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

The data sets come from this paper

PLOS ONE

RESEARCH ARTICLE

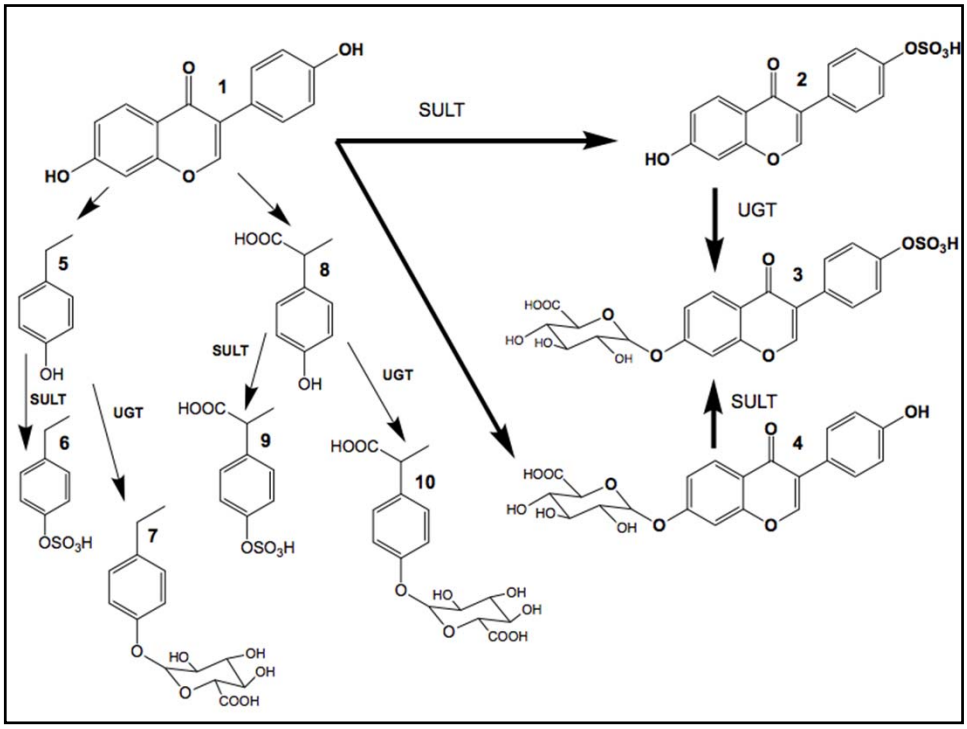
Impact of genistein on the gut microbiome of humanized mice and its role in breast tumor inhibition

Bidisha Paul[1], Kendra J. Royston[1,2], Yuanyuan Li[1,2], Matthew L. Stoll[3], Christine F. Skibola[4], Landon S. Wilson[5], Stephen Barnes[2,5,6,7,8], Casey D. Morrow[9], Trygve O. Tollefsbol[1,2,7,8,10]*

1 Department of Biology, University of Alabama at Birmingham, Birmingham, Alabama, United States of America, **2** Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, Alabama, United States of America, **3** Division of Pediatric Rheumatology, University of Alabama at Birmingham, Birmingham, Alabama, United States of America, **4** Department of Hematology and Medical Oncology, Emory University School of Medicine, Atlanta, Georgia, United States of America, **5** Targeted Metabolomics and Proteomics Laboratory, University of Alabama at Birmingham, Birmingham, Alabama, United States of America, **6** Department of Pharmacology and Toxicology, University of Alabama at Birmingham, Birmingham, Alabama, United States of America, **7** Comprehensive Center for Healthy Aging, University of Alabama at Birmingham, Birmingham, Alabama, United States of America, **8** Nutrition Obesity Research Center, University of Alabama at Birmingham, Birmingham, Alabama, United States of America, **9** Department of Cell, Development & Integrative Biology, University of Alabama at Birmingham, Birmingham, Alabama, United States of America, **10** Comprehensive Diabetes Center, University of Alabama at Birmingham, Birmingham, Alabama, United States of America

Check for updates

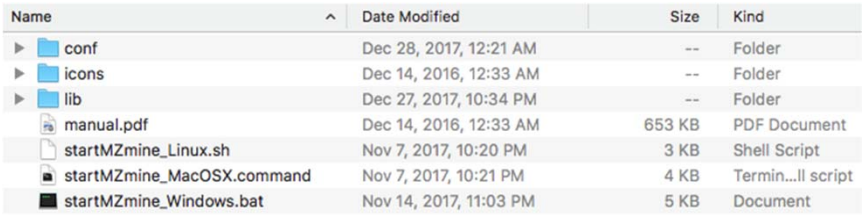
OPEN ACCESS



Download MZmine 2.30

- 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

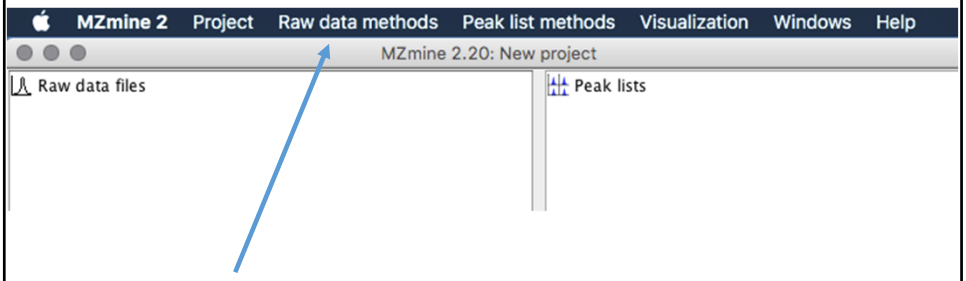
Starting point for MZmine



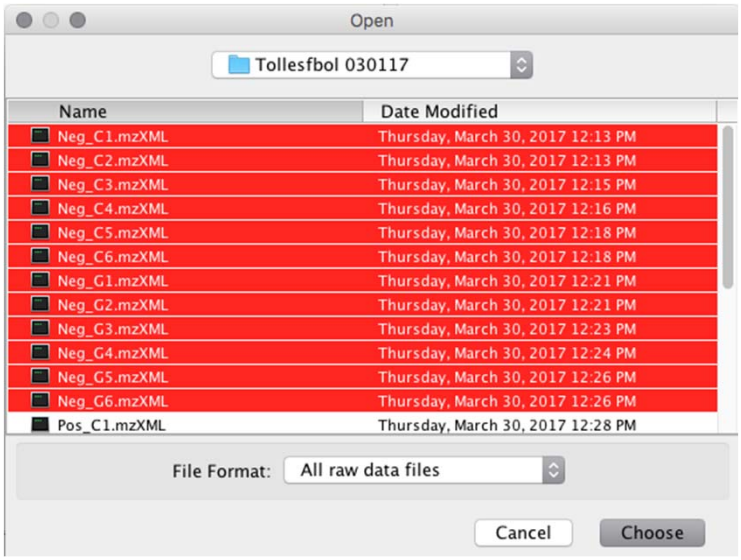
Name	Date Modified	Size	Kind
conf	Dec 28, 2017, 12:21 AM	--	Folder
icons	Dec 14, 2016, 12:33 AM	--	Folder
lib	Dec 27, 2017, 10:34 PM	--	Folder
manual.pdf	Dec 14, 2016, 12:33 AM	653 KB	PDF Document
startMZmine_Linux.sh	Nov 7, 2017, 10:20 PM	3 KB	Shell Script
startMZmine_MacOSX.command	Nov 7, 2017, 10:21 PM	4 KB	Termin...ll script
startMZmine_Windows.bat	Nov 14, 2017, 11:03 PM	5 KB	Document

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

Starting off

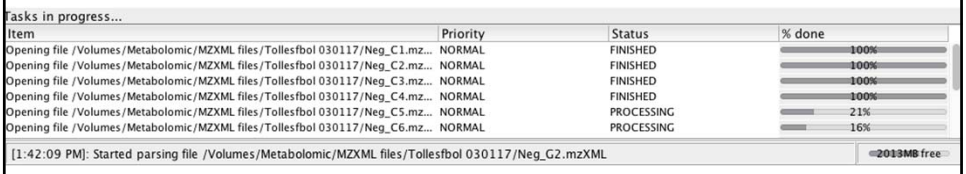


Click on Raw data methods and then select "raw data input" from the drop down box



