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

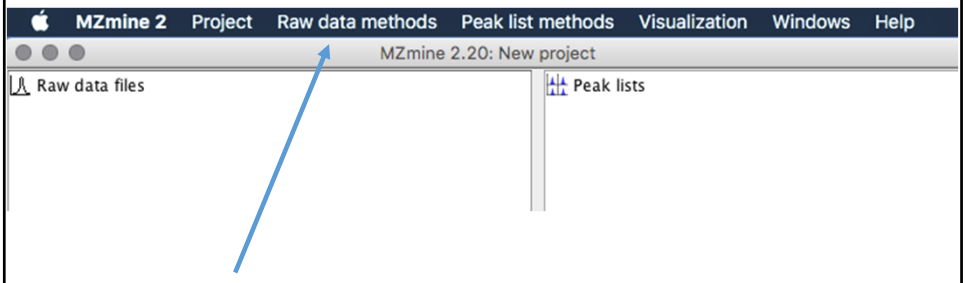
- **Go to http://mzmine.github.io/**
- **Download**
- **Unzip the file and move the folder into Applications**
  - **There are three starting methods**
    - **Linux - startMZmine_Linux.sh**
    - **Mac - startMZmine_MacOSX.command**
    - **Windows - startMZmine_Windows.bat**
  - **Double click to start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

Starting point for Mzmine

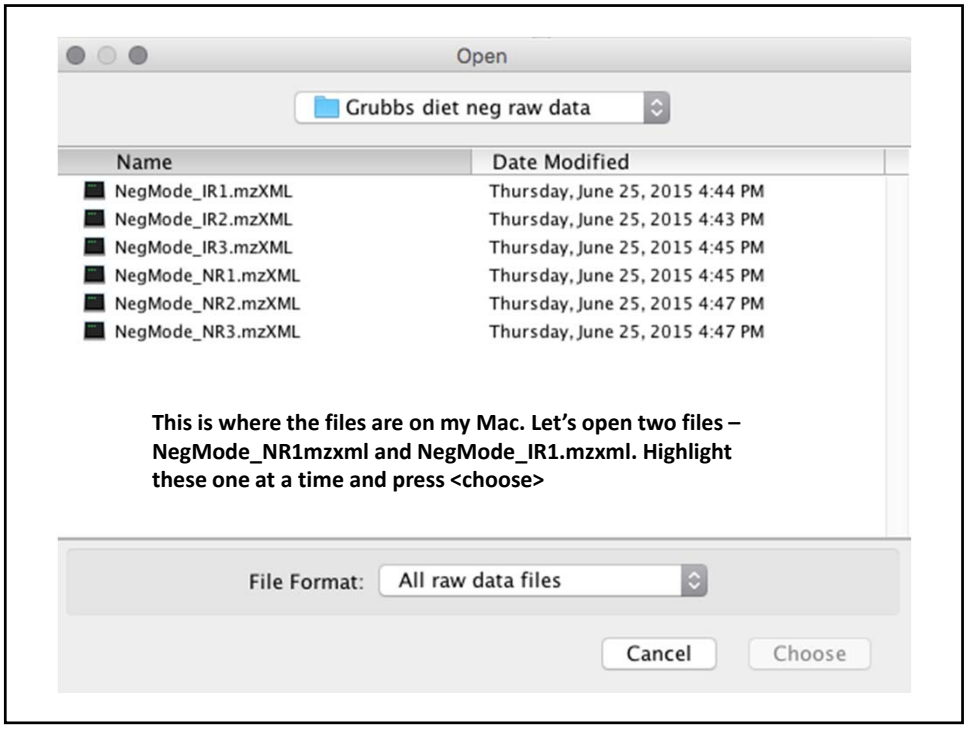
Name	Date Modified	Size	Kind
conf	Apr 3, 2016, 11:29 PM	--	Folder
icons	Aug 14, 2015, 3:15 AM	--	Folder
lib	Apr 3, 2016, 10:42 PM	--	Folder
manual.pdf	Aug 13, 2015, 2:51 AM	653 KB	PDF Document
startMZmine_Linux.sh	Feb 13, 2016, 3:04 PM	3 KB	Shell Script
startMZmine_MacOSX.command	Feb 13, 2016, 3:03 PM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

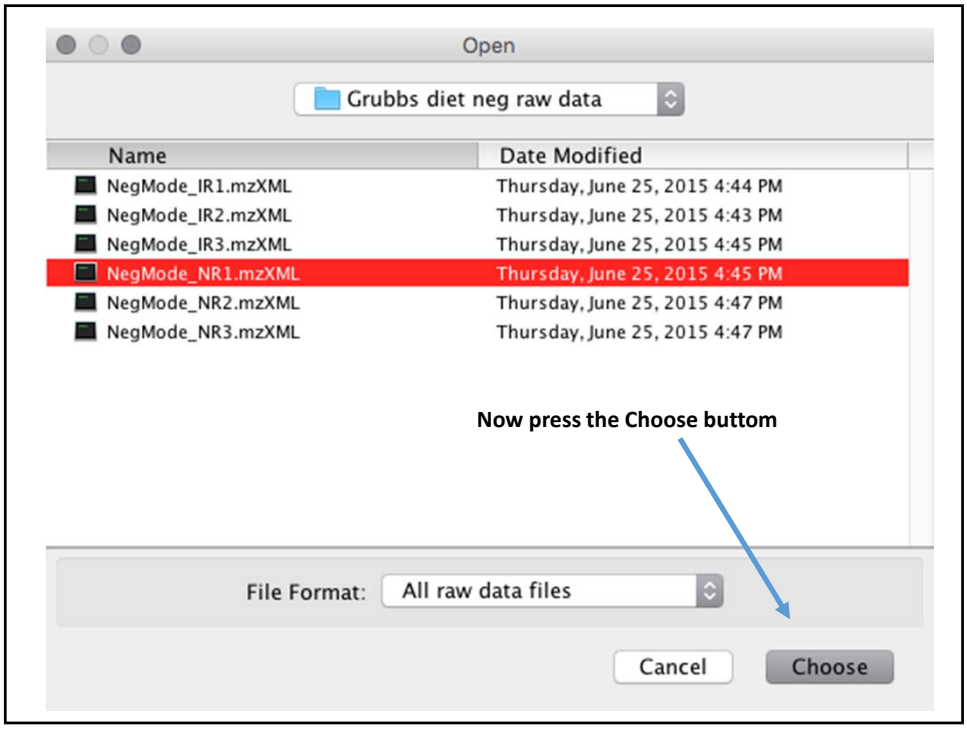
You will see Terminal open and the program load. Then Java will take over.

Starting off

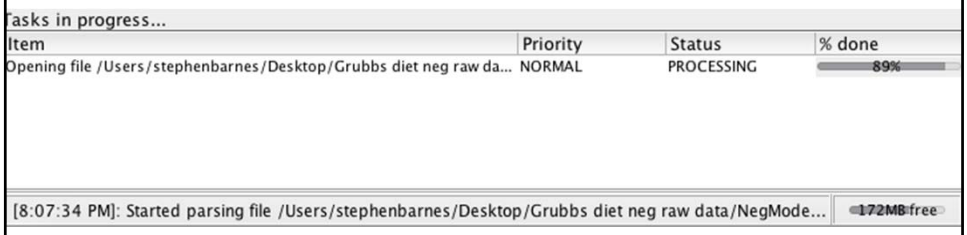


Click on Raw data methods and then select "raw data input" from the drop down box

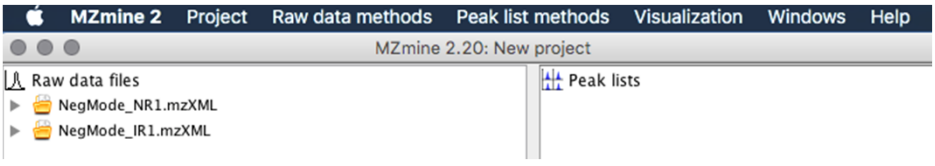

Open
Grubbs diet neg raw data
Name
Date Modified
NegMode_IR1.mzXML
Thursday, June 25, 2015 4:44 PM
NegMode_IR2.mzXML
Thursday, June 25, 2015 4:43 PM
NegMode_IR3.mzXML
Thursday, June 25, 2015 4:45 PM
NegMode_NR1.mzXML
Thursday, June 25, 2015 4:45 PM
NegMode_NR2.mzXML
Thursday, June 25, 2015 4:47 PM
NegMode_NR3.mzXML
Thursday, June 25, 2015 4:47 PM
This is where the files are on my Mac. Let's open two files – NegMode_NR1mzxml and NegMode_IR1.mzxml. Highlight these one at a time and press <choose>
File Format:
All raw data files
Cancel
Choose


