



Knowledge that will change your world

The application of Mzmine 2 to viewing metabolomics raw data

Stephen Barnes, PhD

University of Alabama at Birmingham

sbarnes@uab.edu

Download Mzmine 2.23

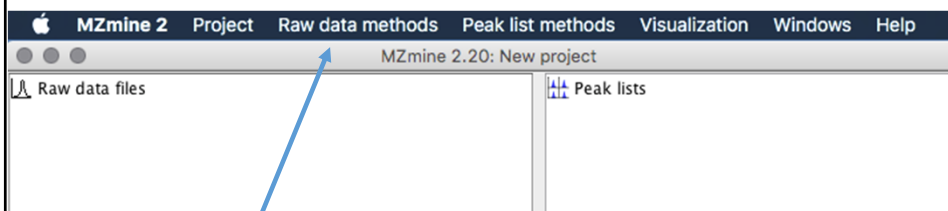
- Go to <http://mzmine.github.io/>
- Download
- Unzip the file and move the folder into Applications
 - There are three starting methods
 - Linux - startMZmine_Linux.sh
 - Mac - startMZmine_MacOSX.command
 - Windows - startMZmine_Windows.bat
 - Double click to start the program

Starting point for Mzmine

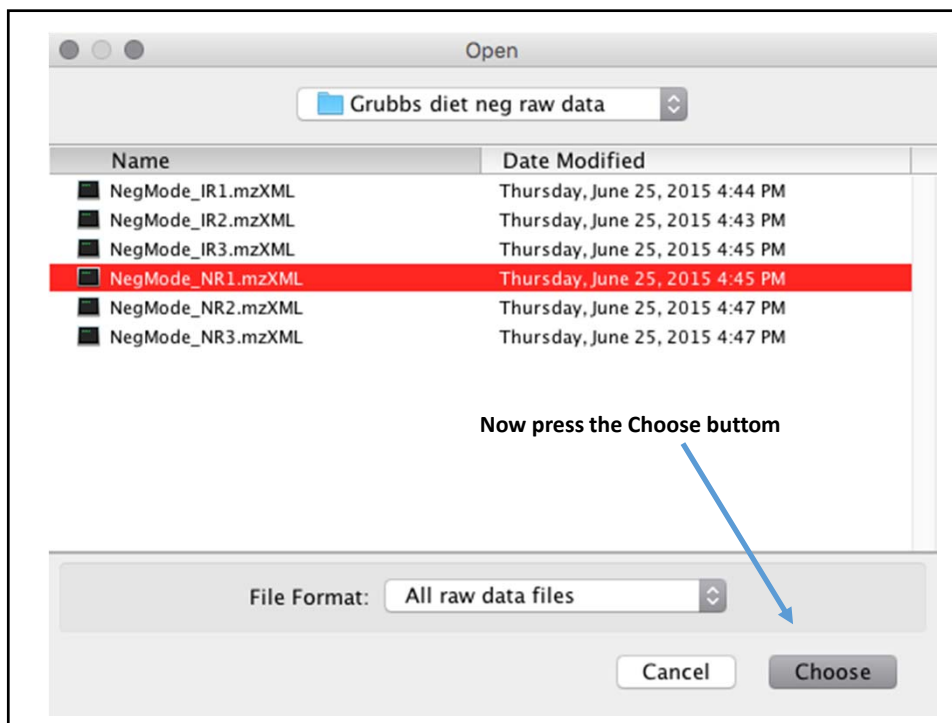
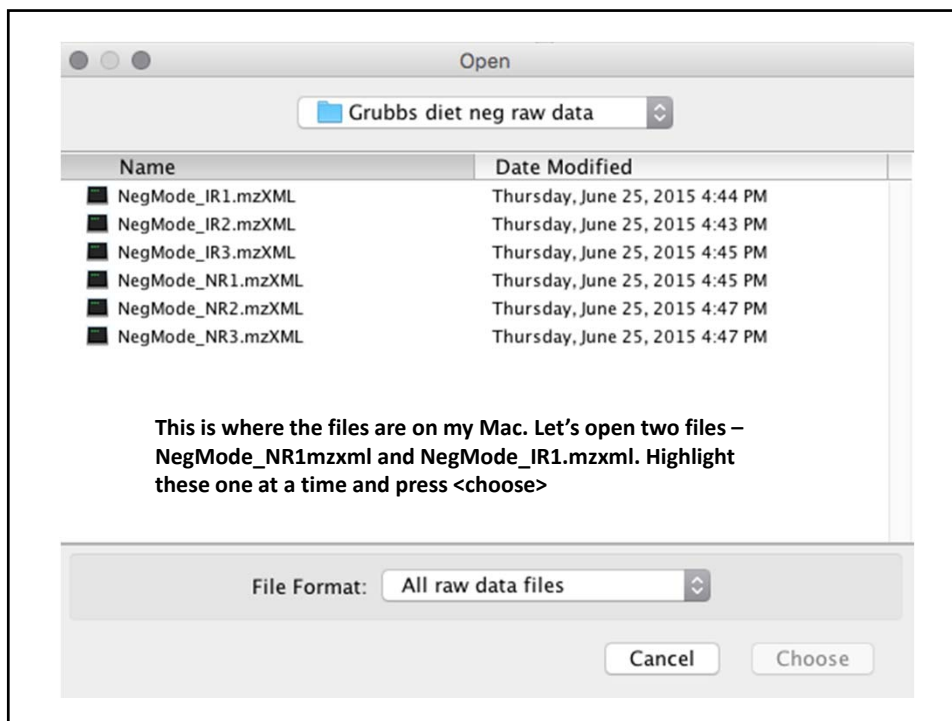
Name	Date Modified	Size	Kind
▶ conf	Apr 3, 2016, 11:29 PM	--	Folder
▶ icons	Aug 14, 2015, 3:15 AM	--	Folder
▶ lib	Apr 3, 2016, 10:42 PM	--	Folder
manual.pdf	Aug 13, 2015, 2:51 AM	653 KB	PDF Document
startMzmine_Linux.sh	Feb 13, 2016, 3:04 PM	3 KB	Shell Script
startMzmine_MacOSX.command	Feb 13, 2016, 3:03 PM	3 KB	Termin...ll script
startMzmine_Windows.bat	Feb 13, 2016, 3:03 PM	5 KB	Document

You will see Terminal open and the program load. Then Java will take over.

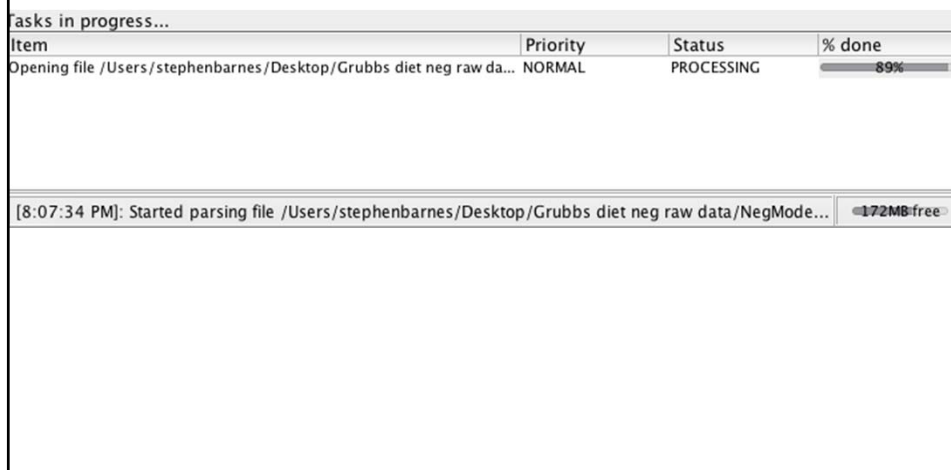
Starting off



Click on Raw data methods and then select "raw data input" from the drop down box



