A stoned cat

Data analysis in metabolomics

Stephen Barnes, PhD
Department of Pharmacology and Toxicology
Targeted Metabolomics and Proteomics Laboratory
BBRB 711
sbarnes@uab.edu; 205-934-7117

1

Goals of the data analysis classes

- **To understand the nature of metabolomics data**
- **To become familiar with three programs**
  - MS-DIAL (current version 4.80)
  - Mzmine (current version 2.53)
  - Metaboanalyst 5.0 (online at http://www.Metaboanalyst.ca)
- **At this time, MS-DIAL is best run on Windows (10) as a GUI**
  - There are Linux and Mac versions
  - Can be run on a Windows box at the command line

2

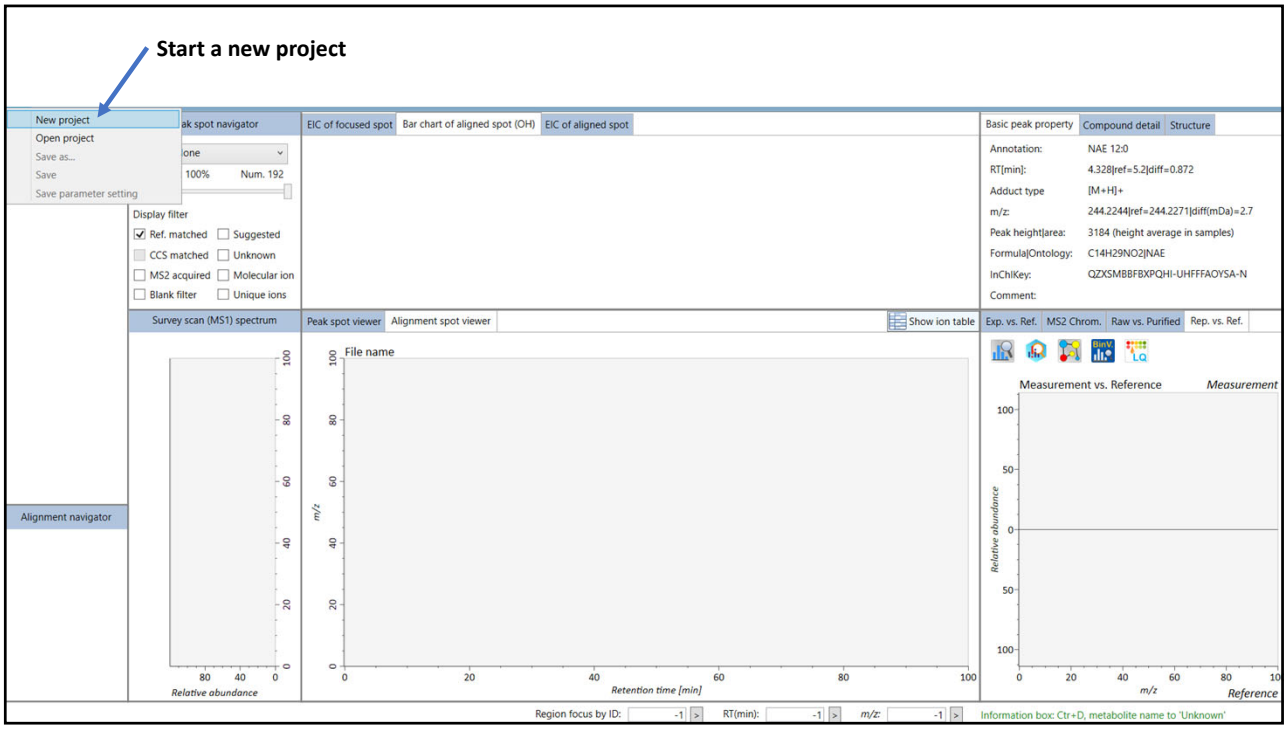
A data set on urines from colleagues at Harvard

- This is part of a study of urines collected from COVID-19 patients, some of whom developed acute kidney injury
- The study has been published in Kidney International Reports (Raines et al.)
- There were 14 patients in each group
- Additional samples run were 3 pooled samples, 3 extraction blanks and one reagent (methanol) blank
- Full scan MS1 data were collected at 4 Hz over a 0-8 min gradient
- The pooled sample was re-run, collecting MS1 data for 100 ms, followed by 8 computer selected, 50 ms MSMS spectra

3

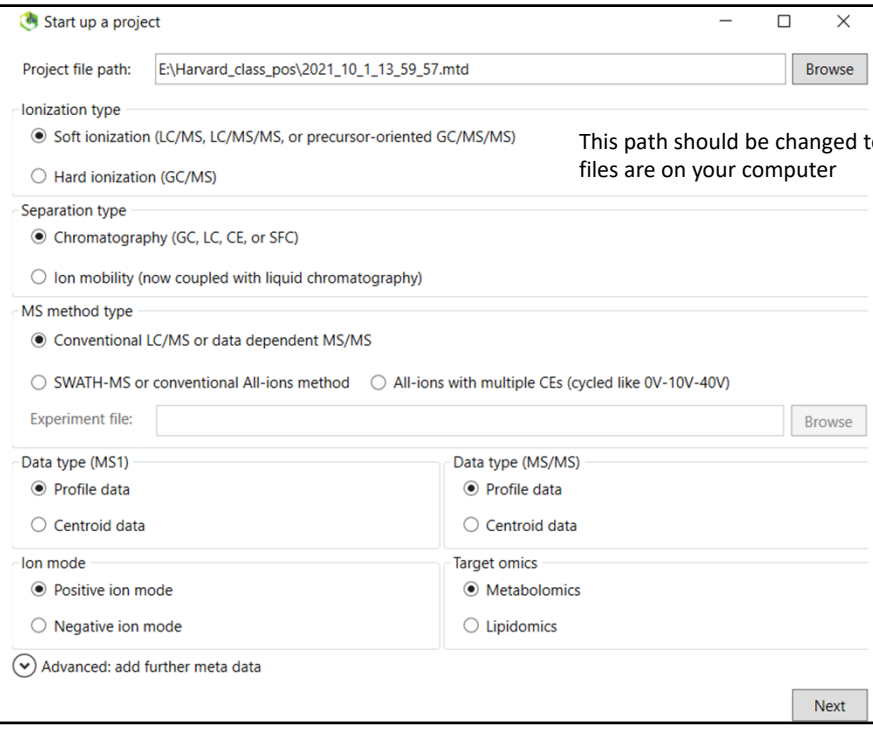
Locate and launch MS-DIAL

4



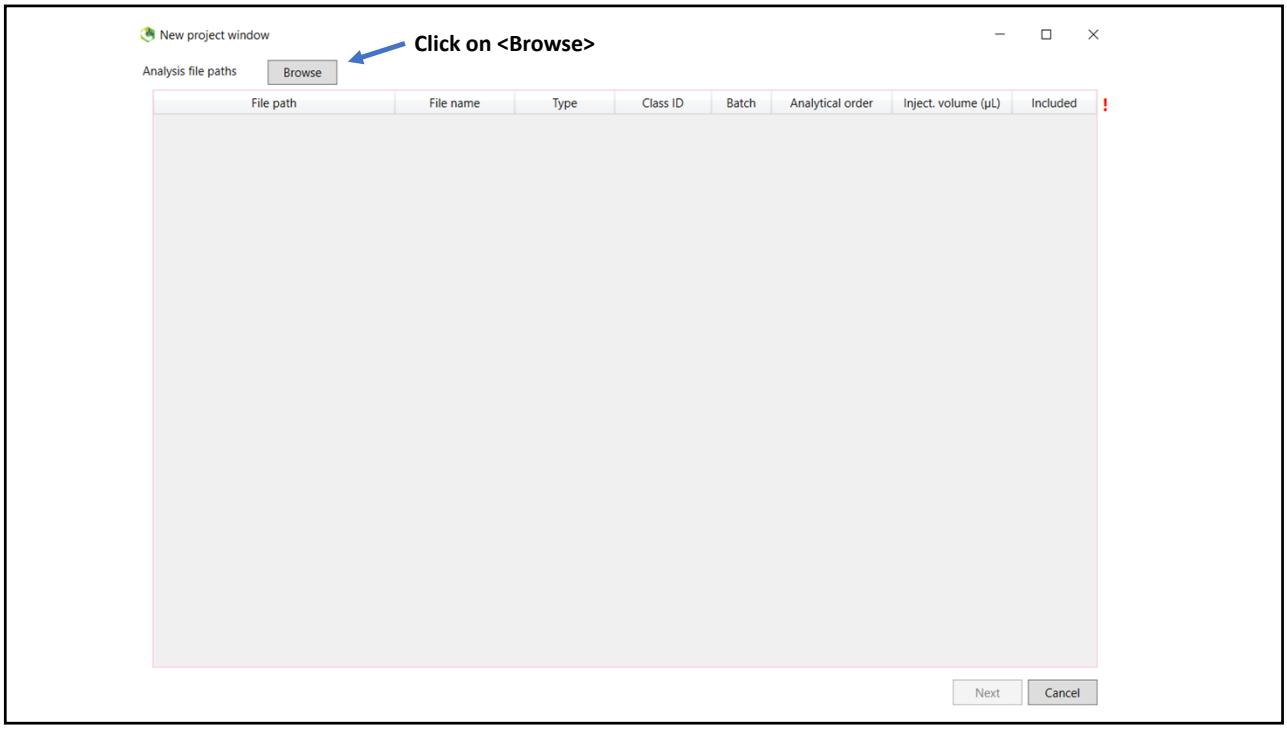
5

Define the path to the files

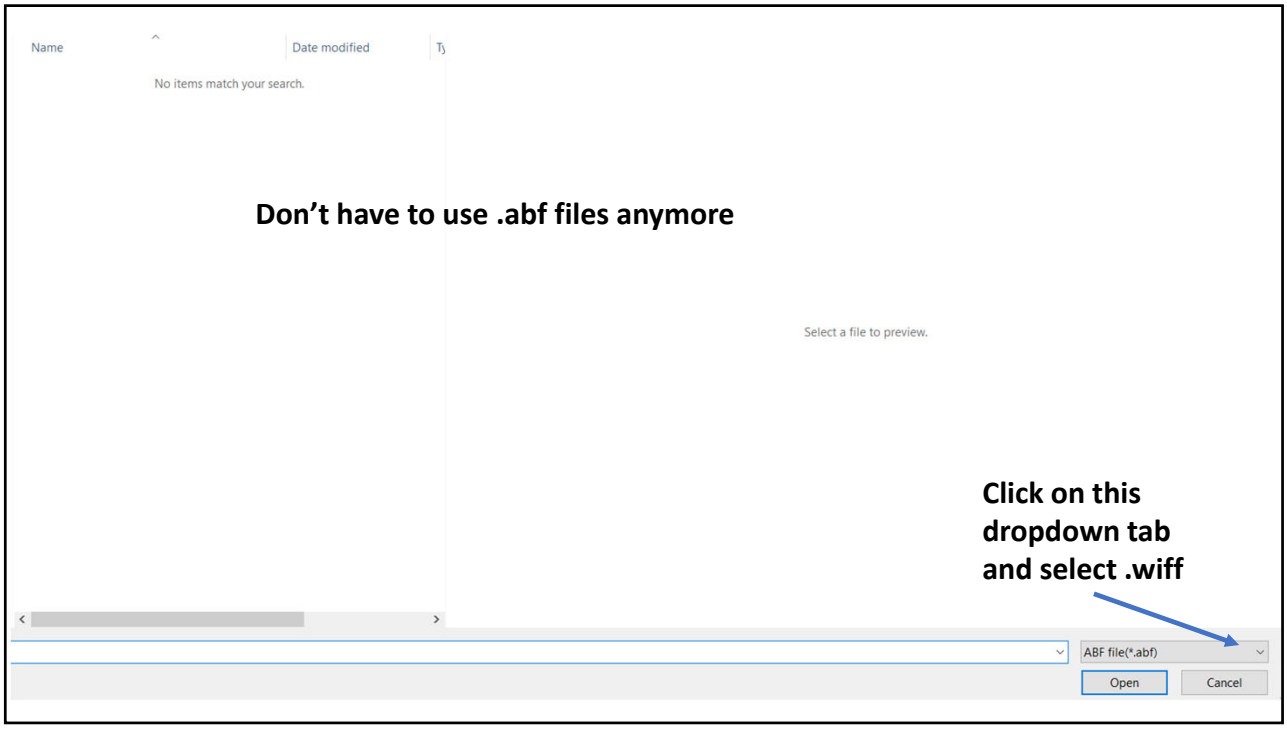


This path should be changed to where the files are on your computer

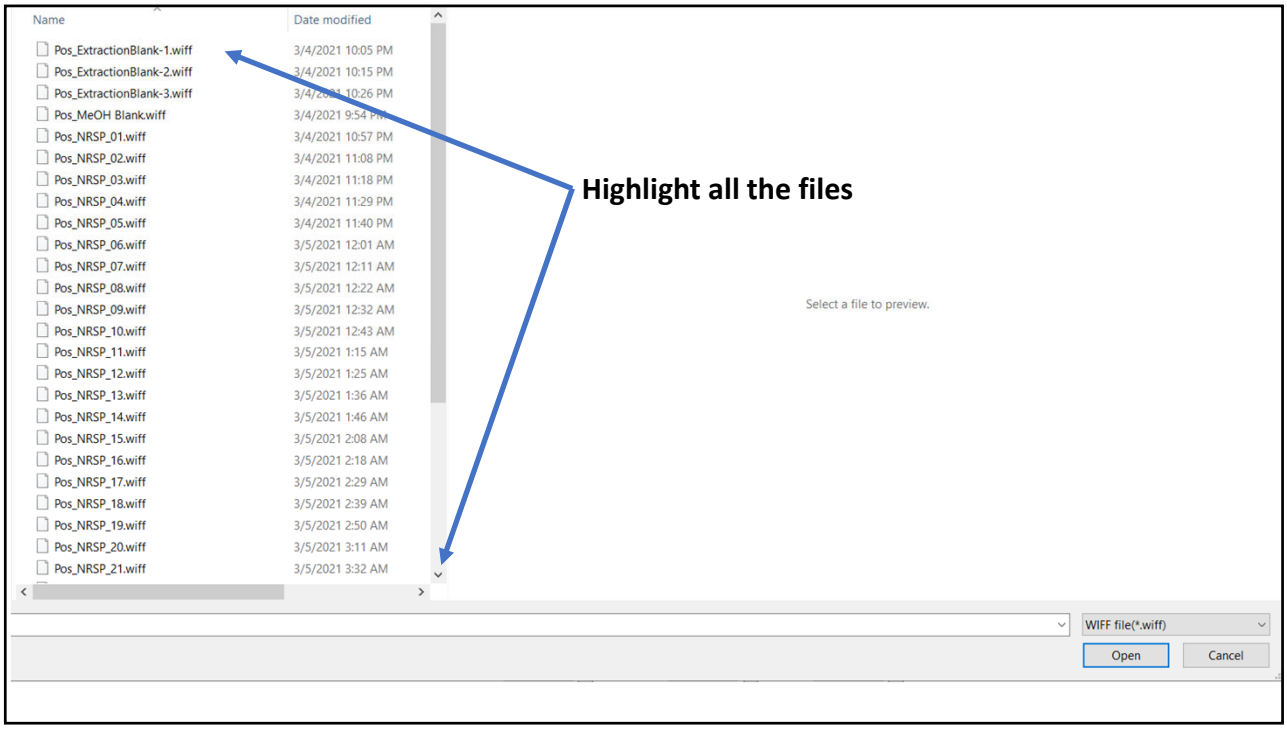
6


New project window
Click on <Browse>
Analysis file paths
Browse
File path
File name
Type
Class ID
Batch
Analytical order
Inject. volume (µL)
Included
Next
Cancel

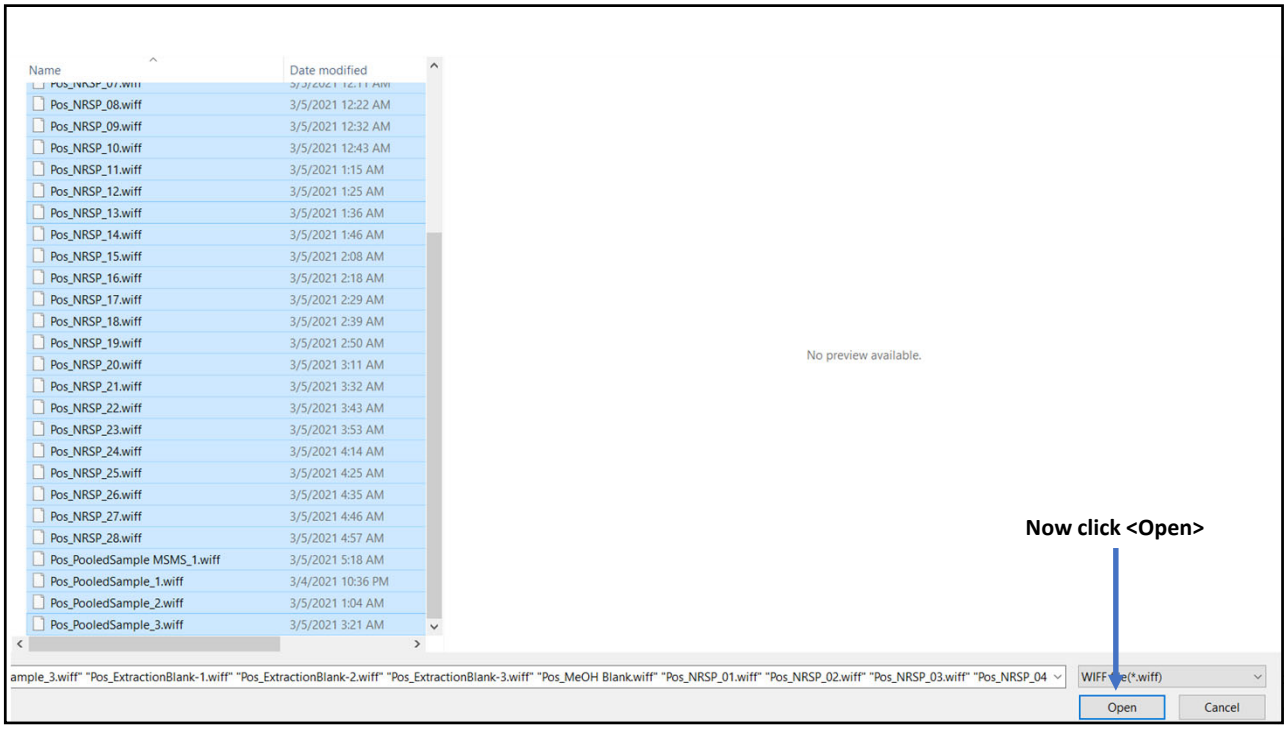
7


Name
Date modified
No items match your search.
Don't have to use .abf files anymore
Select a file to preview.
Click on this dropdown tab and select .wiff
ABF file(*.abf)
Open
Cancel

