High quality curated HRAM MS^n spectral libraries and real time library search for the confident annotation of metabolites

Rahul Deshpande[1], Bashar Amer[1], Daniel Hermanson[1], Brandon Bills[1], Eric Tague[1], Susan Bird[1], Reza Jafari[2], Pedram Rafeie[2] and Andreas Hühmer[1]

[1]Thermo Fisher Scientific, San Jose, California, United States: [2]MetaSci Inc, Toronto, ON, Canada

ABSTRACT

Purpose: Compound class specific deep scan for confident annotation of relevant analytes using a high quality curated High Resolution Accurate Mass (HRAM) MS^n spectral library in conjunction with Real Time Library Search (RTLS).

Methods: A Thermo Scientific™ Vanquish™ Horizon UHPLC system coupled to a Thermo Scientific™ Orbitrap IQ-X™ Tribrid™ mass spectrometer is used for collecting all MS and MS^n data. For data acquisition, MS/MS is always collected with precursor ions detected in the survey MS scan within 1.2 second cycle time. High order MS^n $^{(3\text{-}5)}$ is collected using the built-in spectral library creation templates using LC-MS . The generated MS^n tree data are processed using Thermo Scientific™ Mass Frontier™ 8.0 and Thermo Scientific™ Compound Discoverer™ 3.3 software. Flavonoid standards were used for the library creation and putative flavonoids were annotated in tea extract. Thermo Scientific™ mzVault™ 2.3 software was used for creating the spectral library to be used for RTLS.

Results: Use of standard spectral library and RTLS allows more unknown flavonoid compounds being identified from the natural products. Nearly 100 more putative flavonoids were annotated using this workflow from tea extracts.

INTRODUCTION

Confident compound annotation is essential for translating untargeted metabolomics data into meaningful biological information. Mining of relevant information from untargeted metabolomic data remains a challenge. Local mass spectral libraries consisting of MS^n data from authentic standards provide the opportunity to confidently annotate known and related unknown compounds in biological samples. Here we describe the process of creating and curating a high-quality spectral library of high-resolution accurate mass data using authentic flavonoid standards and the built-in library builder method template on the a Thermo Scientific™ Orbitrap IQ-X™ Tribrid™ Mass Spectrometer (MS) and its utility for the confident annotation of flavonoids in biological samples using Real Time Library Search (RTLS) to automatically guide MS^n data acquisition on the fly.

MATERIALS AND METHODS

Sample Preparation

The Flavonoid Library consisting of 40 standards was obtained from MetaSci (Toron to,On, CA). The library consists of 4 vials each containing 10 authentic standards which could be directly injected for LC-MS analysis. Three varieties of green tea, black tea and herbal tea were obtained from the local market. Three replicates of each were extracted using 240 mL of hot water in commercially available disposable coffee cup. For blank, hot water without tea in similar coffee cup was used. Pooled samples were prepared by adding 100 µL of each extract. This pooled sample was used for QC.

HPLC Conditions

A Thermo Scientific™ Vanquish™ UHPLC system performed separations. The column was a Thermo Scientific™ Accucore™ Vanquish™ C18+ UHPLC column (2.1 x 150mm, 1.5µm) that operated at 55 ℃ and a flow rate of 200 µL/min.

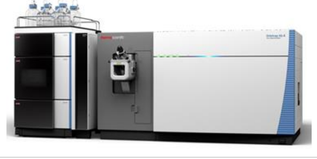
Mobile Phase: (A) 0.1% (v) formic acid (FA) in LC-MS grade water
(B) 0.1% (v) FA in LC-MS grade methanol

HPLC Gradient:

Time	A%	B%	
0.00	100	0	Divert valve: to waste = 0 – 0.2 min
3.00	50	50	to MS = 0.2 – 15.0 min
9.00	2	98	
13.00	2	98	
13.10	100	0	
15.00	100	0	

The injection volume was 2 µL.

