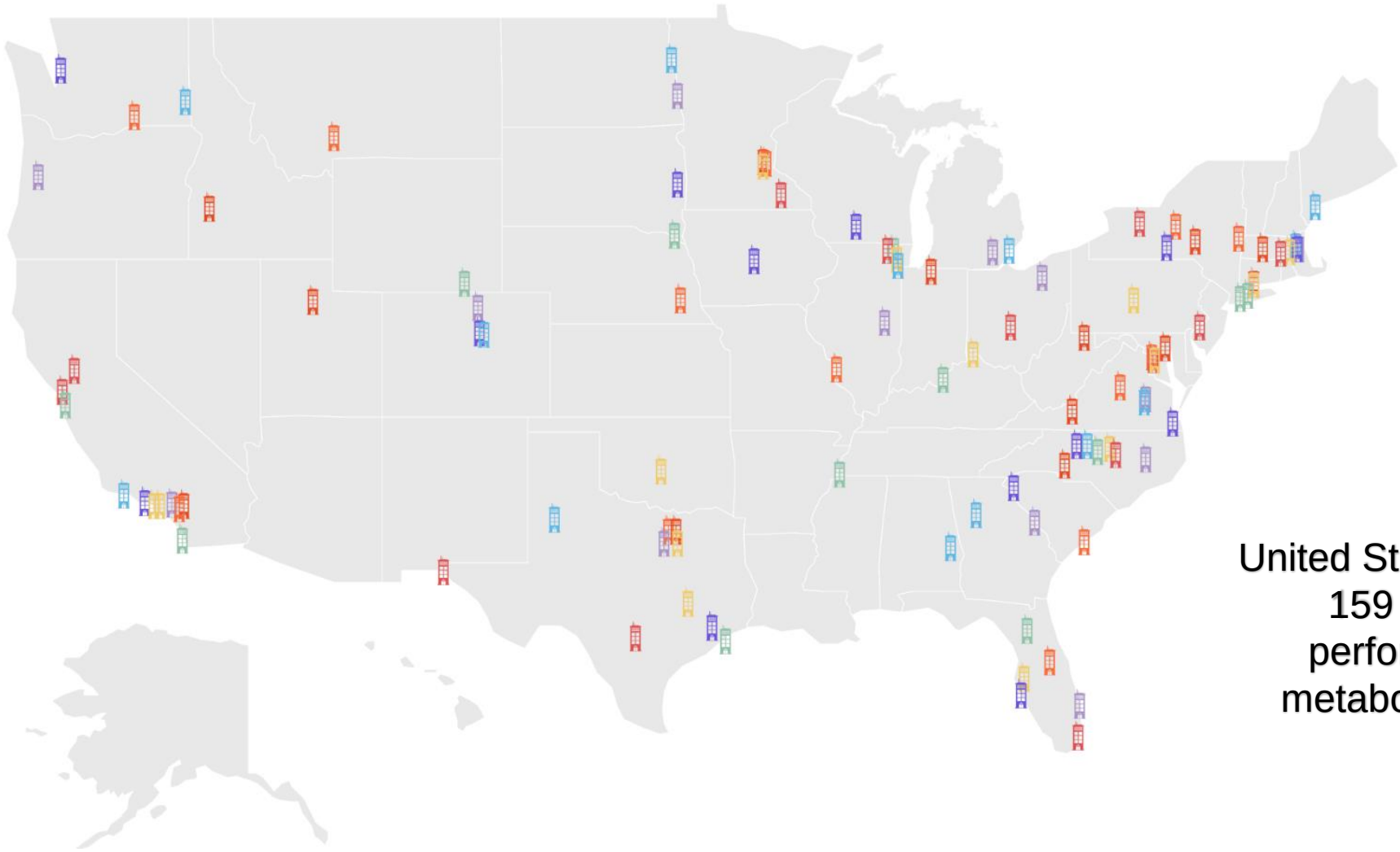
ThermoFisher SCIENTIFIC

Metabolomics in an Identity Crisis? Am I a Feature or a Compound?

Agenda

1. **Unknown Analysis in Metabolomics**
2. **Determining Features and Compounds for Data Reduction**
3. **Multiple Measures for Confident Identification**

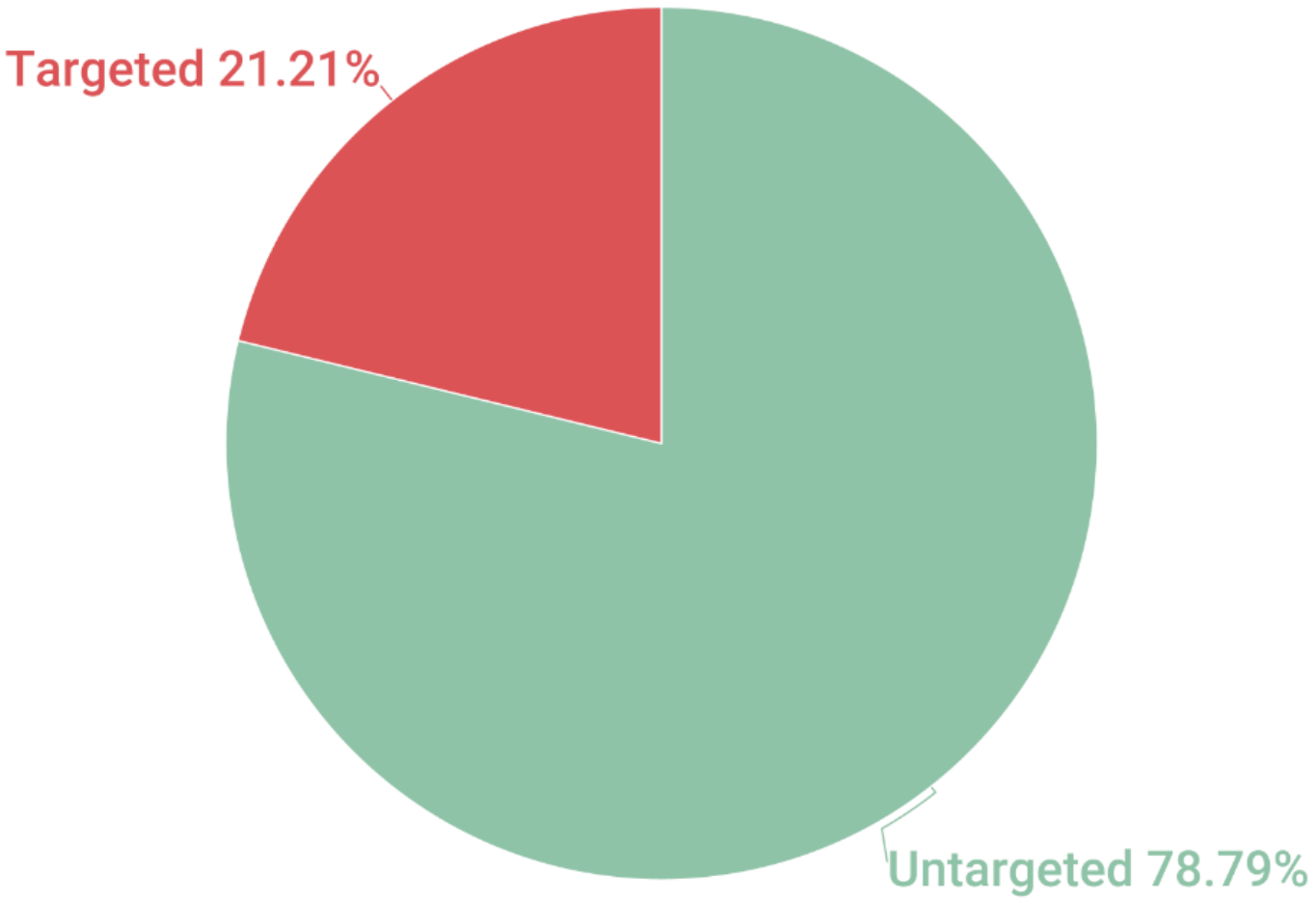
Popularity of Metabolomics



United States only:
159 labs
performing
metabolomics

Untargeted Metabolomics Takes the Lead

99
journal articles featuring Thermo Scientific™ MS instruments



Thermo Scientific™ MS instruments included: *Thermo Scientific™ Q Exactive™ MS, Thermo Scientific™ Q Exactive™ Plus MS, Thermo Scientific™ Q Exactive™ HF MS, Thermo Scientific™ Orbitrap ™ Fusion Tribrid MS, Thermo Scientific™ Orbitrap™ Fusion Tribrid MS, Thermo Scientific™ TSQ Quantiva™ Quadrupole MS*

ThermoFisher SCIENTIFIC

Why untargeted metabolomics?

Advantages

- Hypothesis generating step
- Unbiased peak detection for discovery
- Don’t need to know what to look for ahead of analysis
- Capture the entire metabolome → collect everything, miss nothing
- Minimal method optimization
- Retrospectively mine data

Challenges

- Endogenous metabolites are chemically diverse
- Isobaric/isomeric species
- Databases and libraries are still growing
- Tools for true unknown unknowns
- What measures constitute identification?
- How can we be confident?

Two Types of Unknown Compounds

Known Unknowns

- Knowledge of the sample (biofluid, plant tissue, environmental)
- Expected target compounds
- Search existing databases and spectral libraries for known species