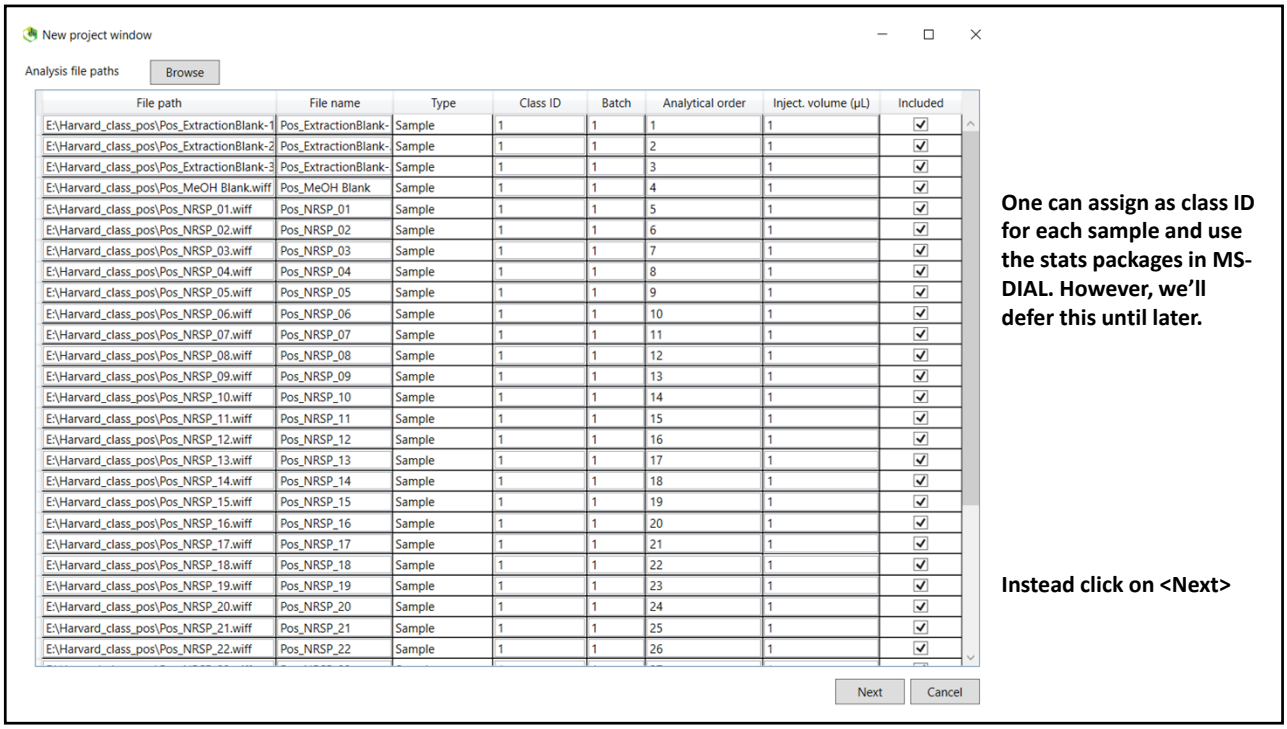
8


Name
Date modified
Pos_ExtractionBlank-1.wiff 3/4/2021 10:05 PM
Pos_ExtractionBlank-2.wiff 3/4/2021 10:15 PM
Pos_ExtractionBlank-3.wiff
Pos_MeOH Blank.wiff
Pos_NRSP_01.wiff 3/4/2021 10:57 PM
Pos_NRSP_02.wiff 3/4/2021 11:08 PM
Pos_NRSP_03.wiff 3/4/2021 11:18 PM
Pos_NRSP_04.wiff 3/4/2021 11:29 PM
Pos_NRSP_05.wiff 3/4/2021 11:40 PM
Pos_NRSP_06.wiff 3/5/2021 12:01 AM
Pos_NRSP_07.wiff 3/5/2021 12:11 AM
Pos_NRSP_08.wiff 3/5/2021 12:22 AM
Pos_NRSP_09.wiff 3/5/2021 12:32 AM
Pos_NRSP_10.wiff 3/5/2021 12:43 AM
Pos_NRSP_11.wiff 3/5/2021 1:15 AM
Pos_NRSP_12.wiff 3/5/2021 1:25 AM
Pos_NRSP_13.wiff 3/5/2021 1:36 AM
Pos_NRSP_14.wiff 3/5/2021 1:46 AM
Pos_NRSP_15.wiff 3/5/2021 2:08 AM
Pos_NRSP_16.wiff 3/5/2021 2:18 AM
Pos_NRSP_17.wiff 3/5/2021 2:29 AM
Pos_NRSP_18.wiff 3/5/2021 2:39 AM
Pos_NRSP_19.wiff 3/5/2021 2:50 AM
Pos_NRSP_20.wiff 3/5/2021 3:11 AM
Pos_NRSP_21.wiff 3/5/2021 3:32 AM
Highlight all the files
Select a file to preview.
WIFF file(*.wiff)
Open
Cancel

9


Name
Date modified
Pos_NRSP_08.wiff 3/5/2021 12:22 AM
Pos_NRSP_09.wiff 3/5/2021 12:32 AM
Pos_NRSP_10.wiff 3/5/2021 12:43 AM
Pos_NRSP_11.wiff 3/5/2021 1:15 AM
Pos_NRSP_12.wiff 3/5/2021 1:25 AM
Pos_NRSP_13.wiff 3/5/2021 1:36 AM
Pos_NRSP_14.wiff 3/5/2021 1:46 AM
Pos_NRSP_15.wiff 3/5/2021 2:08 AM
Pos_NRSP_16.wiff 3/5/2021 2:18 AM
Pos_NRSP_17.wiff 3/5/2021 2:29 AM
Pos_NRSP_18.wiff 3/5/2021 2:39 AM
Pos_NRSP_19.wiff 3/5/2021 2:50 AM
Pos_NRSP_20.wiff 3/5/2021 3:11 AM
Pos_NRSP_21.wiff 3/5/2021 3:32 AM
Pos_NRSP_22.wiff 3/5/2021 3:43 AM
Pos_NRSP_23.wiff 3/5/2021 3:53 AM
Pos_NRSP_24.wiff 3/5/2021 4:14 AM
Pos_NRSP_25.wiff 3/5/2021 4:25 AM
Pos_NRSP_26.wiff 3/5/2021 4:35 AM
Pos_NRSP_27.wiff 3/5/2021 4:46 AM
Pos_NRSP_28.wiff 3/5/2021 4:57 AM
Pos_PooledSample MSMS_1.wiff 3/5/2021 5:18 AM
Pos_PooledSample_1.wiff 3/4/2021 10:36 PM
Pos_PooledSample_2.wiff 3/5/2021 1:04 AM
Pos_PooledSample_3.wiff 3/5/2021 3:21 AM
No preview available.
Now click <Open>
ample_3.wiff" "Pos_ExtractionBlank-1.wiff" "Pos_ExtractionBlank-2.wiff" "Pos_ExtractionBlank-3.wiff" "Pos_MeOH Blank.wiff" "Pos_NRSP_01.wiff" "Pos_NRSP_02.wiff" "Pos_NRSP_03.wiff" "Pos_NRSP_04
Open
Cancel

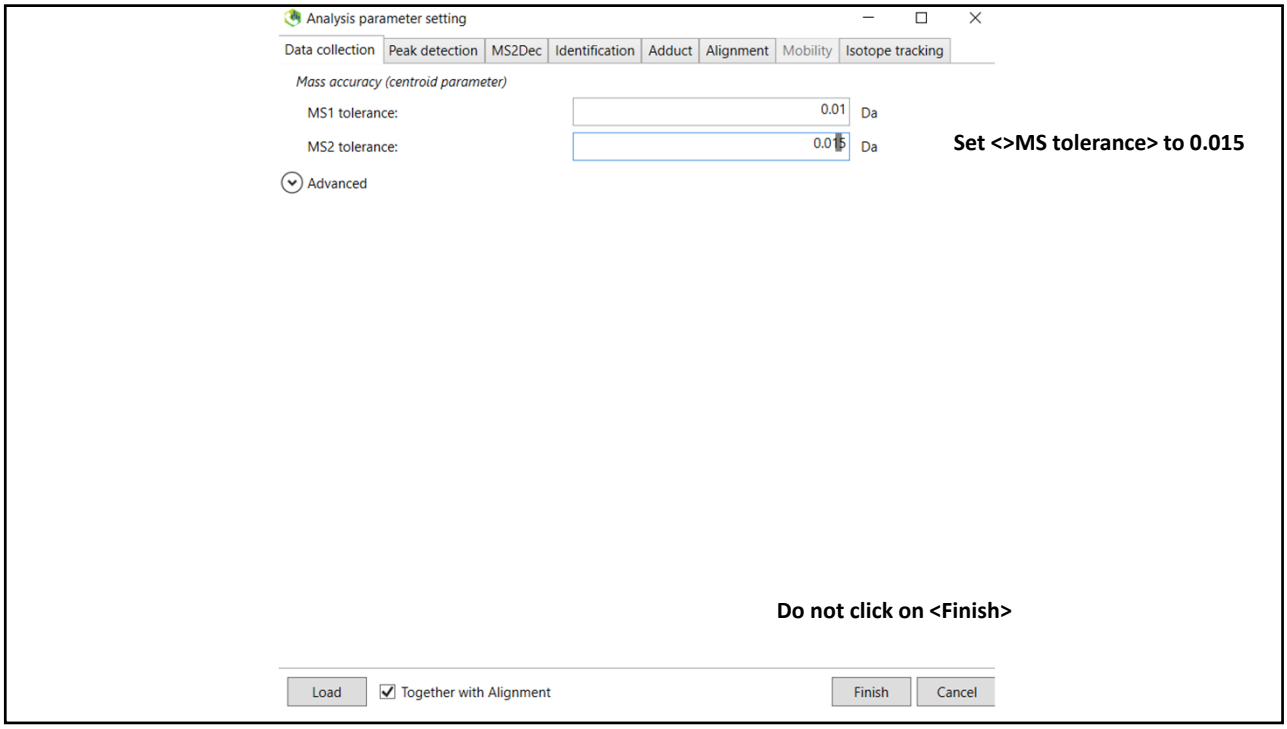
10



One can assign as class ID for each sample and use the stats packages in MS-DIAL. However, we'll defer this until later.

