


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





RESEARCH ARTICLE

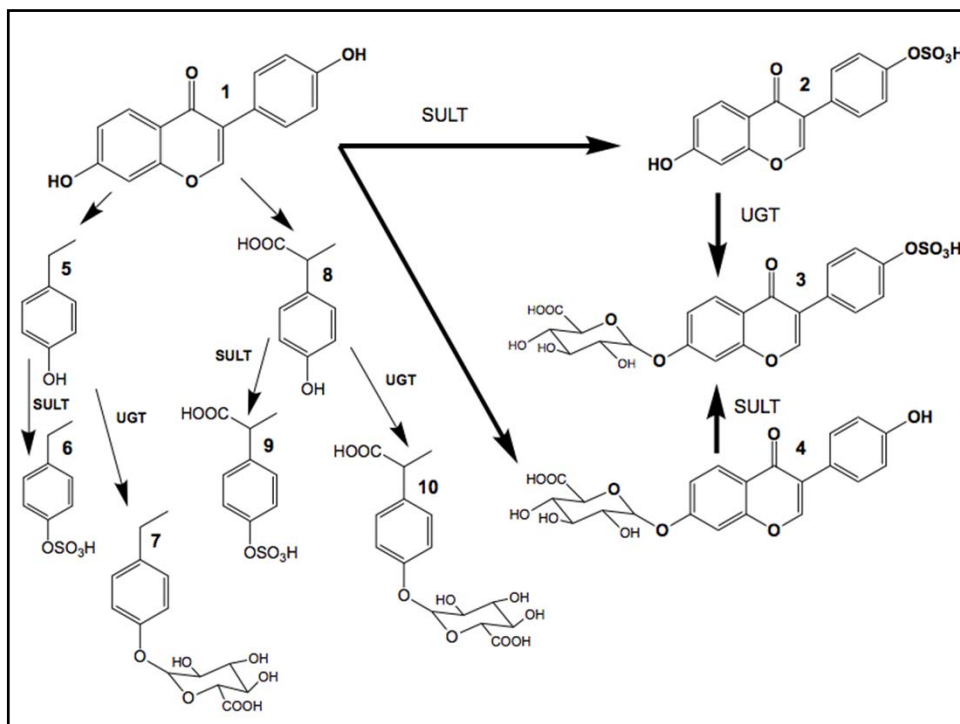
Impact of genistein on the gut microbiome of humanized mice and its role in breast tumor inhibition

Bidisha Paul¹, Kendra J. Royston^{1,2}, Yuanyuan Li^{1,2}, Matthew L. Stoll³, Christine F. Skibola⁴, Landon S. Wilson⁵, Stephen Barnes^{2,5,6,7,8}, Casey D. Morrow⁹, 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

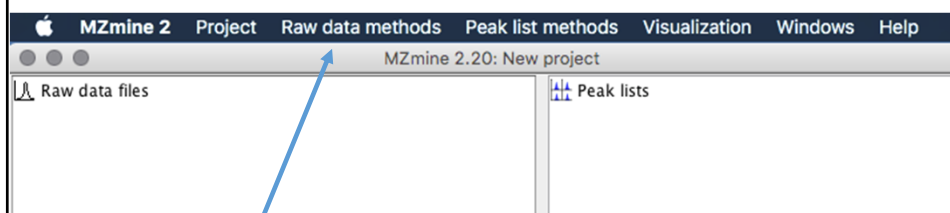
- Go to <http://mzmine.github.io/>
- Download
- Unzip the file and move the folder into Applications
 - There are three starting methods
 - Linux - startMZmine_Linux.sh
 - Mac - startMZmine_MacOSX.command
 - Windows - startMZmine_Windows.bat
 - Double click to start the program

Starting point for MZmine

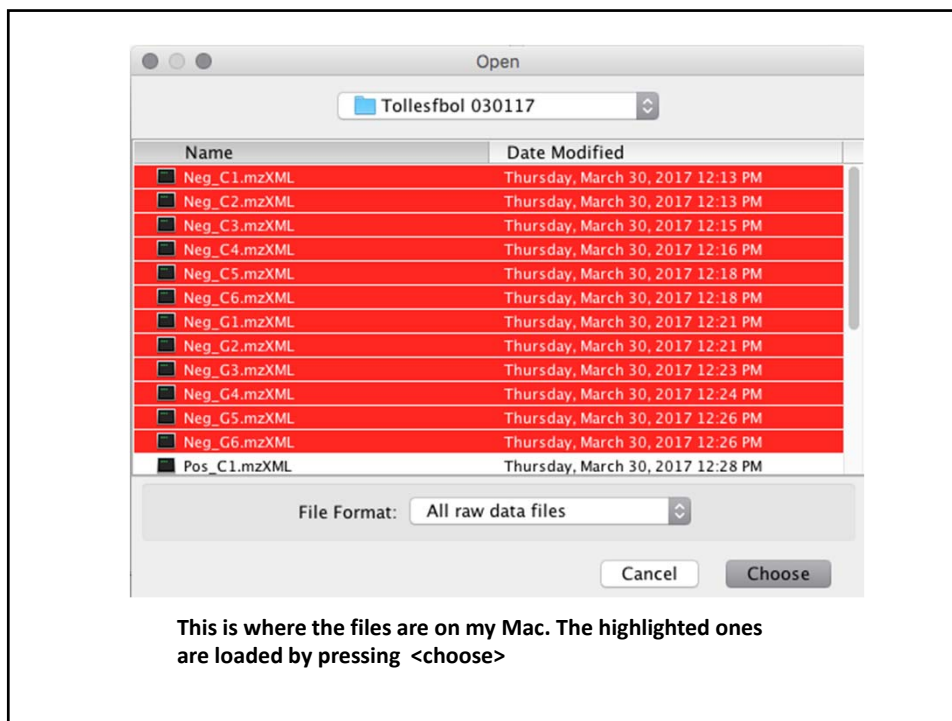
Name	Date Modified	Size	Kind
▶ conf	Dec 28, 2017, 12:21 AM	--	Folder
▶ icons	Dec 14, 2016, 12:33 AM	--	Folder
▶ lib	Dec 27, 2017, 10:34 PM	--	Folder
manual.pdf	Dec 14, 2016, 12:33 AM	653 KB	PDF Document
startMZmine_Linux.sh	Nov 7, 2017, 10:20 PM	3 KB	Shell Script
startMZmine_MacOSX.command	Nov 7, 2017,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



