



# Novel structure-based profiling and annotation workflow—high-throughput analysis of flavonoids using the Thermo Scientific Orbitrap ID-X Tribrid MS

## Authors

Reiko Kiyonami<sup>1</sup>, Iwao Sakane<sup>2</sup>,  
Seema Sharma<sup>1</sup>, Graeme  
McAlister<sup>1</sup>, Caroline Ding<sup>1</sup> and  
Andreas Huhmer<sup>1</sup>

<sup>1</sup>Thermo Fisher Scientific,  
<sup>2</sup>ITO EN, LTD, Tokyo, Japan

## Keywords

Flavonoid, flavonoid class, MS<sup>n</sup>  
spectral tree, sugar neutral loss,  
Mass Frontier 8.0 software, sub-  
tree search, mzCloud spectra  
library, structure annotation,  
Compound Discoverer 3.0  
software, FISH score

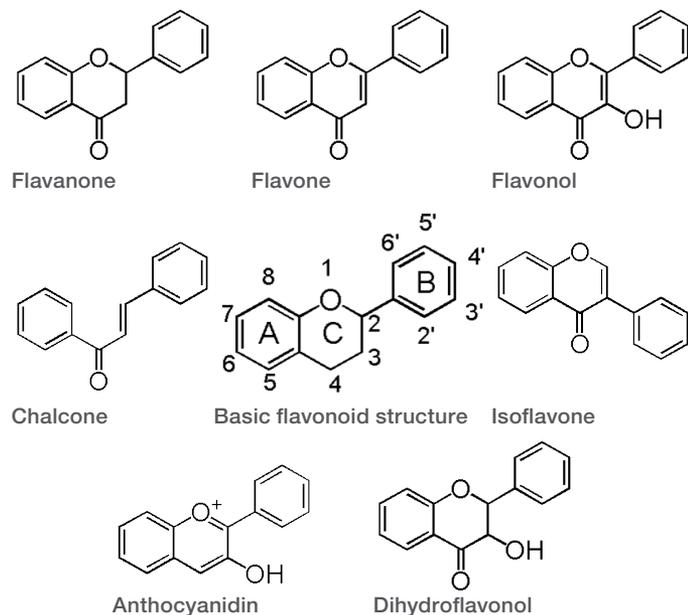
## Introduction

Untargeted metabolomics aims to detect and compare as many metabolites as possible from a sample set. One bottleneck for the untargeted metabolomics approach is how to annotate the identification of unknown compounds from the sample set. Often there are classes of compounds which are difficult to identify and annotate due to the limited availability of authentic standards and structural diversity of these compounds, such as steroids, phospholipids, endocannabinoids and flavonoids. These compounds share basic core structure, but have various modifications in different structure positions of the basic structure, yielding a class of compounds with different molecular weights and diversified structures.

Flavonoids are a good example of this. Flavonoids are widely found in fruits, vegetables, grains, bark, roots, stems, flowers, tea, cocoa and wine, and are powerful antioxidants with anti-inflammatory and immune system benefits. Flavonoids have been widely studied for over a decade because of their diverse and important biological roles. One of the main approaches for flavonoid studies is untargeted flavonoid profiling because it provides insights into their biological functions and potential health benefits for humans. However, comprehensive identification of flavonoids remains challenging because of the structural diversity of this class. Chemically, flavonoids have the general structure of a 15-carbon skeleton, which consists of two phenyl

rings (A and B) and heterocyclic ring (C). They can be subdivided into several classes based on the structural features of the C ring (Figure 1)<sup>1</sup>.

Flavonoids are one class of secondary plant metabolites and are often modified in positions 3, 5, 6, 7, 8, 3', 4', and 5' of the basic flavonoid structure with hydroxylation, methylation, acylation, prenylation, and O- and C-glycosylation via secondary metabolic pathways. This diversity of modifications yields a wide range of isomeric and isobaric species (Figure 2). Currently, more than 10,000 flavonoids have been reported. However, the number of authentic



flavonoid standards that can be used to generate reference spectra for unknown flavonoid identification is very limited. Authentic references for flavonoids with multiple glycoside modifications are especially hard to produce. As such, many unknown flavonoid compounds do not have an exact spectral library match. As a result, the majority of published flavonoid structure characterization studies have been carried out through manual assignment of fragment ions generated from MS<sup>2</sup> and higher order of MS<sup>n</sup> data<sup>2,3</sup>. This type of manual analysis requires extensive knowledge of the flavonoid fragmentation rules and a tremendous amount of time to interpret the data. Plus, for most flavonoid glycoconjugates, MS<sup>2</sup> does not provide sufficient structurally relevant fragment ion information for characterizing aglycone structures<sup>4</sup>.

MS<sup>n</sup> (multiple stage mass spectrometry) can provide more structurally relevant fragment ion information by the systematic breakdown of a compound and may be employed to generate a spectral tree to facilitate the unknown compound annotation. Here we present a new class or structure-based flavonoid profiling workflow analyzing fruit and vegetable juices that uses comprehensive fragment ion information from both HCD (higher-energy collisional dissociation) and CID (collisional induced dissociation) Fourier Transform (FT) MS<sup>2</sup>, and higher order CID FT MS<sup>n</sup>, for rapid flavonoid annotation on a Thermo Scientific™ Orbitrap ID-X™ Tribrid™ mass spectrometer. This workflow is demonstrated on glycoconjugates, but works on other transformation products of secondary metabolism.

Figure 1. Basic structure of the flavonoid and major structure classes.

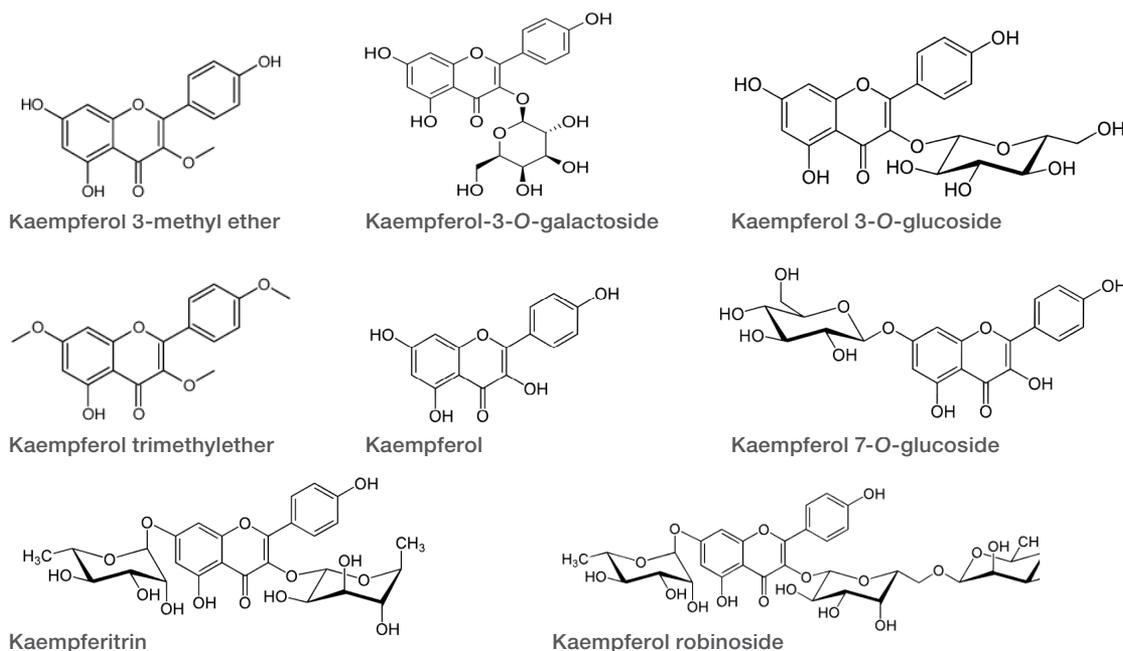


Figure 2. Various modifications on kaempferol aglycone generate complex secondary metabolites.