Open
Grubbs diet neg raw data
Name
Date Modified
NegMode_IR1.mzXML
Thursday, June 25, 2015 4:44 PM
NegMode_IR2.mzXML
Thursday, June 25, 2015 4:43 PM
NegMode_IR3.mzXML
Thursday, June 25, 2015 4:45 PM
NegMode_NR1.mzXML
Thursday, June 25, 2015 4:45 PM
NegMode_NR2.mzXML
Thursday, June 25, 2015 4:47 PM
NegMode_NR3.mzXML
Thursday, June 25, 2015 4:47 PM
Now press the Choose buttom
File Format:
All raw data files
Cancel
Choose

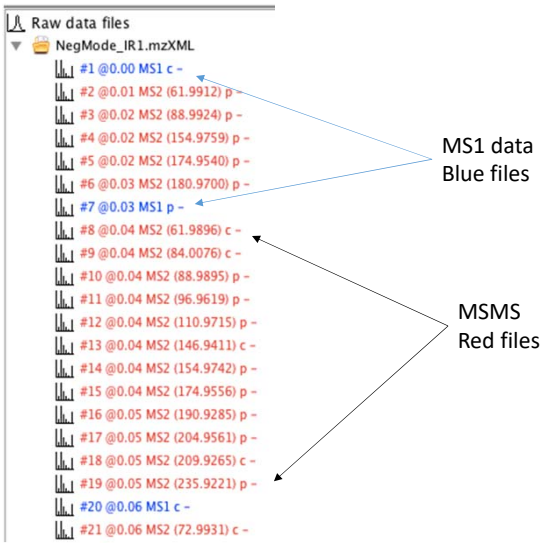
Showing uploading process



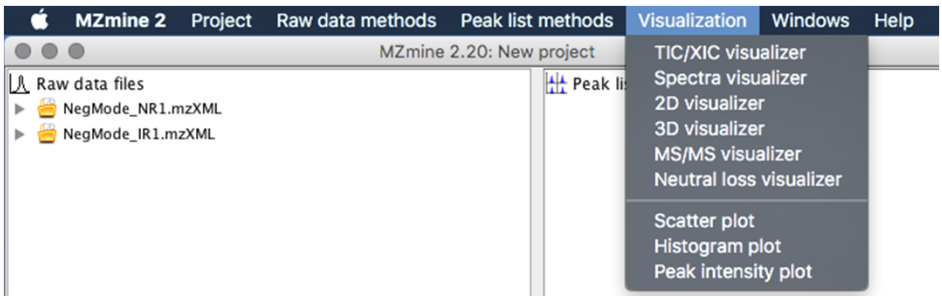
Two files are now loaded



Content of the .mzxml files

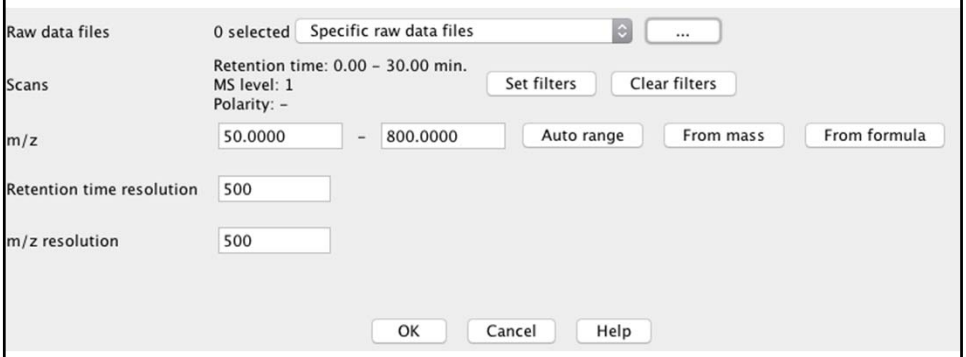


Visualization toolbar



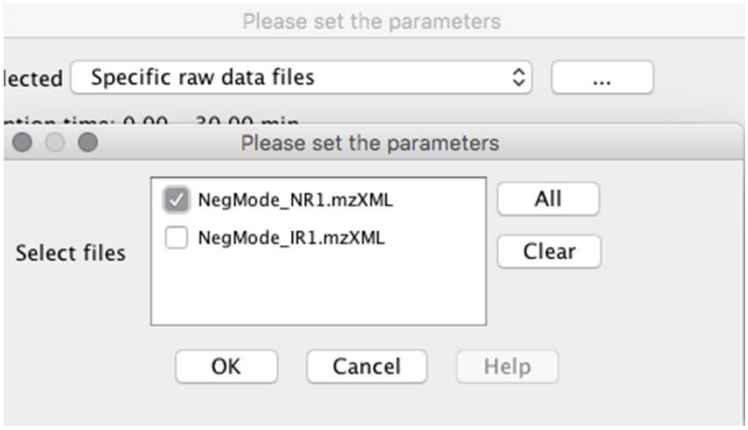
We'll start by selecting "3D visualizer" – this allows us to look at all the data

Selecting all the data

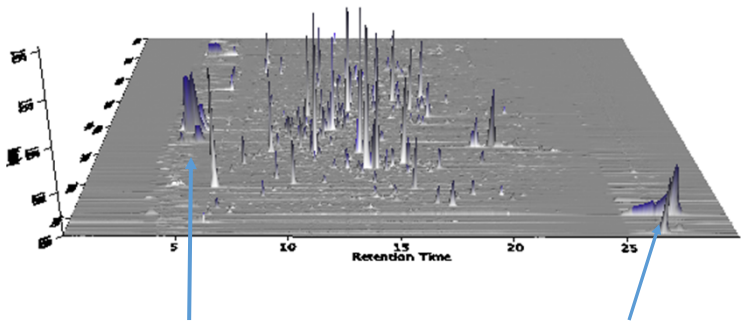


We'll select the file to open next. The 3D visualizer allows one at a time

Selecting the NegMode_NR1 file first



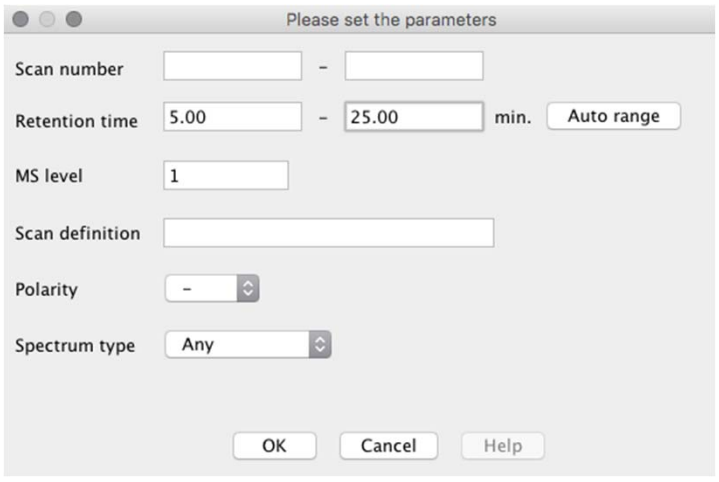
3D view of NegMode_NR1



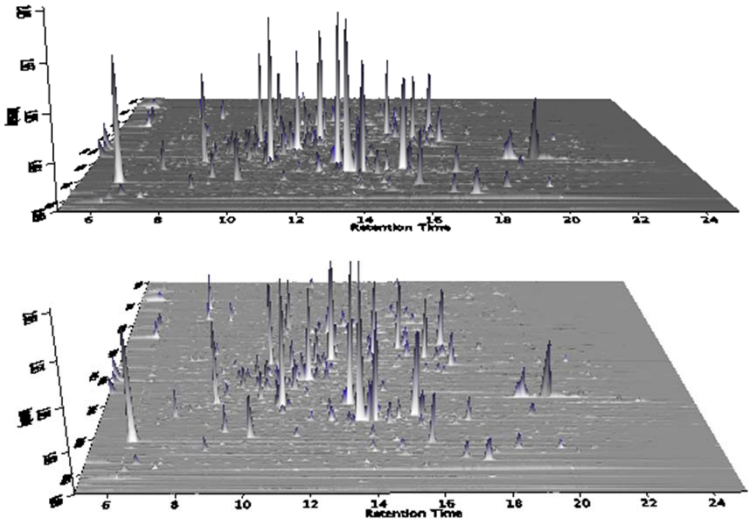
0-5 min – metabolites that didn't bind to the column are poorly resolved