Unknown Unknowns

- True unknown
- Novel compound
- Structure elucidation and characterization

Process Overview

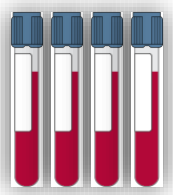
Sample Collection

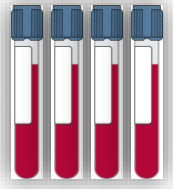
Data Acquisition

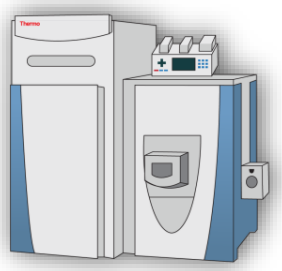
Data Analysis

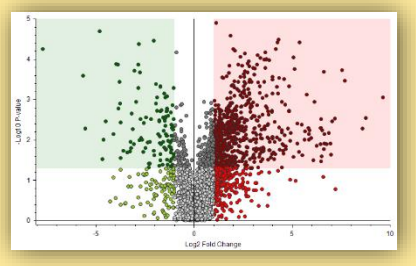
Compound Annotation

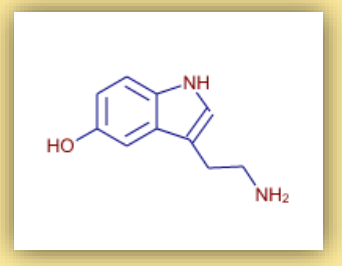
Healthy vs Disease





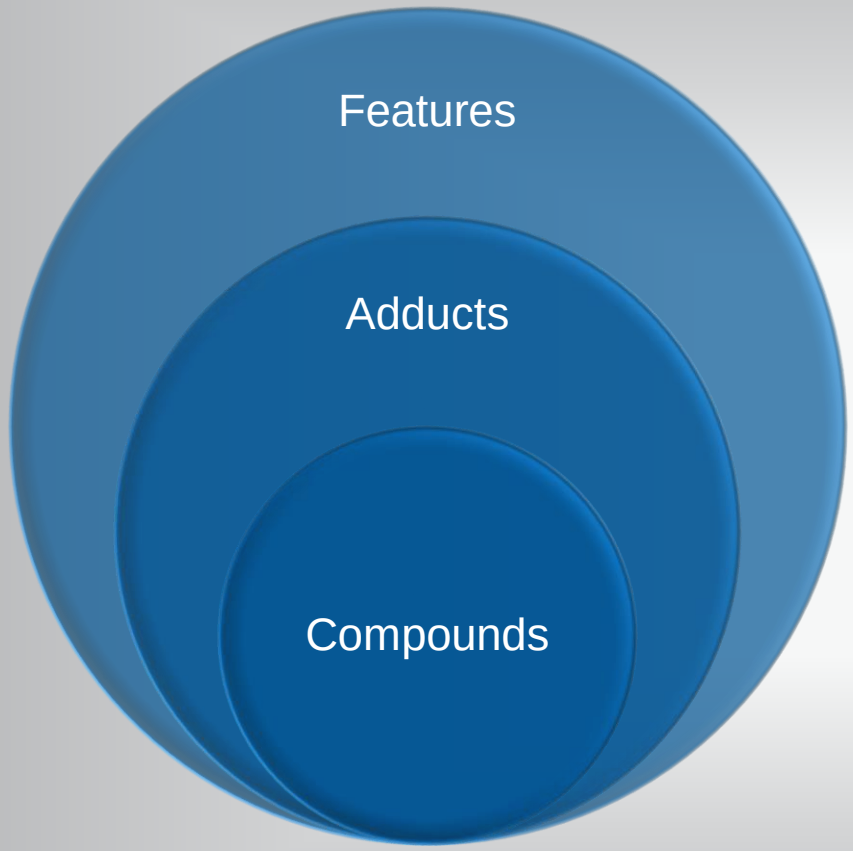




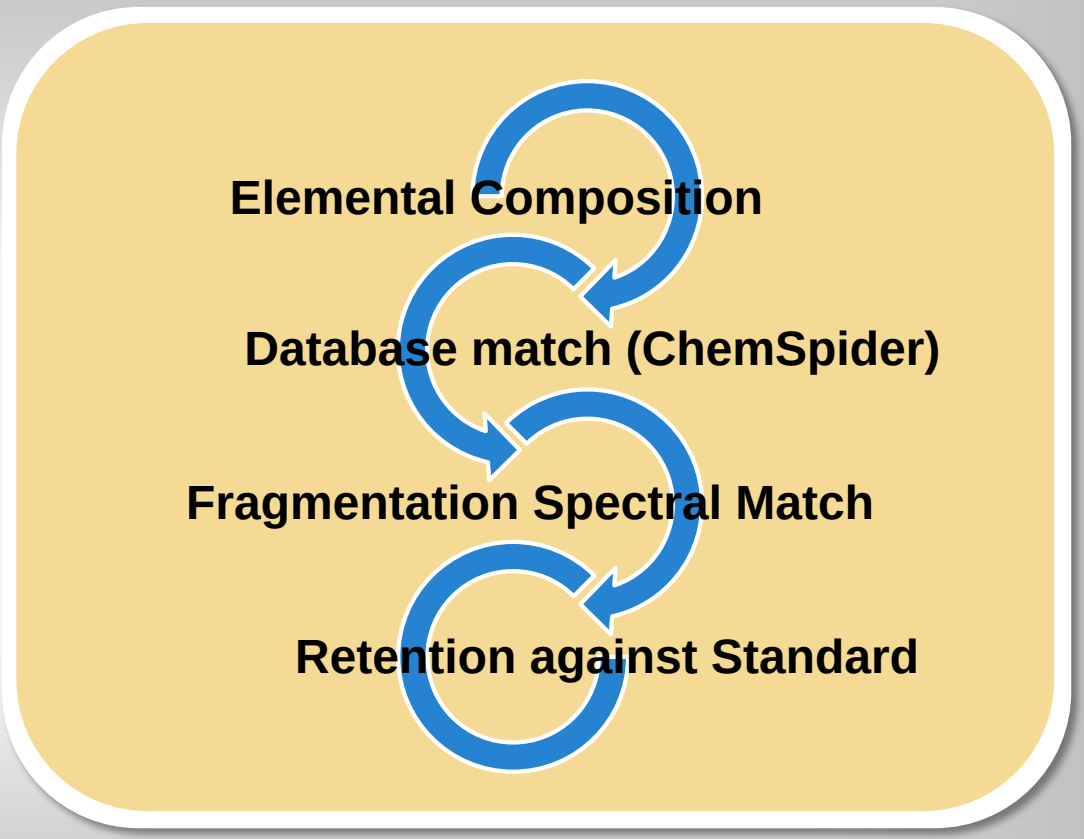


Conquering the Unknown Identification Space with Confidence

Data Reduction



Confident Identification



Data Refinement Process

Identification Tools for Unknown Analysis

Fully Integrated

ThermoFisher SCIENTIFIC

ThermoFisher SCIENTIFIC

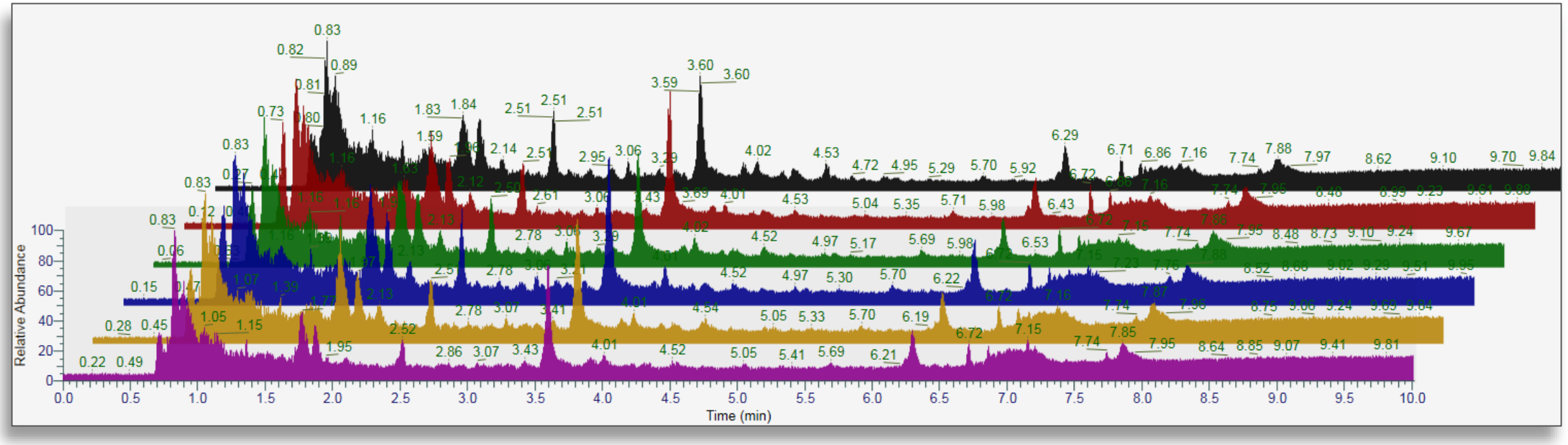
Example Analysis with Sake: Data Reduction

Untargeted Metabolomics with Sake

- Sake → Japanese rice wine
- Six bottles selected from local CA market for comparison
- Are they different biochemically?
- Thermo Scientific Ultimate 3000 RSLC - Q Exactive Plus MS
- Example data available with Compound Discoverer

Untargeted Acquisition Aims to Collect All Data Points

MS1 Full Scan Mode

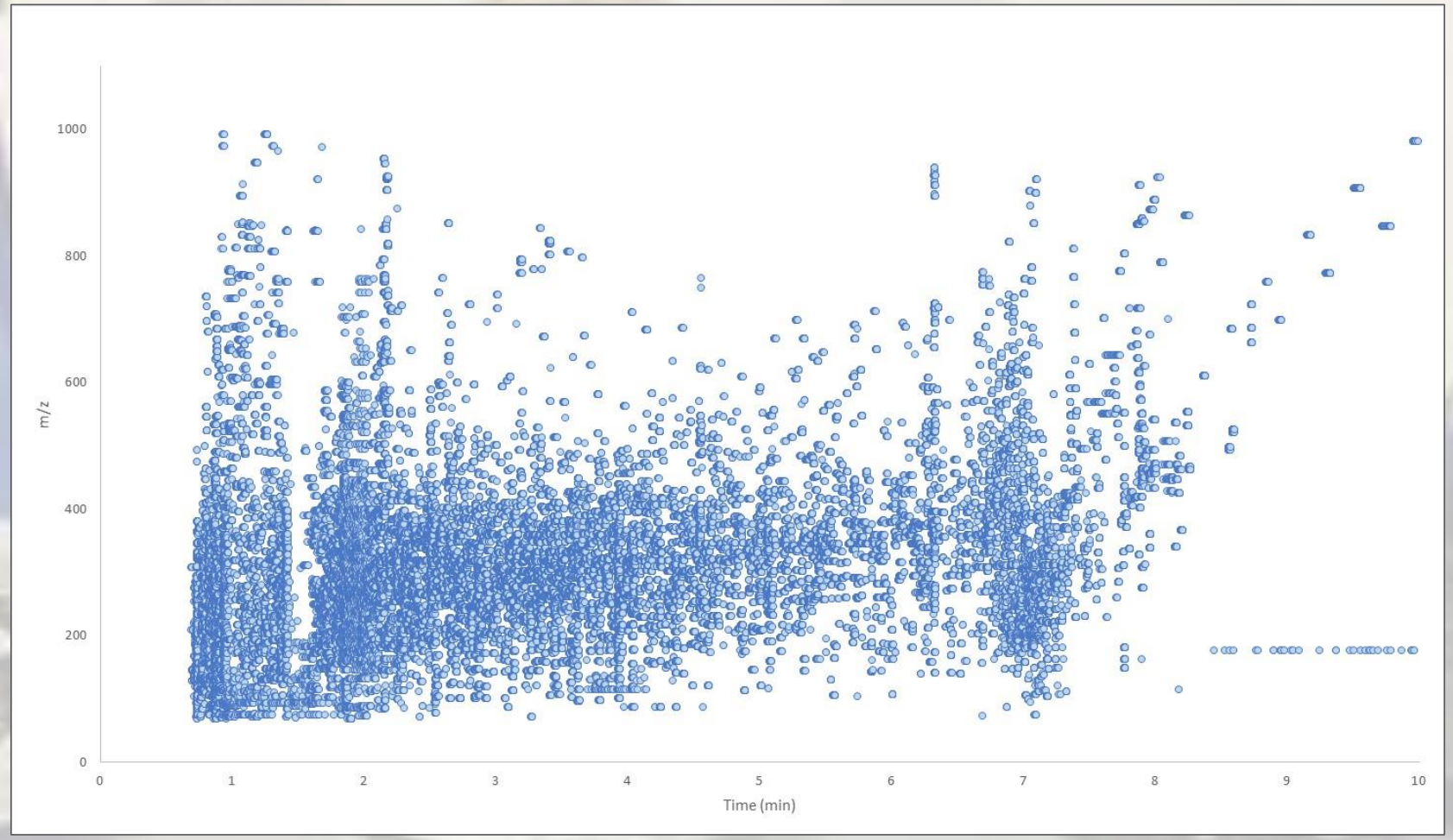


Chromatographic Overlay for Six Different Sake Bottles

ThermoFisher SCIENTIFIC

Unbiased Peak Detection

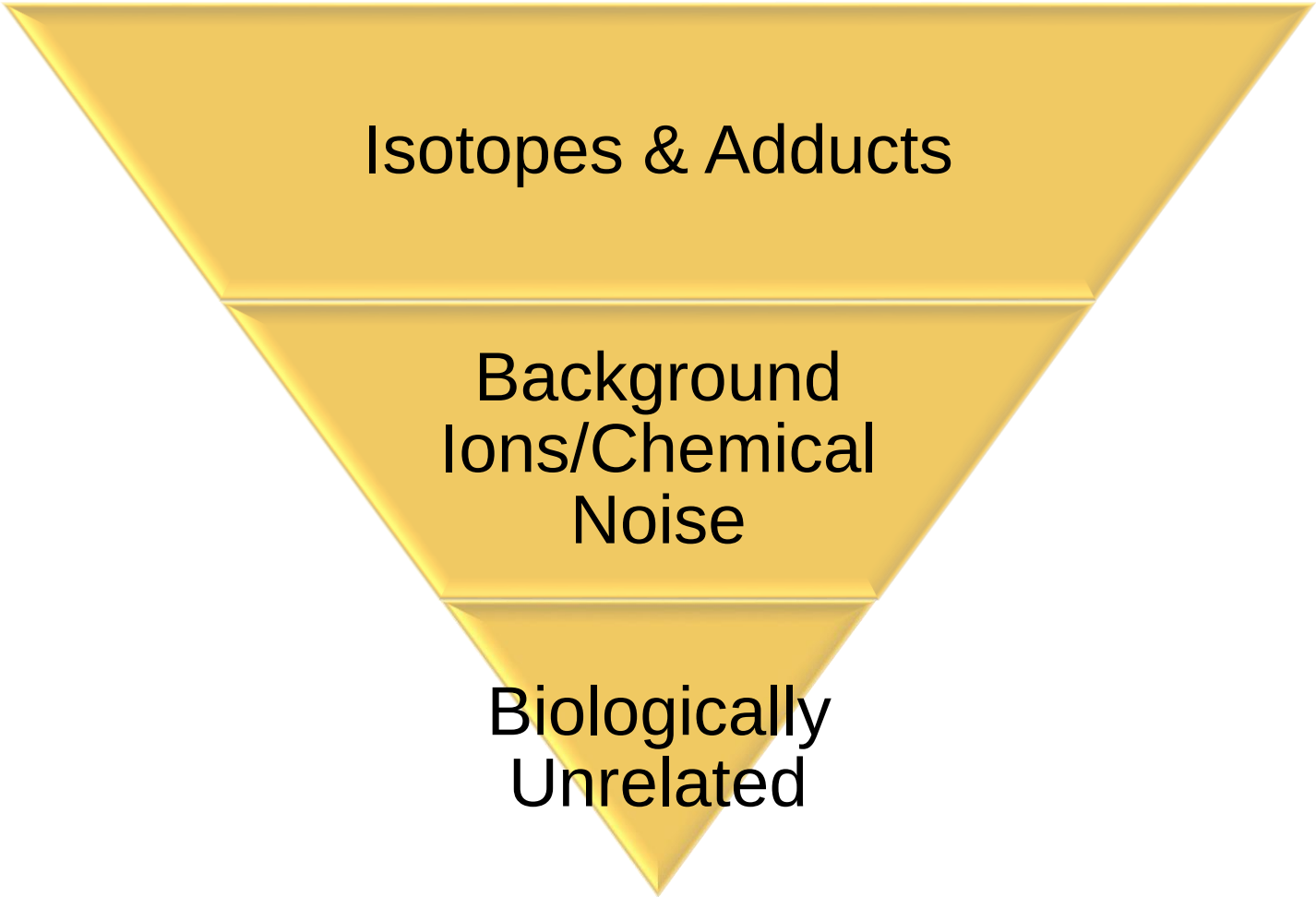
Full Scan Mode Generates Many Data Points



