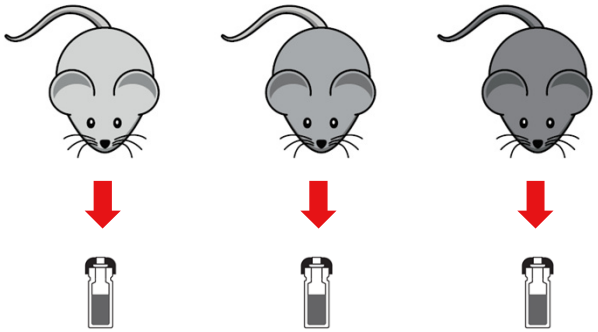
# HRMS data elaboration: from features to compounds

Igor Fochi
Application Specialist
LSMS Support Team, Southern Europe

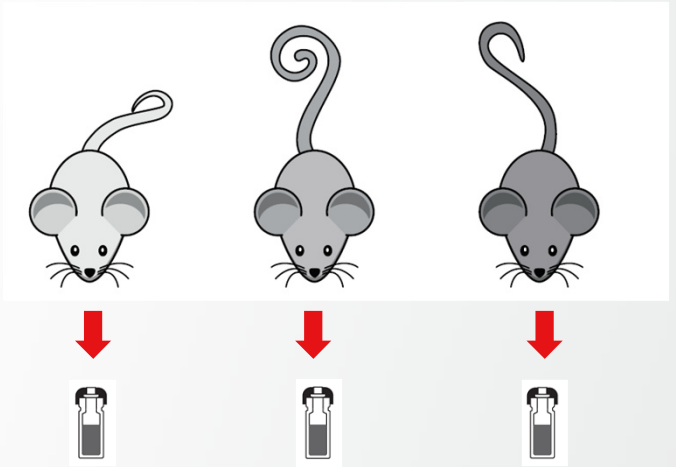
The world leader in serving science

# Untargeted Approach Scenario

**Condition A**



**Condition B**

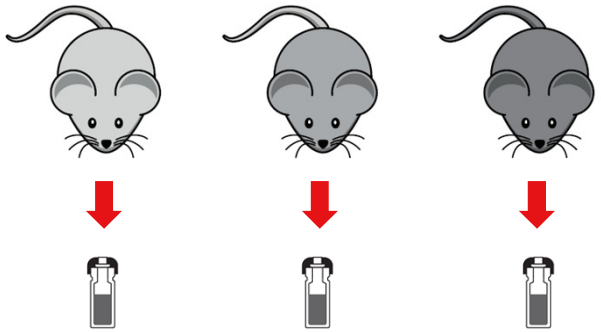


***Find and identify the differences !***

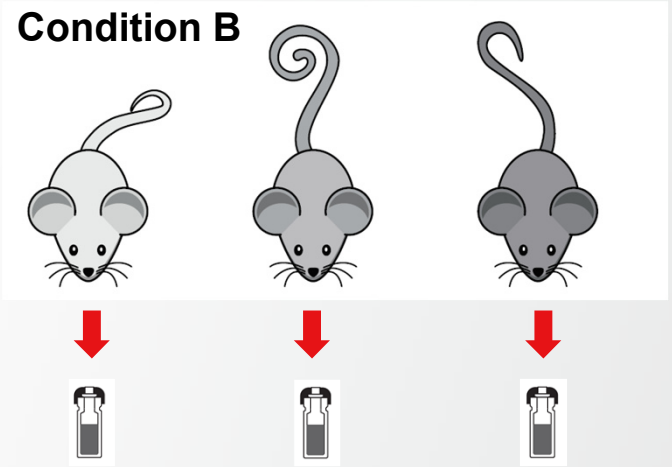
# Untargeted Approach Scenario

Thermo Fisher SCIENTIFIC

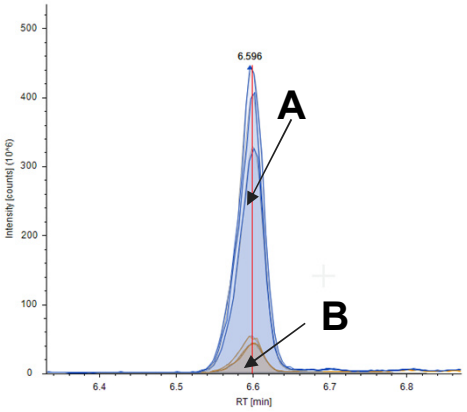
## Condition A



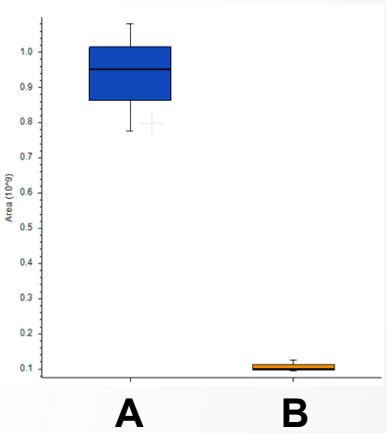
## Condition B



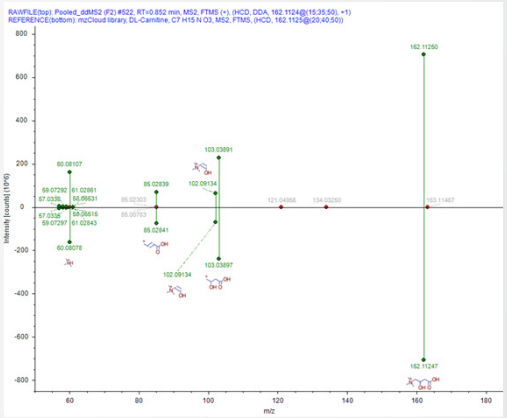
## *Compounds detected*




## *Statistics*




## *Identification*

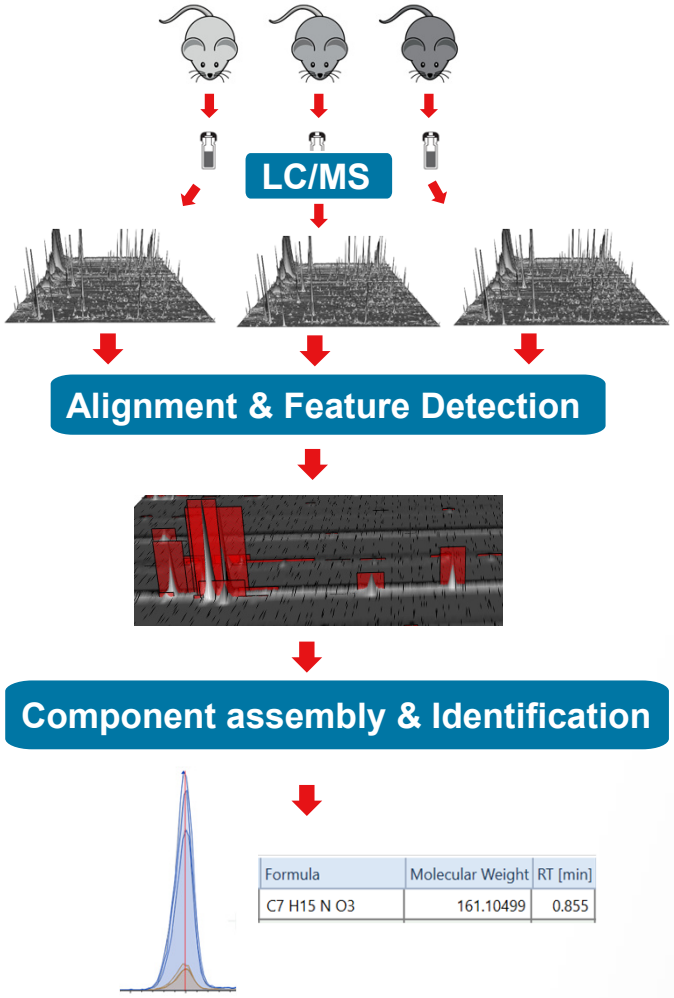


# Software

## Thermo Scientific™ Compound Discoverer™ Software

Metabolomics, Metabolites and Biomarkers Discovery, Environmental research, Clinical and Forensics research, Foodomics, …

# Data Reduction




**100,000,000**

**Raw Data Points** (thousands of spectra in each file)

**500,000 m/z**

e.g. m/z 162.1120 @ 0.88 min

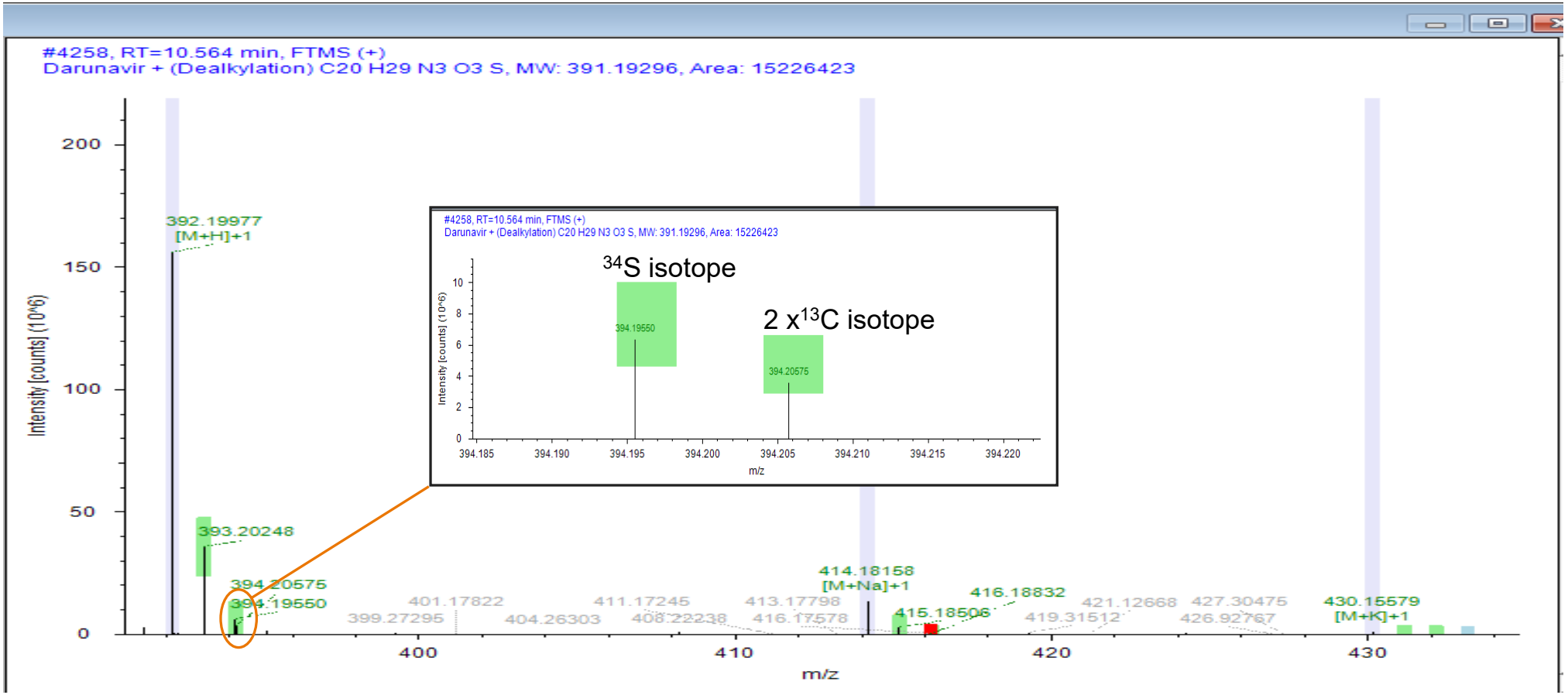
m/z 163.1155 @ 0.89 min

…

**1,000 Compounds**

e.g. MW 161.10 @ 0.85 min

# Features


#4258, RT=10.564 min, FTMS (+)
Darunavir + (Dealkylation) C20 H29 N3 O3 S, MW: 391.19296, Area: 15226423
Intensity [counts] (10^6)
m/z
392.19977
[M+H]+1
393.20248
394.20575
394.19550
414.18158
[M+Na]+1
416.18832
415.18506
430.15579
[M+K]+1
401.17822
399.27295
404.26303
411.17245
408.22238
413.17798
416.17578
421.12668
419.31512
427.30475
426.92767
34S isotope
2 x13C isotope