Showing uploading process

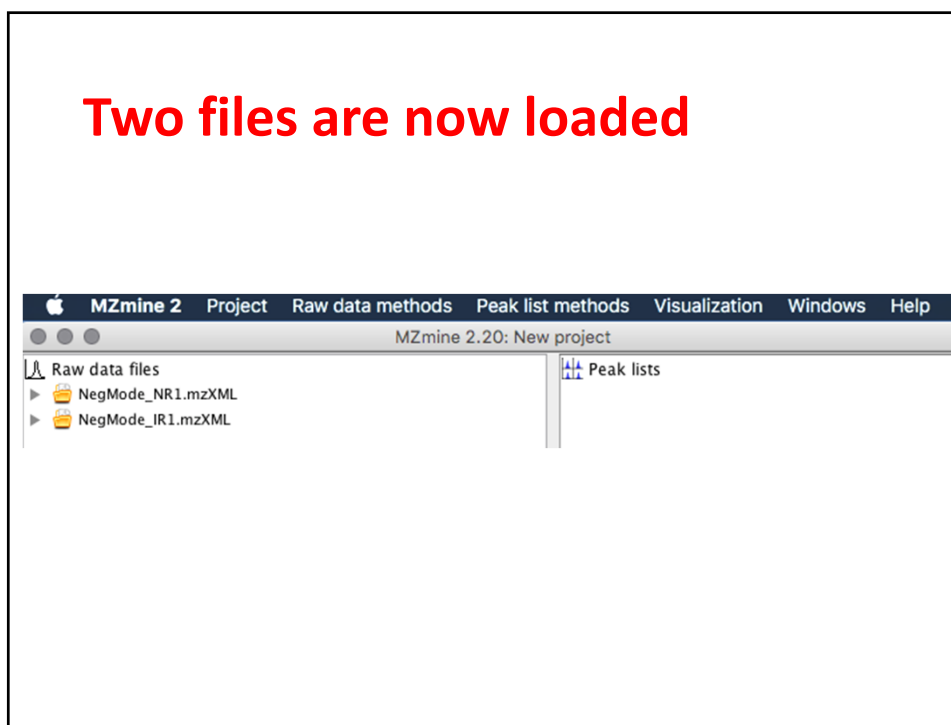


Tasks in progress...

Item	Priority	Status	% done
Opening file /Users/stephenbarnes/Desktop/Grubbs diet neg raw da...	NORMAL	PROCESSING	89%

[8:07:34 PM]: Started parsing file /Users/stephenbarnes/Desktop/Grubbs diet neg raw data/NegMode... 172MB free

Two files are now loaded



MZmine 2 Project Raw data methods Peak list methods Visualization Windows Help

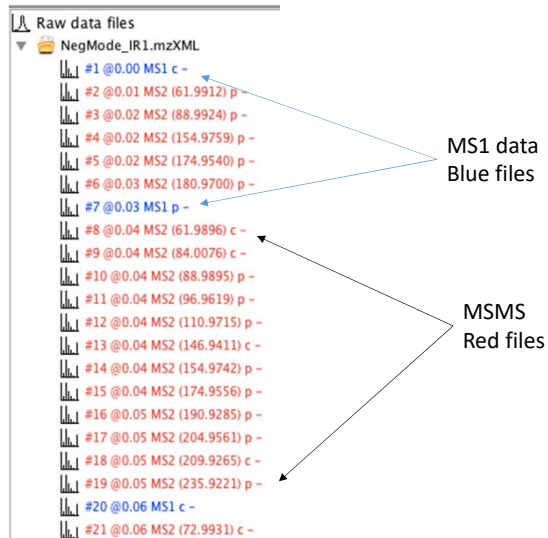
MZmine 2.20: New project

Raw data files

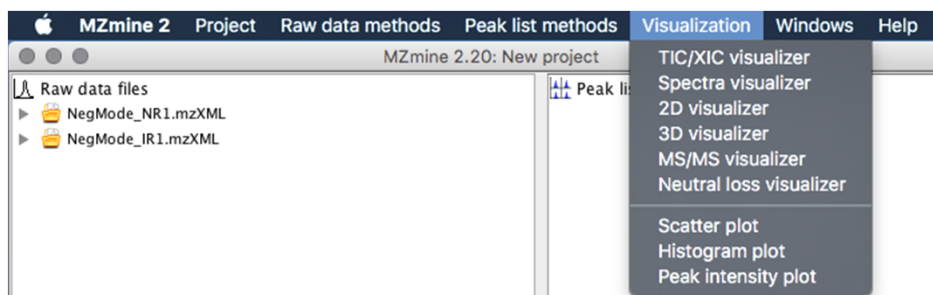
- NegMode_NR1.mzXML
- NegMode_IR1.mzXML

Peak lists

Content of the .mzxml files



Visualization toolbar



We'll start by selecting "3D visualizer" – this allows us to look at all the data

Selecting all the data

Raw data files 0 selected Specific raw data files ...

Scans Retention time: 0.00 – 30.00 min.
MS level: 1 Set filters Clear filters
Polarity: -

m/z 50.0000 - 800.0000 Auto range From mass From formula

Retention time resolution 500

m/z resolution 500

OK Cancel Help

We'll select the file to open next. The 3D visualizer allows one at a time

Selecting the NegMode_NR1 file first

Please set the parameters

Selected Specific raw data files ...

Retention time: 0.00 – 30.00 min.

Please set the parameters

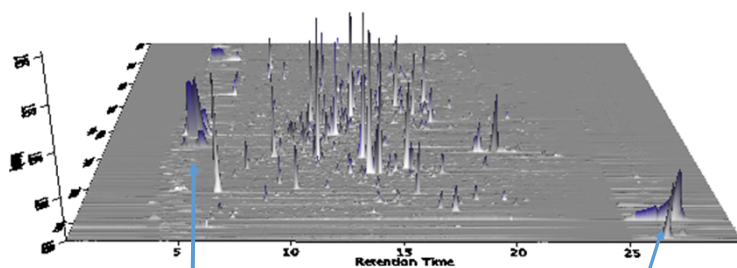
Select files

- NegMode_NR1.mzXML
- NegMode_IR1.mzXML

All Clear

OK Cancel Help

3D view of NegMode_NR1



0-5 min – metabolites that didn't bind to the column are poorly resolved

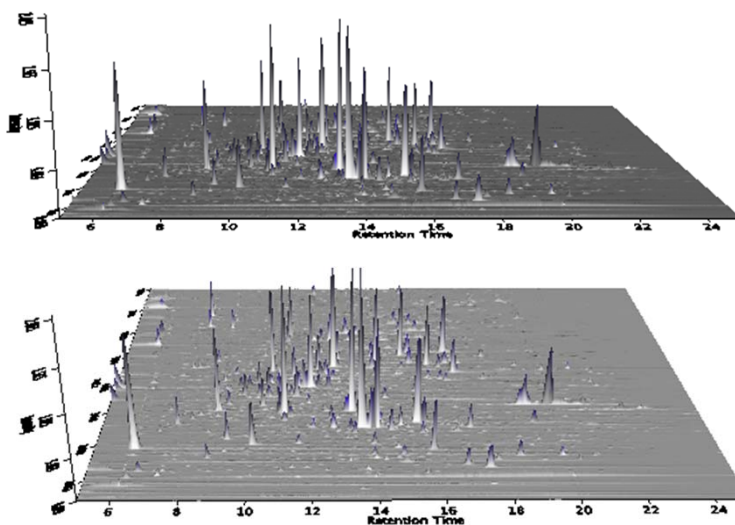
25-30 min – metabolites that stuck too hard to the column