ThermoFisher SCIENTIFIC

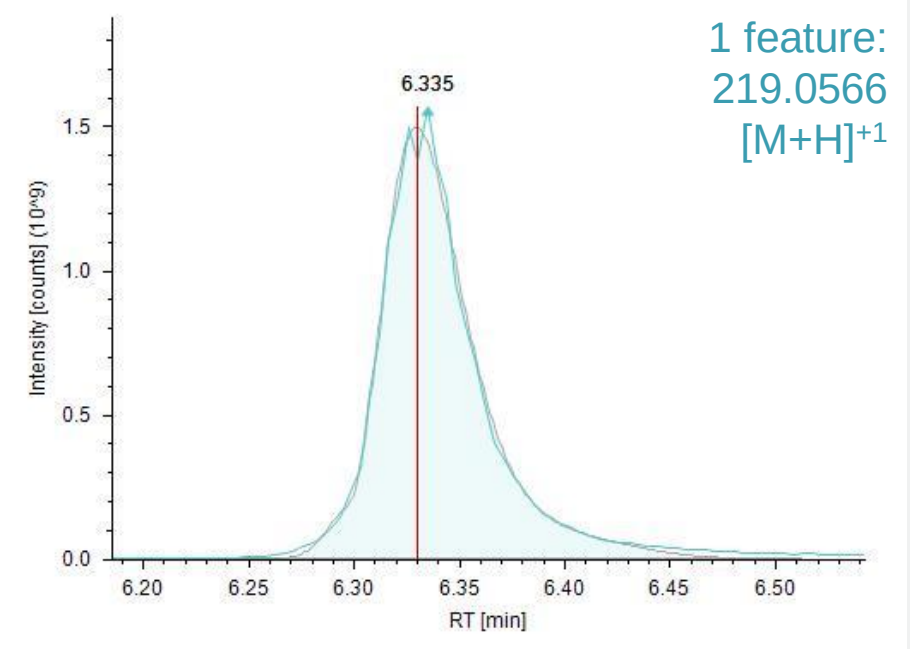
Data Reduction with Intelligent Mining

Not All Unknowns Are Equal



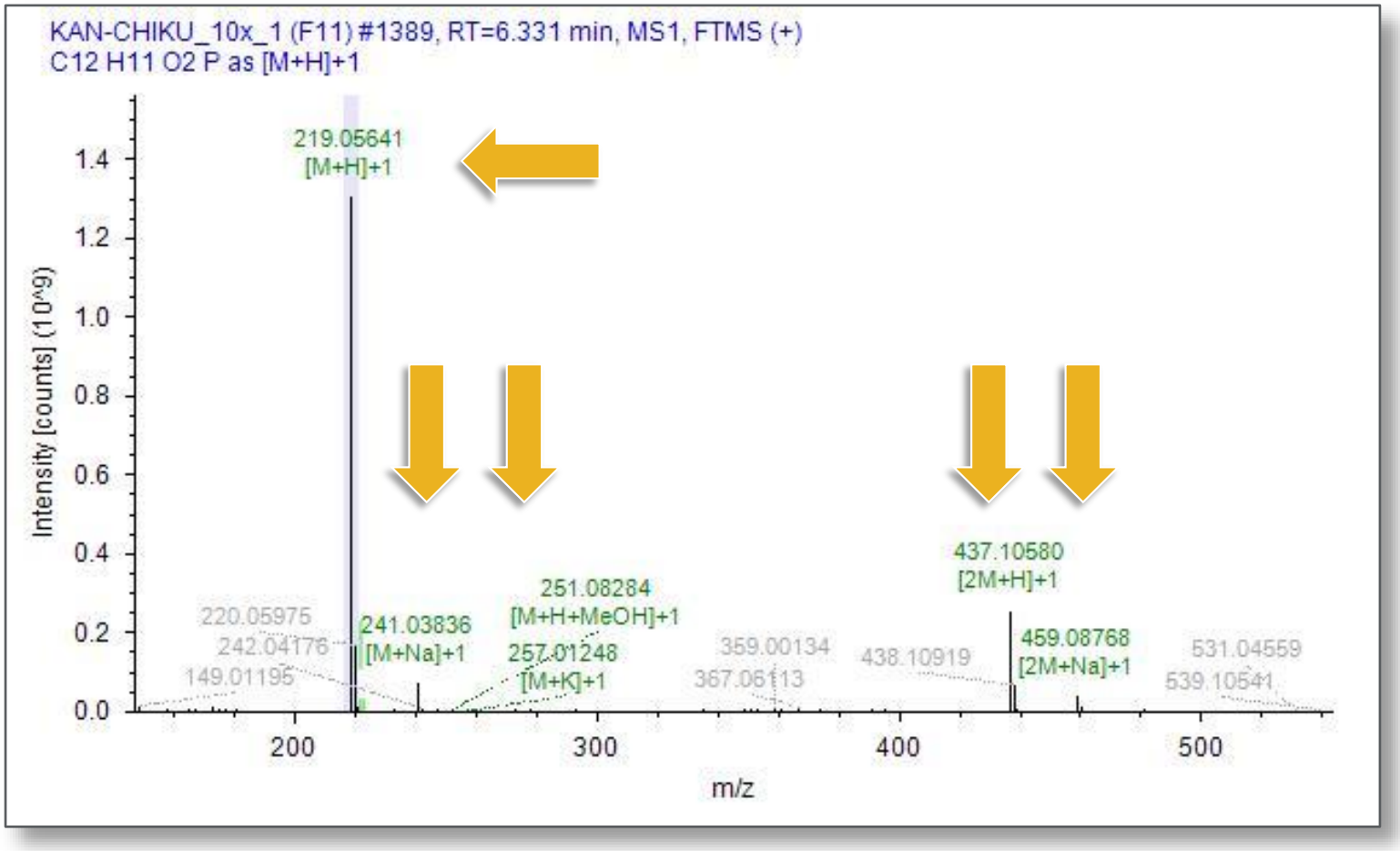
What Is a Feature?

Feature



Distinct m/z at a given retention time

Feature Assembly



$[M+H]^{+1}$

$[M+Na]^{+1}$

$[M+K]^{+1}$

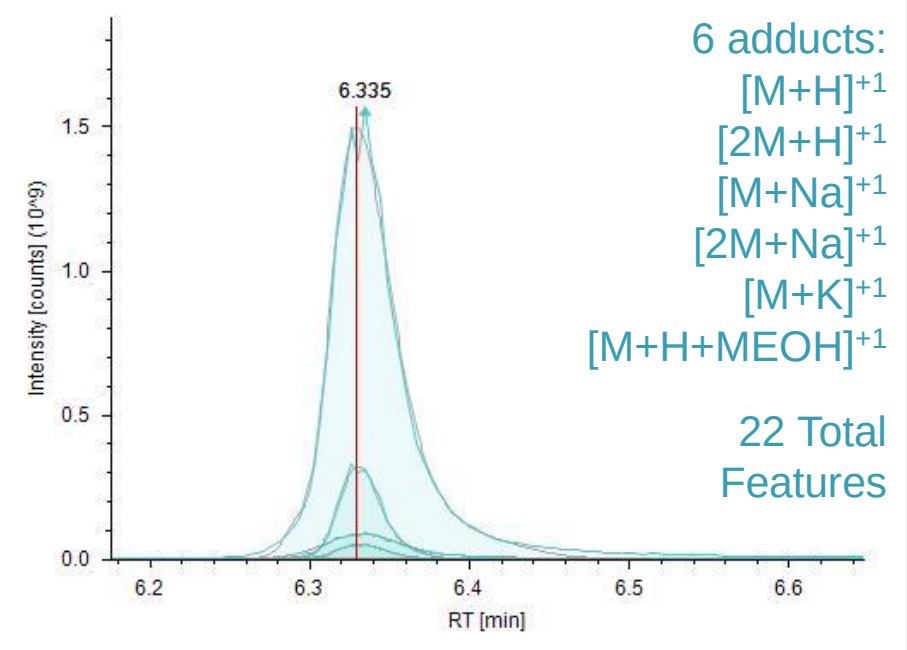
$[M+H+MEOH]^{+1}$

$[2M+H]^{+1}$

$[2M+Na]^{+1}$

What is a Compound?

Compound

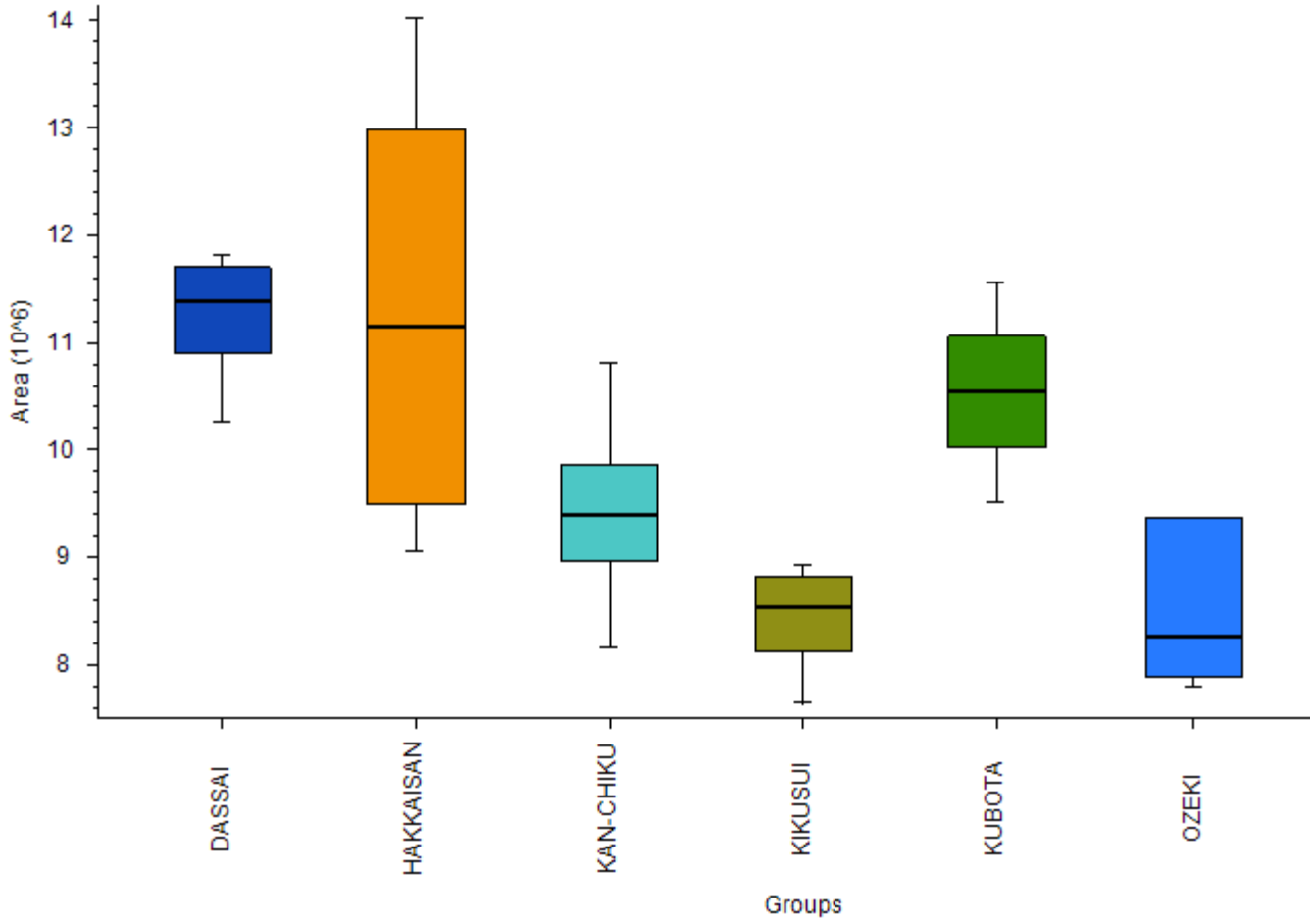


6 adducts:
$[M+H]^{+1}$
$[2M+H]^{+1}$
$[M+Na]^{+1}$
$[2M+Na]^{+1}$
$[M+K]^{+1}$
$[M+H+MEOH]^{+1}$

22 Total Features

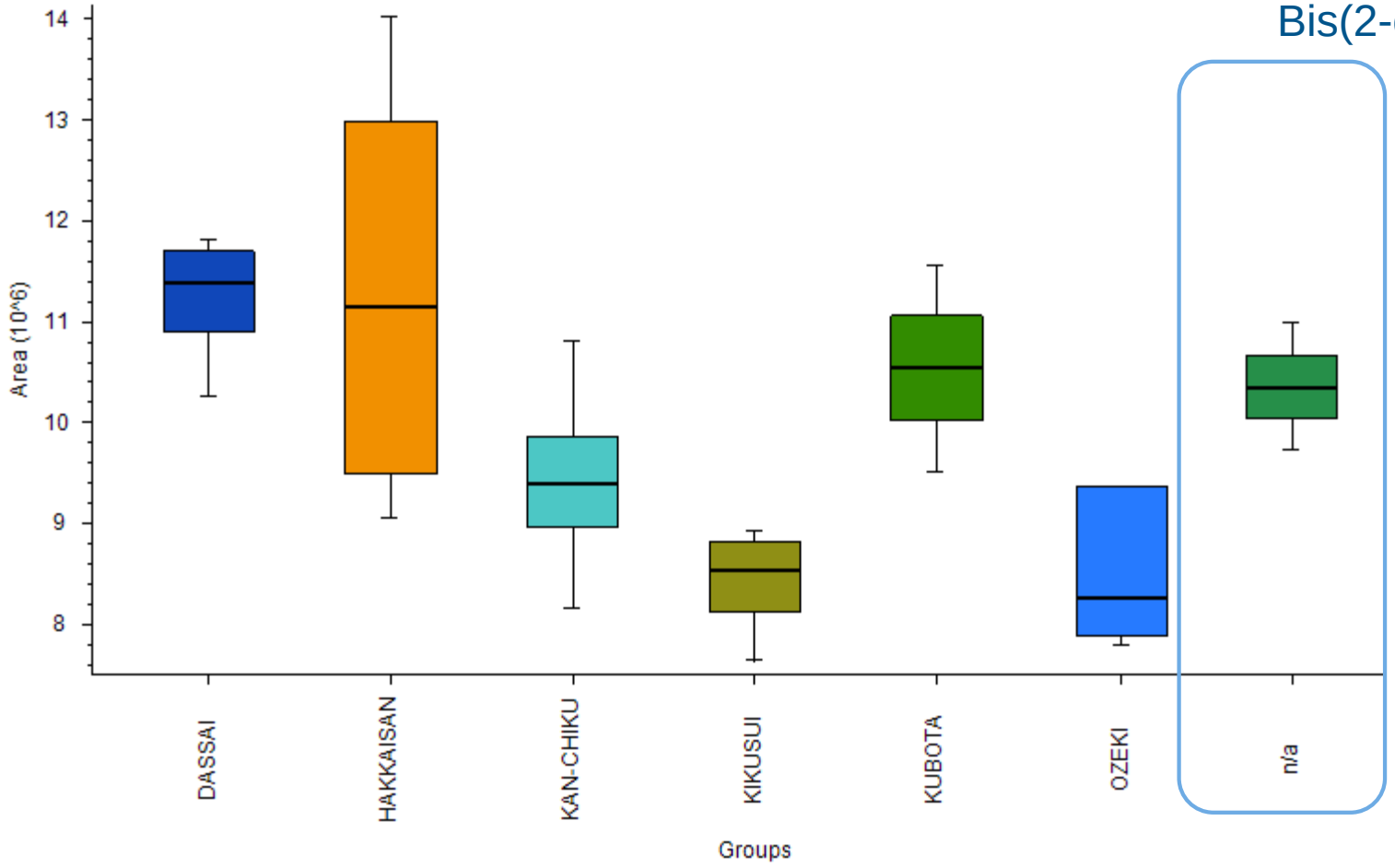
Combination of associated features at a given retention

Is This Compound Important?