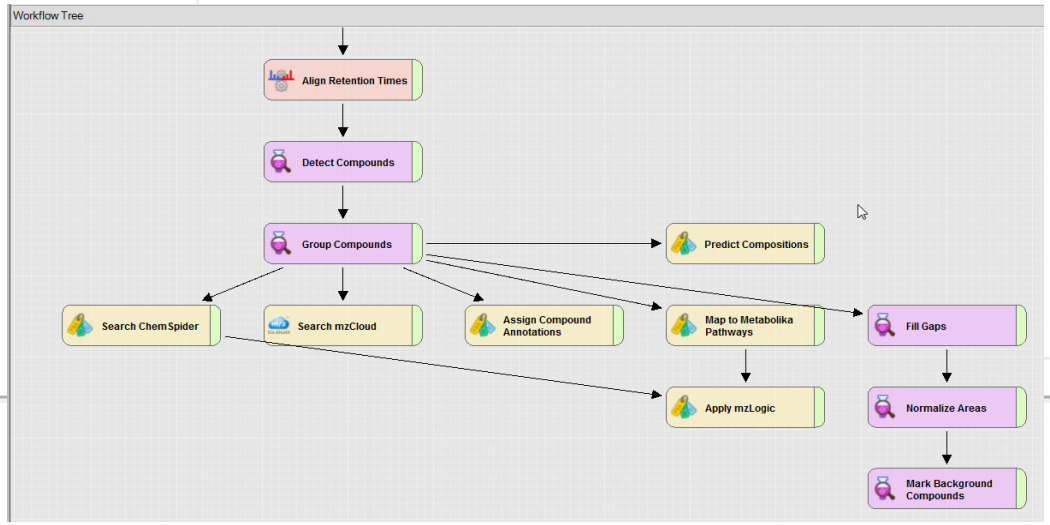
# Workflow Details

Processing

Workflow: WorkflowTemplates \ Metabolomics \ Untargeted Metabolomics with Statistics Detect Unknowns with ID using Online Databases and mzLogic

Workflow Description:

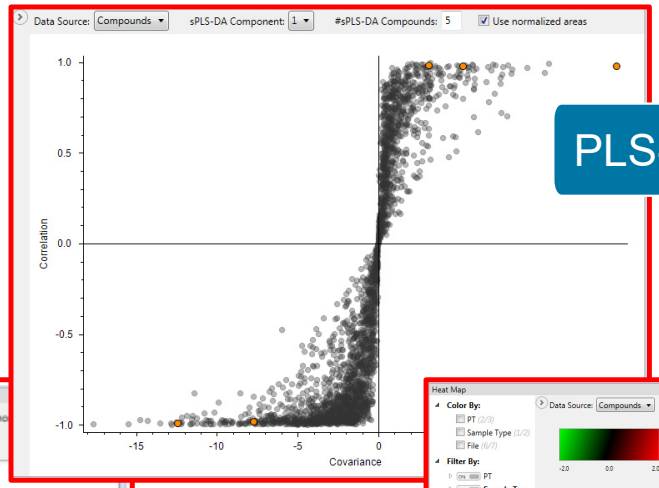
Untargeted Metabolomics workflow: Find and identify the differences between samples.
- Performs retention time alignment, unknown compound detection, and compound grouping across all samples. Predicts elemental compositions for all compounds, fills gaps across all samples, and hides chemical background (using Blank samples). Identifies compounds using mzCloud (ddMS2) and ChemSpider (formula or exact mass). Also performs similarity search for all compounds with ddMS2 data using mzCloud. Applies mzLogic algorithm to rank order ChemSpider results. Maps compounds to biological pathways using Metabolika. Applies QC-based batch normalization if QC samples are available. Calculates differential analysis (t-test or ANOVA), determines p-values, adjusted p-values, ratios, fold change, CV, etc.).



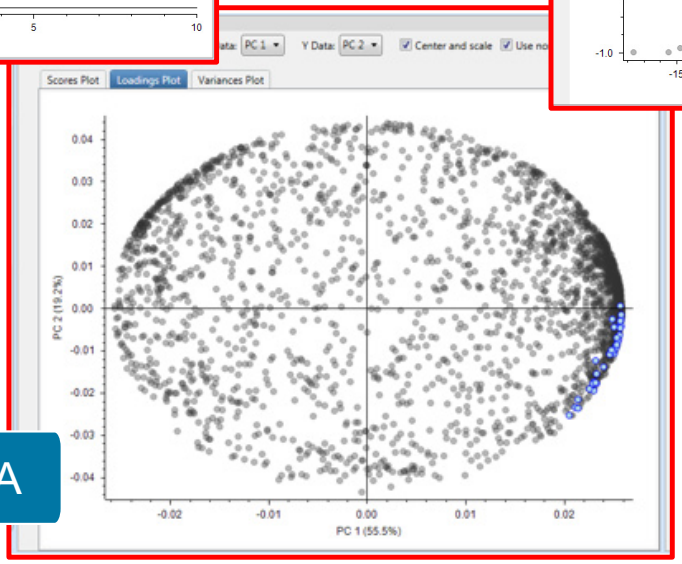

- Flexible data processing workflows
- Use drag & drop, “smart” connection
- Work with predefined workflows or create your own

# Statistics

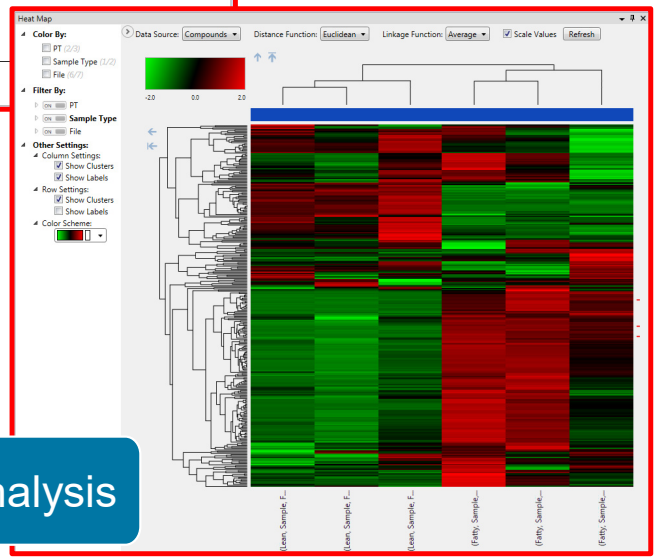

Volcano Plot




PLS-DA

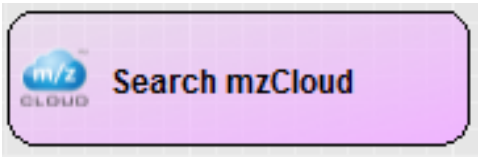



PCA

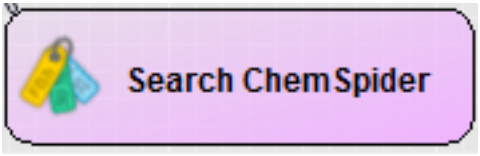



Cluster Analysis

# Spectral Libraries and Compound Databases

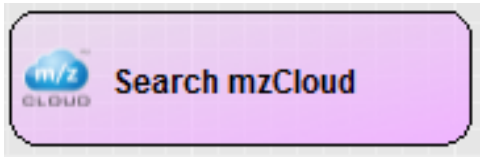



Online spectral library
> 9 million spectra

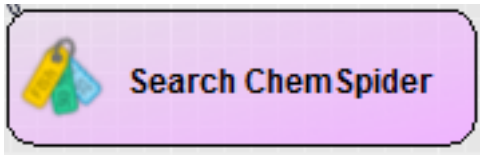



Chemical structure database
> 110 million structures

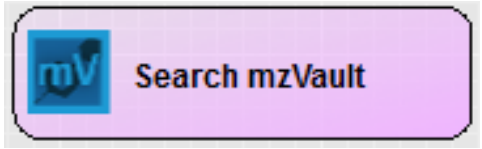
# Spectral Libraries and Compound Databases



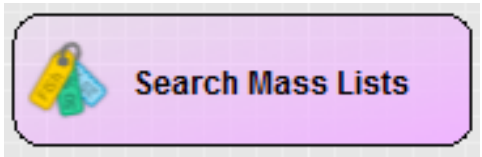

Online spectral library
> 9 million spectra




Chemical structure database
> 110 million structures




Local msms spectral libraries




Local databases

# What is mzCloud?

**mzCloud** is a state of the art mass spectral database that assists analysts in identifying compounds in areas such as life sciences, metabolomics, pharmaceutical research, toxicology, forensic investigations, environmental analysis, food control and various industrial applications. mzCloud™ features a freely searchable collection of high resolution/accurate mass spectra using a new third generation spectra correlation algorithm.

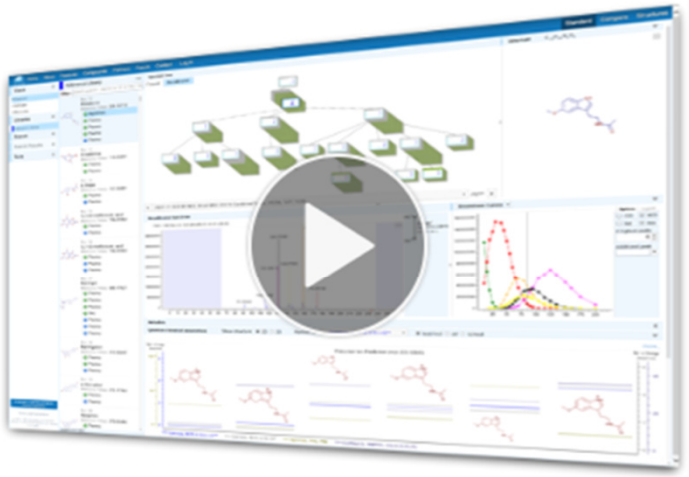
Online access to the database is free of charge and no registration is required.

read more...

Enter Database

Google play

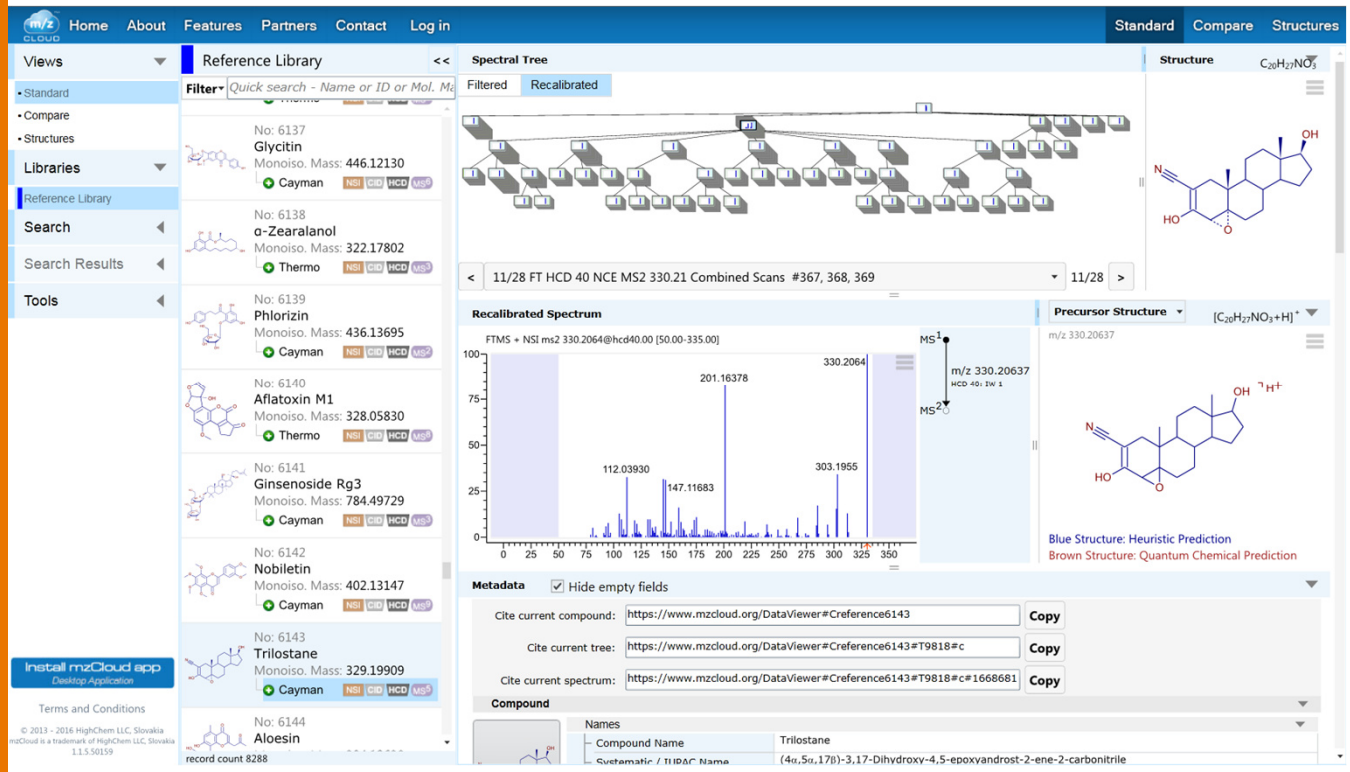
New mzCloud App!