## Materials And Methods

### Sample preparation

Three commercially available fruit and vegetable juice samples (Naked® Kale Blazer, Odwalla® Berries Gomega®, and Odwalla® Red Rhapsody®) were analyzed in this study. Each juice sample was filtered and diluted two times with methanol.

### HPLC conditions

A Thermo Scientific™ Vanquish™ UHPLC system performed separations using the gradient conditions shown in Table 1. Mobile phase A was water with 0.1% formic acid and mobile phase B was methanol with 0.1% formic acid. The column was a Thermo Scientific™ Hypersil Gold™ (2.1 × 150mm, 1.9µm), which we operated at 45 °C and a flow rate of 200 µL/min. The injection volume was 2 µL. Each sample was analyzed in triplicate.

Table 1. HPLC gradient.

Vanquish UHPLC		
Time (min)	%A	%B
0	99.5	0.5
1	90	10
10	70	30
18	50	50
22	1	99
25	1	99
25.1	99.5	0.5
30	99.5	0.5

### MS conditions

All the data was collected on an Orbitrap ID-X Tribrid mass spectrometer. The mass spectrometer source conditions and method are shown in Table 2. The instrument template—structure specific MS<sup>4</sup> (monosaccharide loss) was used. This instrument template aims to collect an MS<sup>n</sup> spectral tree (intuitively organized multi-stage tandem mass spectra) by systematically breaking down fragment ions from multiple stages in order to get structure relevant fragmentation pathways for unknown flavonoid structure annotation. For the purpose of establishing a fragmentation pathway, it is preferred to have fragment ions generating from a precursor ion remain stable without further fragmentation. One of the

unique benefits that the Orbitrap ID-X Tribrid MS offers is multiple fragmentation techniques, including HCD providing relatively higher collision energy and CID providing relatively lower collision energy. In the case of flavonoid glycoconjugates, the softer CID is able to preserve the fragment ions with sugar neutral loss, providing a systematic fragmentation pathway to facilitate the structure annotation.

Table 2. Orbitrap ID-X Tribrid MS instrument set up.

ESI source	Orbitrap ID-X
Sheath gas 35	Pos ion (150-1200 amu)
Aux gas 5	MS: R=60K (FWHM at <i>m/z</i> 200)
Spray volt. 3.4 kV	MS <sup>n</sup> : R=15K (FWHM at <i>m/z</i> 200)
RF-Lens 40	Cycle time: 1.2 second
Cap. temp. 300 °C	MS <sup>2</sup> Isolation width: 1.6 Da
Heater temp. 300 °C	MS <sup>n</sup> Isolation width: 1.6 Da (MS <sup>2</sup> ) → 2.0 Da (MS <sup>n</sup> )

As our goal is to carry out flavonoid identification using FT MS<sup>n</sup> spectral tree data and simultaneous flavonoid quantitation using FT MS data, the instrument template is designed to collect the maximum amount of meaningful MS<sup>n</sup> spectral tree data for the unknown flavonoid structure annotation within a short cycle time (1.2 seconds) to maintain enough MS scan points across a peak for precise quantitation. Because HCD MS<sup>2</sup> already provides sufficient fragment ions for structure annotation when the flavonoid compounds do not have glycol modifications, for the precursor ion mass range between 150–420 *m/z*, only HCD MS<sup>2</sup> data are collected. For the precursor ion mass range between 420–1200 *m/z*, glycol modifications are expected and an intelligent product ion-dependent MS<sup>n</sup> approach was used, in which a high-resolution accurate mass (HRAM) full MS scan was followed by CID MS<sup>2</sup> scans. The product ions generated from each MS<sup>2</sup> scan were monitored by the mass spectrometer and an MS<sup>3</sup> scan was triggered if one or multiple pre-defined neutral sugar molecules were detected. An additional MS<sup>4</sup> scan was triggered if pre-defined neutral sugar molecules were detected from the MS<sup>3</sup> scan. Figure 3 shows the flowchart of the developed product ion-dependent MS<sup>n</sup> data acquisition instrument method.

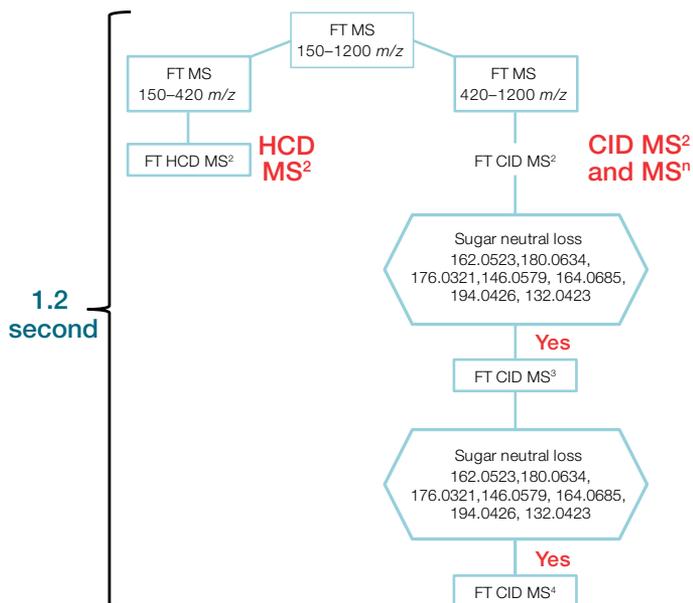


Figure 4 shows an example MS<sup>3</sup> spectral tree generated from the reference standard for the flavonoid rutin. Two major fragment ions in the CID MS<sup>2</sup> spectrum were detected. The *m/z* 465.1022 fragment ion represents the loss of one rhamnose sugar moiety from the intact structure while the *m/z* 303.0495 fragment ion represents the loss of two sugar moieties, rhamnose and hexose. At the MS<sup>3</sup> stage, the major fragment ion from the *m/z* 465.1022 precursor ion was the *m/z* 303.0495, confirming that the 303.0495 did come from the loss of two sugar moieties (rhamnose and hexose) from the starting structure, while the fragment ions of the precursor *m/z* 303.0495 provided sufficient structure relevant fragment ions which can be used for the aglycon structure annotation.

Saccharide	Neutral Loss	Composition
Pentose (xylose, arabinose)	132.04226	C <sub>5</sub> H <sub>8</sub> O <sub>4</sub>
Deoxyhexose (rhamnose)	146.05791	C <sub>6</sub> H <sub>10</sub> O <sub>4</sub>
Hexose (glucose, galactose)	162.05282	C <sub>6</sub> H <sub>10</sub> O <sub>5</sub>
Glucuronide	176.03209	C <sub>6</sub> H <sub>8</sub> O <sub>6</sub>
Glucuronic acid	194.04265	C <sub>6</sub> H <sub>10</sub> O <sub>7</sub>

Figure 3. Flowchart of sugar neutral loss triggered high order MS<sup>n</sup> data acquisition on the Orbitrap ID-X Tribrid MS instrument.

