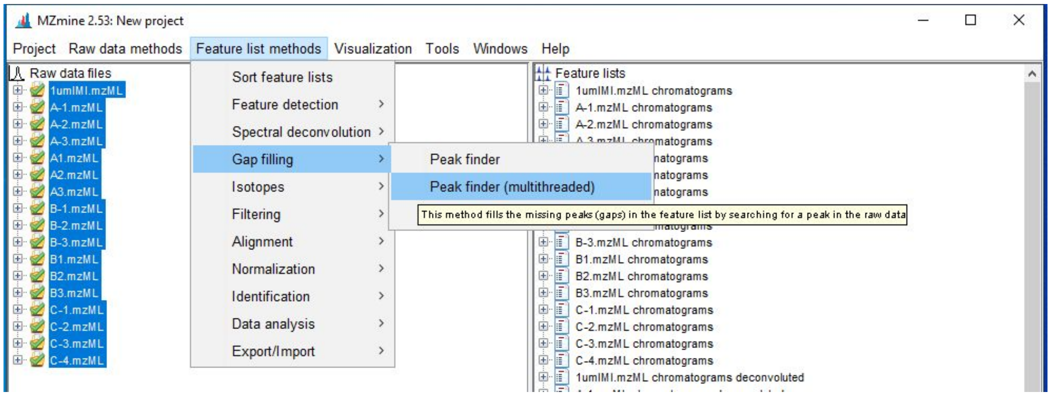
Remove source feature list after filtering: □

OK Cancel Help

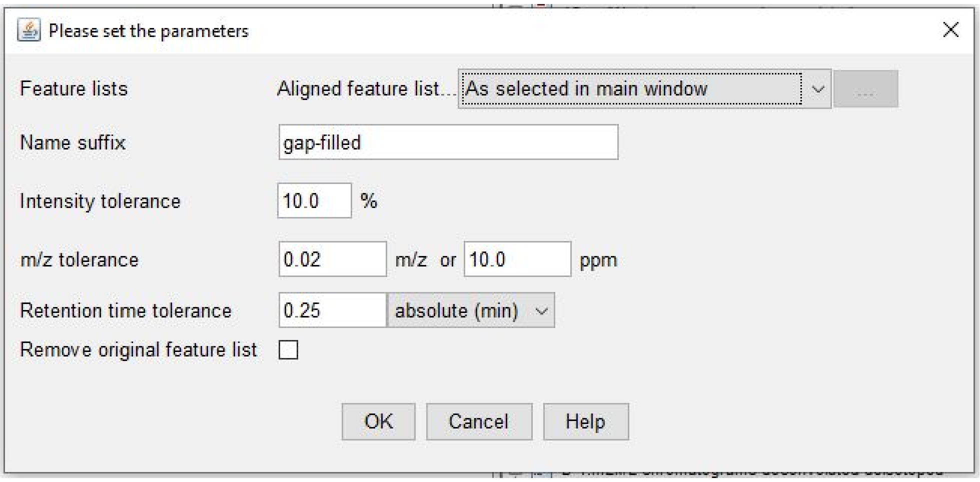
Jump back to table of contents

Gap-filling:

- The gap-filling algorithm (“Peak finder”) searches through the original raw data to find spectra that may have been mis-aligned during the alignment step or spectra that were filtered out due to low peak intensity (or a previous algorithm may have accidentally removed an important identifying peak).
- For our samples, we’ll use the following parameters, and the gap-filling tool can be found by the following path: “Feature list methods” → “Gap filling” → “Peak finder”
  - *The Peak finder tool and the multithreaded peak finder tool use the same parameters and are interchangeable, but the speed of this step can be sped up by allocated tasks to more CPU cores with the multithreaded tool.
  - Intensity tolerance: 10%
  - m/z tolerance: 0.02 or 10 ppm
  - RT tolerance: 0.25



This method fills the missing peaks (gaps) in the feature list by searching for a peak in the raw data



Please set the parameters

Feature lists	Aligned feature list... As selected in main window
Name suffix	gap-filled
Intensity tolerance	10.0 %
m/z tolerance	0.02 m/z or 10.0 ppm
Retention time tolerance	0.25 absolute (min)
Remove original feature list	□

- Visually inspecting the resulting list “Aligned feature list filtered gap-filled” shows which peaks are “estimated” with a yellow status indicator vs the typical green and red status indicator (depending on your gap-filling parameters) as shown in the original aligned feature list.