Resetting the parameters

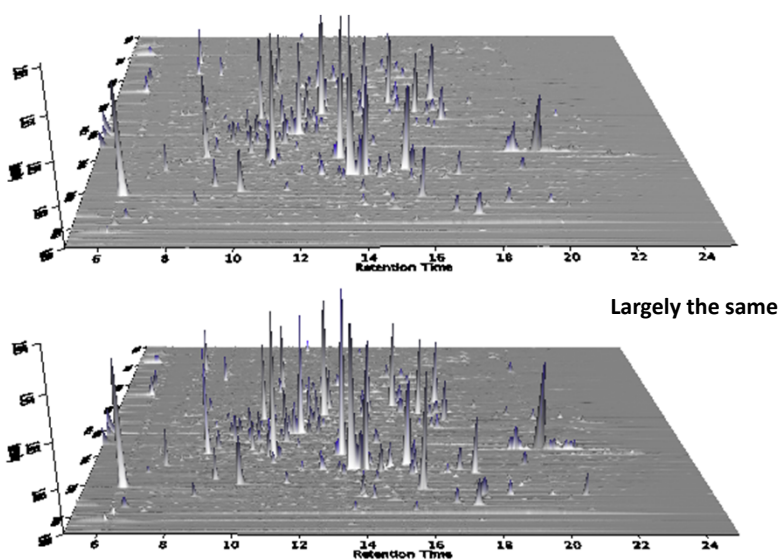
Please set the parameters

Scan number	<input type="text"/>	-	<input type="text"/>
Retention time	<input type="text" value="5.00"/>	-	<input type="text" value="25.00"/> min. <input type="button" value="Auto range"/>
MS level	<input type="text" value="1"/>		
Scan definition	<input type="text"/>		
Polarity	<input type="text" value="-"/>		
Spectrum type	<input type="text" value="Any"/>		

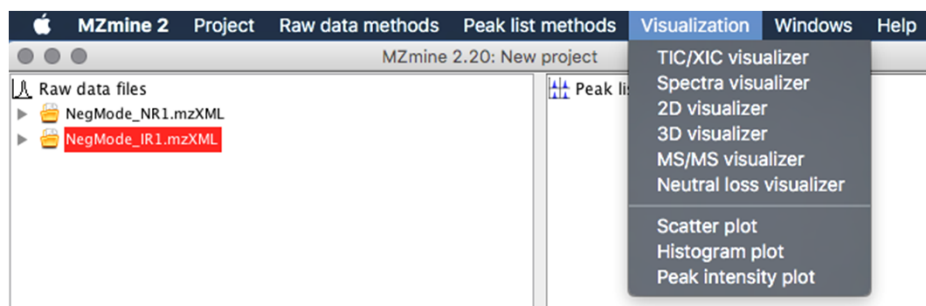
Data from 5-25 min



Comparing NR1 and IR1

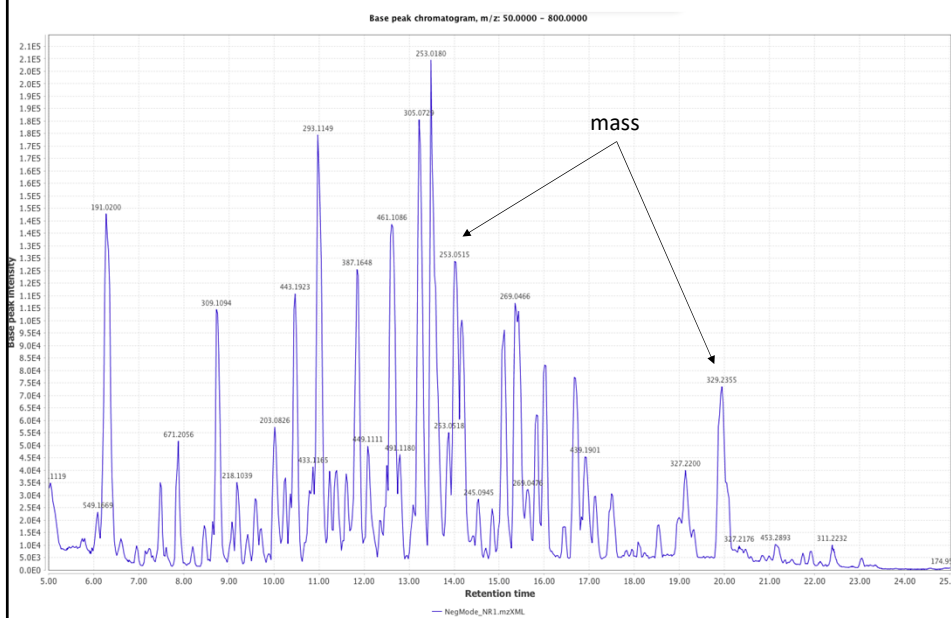


Back to visualization options

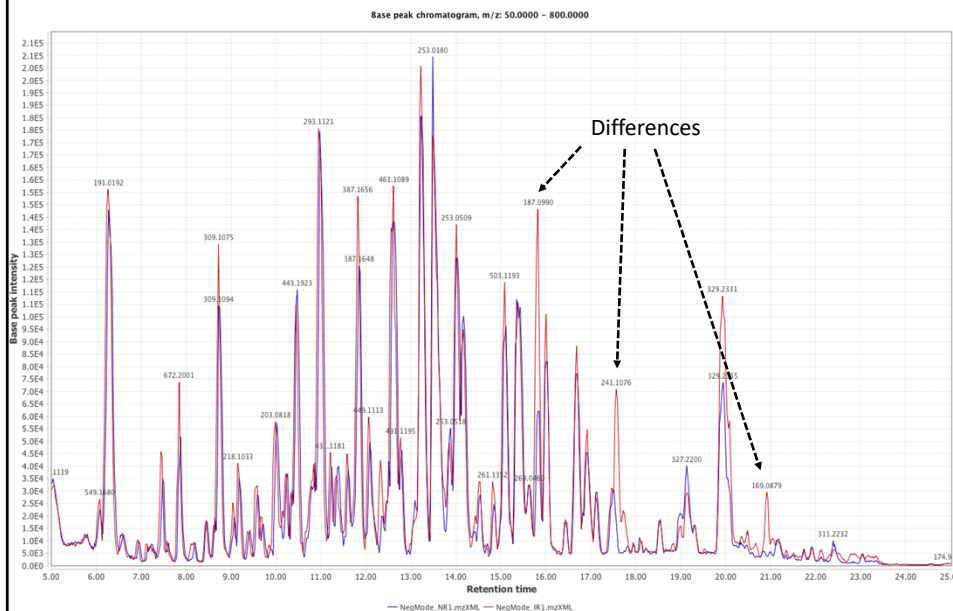


Now we can go to the TIC/XIC visualizer

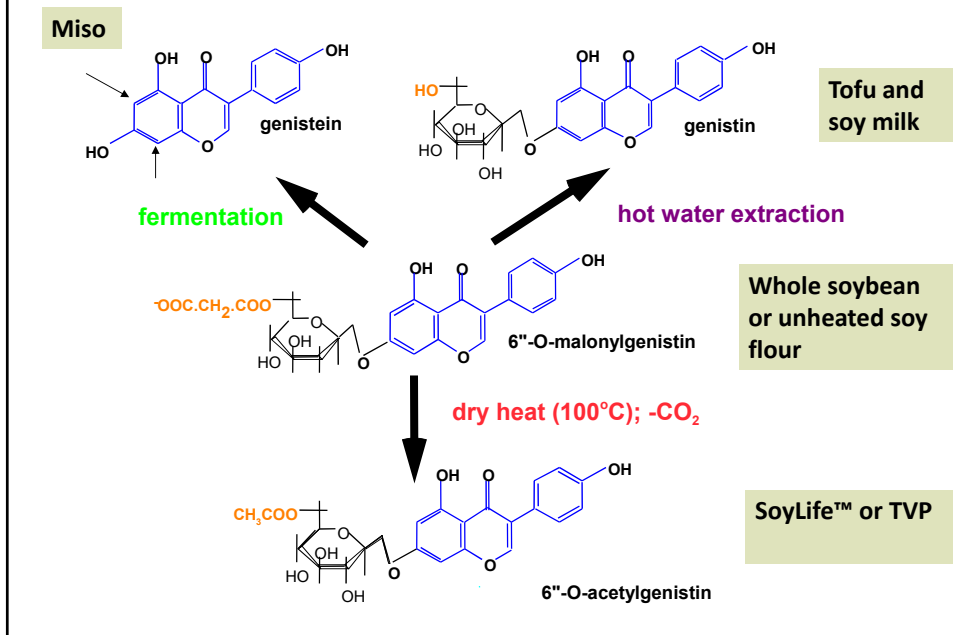
TIC of all ions from m/z 50-800



TIC m/z 50-800 NR/IR comparison



Changing chemistry of isoflavones in soy foods



Let's calculate the mass of genistein [M-H]⁻

- The empirical formula of genistein is C₁₅H₁₀O₅
- If you open the mass calculator Excel file