| 19,939 (+22) | 27,539 (+22) | 9,058,715 (+27,724) |
|---|---|---|
| compounds | trees | spectra |

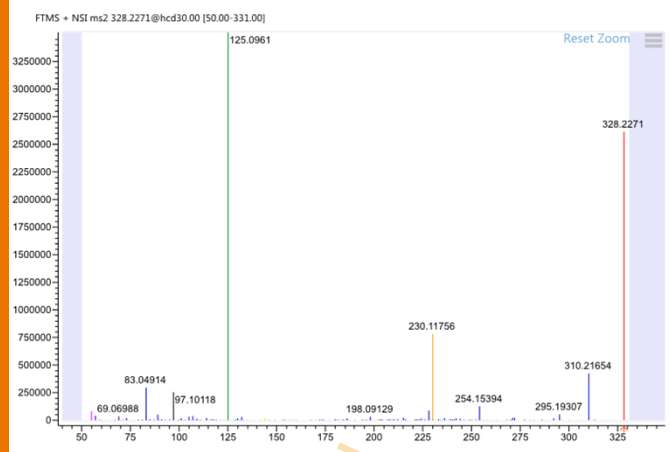
# mzCloud™ - World's largest spectral library

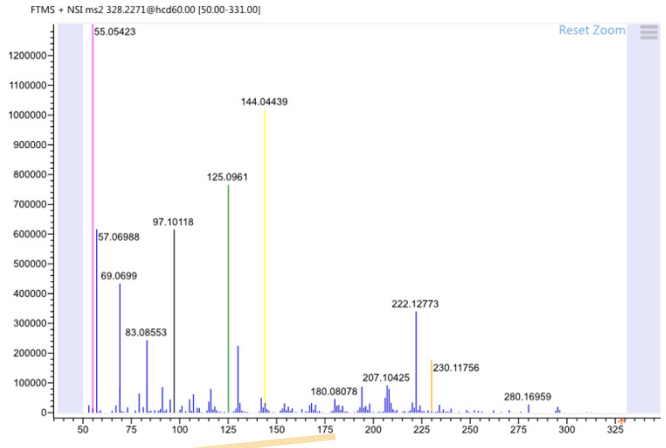
- HRAM MS/MS and $MS^n$
- High quality curated data
- Wide chemical diversity
- Searchable web UI
- Integrated into TFS software

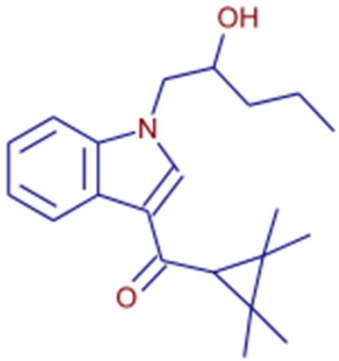


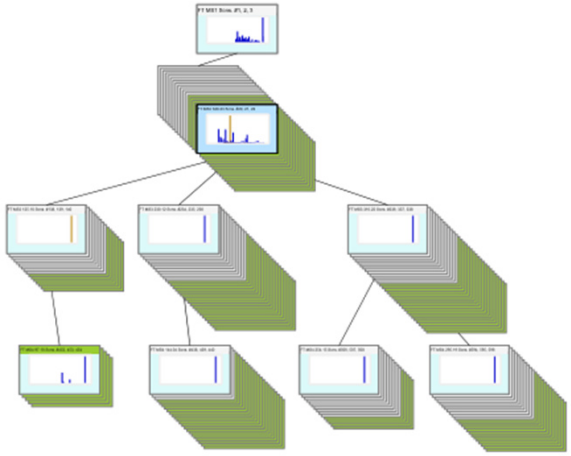
# mzCloud™ - World's largest spectral library

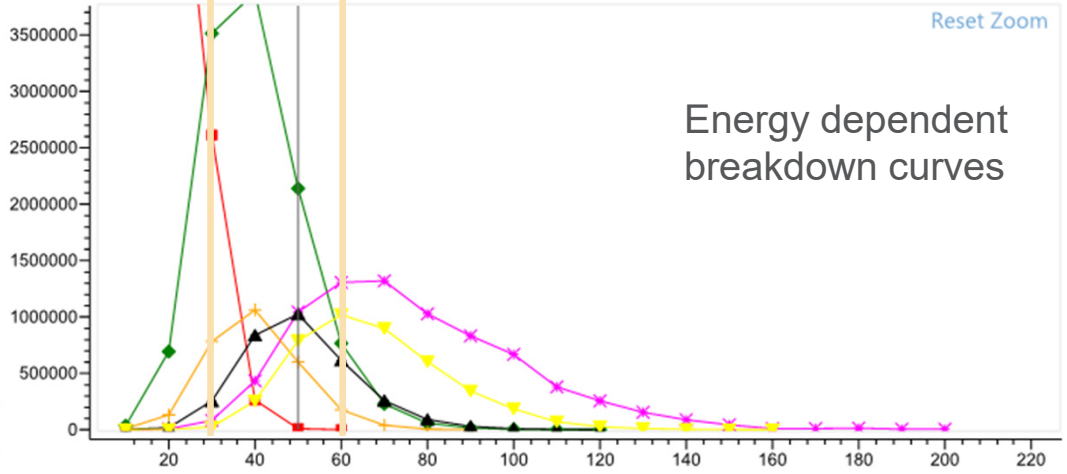
- Spectral data at more than 10 collision energies
- Easily match spectra across different platforms
- Knowledge of energy dependent fragmentation