***Manual Validation**

- Unfortunately, this topic is beyond the scope of this introductory guide.
- In the simplest terms, you should manually inspect the peaks (as shown under “peak” in feature lists to ensure there are no major abnormalities in the peaks identified (e.g. a broad, undefined peak vs. a distinct peaks

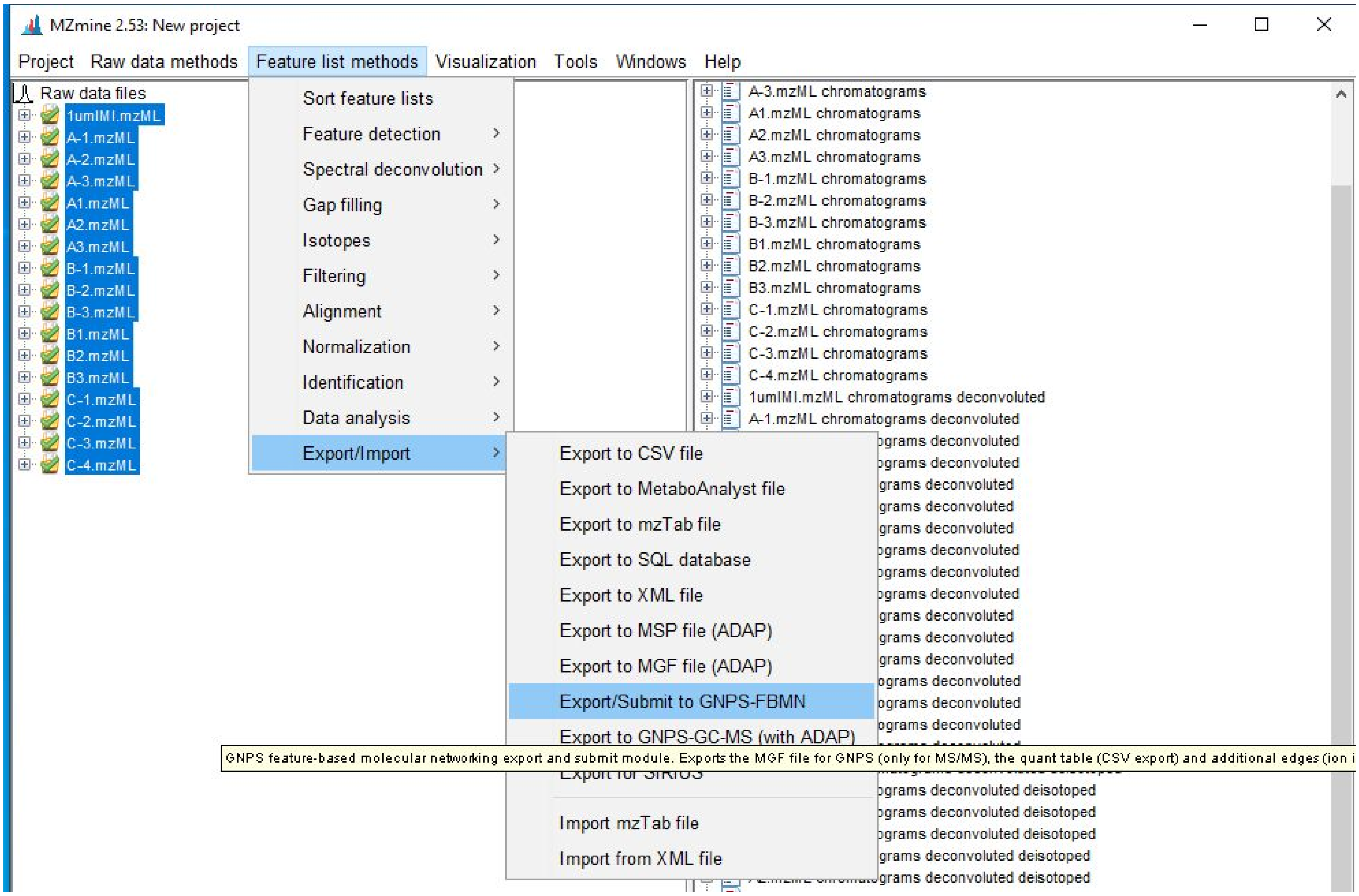
Export and Uploading Data to GNPS:

- The team behind the development of GNPS have great documentation on this part, but we briefly outline the steps here for documentation purposes for our team.
    - This video on YouTube walks through the process of uploading data to GNPS after processing data on MZmine: https://youtu.be/vFcGG7T_44E
    - This documentation page walks through another example project and walks through the process of FBMN workflow: https://ccms-ucsd.github.io/GNPSDocumentation/featurebasedmolecularnetworking-with-mzmine2/
        - The process outlined in this page is largely similar, with the exception of the addition of Gap filling and MS row filtering found in our workflow.

Jump back to table of contents

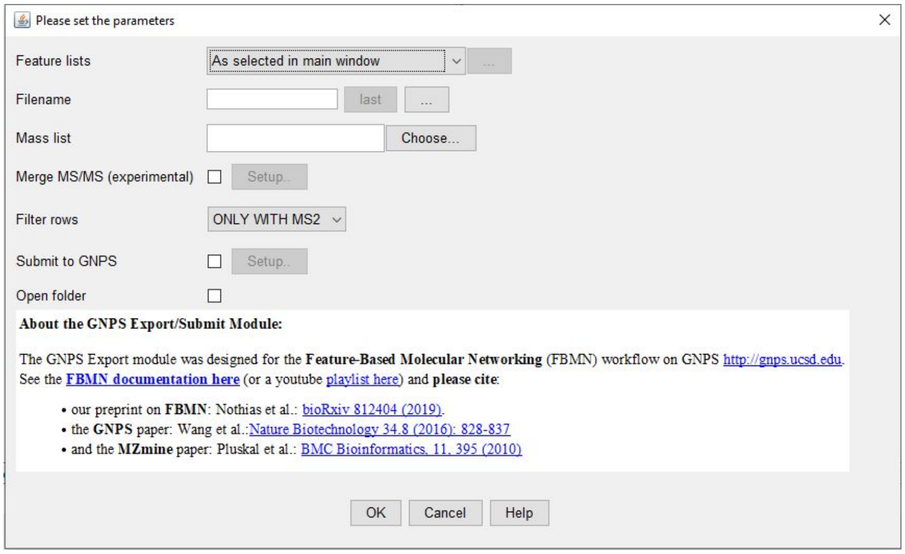
Export:

- MZmine has several convenient exporting tools that will output data in specific file formats or even packages of data to be analyzed in external tools like GNPS and SIRIUS.
- For our purposes we'll use the GNPS option which can be found through the following path: "Feature list methods" → "Export/Import" → "Export/Submit to GNPS-FBMN"



GNPS feature-based molecular networking export and submit module. Exports the MGF file for GNPS (only for MS/MS), the quant table (CSV export) and additional edges (ion i

- The following pop-up should appear; and for our example, we'll use the following parameters:
    - Feature lists: as selected
    - Mass list: masses (*click "Choose..." button like in previous steps and select masses")
    - Filter rows: ALL (you can export a smaller file and just export the MS2 spectra, but including MS1 may improve your odds of features being identified in GNPS)



Feature lists: As selected in main window

Filename:

Mass list:

Merge MS/MS (experimental) □

Filter rows: ONLY WITH MS2

Submit to GNPS □

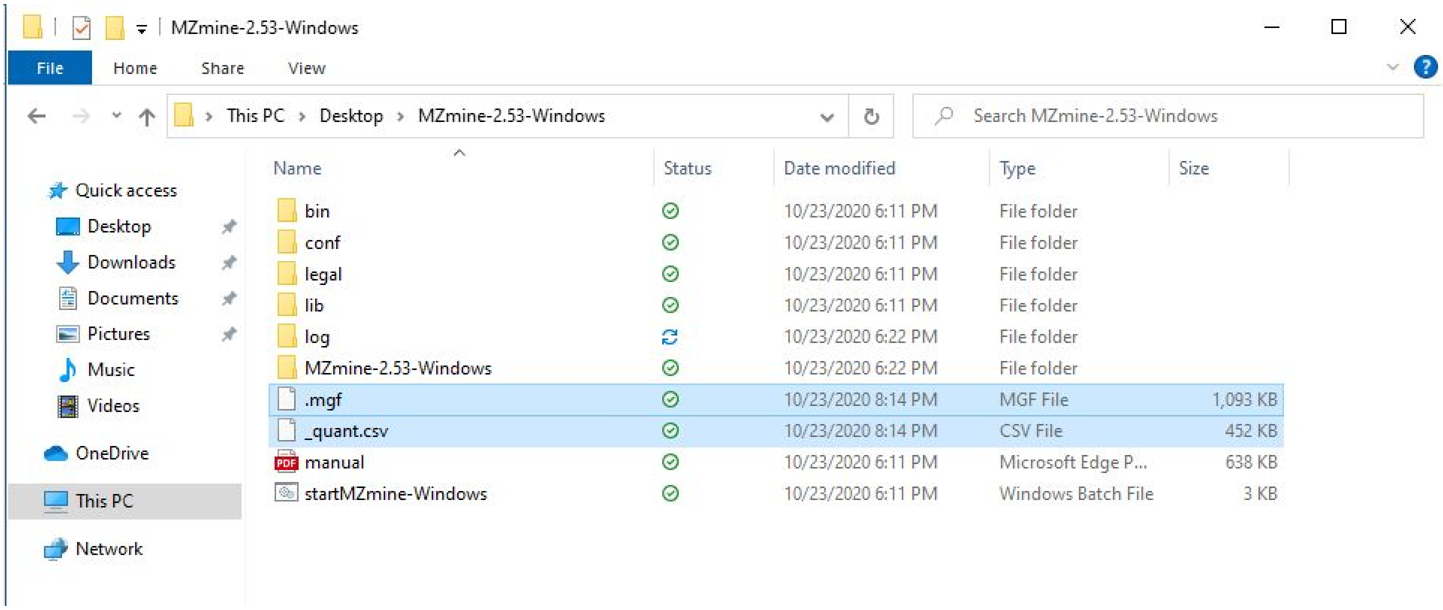
Open folder □

About the GNPS Export/Submit Module:

The GNPS Export module was designed for the **Feature-Based Molecular Networking** (FBMN) workflow on GNPS http://gnps.ucsd.edu. See the **FBMN documentation here** (or a youtube playlist here) and **please cite**:

- our preprint on **FBMN**: Nothias et al.: bioRxiv 812404 (2019).
- the **GNPS** paper: Wang et al.:Nature Biotechnology 34.8 (2016): 828-837
- and the **MZmine** paper: Pluskal et al.: BMC Bioinformatics, 11, 395 (2010)