The associated name and composition for each targeted sugar neutral loss are described in the table.

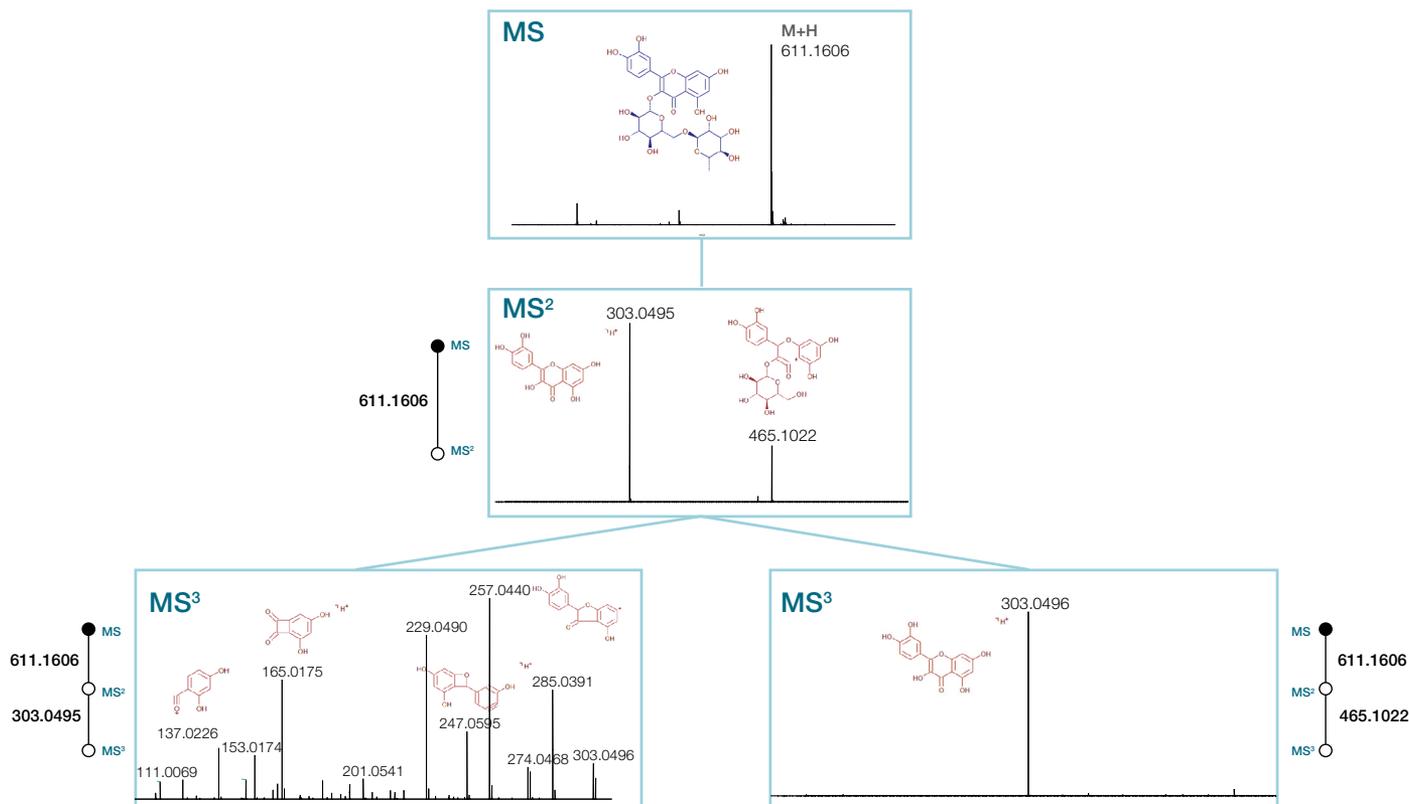


Figure 4. CID MS<sup>3</sup> spectral tree generated from the rutin reference standard.

## Results and discussion

### Addressing the flavonoid annotation challenges using a structure-based MS<sup>n</sup> approach

The MS<sup>n</sup> fragmentation approach provides a systematic breakdown of a compound and generates a spectral tree, yielding more structurally relevant fragment ion information compared to MS/MS fragmentation (Figure 3). Figure 5 shows an MS<sup>3</sup> spectral tree collected from an unknown compound detected from the Kale Blazer juice sample. The MS<sup>2</sup> spectrum alone for the *m/z* 641.1720 did not provide a library match to the mzCloud spectral

library. More structurally relevant fragment ions were detected by further fragmenting the MS<sup>2</sup> product ion at *m/z* 317.0657, which matched the reference compound Isohammetin in the mzCloud spectral library, providing confident sub-structure annotation (aglycon) of this unknown compound. As a result, this unknown compound can be annotated as a flavonoid class-related compound whereby the aglycon sub-structure was detected by using MS<sup>3</sup> spectral data.