Instead click on <Next>

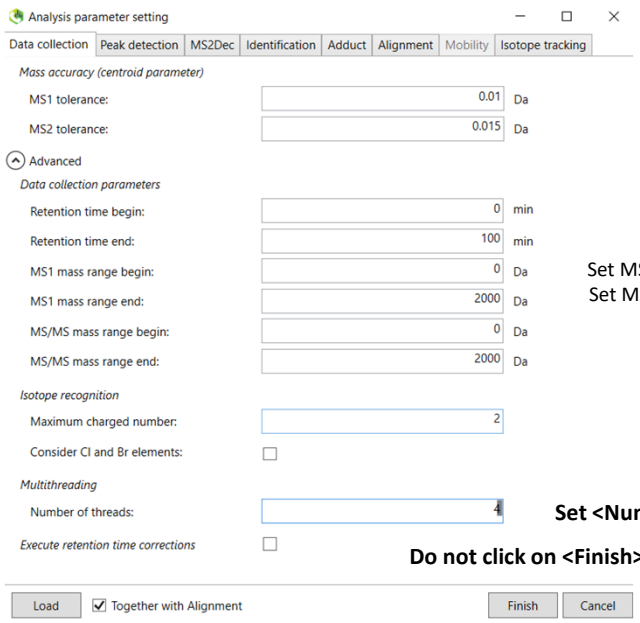
11



Set <>MS tolerance> to 0.015

Do not click on <Finish>

12

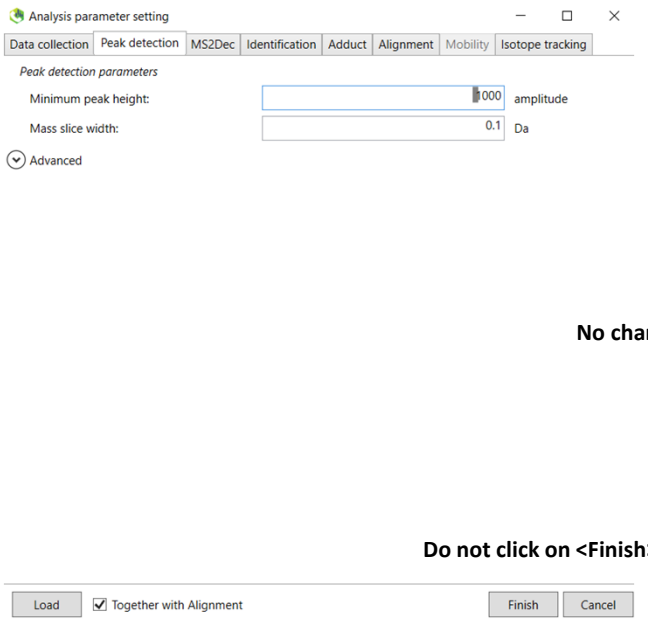


Set MS1 begin to 50
Set MS1 end to 1000

Set <Number of threads> to 4

Do not click on <Finish>

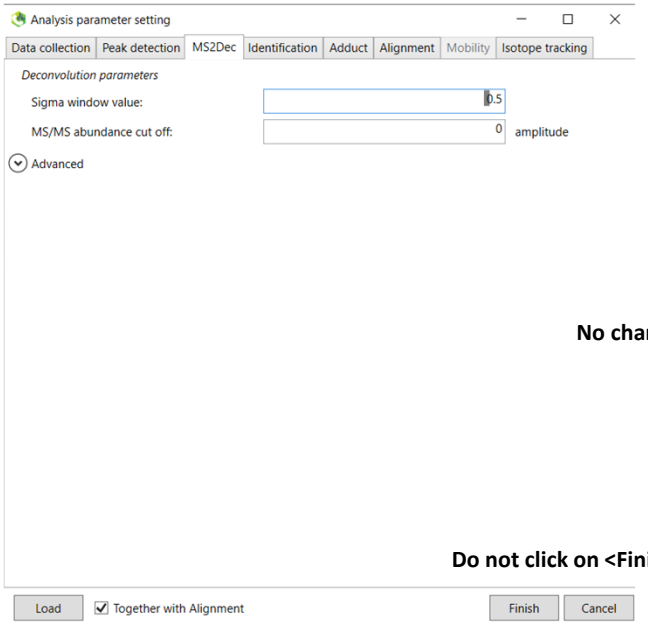
13



No changes from default values

Do not click on <Finish>

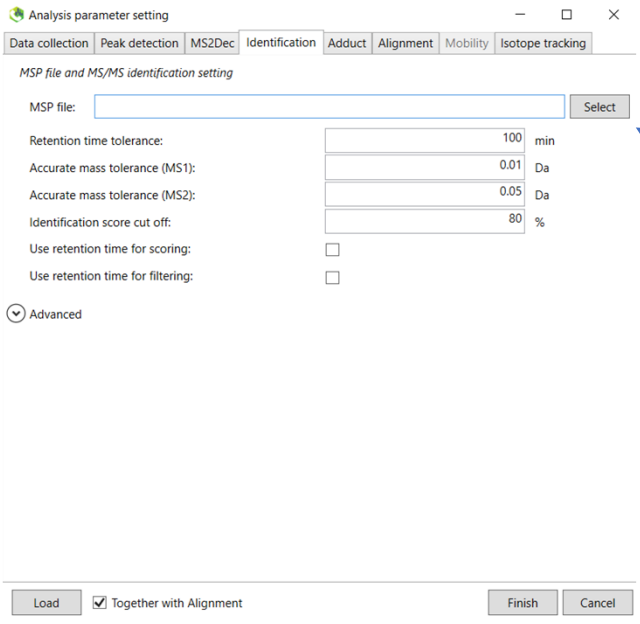
14



No changes from default values

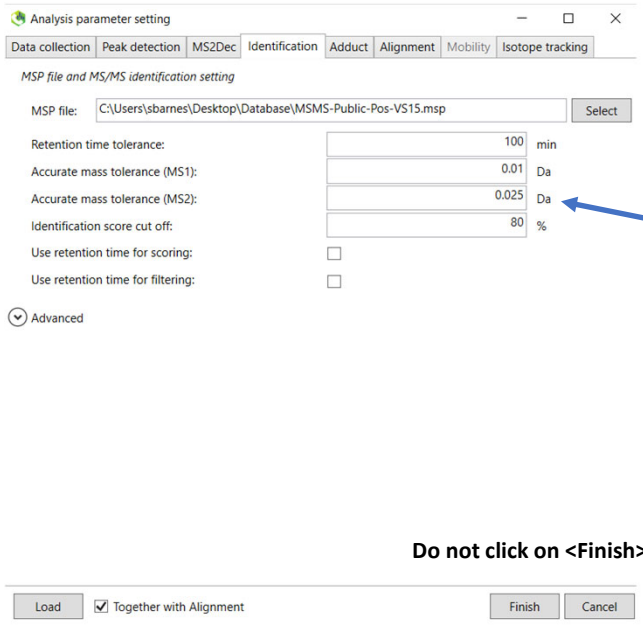
Do not click on <Finish>

15



Find the Dbase

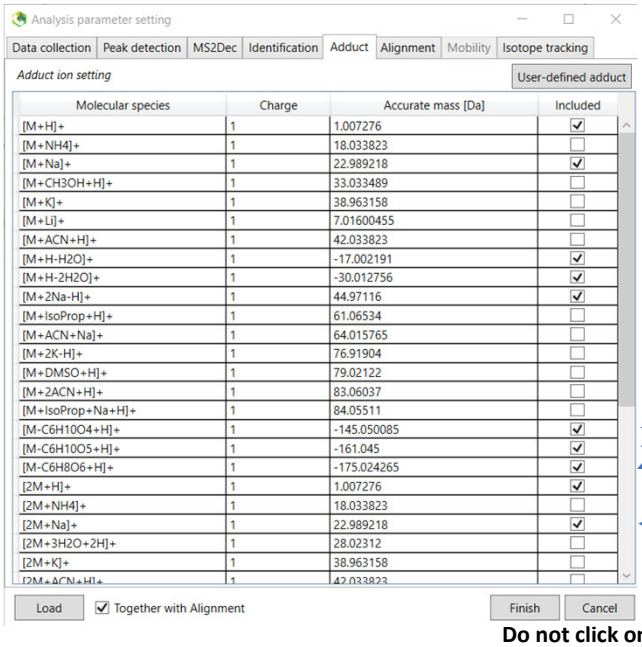
16



Change the MS2 tolerance to 0.025

Do not click on <Finish>

17



Select the likely adducts besides $[M+H]^+$

Na^+ adduct

Loss of water

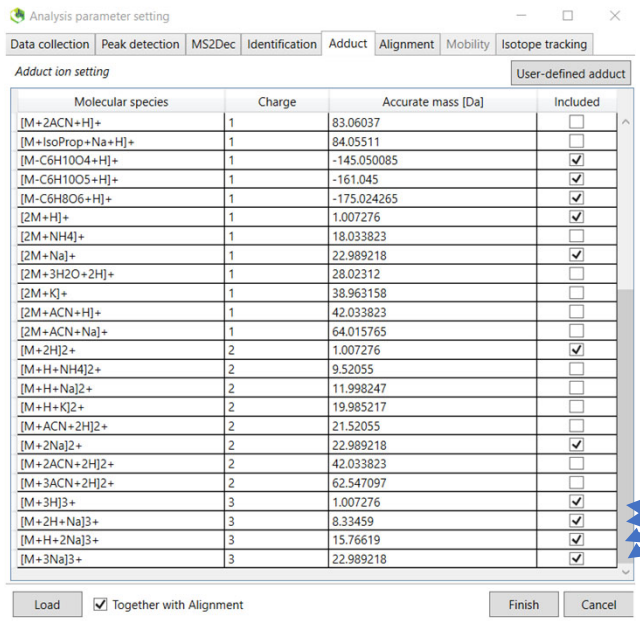
2 Na^+ adduct

Polar metabolites found in urine

Dimer

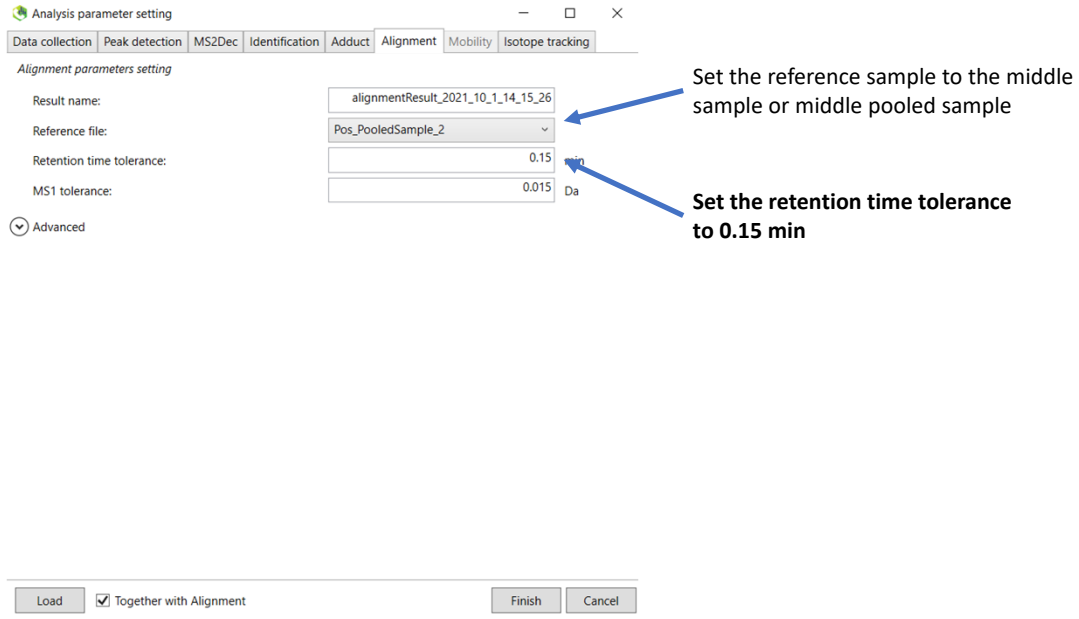
Do not click on <Finish>

18



Do not click on <Finish>

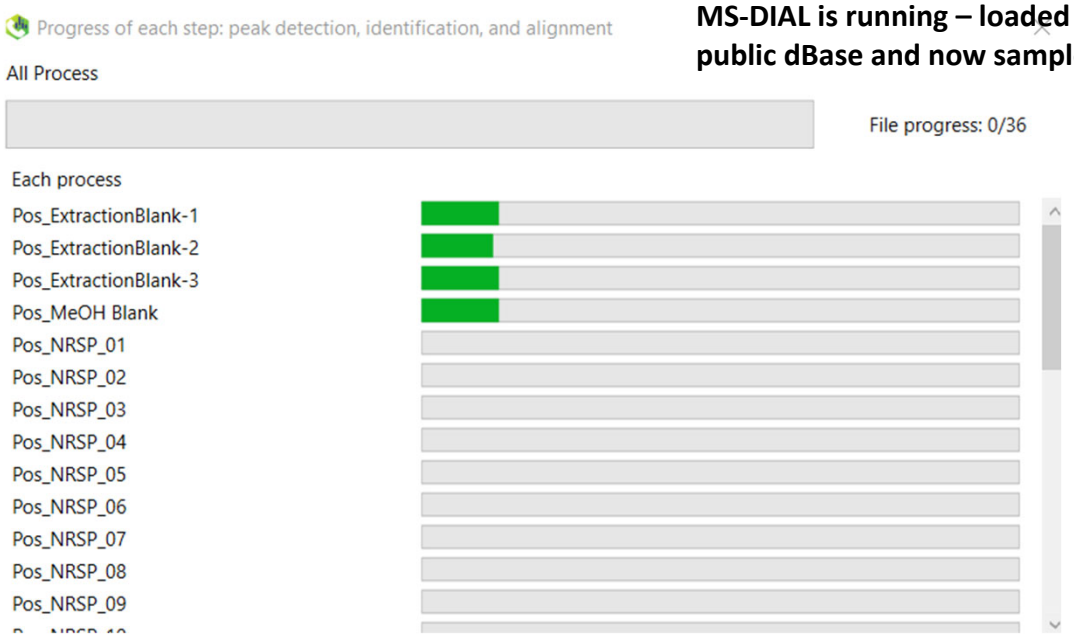
19



Now you can click on <Finish>

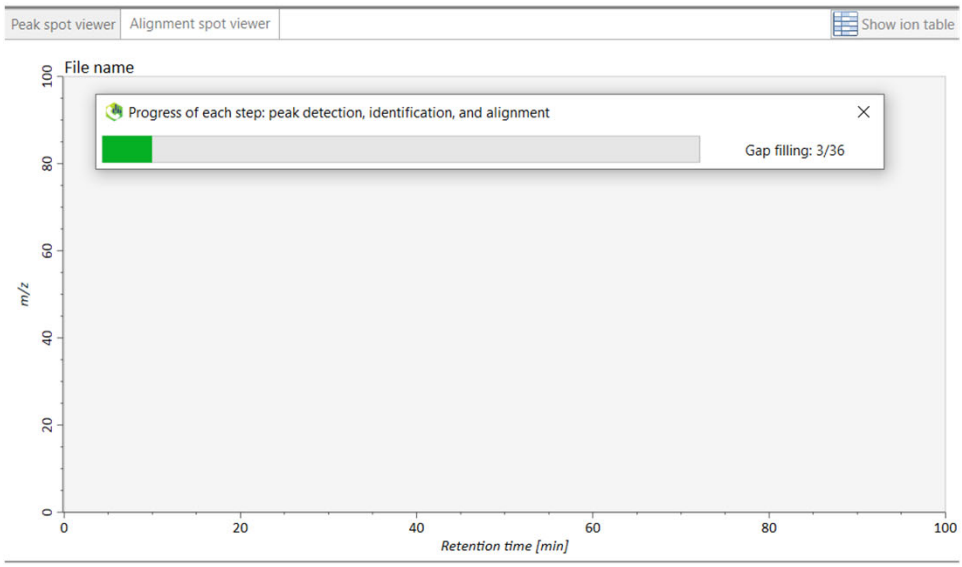
20

MS-DIAL is running – loaded public dBase and now samples

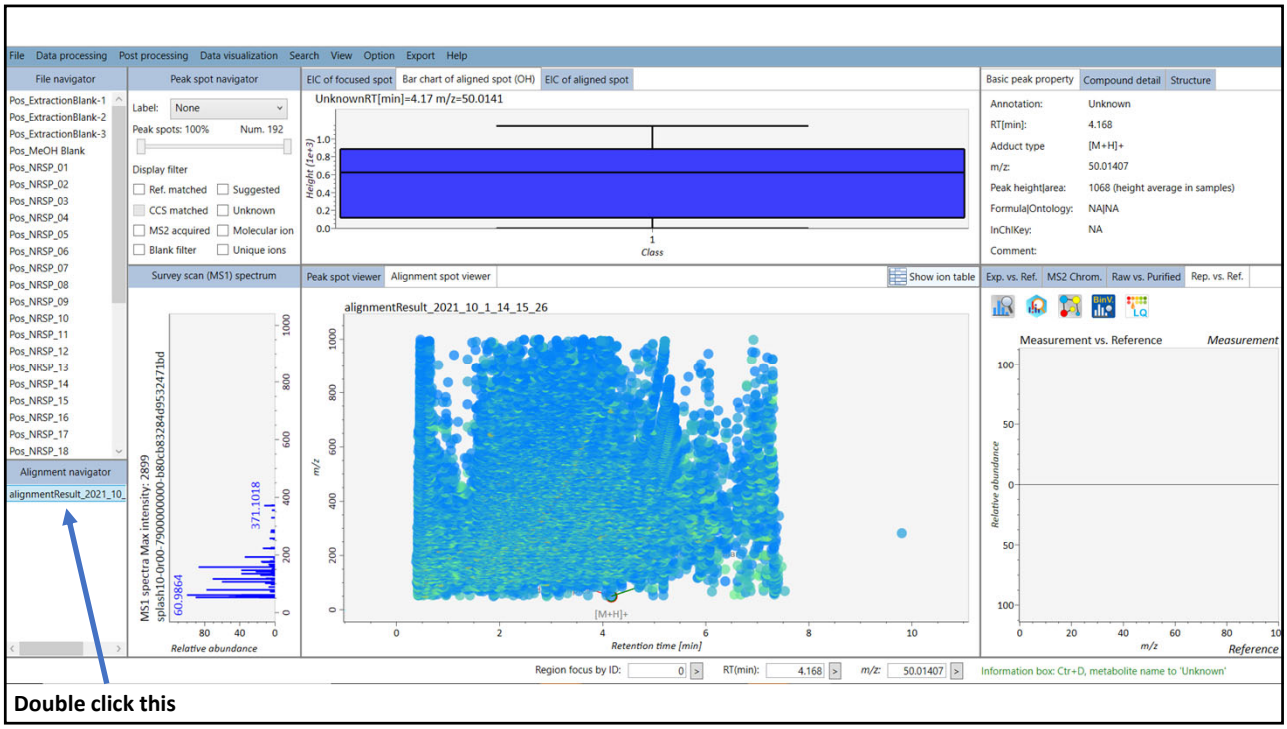


21

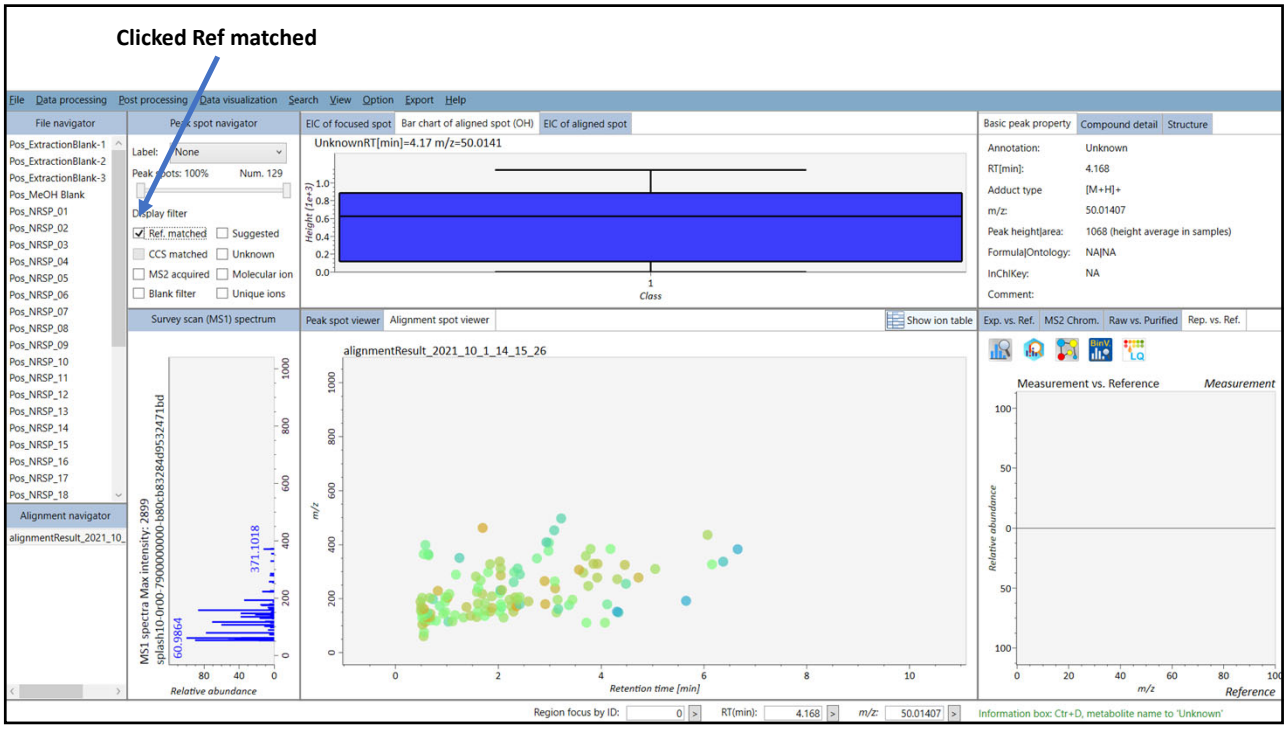
MS-DIAL is detecting peaks, identifying them from the dBase, aligning and gap filling



22

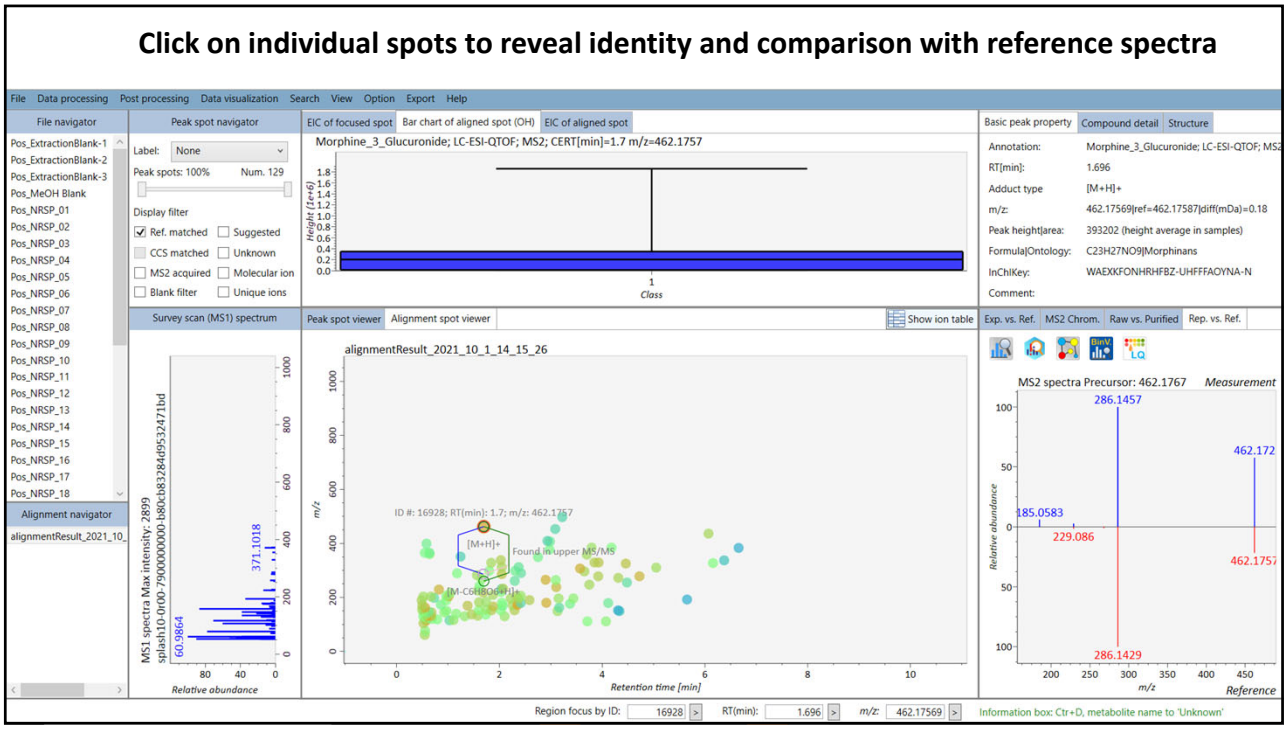

File navigator
Peak spot navigator
EIC of focused spot
Bar chart of aligned spot (OH)
EIC of aligned spot
Basic peak property
Compound detail
Structure
UnknownRT[min]=4.17 m/z=50.0141
Survey scan (MS1) spectrum
Peak spot viewer
Alignment spot viewer
Show ion table
Exp. vs. Ref.
MS2 Chrom.
Raw vs. Purified
Rep. vs. Ref.
alignmentResult_2021_10_1_14_15_26
Measurement vs. Reference
Alignment navigator
Double click this

23

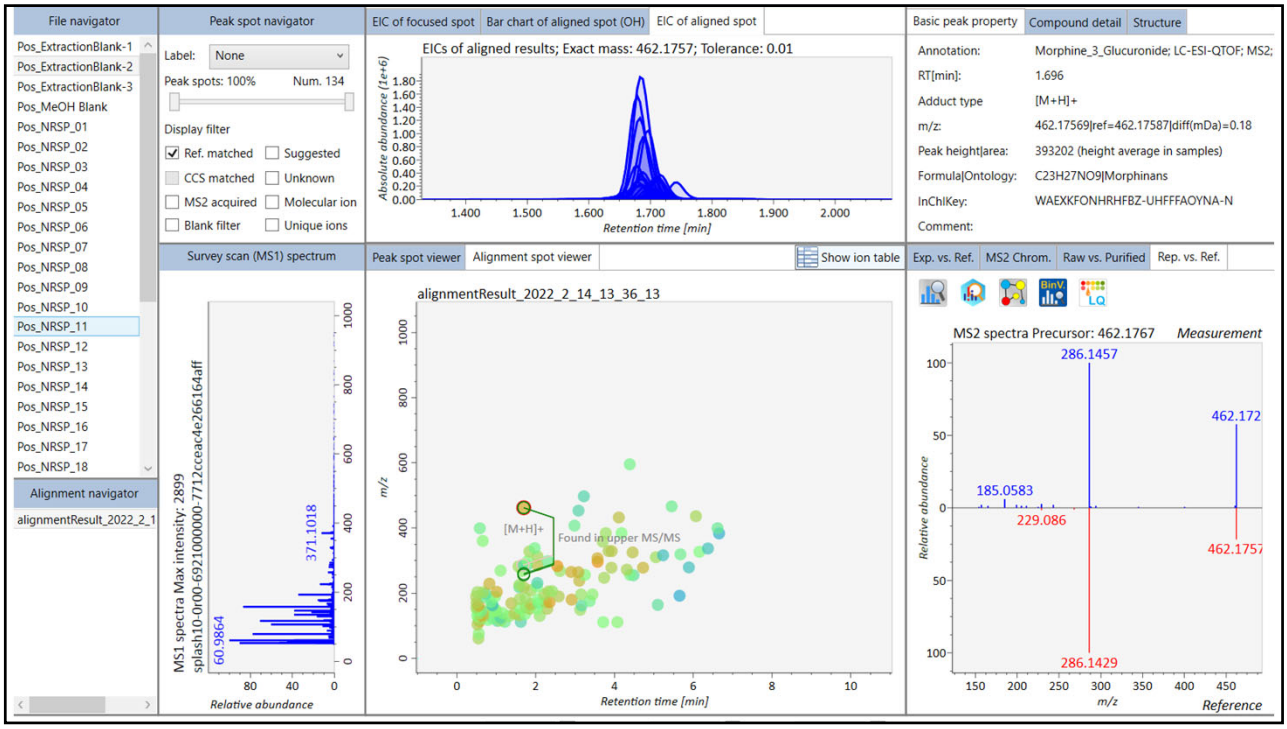

Clicked Ref matched
File navigator
Peak spot navigator
EIC of focused spot
Bar chart of aligned spot (OH)
EIC of aligned spot
Basic peak property
Compound detail
Structure
UnknownRT[min]=4.17 m/z=50.0141
Survey scan (MS1) spectrum
Peak spot viewer
Alignment spot viewer
Show ion table
Exp. vs. Ref.
MS2 Chrom.
Raw vs. Purified
Rep. vs. Ref.
alignmentResult_2021_10_1_14_15_26
Measurement vs. Reference
Alignment navigator