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

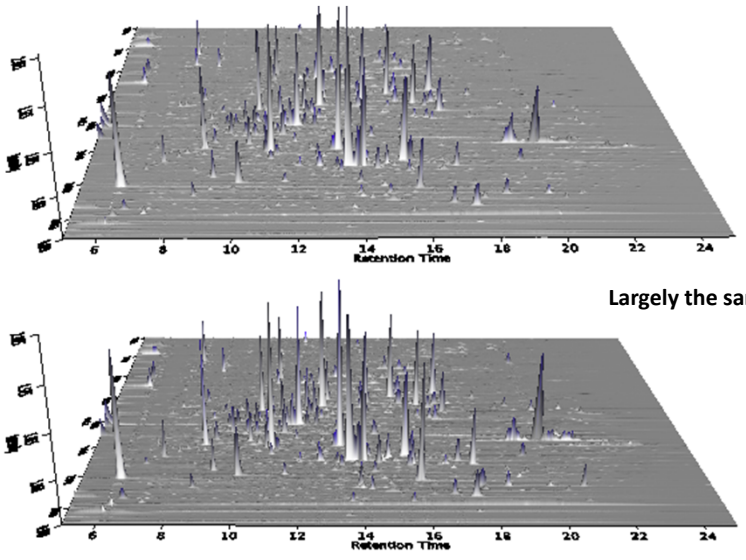
Resetting the parameters



Data from 5-25 min

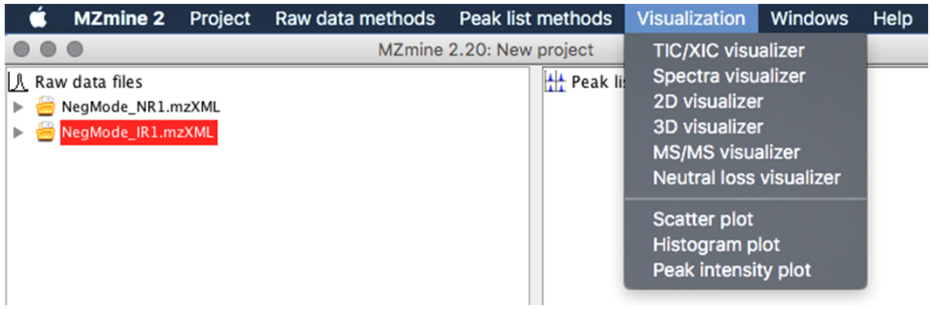


Comparing NR1 and IR1



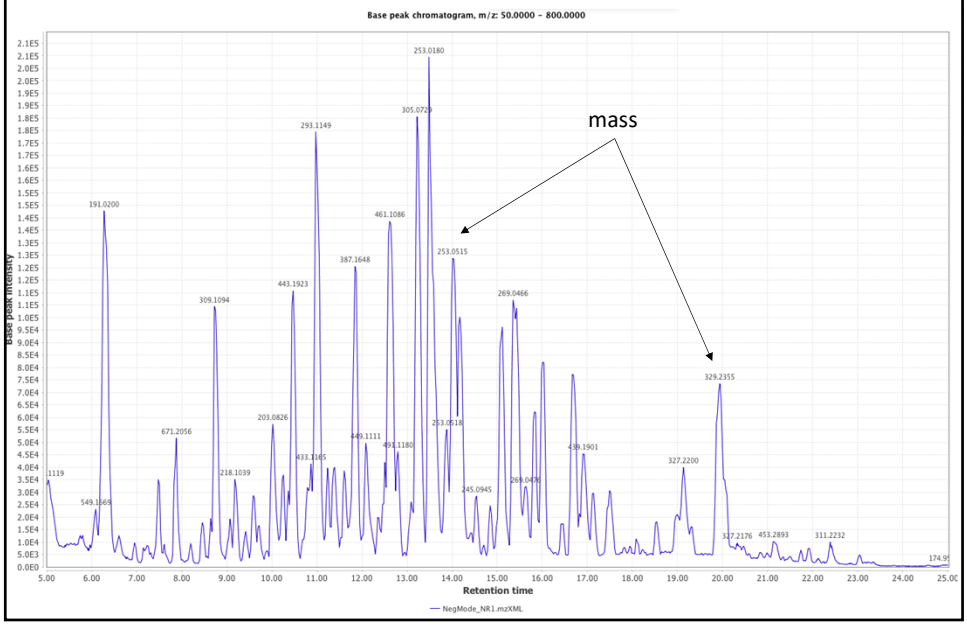
Largely the same

Back to visualization options

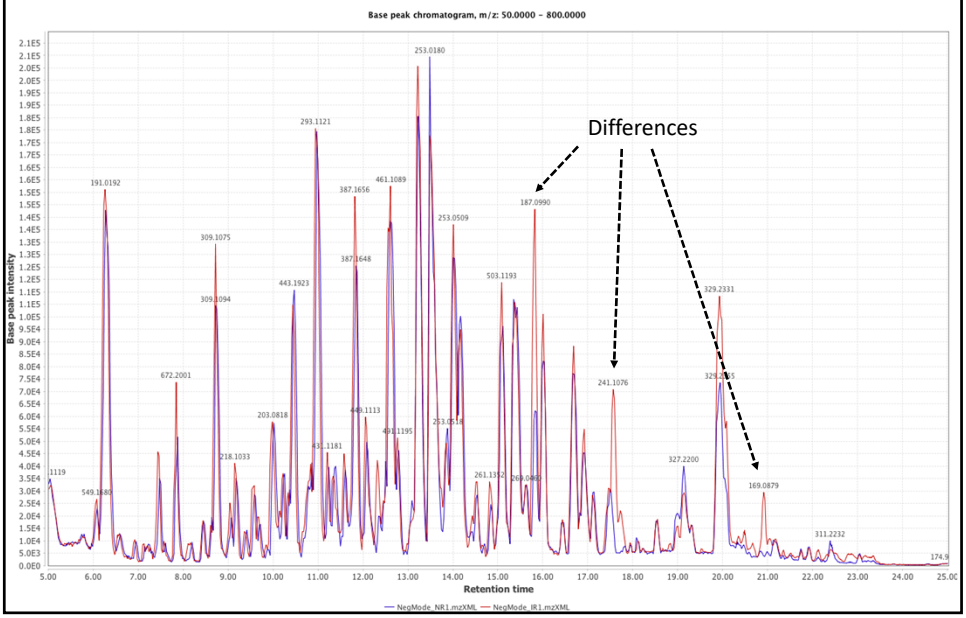


Now we can go to the TIC/XIC visualizer

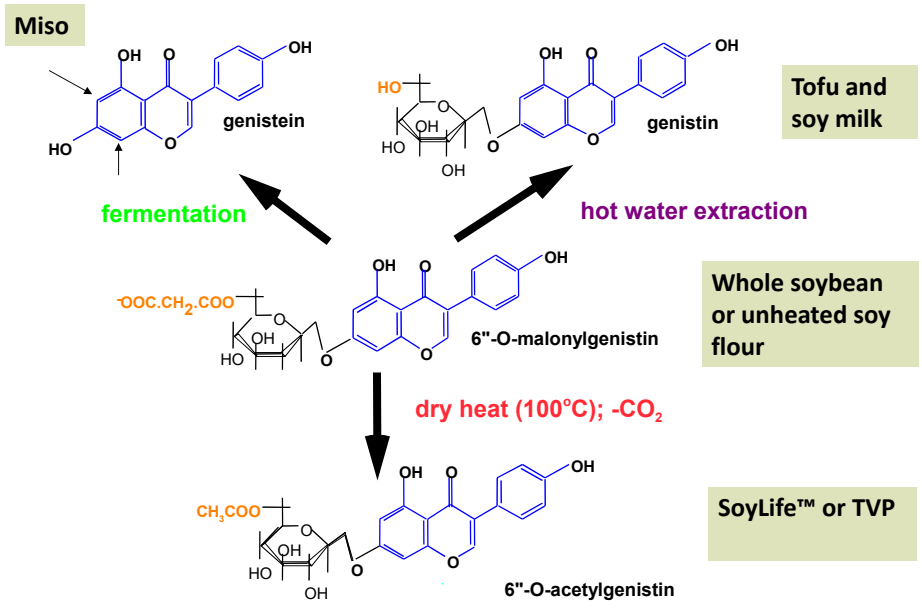
TIC of all ions from *m/z* 50-800



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 $C_{15}H_{10}O_5$**
- **If you open the mass calculator Excel file**