This is where the files are on my Mac. The highlighted ones are loaded by pressing <choose>

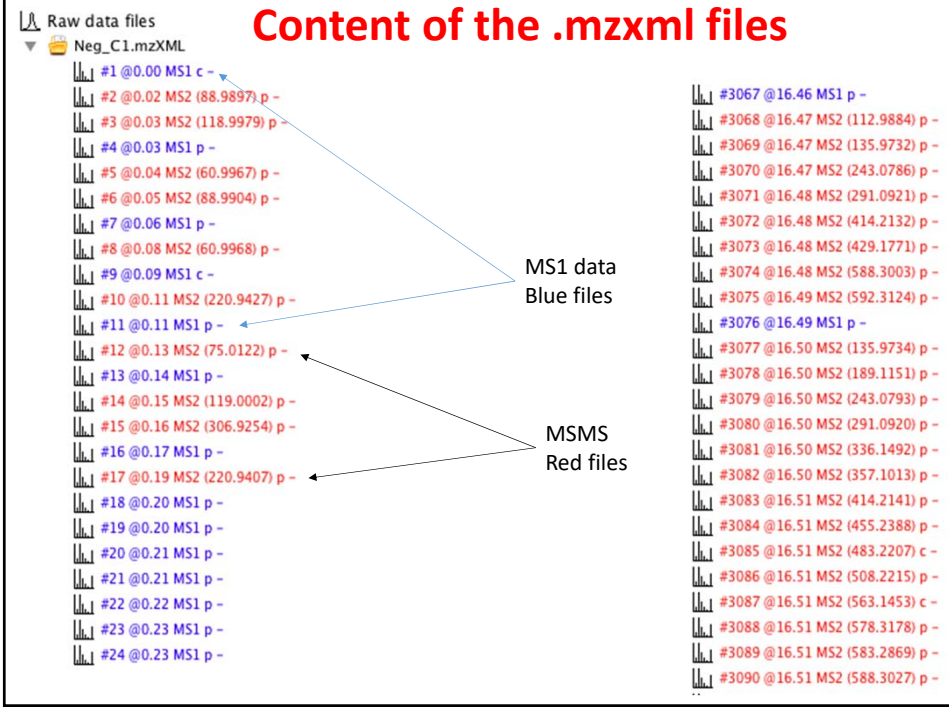
Showing uploading process



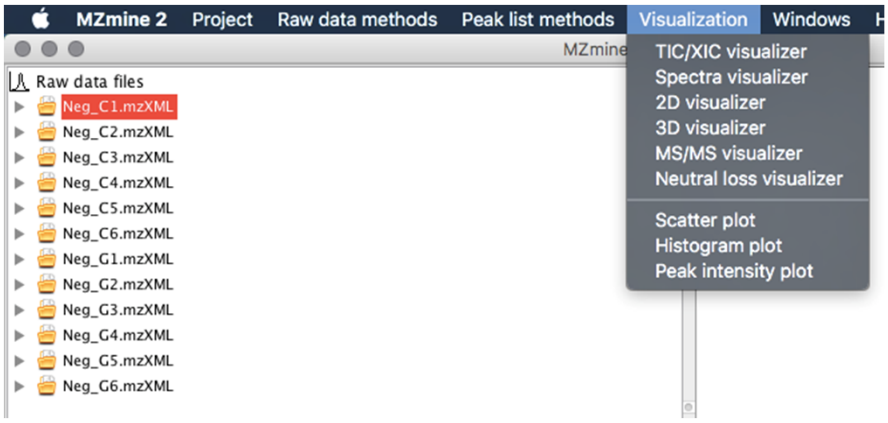
Files are now loaded

Highlight the files **Re-order the files**

Content of the .mzxml files

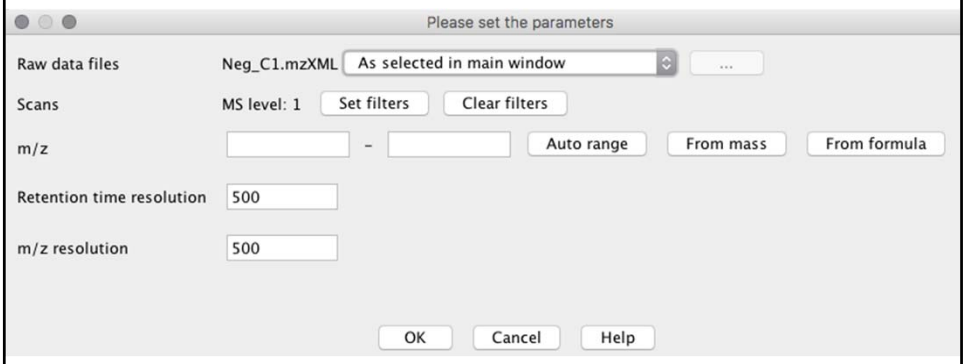


Visualization toolbar



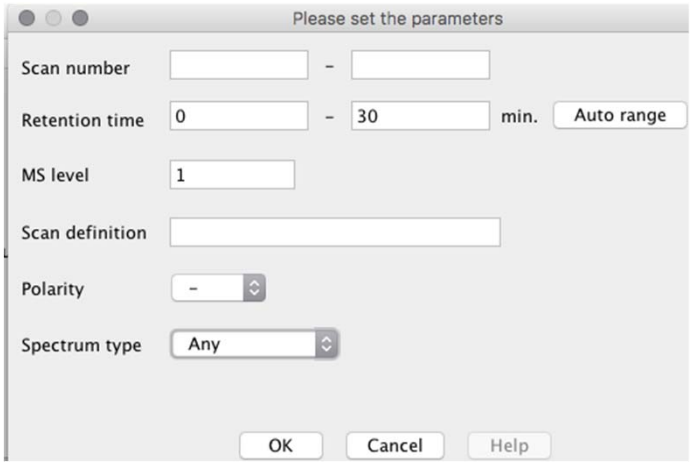
We'll start by selecting "3D visualizer" – this allows us to look at all the data