A	B	C	D	E	F
		C	H	N	O
Name	Empirical formula	12.000000000	1.007825032	14.003074004	15.9949146
hexanol	C6H14O	6	14	0	1
glucose	C6H12O6	6	12		6
genistein	C15H10O5	15	10		5

MW	[M+H] ⁺	[M-H] ⁻
102.104469	103.111745	101.097193
180.063391	181.070668	179.056115
270.052826	271.060103	269.04555

Ions of genistein and its conjugates

Name	Empirical formula	Mass (M)	[M-H] ⁻	[M+HCOOH-H] ⁻ Formate adduct
Genistein	C ₁₅ H ₁₀ O ₅	270.0528	269.0455	315.0510
Genistein glucoside	C ₂₁ H ₂₀ O ₁₀	432.1056	431.0984	477.1039
Genistein acetylglucoside	C ₂₃ H ₂₂ O ₁₁	474.1162	473.1089	519.1144
Genistein malonylglucoside	C ₂₄ H ₂₂ O ₁₃	518.1060	517.0988	563.1043

Setting the mass window

The screenshot shows a dialog box titled "Please set the parameters" with the following fields and controls:

- Raw data files: NegMode_IR1.mzXML, As selected in main window (dropdown), ... (button)
- Scans: Retention time: 5.00 - 25.00 min., MS level: 1, Polarity: - (text), Set filters (button), Clear filters (button)
- Plot type: Base peak intensity (dropdown)
- m/z: 269.0350 - 269.0550 (text), Auto range (button), From mass (button), From formula (button)
- Peaks: [Empty box] (text), All (button), Clear (button)
- Buttons: OK, Cancel, Help

Selecting the mass with the empirical formula

The screenshot shows a dialog box titled "Please set the parameters" with the following fields and controls:

- Formula: C15H10O5 (text)
- Ionization type: [M-H]- (dropdown)
- Charge: 1 (text)
- m/z tolerance: 0.01 (text), m/z or 5.0 (text), ppm (text)
- Buttons: OK, Cancel, Help

Selected mass

Please set the parameters

Raw data files: NegMode_IR1.mzXML As selected in main window ...

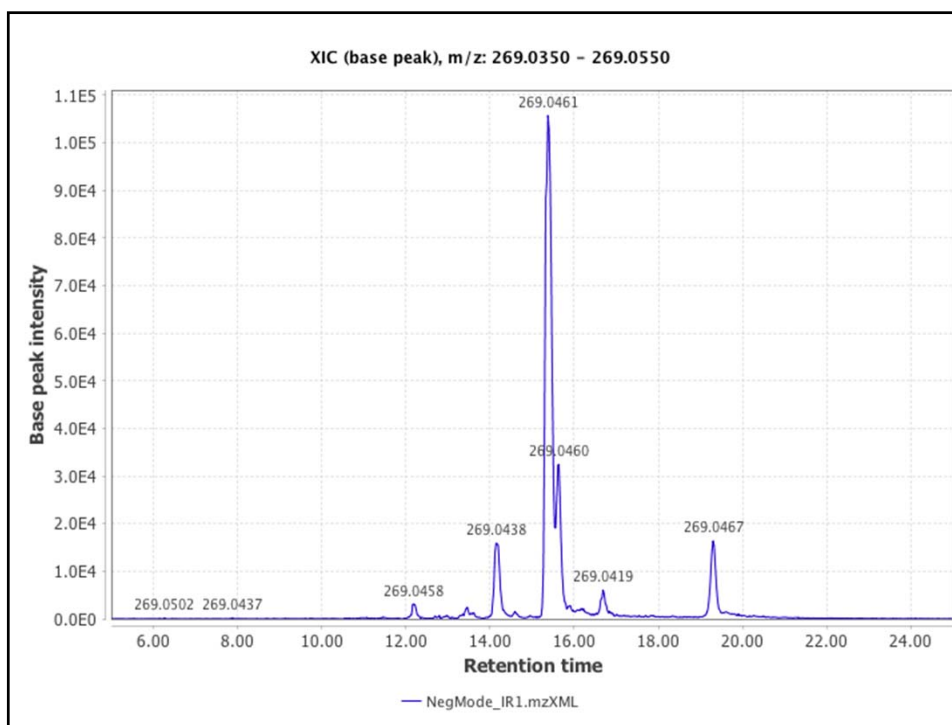
Scans: Retention time: 5.00 - 25.00 min. MS level: 1 Polarity: -
Set filters Clear filters

Plot type: Base peak intensity

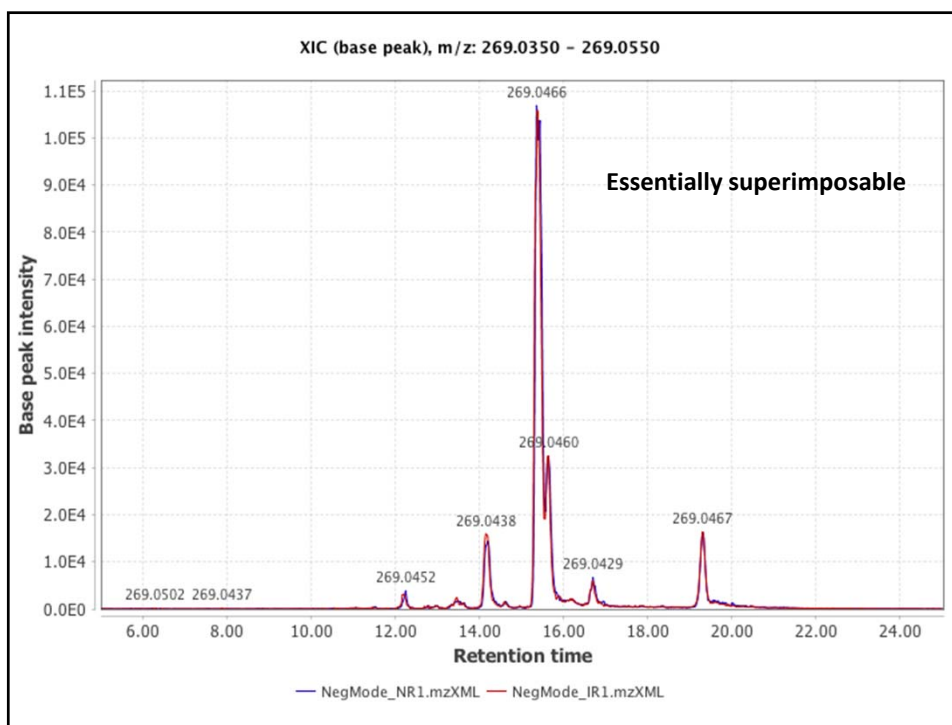
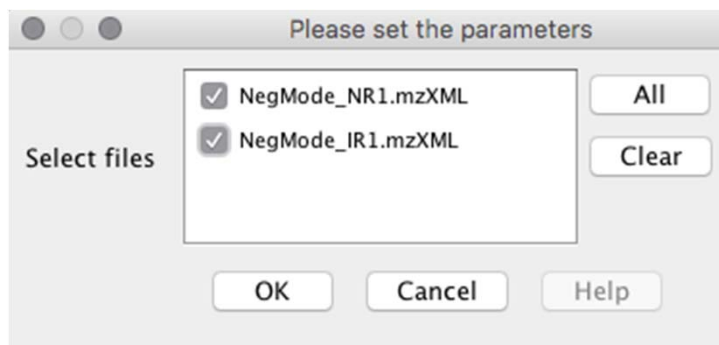
m/z: 269.0344 - 269.0545 Auto range From mass From formula

Peaks: [Empty list] All Clear

OK Cancel Help



Selecting both files



Getting MS/MS data

Select TIC/XIC and reset the parameters

Please set the parameters

Scan number -

Retention time - min.