Showing uploading process

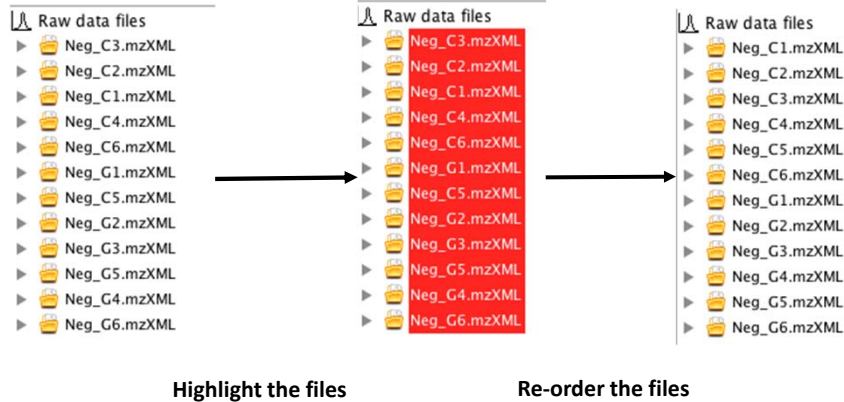
Tasks in progress...

Item	Priority	Status	% done
Opening file /Volumes/Metabolomic/MZXML files/Tollesfbol 030117/Neg_C1.mz...	NORMAL	FINISHED	100%
Opening file /Volumes/Metabolomic/MZXML files/Tollesfbol 030117/Neg_C2.mz...	NORMAL	FINISHED	100%
Opening file /Volumes/Metabolomic/MZXML files/Tollesfbol 030117/Neg_C3.mz...	NORMAL	FINISHED	100%
Opening file /Volumes/Metabolomic/MZXML files/Tollesfbol 030117/Neg_C4.mz...	NORMAL	FINISHED	100%
Opening file /Volumes/Metabolomic/MZXML files/Tollesfbol 030117/Neg_C5.mz...	NORMAL	PROCESSING	21%
Opening file /Volumes/Metabolomic/MZXML files/Tollesfbol 030117/Neg_C6.mz...	NORMAL	PROCESSING	16%

[1:42:09 PM]: Started parsing file /Volumes/Metabolomic/MZXML files/Tollesfbol 030117/Neg_G2.mzXML

©2013MB free

Files are now loaded



Content of the .mzxml files

Raw data files

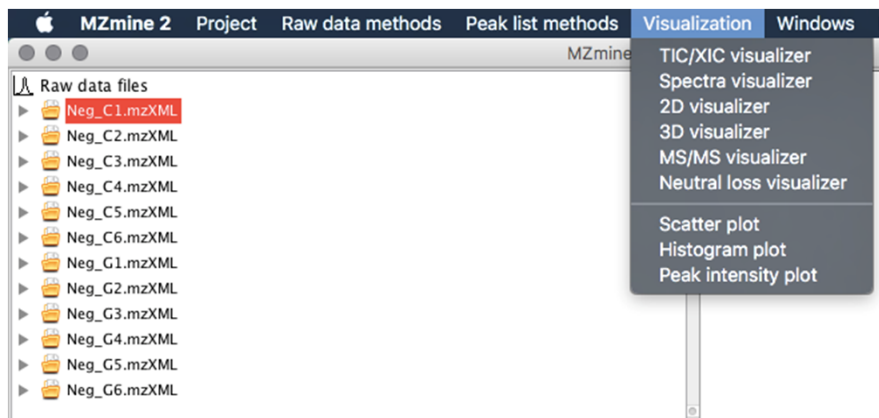
- ▼ Neg_C1.mzXML
 - #1 @0.00 MS1 c -
 - #2 @0.02 MS2 (88.9897) p -
 - #3 @0.03 MS2 (118.9979) p -
 - #4 @0.03 MS1 p -
 - #5 @0.04 MS2 (60.9967) p -
 - #6 @0.05 MS2 (88.9904) p -
 - #7 @0.06 MS1 p -
 - #8 @0.08 MS2 (60.9968) p -
 - #9 @0.09 MS1 c -
 - #10 @0.11 MS2 (220.9427) p -
 - #11 @0.11 MS1 p -
 - #12 @0.13 MS2 (75.0122) p -
 - #13 @0.14 MS1 p -
 - #14 @0.15 MS2 (119.0002) p -
 - #15 @0.16 MS2 (306.9254) p -
 - #16 @0.17 MS1 p -
 - #17 @0.19 MS2 (220.9407) p -
 - #18 @0.20 MS1 p -
 - #19 @0.20 MS1 p -
 - #20 @0.21 MS1 p -
 - #21 @0.21 MS1 p -
 - #22 @0.22 MS1 p -
 - #23 @0.23 MS1 p -
 - #24 @0.23 MS1 p -

MS1 data
Blue files

MSMS
Red files

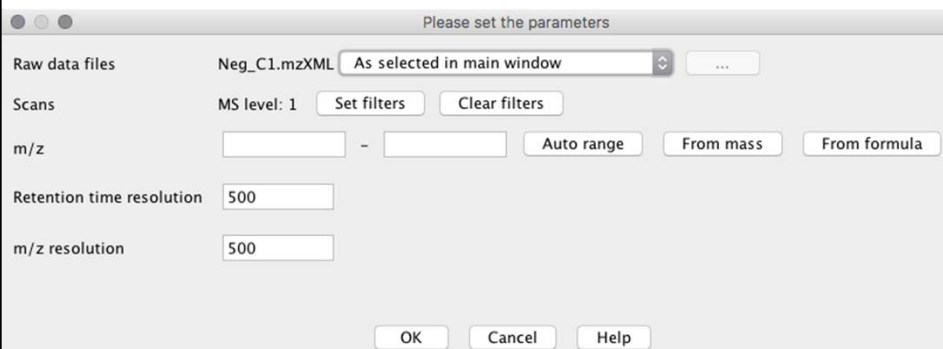
- #3067 @16.46 MS1 p -
- #3068 @16.47 MS2 (112.9884) p -
- #3069 @16.47 MS2 (135.9732) p -
- #3070 @16.47 MS2 (243.0786) p -
- #3071 @16.48 MS2 (291.0921) p -
- #3072 @16.48 MS2 (414.2132) p -
- #3073 @16.48 MS2 (429.1771) p -
- #3074 @16.48 MS2 (588.3003) p -
- #3075 @16.49 MS2 (592.3124) p -
- #3076 @16.49 MS1 p -
- #3077 @16.50 MS2 (135.9734) p -
- #3078 @16.50 MS2 (189.1151) p -
- #3079 @16.50 MS2 (243.0793) p -
- #3080 @16.50 MS2 (291.0920) p -
- #3081 @16.50 MS2 (336.1492) p -
- #3082 @16.50 MS2 (357.1013) p -
- #3083 @16.51 MS2 (414.2141) p -
- #3084 @16.51 MS2 (455.2388) p -
- #3085 @16.51 MS2 (483.2207) c -
- #3086 @16.51 MS2 (508.2215) p -
- #3087 @16.51 MS2 (563.1453) c -
- #3088 @16.51 MS2 (578.3178) p -
- #3089 @16.51 MS2 (583.2869) p -
- #3090 @16.51 MS2 (588.3027) p -

