Escape the Metabolomics Identity Crisis with Intelligence Driven Mass Spectrometry

Amanda Souza, Ioanna Ntai, Reiko Kiyonami, David Peake, Ralf Tautenhahn, Tim Stratton, Sally Webb, and Andreas Hühmer
Thermo Fisher Scientific, 355 River Oaks Parkway, San Jose, CA, USA, 95134

ORBITRAP-BASED MASS SPECTROMETRY FOR METABOLOMICS

High Quality Data for High Quality Results

Metabolomics is the comprehensive, qualitative, and quantitative study of all endogenous small molecules in a given biological system.[1] The collection of these small molecules reflects the biochemical phenotype, which in turn aids in understanding physiological function and associated pathologies. Yet, the physical-chemical properties of these endogenous compounds are vastly diverse including a range of molecular weight, polarity, structural possibilities and concentration creating analytical challenges in any metabolomics analysis. The detection of metabolites by electrospray ionization mass spectrometry is further challenged with spectra containing molecular features derived from external sources that are experimentally unrelated or multiple ion species reflecting the same molecule such as adduct formation. Metabolomics analysis requires high resolution to distinguish closely related masses in complex matrices, accurate mass measurements for confident spectral peak assignments, and consistent results from scan-to-scan and run-to-run over extended periods.

Leading Orbitrap-based mass spectrometers (Figure 1) provide high-resolution accurate mass (HRAM) measurements and sensitivity required to measure metabolites in complex matrices. High resolution distinguishes spectral features of similar mass, which is required to differentiate isobaric species and determining fine-isotopic pattern (Figure 2). Accurate mass measurements are paramount for confident spectral assignment (Figure 3).

When combined with advanced separations, high throughput and quantitative capabilities expand the scope of what we know about metabolites and their role in several different areas of study (Figure 1).

Figure 1. Thermo Scientific™ Orbitrap™ based Mass Spectrometers provide HRAM measurements suitable for metabolomics analysis of sample types including animal, plant and cellular components. The Thermo Scientific™ Q Exactive™ Series instruments are equipped with a quadrupole mass filter and Orbitrap mass analyzer for HRAM MS and MS/MS while the Thermo Scientific™ Orbitrap Tribrid™ Series instruments include a dual pressure linear ion trap in addition to a quadrupole mass filter and Orbitrap mass analyzer for HRAM MS, MS/MS and MSⁿ spectra.

Thermo Scientific™ Vanquish™ Duo UHPLC System

Thermo Scientific™ Q Exactive™ HF MS

Thermo Scientific™ Orbitrap ID-X ™ Tribrid™ MS

Figure 2. Fine isotopic distribution for the A1 ion cluster of methionine. Data collection was on an organic extraction of human plasma reference material, NIST SRM 1950, using a Q Exactive HF MS with a resolution setting of 120K FWHM @ 200 *m/z*.

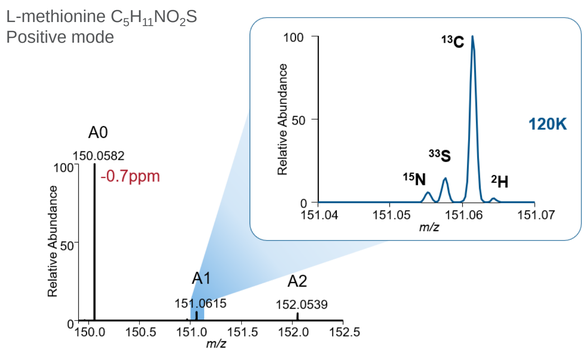
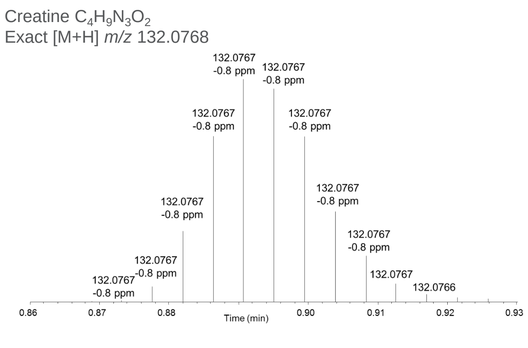