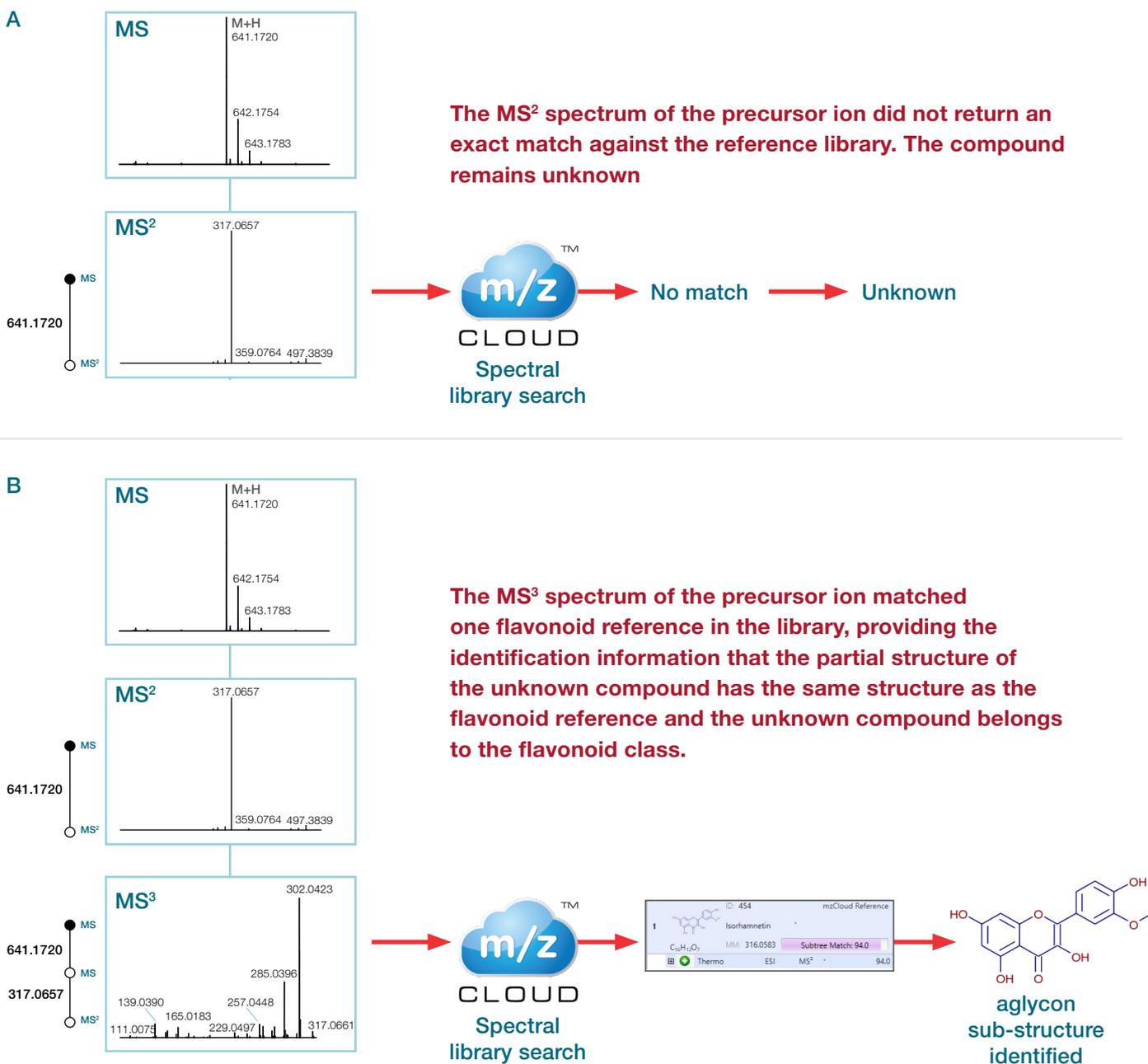


Figure 5. MS<sup>2</sup> (A) and MS<sup>3</sup> (B) spectral trees generated on an unknown compound detected (M+H: 641.1720) from kale blazer juice sample.

## MS<sup>n</sup> spectral tree data acquisition



## MS<sup>n</sup> spectral tree data processing



**Figure 6. Workflow of structure-based MS<sup>n</sup> approach to address the flavonoid annotation challenges.**

Taking advantage of the fact that MS<sup>n</sup> fragmentation enables sub-structure identification, as shown in above example, we developed a structure-based MS workflow to address the flavonoid annotation challenges. Figure 6 shows the workflow of this approach, which acquires the MS<sup>n</sup> spectral tree data using the instrument template discussed in the MS conditions section. The collected MS<sup>n</sup> spectral tree data are first processed by Thermo Scientific™ Mass Frontier™ 8.0 software to determine which compounds include the basic flavonoid structure thus belonging to the flavonoid class. These detected flavonoid-related compounds are then further annotated using a flavonoid structure database and structure ranking tools within Thermo Scientific™ Compound Discoverer™ 3.0 software.

### Flavonoid class compound identification using Mass Frontier 8.0 spectral interpretation software

Mass Frontier 8.0 software was used to process the MS<sup>n</sup> tree data collected from juice samples. The software detects unknown compounds from each juice raw file using the Joint Components Detection (JCD) algorithm. All detected compounds and associated spectral trees were then queried against the mzCloud MS<sup>n</sup> spectral library containing mass spectra generated from authentic reference material (Box 1) using the “Subtree Search” approach. A Subtree search compares the experimental MS<sup>n</sup> tree against MS<sup>n</sup> trees in the mzCloud library (Figure 7). By selecting “Search All”, all detected compounds are automatically searched against the references in the mzCloud library in batch mode (Figure 8).

#### Box 1.

##### mzCloud MS<sup>n</sup> spectral library

The mzCloud library is an extensively curated database of high-resolution tandem mass spectra that are arranged into spectral trees. MS<sup>2</sup> and multi-stage MS<sup>n</sup> spectra were acquired at various collision energies, precursor *m/z*, and isolation widths using CID and HCD. Each raw mass spectrum was filtered and recalibrated. It is a fully searchable library. mzCloud’s website <https://www.mzcloud.org/> is freely accessible to query MS<sup>n</sup> spectra data and retrieve the library reference match manually. The users of Compound Discover (2.1, 3.0) software and Mass Frontier 8.0 software can search all MS<sup>n</sup> spectra data automatically to get library reference match in a batch mode.

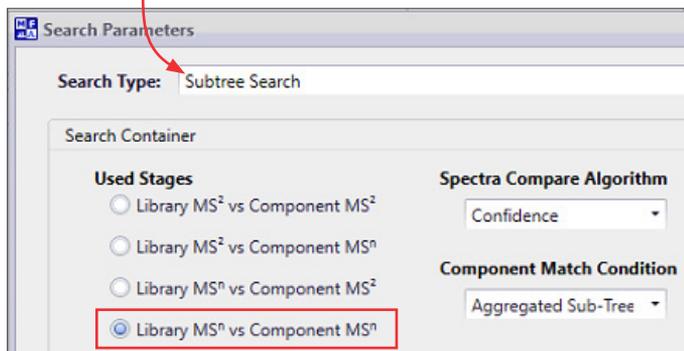
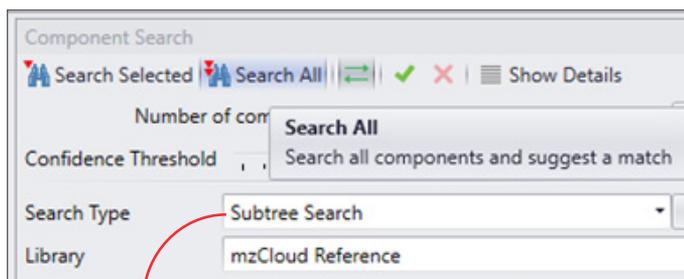
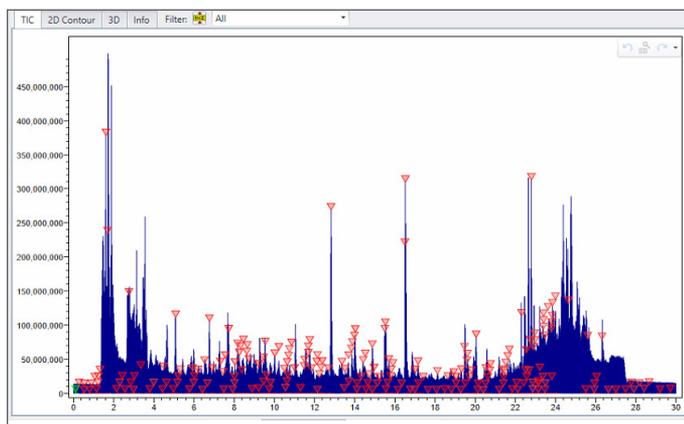


Figure 7. Partial MS<sup>n</sup> spectral tree search results from the Kale juice sample.