MS level ←

Scan definition

Polarity

Spectrum type

Ready to go

Please set the parameters

Raw data files NegMode_IR1.mzXML ...

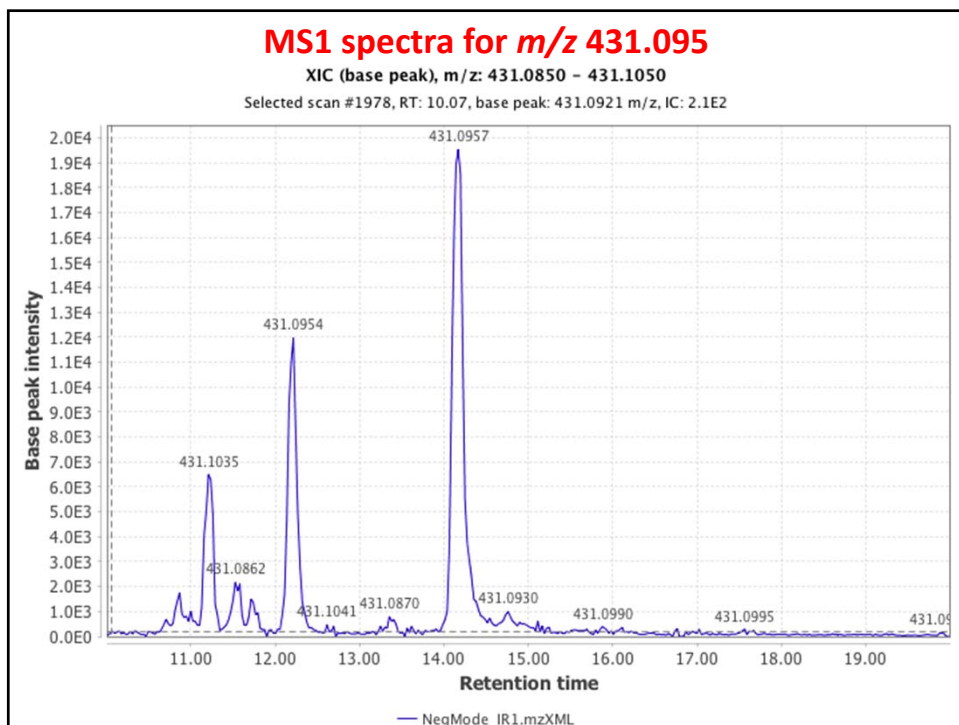
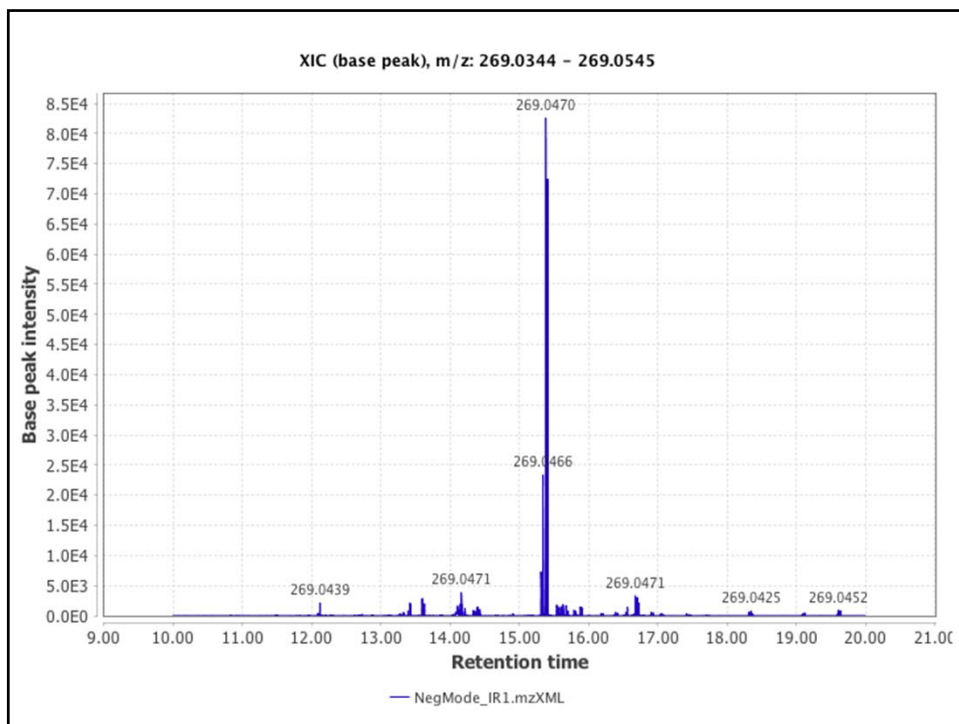
Scans Retention time: 10.00 - 20.00 min.

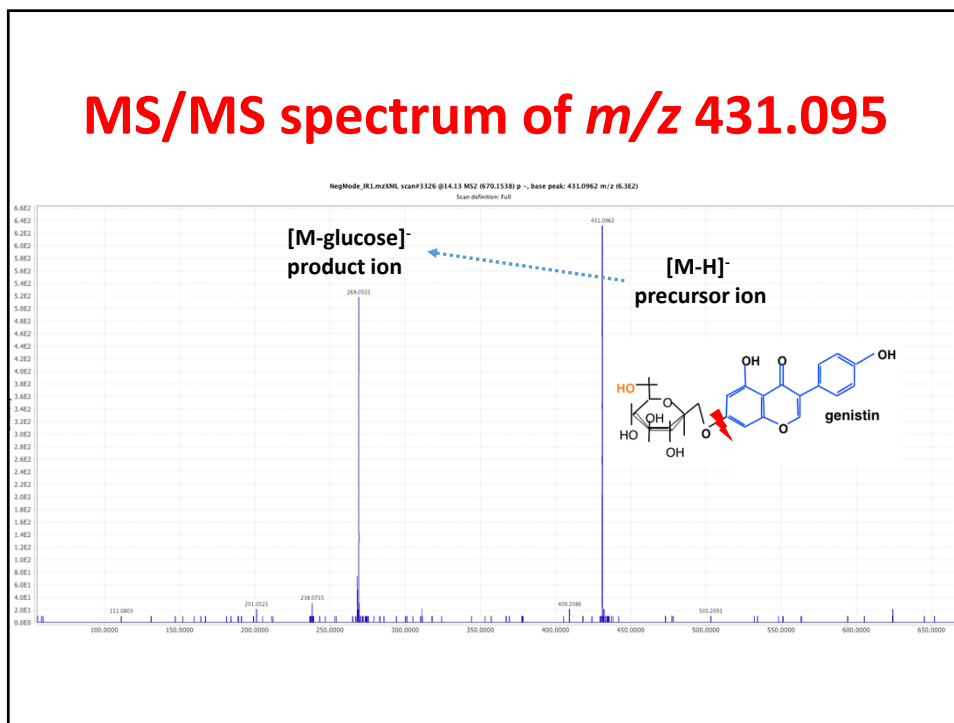
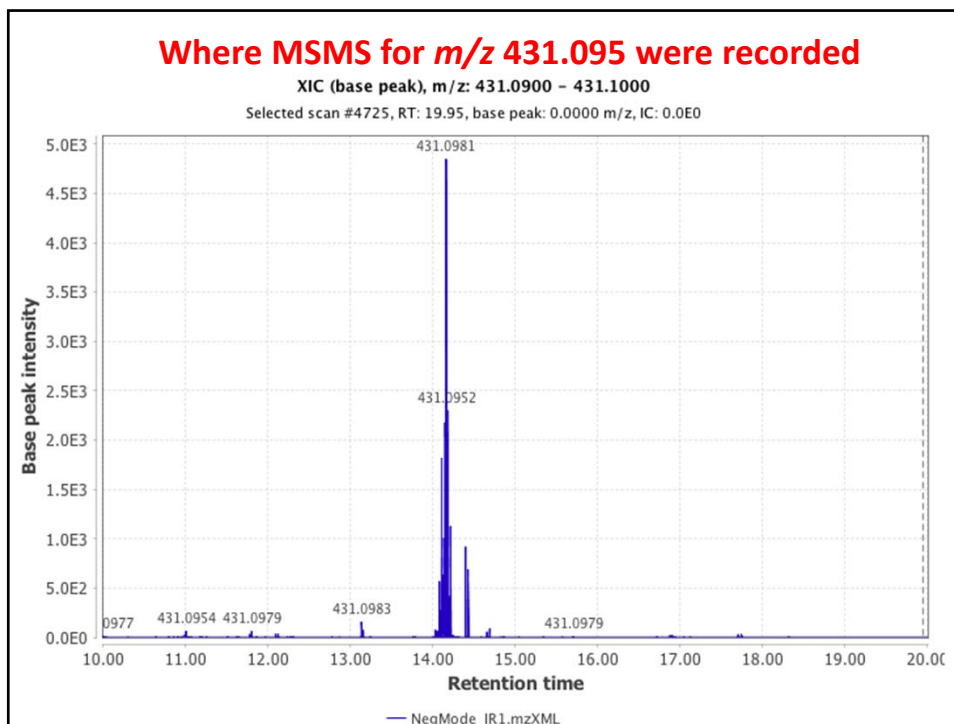
MS level: 2
Polarity: -