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

A dialog box titled "Please set the parameters" with the following fields and controls:

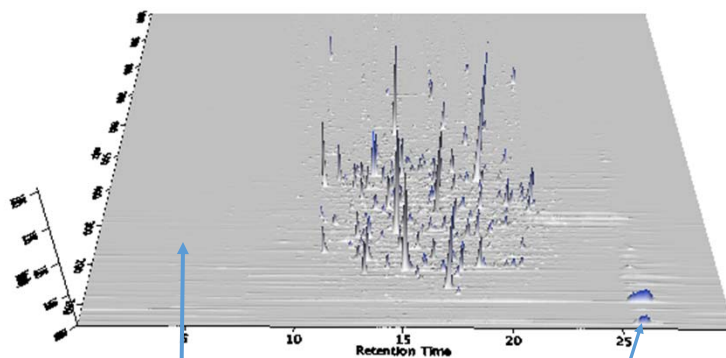
- Scan number: Two empty text input boxes separated by a hyphen.
- Retention time: Text input boxes containing "0" and "30", followed by "min." and an "Auto range" button.
- MS level: Text input box containing "1".
- Scan definition: Empty text input box.
- Polarity: Dropdown menu showing "-".
- Spectrum type: Dropdown menu showing "Any".
- Buttons: "OK", "Cancel", and "Help" at the bottom.

Ready to view the 3D-plot

A dialog box titled "Please set the parameters" with the following fields and controls:

- Raw data files: "Neg_C1.mzXML" and a dropdown menu showing "As selected in main window" with a "..." button.
- Scans: "Retention time: 0.00 - 30.00 min.", "MS level: 1", and "Polarity: -". Includes "Set filters" and "Clear filters" buttons.
- m/z: Text input boxes containing "49.9908" and "1000.0003", followed by "Auto range", "From mass", and "From formula" buttons.
- Retention time resolution: Text input box containing "500".
- m/z resolution: Text input box containing "500".
- Buttons: "OK", "Cancel", and "Help" at the bottom.

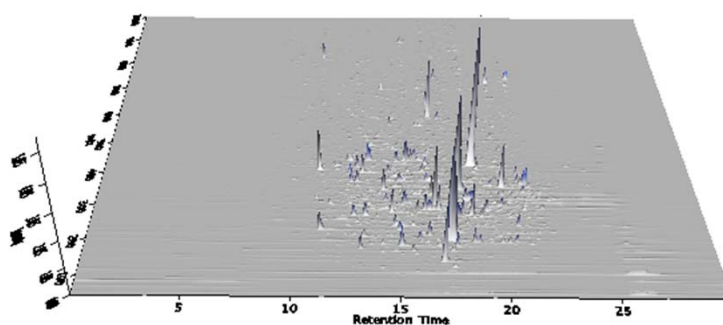
3D view of NegMode_C1



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

Please set the parameters

Scan number -

Retention time - min.