## Unknown compound detection



## Match result

Name	Scan No.	Precursor m/z	Match	MS <sup>n</sup>	t <sub>R</sub> (min)	Abundance
<b>Components</b>						
Component 2766	11336	595.2028	94 Poncirin	4	20.040	30,625,312
Component 2423	9553	641.1722	94 Isorhamnetin	4	17.005	4,016,400
Component 2358	9287	611.1976	93 Neoesperidin	4	16.516	136,944,496
Component 2268	8931	611.1980	73 Neoesperidin	4	15.897	375,730
Component 2213	8718	581.1872	87 Naringin	4	15.520	41,592,528
Component 1985	7884	611.1611	98 Robinin	4	14.033	16,899,672
Component 1934	7698	597.1822	79 Hesperidin	4	13.695	14,091,042
Component 1913	7603	611.1965	86 Hesperidin	4	13.522	5,888,135
Component 1740	6954	627.1563	98 Rutin	4	12.389	2,455,858
Component 1705	6785	741.2244	91 Rhoifolin	4	12.114	707,318
Component 1660	6584	581.1873	84 Naringin	4	11.763	14,305,817
Component 1585	6325	787.2299	94 Isorhamnetin	4	11.304	3,313,007
Component 1548	6164	803.2244	94 Isorhamnetin	4	11.010	1,965,657
Component 1523	6085	757.2191	96 Kaempferitrin	4	10.856	814,939
Component 1505	6004	935.2670	76 Kaempferitrin	4	10.709	1,112,348
Component 1496	5971	773.2142	92 Robinin	4	10.644	2,863,759
Component 1285	5148	773.2141	93 Rutin	4	9.217	2,942,259
Component 1281	5147	188.0707	97 Rutin	4	9.215	2,378,850
Component 1165	4721	935.2456	84 Kymurenic acid	4	8.461	11,755,922

Figure 8. Partial MS<sup>n</sup> spectral tree search results from the Kale juice sample.

## Breakdown of MS<sup>n</sup> subtree search results:

Subtree search calculates the largest overlap between the potentially large component spectral tree against the library. It not only provides exact compound matches based on an MS<sup>n</sup> tree match, but also provides substructure/subtree matches when the unknown compound does not exist in the reference library. Subtree search results can include the following two types:

Match result type 1 (exact MS<sup>n</sup> tree match): The MS<sup>2</sup> precursors of the unknown compound and library reference matches and the spectral tree match between the unknown compound (MS<sup>2</sup>, MS<sup>3</sup> and MS<sup>4</sup>) and reference (confidence score >60). The unknown compound is confidently identified with full annotation of tree match (Figure 9).

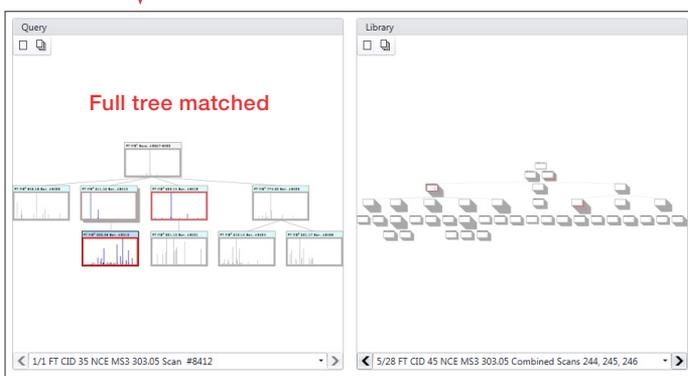
Match result type 2 (partial MS<sup>n</sup> tree match): In most cases, the MS<sup>2</sup> precursor and MS<sup>2</sup> spectra of the unknown compound do not match any library references because of limited reference flavonoid standards. The Subtree search is not limited by this factor. Partial MS<sup>n</sup> spectral tree match provides valuable substructure information for true unknown compounds. When subtree matches between unknown and reference, the sub-structure of the unknown compound is identified to match the reference structure or its substructure, in this case, the flavonoid reference structure (Figure 10). With subtree search, Mass Frontier 8.0 software was able to detect true unknown compounds that belong to the flavonoid compound class with molecular weight, retention time and sub-structure information.

All detected flavonoid class compounds can be exported to a cvs file using the "Export to Grid" tool in the Mass Frontier 8.0.

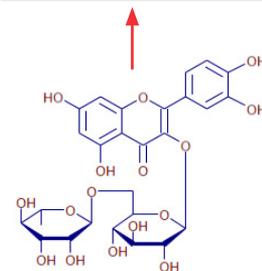
### Detected compound

Name	Scan N...	Precursor...	Match	MSn	t <sub>R</sub> (min)	Abundance
Component 68	8425	611.1616	98 Rutin	3	14.985	2,394,462

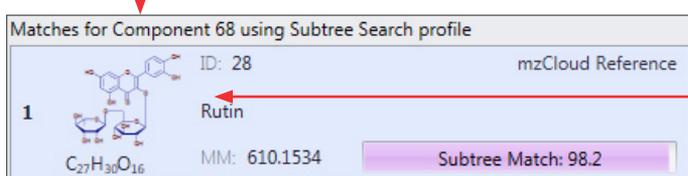
### MS<sup>n</sup> tree library search



Structure of query compound  
M+H: 611.1616



### Library match result



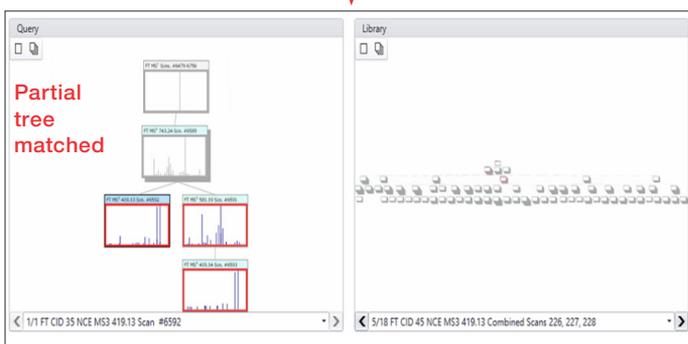
Structure of matched library reference  
M+H: 611.1616

**Figure 9. Full tree match using the Mass Frontier software.** The unknown compound ([M+H]<sup>+</sup>: *m/z* 611.1616) from the juice sample was identified as rutin, with complete structure annotation by an exact MS<sup>n</sup> spectral tree match with the flavonoid rutin reference in the library.

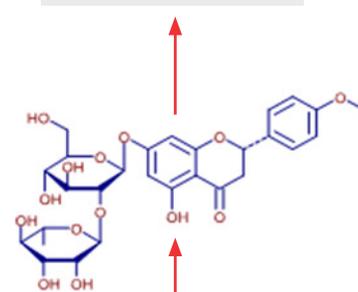
### Detected compound

Name	Scan No.	Precursor <i>m/z</i>	Match	MS <sup>n</sup>	t <sub>R</sub> (min)	Abundance
Component 998	6589	743.2400	83 Naringin	4	11.773	1,787,997

### MS<sup>n</sup> tree library search



Sub-structure of query compound  
M+H: 743.2400



### Library match result



Structure of library reference  
M+H: 581.1876

**Figure 10. Partial tree match using Mass Frontier software.** The unknown compound ([M+H]<sup>+</sup>: *m/z* 743.2394) from the juice sample was identified as a member of the flavonoid class by sub-structure annotation from a partial MS<sup>n</sup> spectral tree match with the flavonoid naringin reference in the library.

## Box 2.

### Arita Lab 6549 flavonoid structure database

A dedicated flavonoid structure database built from 6549 Mol files contributed by Professor Masanori Arita from National Institute of Genetics, Japan.