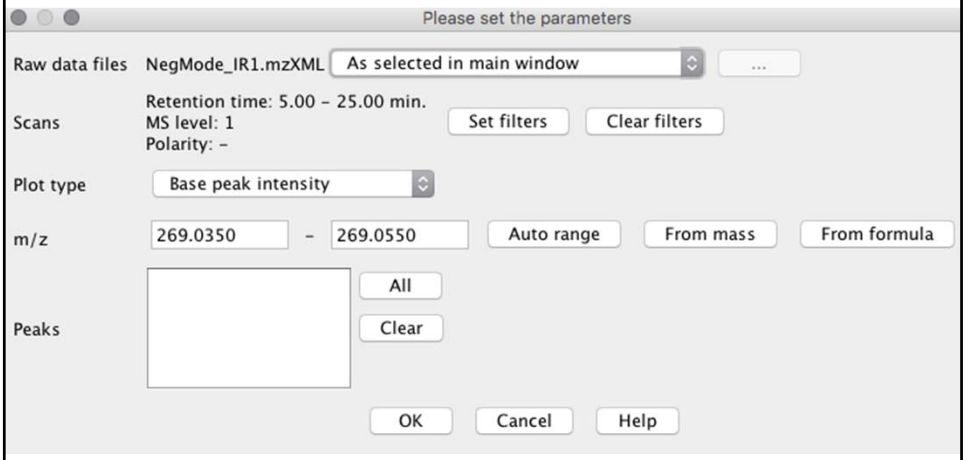
	A	B	C	D	E	F
			C	H	N	O
	Name	Empirical formula	12.000000000	1.007825032	14.003074004	15.9949146
	hexanol	C6H14O	6	14	0	1
	glucose	C6H12O6	6	12		6
	genistein	C15H10O5	15	10		5