Figure 1. A Thermo Scientific™ Vanquish™ Horizon UHPLC system coupled to a Thermo Scientific™ Orbitrap IQ-X™ Tribrid™ mass spectrometer



Mass Spectrometer

Thermo Scientific™ Orbitrap IQ-X™ Tribrid™ Mass Spectrometer was used.

MS Conditions-Spectral Library Building

The inbuilt template for library building on the IQ-X™ using LC-MS was used for the flavonoid standards. The template allows for the collection of MS^n spectral data. Data was also collected with a modified template allowing for collection of MS^2 data utilizing different fragmentation strategies (HCD, CID and UVPD).

MS Conditions-RTLS

The inbuilt template for using RTLS on IQ-X™ (named Met-IQ) was used for data acquisition on the pooled sample. The template allows for similarity search of compounds present in the library and a trigger for MS^3 for compounds similar to those in the library.

Figure 2. Modified spectral library data acquisition process utilizing different fragmentation methods (HCD, CID and UVPD).

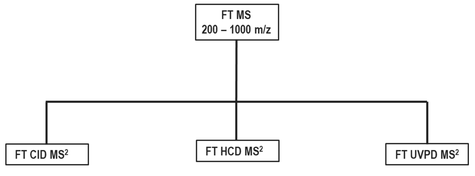
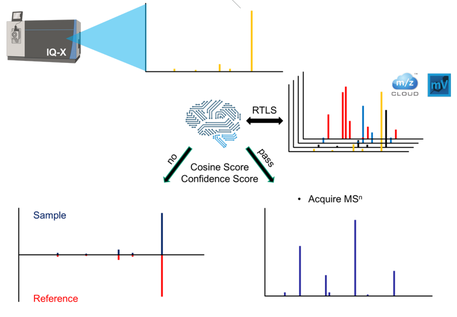


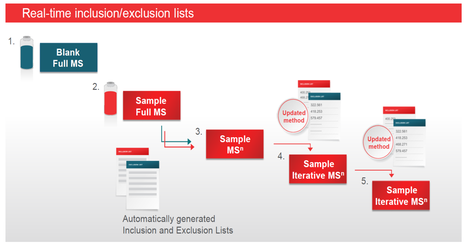
Figure 3. Real Time Library Search on IQ-X™. MS/MS spectra are searched against offline mzVault library in real time. If the spectra passes the criteria subsequent MSn is triggered for improved annotation.



MS Conditions-Data Aquisition

Deep Scan Acquire-X acquisition workflow was used for analysis. This workflow automatically creates an exclusion list from background ions (Blank) and an inclusion list of metabolites of interest from the reference sample (Pooled) with iterative updating in between each injection after performing MS^2. A single injection was was Full scan MS1 (200-1000 m/z) was performed on all the tea samples.

Figure 4. Thermo Scientific™ AcquireX Deep Scan mode for intelligent data acquisition to maximize the number of relevant compounds interrogated by MS/MS.

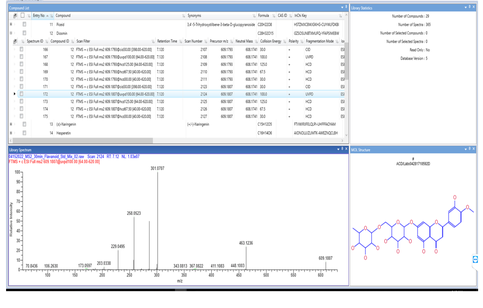


RESULTS

Spectral Library

Thermo Scientific™ Compound Discoverer™ 3.3 software was used to process the LC-MS data from the standards. The spectral data with the various fragmentation strategies from the standards was transferred to a mzVault™ library. The mzVault™ library containing the spectral fragmentation information of the flavonoid standards was populated with the structural and chemical information of individual compounds. This library was used for the RTLS on the instrument.

Figure 5. Thermo Scientific™ mzVault™ library showing the spectra for an example standard flavonoid.