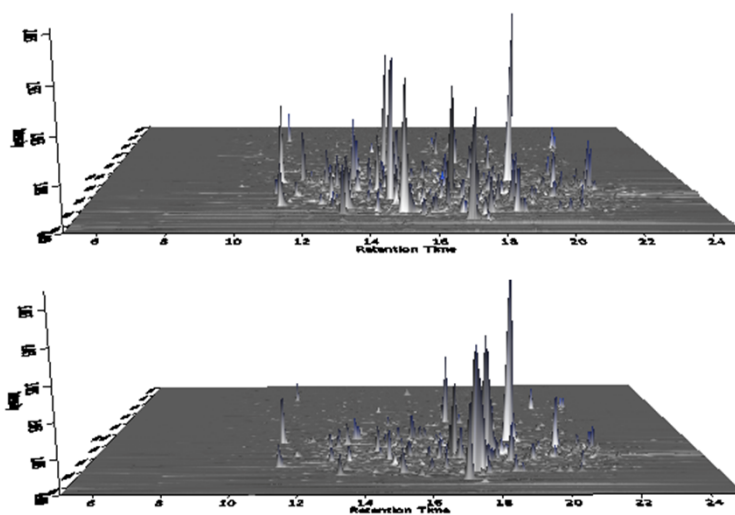
MS level

Scan definition

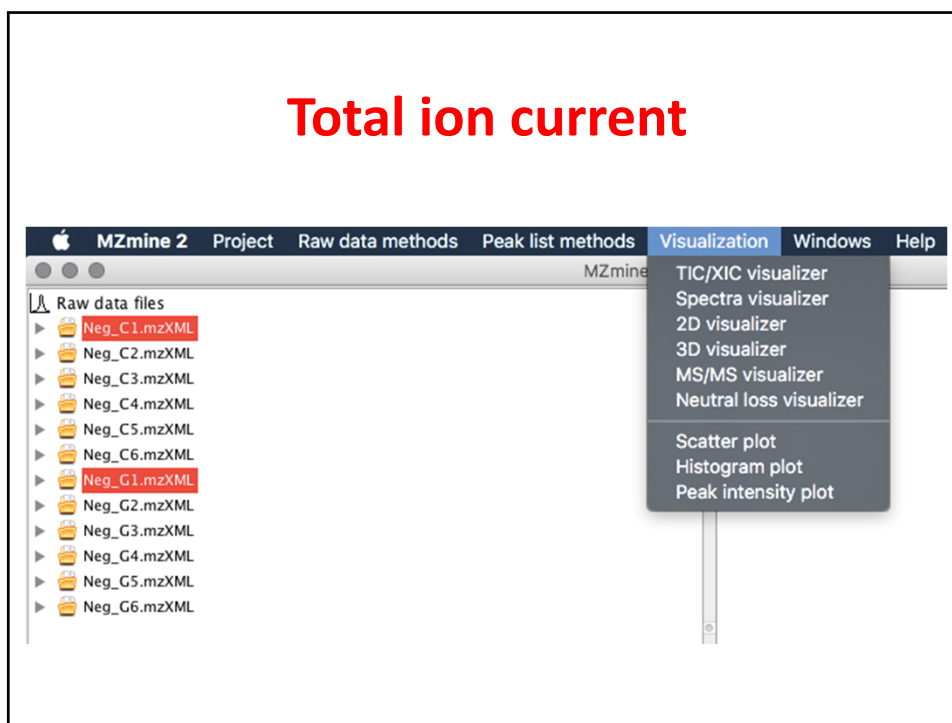
Polarity

Spectrum type

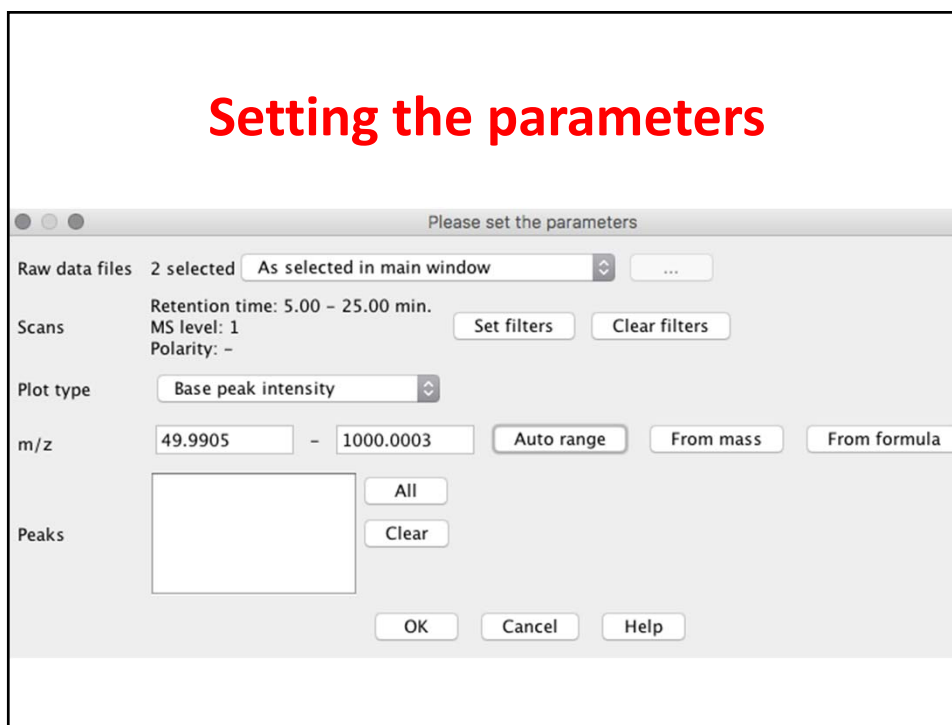
Data from 5-25 min



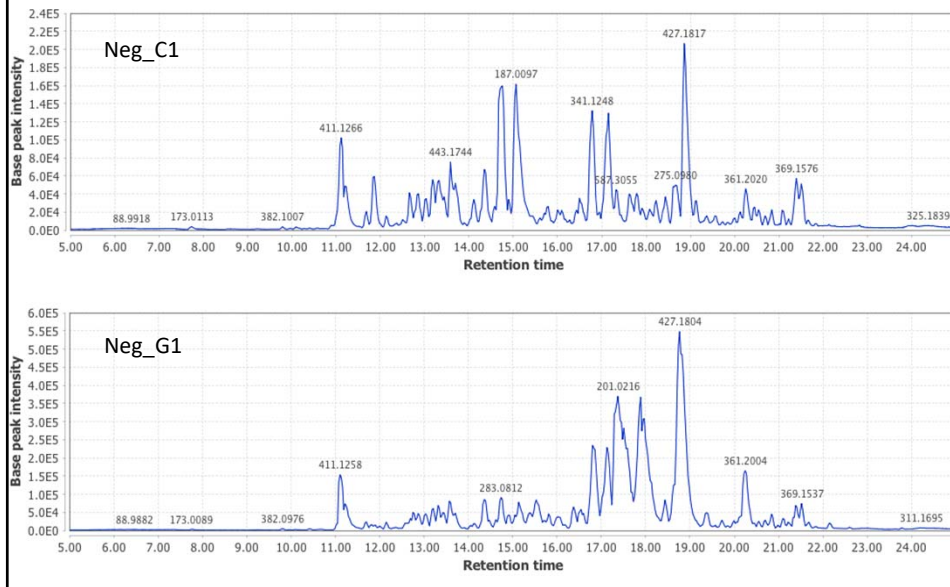
Total ion current



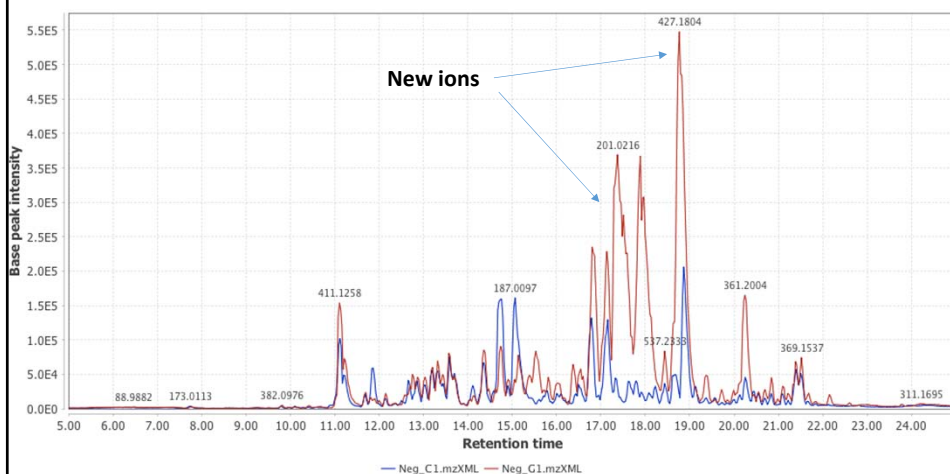
Setting the parameters



TIC of all ions from m/z 50-1000

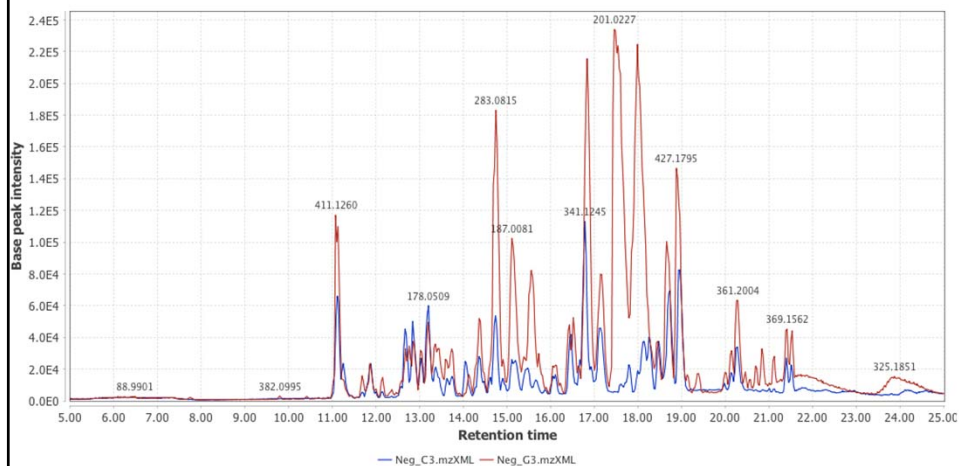


TIC m/z 50-1000 C1/G1 comparison



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

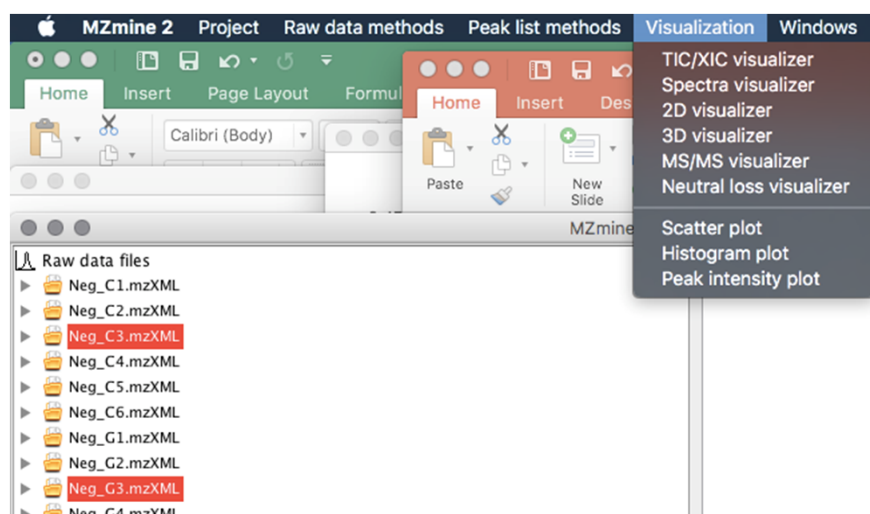
A	B	C	D	E	F
Name	Empirical formula	C	H	N	O
hexanol	C6H14O	6	14	0	1
glucose	C6H12O6	6	12	0	6
genistein	C15H10O5	15	10	0	5