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

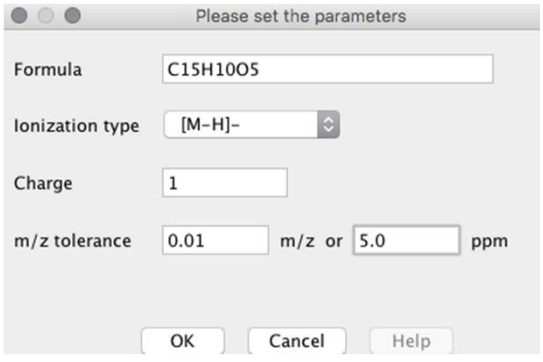
Ions of genistein and its conjugates

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

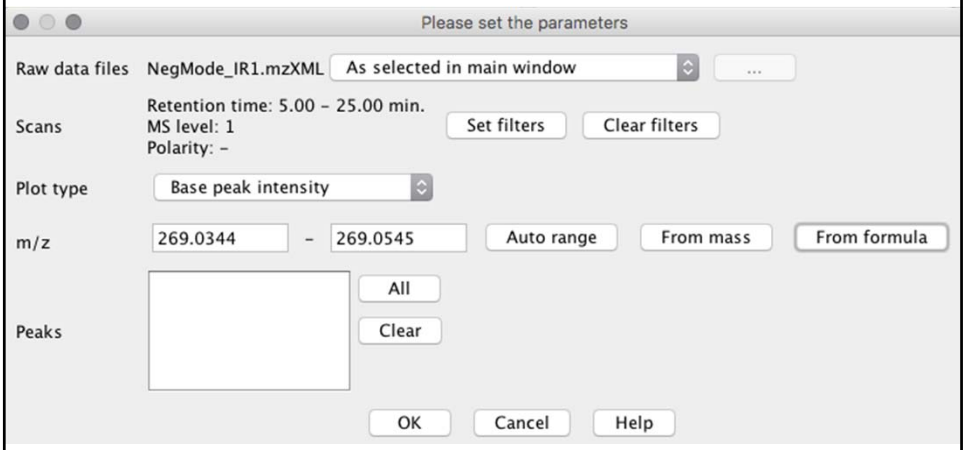
Setting the mass window

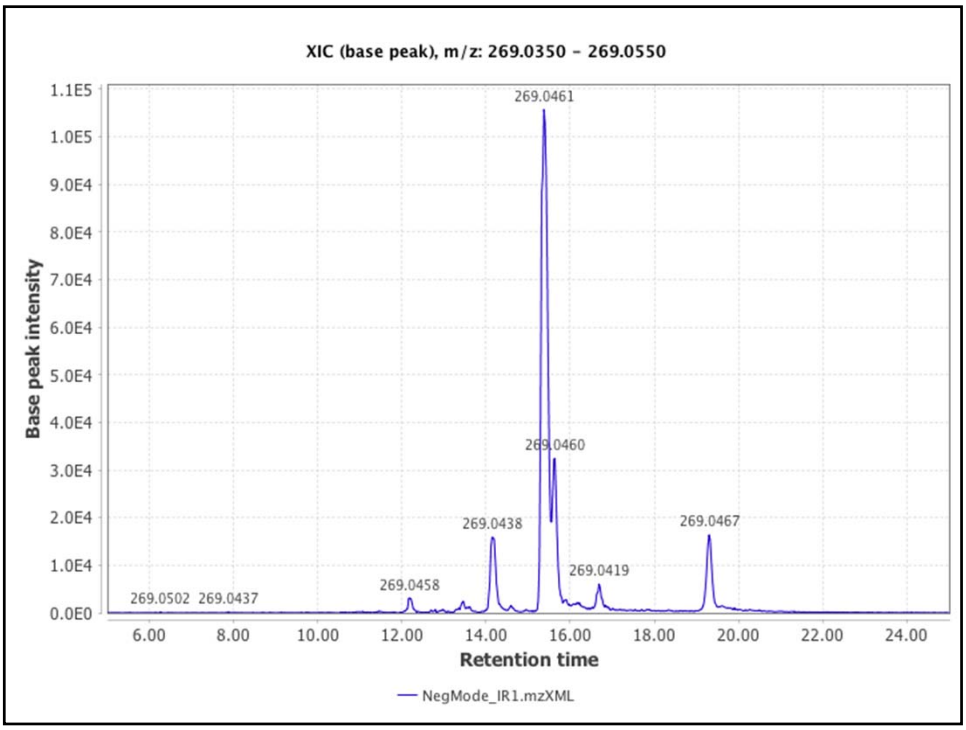


Selecting the mass with the empirical formula

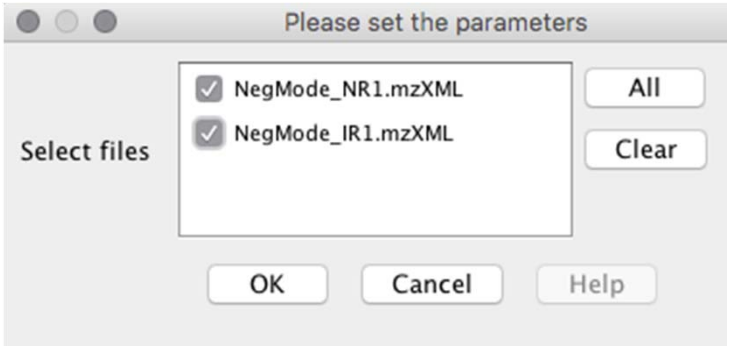


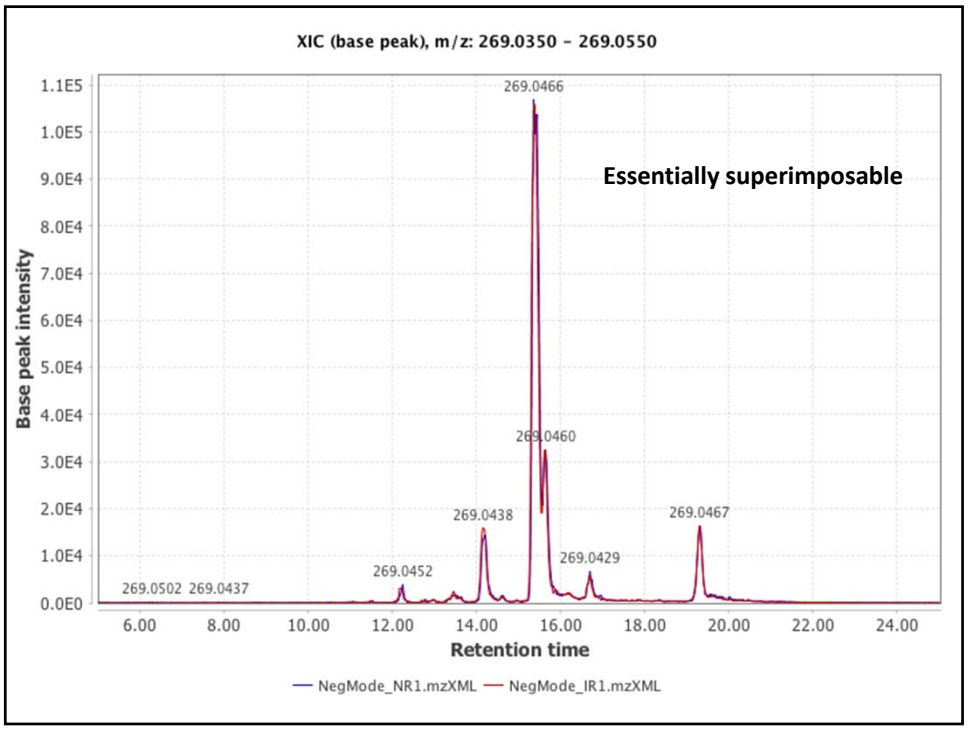
Selected mass





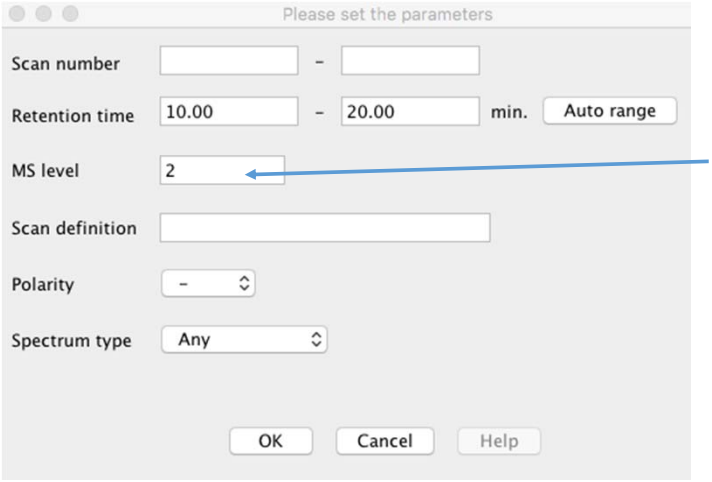
Selecting both files



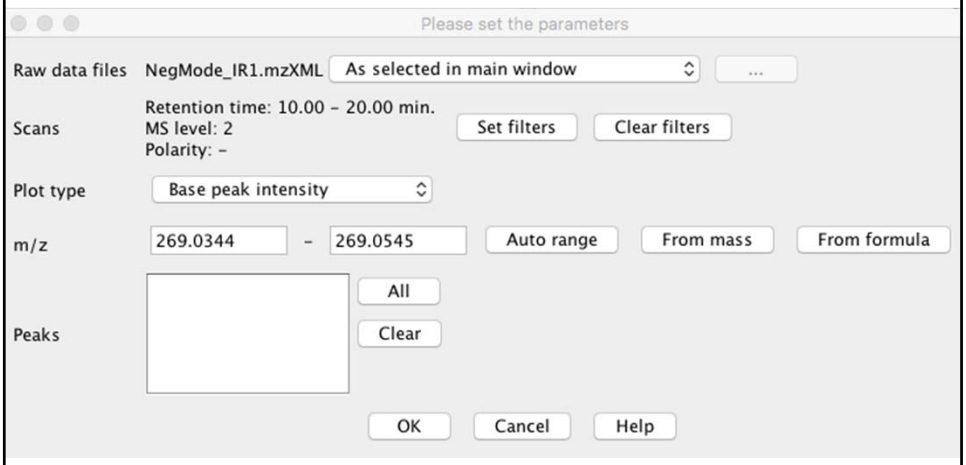


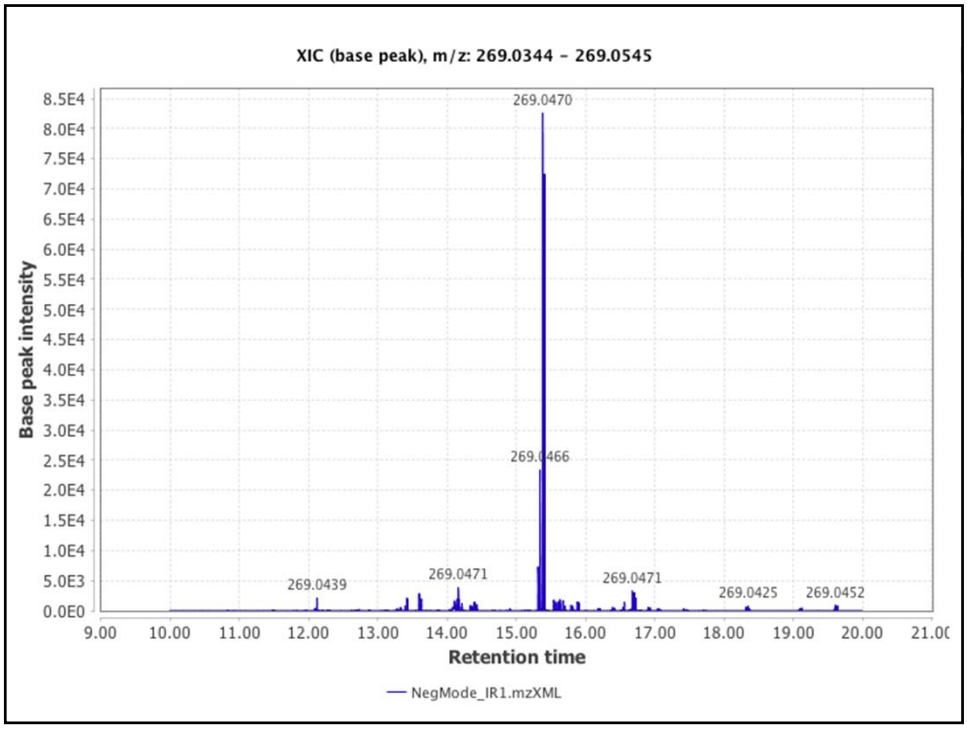
Getting MS/MS data

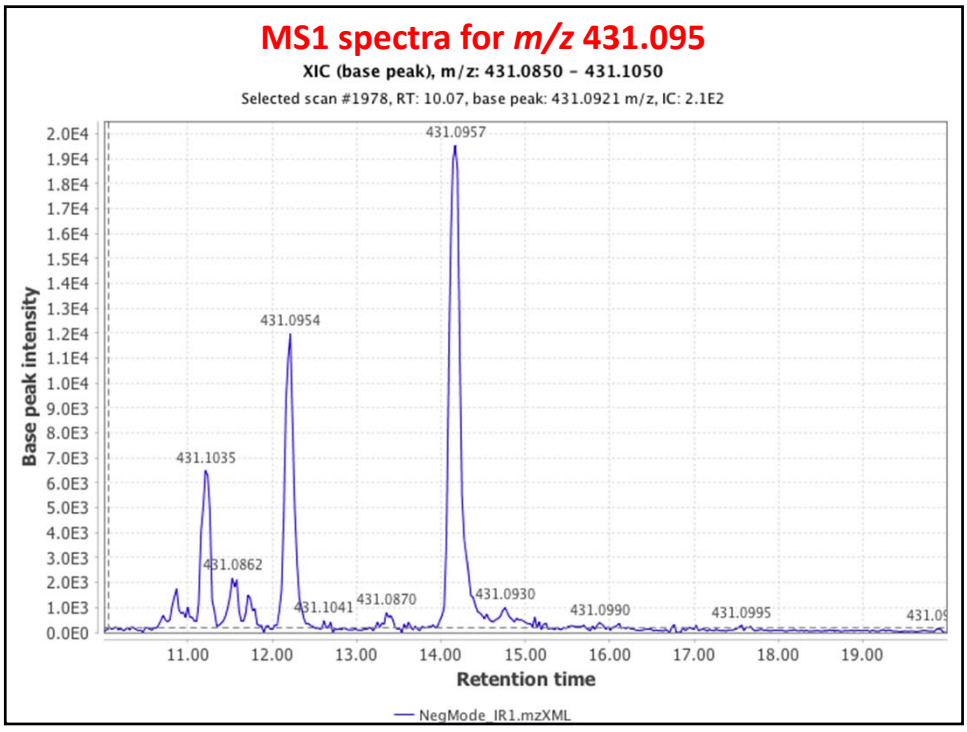