Professor Arita's laboratory has worked extensively to build a web-searchable, flavonoid database. This database collects original references that report the identification of flavonoid in various plant species. The database consists of three major resources: (flavonoid) compounds, plant species, and references. Currently, 6961 flavonoid structures, 3961 plant species, and 5215 references describing a total of 19,861 metabolite-species relationships are registered. More details can be found on the website [http://metabolomics.jp/wiki/Main\\_Page](http://metabolomics.jp/wiki/Main_Page).

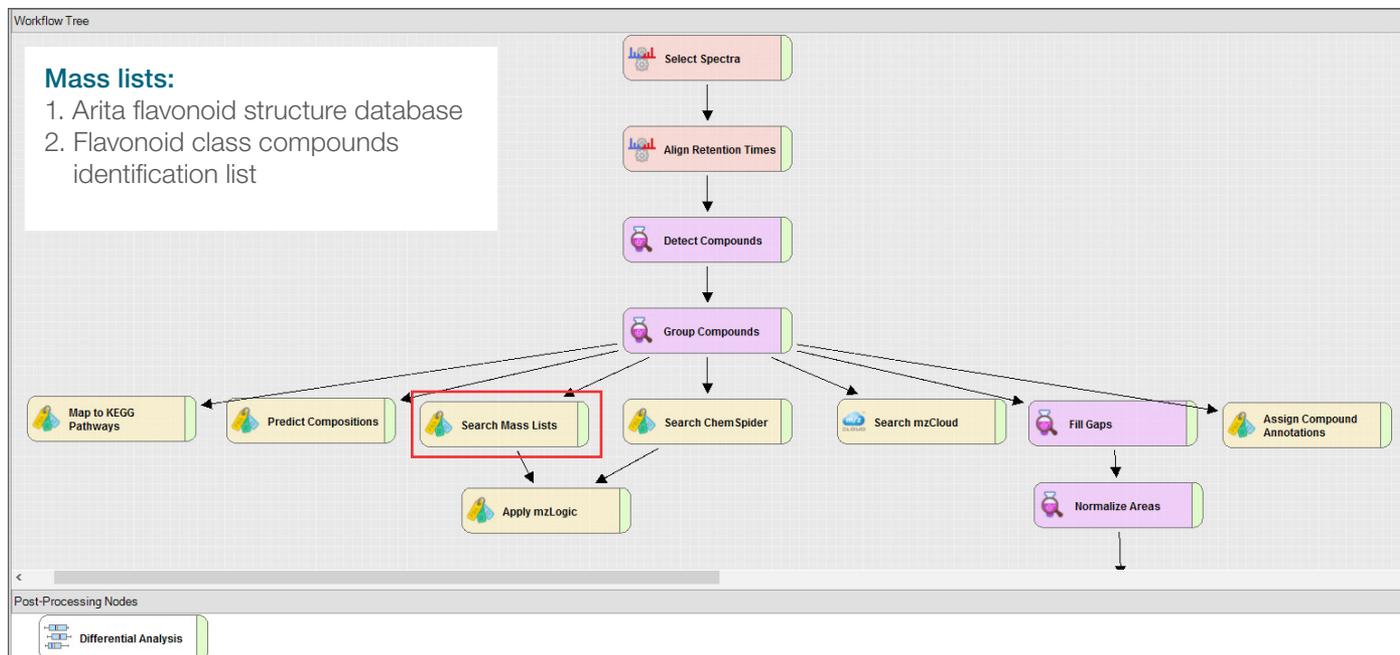


Figure 11. The Compound Discoverer 3.0 software processing workflow used for flavonoid annotation and statistical analysis.

The detected compounds that matched both mass lists (molecular weight from both lists and retention time from juice flavonoid class compound identification result list) are selected for further flavonoid structure annotation. Figure 12 shows that a detected compound with the molecular weight of 742.2320 matched both mass lists. The two isomeric flavonoid structures from the Arita lab 6549 Flavonoid structure database were selected as structure candidates of the compound. In addition, three isomeric flavonoid structures from ChemSpider database were also selected as structure candidates (Figure 13). The ranking of total five structure candidates was carried

out using the FISH (Fragment Ion Search) scoring algorithm (Box 3). To rank the isomeric structures, the software predicts fragmentation of five structure candidates based on known fragmentation rules first, then calculates the FISH scores by matching the predicted fragment ions with the observed fragment ions from MS<sup>n</sup> data. The proposed structure with the highest FISH score represents the best match with the observed fragment ions from the MS<sup>n</sup> data and is the best structure candidate for the unknown flavonoid class compound. Figure 14 shows that this flavonoid is annotated as Narirutin 4'-glucoside based on the FISH score ranking.

**Box 3.****FISh scoring algorithm included in Compound Discoverer 3.0 software**

The FISh scoring algorithm attempts to match the fragment structures in a list of predicted fragments to the centroids in the fragmentation scans of the precursor ions.

When a precursor ion scan is followed by only one fragmentation scan, it calculates the FISh coverage score as follows:

$$\text{FISh coverage score} = \frac{\# \text{ matched centroids}}{\# \text{ used centroids}} \times 100$$

where:

# used (matched + unmatched) centroids represents the number of centroids in the fragmentation scan that are above the user-specified signal-to-noise threshold. The algorithm skips centroids below the user-specified signal to noise threshold.

When a precursor scan is followed by more than one fragmentation scan, it calculates a composite score as follows:

$$\text{FISh coverage score} = \frac{(\sum_{\text{per all scans}} \# \text{ matched centroids})}{(\sum_{\text{per all scans}} \# \text{ used centroids})} \times 100$$

The FISh scoring algorithm annotates the centroids in the fragmentation scans with the matching fragment structures. It also provides a FISh Coverage score for data-dependent scans in the Mass Spectrum view legend and a FISh Coverage score in the results table.

Checked	Name	Formula	Annotation Sc	FISh	Molecular Weight	RT [min]	Area (Max.)	# ChemSpi	# mzCloud	Mass List Matches
<input checked="" type="checkbox"/>	[Similar to: Naringin; ΔMass: -162.0535 Da]	C33 H42 O19	■ ■ ■ ■ ■		742.23274	11.778	788500	30	1	■ ■ ■ ■ ■

Structure Proposals	Compounds per File	Predicted Compositions	Metabolika Results	mzCloud Results	ChemSpider Results	Mass List Search Results	Metabolika Pathways		
	Structure	Compound Match	Name	Formula	Molecular Weight	ΔMass [Da]	ΔMass [ppm]	RT [min]	Reference List Name
1		■	Component 1816		742.23225	-0.00049	-0.66	11.775	072618_threejuice_masslist
2		■	Naringin 4'-glucoside	C33 H42 O19	742.23203	-0.00071	-0.96		Arita Lab 6549 Flavonoid Structure Database
3		■	Naringin 4'-glucoside	C33 H42 O19	742.23203	-0.00071	-0.96		Arita Lab 6549 Flavonoid Structure Database

Figure 12. An example showing unknown flavonoid compound which matched both mass lists.