24

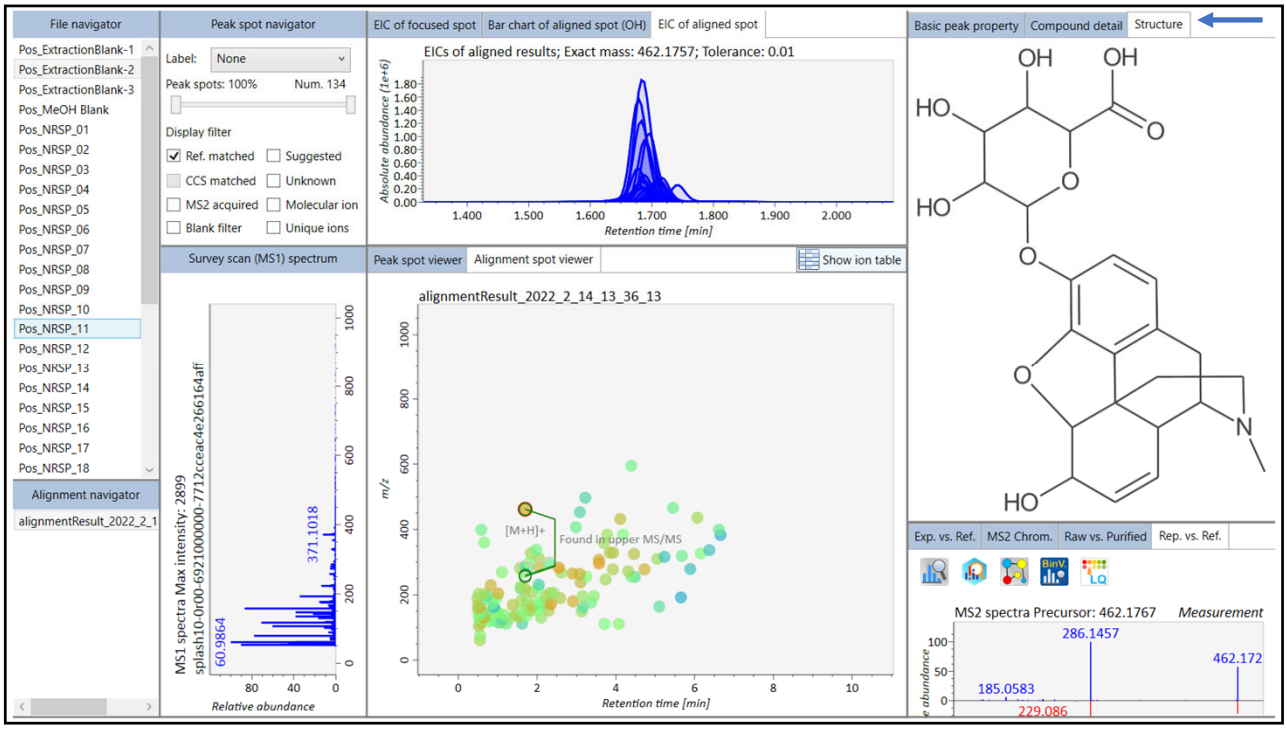
Click on individual spots to reveal identity and comparison with reference spectra