Select TIC/XIC and reset the parameters



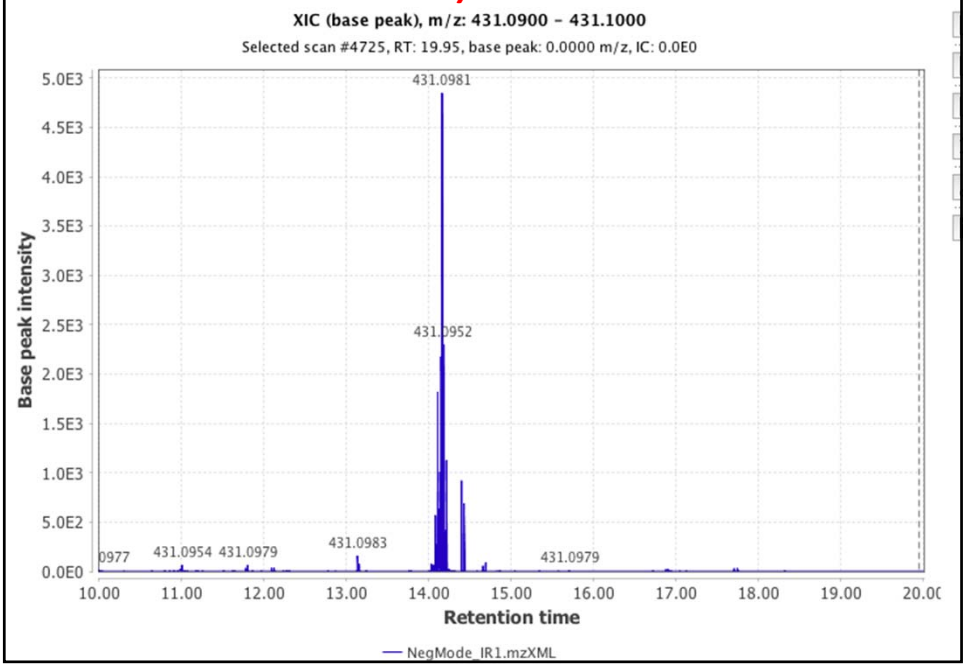
Ready to go



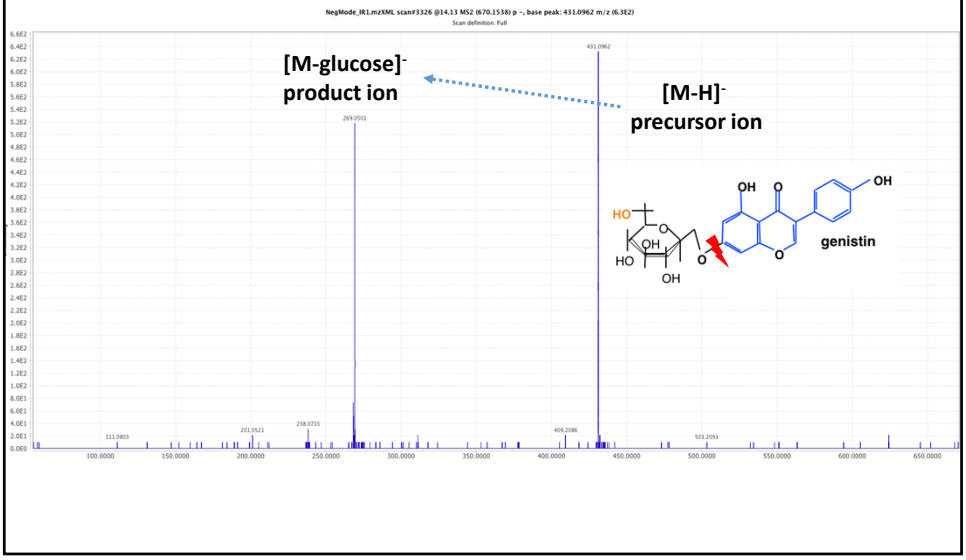

XIC (base peak), m/z: 269.0344 - 269.0545
Base peak intensity
Retention time
269.0470
269.0466
269.0439
269.0471
269.0471
269.0425
269.0452
NegMode_IR1.mzXML


MS1 spectra for m/z 431.095
XIC (base peak), m/z: 431.0850 - 431.1050
Selected scan #1978, RT: 10.07, base peak: 431.0921 m/z, IC: 2.1E2
Base peak intensity
Retention time
431.0957
431.0954
431.1035
431.0862
431.1041
431.0870
431.0930
431.0990
431.0995
NegMode_IR1.mzXML

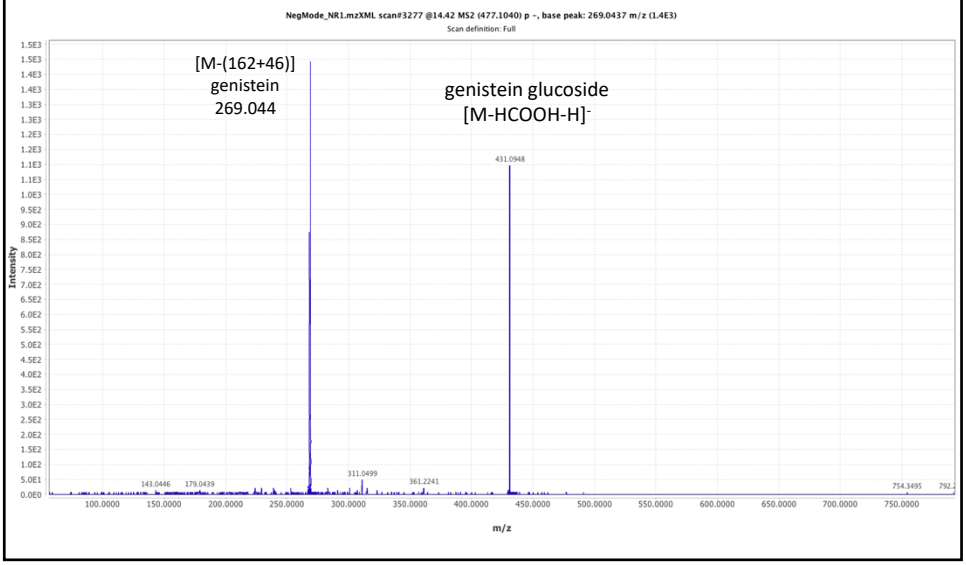
Where MSMS for *m/z* 431.095 were recorded