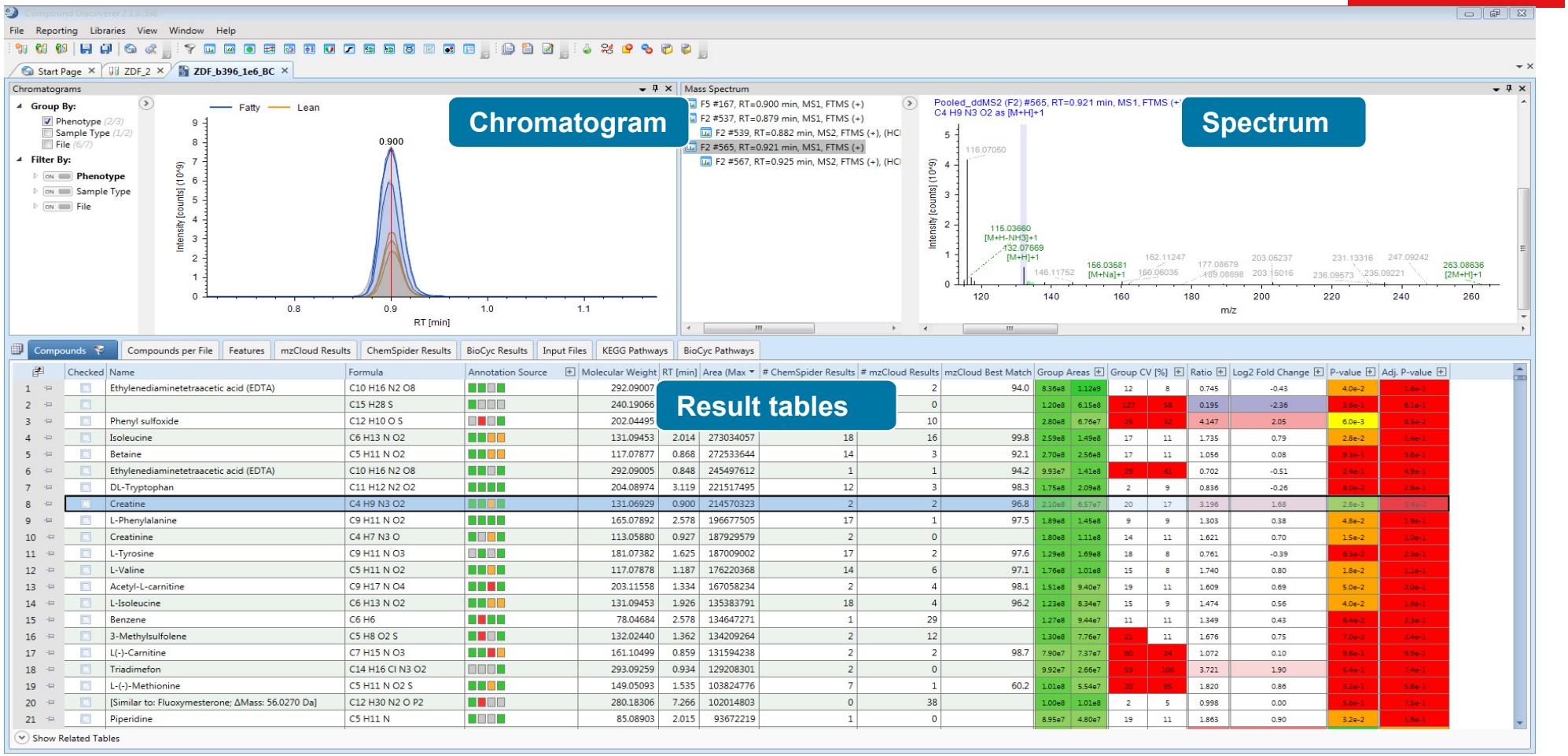



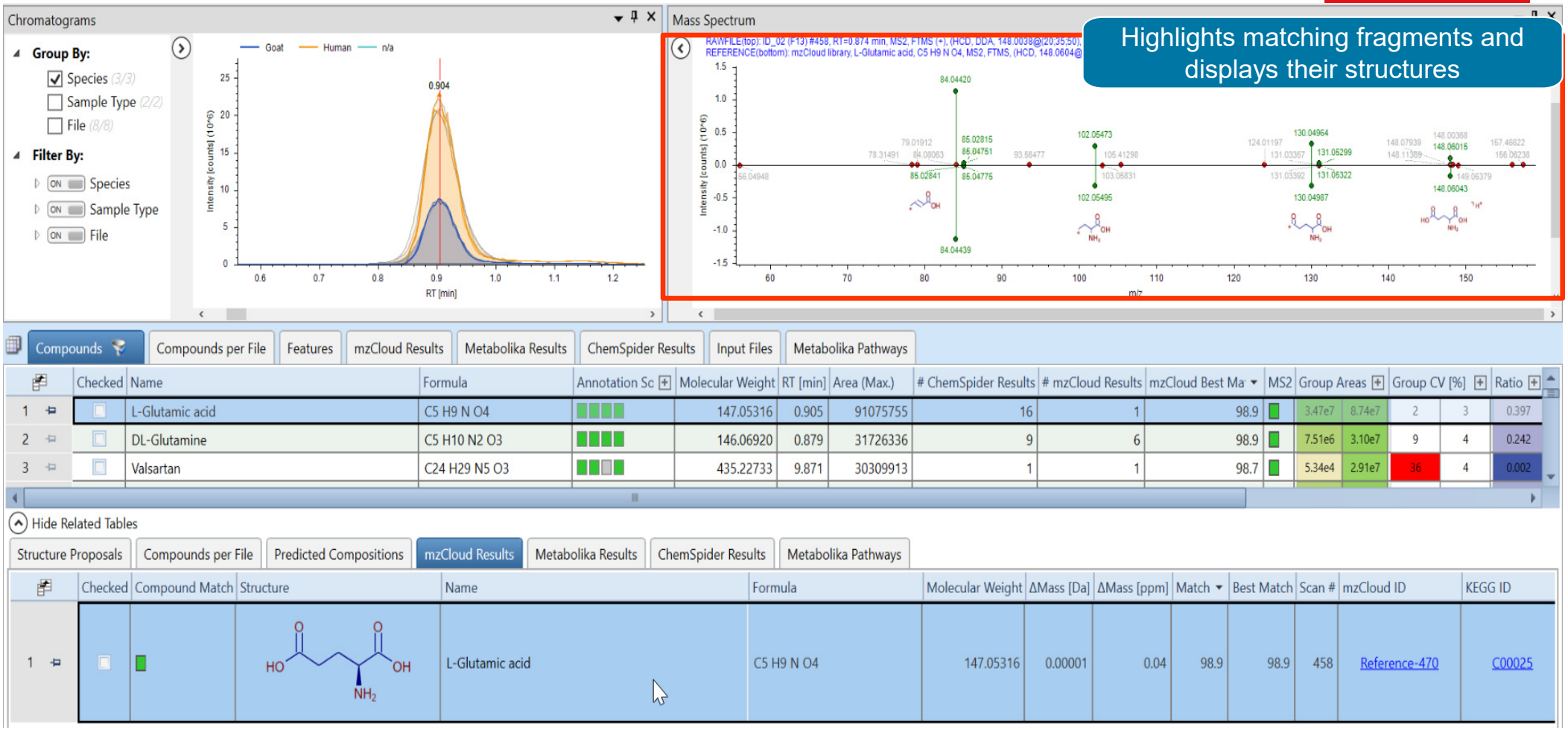





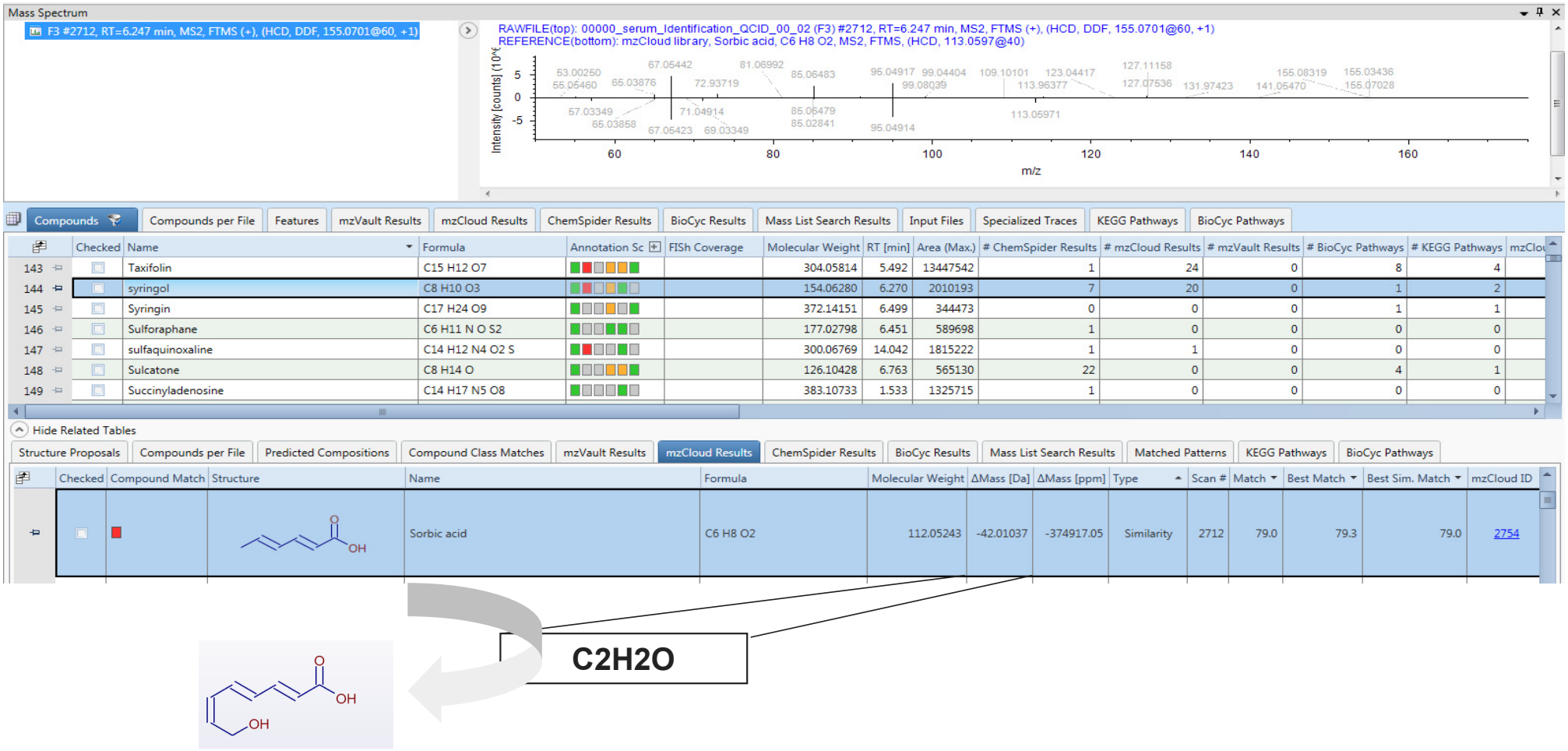
# Interpreting the Results



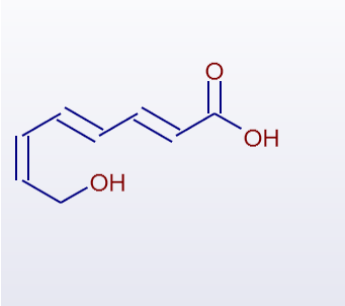
# mzCloud search

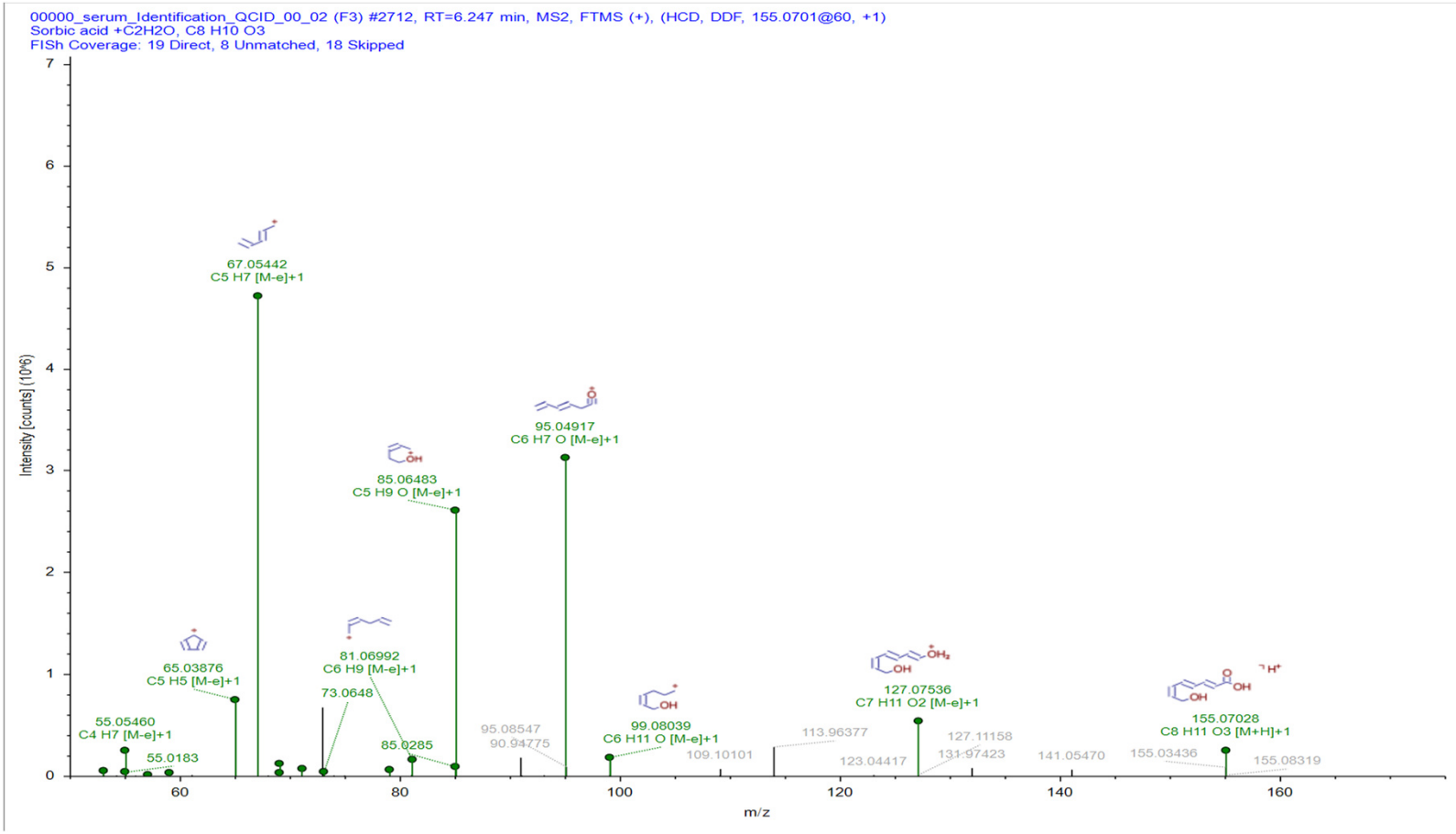


# Similarity search: mzCloud

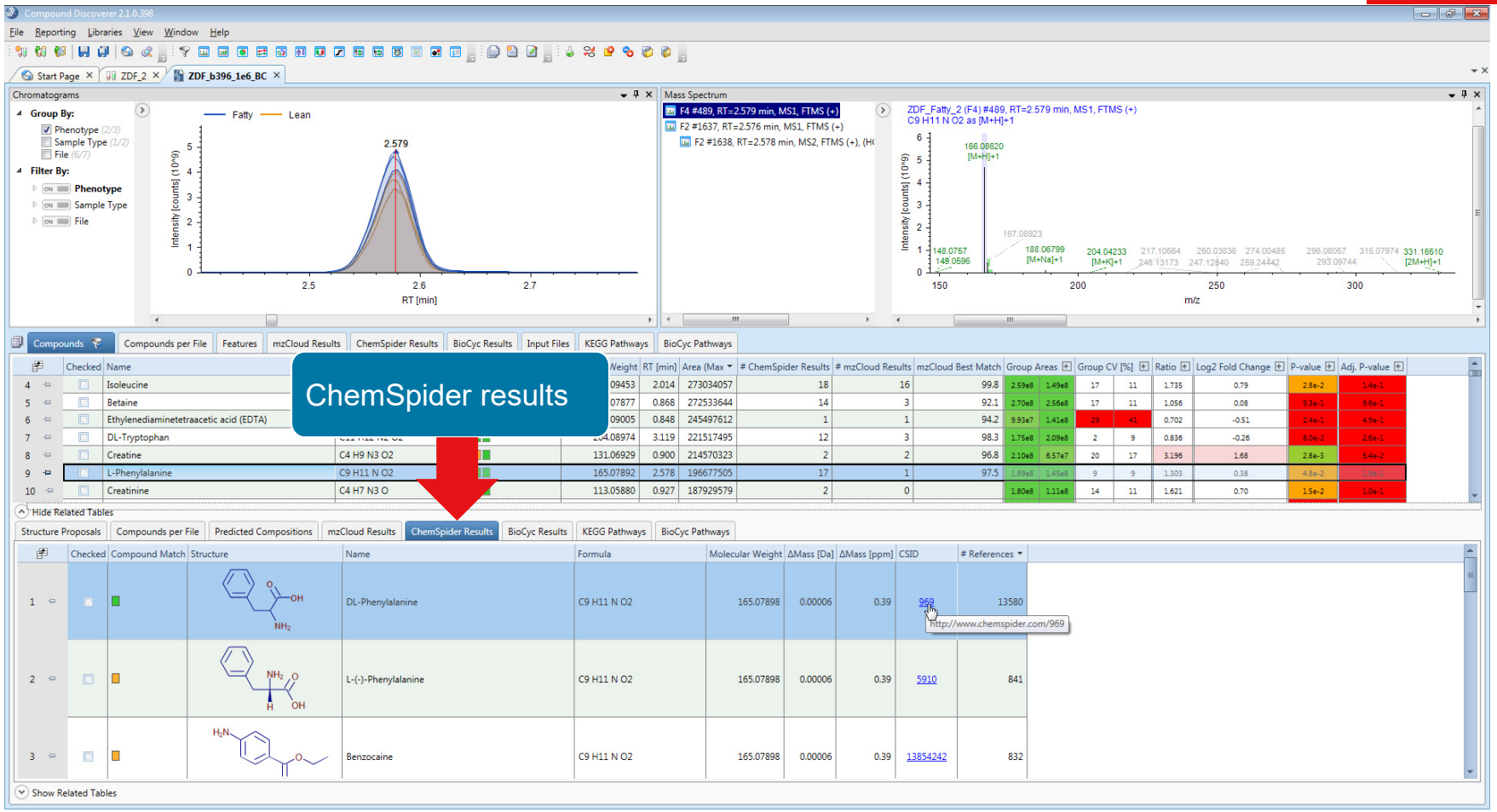