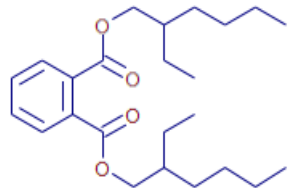


Nope. It's a Background Compound

Found in all samples, including solvent blank.



Bis(2-ethylhexyl) phthalate



n/a: water blanks

ThermoFisher SCIENTIFIC

Nope. It's a Background Compound

Found in all samples, including solvent blank.

Compound Discoverer Flags Background Compounds for Easy Filtering!

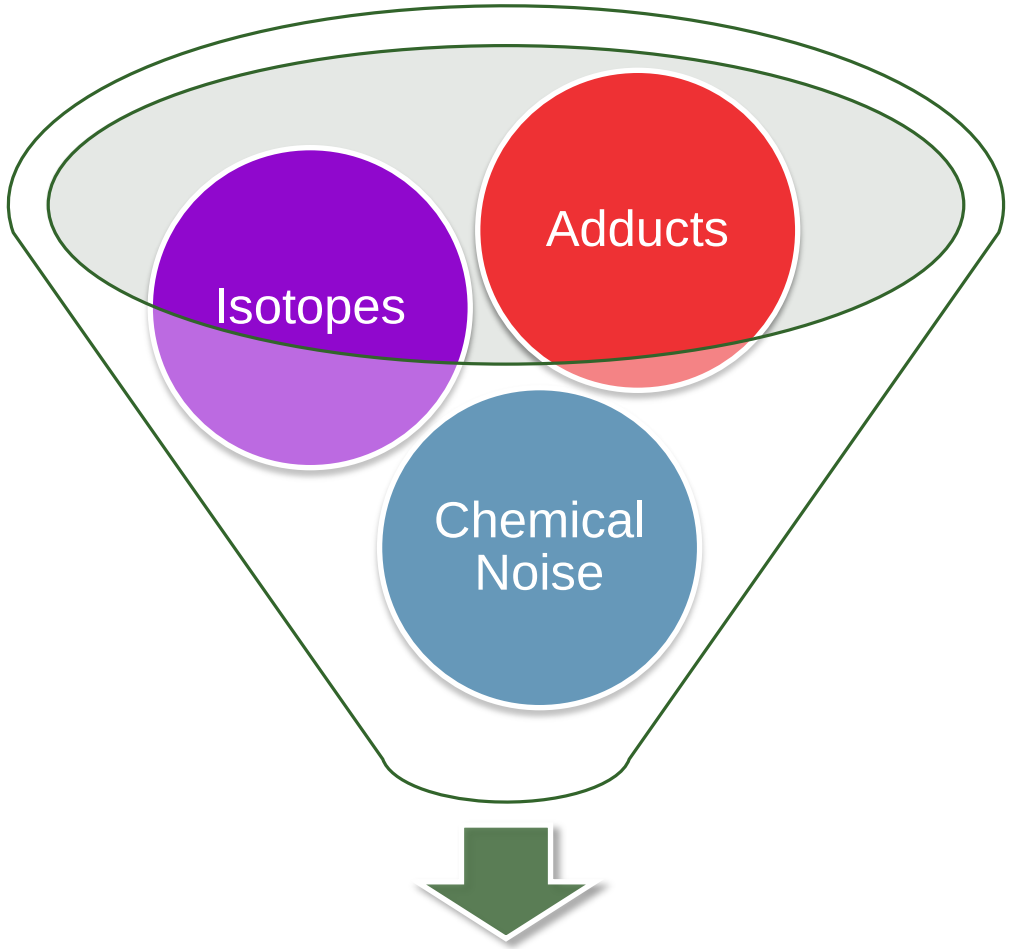
n/a: water blanks

ThermoFisher SCIENTIFIC

Expedited Data Reduction

Stop Chasing Nonsense

Focus on What Is Truly Valid



Component Assembly – Detecting Real Peaks

Over 219,000 unique features

Over 100,000 unique isotope groups

6,230 unique compounds

new

Compound Discoverer Software

Integrated solutions for small molecule research applications in LC- HRAM MS

Compound Discoverer 2.1 Software

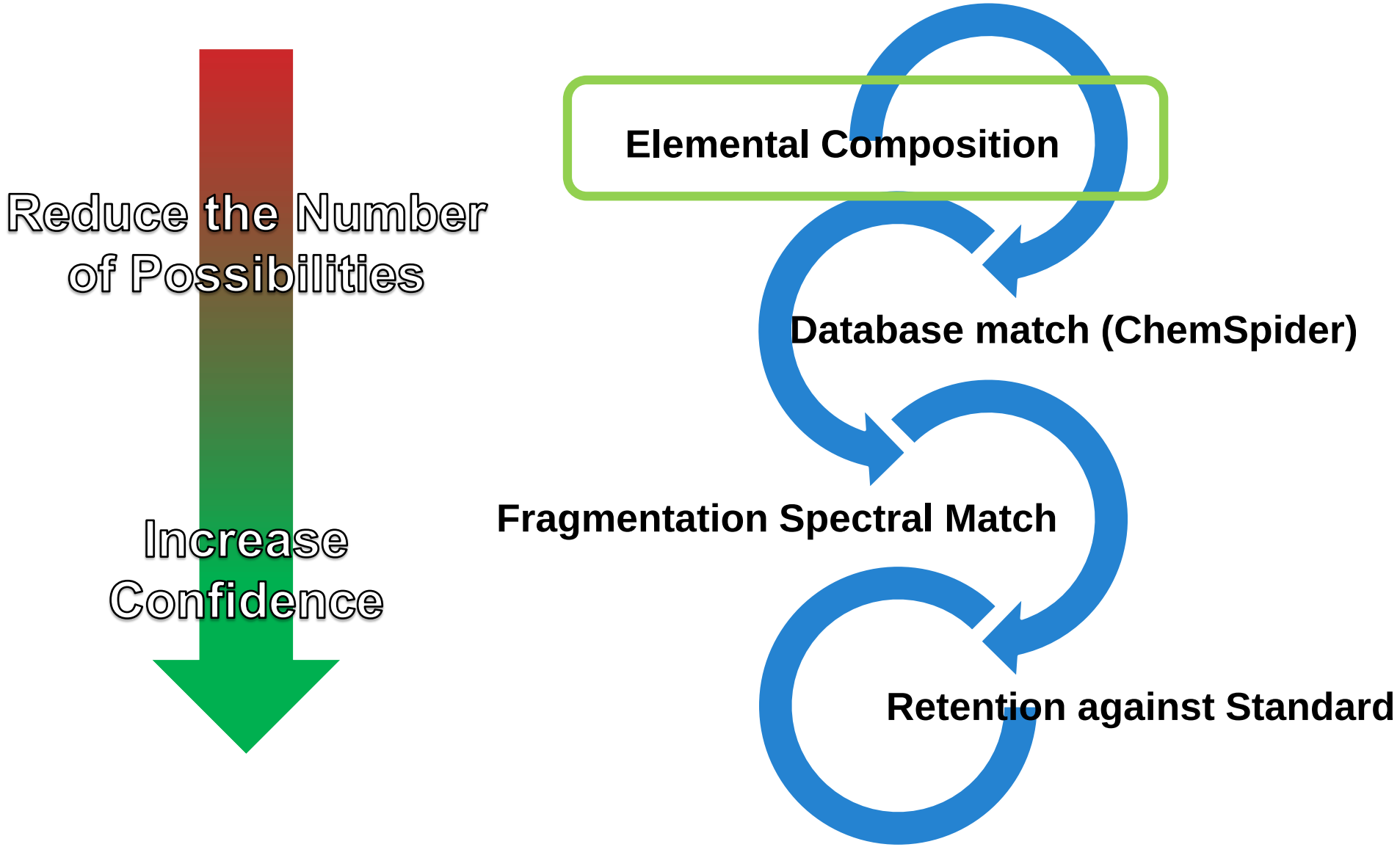
Small molecule research software for a wide range of applications.

ThermoFisher SCIENTIFIC

ThermoFisher SCIENTIFIC

Identification of Unknown Compounds in Sake Samples

The Path to Confident Identification



Value of Mass Accuracy in Predicting Elemental Composition

Reduce Potential Candidates with High Mass Accuracy

Possible Empirical Formulae
(C, H, N, O, P, S)

Metabolite	$[M+H]^{+1}$ (m/z)	10 ppm	5 ppm	3 ppm	1 ppm
Proline	116.0706	1	1	1	1
Acetylcarnitine	204.1230	4	1	1	1
Saccharopine	277.1394	15	9	5	2
S-adenosyl methionine	399.1437	85	44	26	9
Bilirubin	585.2707	224	111	67	24

Mass error (ppm) = exact mass – accurate mass/MW*10^6

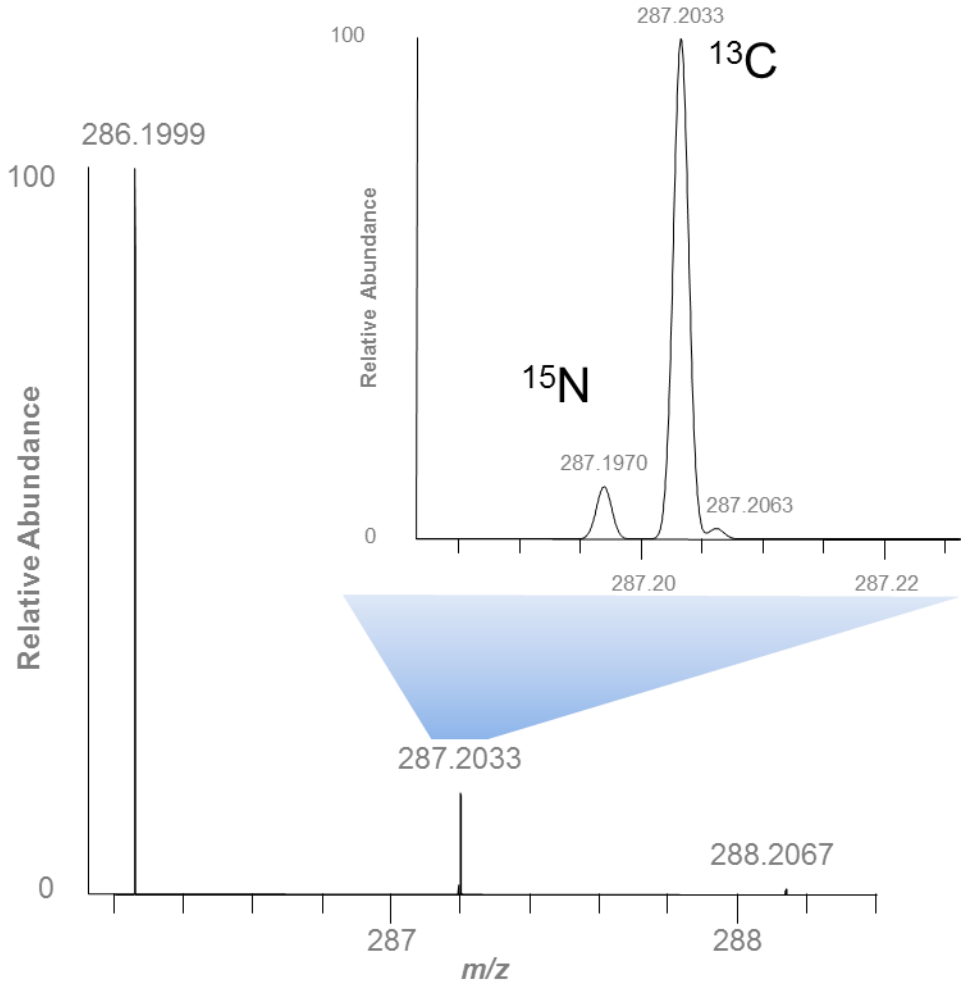
Using Very High Resolution Accurate Mass