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 ₈ S	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

Please set the parameters

Raw data files 2 selected ...

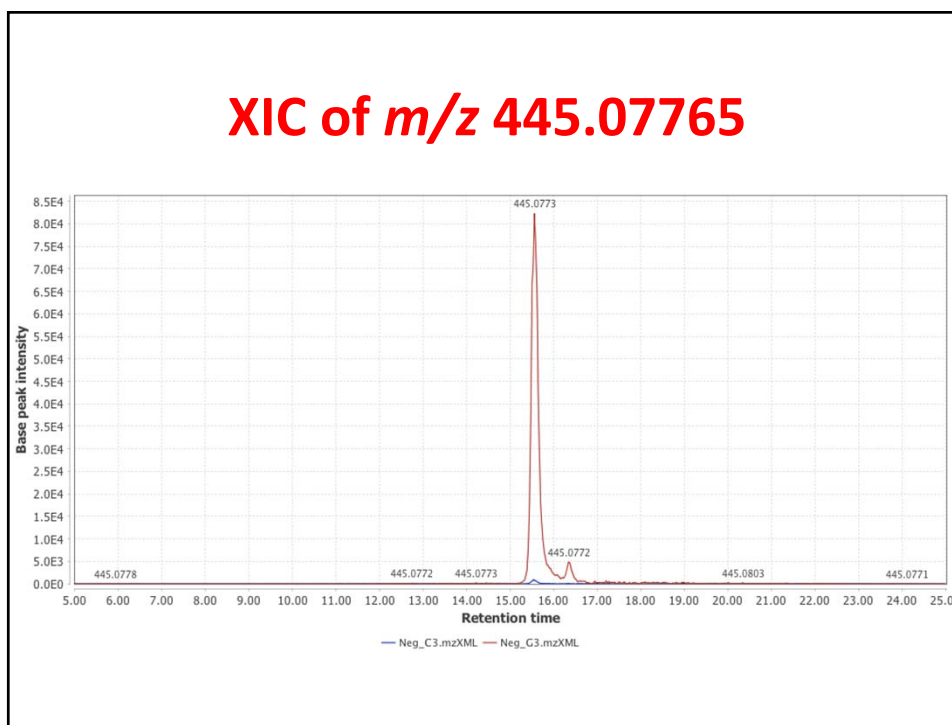
Scans Retention time: 5.00 - 25.00 min.
MS level: 1
Polarity: -

Plot type

m/z -

Peaks **C₂₁H₁₈O₁₁**

XIC of m/z 445.07765



Getting MS/MS data

Select TIC/XIC and reset the parameters

Please set the parameters

Scan number -

Retention time - min.

MS level Retention time range in minutes

Scan definition

Polarity

Spectrum type

Ready to go

Please set the parameters

Raw data files 2 selected ...

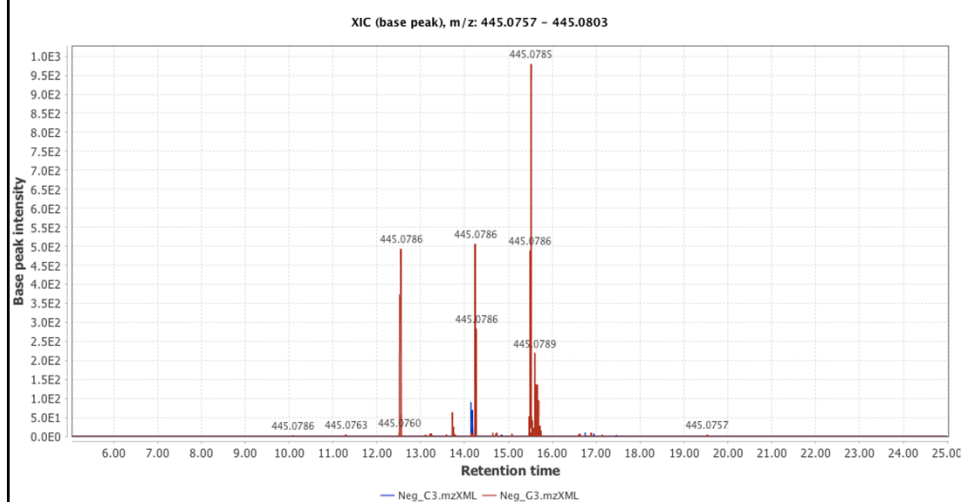
Scans Retention time: 5.00 - 25.00 min.
MS level: 2
Polarity: -