Selecting all the data

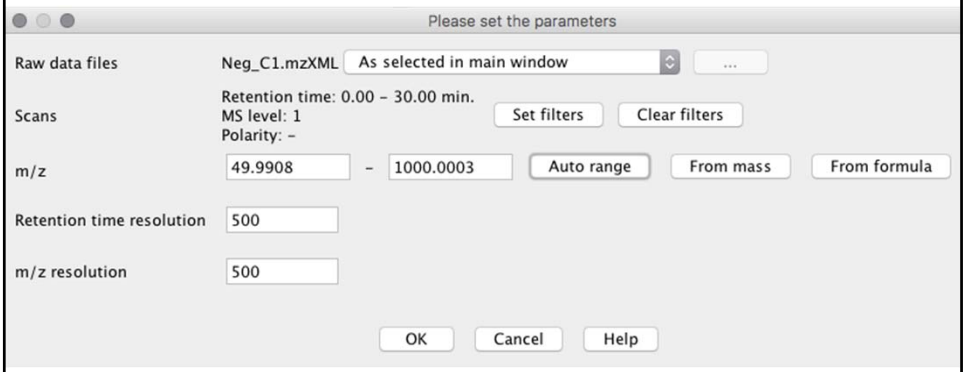


We'll select the file to open next. The 3D visualizer allows one at a time

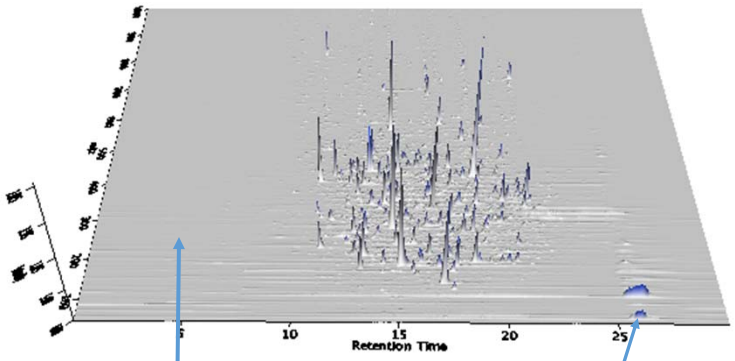
Setting 3D-parameters



Ready to view the 3D-plot



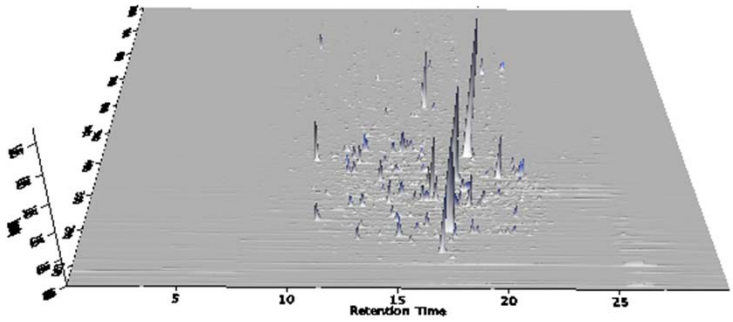
3D view of NegMode_C1



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

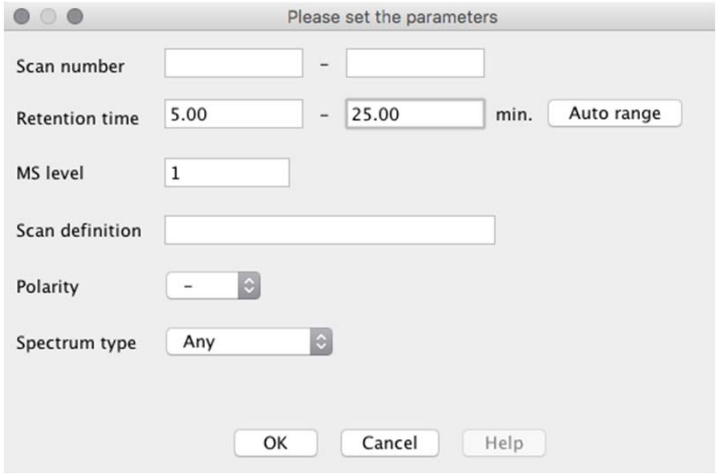
25-30 min – metabolites that stuck too hard to the column