25

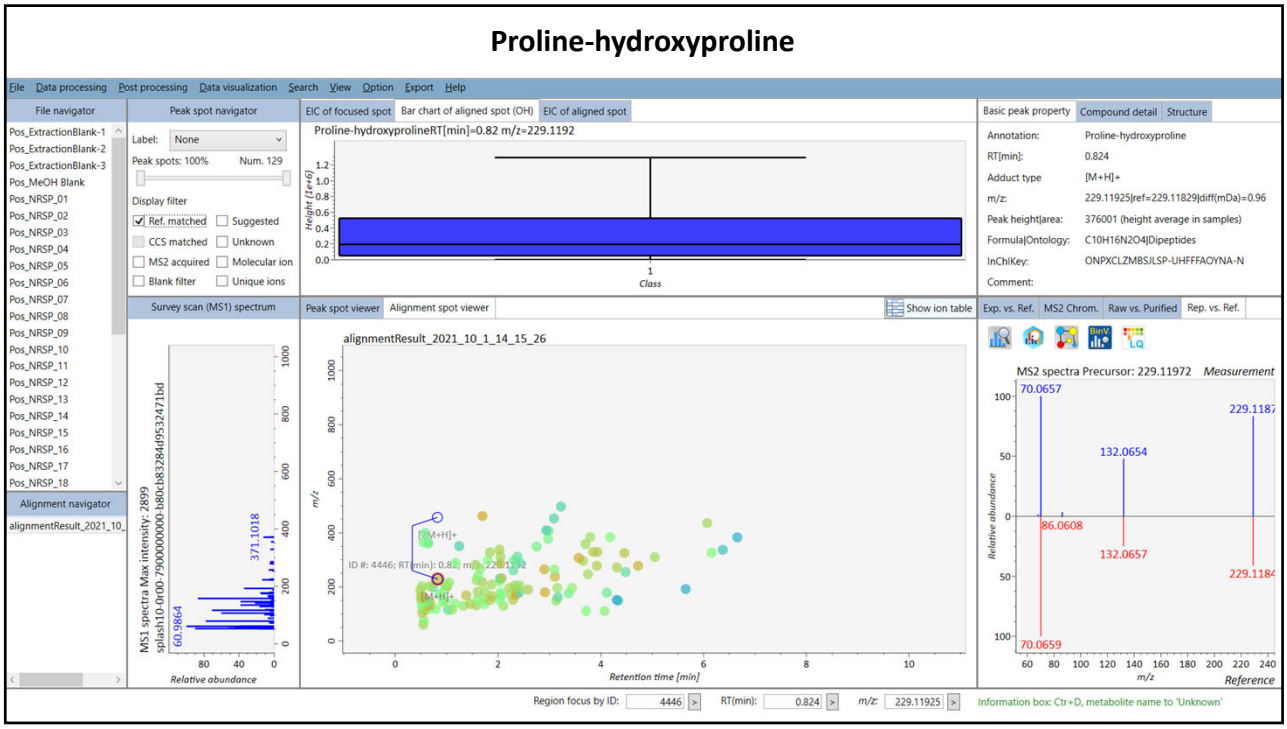


26



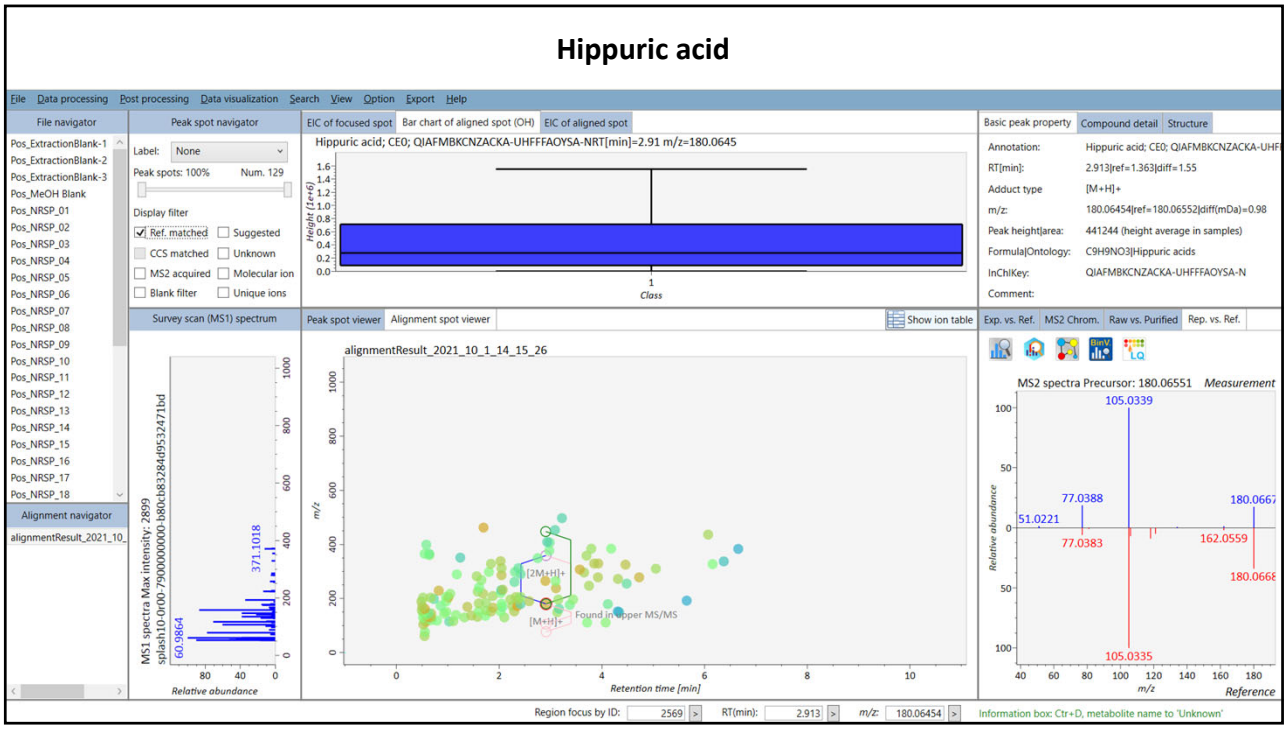
27

Proline-hydroxyproline



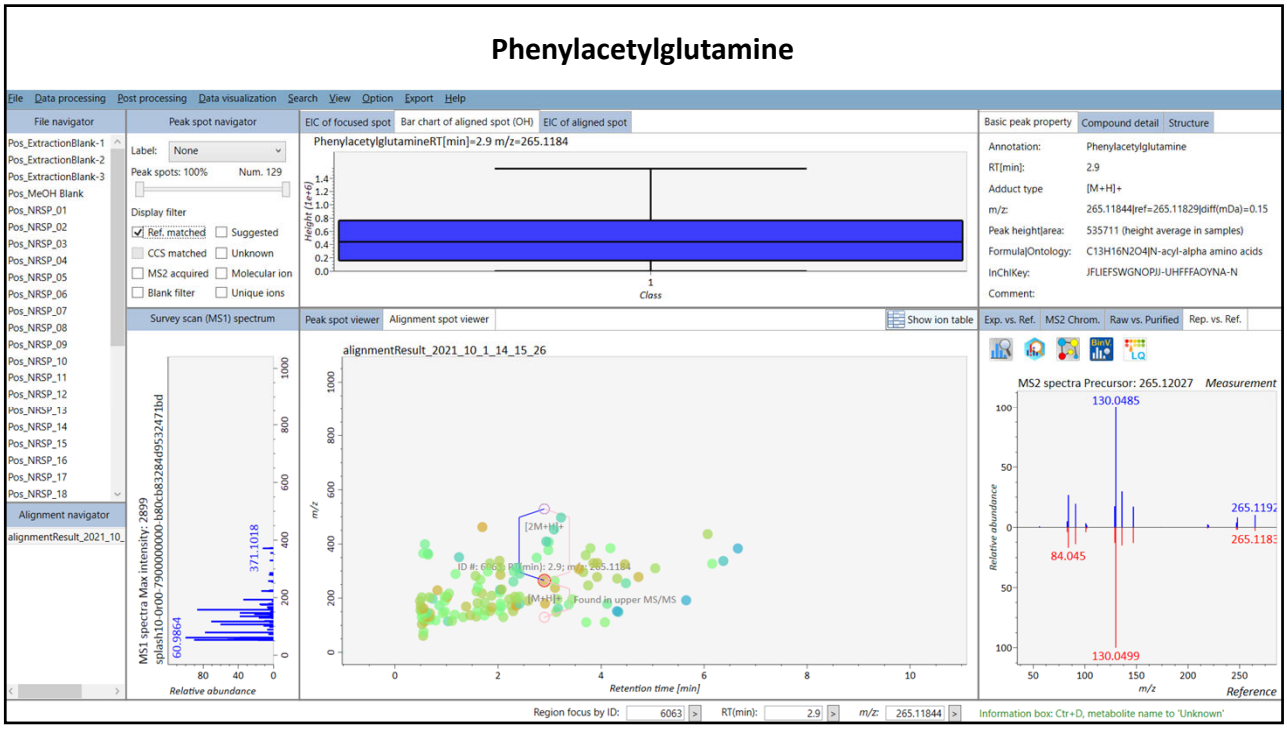
28

Hippuric acid



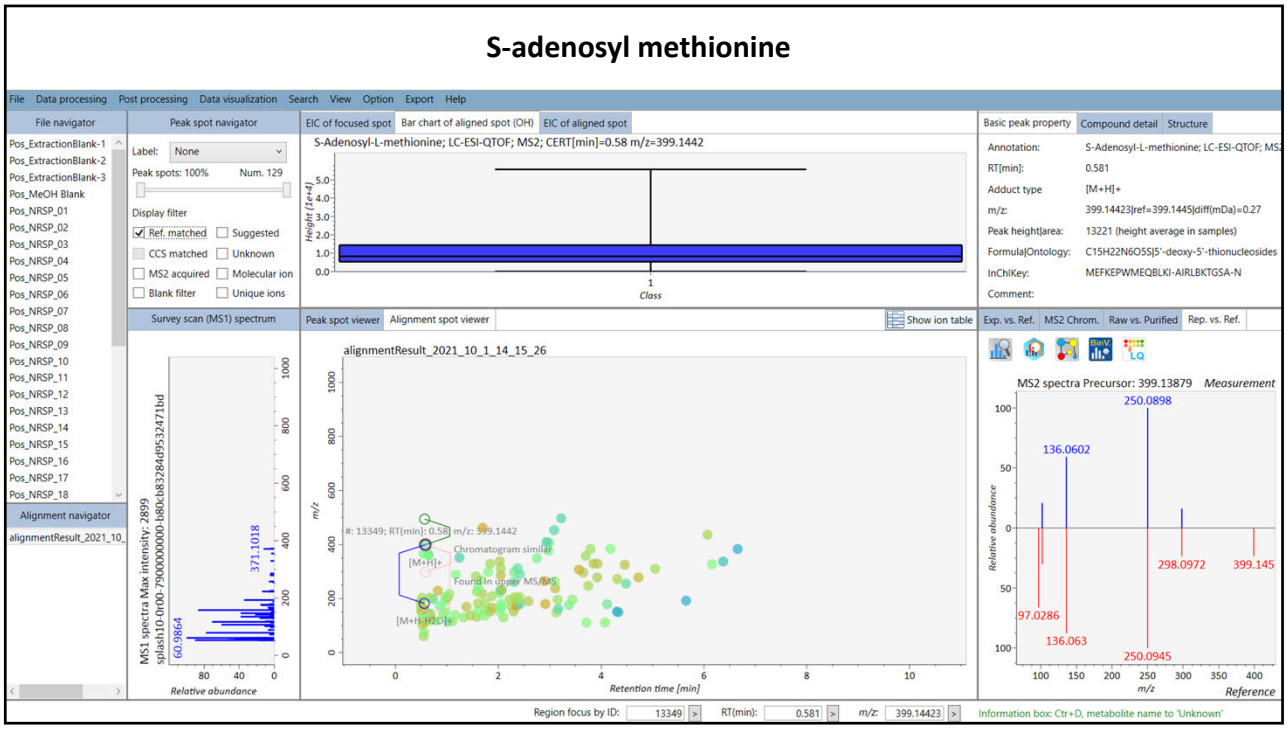
29

Phenylacetylglutamine



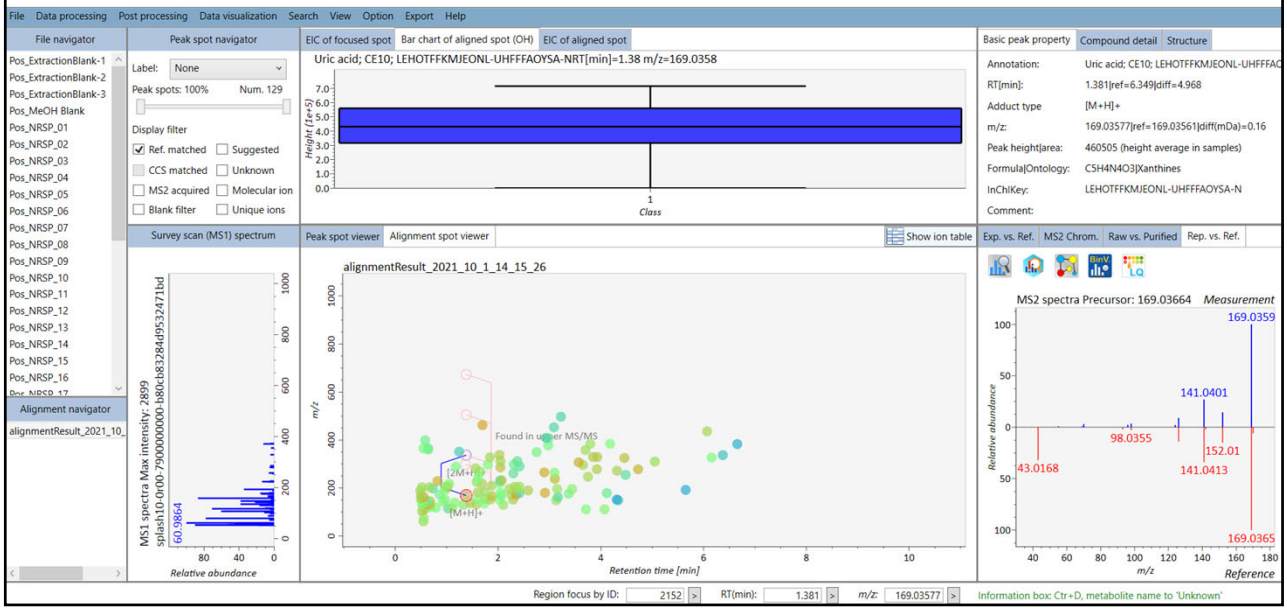
30

S-adenosyl methionine



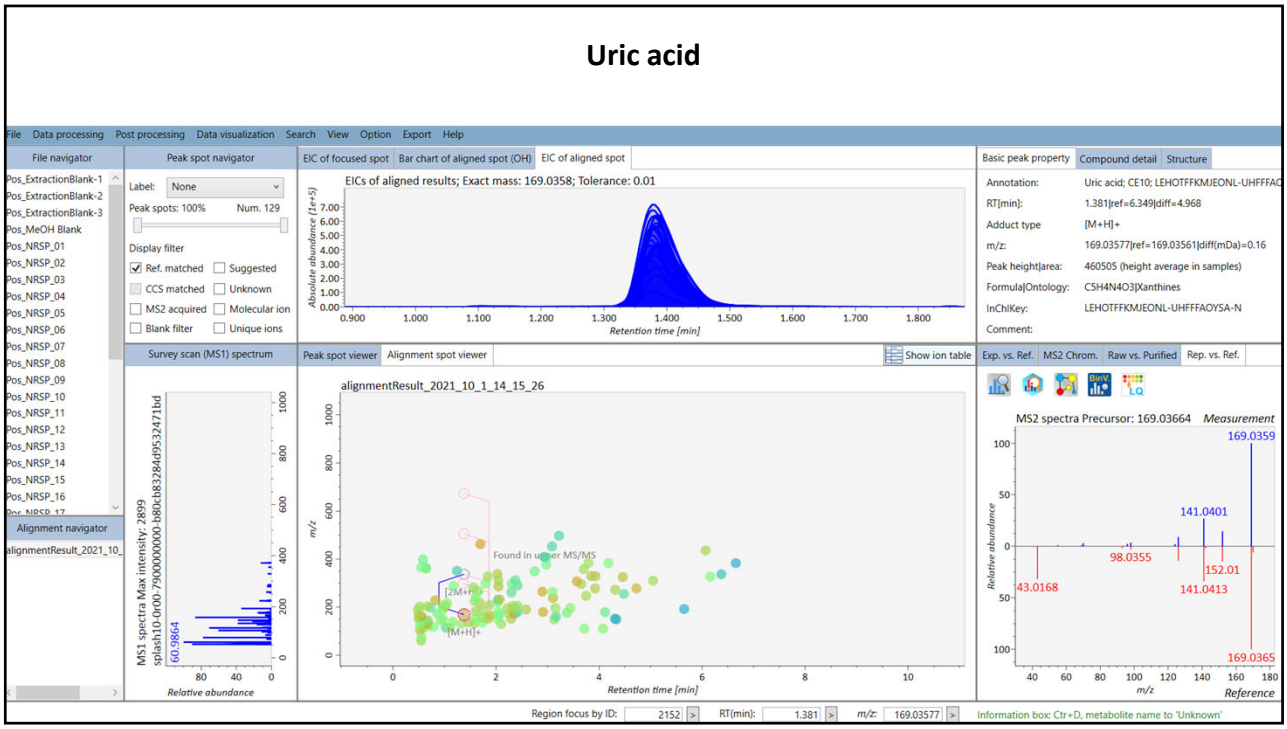
31

Uric acid



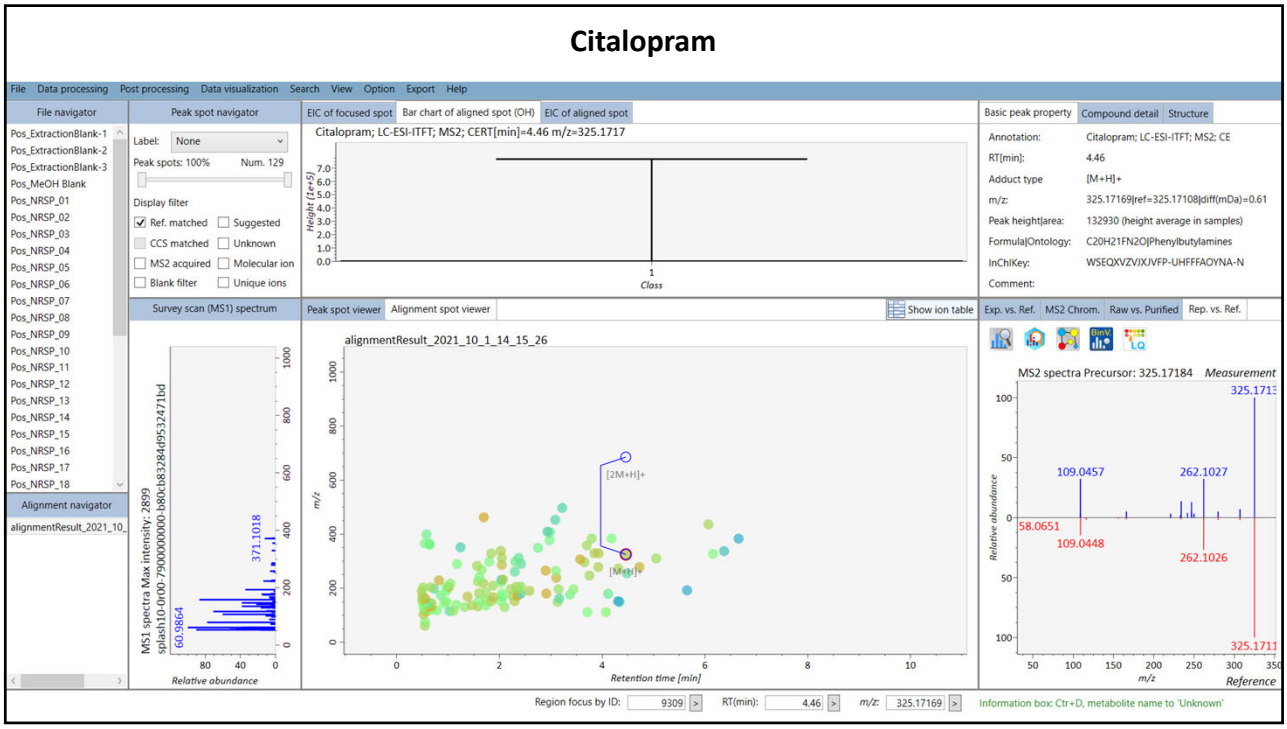
32

Uric acid



33

Citalopram

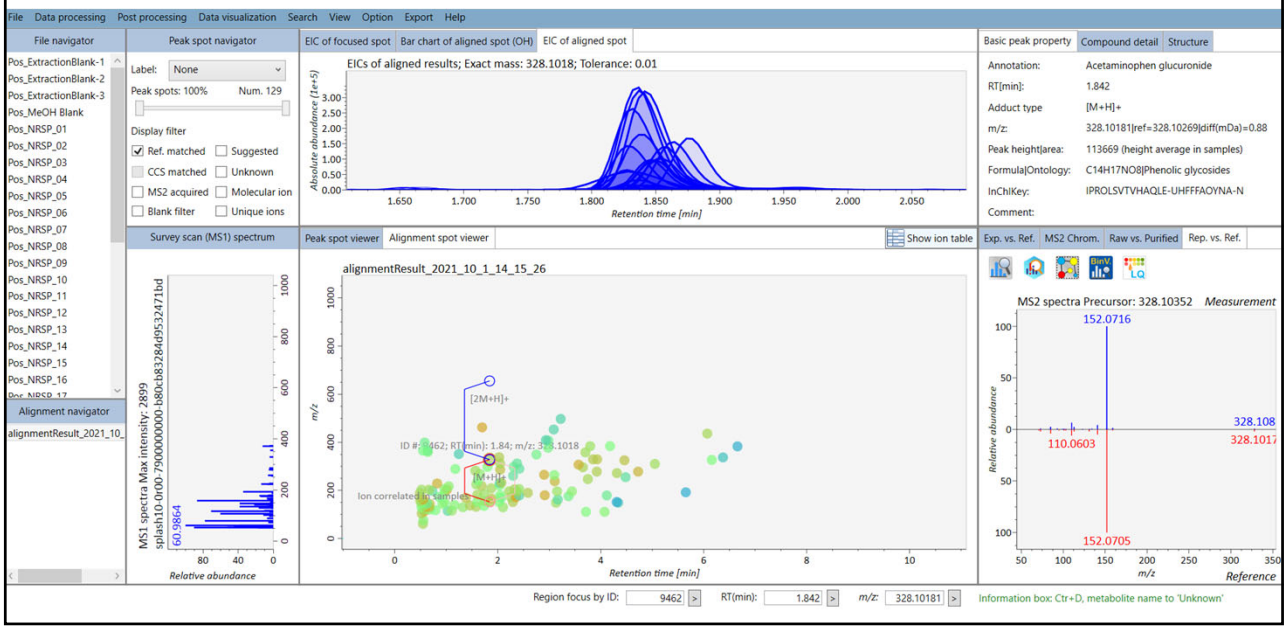


34

Valsartan

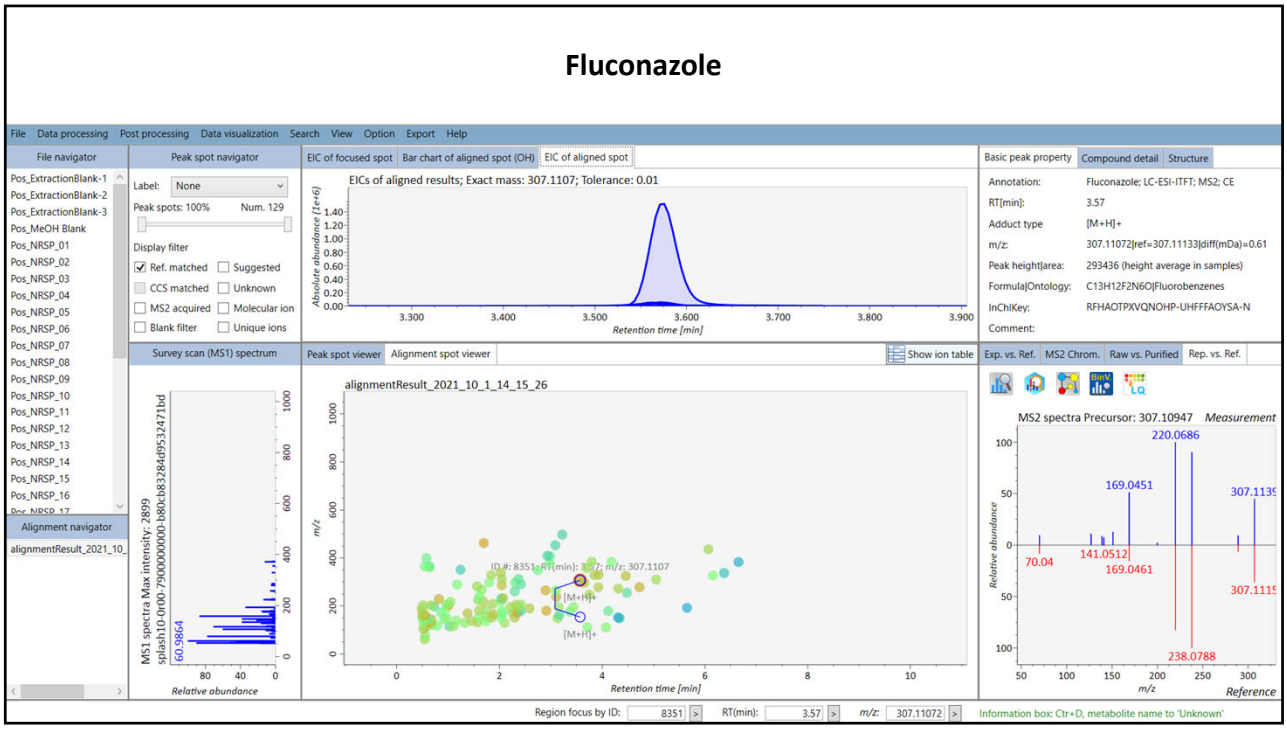
35

β-glucuronide of acetaminophen

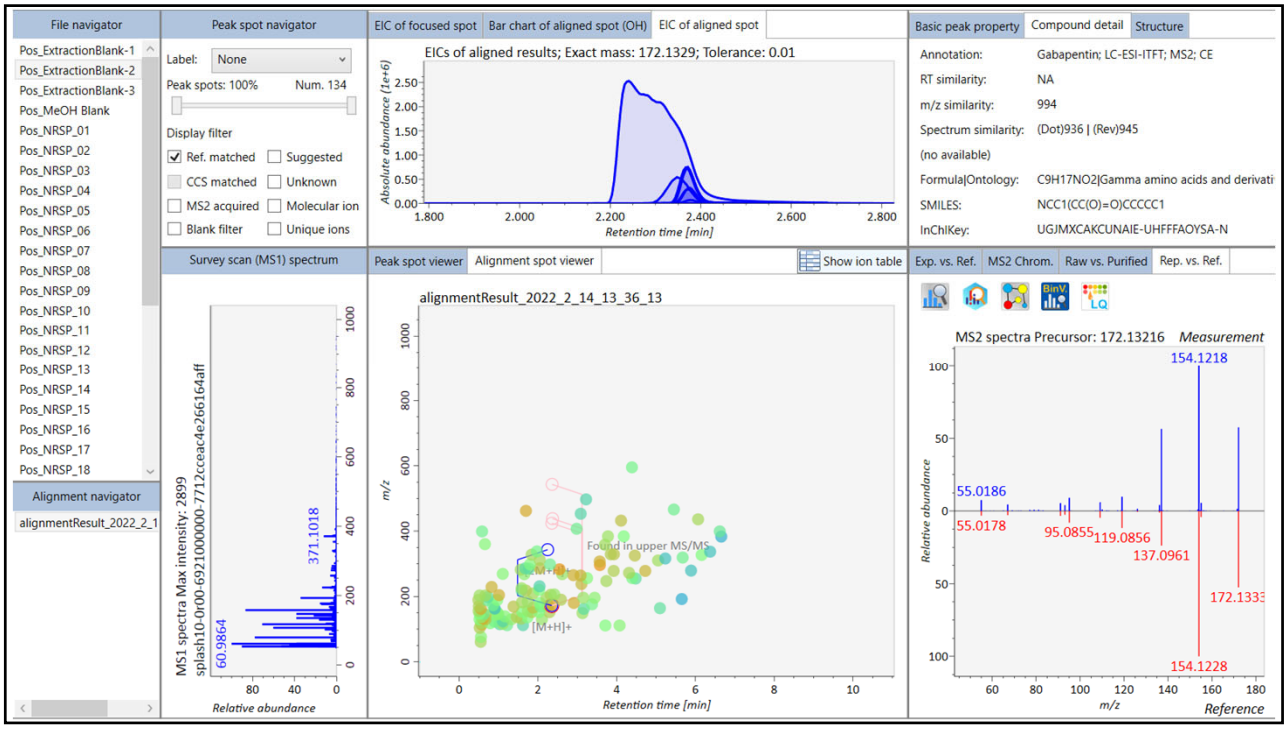


36

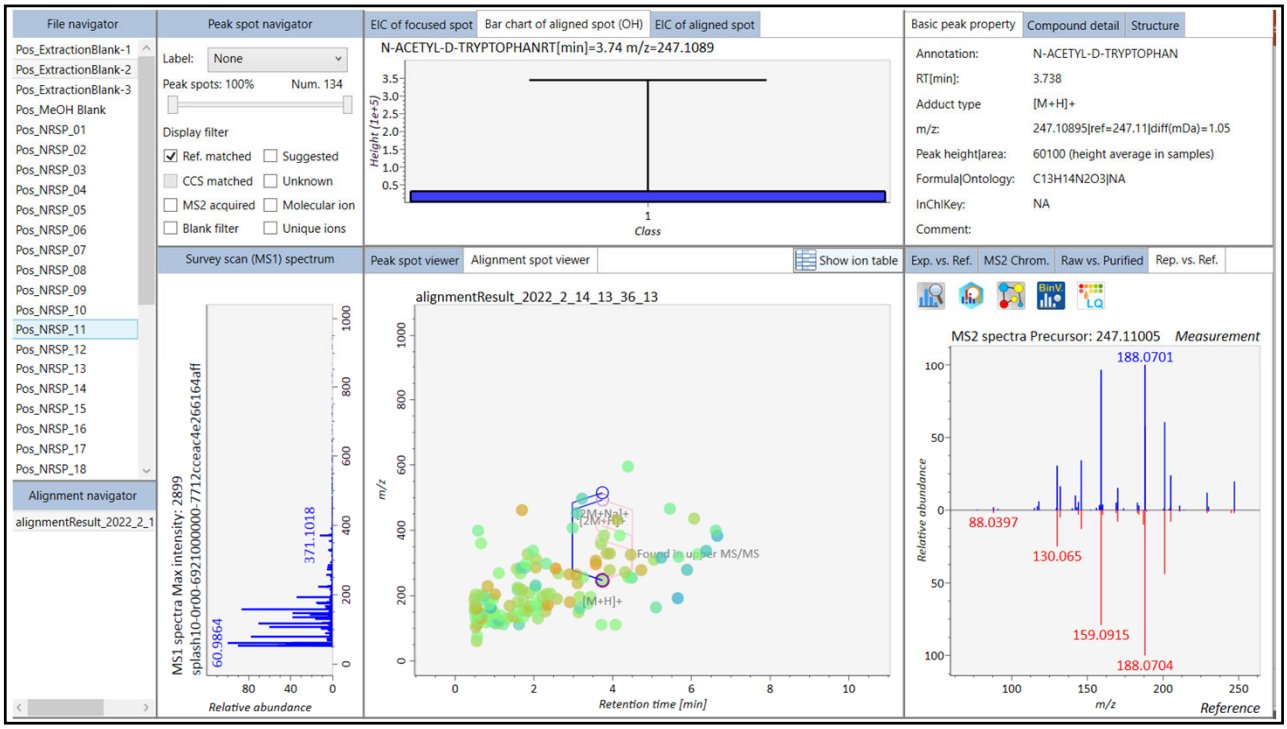
Fluconazole



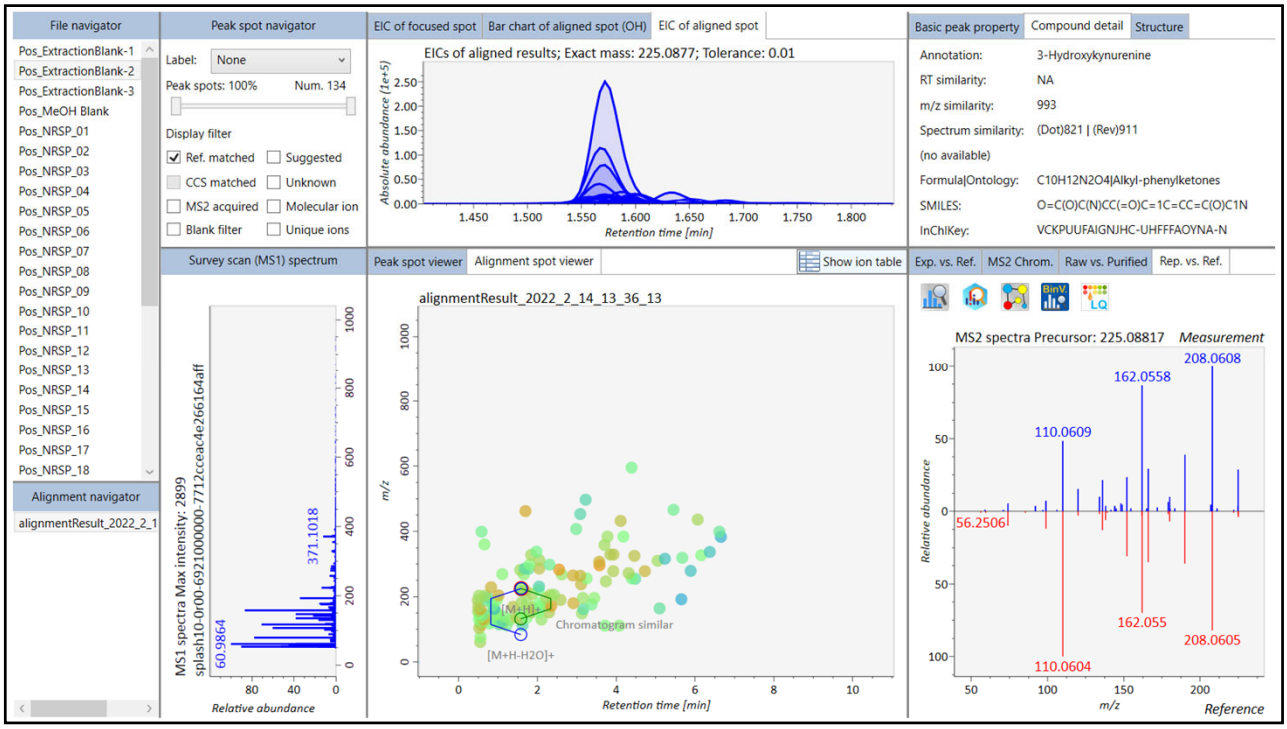
37



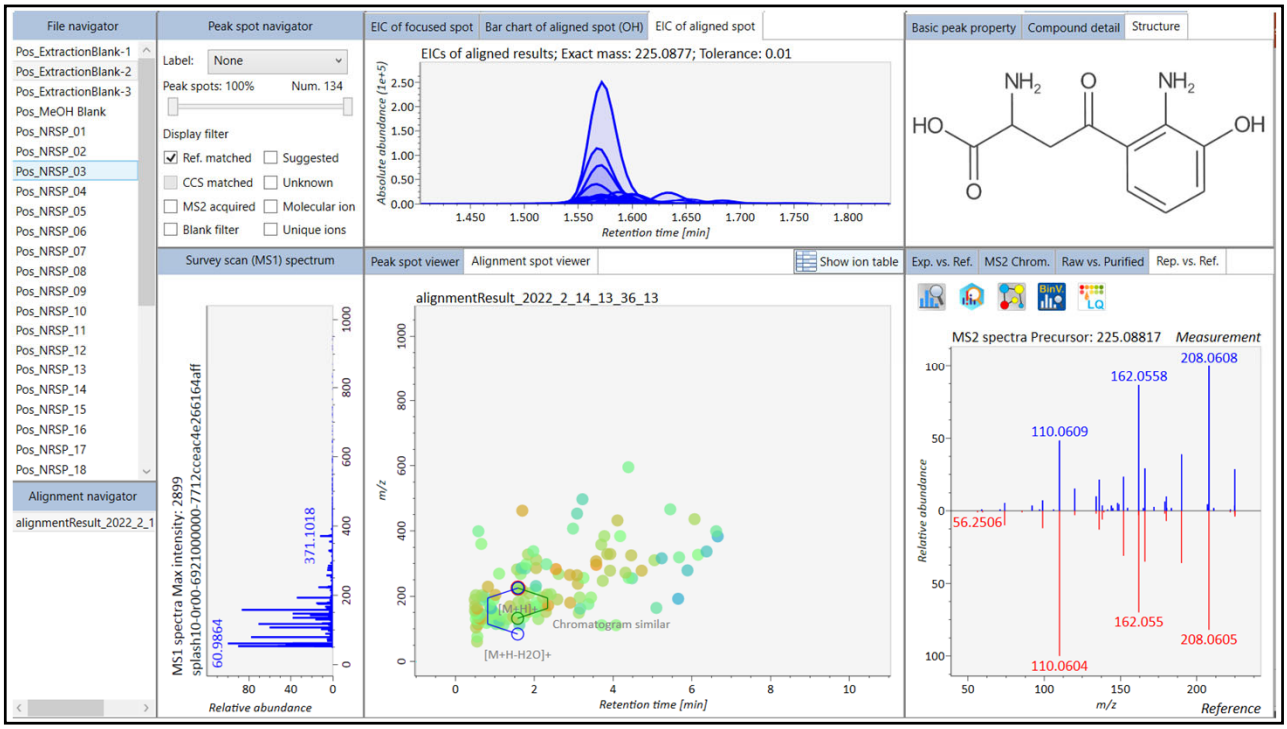
38


File navigator
Pos_ExtractionBlank-1
Pos_ExtractionBlank-2
Pos_ExtractionBlank-3
Pos_MeOH Blank
Pos_NRSP_01
Pos_NRSP_02
Pos_NRSP_03
Pos_NRSP_04
Pos_NRSP_05
Pos_NRSP_06
Pos_NRSP_07
Pos_NRSP_08
Pos_NRSP_09
Pos_NRSP_10
Pos_NRSP_11
Pos_NRSP_12
Pos_NRSP_13
Pos_NRSP_14
Pos_NRSP_15
Pos_NRSP_16
Pos_NRSP_17
Pos_NRSP_18
Alignment navigator
alignmentResult_2022_2_1
Peak spot navigator
Label: None
Peak spots: 100% Num. 134
Display filter
Ref. matched
Suggested
CCS matched
Unknown
MS2 acquired
Molecular ion
Blank filter
Unique ions
Survey scan (MS1) spectrum
MS1 spectra Max intensity: 2899
splash10-0r00-6921000000-7712cceac4e266164aff
371.1018
60.9864
Relative abundance
EIC of focused spot
Bar chart of aligned spot (OH)
EIC of aligned spot
N-ACETYL-D-TRYPTOPHANRT[min]=3.74 m/z=247.1089
Height (1e+5)
Class
Peak spot viewer
Alignment spot viewer
Show ion table
alignmentResult_2022_2_14_13_36_13
m/z
Retention time [min]
[M+Na]+
[2M+Na]+
[M+H]+
Found in upper MS/MS
Basic peak property
Compound detail
Structure
Annotation: N-ACETYL-D-TRYPTOPHAN
RT[min]: 3.738
Adduct type [M+H]+
m/z: 247.10895|ref=247.11|diff(mDa)=1.05
Peak height|area: 60100 (height average in samples)
Formula|Ontology: C13H14N2O3|NA
InChIKey: NA
Comment:
Exp. vs. Ref.
MS2 Chrom.
Raw vs. Purified
Rep. vs. Ref.
MS2 spectra Precursor: 247.11005 Measurement
188.0701
88.0397
130.065
159.0915
188.0704
Relative abundance
m/z
Reference

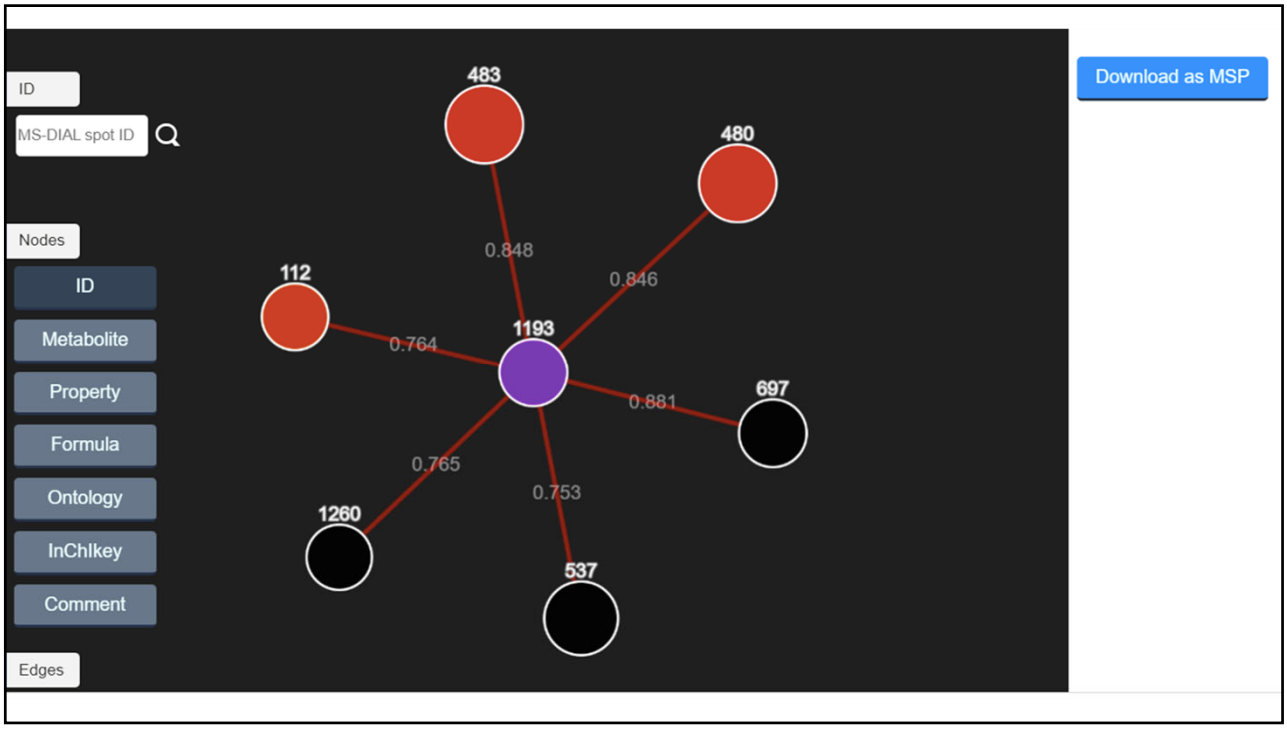
39


File navigator
Pos_ExtractionBlank-1
Pos_ExtractionBlank-2
Pos_ExtractionBlank-3
Pos_MeOH Blank
Pos_NRSP_01
Pos_NRSP_02
Pos_NRSP_03
Pos_NRSP_04
Pos_NRSP_05
Pos_NRSP_06
Pos_NRSP_07
Pos_NRSP_08
Pos_NRSP_09
Pos_NRSP_10
Pos_NRSP_11
Pos_NRSP_12
Pos_NRSP_13
Pos_NRSP_14
Pos_NRSP_15
Pos_NRSP_16
Pos_NRSP_17
Pos_NRSP_18
Alignment navigator
alignmentResult_2022_2_1
Peak spot navigator