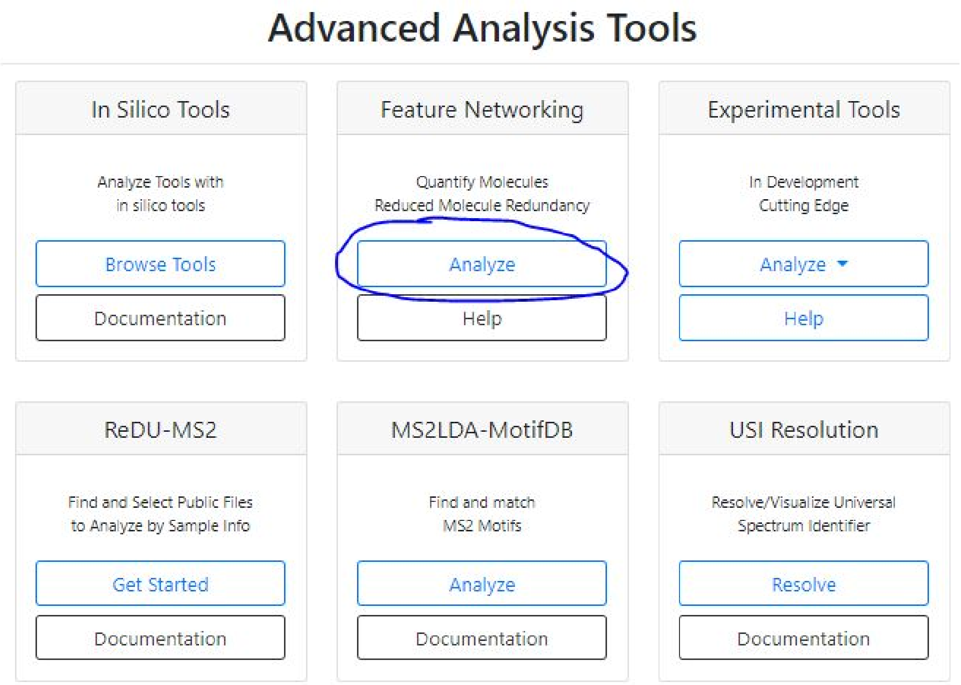
- The two output files required for GNPS-FBMN will appear in the folder in which you un-zipped MZmine:



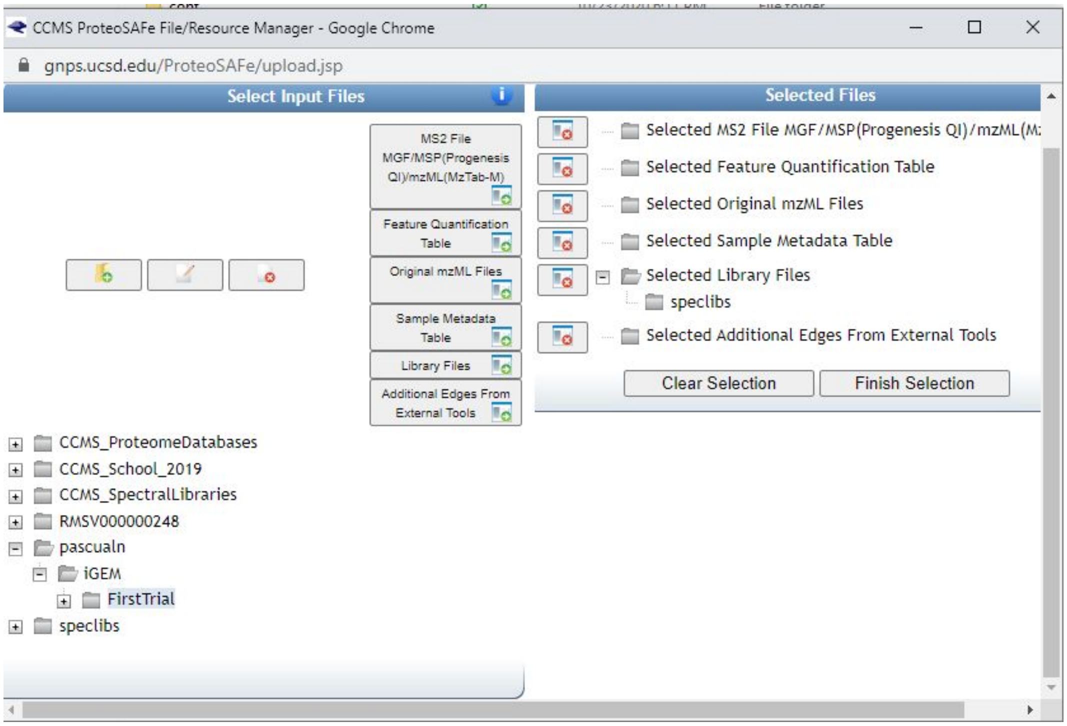
Name	Status	Date modified	Type	Size
bin	✓	10/23/2020 6:11 PM	File folder	
conf	✓	10/23/2020 6:11 PM	File folder	
legal	✓	10/23/2020 6:11 PM	File folder	
lib	✓	10/23/2020 6:11 PM	File folder	
log	⟳	10/23/2020 6:22 PM	File folder	
MZmine-2.53-Windows	✓	10/23/2020 6:22 PM	File folder	
.mgf	✓	10/23/2020 8:14 PM	MGF File	1,093 KB
_quant.csv	✓	10/23/2020 8:14 PM	CSV File	452 KB
manual	✓	10/23/2020 6:11 PM	Microsoft Edge P...	638 KB
startMZmine-Windows	✓	10/23/2020 6:11 PM	Windows Batch File	3 KB