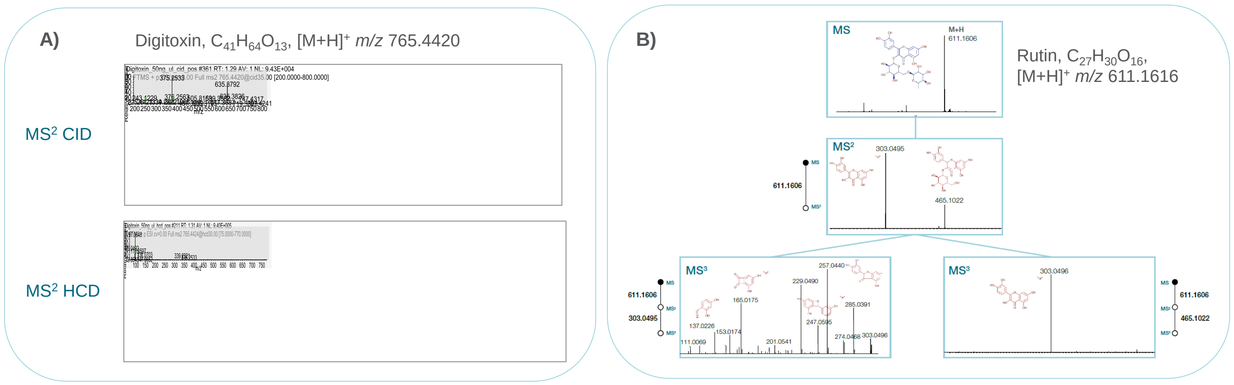
Figure 3. Excellent scan-to-scan mass measurement accuracy obtained with the Orbitrap ID-X Tribrid MS. Sub-ppm mass measurement accuracy for creatine detected over the LC-MS elution profile.



Multiple Dissociation for More Spectral Information

Unique to the Tribrid MS platform are multiple fragmentation techniques (higher-energy collision dissociation (HCD) in a high-pressure collision cell or collisional-induced dissociation (CID) in an ion trap plus multi-stage fragmentation or stepwise MS^n) for annotating unknown metabolite structures in untargeted metabolomics and lipidomics experiments.

Figure 4. Multiple dissociation techniques provides more fragment ion information for structure characterization and elucidation of unknowns using the Orbitrap ID-X Tribrid MS. A) CID and HCD dissociation provides complementary spectral information for the cardio glycoside digitoxin. B) Higher order fragmentation generates MS^n spectral trees enabling the systematic breakdown of the flavonoid rutin.

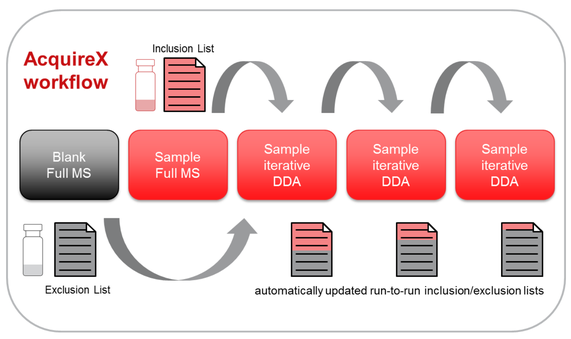


INTELLIGENCE DRIVEN MASS SPECTROMETRY

AcquireX Data Acquisition to Collect More Meaningful Data

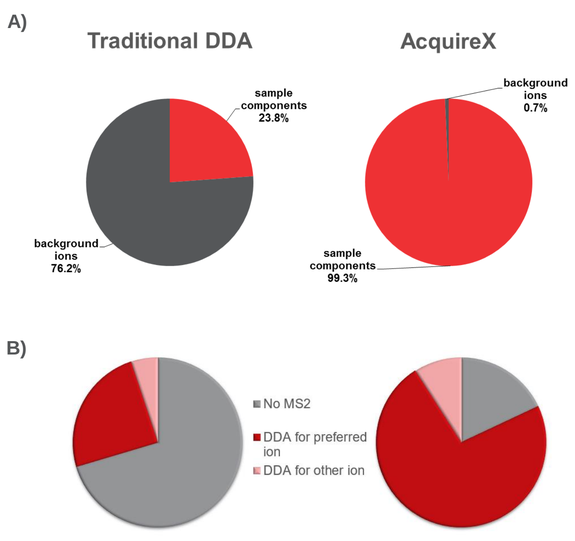
Insufficient metabolome annotation has limited the biological interpretation of untargeted metabolomics studies. Fragmentation spectra provide more spectral information to increase the confidence of unknown annotations. The AcquireX intelligent acquisition generates more fragmentation spectra for sample relevant compounds, avoids unrelated background ions, and removes redundancies by using fully automated iterative inclusion and exclusion lists. Available on the Orbitrap Tribrid MS systems, AcquireX takes advantage of knowledge to drive acquisition where the blank sample generates a list of ions to exclude for subsequent fragmentation and a matrix sample generates a list of true sample components to prioritize for data dependent MS^2 and MS^n acquisition (Figure 5). AcquireX leads to information-rich fragmentation of more experimentally relevant compounds (Figure 6).

Figure 5. The AcquireX Deep Scan acquisition workflow for improved compound annotation.