# Methyl Hydroxy Sorbic Acid fragments explanation: FISh (Fragment Ion Search) scoring

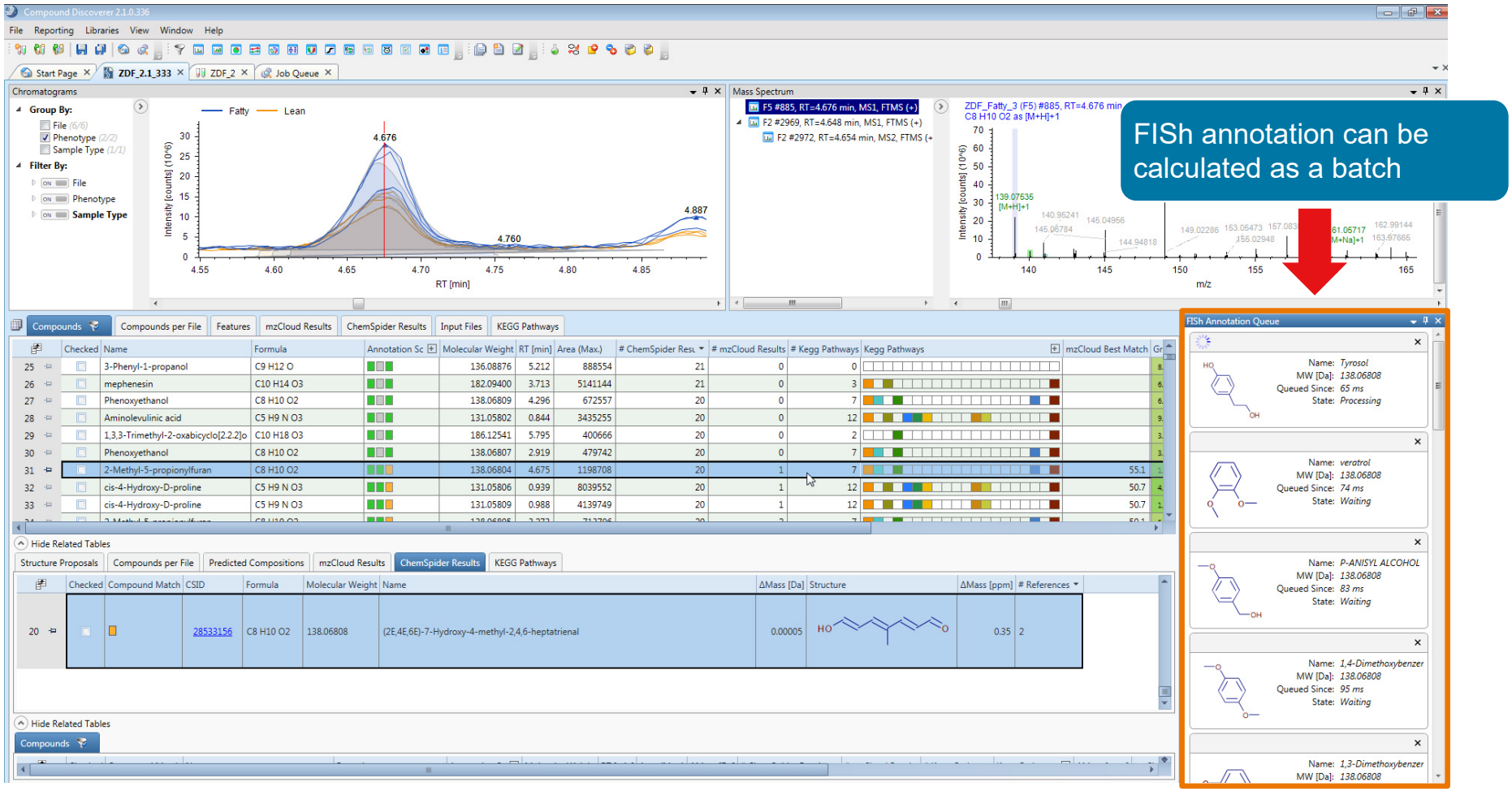


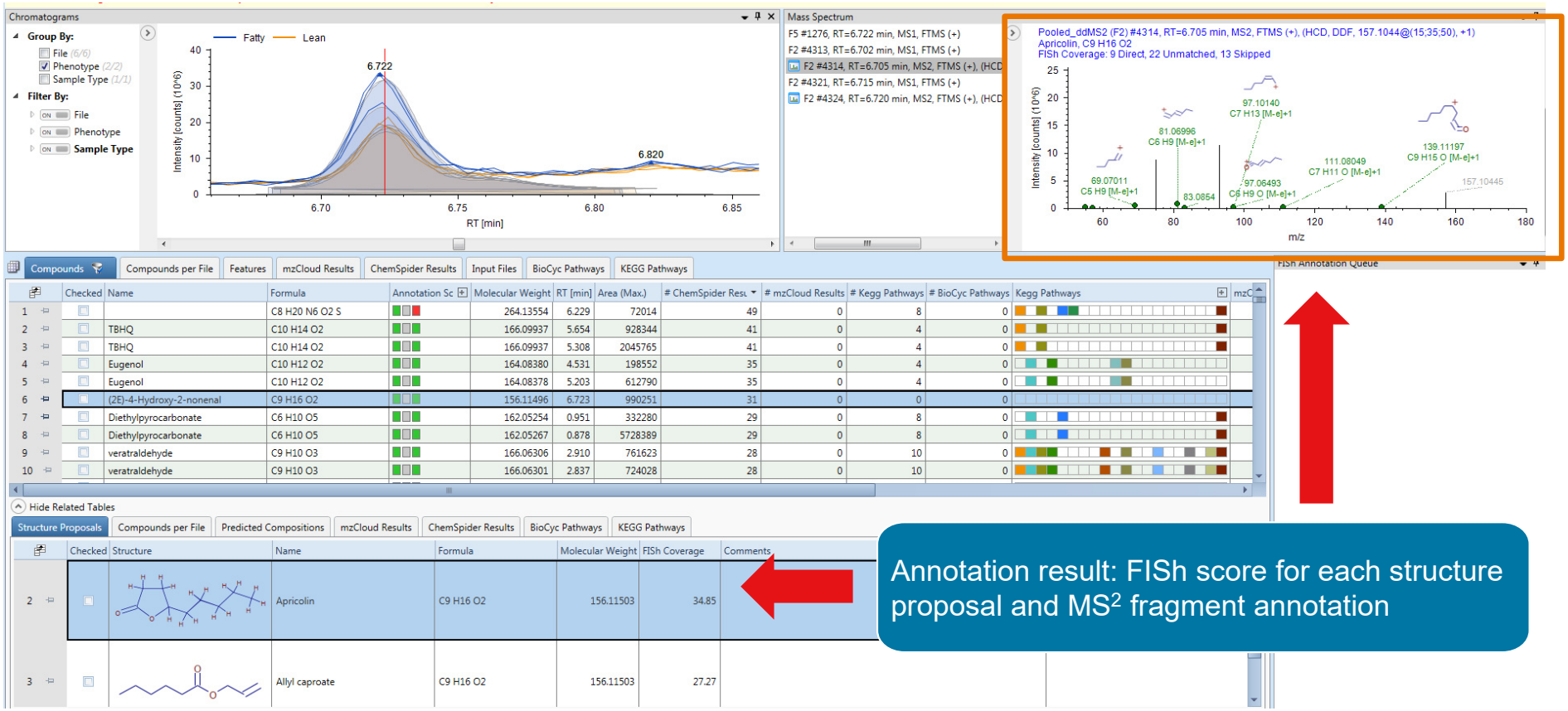
# ChemSpider



# FISh score for Structure Proposals

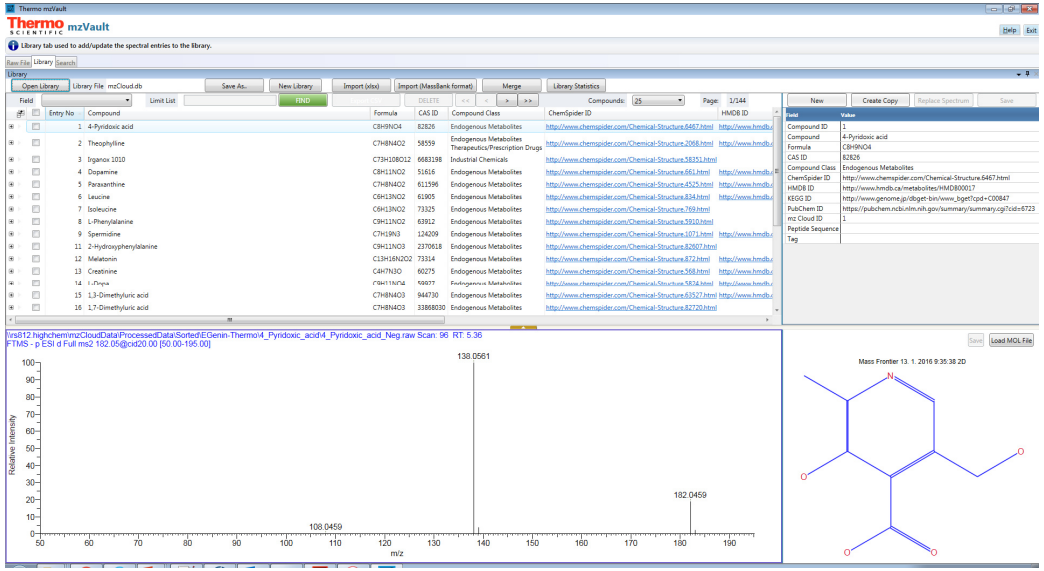


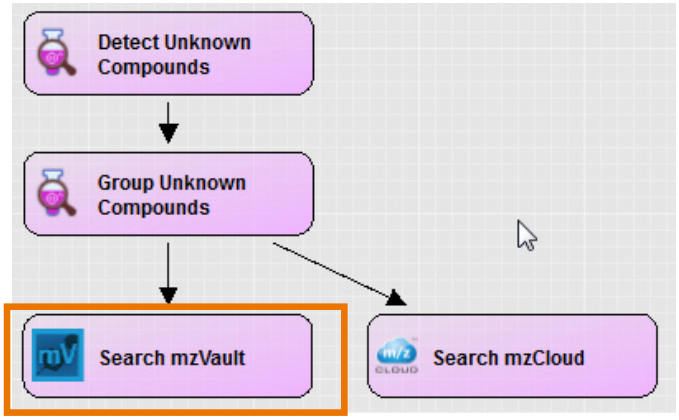
# FISh Annotation for Structure Proposals