$C_{11}H_{24}O_2N_7$

or

~~$C_{15}H_{28}O_4N$~~

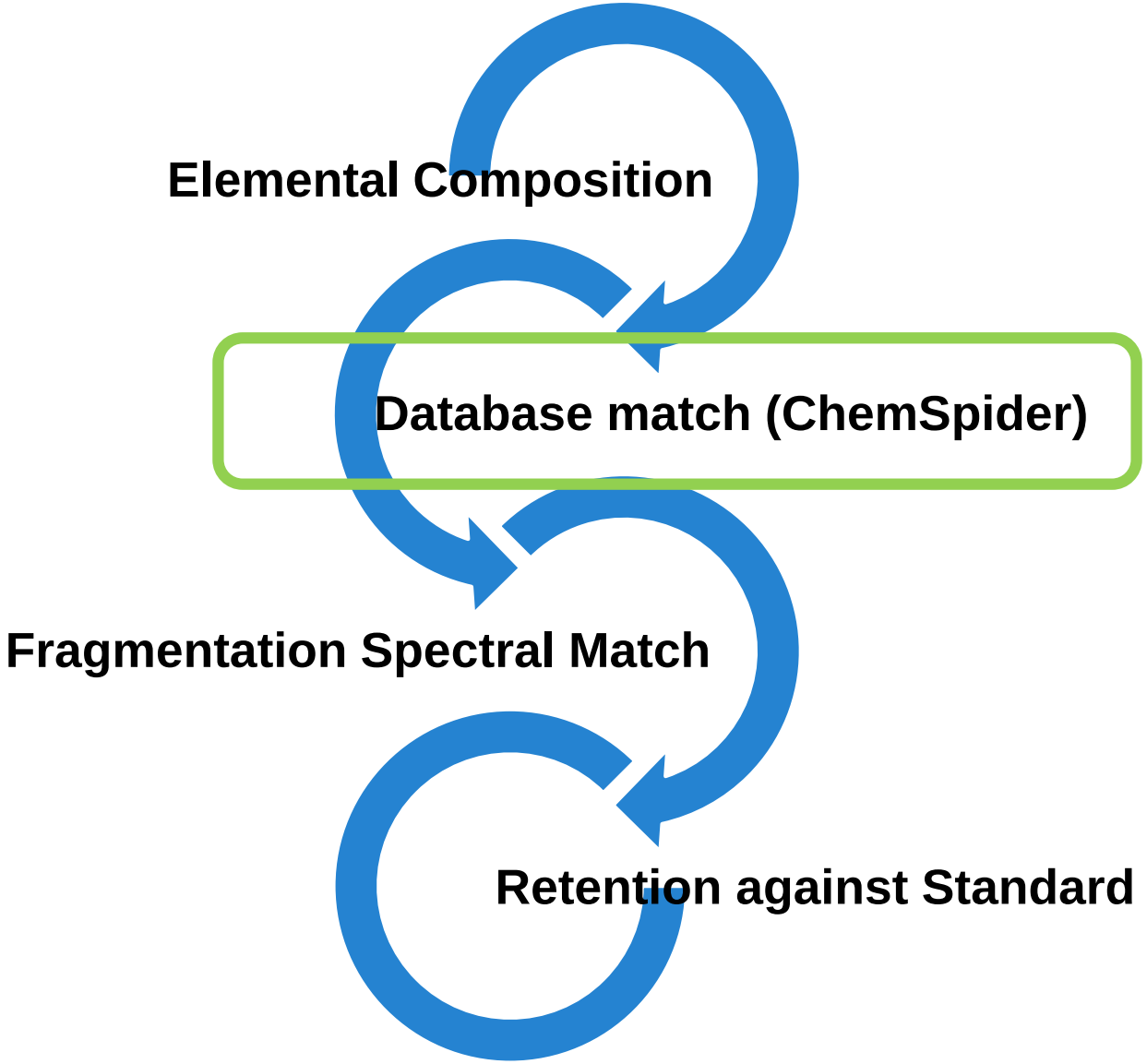
- Accurate mass and fragmentation are not the only tools available to us
- Accurate mass gives us access to elemental composition
- Very high resolutions give fine isotopic structure to help determine correct elemental composition



VHRAM

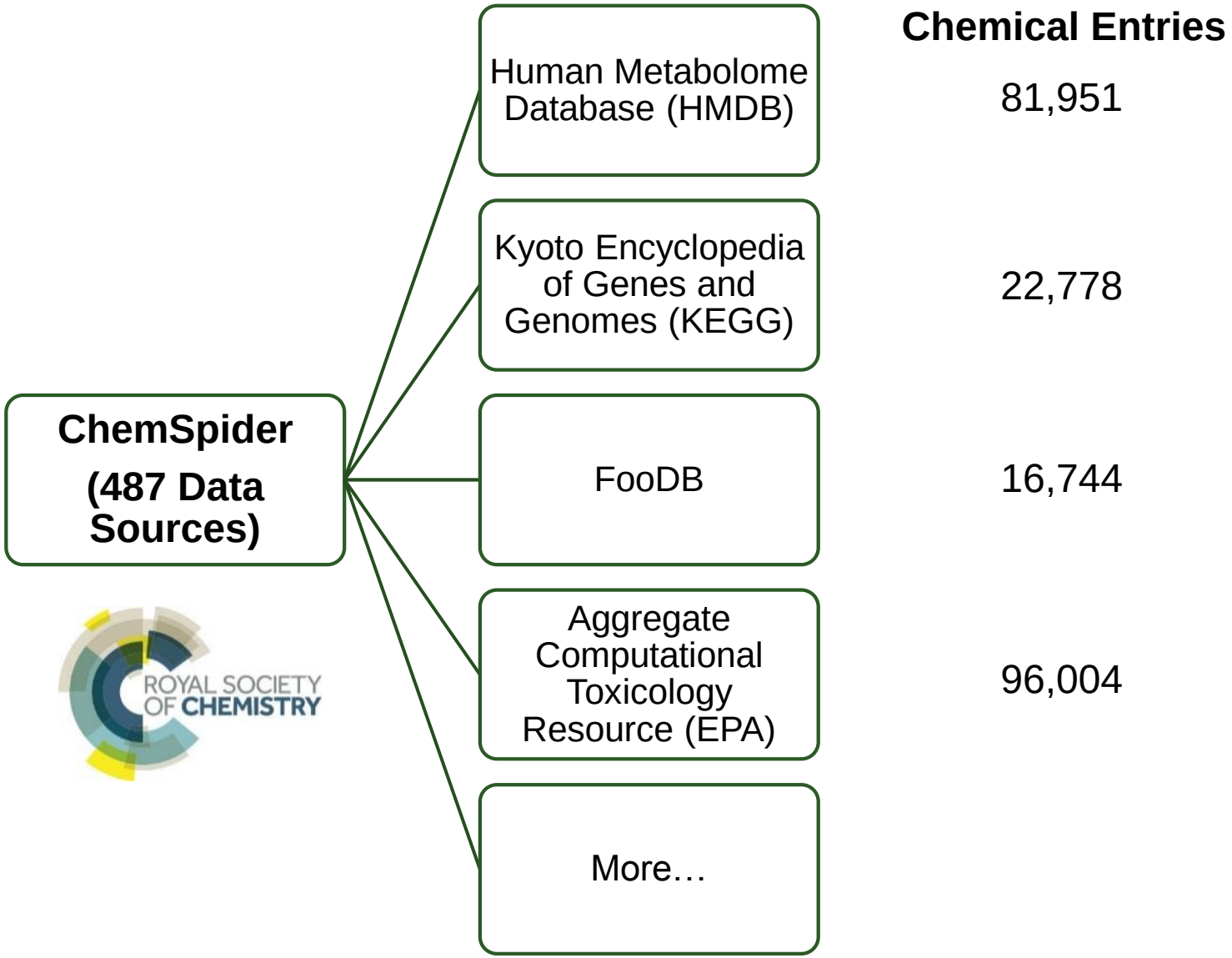
ThermoFisher SCIENTIFIC

The Path to Confident Identification



Search Existing Databases

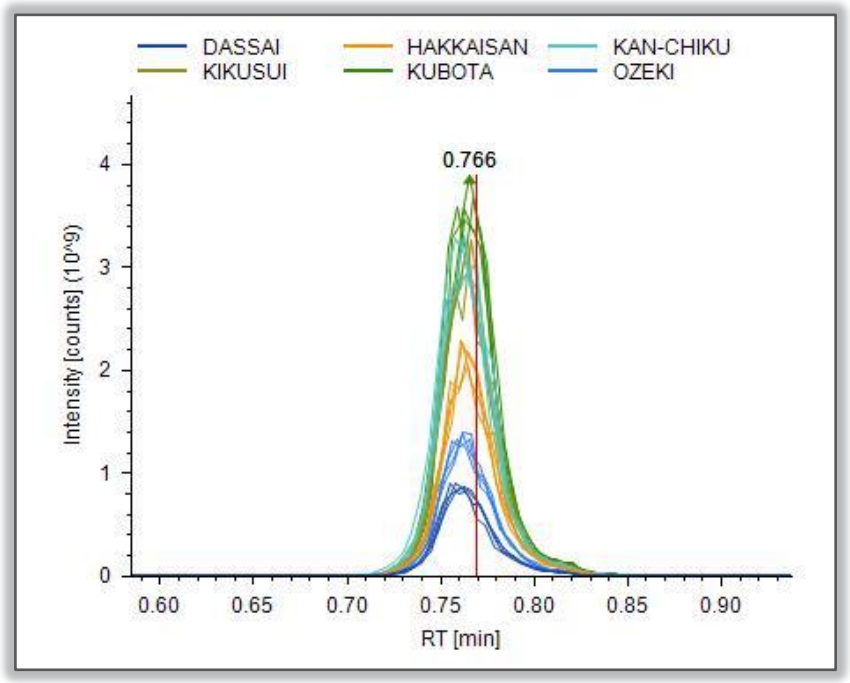
Select Experimentally Relevant Data Sources to Reduce Erroneous Candidates



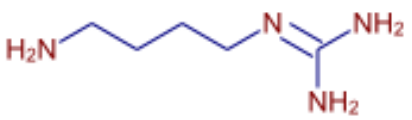
ThermoFisher SCIENTIFIC

The Potential Peril of Database-only “ID”

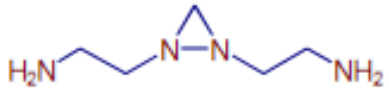
Molecular Weight → Elemental Composition → ChemSpider Search



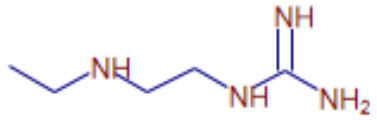
Multiple Candidates



Agmatine



1,2-Diaziridinediethanamine



1-[2-Ethylamino)ethyl]guanidine

ThermoFisher SCIENTIFIC

The Potential Peril of Database-only “ID”

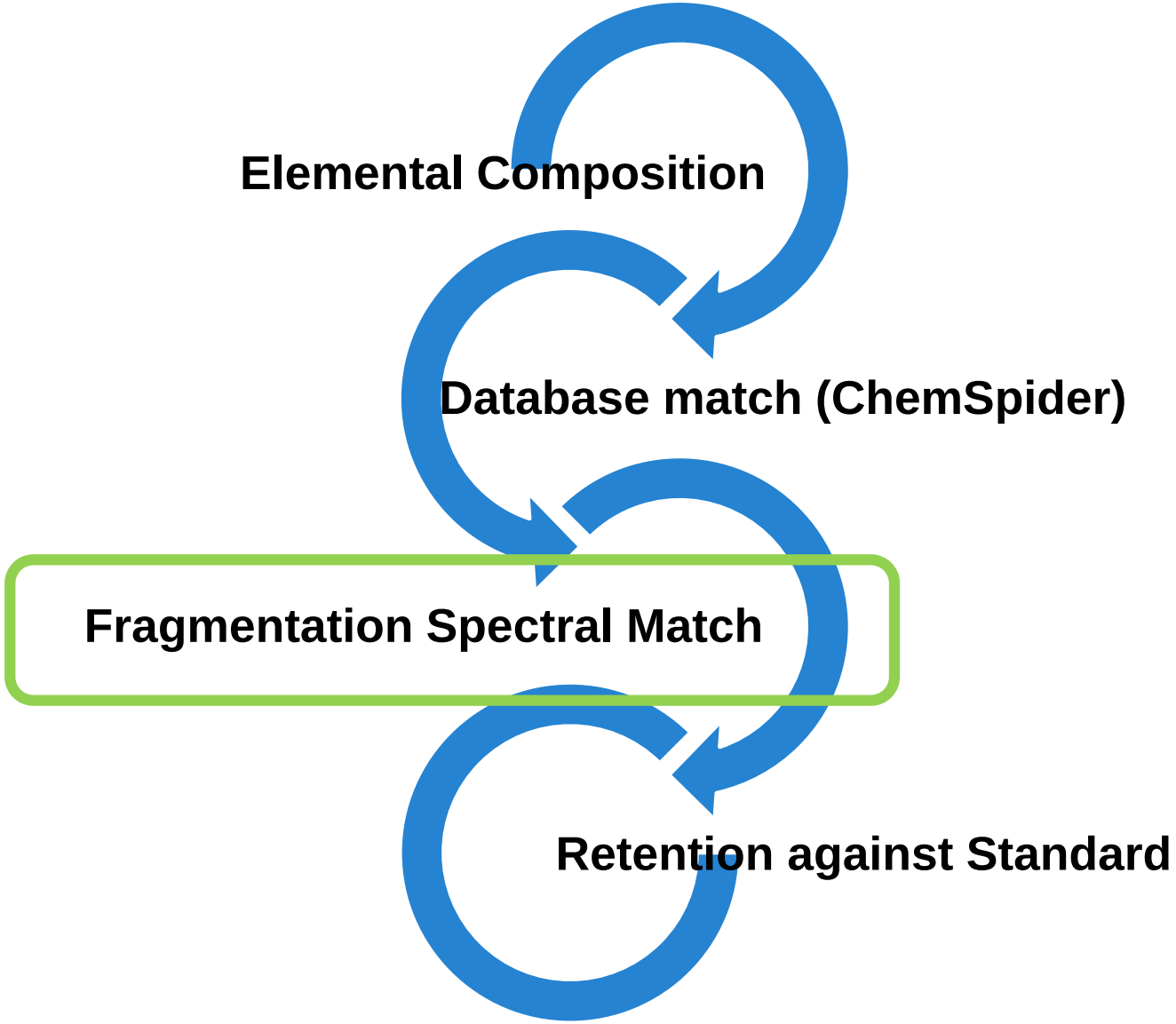
Molecular Weight → Elemental Composition → ChemSpider Search

Multiple Candidates

Which one is it?

Let’s fragment!