Plot type

m/z -

Peaks

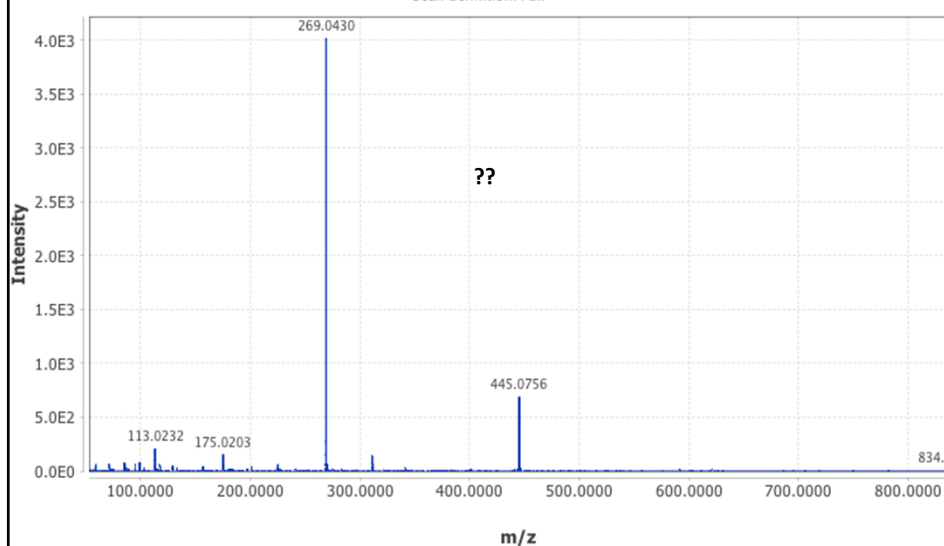
MSMS spectra captured of m/z 445.077



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

Neg_G3.mzXML scan#1726 @12.55 MS2 (621.1085) p -, base peak: 269.0430 m/z (4.0E3)

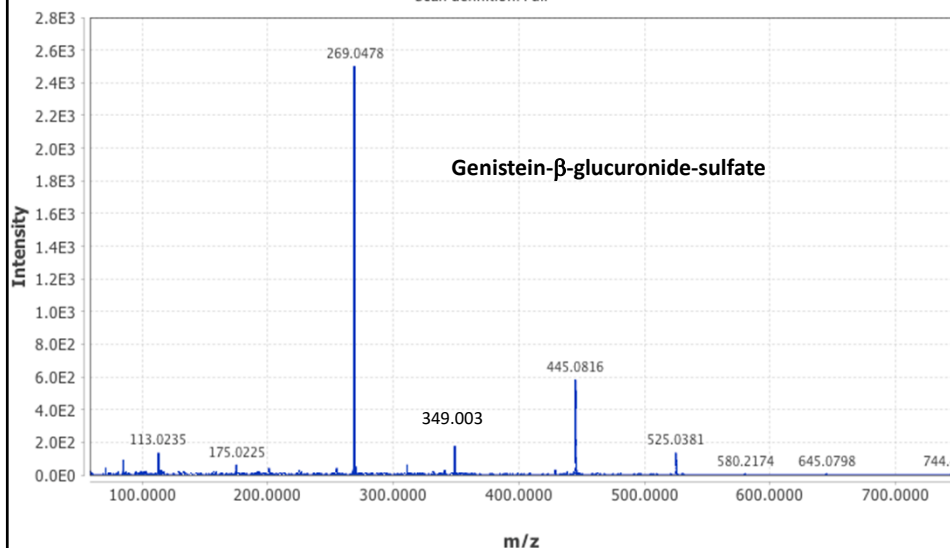
Scan definition: Full



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

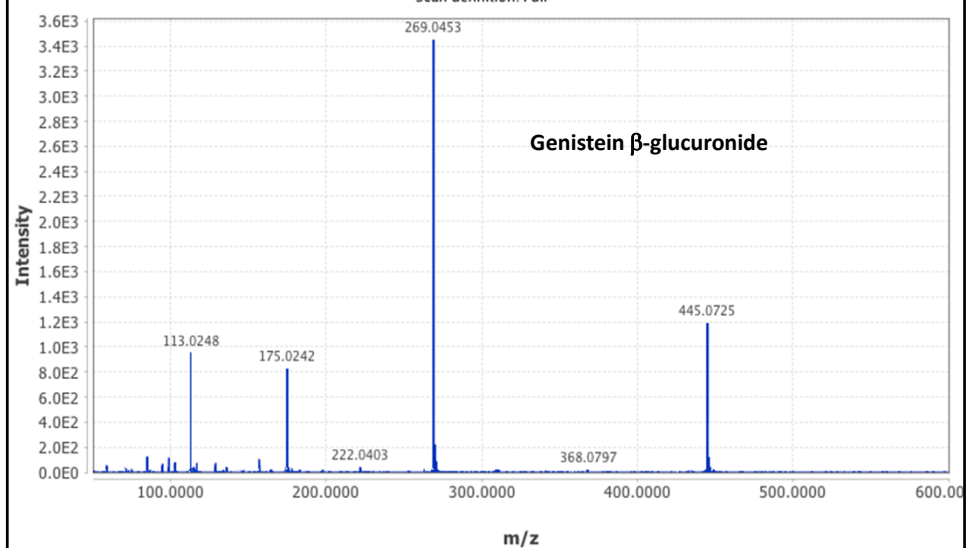
Neg_G3.mzXML scan#2200 @14.24 MS2 (525.0334) p -, base peak: 269.0478 m/z (2.5E3)