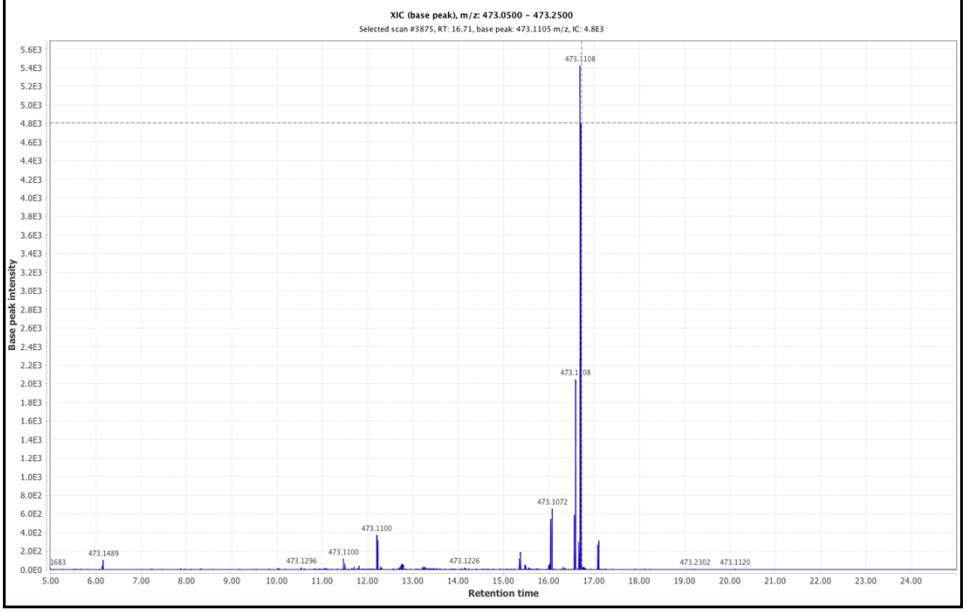
MS/MS spectrum of *m/z* 431.095



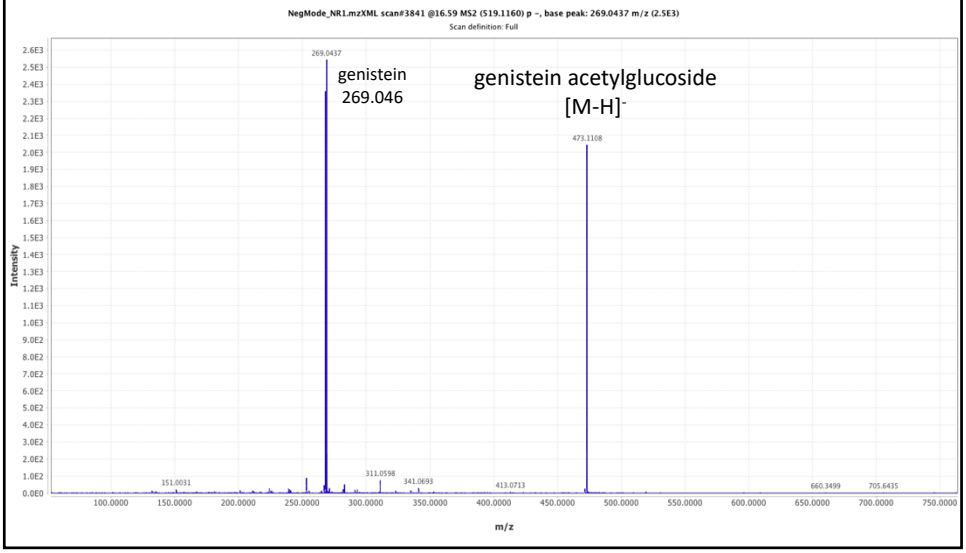
MSMS of *m/z* 477.10 – formate adduct



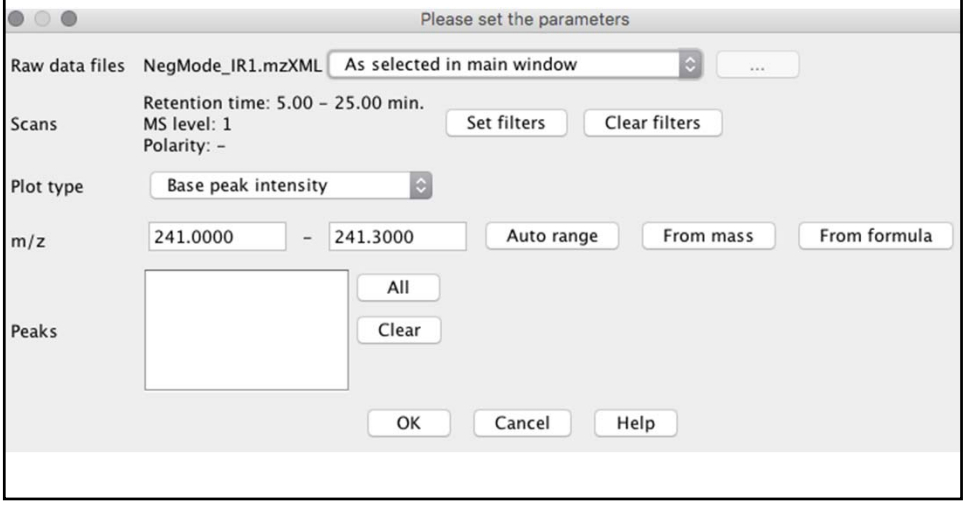
MSMS spectra *m/z* 473.11 collected in run