The Path to Confident Identification



www.mzCloud.org

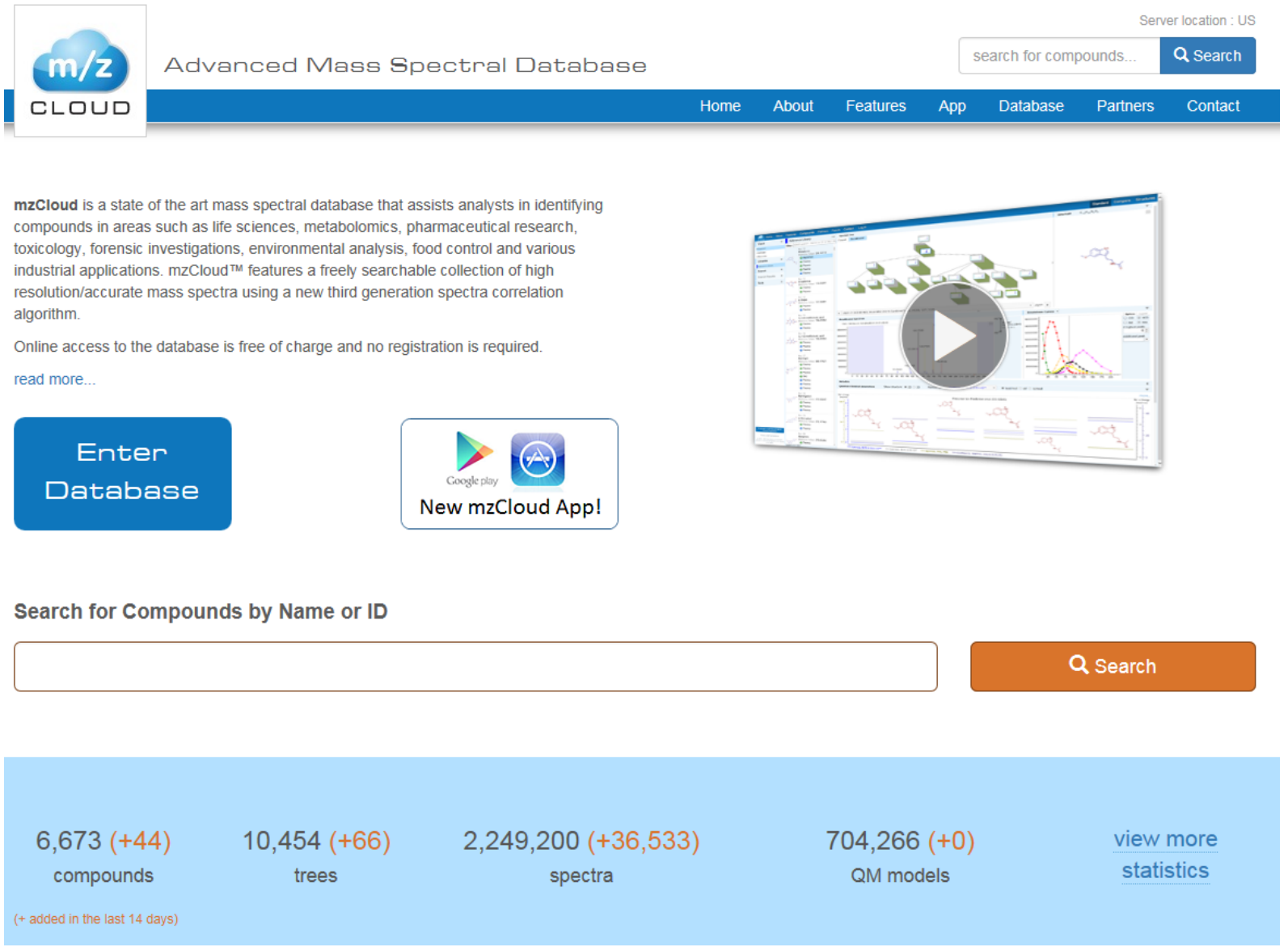
mzCloud

Ultra high quality online reference MS/MS and MS^n spectral library

Highly curated data – superior quality

High resolution accurate mass MS/MS and MS^n data – identify more unknowns with similarity searching.

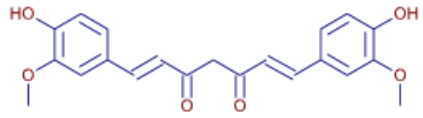
mzCloud (www.mzCloud.org)



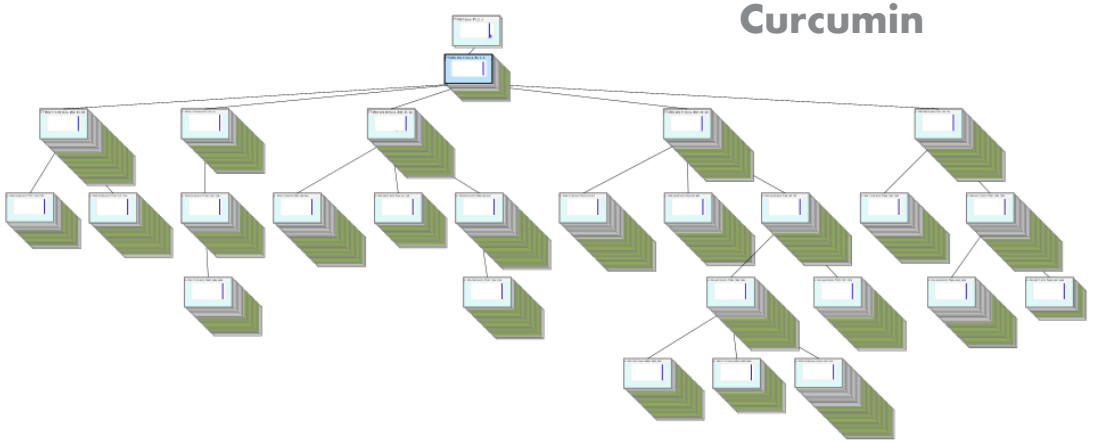
ThermoFisher SCIENTIFIC

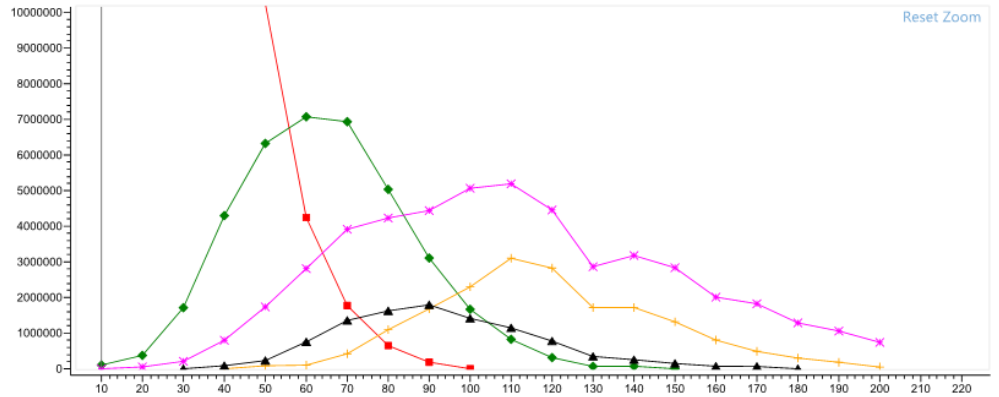
Extensive MS/MS and MS^n

- Dozens to thousands of spectra per compound
- 10-20 Different HCD Energies
- Dynamically optimized trap CID energy
- No limits on how you run your instrument

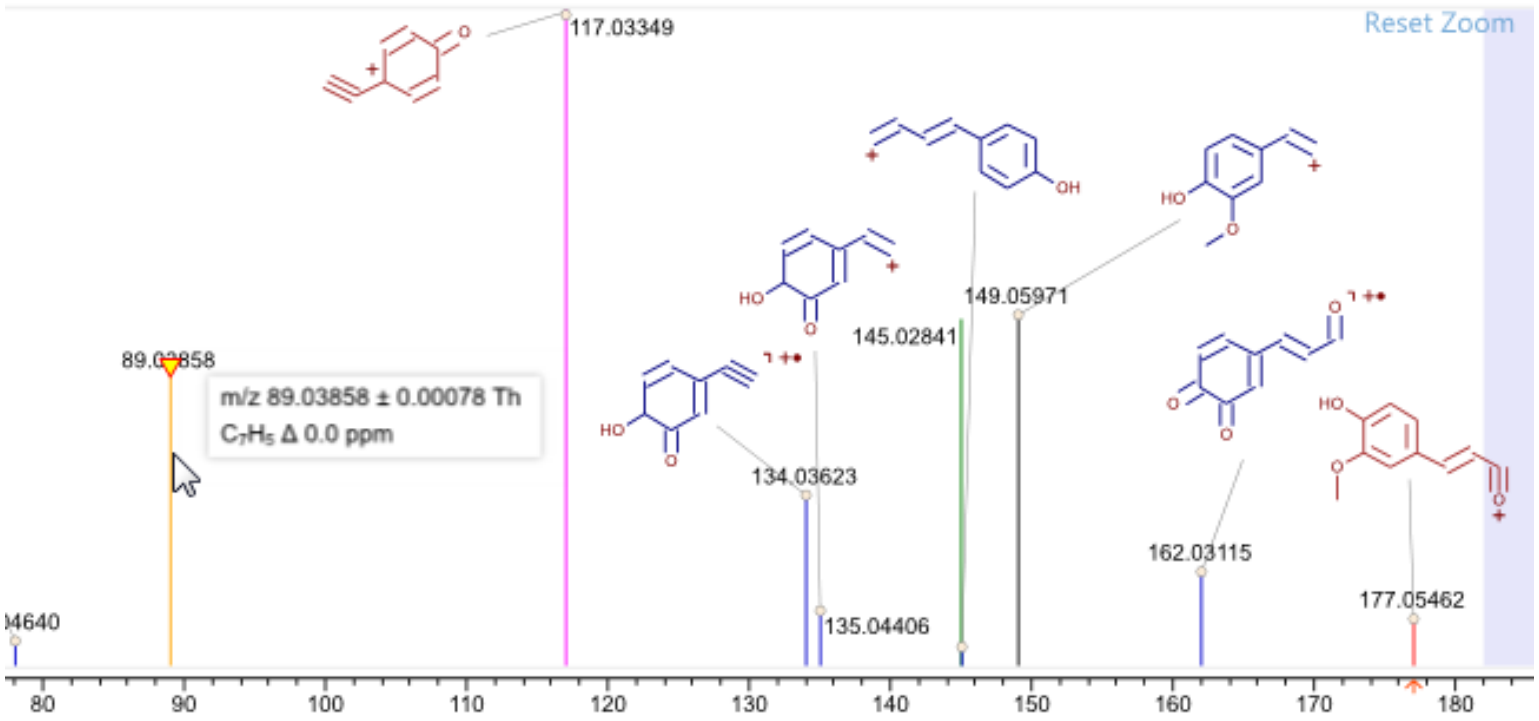


Curcumin





Highest Quality



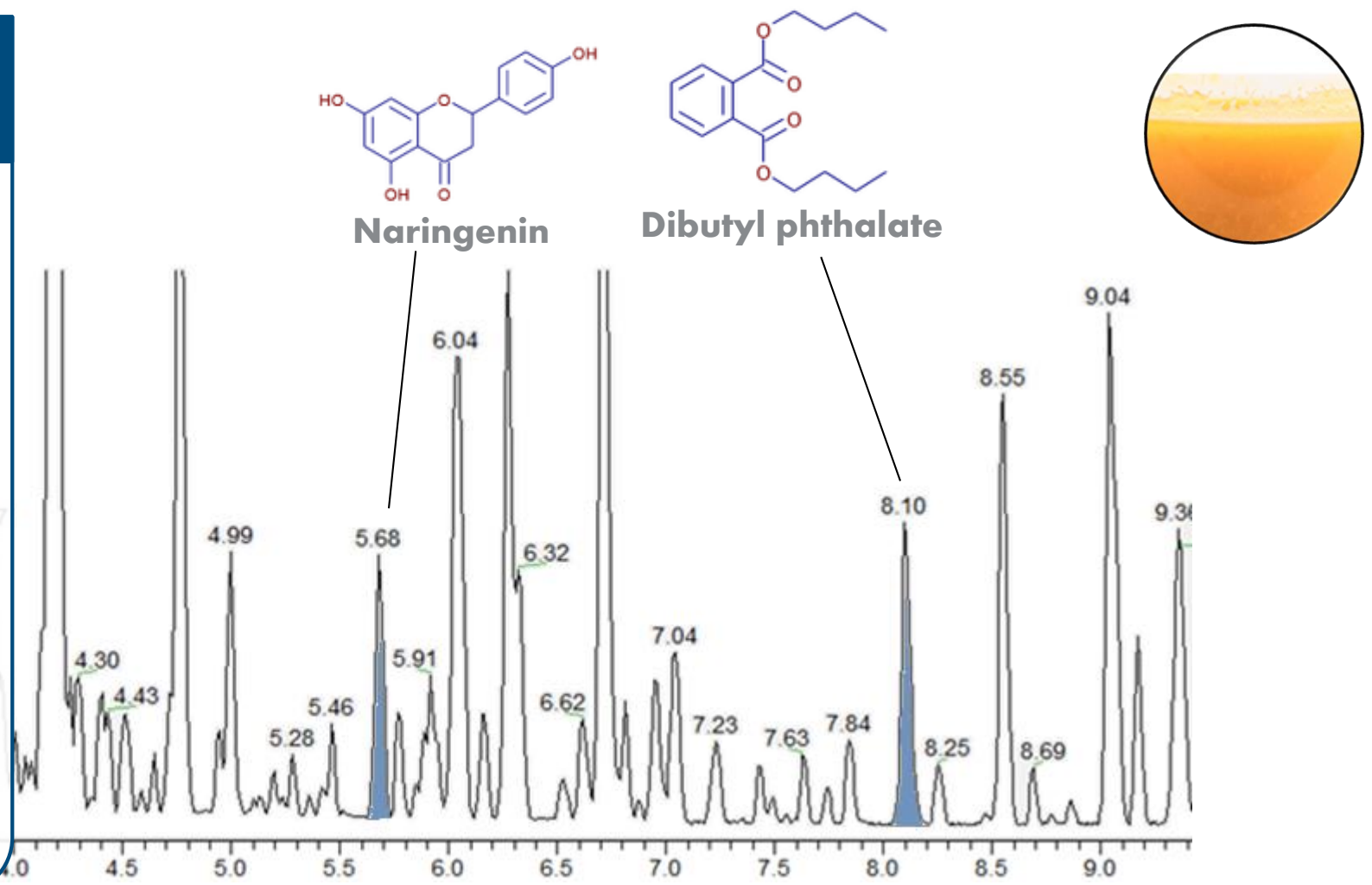
Highly Curated Data

- **Replicate spectra acquired for each energy**
- **Noise filtered**
- **Every spectra is recalibrated**
- **Extensive annotation of structure, formula, and neutral loss**