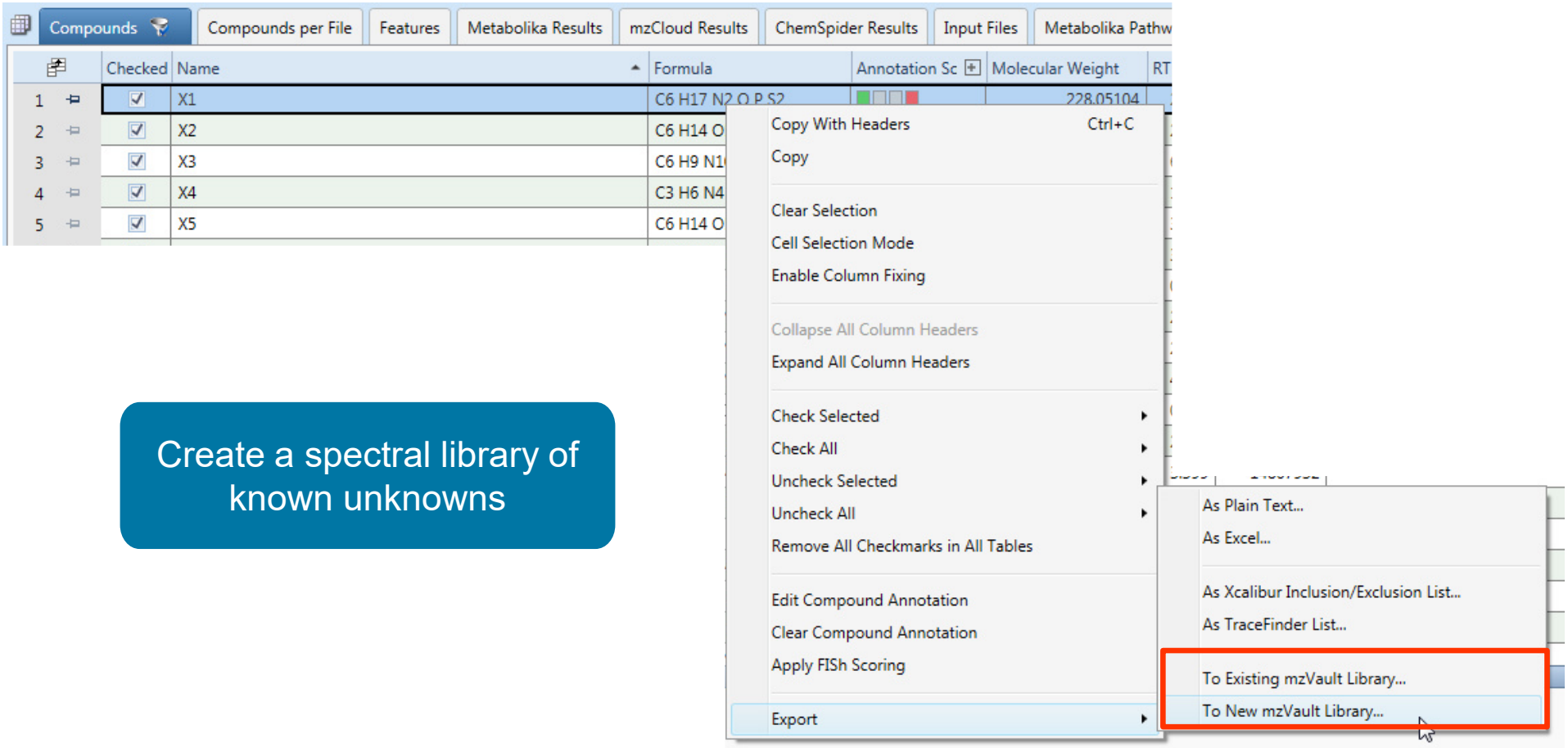
# Local Spectral Libraries (mzVault)

- **mzVault** – support for local spectral libraries
  - Compound Discoverer software comes with a local version of the mzCloud™ database
  - Custom spectral libraries can created and edited using mzVault application