Plot type

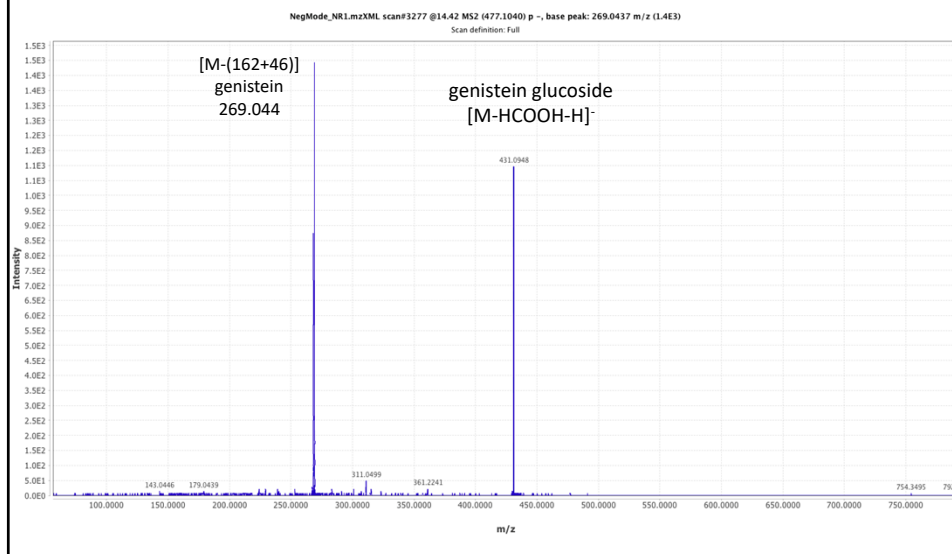
m/z -

Peaks

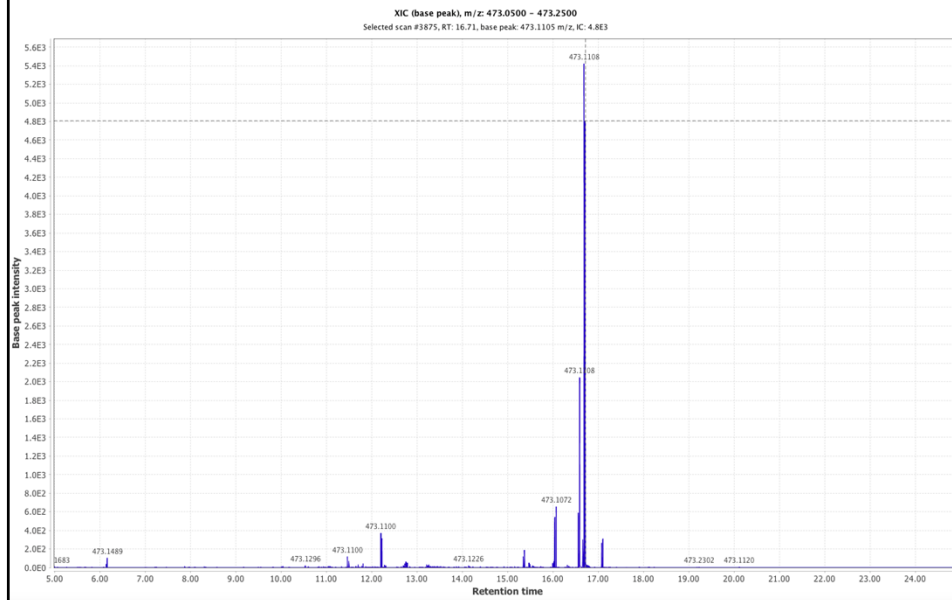




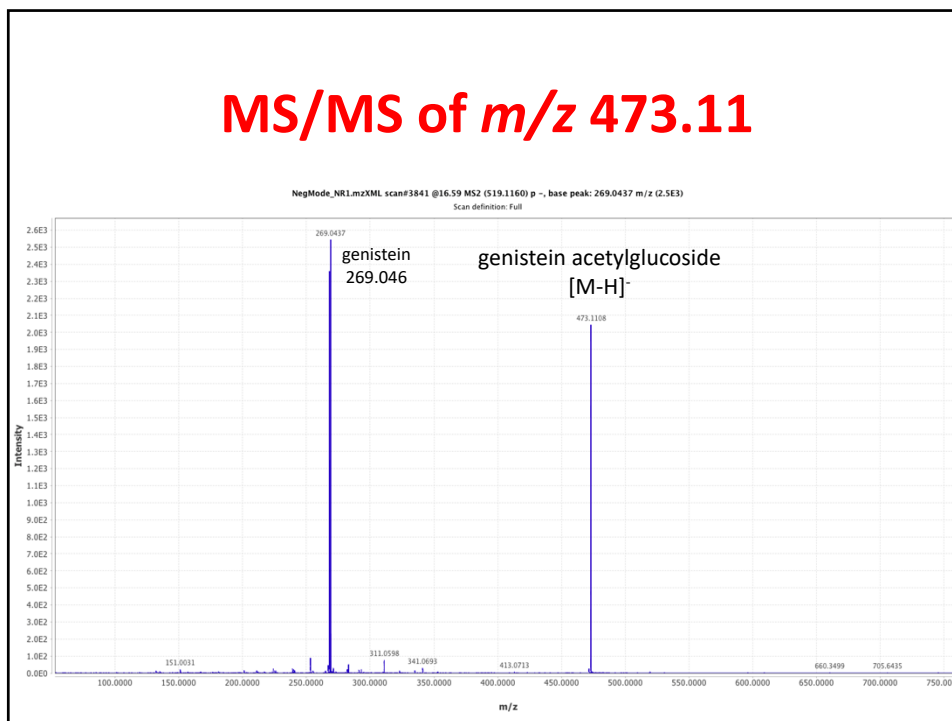
MSMS of m/z 477.10 – formate adduct



MSMS spectra m/z 473.11 collected in run



MS/MS of m/z 473.11



Stats analysis (coming later) reveals m/z 241 as having a big change

Please set the parameters

Raw data files NegMode_IR1.mzXML As selected in main window ...