3D view of NegMode_G1

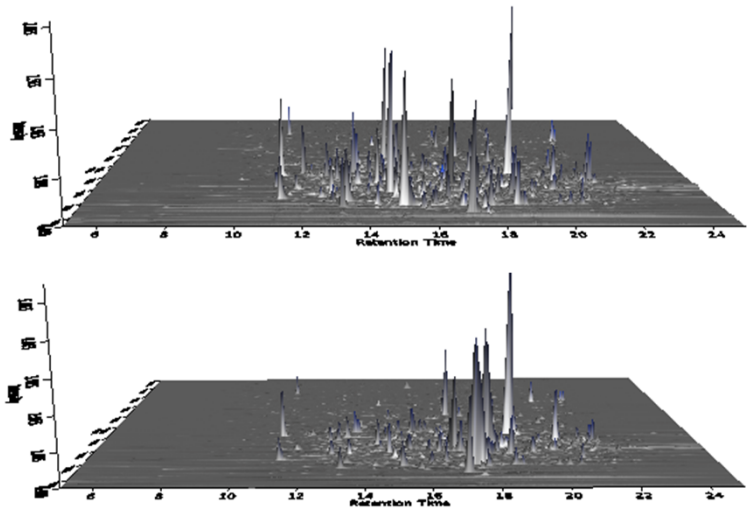


Note the large peaks – these are genistein metabolites

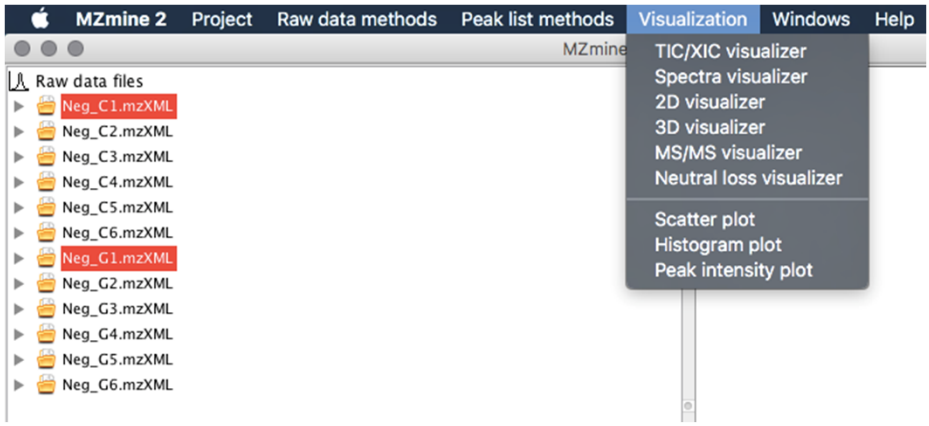
Resetting the parameters



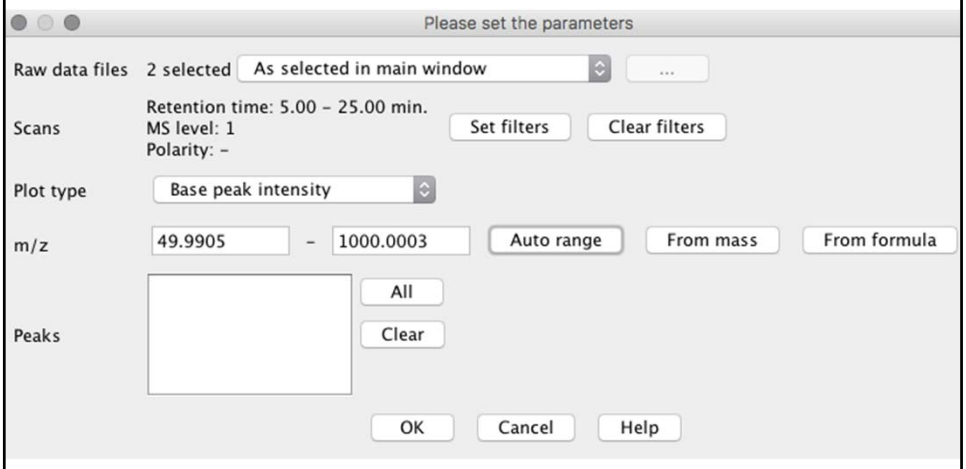
Data from 5-25 min



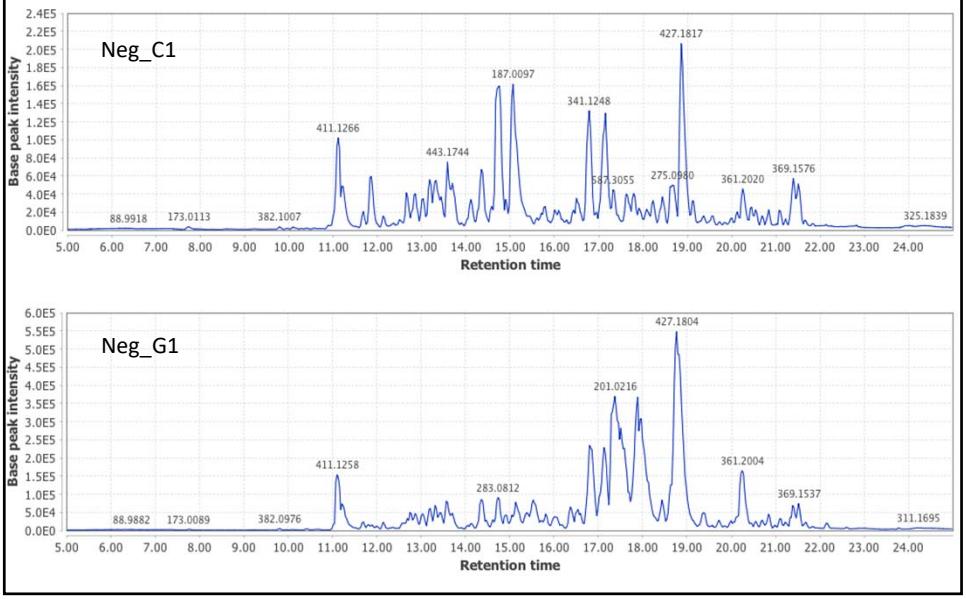
Total ion current



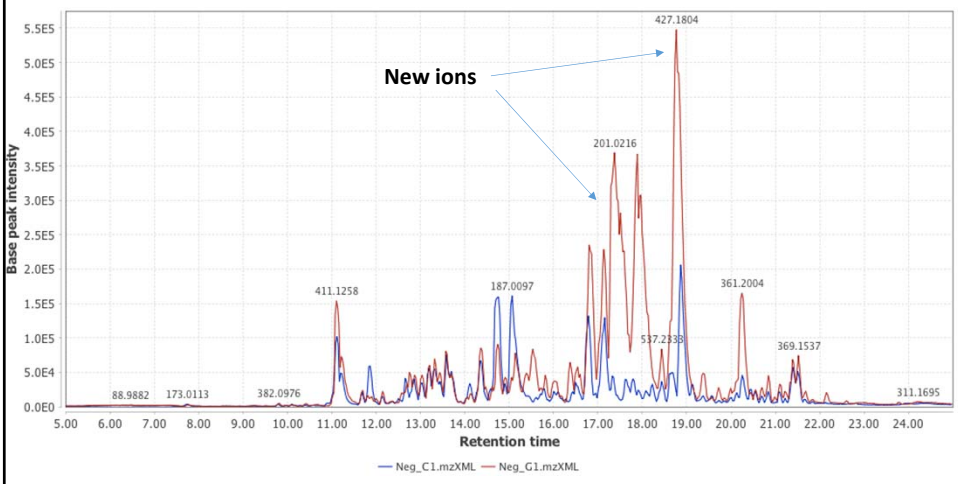
Setting the parameters



TIC of all ions from *m/z* 50-1000

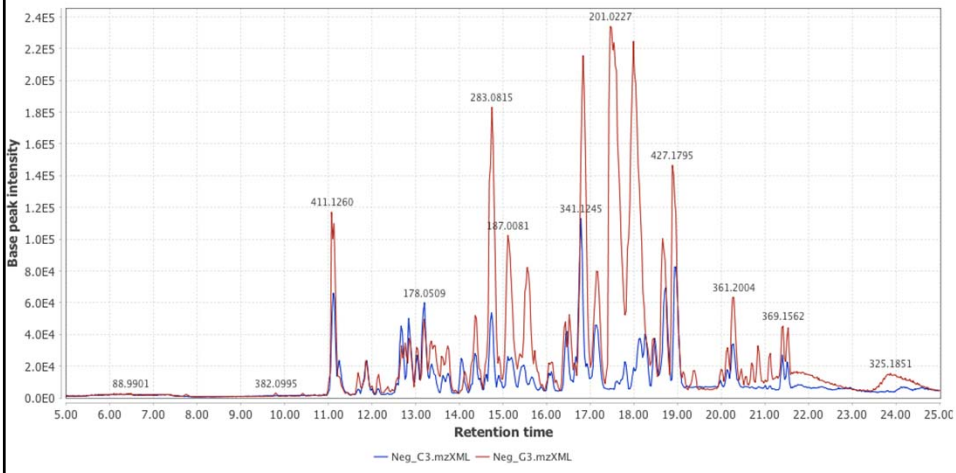


TIC *m/z* 50-1000 C1/G1 comparison



To get this, highlight the files you want to compare, before invoking the XIC module

TIC *m/z* 50-1000 C3/G3 comparison



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

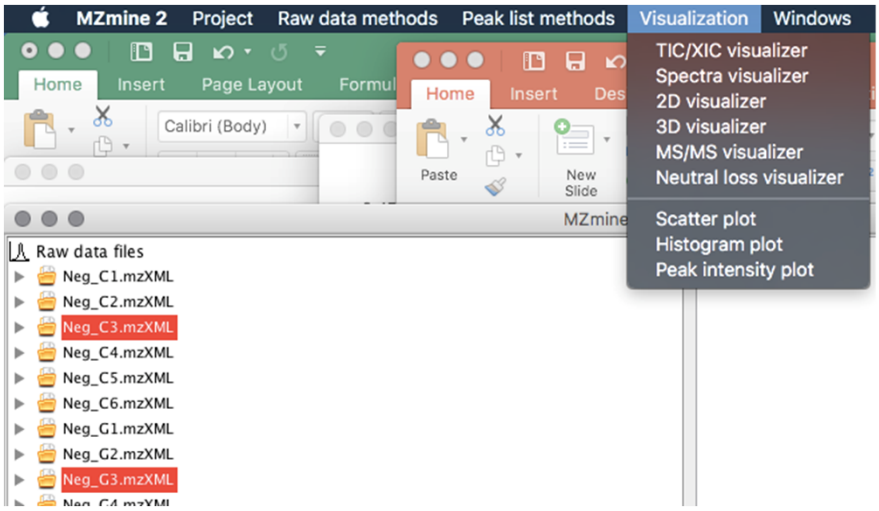
A	B	C	D	E	F
		C	H	N	O
Name	Empirical formula	12.000000000	1.007825032	14.003074004	15.9949146
hexanol	C6H14O	6	14	0	1
glucose	C6H12O6	6	12		6
genistein	C15H10O5	15	10		5

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

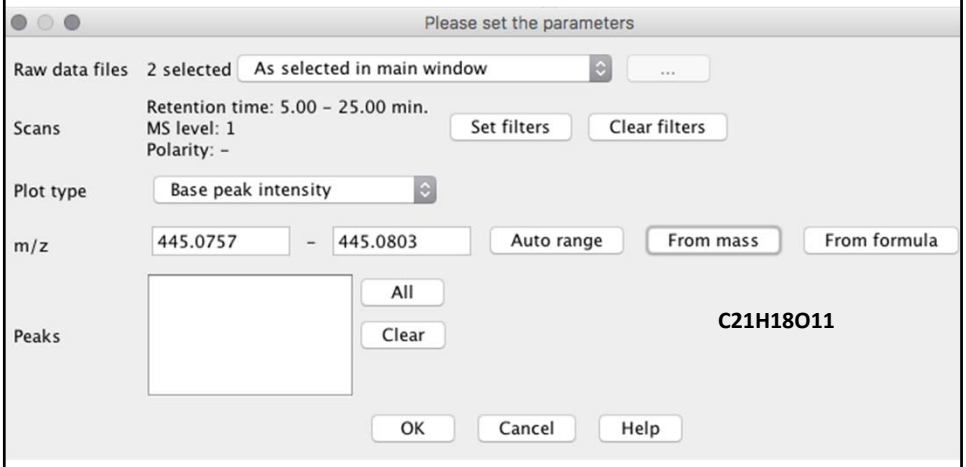
Ions of genistein and its conjugates

Name	Empirical formula	Mass (M)	[M-H]-
Genistein	$C_{15}H_{10}O_5$	270.05282	269.04557
Genistein sulfate	$C_{15}H_{10}O_8S$	350.00963	349.00238
Genistein β-glucuronide	$C_{21}H_{18}O_{11}$	446.08490	445.07765
Genistein β-glucuronide/sulfate	$C_{21}H_{18}O_{14}S$	526.04172	525.03444