- First, the AcquireX process obtains the LC-MS data for the blank and a pooled sample
- The AcquireX process creates an exclusion list from the blank and an inclusion list from the sample data
- The first data dependent MS^2 run is acquired and the inclusion/exclusion lists are updated after the run
- On the second injection, MS^2 spectra are acquired for compounds remaining on the inclusion list
- This process is repeated for a user-specified number of injections

Figure 6. A) Obtaining MS^2 information on compounds vs. background with traditional data dependent acquisition (DDA) compared to AcquireX. B) Sample components with no MS^2 or MS^2 selected for the preferred ion or associated ion for traditional DDA compared to AcquireX.



Human plasma (NIST SRM1950), C18, 15 min gradient, data dependent LC-MS^2

SPECTRAL LIBRARIES TO ANNOTATE MORE

mzCloud™ Mass Spectral Library

The mzCloud mass spectral library is a highly curated, public library of endogenous and exogenous small molecules containing over 5 million fragmentation spectra. Each compound entry in the library generally includes two fragmentation techniques: HCD and CID. For each dissociation technique, fragmentation spectra span a wide range of collision energies in iterations of 5 or 10%. This systematic collection of spectra eliminates constraints in how data are acquired in terms of collision energy for subsequent spectral library matching. Ion trap technology provides MS^n capabilities with repeated isolation and fragmentation of product ions beyond MS^2 producing MS^n spectral trees for each library entry (Figure 7). This highly diverse spectral library (Figure 8) contains true mass spectra generated from purified reference standards. Each spectrum is recalibrated for exact mass and noise removed and is further structurally annotated making this an ultra high-quality spectral library. The mzCloud library is fully integrated into the Thermo Scientific™ Compound Discoverer™ small molecule processing software and the Thermo Scientific™ Mass Frontier™ spectrum interpretation software.

Figure 7. Representation of an MS^n spectral tree from the mzCloud spectral library. Stacked spectra reflect separate collision energies covering a wide range.

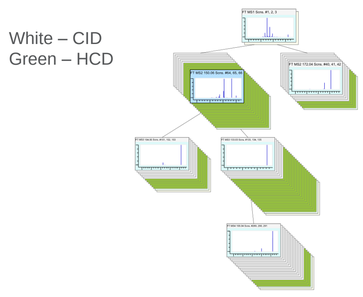
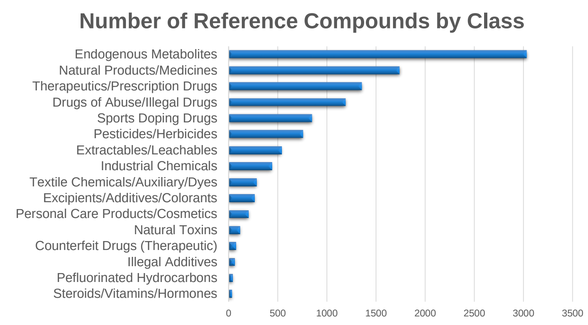


Figure 8. The reference library within the mzCloud spectral library consist of 16 categories of small molecules including endogenous metabolites. Sourced from mzCloud.org on 5/11/2019.



Spectral Libraries for More Than an Identity Match

In addition to HRAM measurements of the precursor ion, fragmentation spectra provide an additional layer of knowledge about the molecular makeup of a compound and subsequently increases unknown annotation confidence. Fragmentation spectra generated from the experimental results are compared against reference spectra within the library. A direct match, where both precursor and product ions match, results in an "identity match" (Figure 9). There are instances when the library search results in not matched due to the absence of the reference standard in the library. At the same time, a similarity search can be applied. With this search type, match criteria are less stringent where only product ions need to match (Figure 10). This allows for associations based on structural relationships to the parent structure in the reference library further enabling unknown annotations.

Figure 9. Identity match for the endogenous compound nicotinamide adenine dinucleotide (NAD) with MS^2 experimental data collected from the SRM 1950 plasma extract against the mzCloud spectral library.

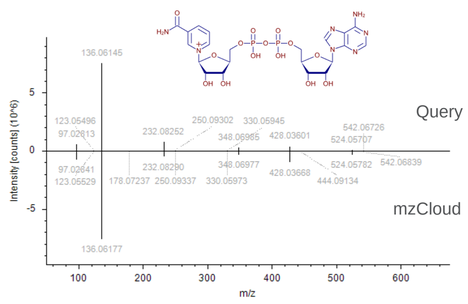
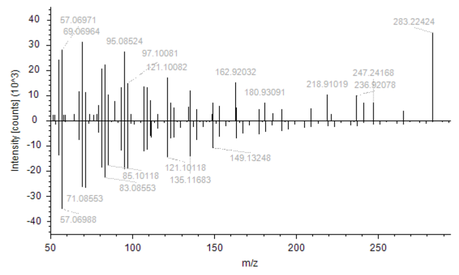


Figure 10. Similarity match to nervonic acid with MS^2 experimental data collected from the SRM 1950 plasma extract against the mzCloud spectral library. The precursor mass differs by 98.2 Daltons. While not exact, these compounds are highly similar.