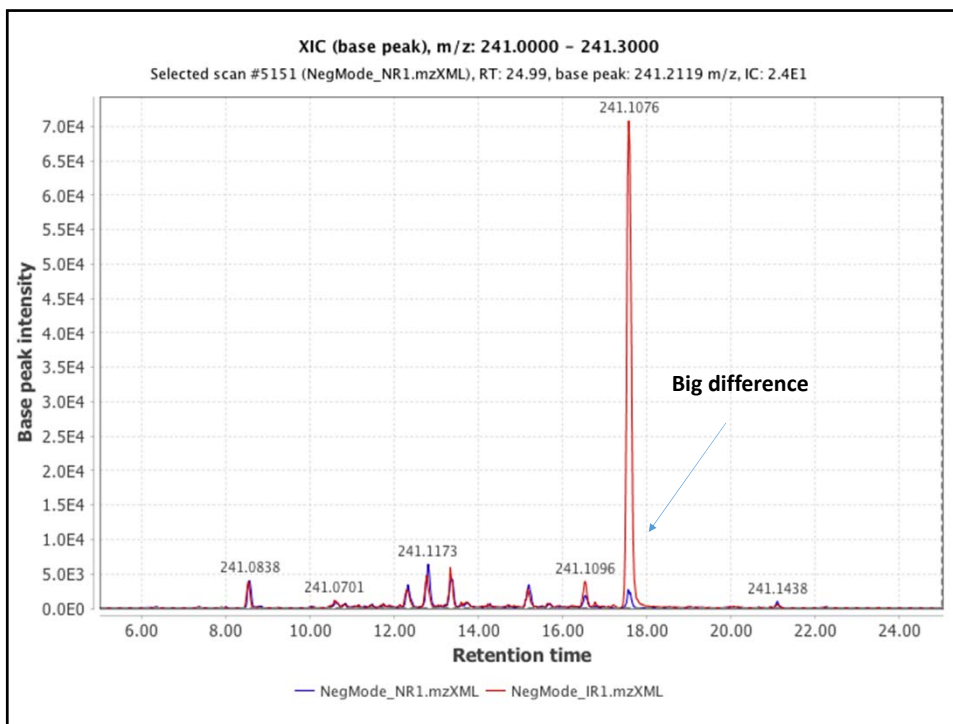
Scans Retention time: 5.00 - 25.00 min. MS level: 1 Polarity: - Set filters Clear filters

Plot type Base peak intensity

m/z 241.0000 - 241.3000 Auto range From mass From formula

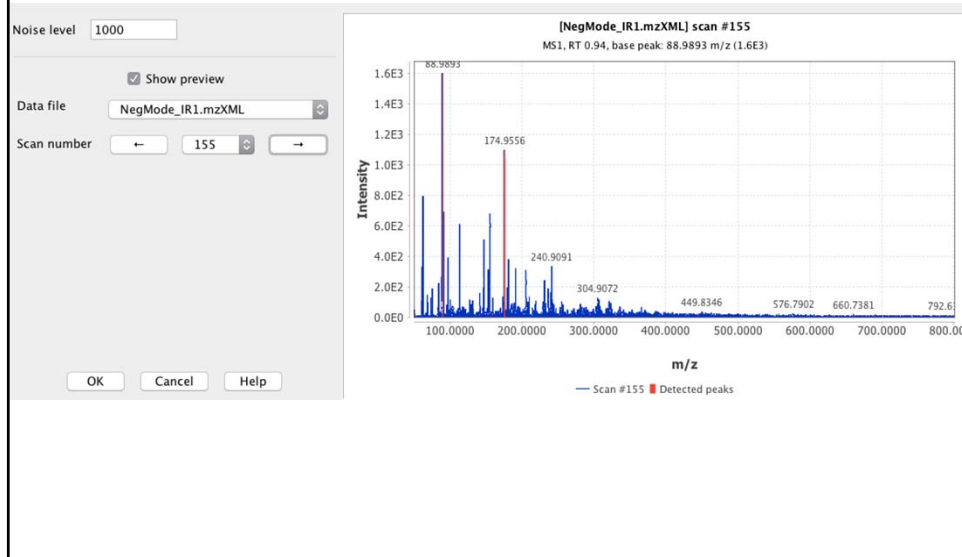
Peaks All Clear

OK Cancel Help

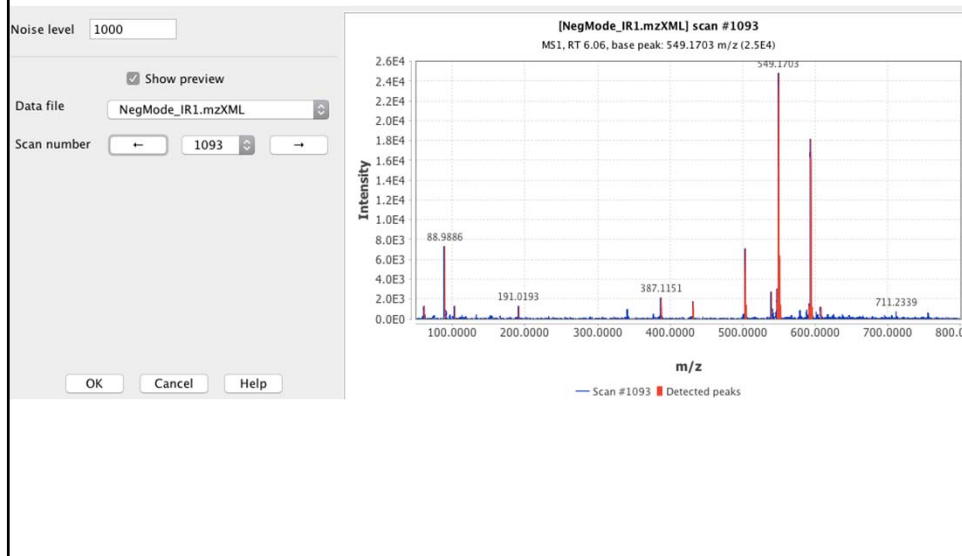


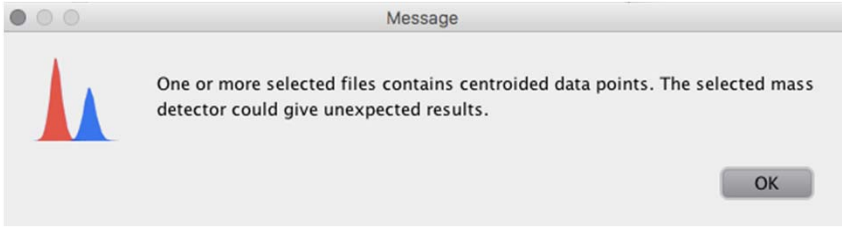
Identifying the masses in .mzxml files

Setting the background



Setting the background





Message

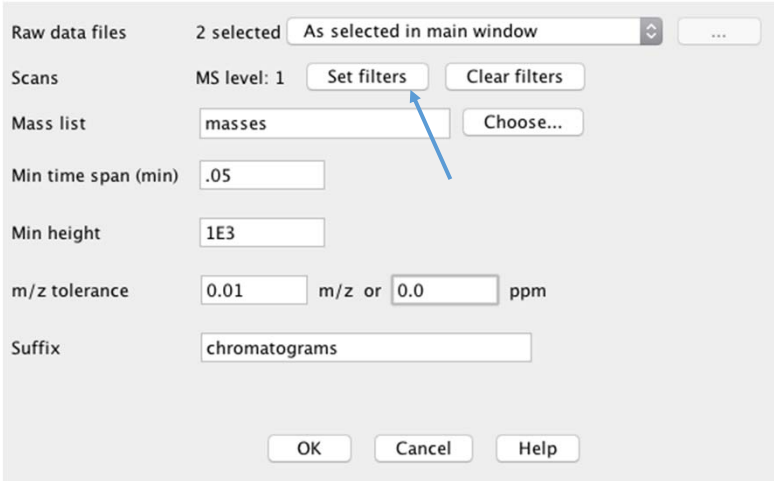
One or more selected files contains centroided data points. The selected mass detector could give unexpected results.

OK

You can ignore this – it's due to the MSMS data in the file

Using the masses to create chromatograms

Go to *Raw data methods, peak detection, chromatogram builder*



Raw data files 2 selected As selected in main window ...

Scans MS level: 1 Set filters Clear filters

Mass list masses Choose...