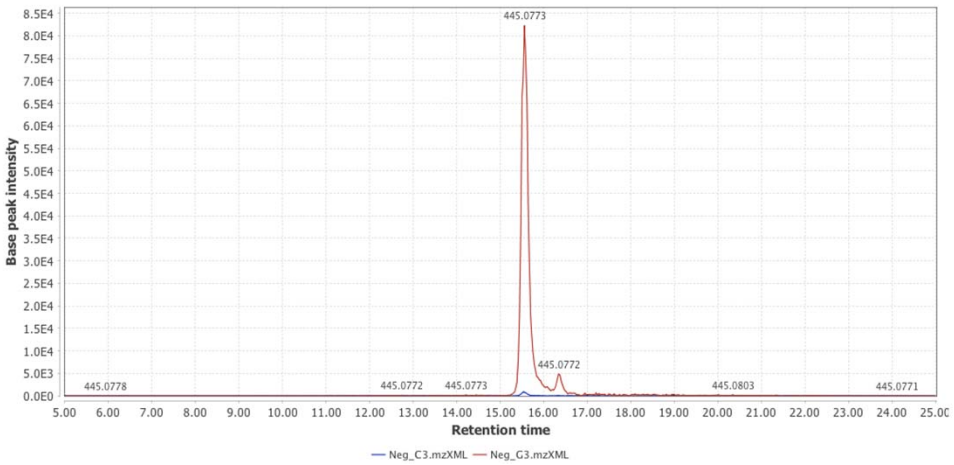
Setting the mass window



Setting the mass window

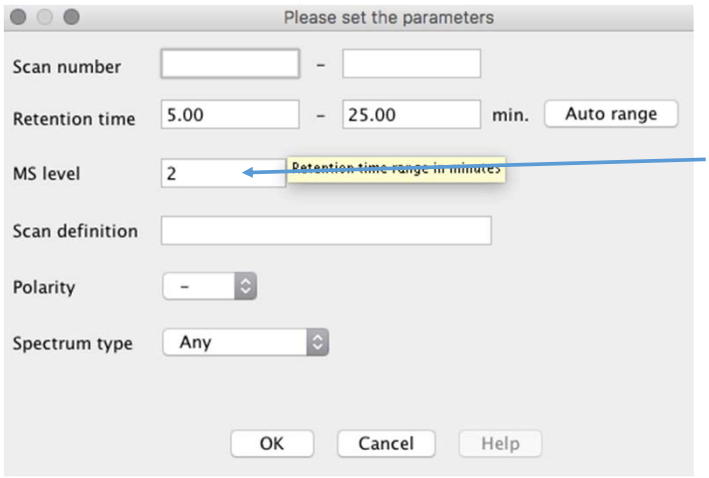


XIC of *m/z* 445.07765

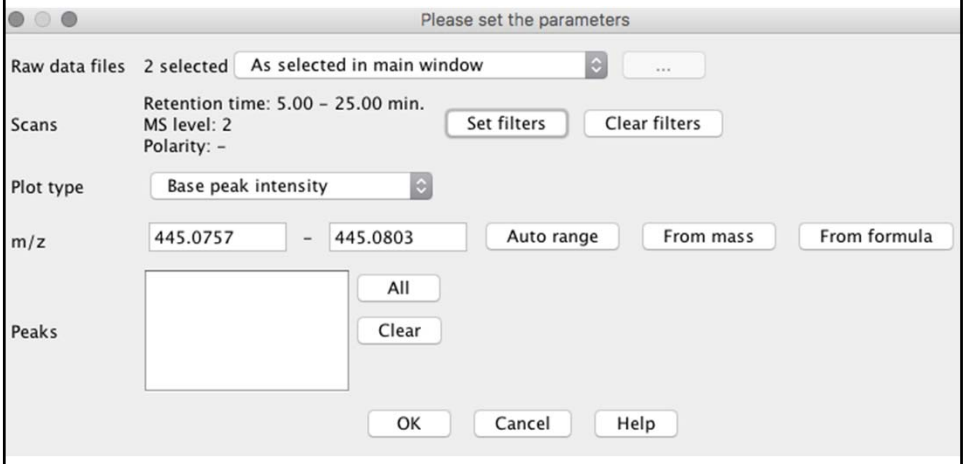


Getting MS/MS data

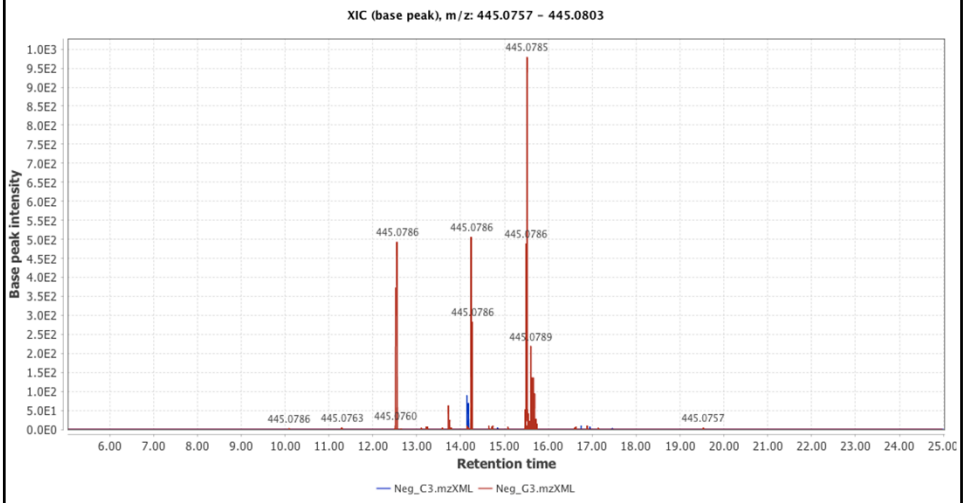
Select TIC/XIC and reset the parameters



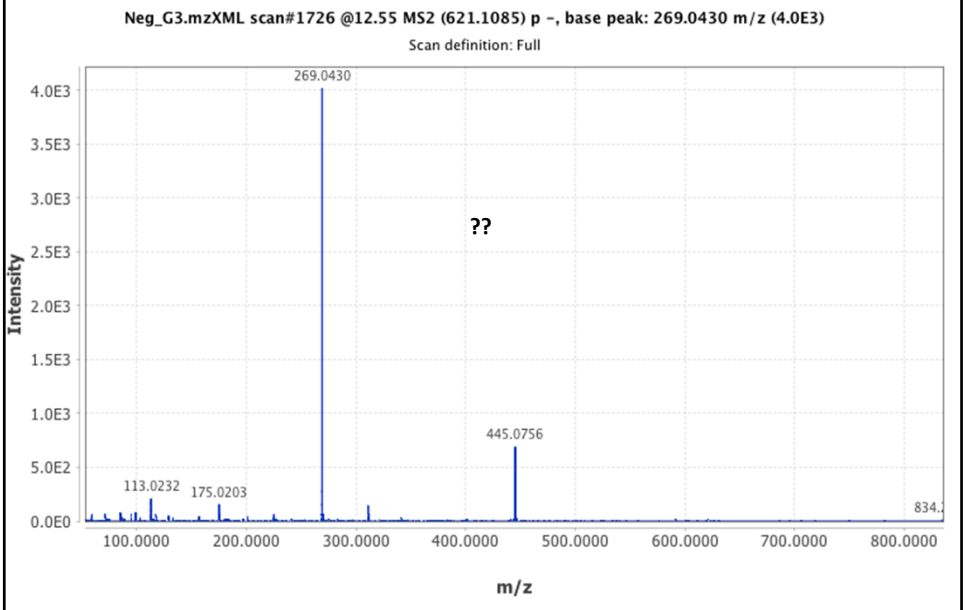