Data Processing

Thermo Scientific™ Compound Discoverer™ 3.3 software was used to process the LC-MS data from the samples. Differential analysis and compound annotation revealed relative differences between the various tea samples. Several annotation tools are available in the Compound Discoverer 3.3 software including database searching and spectral library matching against the mzCloud™ spectral library at the MS^n level. In parallel it has the ability to search custom mass lists with corresponding chemical structures. Here the data was processed using the Arita Flavonoid Database which has the mass and structural information of around 6500 Flavonoids.

Figure 6. Compound Discoverer™ workflow to annotate structures of putative flavonoids which are identified based on the sub-structure match with flavonoid references in the mzCloud library and carrying out the statistical analysis

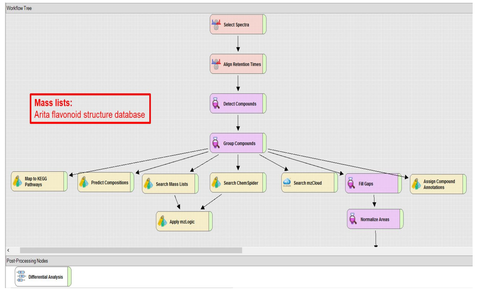
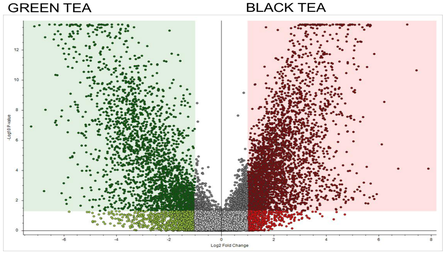


Figure 7. Volcano plot of the metabolomics data showing all the compounds detected by Compound Discoverer™ 3.3 software. The x-axis is the mean ratio fold-change (plotted on a log 2 scale) of the relative abundance of each metabolite between green tea and black tea



Data Filtering

The data obtained from Compound Discoverer™ 3.3 was filtered for showing the utility of RTLS to trigger more relevant compounds (structurally similar to flavonoids) for MS^n fragmentation. The filter applied extracted only compounds in the Arita Flavonoid Database which had a subsequent MS^3 fragmentation.

Figure 8. Filter applied to Compound Discoverer™ 3.3 data to extract only relevant compounds for utility of RTLS

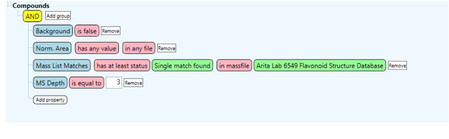
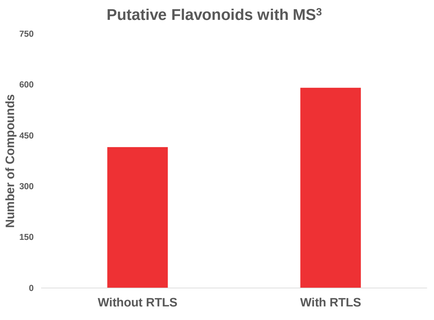


Figure 6. Filtered data from Compound Discoverer™ showing compounds in the Arita Flavonoid Database which had a subsequent MS^3 fragmentation (putative flavonoids).



RTLS FOR COMPOUND CLASS SPECIFIC DEEP SCAN

Two major challenges in untargeted metabolomics are

1. Extracting biologically relevant information
2. Confident compound annotation with structural information

The workflow described here addresses both challenges by utilizing RTLS. RTLS can be used to identify and confidently annotate analytes belonging to certain biological classes. The annotation is done by getting structural information utilizing MS^n and multiple fragmentation methods in combination with using Compound Discoverer™ 3.3.

CONCLUSION

A workflow for utilization of RTLS is for compound class (flavonoids) specific identification of analytes in biological matrices (Tea) is described. It involves

1. Creation of spectral library using authentic flavonoid standards.
2. Utilizing the created spectral library for RTLS of related compounds in biological matrices (tea).

More putative flavonoids were identified in tea using RTLS in comparison with samples without acquired without RTLS.

The challenge of obtaining deep and relevant metabolome coverage can be addressed using intelligent data acquisition strategies such as AcquireX in combination with RTLS.

TRADEMARKS/LICENSING