Min time span (min) .05

Min height 1E3

m/z tolerance 0.01 m/z or 0.0 ppm

Suffix chromatograms

OK Cancel Help

Raw data files 2 selected As selected in main window

Scans Retention time: 5.00 – 25.00 min.
MS level: 1
Polarity: -

Mass list masses Choose...

Min time span (min) .05

Min height 1E3

m/z tolerance 0.01 m/z or 0.0 ppm

Suffix chromatograms

OK Cancel Help

This stops ppm being used

Chromatogram deconvolution

Go to *Peak list methods, peak detection, chromatogram deconvolution*

Please set the parameters

Peak lists 2 selected As selected in main window

Suffix deconvoluted

Algorithm Wavelets (XCMS)

Remove original peak list

OK Cancel Help

S/N threshold 10

Wavelet scales 0.25 – 5.00

Peak duration range 0.05 – 0.50

Peak integration method Use raw data

Show preview

OK Cancel Help

Chromatogram deconvolution output

Raw data files

- NegMode_IR1.mzXML
- NegMode_NRI1.mzXML

NegMode_NRI1.mzXML chromatograms deconvoluted

- #1 60.9941 m/z @6.04
- #2 75.0093 m/z @5.25
- #3 80.9658 m/z @13.51
- #4 89.0245 m/z @6.17
- #5 93.0346 m/z @17.49
- #6 111.0074 m/z @6.30
- #7 116.0504 m/z @10.02
- #8 117.0194 m/z @7.23
- #9 119.0500 m/z @13.99
- #10 121.0290 m/z @13.70
- #11 121.0295 m/z @15.54
- #12 125.0972 m/z @15.83
- #13 128.0358 m/z @6.60
- #14 130.0868 m/z @7.09
- #15 130.0869 m/z @7.37
- #16 131.0711 m/z @13.72
- #17 131.0709 m/z @13.96
- #18 137.0247 m/z @17.49
- #19 138.0280 m/z @17.49
- #20 138.0560 m/z @11.34

Peak lists

- NegMode_NRI1.mzXML chromatograms
- NegMode_IR1.mzXML chromatograms
- NegMode_NRI1.mzXML chromatograms deconvoluted
- NegMode_IR1.mzXML chromatograms deconvoluted

Pseudo-spectrum #363
#367 431.0966 m/z @14.20

File Name	Mass	RT	Height	Area
NegMode_NRI1.mzXML	431.0966	14.20	2.2E4	8.6E4

Identifying isotopes - CAMERA

Go to *peak list methods, peak identification, CAMERA search*

Peak lists 2 selected As selected in main window

FWHM sigma 0.2

FWHM percentage 1.0 %

Isotopes max. charge 3

Isotopes max. per cluster 4

Isotopes mass tolerance 0.010 m/z or 10.0 ppm

Correlation threshold 0.9

Correlation p-value 0.05

OK Cancel Help

Peak and chromatogram alignment

Go to *Peak list method, alignment, join aligner*

Peak lists 2 selected As selected in main window

Peak list name Aligned peak list

m/z tolerance 0.015 m/z or 5.0 ppm

Weight for m/z 0.5

Retention time tolerance 0.5 absolute (min)

Weight for RT 0.5

Require same charge state

Require same ID

Compare isotope pattern Setup..

OK Cancel Help

Aligned data

Aligned peak list

- #1 60.9943 m/z @6.04 Pseudo-spectrum #215
- #2 75.0093 m/z @5.23 Pseudo-spectrum #342
- #3 80.9657 m/z @13.51 Pseudo-spectrum #358
- #4 89.0251 m/z @6.15 Pseudo-spectrum #111
- #5 93.0349 m/z @17.47 Pseudo-spectrum #030
- #6 111.0079 m/z @6.30 Pseudo-spectrum #423
- #7 116.0503 m/z @10.00 Pseudo-spectrum #014
- #8 117.0193 m/z @7.22 Pseudo-spectrum #138
- #9 119.0500 m/z @13.99 Pseudo-spectrum #404
- #10 121.0291 m/z @13.68 Pseudo-spectrum #103
- #11 121.0295 m/z @15.55 Pseudo-spectrum #100
- #12 125.0975 m/z @15.81 Pseudo-spectrum #083
- #13 128.0356 m/z @6.60 Pseudo-spectrum #089
- #14 130.0868 m/z @7.09 Pseudo-spectrum #324
- #15 130.0867 m/z @7.35 Pseudo-spectrum #296
- #16 131.0710 m/z @13.71 Pseudo-spectrum #160
- #17 131.0707 m/z @13.97 Pseudo-spectrum #110

Double click on Aligned peak list

ID	Average		...	Peak shape	NegMode_NR1.mzXML			NegMode_IR1.mzXML		
	m/z	RT			Status	Height	Area	Status	Height	Area
1	60.9943	6.04	P...		●	2.1E3	1.1E4	●	1.4E3	9.8E3
2	75.0093	5.23	P...		●	1.5E3	8.8E3	●	1.6E3	1.4E4
3	80.9657	13.51	P...		●	2.6E3	1.6E4	●	2.7E3	1.3E4
4	89.0251	6.15	P...		●	1.2E4	4.0E4	●	1.0E4	3.1E4
5	93.0349	17.47	P...		●	1.9E3	1.3E4	●	2.1E3	1.3E4
6	111.0079	6.30	P...		●	8.5E3	8.1E4	●	9.1E3	8.2E4
7	116.0503	10.00	P...		●	3.7E3	1.6E4	●	3.4E3	1.6E4
8	117.0193	7.22	P...		●	8.7E3	6.2E4	●	8.1E3	5.3E4
9	119.0500	13.99	P...		●	2.1E3	8.9E3	●		
10	121.0291	13.68	P...		●	4.2E3	3.0E4	●	1.0E4	9.2E4
11	121.0295	15.55	P...		●	7.3E3	4.1E4	●	1.1E4	7.6E4
12	125.0975	15.81	P...		●	1.9E3	7.8E3	●	3.1E3	2.1E4
13	128.0356	6.60	P...		●	1.3E4	1.1E5	●	1.3E4	1.2E5
14	130.0868	7.09	P...		●	1.8E3	5.3E3	●		
15	130.0867	7.35	P...		●	3.2E3	1.4E4	●	2.3E3	1.0E4
16	131.0710	13.71	P...		●	6.2E3	5.5E4	●	3.7E3	2.7E4

Organizing by retention time

ID	Average		...	Peak shape	NegMode_NR1.mzXML			NegMode_IR1.mzXML		
	m/z	RT			Status	Height	Area	Status	Height	Area
748	270.0457	14.14	P...		●			●	3.4E3	1.5E4
853	477.1528	14.14	P...		●			●	7.6E3	2.0E4
921	647.1545	14.14	P...		●			●	2.6E3	1.1E4
471	490.1299	14.14	P...		●	4.1E3	1.9E4	●		
375	432.0996	14.16	P...		●	6.0E3	3.9E4	●	4.9E3	4.0E4
462	478.1060	14.16	P...		●	2.3E4	2.1E5	●	2.7E4	1.9E5
463	479.1084	14.16	P...		●	6.4E3	4.3E4	●	9.0E3	4.9E4
852	477.1022	14.17	P...		●			●	9.5E4	8.7E5
854	477.1674	14.17	P...		●			●	2.3E3	1.2E4
138	269.0447	14.17	P...		●	1.4E4	1.4E5	●	1.6E4	1.5E5
367	431.0963	14.18	P...		●	2.2E4	8.6E4	●	2.0E4	1.8E5
636	670.1571	14.20	P...		●	3.9E3	2.0E4	●		
156	283.0604	14.22	P...		●	1.3E3	5.4E3	●		

Class exercise

- Load the NR1_pos and ir1_pos .mzxml files
- Locate the ions that have the genistein ion (in positive - what is its m/z value?)
 - Get MSMS spectra of each one
- As for the negative data, run the masses, generate chromatograms, deconvolute the chromatograms, run CAMERA, and align the data