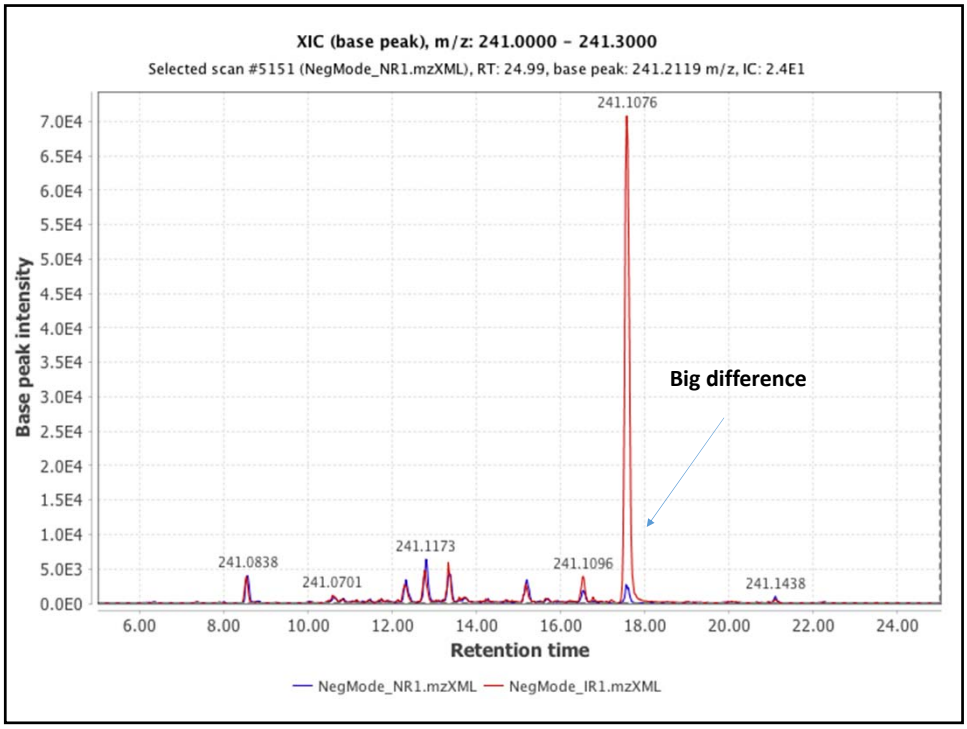


MS/MS of *m/z* 473.11

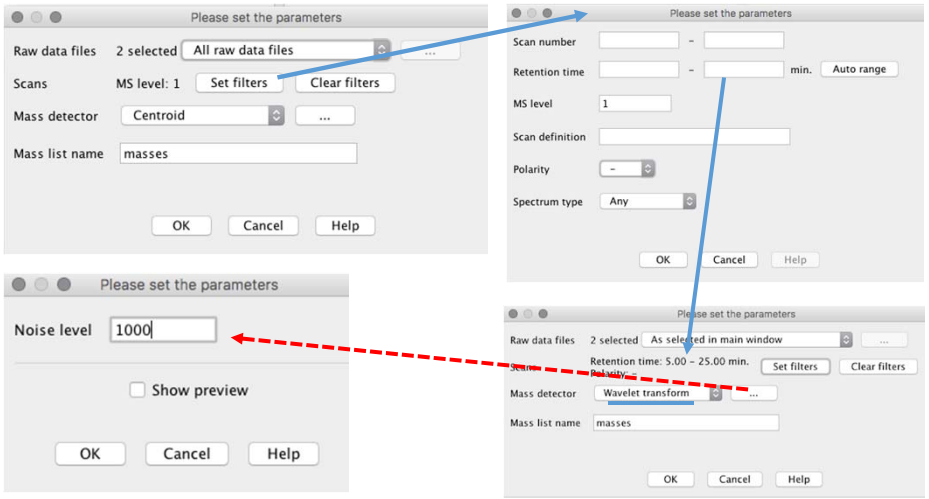


Stats analysis (coming later) reveals *m/z* 241 as having a big change

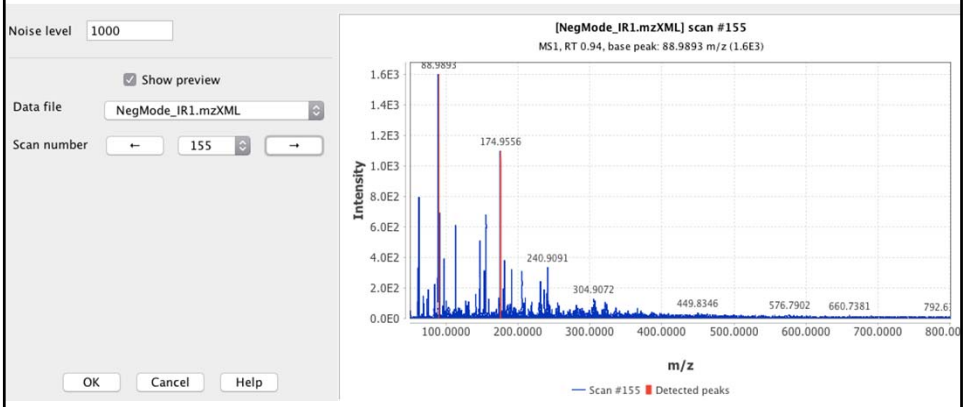




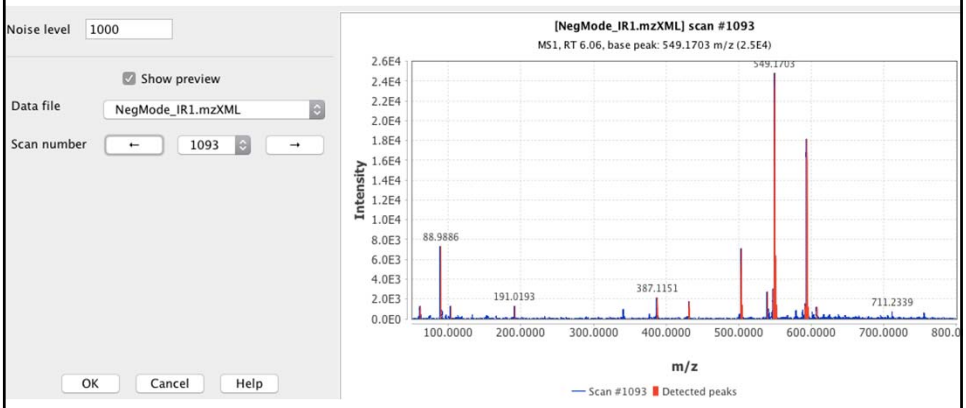
Identifying the masses in .mzxml files



Setting the background



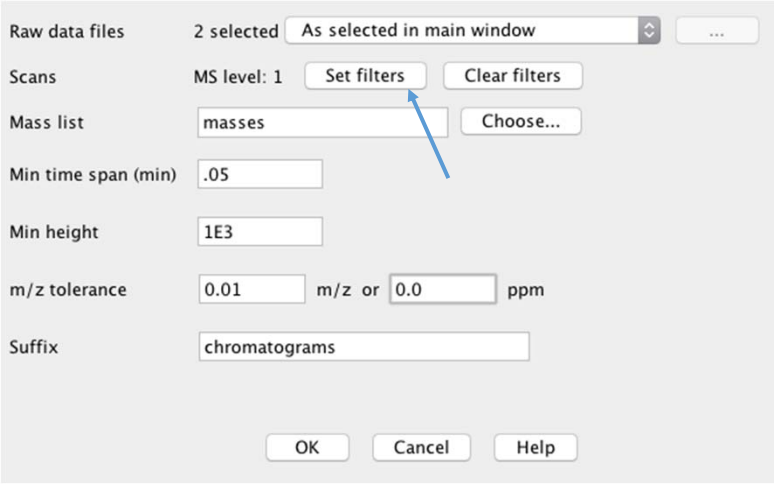
Setting the background

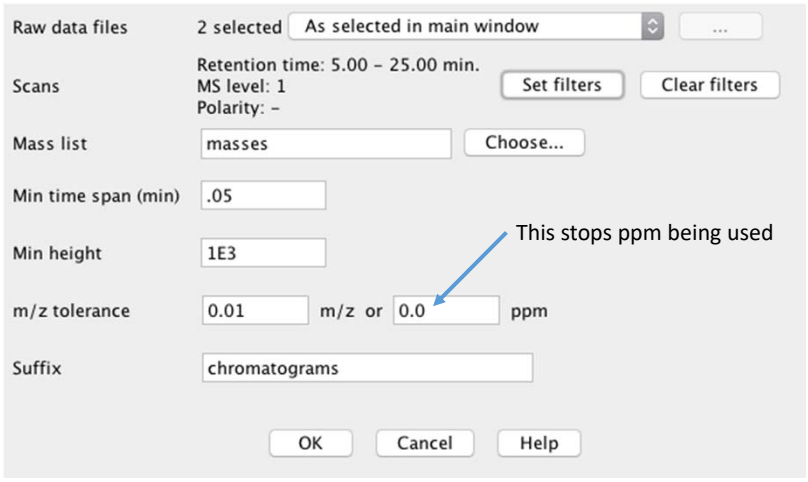


You can ignore this – it's due to the MSMS data in the file

Using the masses to create chromatograms

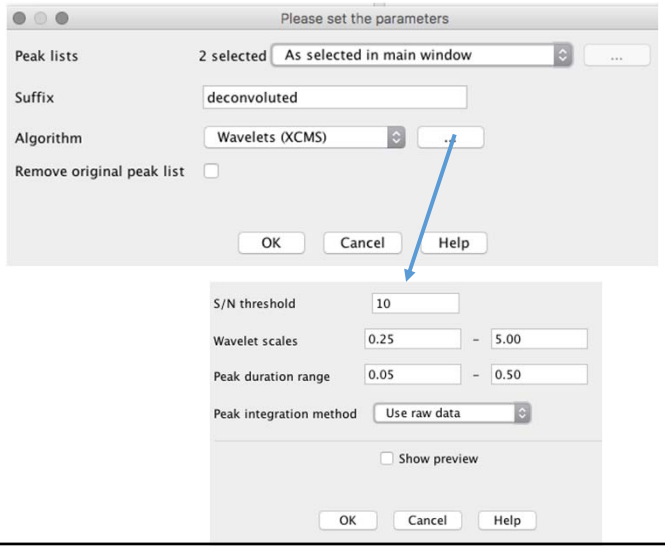
Go to *Raw data methods, peak detection, chromatogram builder*



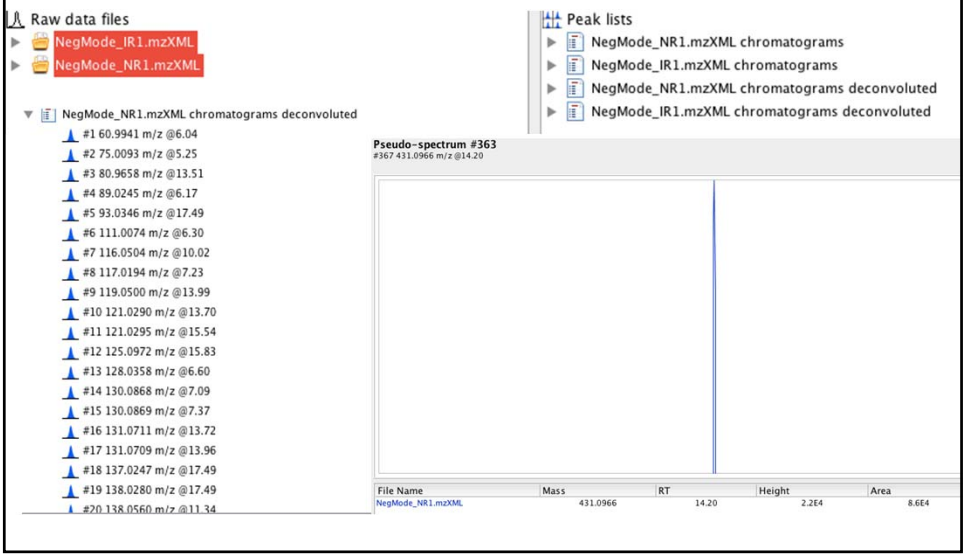


Chromatogram deconvolution

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

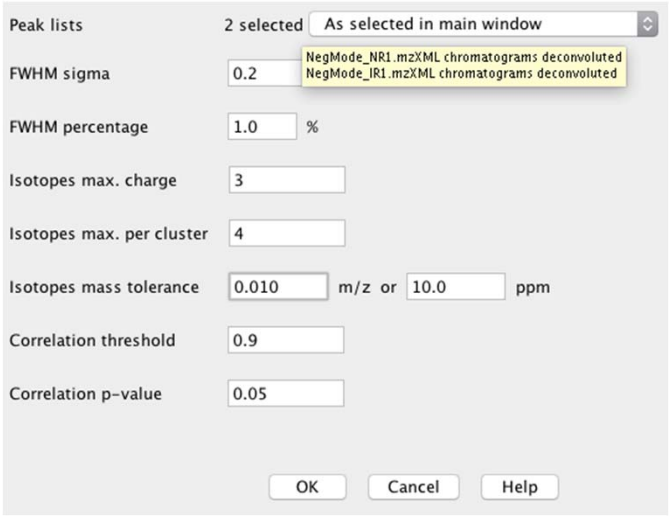


Chromatogram deconvolution output



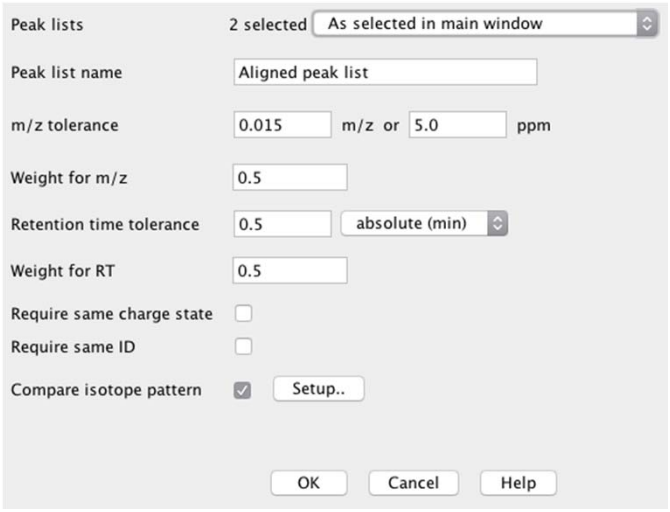
Identifying isotopes - CAMERA

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

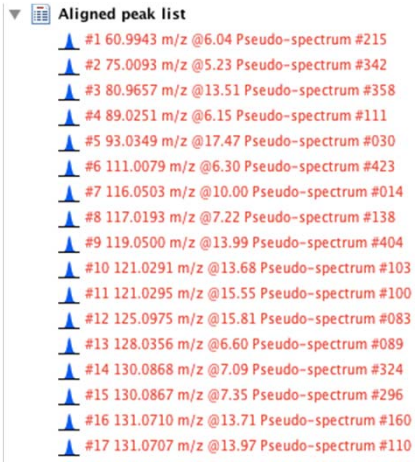


Peak and chromatogram alignment

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



Aligned data



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