Label: None
Peak spots: 100% Num. 134
Display filter
Ref. matched
Suggested
CCS matched
Unknown
MS2 acquired
Molecular ion
Blank filter
Unique ions
Survey scan (MS1) spectrum
MS1 spectra Max intensity: 2899
splash10-0r00-6921000000-7712cceac4e266164aff
371.1018
60.9864
Relative abundance
EIC of focused spot
Bar chart of aligned spot (OH)
EIC of aligned spot
EICs of aligned results; Exact mass: 225.0877; Tolerance: 0.01
Absolute abundance (1e+5)
Retention time [min]
Peak spot viewer
Alignment spot viewer
Show ion table
alignmentResult_2022_2_14_13_36_13
m/z
Retention time [min]
[M+H]+
Chromatogram similar
[M+H-H2O]+
Basic peak property
Compound detail
Structure
Annotation: 3-Hydroxykynurenine
RT similarity: NA
m/z similarity: 993
Spectrum similarity: (Dot)821 | (Rev)911
(no available)
Formula|Ontology: C10H12N2O4|Alkyl-phenylketones
SMILES: O=C(O)C(N)CC(=O)C=1C=CC=C(O)C1N
InChIKey: VCKPUUFAIGNJHC-UHFFFAOYNA-N
Exp. vs. Ref.
MS2 Chrom.
Raw vs. Purified
Rep. vs. Ref.
MS2 spectra Precursor: 225.08817 Measurement
208.0608
162.0558
110.0609
56.2506
162.055
208.0605
110.0604
Relative abundance
m/z
Reference

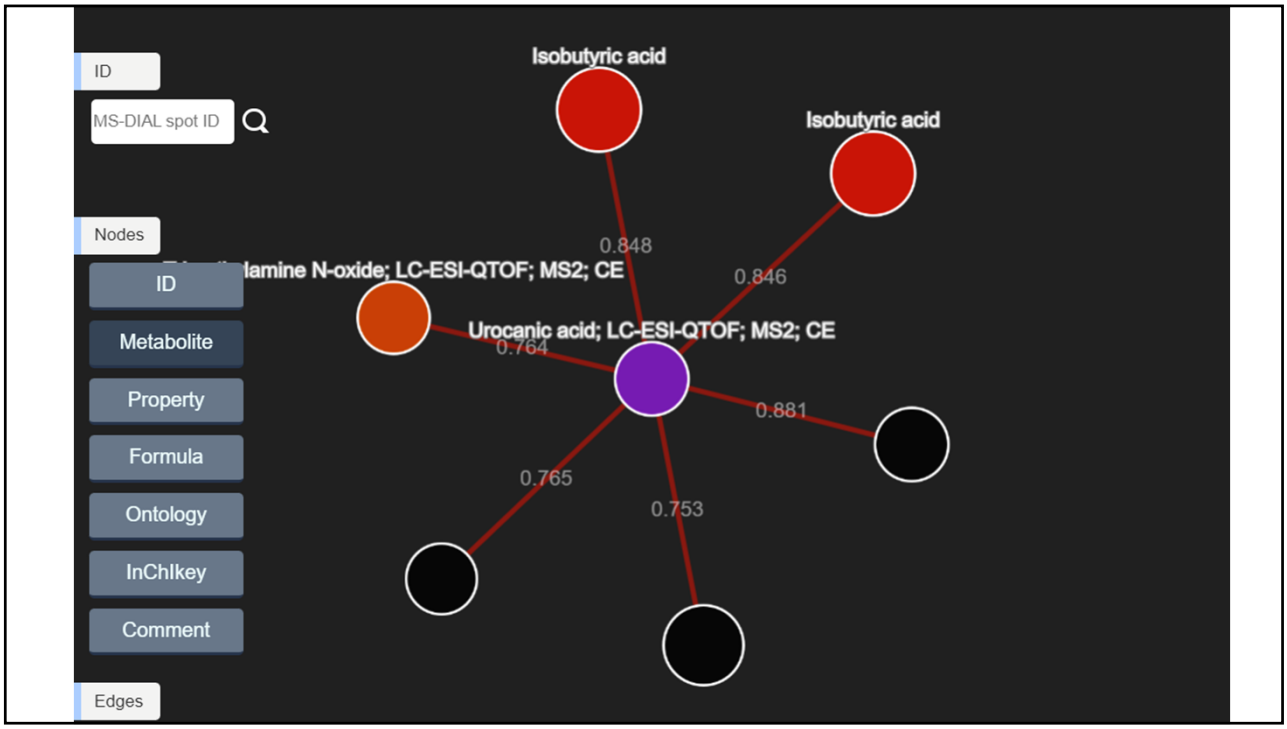
40


File navigator
Pos_ExtractionBlank-1
Pos_ExtractionBlank-2
Pos_ExtractionBlank-3
Pos_MeOH Blank
Pos_NRSP_01
Pos_NRSP_02
Pos_NRSP_03
Pos_NRSP_04
Pos_NRSP_05
Pos_NRSP_06
Pos_NRSP_07
Pos_NRSP_08
Pos_NRSP_09
Pos_NRSP_10
Pos_NRSP_11
Pos_NRSP_12
Pos_NRSP_13
Pos_NRSP_14
Pos_NRSP_15
Pos_NRSP_16
Pos_NRSP_17
Pos_NRSP_18
Alignment navigator
alignmentResult_2022_2_1
Peak spot navigator
Label: None
Peak spots: 100% Num. 134
Display filter
Ref. matched
Suggested
CCS matched
Unknown
MS2 acquired
Molecular ion
Blank filter
Unique ions
EIC of focused spot
Bar chart of aligned spot (OH)
EIC of aligned spot
EICs of aligned results; Exact mass: 225.0877; Tolerance: 0.01
Absolute abundance (1e+5)
Retention time [min]
Basic peak property
Compound detail
Structure
Survey scan (MS1) spectrum
MS1 spectra Max intensity: 2899
splash10-0r00-6921000000-7712cceac4e266164aff
371.1018
60.9864
Relative abundance
Peak spot viewer
Alignment spot viewer
Show ion table
alignmentResult_2022_2_14_13_36_13
[M-H]-
Chromatogram similar
[M+H-H2O]+
m/z
Retention time [min]
Exp. vs. Ref.
MS2 Chrom.
Raw vs. Purified
Rep. vs. Ref.
MS2 spectra Precursor: 225.08817 Measurement
208.0608
162.0558
110.0609
56.2506
162.055
208.0605
110.0604
Relative abundance
m/z
Reference

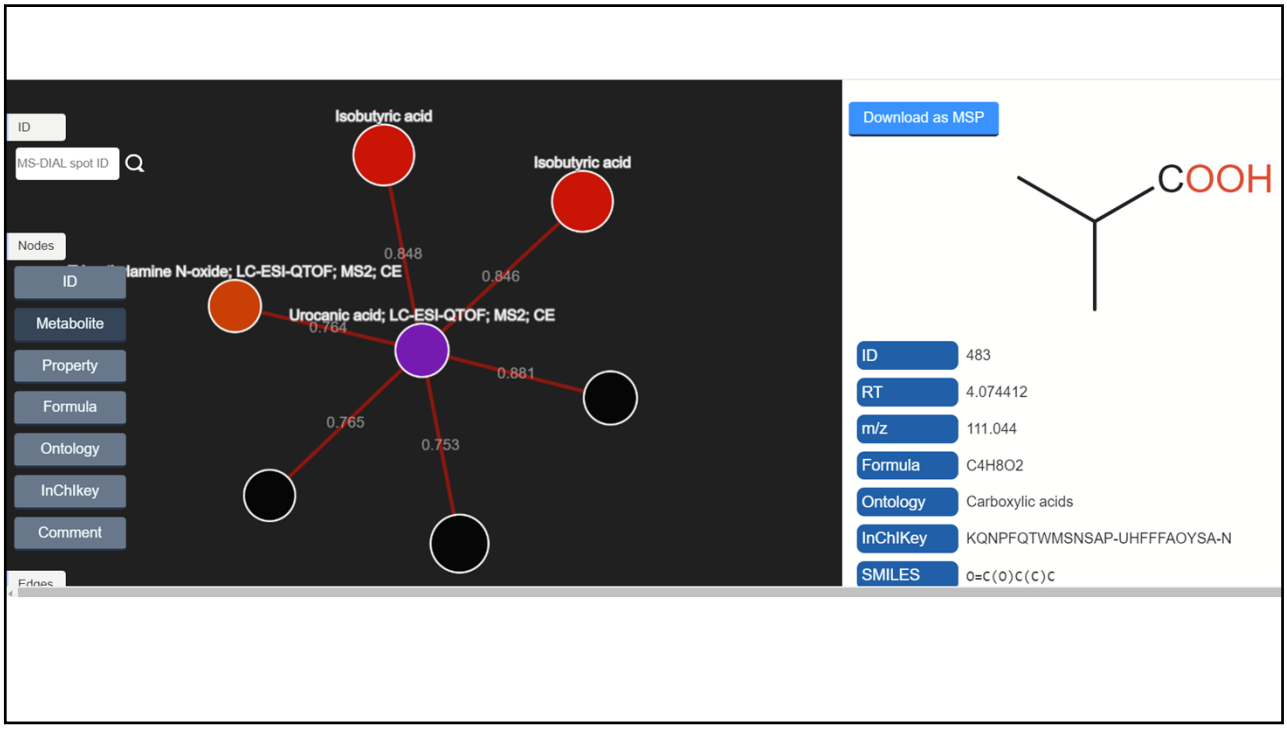
41


ID
MS-DIAL spot ID
Nodes
ID
Metabolite
Property
Formula
Ontology
InChIkey
Comment
Edges
Download as MSP
483
480
112
1193
697
1260
537
0.848
0.846
0.764
0.881
0.765
0.753