Ready to go



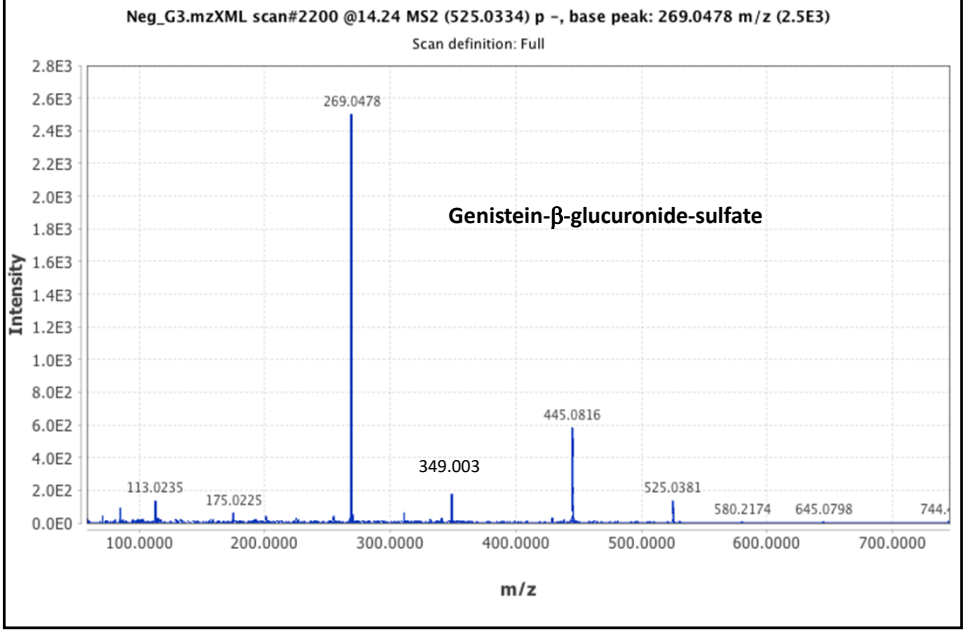
MSMS spectra captured of *m/z* 445.077



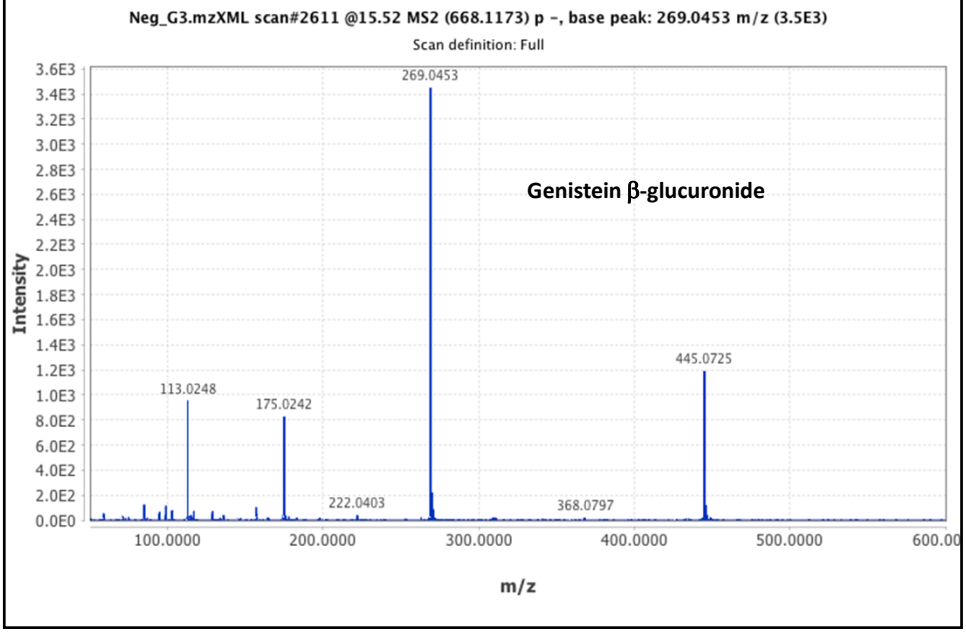
Where MSMS for *m/z* 445.079 were recorded (1)



Where MSMS for *m/z* 445.079 were recorded (2)

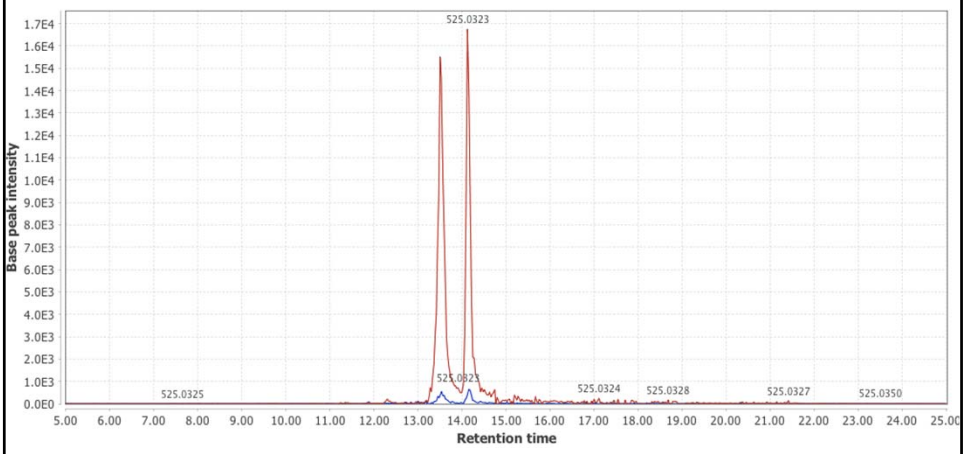


Where MSMS for *m/z* 445.079 were recorded (3)