## Information from MS<sup>n</sup> tree data search results

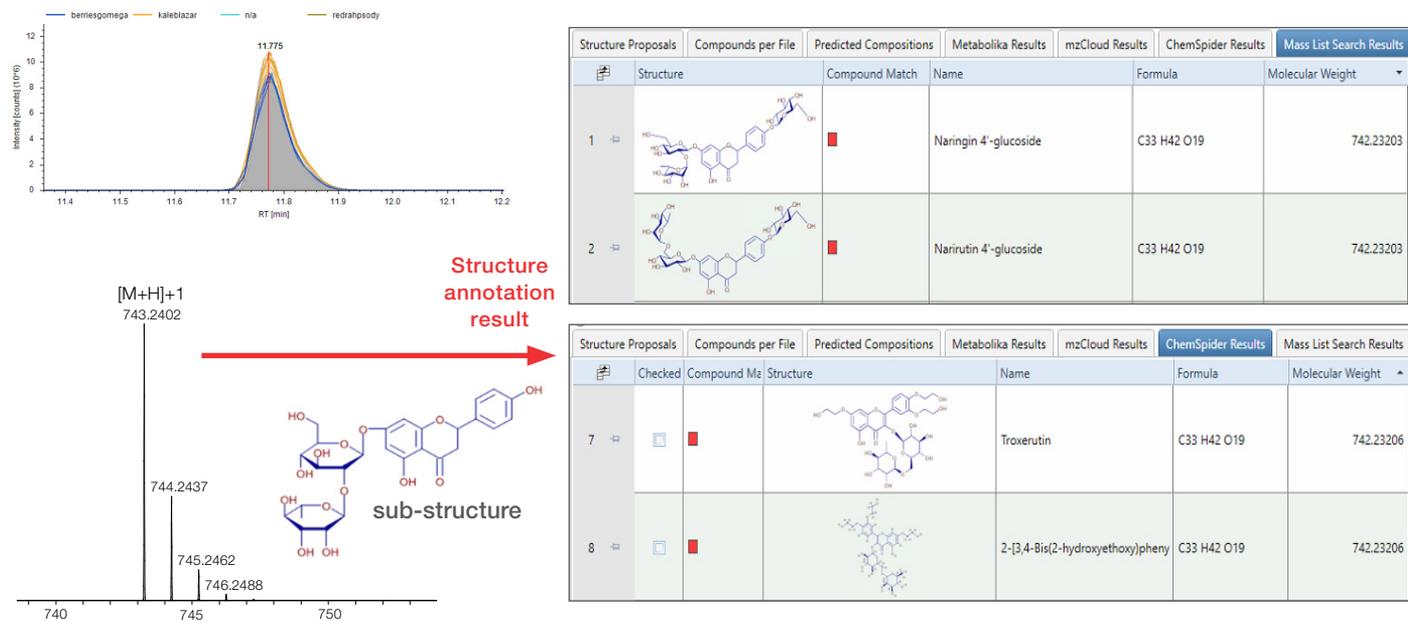


Figure 13. Structure candidates proposed using the Arita flavonoid structure database and the ChemSpider database for the identified flavonoid class compound (MW: 742.2320).

## Information from MS<sup>n</sup> tree data search results

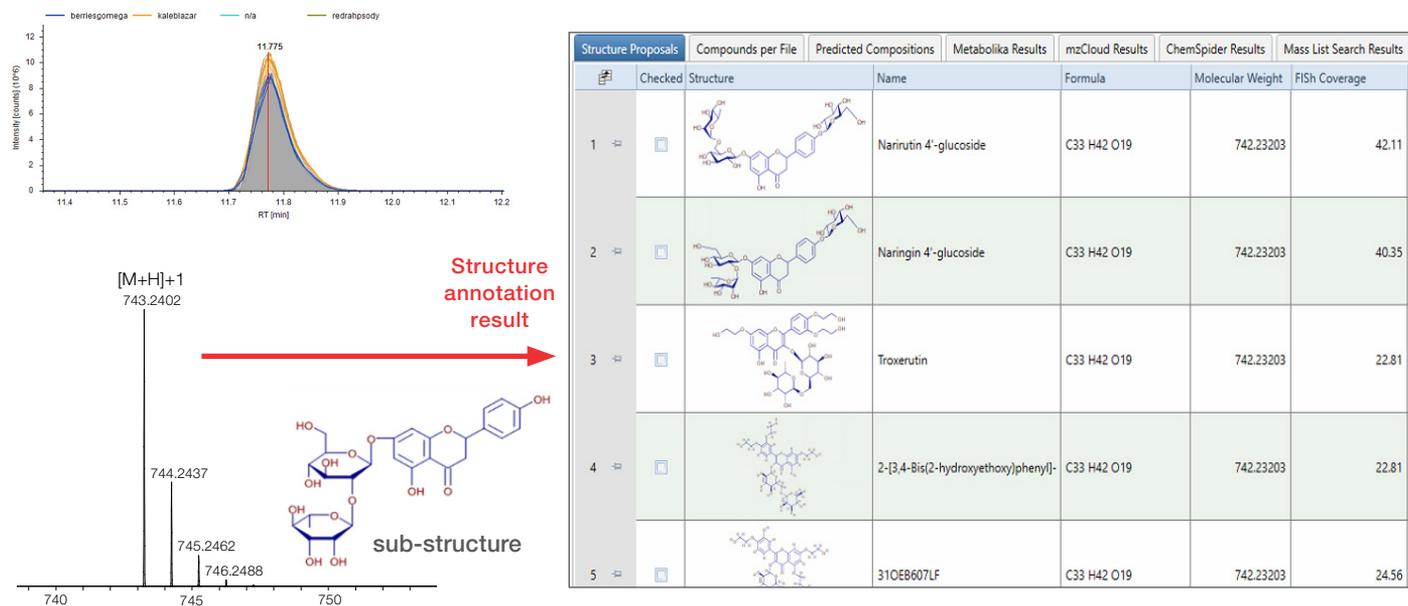


Figure 14. The detected flavonoid class compound (MW: 742.2320) is annotated as Narirutin 4'-glucoside based on the FISH score ranking.

## Discussion

### The new structure-based MS<sup>n</sup> approach enables increased flavonoid identification coverage

Although it has been known that MS<sup>n</sup> spectral tree data provide deeper and more detailed fragmentation pathways, thus enabling more structural information for flavonoid annotation<sup>4</sup>, the use of an MS<sup>n</sup> workflow for flavonoid annotation was historically hindered by two factors: (i) the MS<sup>n</sup> instrument method setup was difficult for non-mass-

spectrometer-expert users, and (ii) the MS<sup>n</sup> spectral tree data processing was a bottleneck because it required manual fragment ion assignments that were based upon extensive expert knowledge about flavonoid chemical structure and fragmentation rules. Our newly developed structure-based MS<sup>n</sup> approach addresses both of these challenges and allows unknown flavonoid annotation with increased throughput and coverage.

Firstly, by offering a ready-to-use structure-specific MS<sup>n</sup> instrument template, everyone can easily acquire high-quality of MS<sup>n</sup> data using this template.

Secondly, by offering new software tools including the sub-tree search process in Mass Frontier 8.0 software, new Mass List format allowing structure attachment and FiSh score calculated using an MS<sup>n</sup> spectra, the rich structure relevant fragment ion information from MS<sup>n</sup> spectra tree can be processed in an automatic fashion without need to know any specific fragmentation rules.