Scan definition: Full

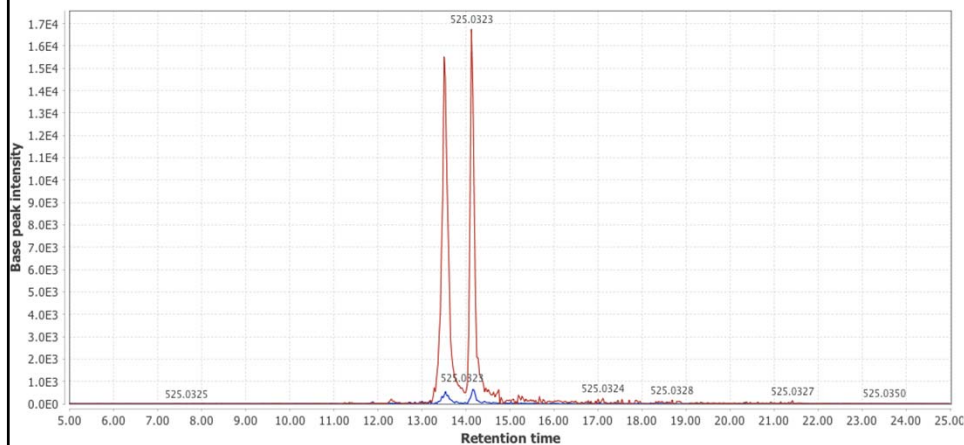
**Where MSMS for m/z 445.079 were recorded (3)**

Neg_G3.mzXML scan#2611 @15.52 MS2 (668.1173) p -, base peak: 269.0453 m/z (3.5E3)

Scan definition: Full



XIC of m/z 525.033 genistein β -glucuronide/sulfate

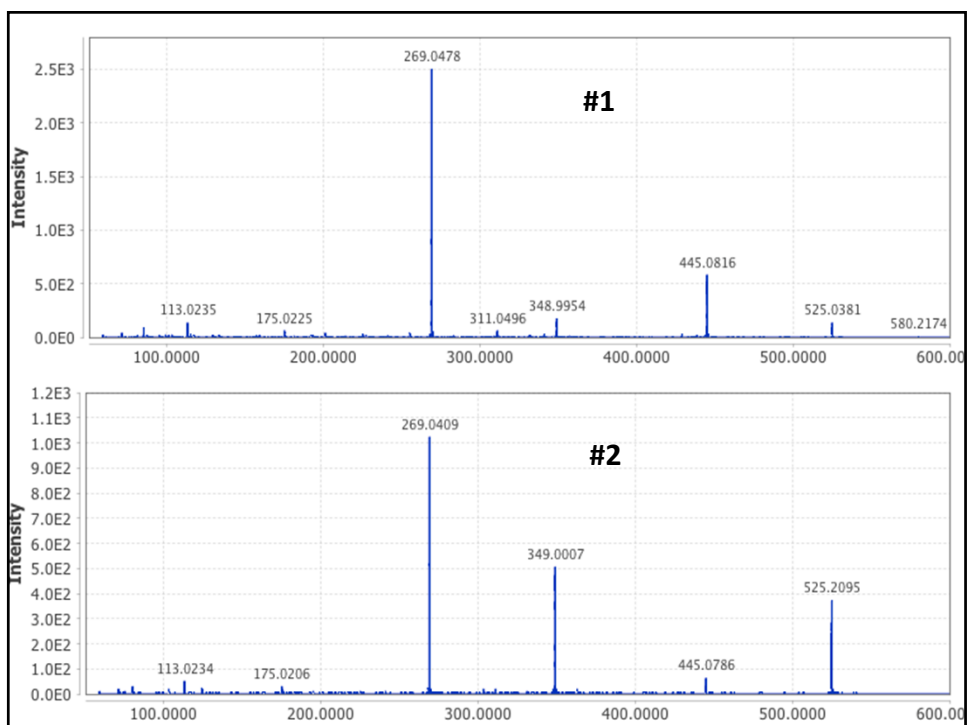
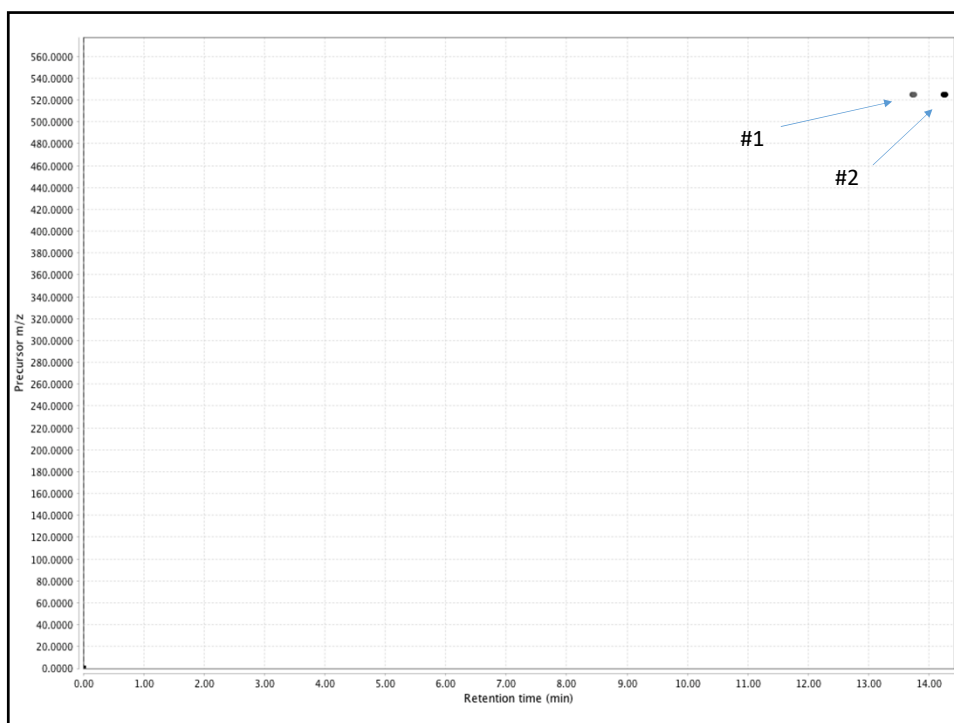


Another way to get MSMS spectra Select MSMS visualizer

Please set the parameters

Raw data files	Neg_G3.mzXML	As selected in main window	...
Retention time	10.00	-	20.00 min. Auto range
m/z	525.0313	-	525.0354 Auto range From mass From formula
Intensity	Total intensity in MS/MS scan		
Normalize by	All data points		
Min. MS/MS peak intensity	1		