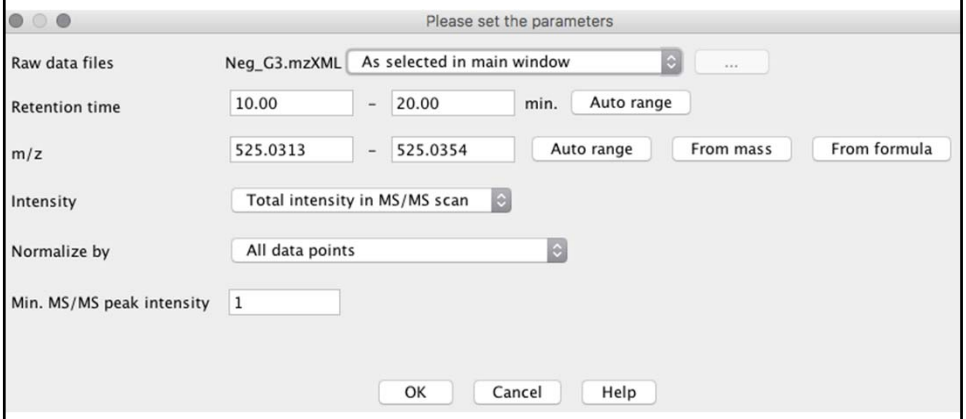
XIC of *m/z* 525.033

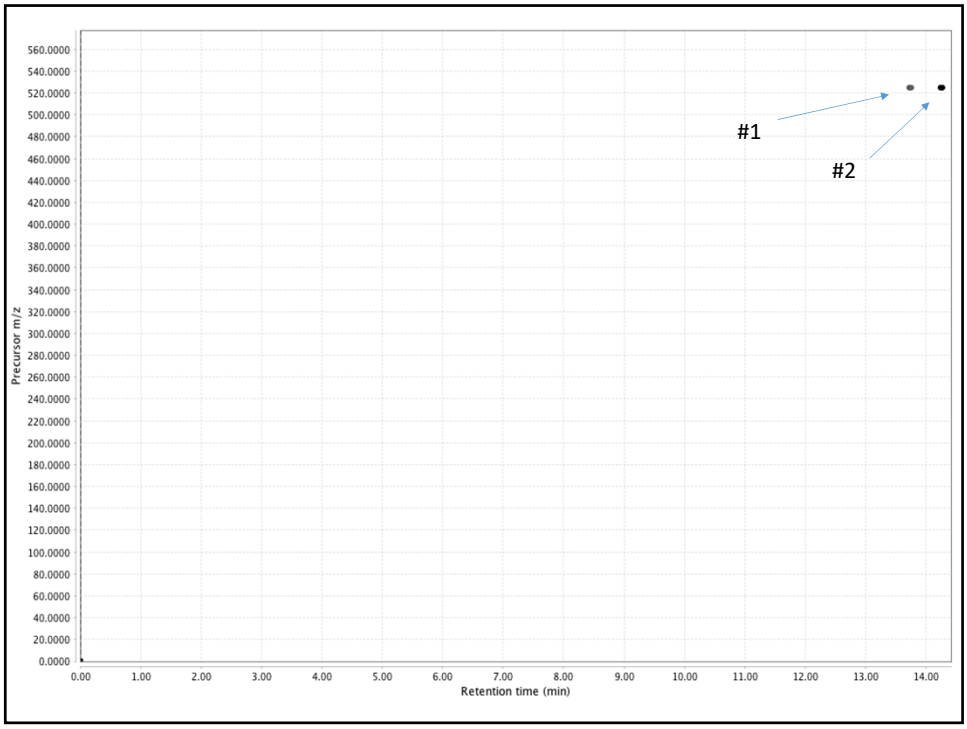
genistein β-glucuronide/sulfate

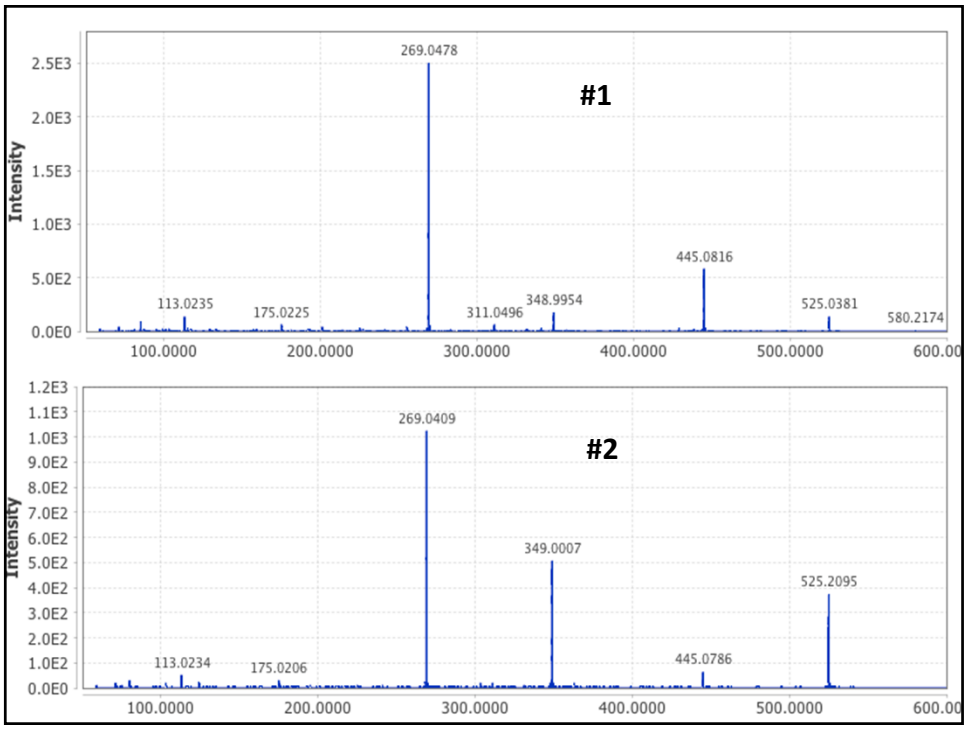


Another way to get MSMS spectra

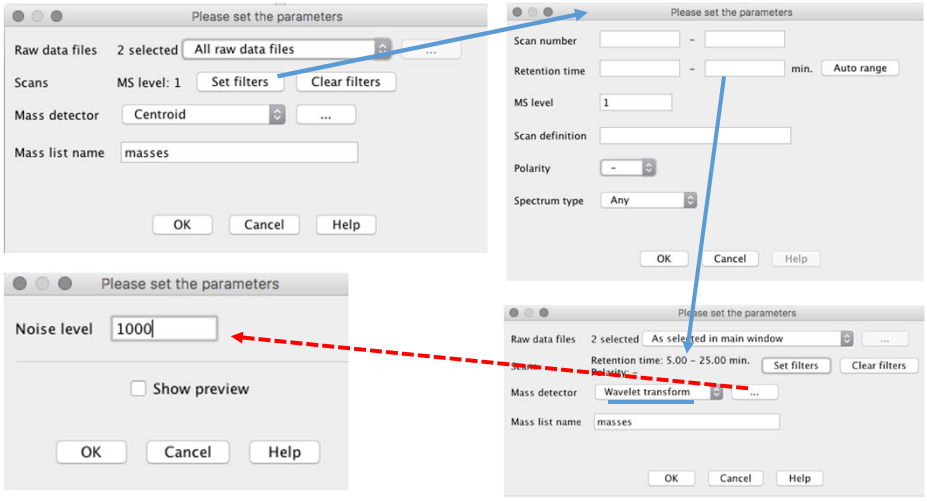
Select MSMS visualizer



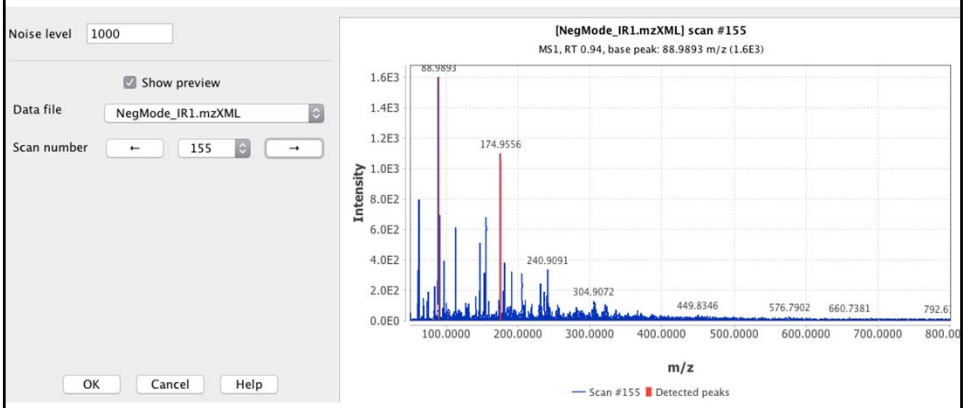

#1
#2
Precursor m/z
Retention time (min)


#1
269.0478
445.0816
113.0235
175.0225
311.0496
348.9954
525.0381
580.2174
#2
269.0409
349.0007
525.2095
113.0234
175.0206
445.0786
Intensity