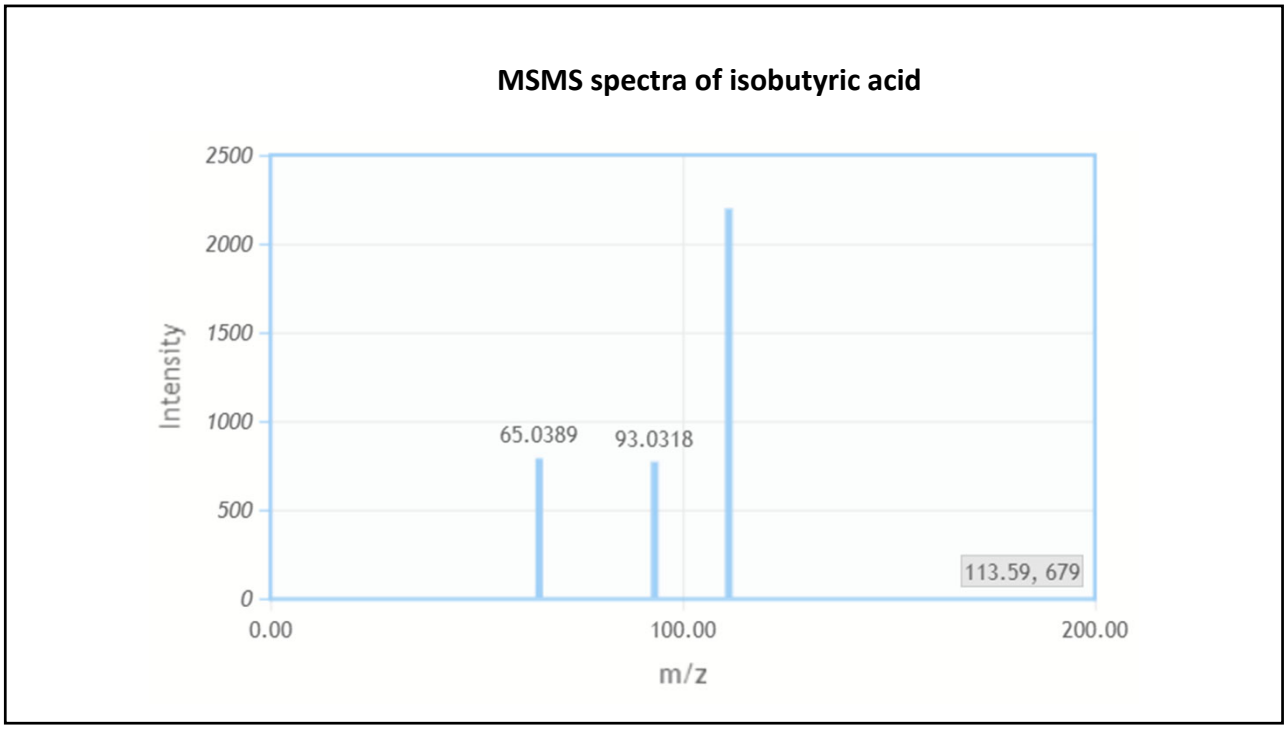
42


ID
MS-DIAL spot ID
Nodes
ID
Metabolite
Property
Formula
Ontology
InChIkey
Comment
Edges
Isobutyric acid
Isobutyric acid
0.848
0.846
lamine N-oxide; LC-ESI-QTOF; MS2; CE
Urocanic acid; LC-ESI-QTOF; MS2; CE
0.764
0.881
0.765
0.753

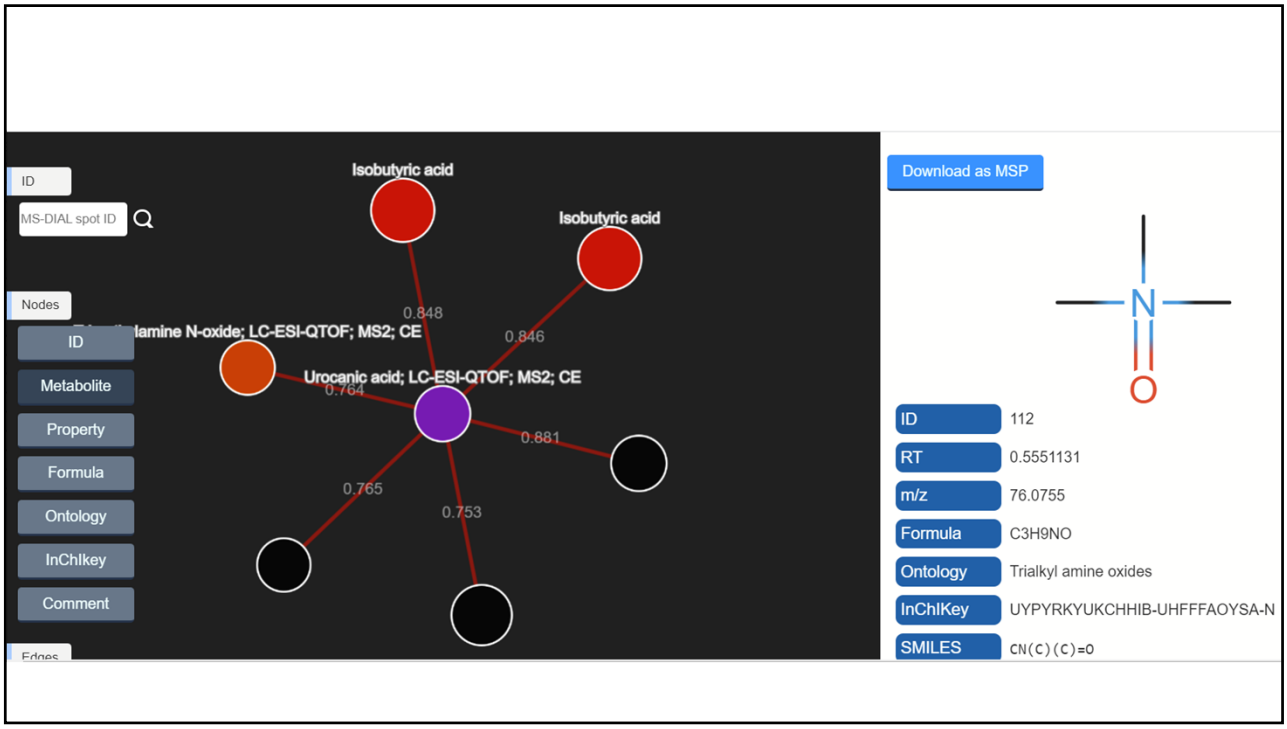
43


ID
MS-DIAL spot ID
Nodes
ID
Metabolite
Property
Formula
Ontology
InChIkey
Comment
Edges
Isobutyric acid
Isobutyric acid
0.848
0.846
lamine N-oxide; LC-ESI-QTOF; MS2; CE
Urocanic acid; LC-ESI-QTOF; MS2; CE
0.764
0.881
0.765
0.753
Download as MSP
COOH
ID 483
RT 4.074412
m/z 111.044
Formula C4H8O2
Ontology Carboxylic acids
InChIKey KQNPFQTWMSNSAP-UHFFFAOYSA-N
SMILES O=C(O)C(C)C

44

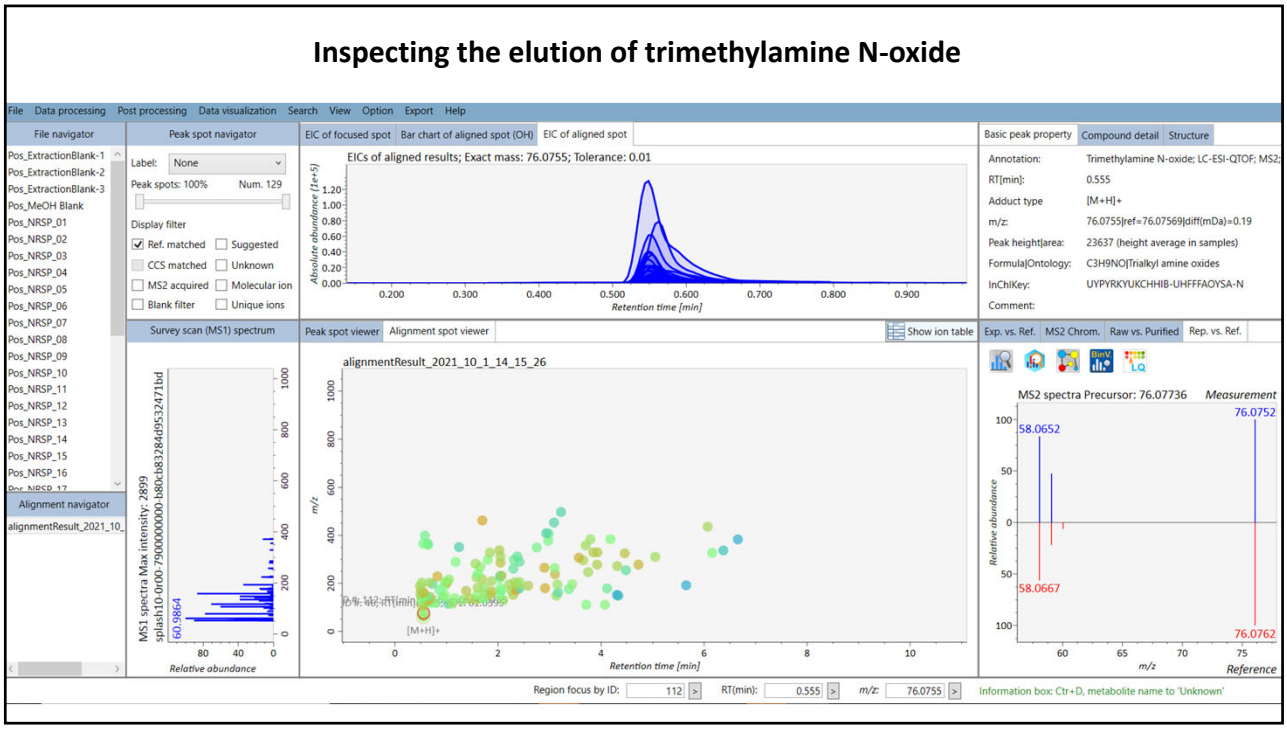

MSMS spectra of isobutyric acid
2500
2000
1500
1000
500
0
Intensity
65.0389
93.0318
113.59, 679
0.00
100.00
200.00
m/z

45


ID
MS-DIAL spot ID
Nodes
ID
Metabolite
Property
Formula
Ontology
InChIkey
Comment
Edges
Isobutyric acid
Isobutyric acid
lamine N-oxide; LC-ESI-QTOF; MS2; CE
Urocanic acid; LC-ESI-QTOF; MS2; CE
0.848
0.846
0.764
0.881
0.765
0.753
Download as MSP
N
O
ID 112
RT 0.5551131
m/z 76.0755
Formula C3H9NO
Ontology Trialkyl amine oxides
InChIKey UYPYRKYUKCHHIB-UHFFFAOYSA-N
SMILES CN(C)(C)=O

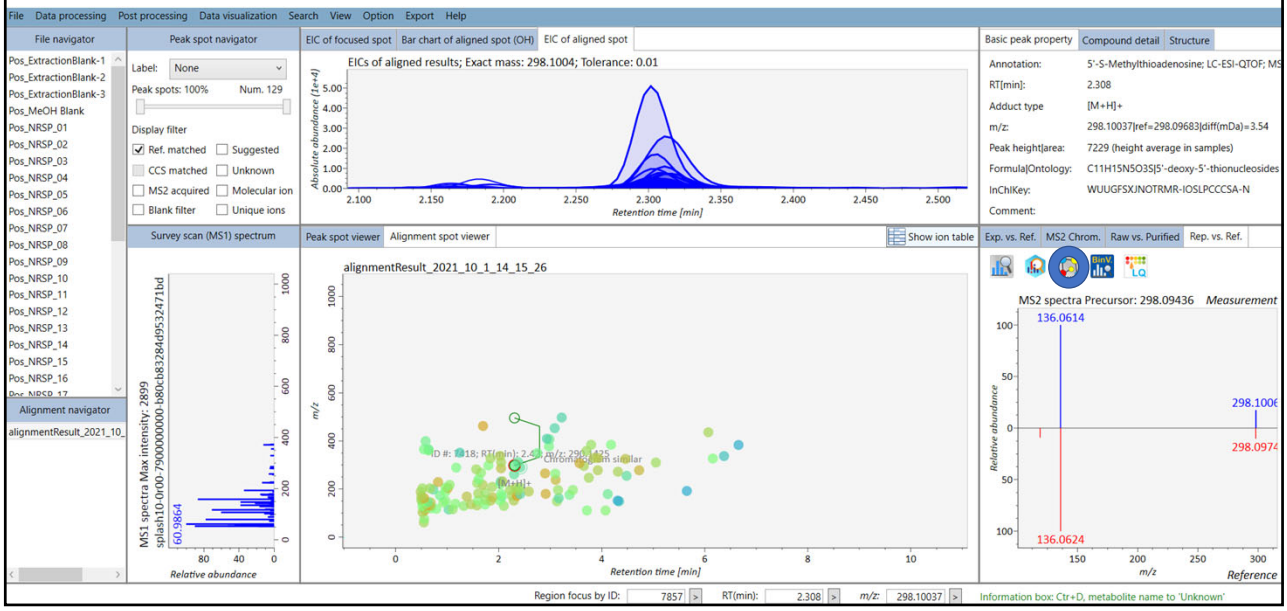
46

Inspecting the elution of trimethylamine N-oxide

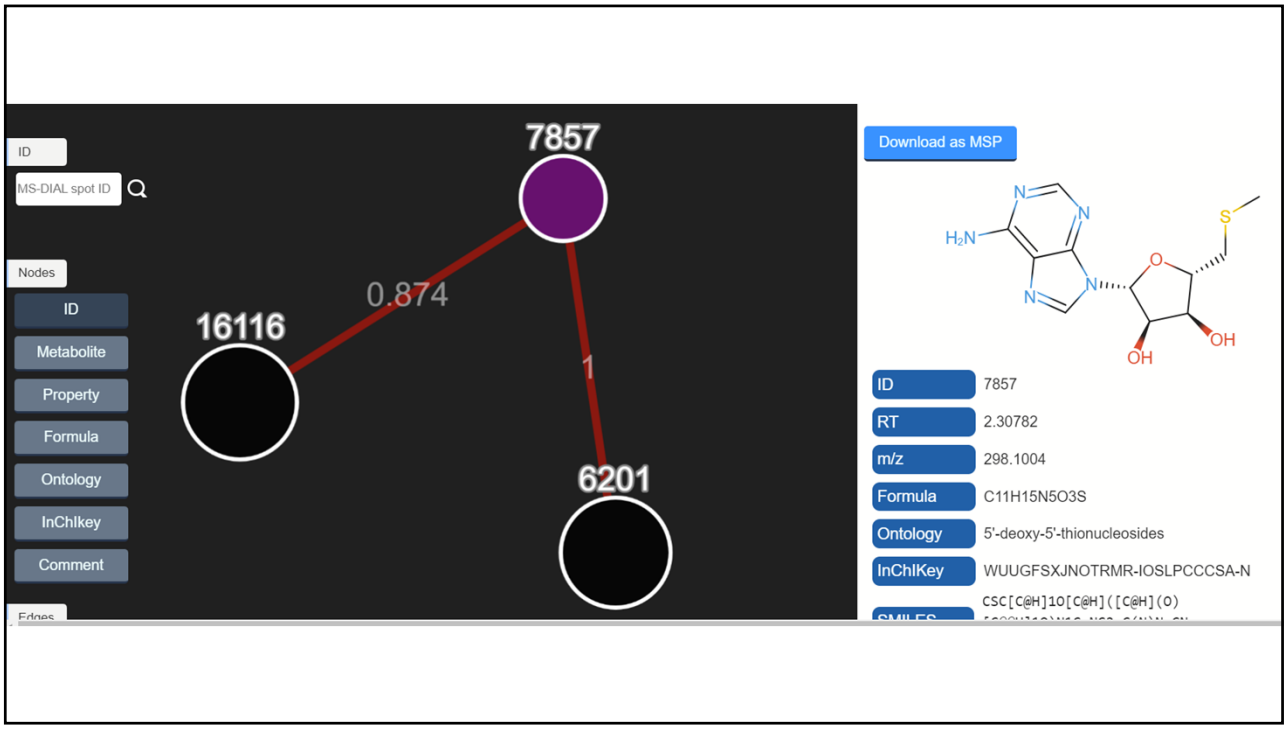


47

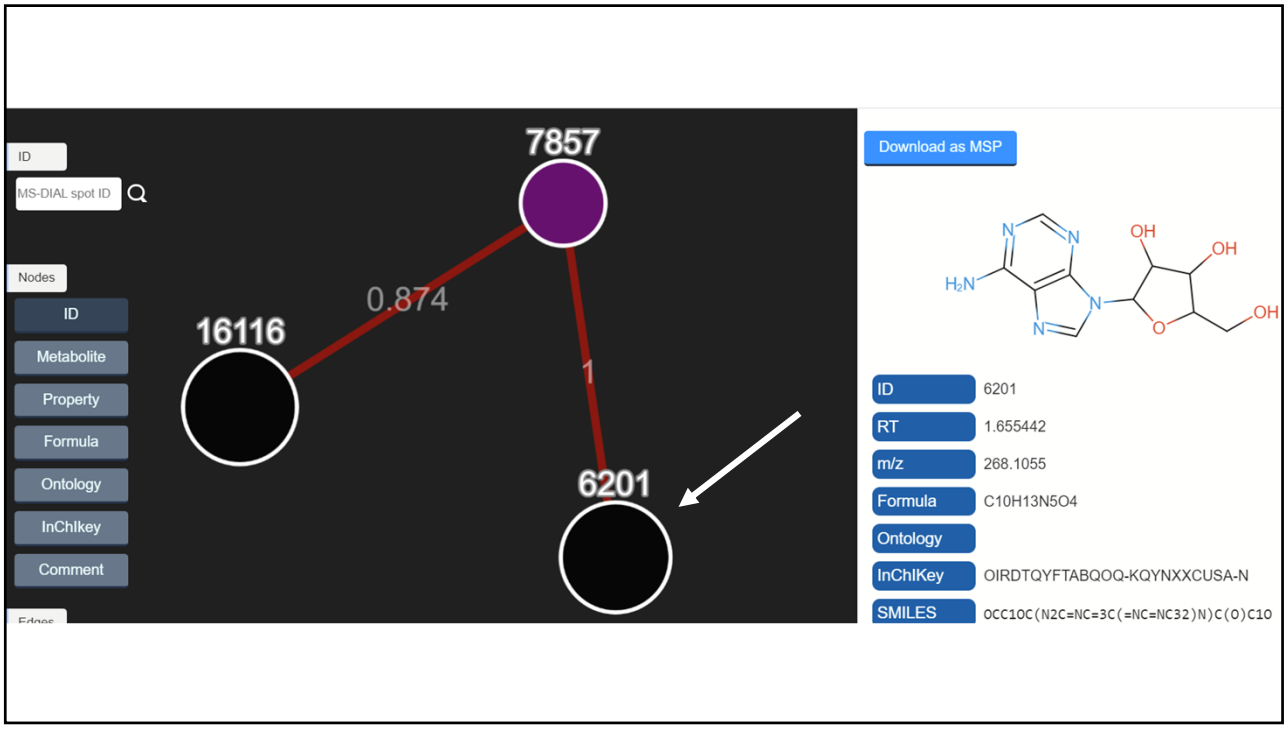
5-5’-methylthioadenosine



48

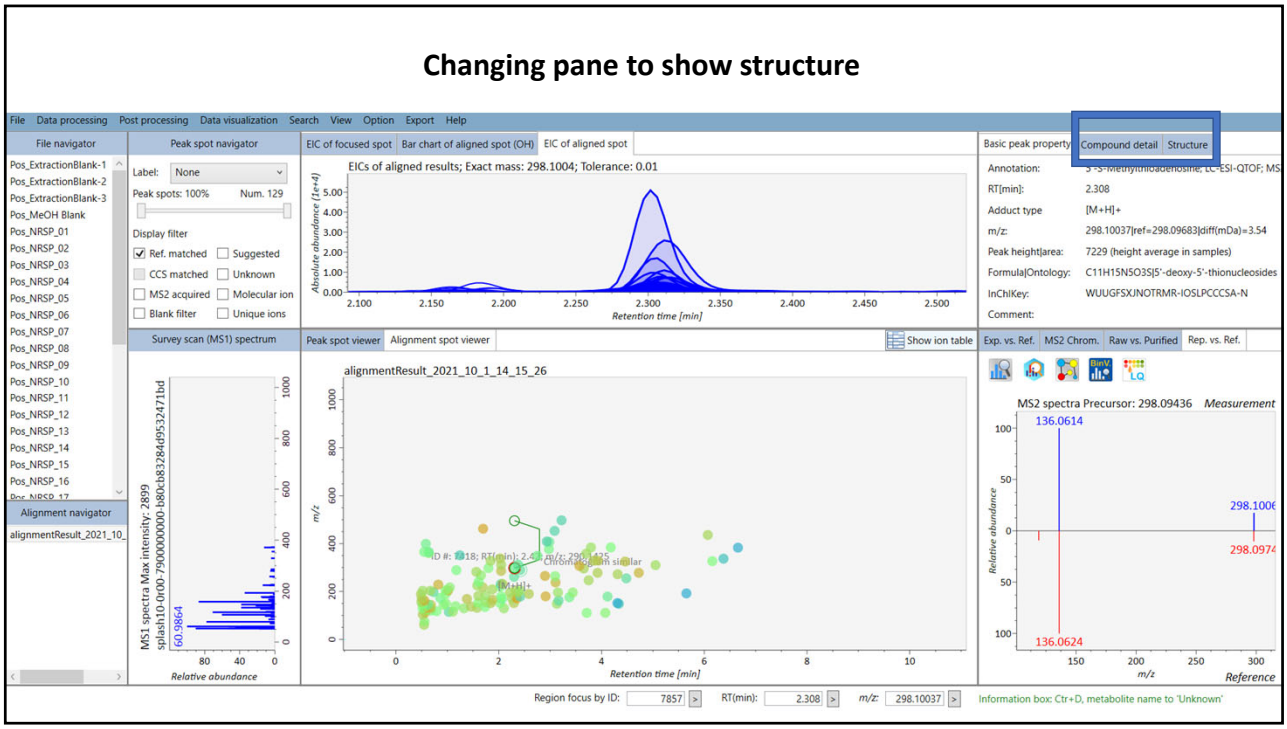

ID
MS-DIAL spot ID
Nodes
ID
Metabolite
Property
Formula
Ontology
InChIkey
Comment
Edges
7857
0.874
16116
1
6201
Download as MSP
ID 7857
RT 2.30782
m/z 298.1004
Formula C11H15N5O3S
Ontology 5'-deoxy-5'-thionucleosides
InChIKey WUUGFSXJNOTRMR-IOSLPCCCSA-N
SMILES CSC[C@H]1O[C@H]([C@H](O)