Diversity is Key

Wide Chemical Space

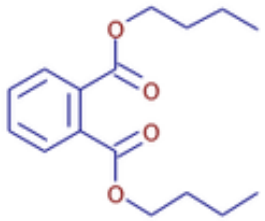
- 16 different broad chemical classes based on application
- Identify a more diverse range of compounds
- Handle complex samples
- Know more unknowns



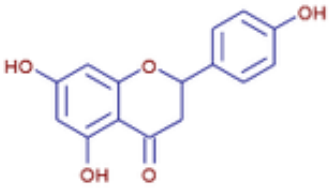
ThermoFisher SCIENTIFIC

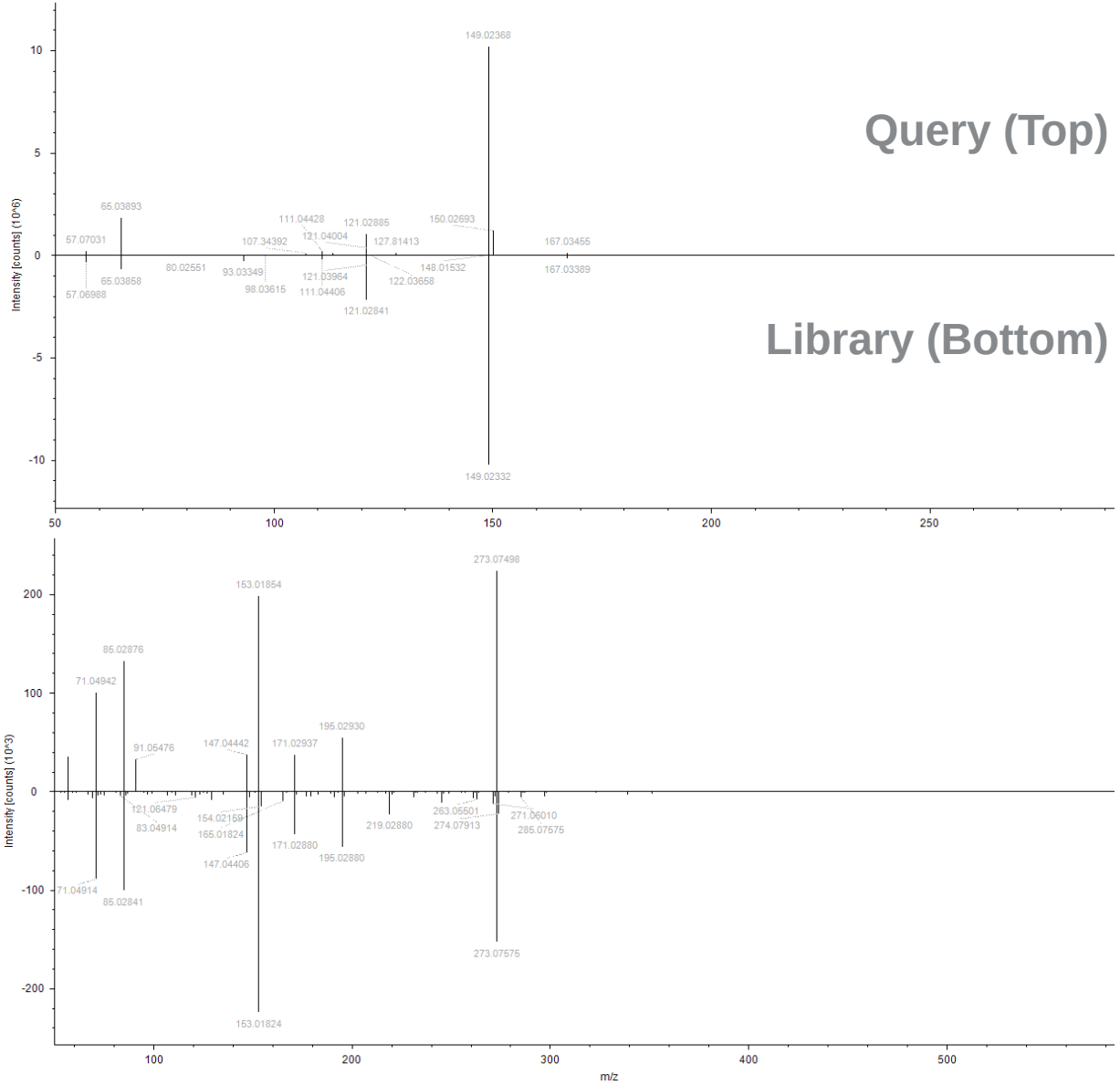
Identify More Than Just Endogenous Species

Dibutyl phthalate
m/z 279.1591



Naringenin
m/z 273.0757

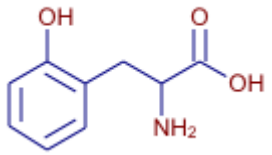


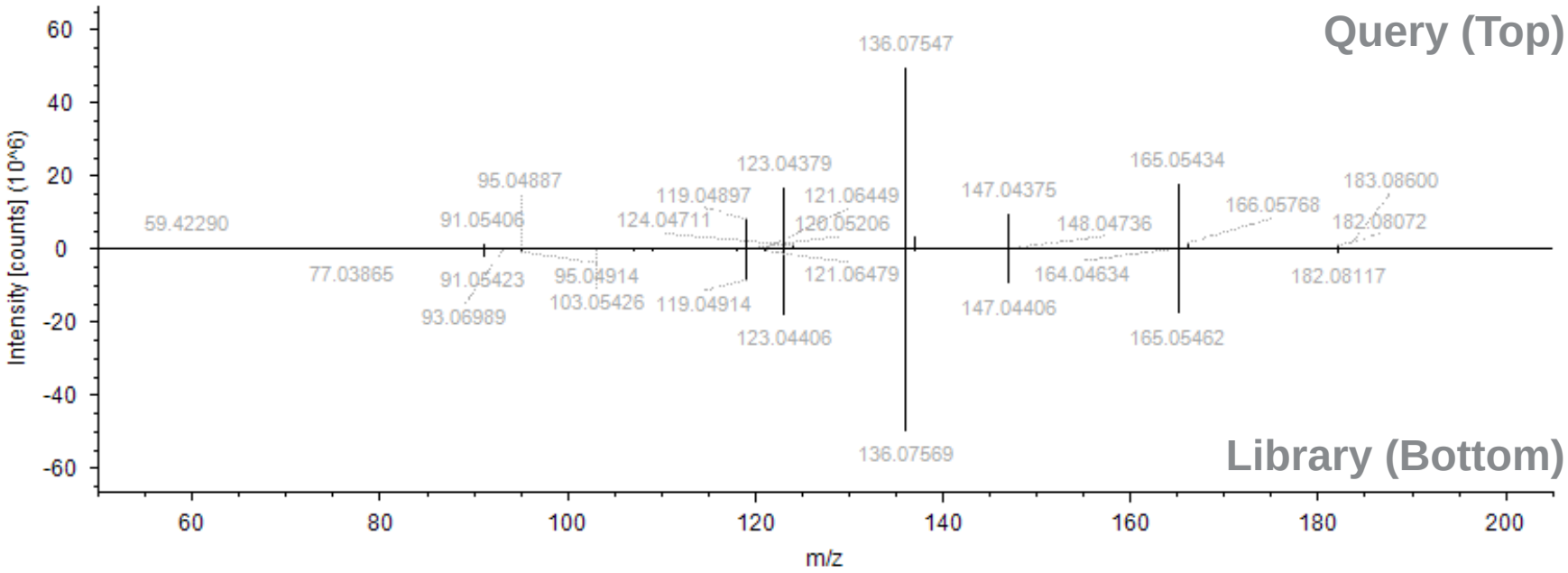


More than just a "targeted library", mzCloud helps to even identify compounds you can comfortably "ignore"

Fragmentation Match – Identity Search

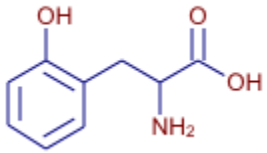
Tyrosine
m/z 182.0807



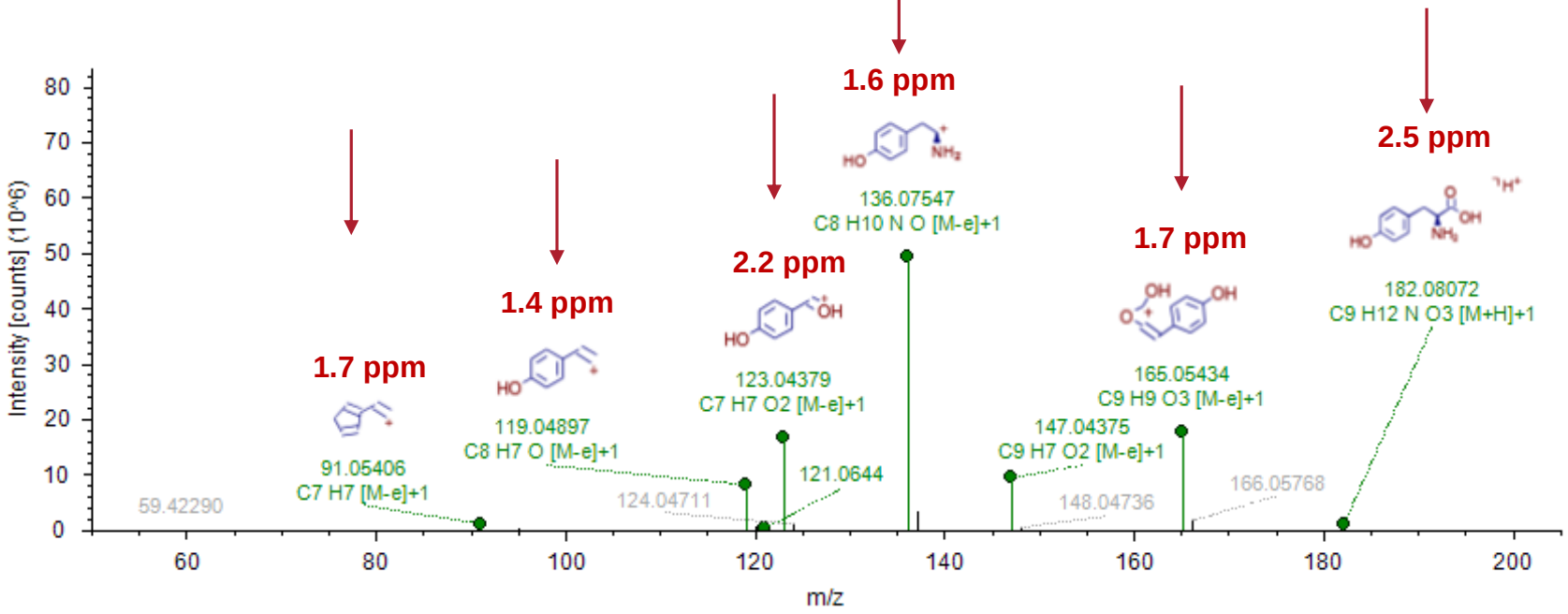


ThermoFisher SCIENTIFIC

Structural Annotation to Increase Confidence



Tyrosine
m/z 182.0807



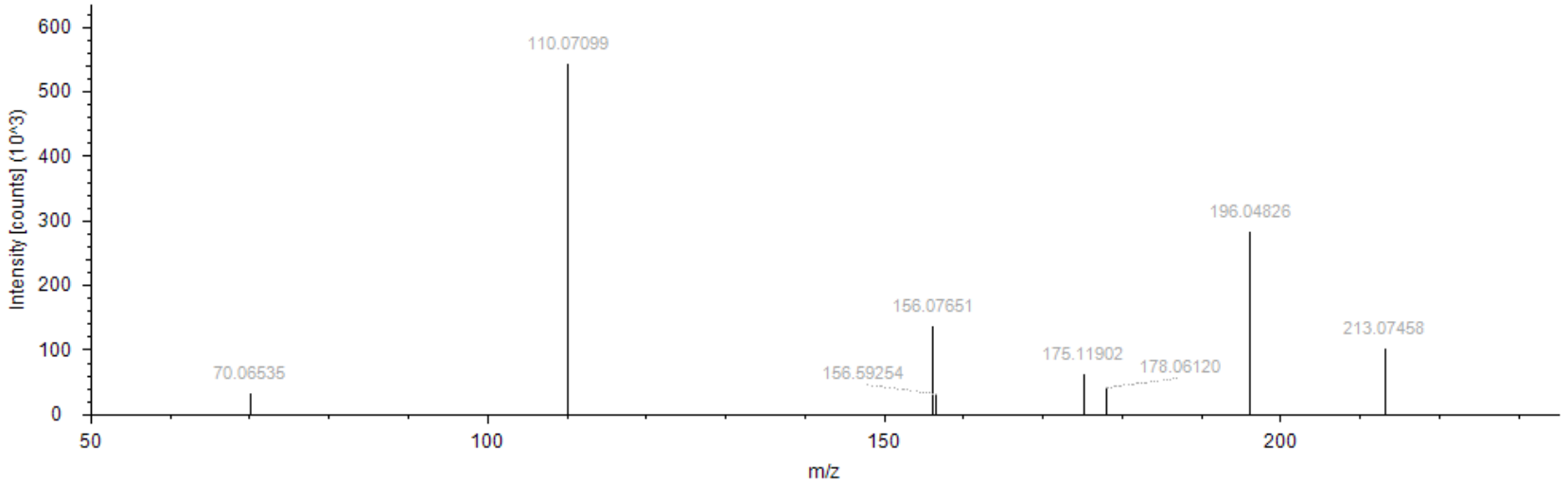
Green = Matched Ion

ThermoFisher SCIENTIFIC

Fragmentation Match – Similarity Search

Molecular Weight → Elemental Composition → ~~ChemSpider Search~~ → ~~mzCloud~~

Unknown Precursor Ion: m/z 213.0745

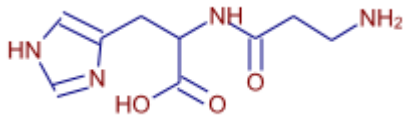


Are there any compounds that share similar fragments? Yes!