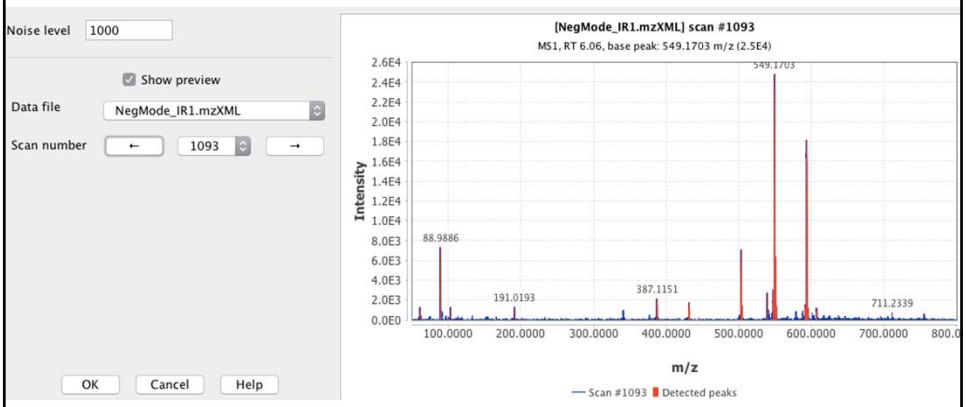
Identifying the masses in .mzxml files



Setting the background



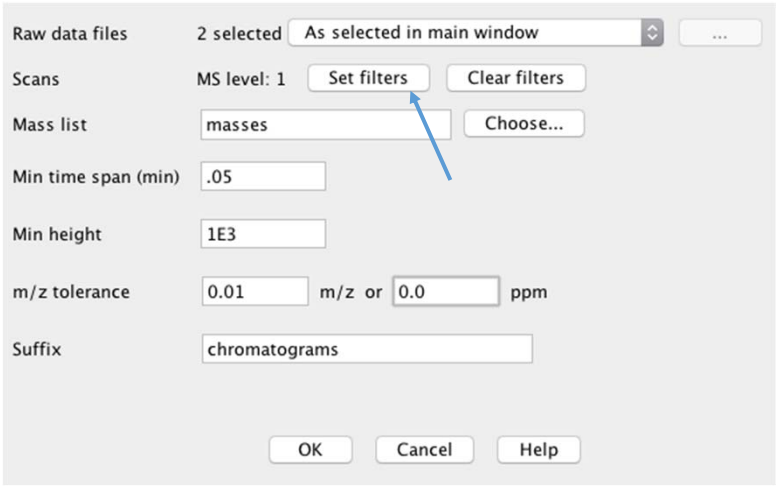
Setting the background

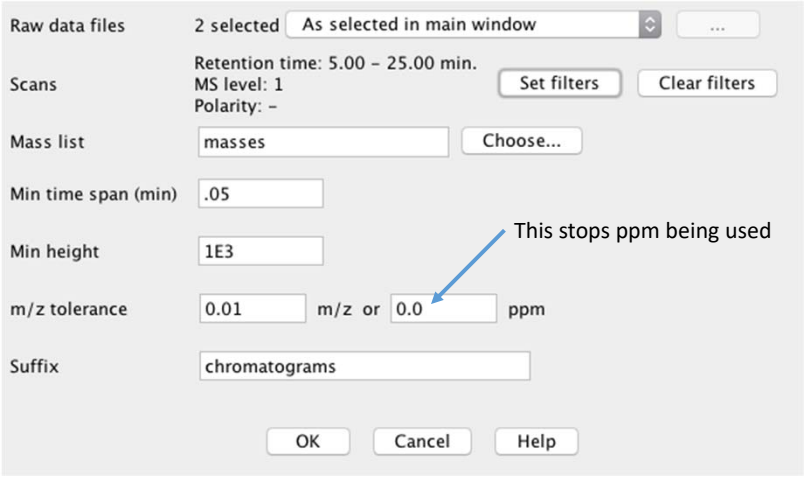


You can ignore this – it's due to the MSMS data in the file

Using the masses to create chromatograms

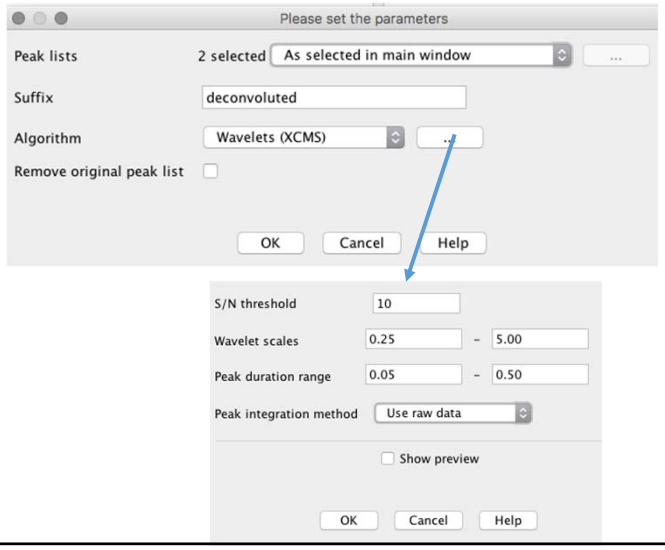
Go to *Raw data methods, peak detection, chromatogram builder*



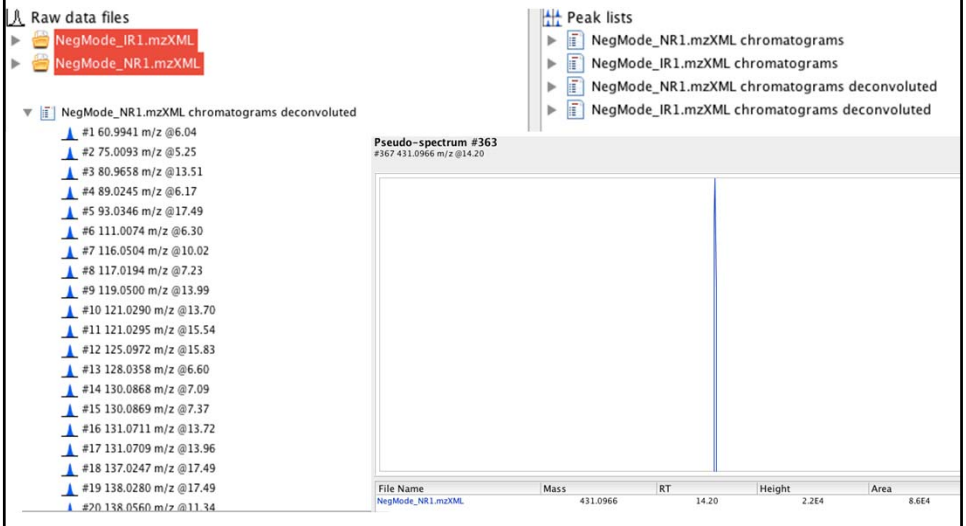


Chromatogram deconvolution

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

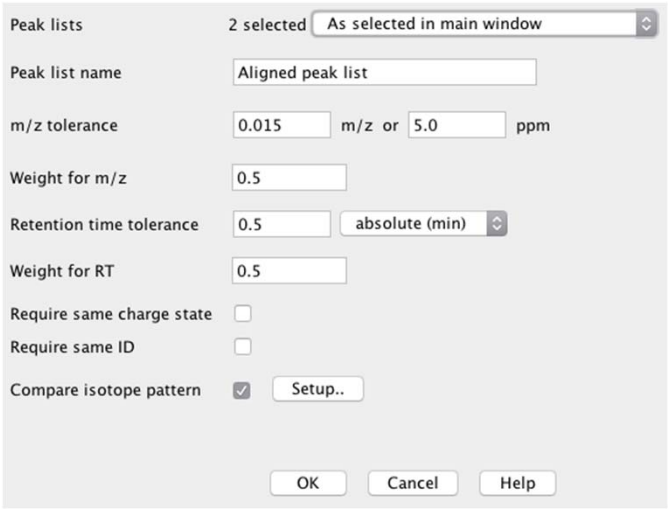


Chromatogram deconvolution output

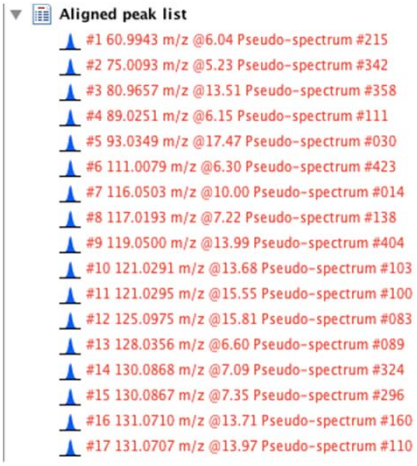


Peak and chromatogram alignment

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 C1-C6 and G1-GS .mzxml files
- Locate the ions that have the genistein ion (in negative - what is its *m/z* value?)
  - Get MSMS spectra of each one
- Identify all the *masses* in each file (see slide 39) – from these generate chromatograms, and then deconvolute the chromatograms
- Output the data into a .csv file (choose row ID, m/z, retention time, peak height, peak area and FWHM)
- Sort the file by retention time – identify ions that are co-eluting and are isotopes.