Uploading files to GNPS:

- Using the file transfer client of choice, upload the .mgf data and _quant.csv quantification table to your personal storage account
- On GNPS's main home page (after logging into your account in your web-browser), fin "Feature Networking" tool under "Advanced Analysis Tools":

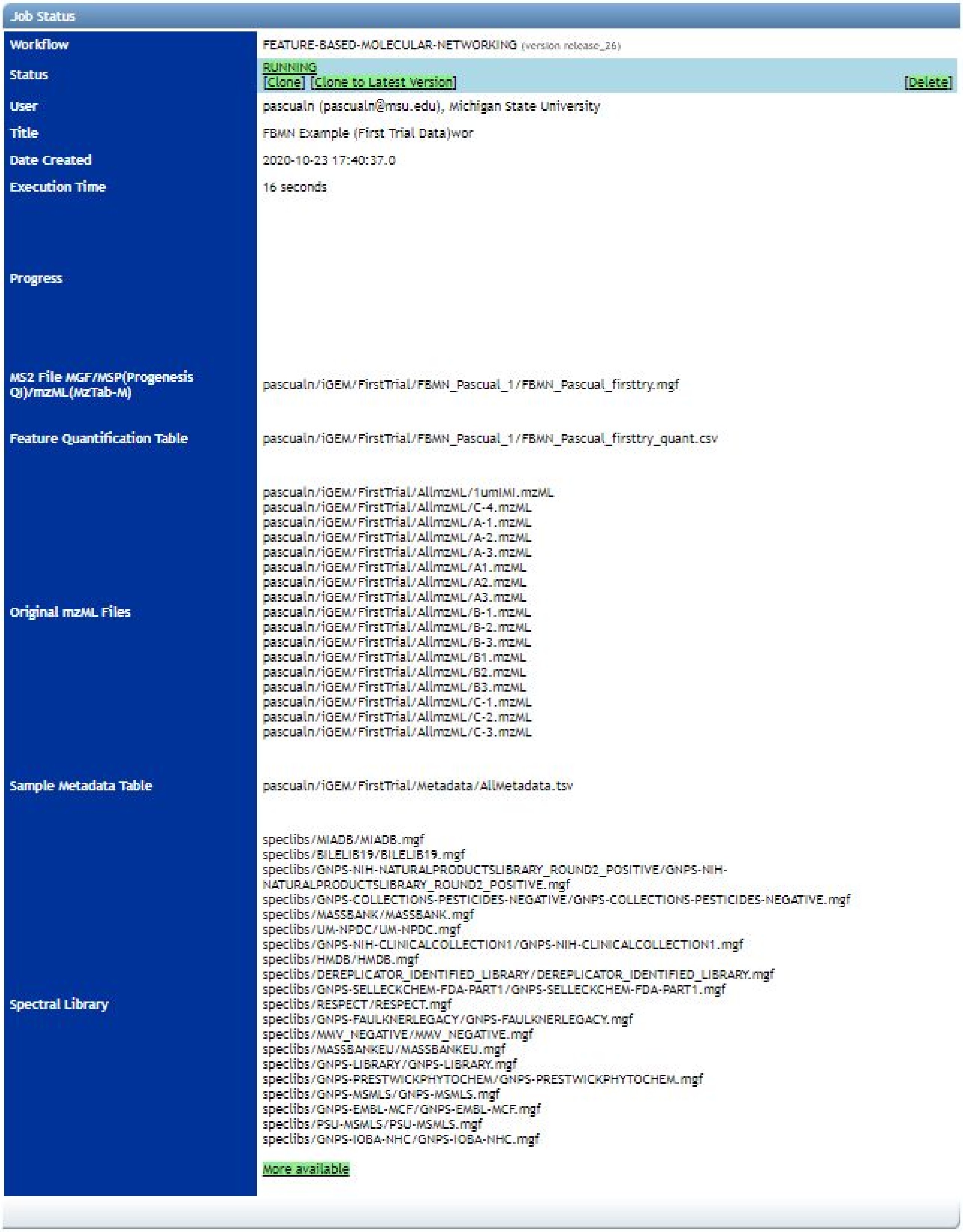

Advanced Analysis Tools
In Silico Tools
Analyze Tools with in silico tools
Browse Tools
Documentation
Feature Networking
Quantify Molecules Reduced Molecule Redundancy
Analyze
Help
Experimental Tools
In Development Cutting Edge
Analyze
Help
ReDU-MS2
Find and Select Public Files to Analyze by Sample Info
Get Started
Documentation
MS2LDA-MotifDB
Find and match MS2 Motifs
Analyze
Documentation
USI Resolution
Resolve/Visualize Universal Spectrum Identifier
Resolve
Documentation

- Give your job a title (under "Workflow Selection")
- Under "File Selection", when you click "Select Input Files", you can select input files front he files you uploaded in the first step of this section with the following pop-up browser


CCMS ProteoSAFe File/Resource Manager - Google Chrome
gnps.ucsd.edu/ProteoSAFe/upload.jsp
Select Input Files
MS2 File MGF/MSP(Progenesis QI)/mzML(MzTab-M)
Feature Quantification Table
Original mzML Files
Sample Metadata Table
Library Files
Additional Edges From External Tools
CCMS_ProteomeDatabases
CCMS_School_2019
CCMS_SpectralLibraries
RMSV000000248
pascualn
iGEM
FirstTrial
speclibs
Selected Files
Selected MS2 File MGF/MSP(Progenesis QI)/mzML(M:
Selected Feature Quantification Table
Selected Original mzML Files
Selected Sample Metadata Table
Selected Library Files
speclibs
Selected Additional Edges From External Tools
Clear Selection
Finish Selection

- After selecting your data to run in the job, you can customize a lot of other parameters, but for this example, we will leave the remainder default values.

- After pressing, submit, you should see the following status page:



Job Status	
Workflow	FEATURE-BASED-MOLECULAR-NETWORKING (version release_26)
Status	RUNNING [Clone] [Clone to Latest Version] [Delete]
User	pascualn (pascualn@msu.edu), Michigan State University
Title	FBMN Example (First Trial Data)wor
Date Created	2020-10-23 17:40:37.0
Execution Time	16 seconds
Progress	
MS2 File MGF/MSP(Progenesis QI)/mzML(MzTab-M)	pascualn/iGEM/FirstTrial/FBMN_Pascual_1/FBMN_Pascual_firsttry.mgf
Feature Quantification Table	pascualn/iGEM/FirstTrial/FBMN_Pascual_1/FBMN_Pascual_firsttry_quant.csv
Original mzML Files	pascualn/iGEM/FirstTrial/AllmzML/1umIMI.mzML pascualn/iGEM/FirstTrial/AllmzML/C-4.mzML pascualn/iGEM/FirstTrial/AllmzML/A-1.mzML pascualn/iGEM/FirstTrial/AllmzML/A-2.mzML pascualn/iGEM/FirstTrial/AllmzML/A-3.mzML pascualn/iGEM/FirstTrial/AllmzML/A1.mzML pascualn/iGEM/FirstTrial/AllmzML/A2.mzML pascualn/iGEM/FirstTrial/AllmzML/A3.mzML pascualn/iGEM/FirstTrial/AllmzML/B-1.mzML pascualn/iGEM/FirstTrial/AllmzML/B-2.mzML pascualn/iGEM/FirstTrial/AllmzML/B-3.mzML pascualn/iGEM/FirstTrial/AllmzML/B1.mzML pascualn/iGEM/FirstTrial/AllmzML/B2.mzML pascualn/iGEM/FirstTrial/AllmzML/B3.mzML pascualn/iGEM/FirstTrial/AllmzML/C-1.mzML pascualn/iGEM/FirstTrial/AllmzML/C-2.mzML pascualn/iGEM/FirstTrial/AllmzML/C-3.mzML
Sample Metadata Table	pascualn/iGEM/FirstTrial/Metadata/AllMetadata.tsv
Spectral Library	speclibs/MIADB/MIADB.mgf speclibs/BILELIB19/BILELIB19.mgf speclibs/GNPS-NIH-NATURALPRODUCTSLIBRARY_ROUND2_POSITIVE/GNPS-NIH-NATURALPRODUCTSLIBRARY_ROUND2_POSITIVE.mgf speclibs/GNPS-COLLECTIONS-PESTICIDES-NEGATIVE/GNPS-COLLECTIONS-PESTICIDES-NEGATIVE.mgf speclibs/MASSBANK/MASSBANK.mgf speclibs/UM-NPDC/UM-NPDC.mgf speclibs/GNPS-NIH-CLINICALCOLLECTION1/GNPS-NIH-CLINICALCOLLECTION1.mgf speclibs/HMDB/HMDB.mgf speclibs/DEREPLICATOR_IDENTIFIED_LIBRARY/DEREPLICATOR_IDENTIFIED_LIBRARY.mgf speclibs/GNPS-SELLECKCHEM-FDA-PART1/GNPS-SELLECKCHEM-FDA-PART1.mgf speclibs/RESPECT/RESPECT.mgf speclibs/GNPS-FAULKNERLEGACY/GNPS-FAULKNERLEGACY.mgf speclibs/MMV_NEGATIVE/MMV_NEGATIVE.mgf speclibs/MASSBANKEU/MASSBANKEU.mgf speclibs/GNPS-LIBRARY/GNPS-LIBRARY.mgf speclibs/GNPS-PRESTWICKPHYTOCHEM/GNPS-PRESTWICKPHYTOCHEM.mgf speclibs/GNPS-MSMLS/GNPS-MSMLS.mgf speclibs/GNPS-EMBL-MCF/GNPS-EMBL-MCF.mgf speclibs/PSU-MSMLS/PSU-MSMLS.mgf speclibs/GNPS-IOBA-NHC/GNPS-IOBA-NHC.mgf More available