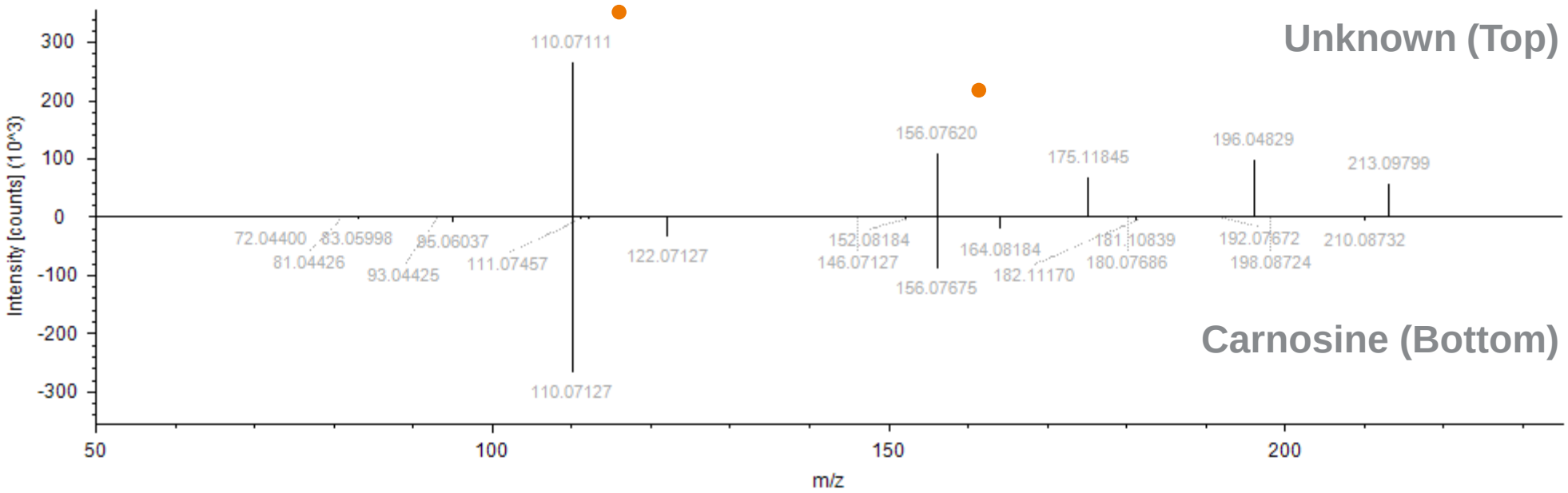
ThermoFisher SCIENTIFIC

Similarity Search for Structure Elucidation

Carnosine
m/z 226.1066



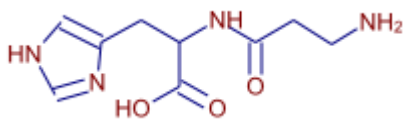
Unknown
m/z 213.0745
Δ 13.0331

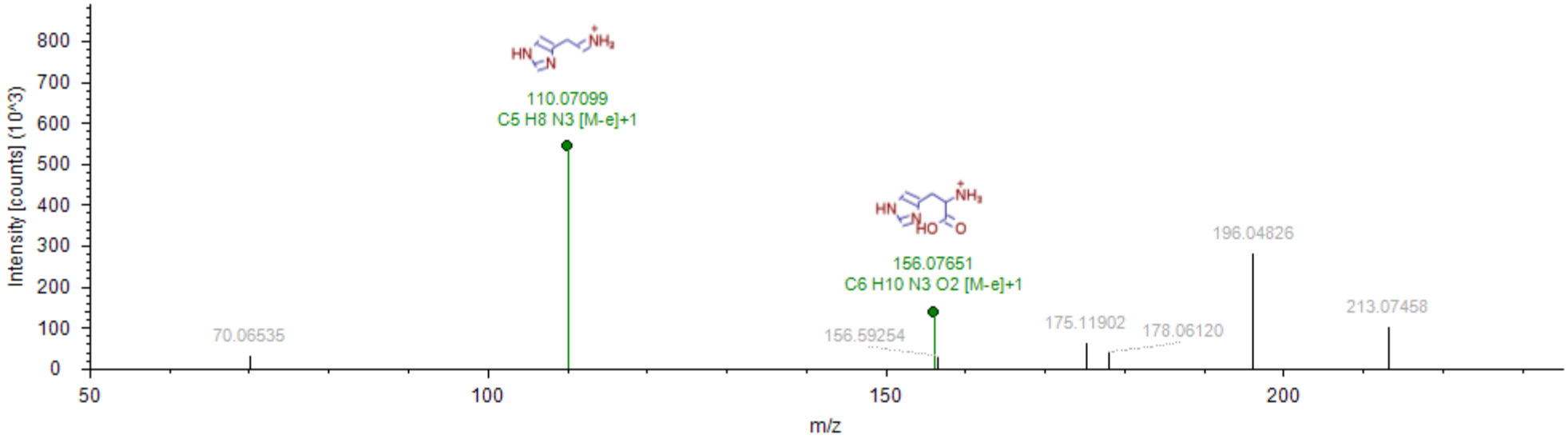


ThermoFisher SCIENTIFIC

Structure Elucidation for Unknown Unknowns

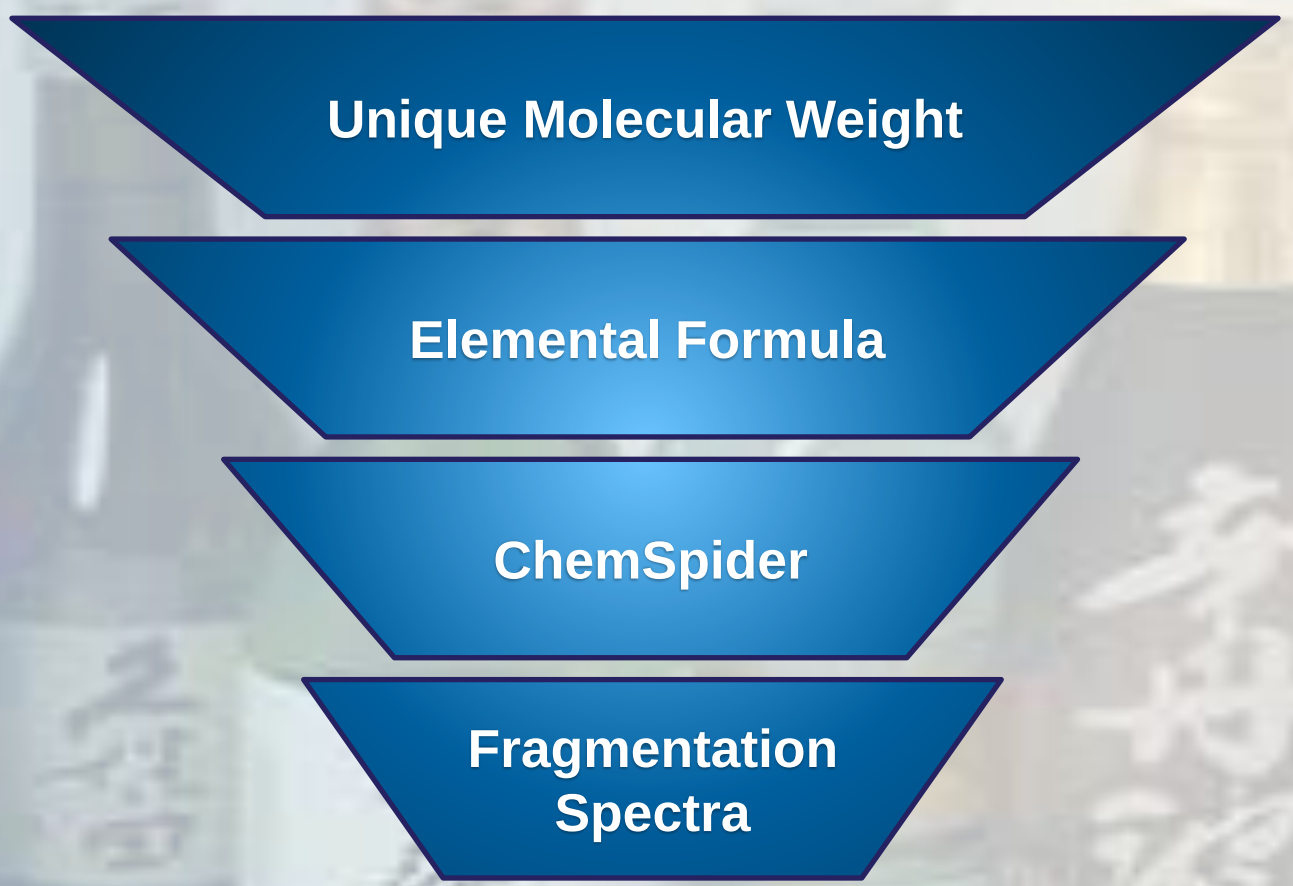
Carnosine
m/z 226.1066





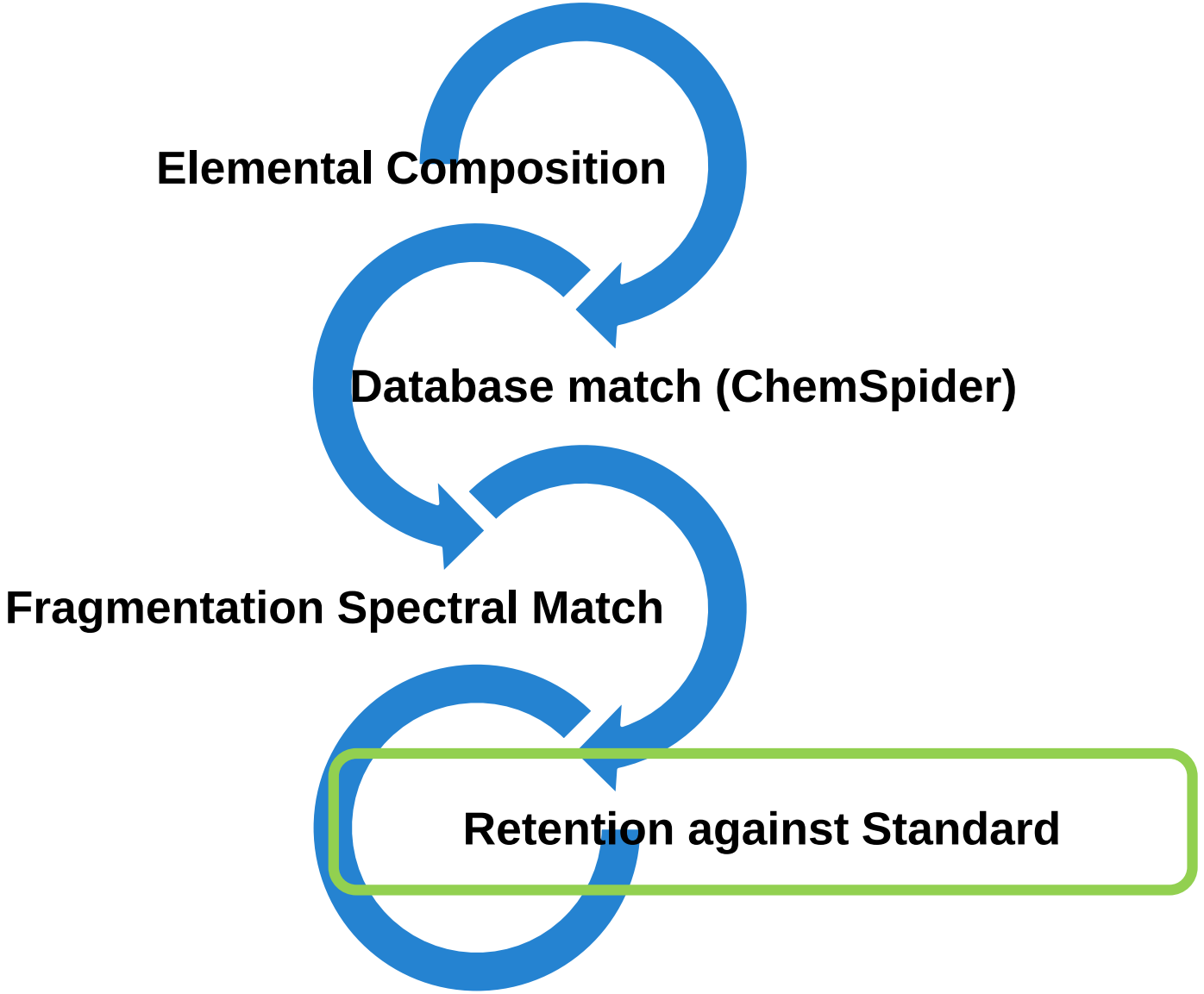
ThermoFisher SCIENTIFIC

Accelerate Unknown Identification



Fewer Candidates. Higher Confidence.

Gold Standard for Confirmation of Identification



Unknown Compound Identification

Compound Discoverer

mzCloud

Summary

- Untargeted metabolomics provides a comprehensive approach for measuring the metabolome
- High resolution and mass accuracy increase confidence of unknown identification
- Compound Discoverer software is powerful data processing software that accelerates data mining and unknown identification
- mzCloud fragmentation library facilitates unknown compound identification and structure elucidation
- Compound Discoverer and mzCloud turn complex untargeted data into meaningful information