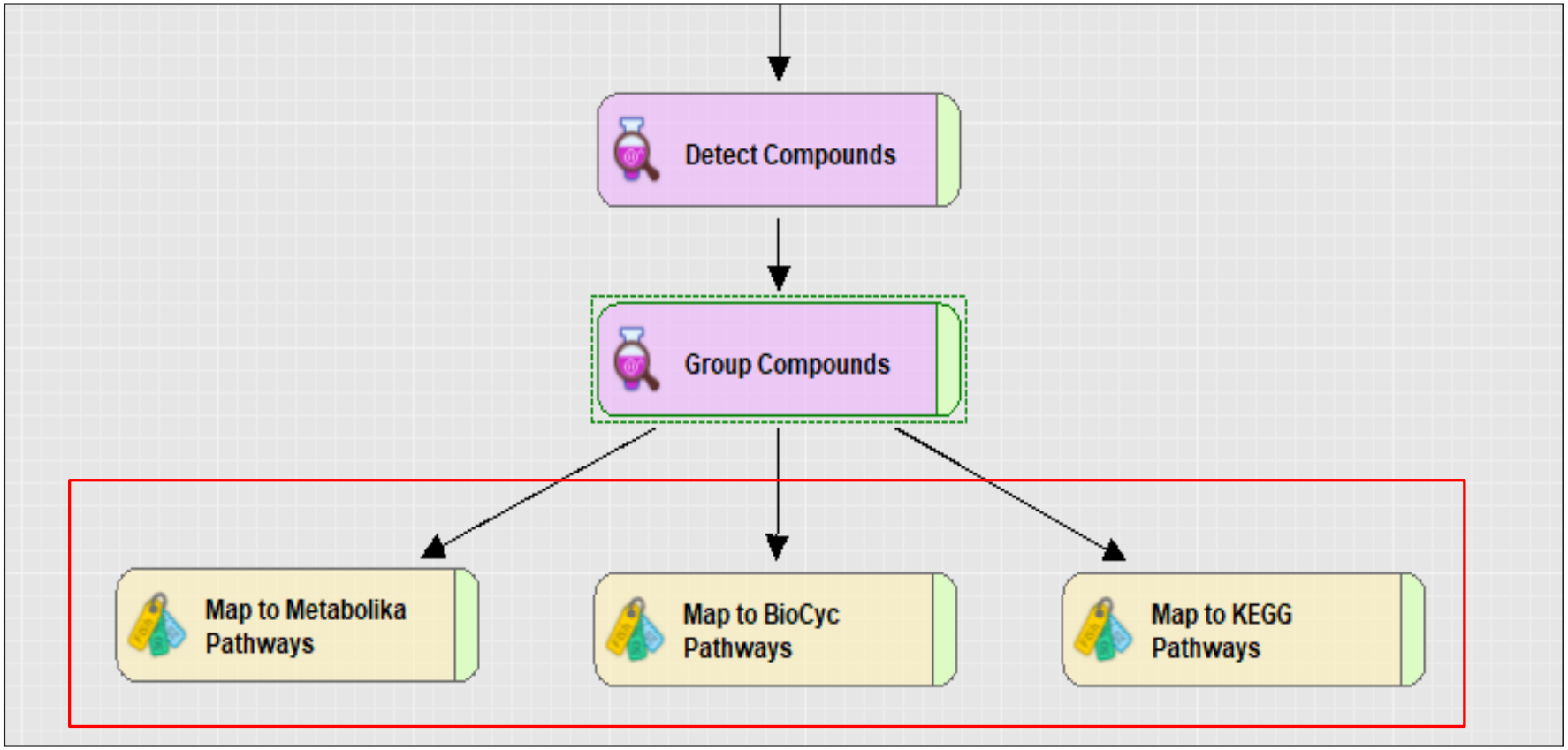



# Export to mzVault

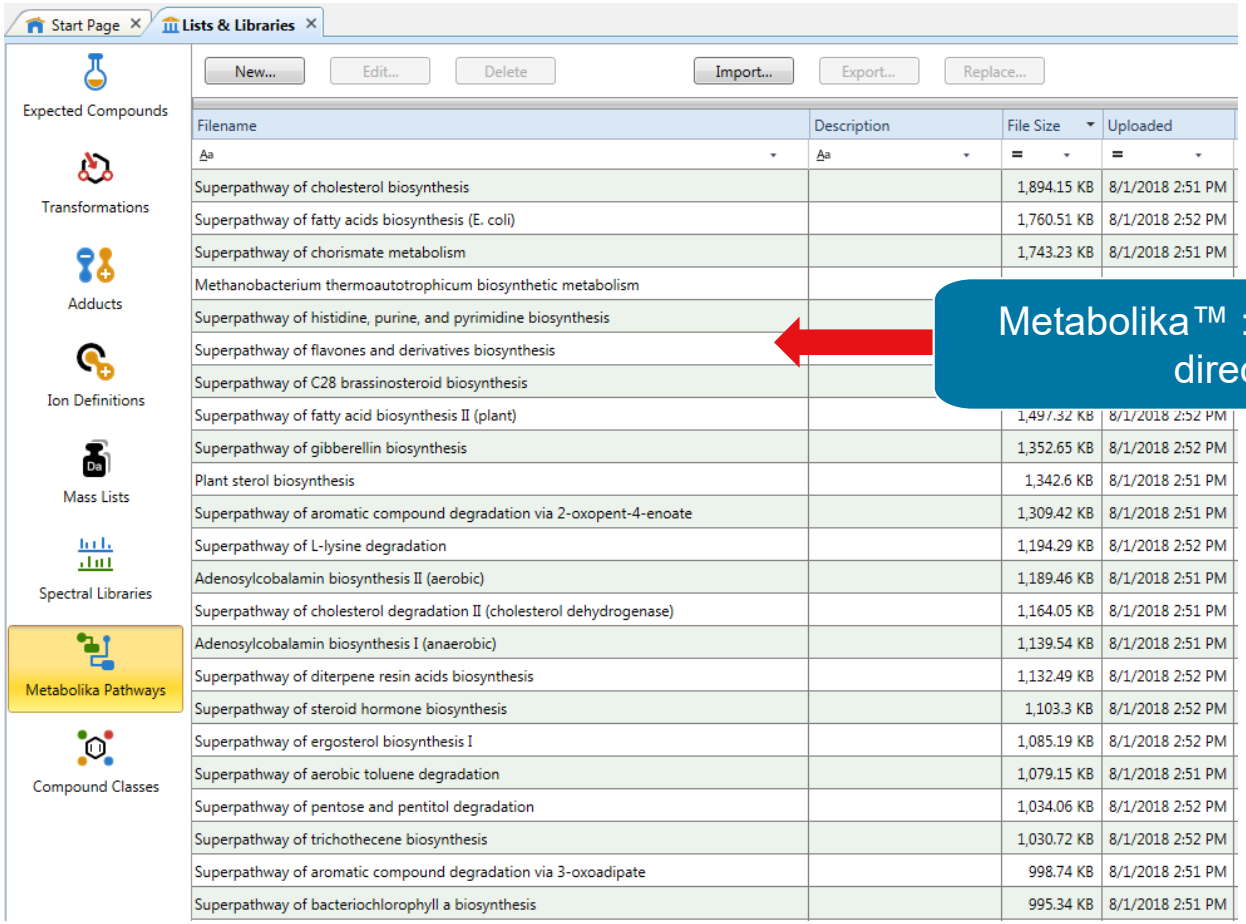



Create a spectral library of known unknowns

# Metabolomics Database

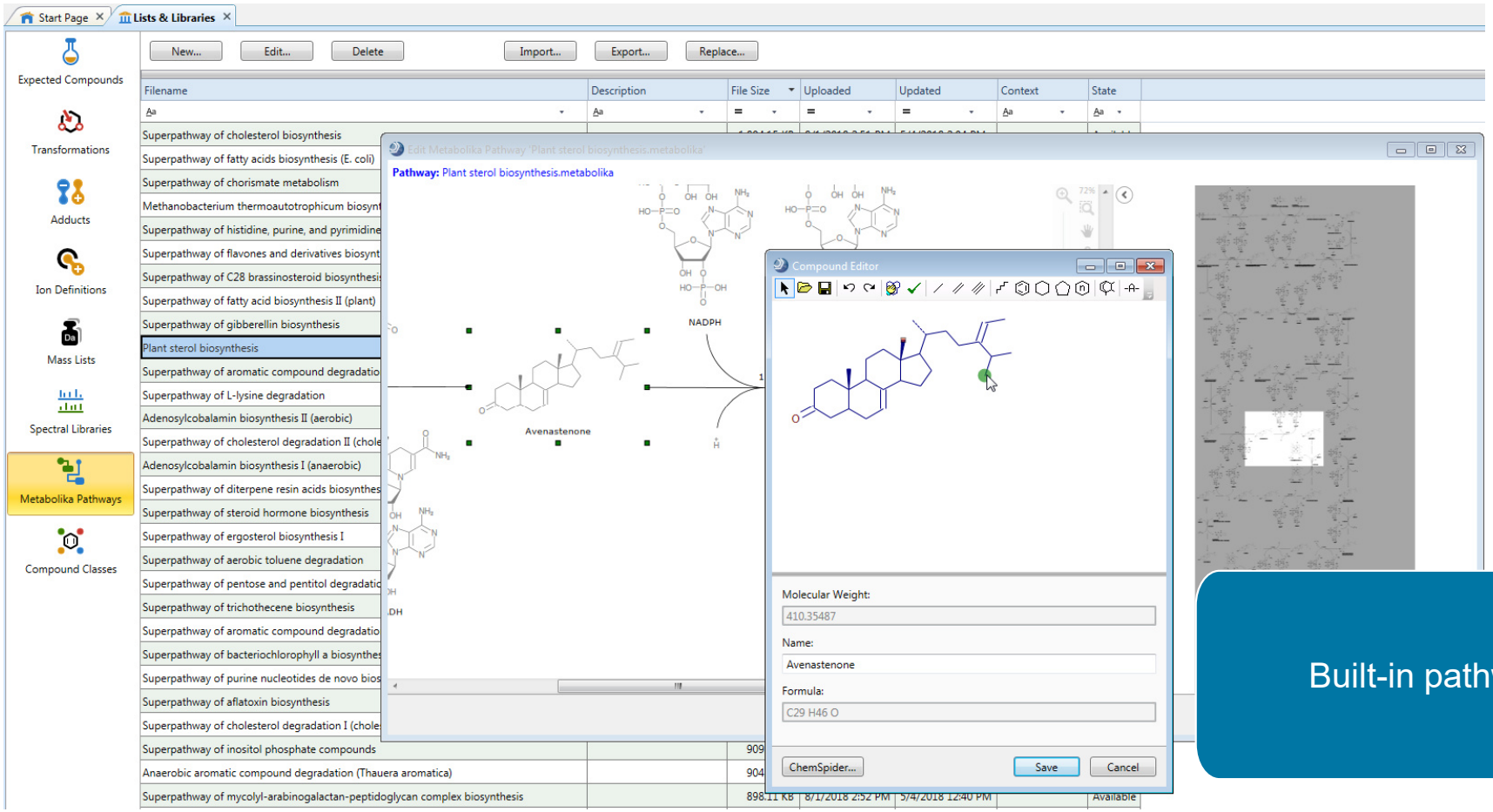


# Metabolika Pathways



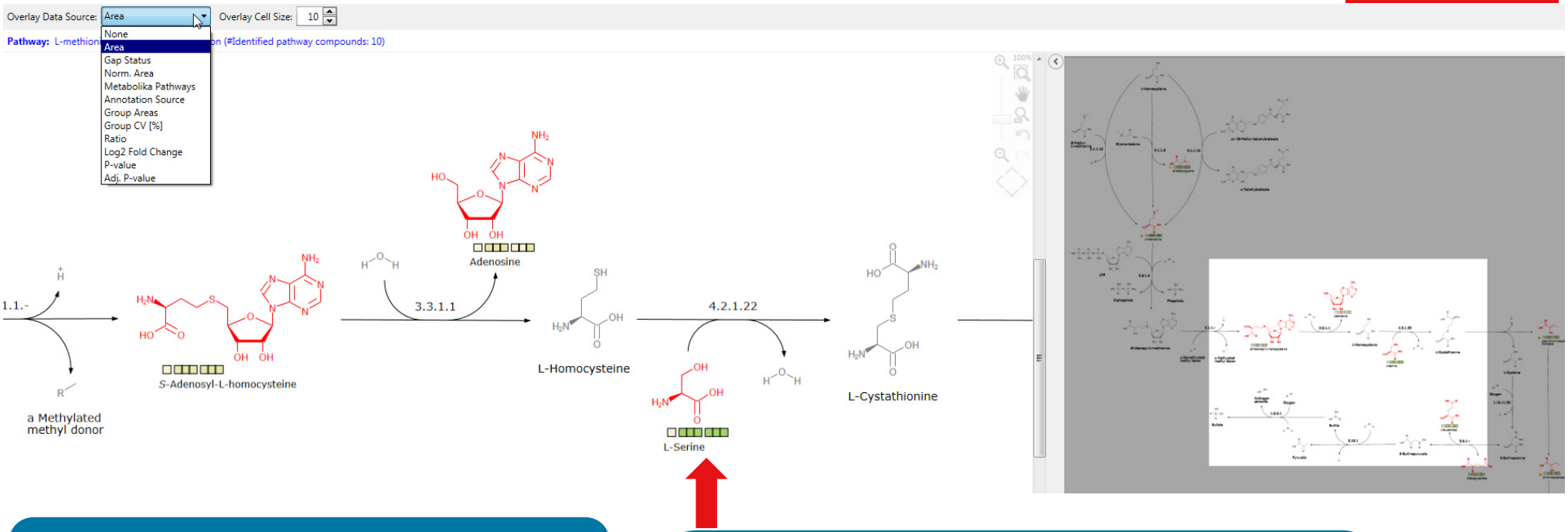

Metabolika™ : set of 378 metabolic pathways, directly integrated into CD

# Metabolika Pathways - Editor




Built-in pathway editor

# Metabolika Pathways

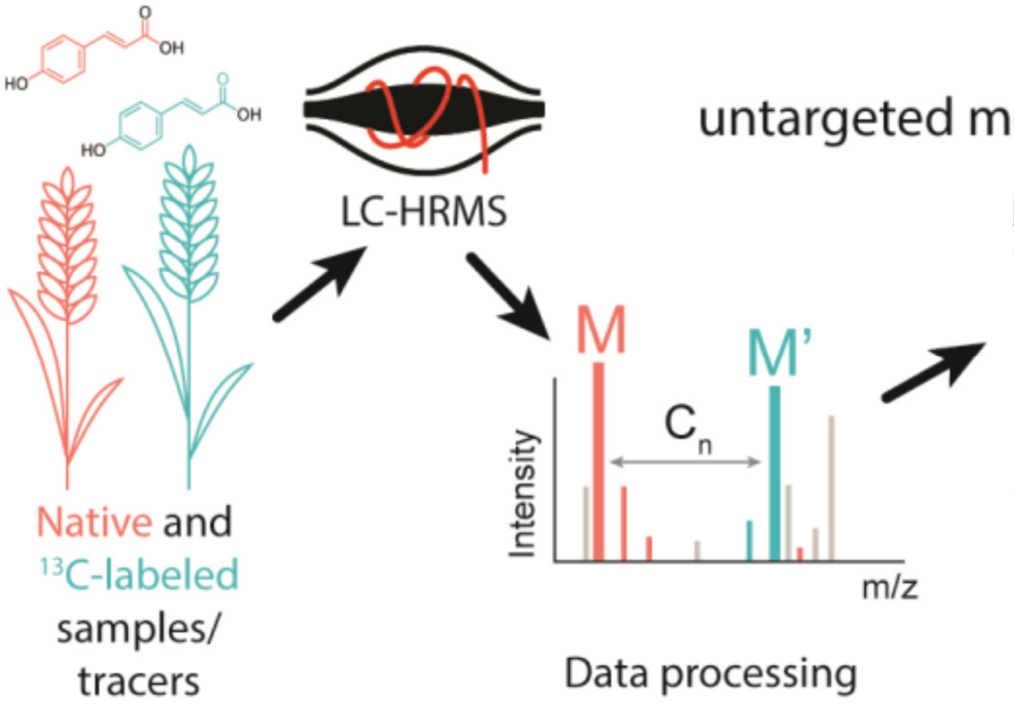



Interactive, customizable pathway maps

Statistical data can be overlaid onto the pathway

# Stable Isotope Labeling

## Stable Isotope Labeling

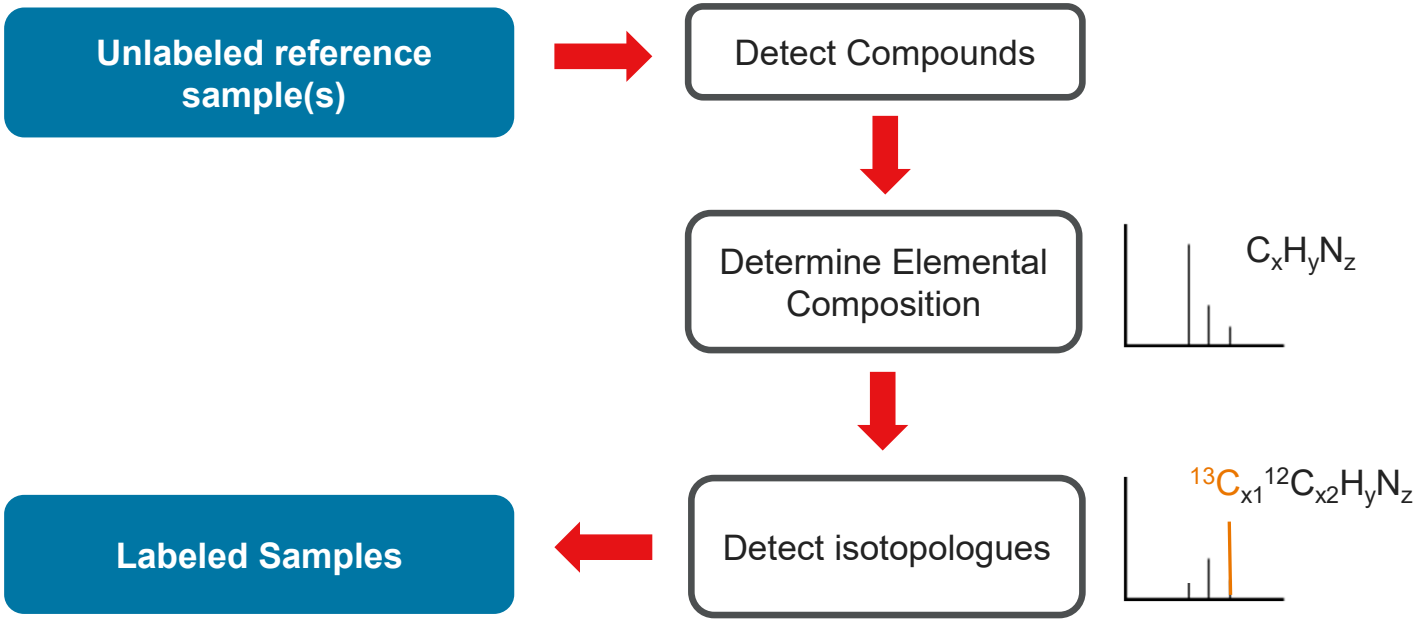



| | Checked | Name | Formula | Annotation | Molecular Weight | RT [min] | Area (Max.) | Rel. Exchange [%] |
|---|---|---|---|---|---|---|---|---|
| 1 | ☑ | Uridine | C9 H12 N2 O6 | | 244.06922 | 7.146 | 362256836 | 0 0 39 67 52 0 |
| 2 | ☑ | Proline | C5 H9 N O2 | | 115.06337 | 11.773 | 70424932 | 0 0 3 3 3 0 |
| 3 | ☑ | Malate | C4 H6 O5 | | 134.02156 | 24.962 | 56234422 | 0 0 8 9 9 0 |
| 4 | ☑ | Glutamine | C5 H10 N2 O3 | | 146.06924 | 13.956 | 48610589 | 0 0 0 0 0 0 |
| 5 | ☑ | Citrate/Isocitrate | C6 H8 O7 | | 192.02713 | 31.567 | 43732087 | 0 0 3 5 4 0 |
| 6 | ☑ | Lactate | C3 H6 O3 | | 90.03163 | 14.231 | 14127499 | 0 0 66 55 69 0 |
| 7 | ☑ | NAD | C21 H27 N7 O14 P2 | | 663.11102 | 20.138 | 5792721 | 0 0 40 39 39 0 |
| 8 | ☑ | Alanine | C3 H7 N O2 | | 89.04760 | 12.561 | 5698362 | 0 0 45 48 44 0 |
| 9 | ☑ | a-Ketogluterate | C5 H6 O5 | | 146.02173 | 25.941 | 6177872 | 1 0 3 3 3 1 |
| 10 | ☑ | Hexose-6-phosphate | C6 H13 O9 P | | 260.02978 | 25.877 | 3922810 | 0 0 83 83 81 0 |
| 11 | ☑ | Serine | C3 H7 N O3 | | 105.04258 | 14.386 | 3159629 | 0 0 13 10 10 0 |
| 12 | ☑ | Glycine | C2 H5 N O2 | | 75.03207 | 13.587 | 4393736 | 0 0 6 6 6 0 |