- And once complete, you can browse through the various GNPS outputs through this link:
    - https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=85753400a82046728290c4190b850bed

References:

Katajamaa, M., Miettinen, J. & Oresic, M. MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. Bioinformatics 22, 634–636 (2006).

Myers OD, Sumner SJ, Li S, Barnes S, Du X: One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms for Constructing Extracted Ion Chromatograms and Detecting Chromatographic Peaks. Anal Chem 2017, DOI: 10.1021/acs.analchem.7b00947

Nothias, L., Petras, D., Schmid, R. et al. Feature-based molecular networking in the GNPS analysis environment. Nat Methods 17, 905–908 (2020). https://doi.org/10.1038/s41592-020-0933-6

Pluskal, T., Castillo, S., Villar-Briones, A. & Oresic, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics 11, 395 (2010).

Ty N.F. Roach, Jenna Dilworth, H. Christian Martin, A. Daniel Jones, Robert Quinn, Crawford Drury. bioRxiv 2020.05.10.087072; doi: https://doi.org/10.1101/2020.05.10.087072

Wang, M. et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat. Biotechnol. 34, 828–837 (2016).

Acknowledgements:

We'd like to thank the time and expertise offered by Dr. Rob Quinn and members of his lab (specifically Dr. Christian Martin, Dr. Kerri Neugebauer, and Kevin Santiago) to analyze our samples in their Thermo QExactiive LC-MS/MS Mass Spectrometer. Dr. Rob Quinn and Dr. Christian Martin were particularly helpful in offering training on the overall concepts of metabolomics networking and troubleshooting FBMN parameters for our project.

Furthermore, we'd like to thank Dr. Michaela TerAvest and members of her lab (specifically Shaylynn Miller, Nicholas Telft, and Kati Ford) for their time and expertise in designing and executing the imidacloprid biotransformation experiments from which the data in this intro guide was collected.

Jump back to table of contents