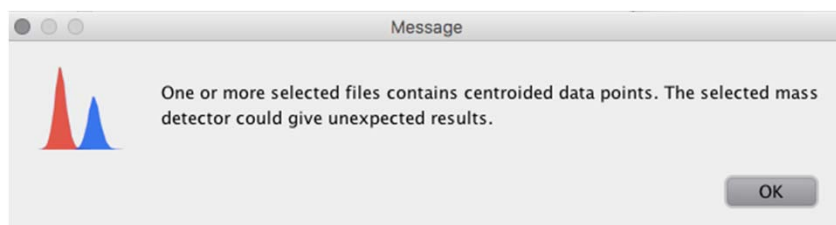
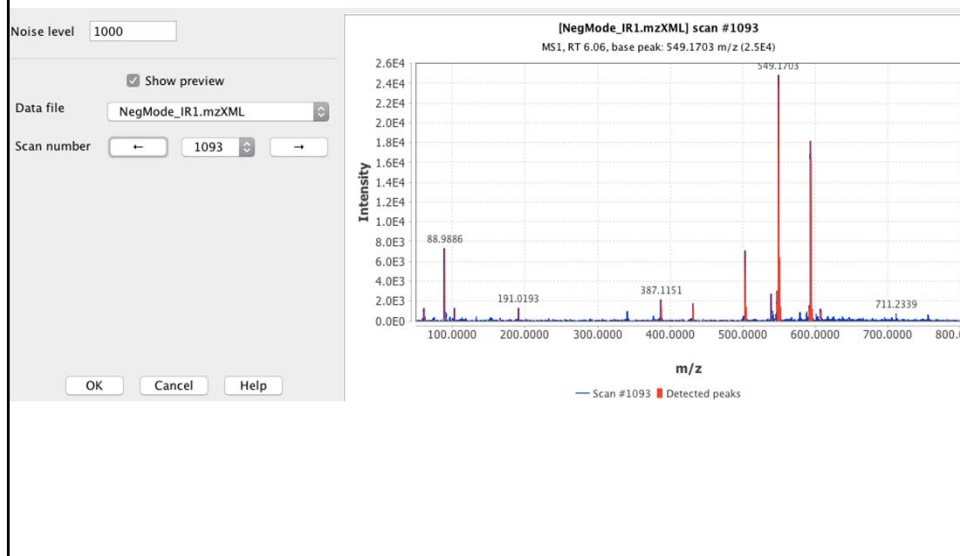
OK Cancel Help



Identifying the masses in .mzxml files

Setting the background

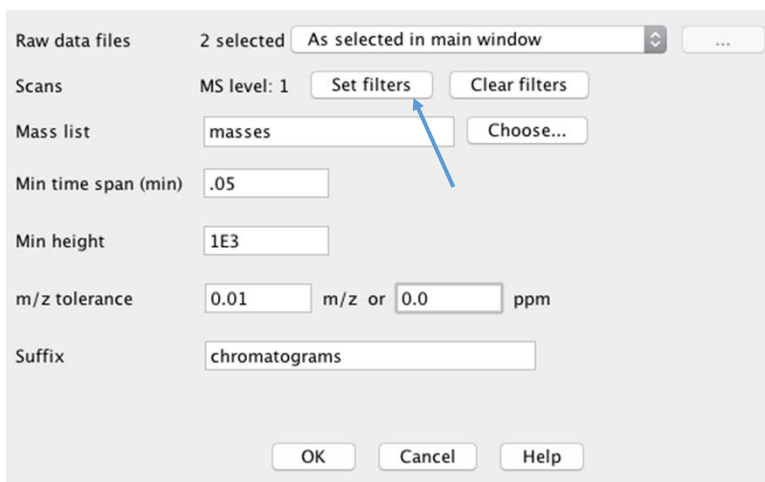
Setting the background



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



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

Scans MS level: 1 Set filters Clear filters

Mass list masses Choose...

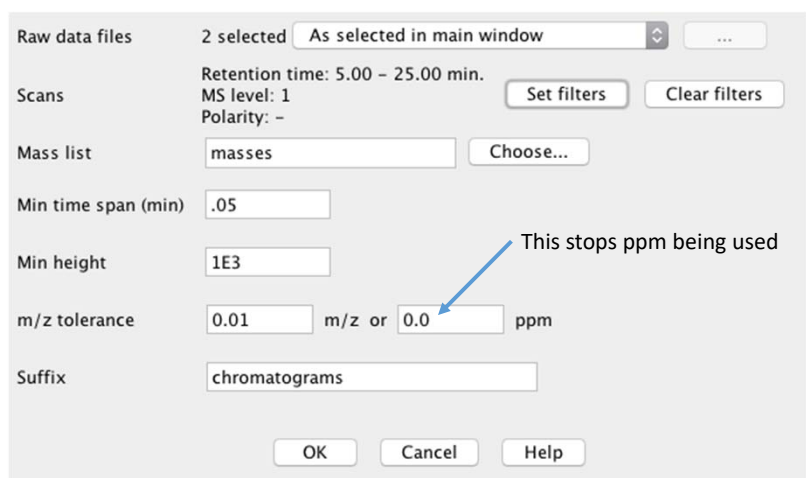
Min time span (min) .05

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 ...

Retention time: 5.00 - 25.00 min.

Scans MS level: 1 Set filters Clear filters
Polarity: -

Mass list masses Choose...

Min time span (min) .05

Min height 1E3

m/z tolerance 0.01 m/z or 0.0 ppm

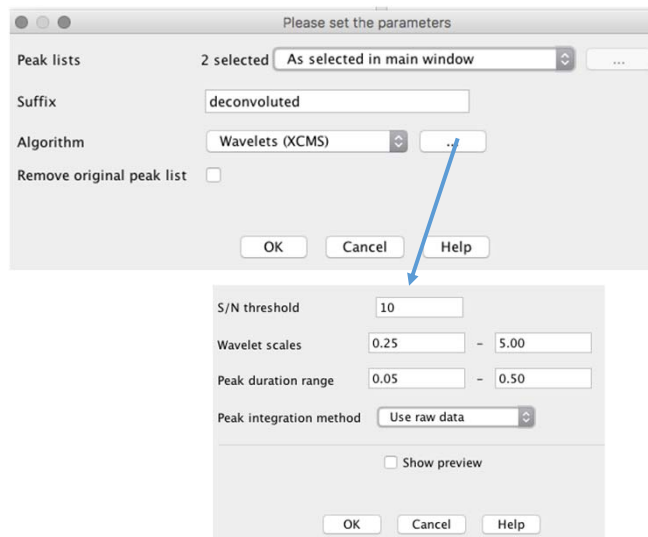
Suffix chromatograms

OK Cancel Help

This stops ppm being used

Chromatogram deconvolution

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



Chromatogram deconvolution output

Raw data files

- NegMode_IR1.mzXML
- NegMode_NR1.mzXML

NegMode_NR1.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_NR1.mzXML chromatograms
- NegMode_IR1.mzXML chromatograms
- NegMode_NR1.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_NR1.mzXML	431.0966	14.20	2.2E4	8.6E4

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 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.