**Detect labeled compounds**

modified from Bueschl et. al. Analytical Chemistry 2017

# Untargeted detection of labeled compounds




**Untargeted Analysis**: CD automatically detects labeled compounds (isotopologues) based on formulas of unlabeled compounds found in reference file(s)

# Untargeted detection of labeled compounds


Study Definition | Input Files | Samples

| Error | Sample | File | Sample Identifier | Sample Type |
|---|---|---|---|---|
| | S1 | F1 | 20150226_HeLa_24_NH2_NEG_12C_Glc_1_07 | Sample |
| | S2 | F2 | 20150226_HeLa_24_NH2_NEG_12C_Glc_3_13 | Sample |
| | S6 | F6 | 20150226_HeLa_24hr_NH2_NEG_12C_Glc_2_11 | Sample |
| | S3 | F3 | 20150226_HeLa_24_NH2_NEG_13C_Glc_1_06 | Labeled |
| | S4 | F4 | 20150226_HeLa_24_NH2_NEG_13C_Glc_2_09 | Labeled |
| | S5 | F5 | 20150226_HeLa_24_NH2_NEG_13C_Glc_3_14 | Labeled |

Unlabeled reference samples

Labeled Samples

Parameters of 'Analyze Labeled Compounds'

Show Advanced Parameters

| | |
|---|---|
| **1. Label Settings** | |
| Label Element | [13]C |
| Max. Exchange | 25 |
| Source Efficiency [%] | 100 |
| **2. Pattern Analysis** | |
| Mass Tolerance [ppm] | 5 ppm |
| Intensity Tolerance [%] | 30 |
| Intensity Threshold [%] | 0.1 |
| SN Threshold | 3 |
| **3. General Settings** | |
| Hide Unprocessed | True |

Label is the only user-entered parameter required for automatic isotopologues detection.

*$^{13}$C SIL data courtesy of Gary Patti, Washington University*

# Stable Isotope Labeling - Details

| | Checked | Name | Formula | Annotation | Molecular Weight | RT [min] | Area (Max.) | 20150226_HeLa_24_NH2_NEG_12C | 20150226_HeLa_24_NH2_NEG_12C_G | 20150226_HeLa_24hr_NH2_NEG_12C | 20150226_HeLa_24_NH2_NEG_13C_G | 20150226_HeLa_24_NH2_NEG_13C_G | 20150226_HeLa_24_NH2_NEG_13C_G |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | ☑ | Lactate | C3 H6 O3 | | 90.03163 | 14.174 | 13540685 | 0 | 0 | 0 | 67 | 54 | 70 |
| 4 | ☑ | Alanine | C3 H7 N O2 | | 89.04760 | 12.577 | 5415885 | 0 | 0 | 0 | 45 | 48 | 44 |
| 5 | ☑ | Hexose-6-phosphate | C6 H13 O9 P | | 260.02978 | 25.819 | 3930432 | 0 | 0 | 0 | 95 | 83 | 81 |

*Total* exchange rate for this compound in each sample

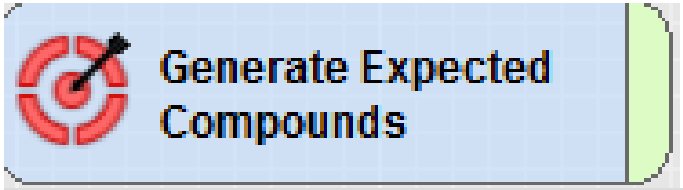
Hide Related Tables

Structure Proposals | Compounds per File | Predicted Compositions | Labeled Compounds per File

| | Checked | Molecular Weight | RT [min] | Max. # MI | Area | Max. Exchange | Avg. Exchange | Rel. Exchange [%] | Status | Exchange Rate [%] | | | | | | | Study File ID |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | ☐ | 260.02978 | 25.840 | 2 | 3930432 | 6 | 0.0 | 0 | | 100 | 0 | 0 | 0 | 0 | 0 | 0 | F1 |
| 2 | ☐ | 260.02978 | 25.850 | 1 | 1498841 | 6 | 0.0 | 0 | | 100 | 0 | 0 | 0 | 0 | 0 | 0 | F2 |
| 3 | ☐ | 260.02978 | 25.839 | 2 | 1043633 | 6 | 0.0 | 0 | | 100 | 0 | 0 | 0 | 0 | 0 | 0 | F6 |
| 4 | ☐ | 260.02978 | 25.831 | 6 | 755352 | 6 | 5.0 | 83 | | 15 | 0 | 0 | 2 | 1 | 4 | 77 | F4 |
| 5 | ☐ | 260.02978 | 25.835 | 5 | 497729 | 6 | 4.9 | 81 | | 17 | 0 | 0 | 2 | 1 | 4 | 76 | F5 |
| 6 | ☐ | 260.02978 | 25.835 | 6 | 436343 | 6 | 5.7 | 95 | | 3 | 0 | 0 | 3 | 0 | 4 | 90 | F3 |

Exchange rates for *each isotopologue* in each sample

# Expected Compound Generator




Parameters of 'Generate Expected Compounds'

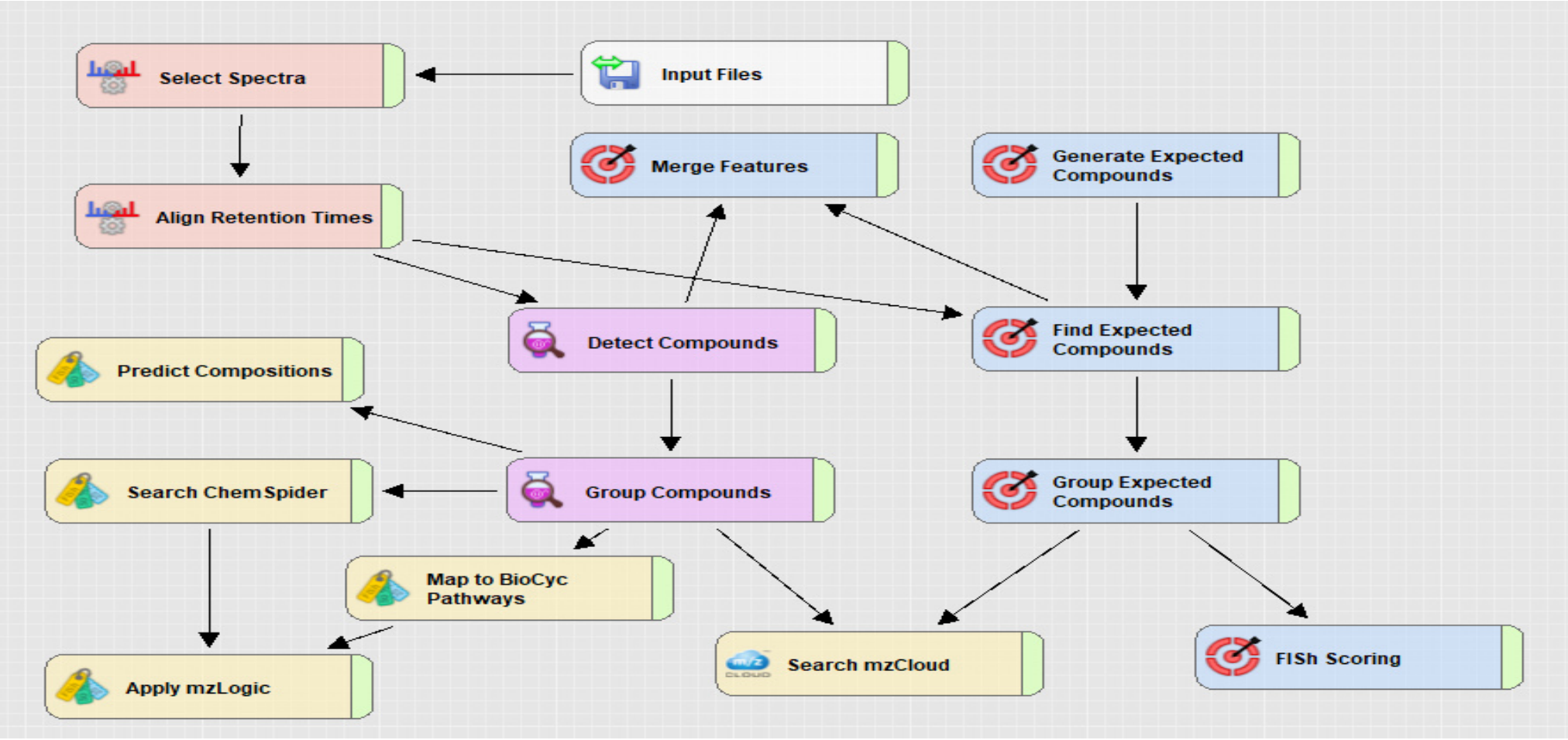
Show Advanced Parameters

| | |
|---|---|
| **1. Compound Selection** | |
| Compound | Omeprazole (C17 H19 N3 O3 S) |
| **2. Dealkylation** | |
| Apply Dealkylation | True |
| Apply Dearylation | True |
| Max. # Steps | 2 |
| Min. Mass [Da] | 150 |
| **3. Transformations** | |
| Phase I | Dehydration (H2 O -> ); Desaturation (H2 -> ); Hydra |
| Phase II | Acetylation (H -> C2 H3 O); Arginine Conjugation (H |
| Others | |
| Max. # Phase II | 1 |
| Max. # All Steps | 2 |
| **4. Ionization** | |
| Ions | [M+H]+1; [M-H]-1 |

**Generate Expected Compounds**

- dealkylation / dearylation
- Combinatorial common transformation list

# Customizable Workflow For Identifying Both Expected and Unexpected



# THANK YOU FOR YOUR ATTENTION

## Urinary metabolomics reveals kynurenine pathway perturbation in newborns with transposition of great arteries after surgical repair


Manuela Simonato[1,2] · Igor Fochi[3] · Luca Vedovelli[2] · Sonia Giambelluca[4] · Cristiana Carollo[1] · Massimo Padalino[5] · Virgilio P. Carnielli[6] · Paola Cogo[7]

1. Anesthesiology and Intensive Care Unit, Department of Medicine-DIMED, University of Padova, Padua, Italy
2. PCare Laboratory, Fondazione Istituto di Ricerca Pediatrica, Padua, Italy
3. Thermo Fisher Scientific, Milan, Italy
4. Women and Child Health Department, University of Padova, Padua, Italy
5. Pediatric and Congenital Cardiac Surgical Unit, Department of Cardiac, Thoracic and Vascular Sciences, Padova University, Padua, Italy
6. Division of Neonatology, Department of Clinical Sciences, Polytechnic University of Marche and Azienda-Ospedaliero Universitaria Ospedali Riuniti, Ancona, Italy
7. Division of Pediatrics, Department of Medicine, Udine University, Udine, Italy