49

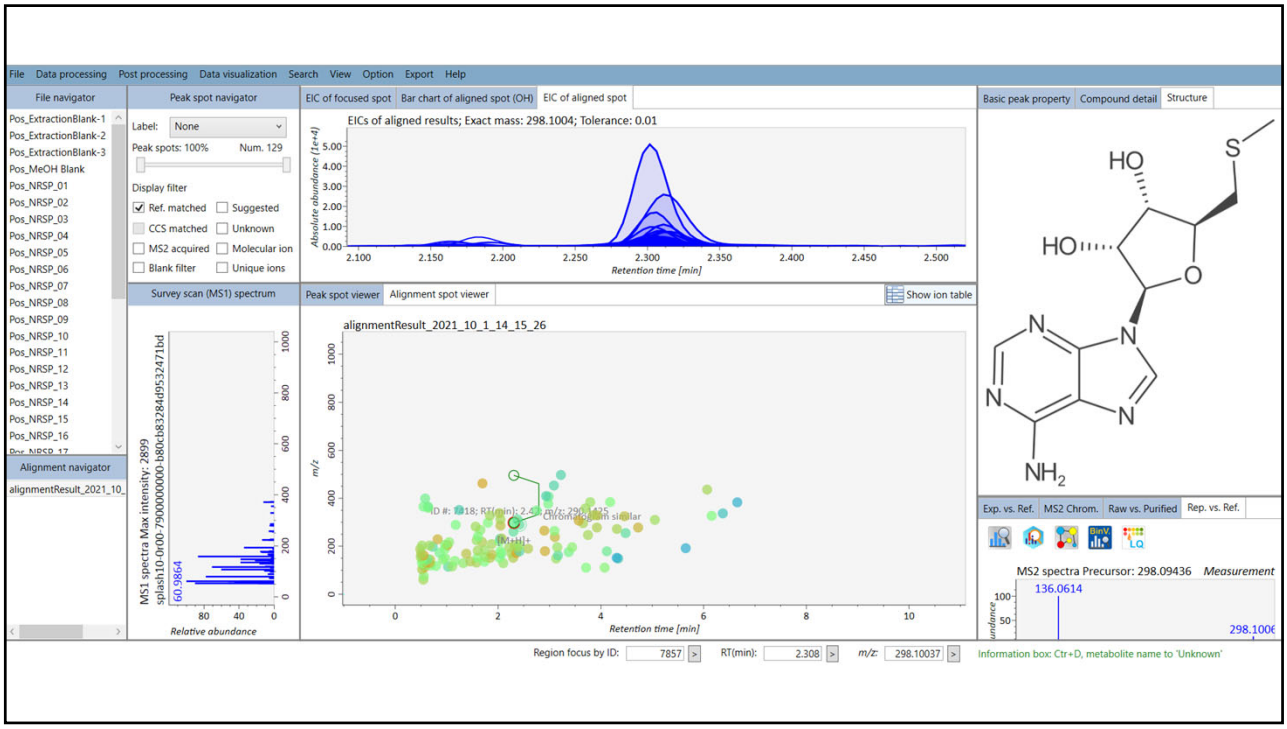

ID
MS-DIAL spot ID
Nodes
ID
Metabolite
Property
Formula
Ontology
InChIkey
Comment
Edges
7857
0.874
16116
1
6201
Download as MSP
ID 6201
RT 1.655442
m/z 268.1055
Formula C10H13N5O4
Ontology
InChIKey OIRDTQYFTABQOQ-KQYNXXCUSA-N
SMILES OCC1OC(N2C=NC=3C(=NC=NC32)N)C(O)C1O

50

Changing pane to show structure

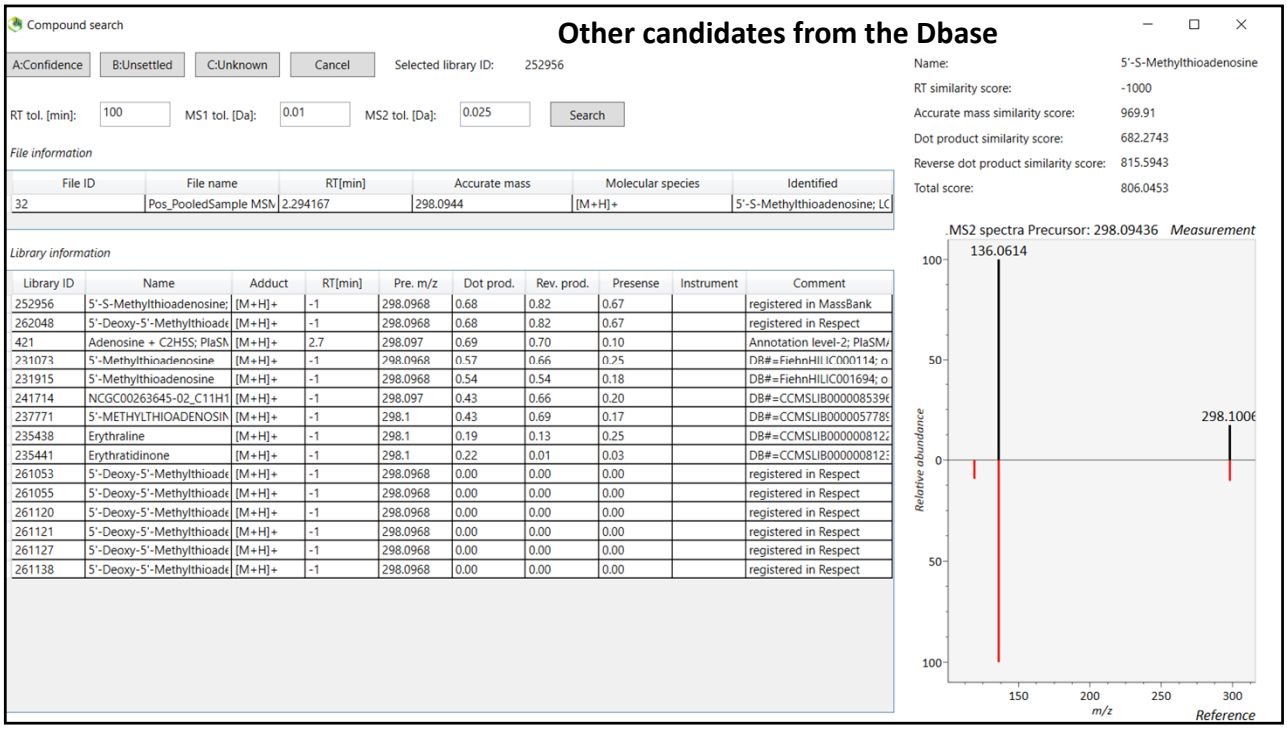


51



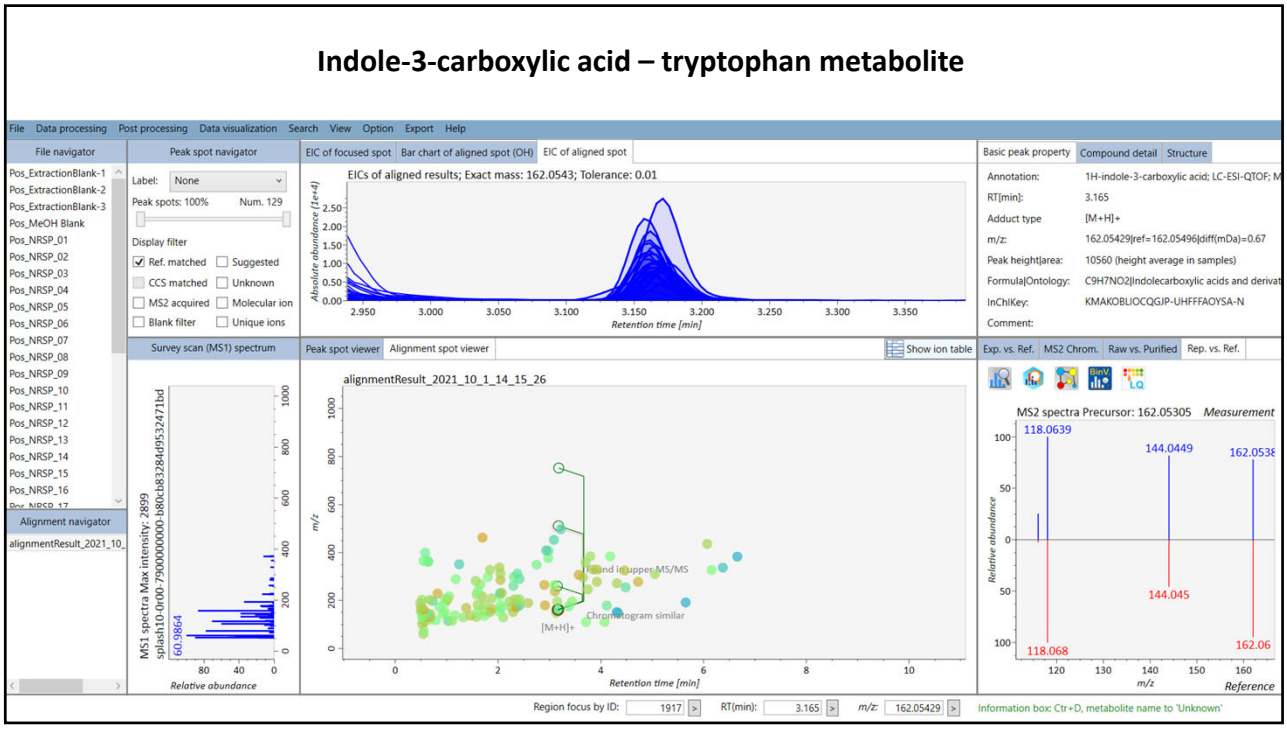
52

Other candidates from the Dbase



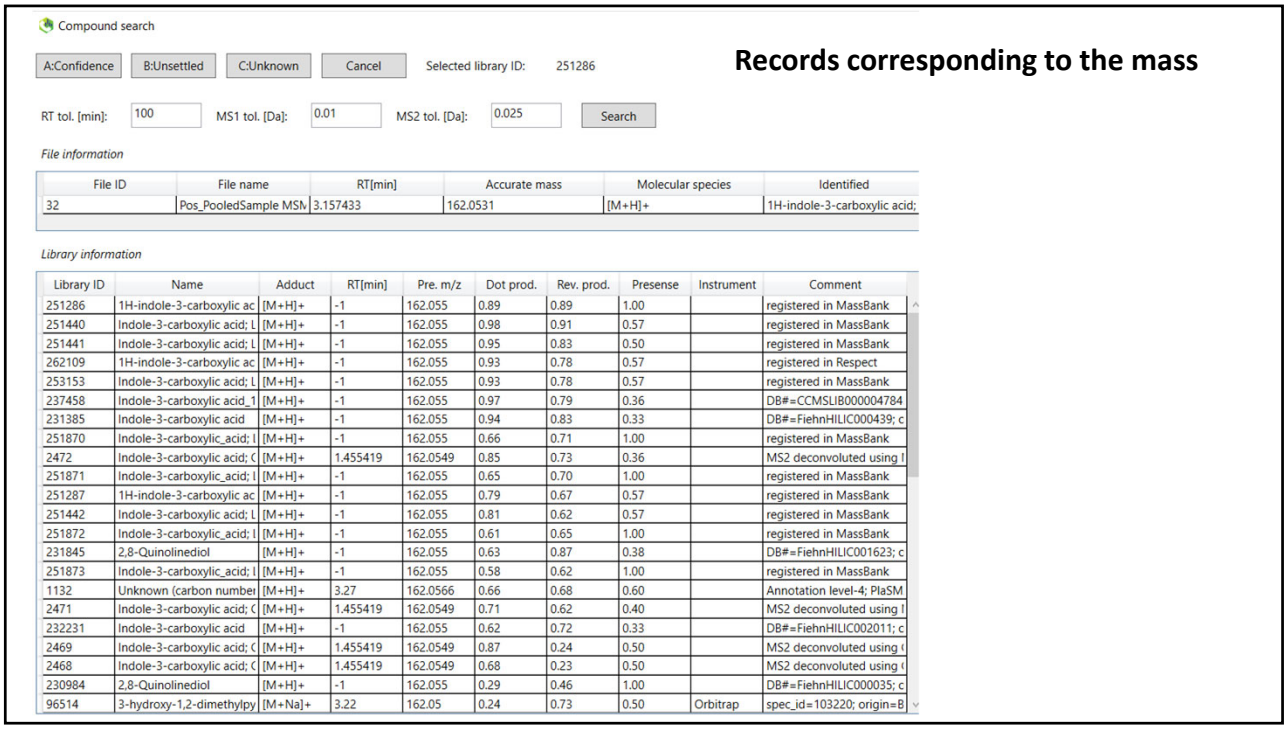
53

Indole-3-carboxylic acid – tryptophan metabolite



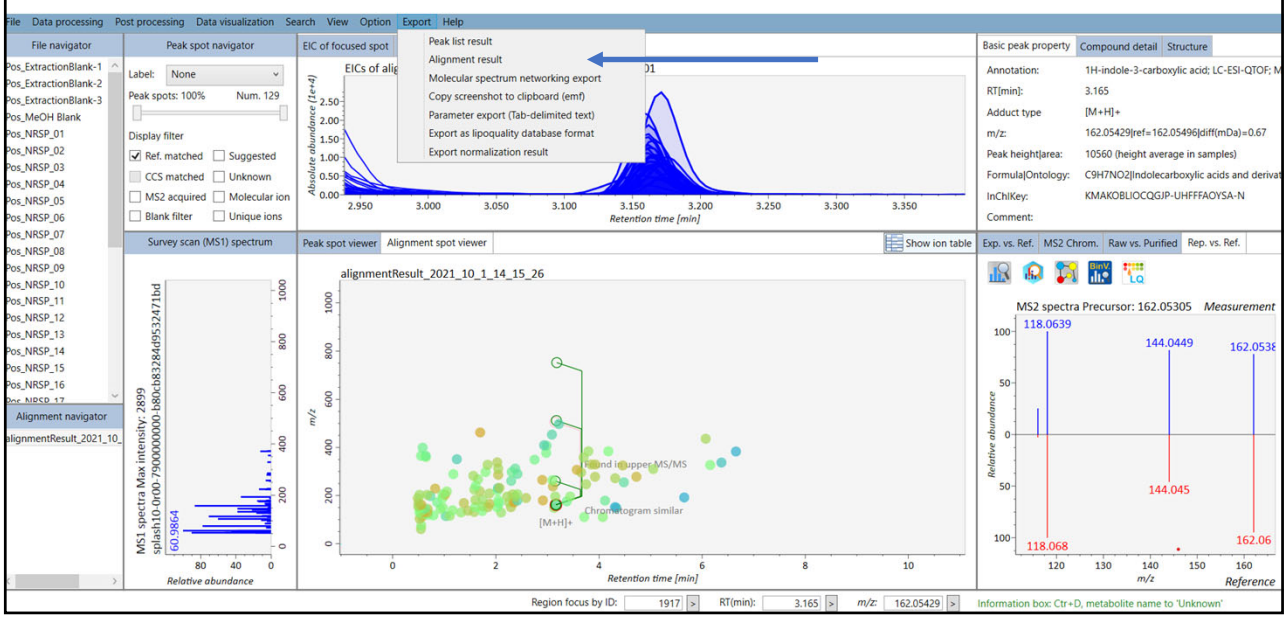
54

Records corresponding to the mass

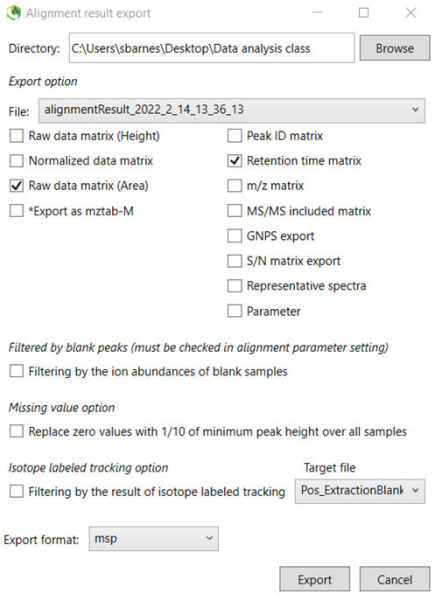


55

Exporting the aligned data



56



Before exiting MS-DIAL, go to Export (top bar) and select <alignment result>. Then select where you are going to save the export file.

De-select raw data matrix (height) and select raw data matrix (area) and retention time matrix.

Then <Export>. The resulting file is in .txt format and can be read into Excel on any platform and then saved in a .xlsx format – we'll do that on Friday.

Also in MS-DIAL, make sure to save the MS-DIAL file – click on <file> (top left corner) and <save>.

57