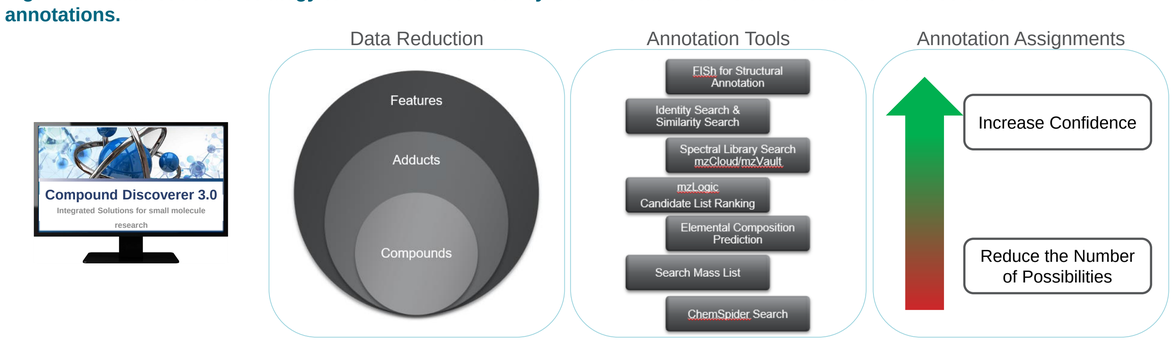


SOFTWARE BUILT TO ANNOTATE UKNOWNS

Compound Discoverer Software

In conjunction with HRAM and intelligence driven data dependent acquisition is the data processing. Compound Discoverer software is a single software platform that streamlines high resolution LC-MS and MS^2 data analysis, identification of primary and secondary metabolites, statistical analysis, and metabolic pathway mapping. The workflow for untargeted metabolomics minimizes false positive reporting by incorporating a data reduction strategy to mark unrelated background ions and reducing resulting spectral features down to sample relevant compounds through deisotoping and adduct association (Figure 11). Multiple annotation tools are applied to increase annotation confidence including prediction of elemental composition from the measured accurate mass, a ChemSpider database search with elemental formula, spectral library searches via mzCloud and/or an in-house spectral library, the mzLogic analysis algorithm to rank order matched candidates from database searching and *in silico* fragmentation of proposed chemical structures (Figure 11). Results from all data sources are used to assess the level of consensus. When spectral library and database annotations are not found, similarity searches are performed.

Figure 11. Data reduction strategy to minimize redundancy and annotation tools to increase confidence of unknown annotations.



Specialized Software for Complementary Unknown Annotation

For specific compound classes or analytical approaches (MS^n), specialized software is used. Thermo Scientific™ LipidSearch™ software provides lipid annotation using LC-MS^2 and MS^3 data by relying on product ion matching to predicted lipid fragments. The lipid database contains >2 million lipid ions and their predicted fragment ions covering 92 lipid sub-classes for a broad range of species. Combined with sophisticated structure-specific data acquisition using HCD/CID, target product-ion and neutral loss modes to distinguish isomeric species, LipidSearch software enables deeper structural characterization for improved lipid annotations.

Mass Frontier interpretation software is especially suited for analysis involving MS^n spectral trees. Mass Frontier software is fully integrated with the mzCloud spectral library and provides search capabilities for identity matching and sub-structure matching where matched product ions can occur at different fragmentation stages for the queried spectra compared to the reference spectra. Partial matches allow for users to annotate more compounds pertaining to structure or class associations. Mass Frontier software also hosts in-house MS^n spectral libraries with the ability to create, curate and store reference spectra.

Thermo Scientific Small Molecule Software Suite for Unknown Annotation

CONCLUSIONS

- Confidence in compound annotation increases when combining HRAM full scan measurements with multiple dissociations (CID/HCD, MS^2 and MS^n)
- Intelligence-driven data acquisition generates more experimentally relevant spectral information and ignores background and redundancies to increase annotation of unknown compounds
- Software with advanced and specialized annotation tools combined with comprehensive spectral libraries improve confidence in unknown annotations

REFERENCES

1. Oliver, S.G., Winson, M.K., Kell, D.B., and Baganz, F. (1998). Systematic functional analysis of the yeast genome. Trends in Biotechnology, 16(9):373-8.

TRADEMARKS/LICENSING