Combining the new instrument template and software tools, the new approach takes full advantage of deeper and more structurally relevant fragment ion information offered by MS<sup>n</sup>, enabling more flavonoid compounds to be annotated compared to an MS<sup>2</sup> only approach. Partial MS<sup>n</sup> spectral tree match results provided valuable sub-structure information for true unknown compounds. With subtree search, Mass Frontier 8.0 software identified true unknown compounds that belong to the flavonoid compound class

which do not have exact references in the mzCloud library. Table 3 shows that six unknown flavonoids which are rutin and its secondary metabolites with different modifications were identified from the juice samples using MS<sup>n</sup> tree data, while MS<sup>2</sup> data only identified rutin. Table 4 shows that five unknown flavonoids which are isorhamnetin and its secondary metabolites with different modifications were identified using MS<sup>n</sup> tree data while MS<sup>2</sup> data only identified isorhamnetin and one of its secondary metabolites. Figure 15 shows a comparison of the number of annotated flavonoids from the three fruit and vegetable juice samples using MS<sup>2</sup> only data and MS<sup>n</sup> spectral tree data with 2D column and Venn diagram formats. The MS<sup>2</sup> only approach annotated sixty-two flavonoids, while the MS<sup>n</sup> approach annotated the same sixty-two flavonoids and plus sixty-seven more flavonoids. As a result, the MS<sup>n</sup> approach showed a two-fold increase in annotations to the MS<sup>2</sup> only approach.

**Table 3. Identified rutin and its secondary metabolites using MS<sup>2</sup> vs MS<sup>n</sup>.**

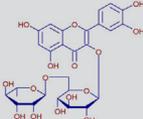
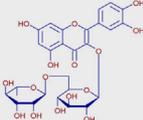
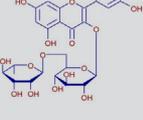
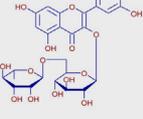
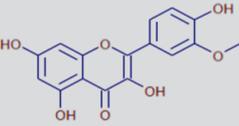
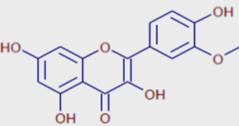
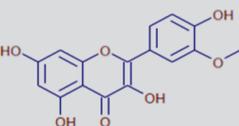
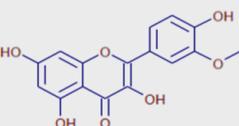
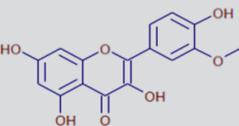
Molecular Weight	ID Structure/Substructure in MF 8.0	Identified with MS <sup>2</sup> in CD 3.0	Identified with MS <sup>n</sup> and FiSh score in CD 3.0
610.1539	C27H30O16 MW: 610.1534 	●	●
626.1490	C27H30O16 MW: 610.1534 	✗	●
756.2120	C27H30O16 MW: 610.1534 	✗	●
772.2071	C27H30O16 MW: 610.1534 	✗	●
788.2023	C27H30O16 MW: 610.1534 	✗	●
950.2328	C27H30O16 MW: 610.1534 	✗	●

Table 4. Identified isorhamnetin and its second metabolites using MS<sup>2</sup> vs MS<sup>n</sup>.

Molecular Weight	ID Structure/Substructure in MF 8.0	Identified with MS <sup>2</sup> in CD 3.0	Identified with MS <sup>n</sup> and FISH score in CD 3.0
316.0590	C16H12O7 MW: 316.0583 	●	●
478.1122	C16H12O7 MW: 316.0583 	✗	●
624.1698	C16H12O7 MW: 316.0583 	●	●
640.1652	C16H12O7 MW: 316.0583 	✗	●
786.2226	C16H12O7 MW: 316.0583 	✗	●

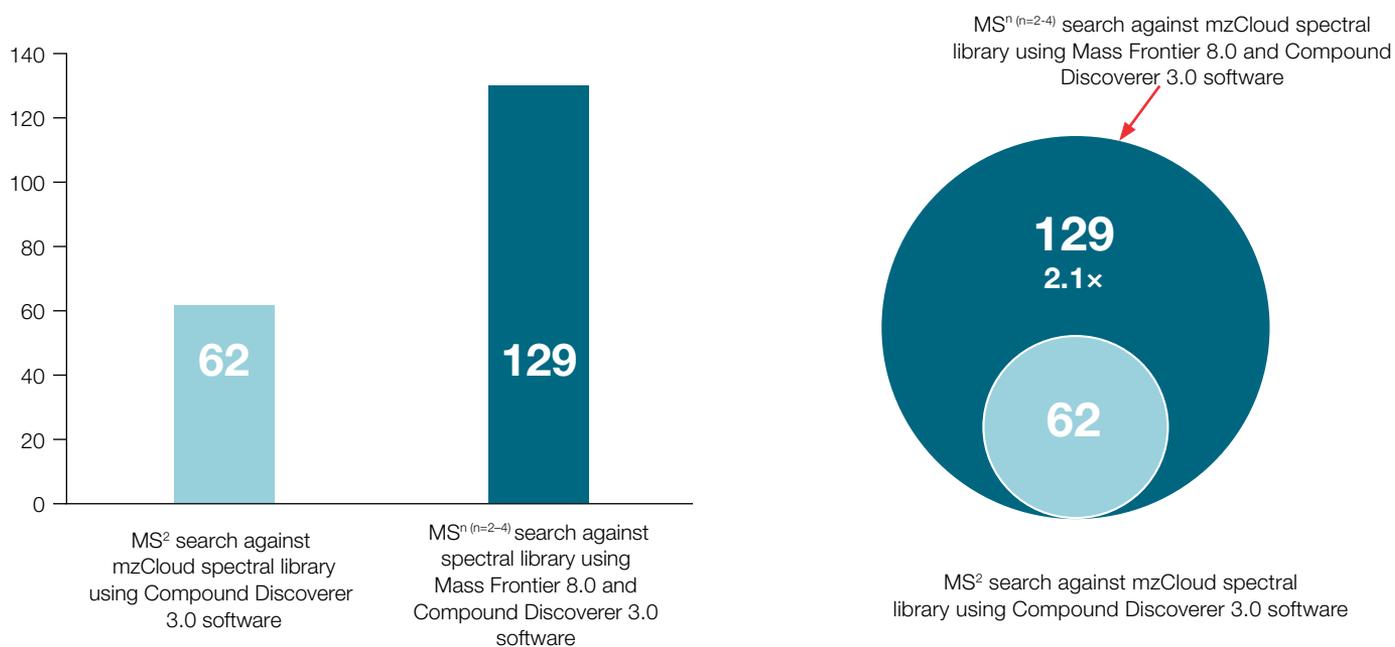


Figure 15. Comparison of the number of detected flavonoids from the juice samples with full structure annotation obtained using MS<sup>2</sup> only and MS<sup>n</sup> tree spectral data.

## Advantage to enable simultaneous quantitation and statistical analysis

Our newly developed structure-based MS<sup>n</sup> approach also enables simultaneous quantitation of identified flavonoid compounds and statistical analysis. We designed the instrument template using a short cycle time (1.2 second) in order to get enough scan points across the chromatographic peak for precise quantitation while collecting MS<sup>n</sup> spectral tree data in the same LC-MS run. Higher annotation coverage of flavonoid compounds using the new developed structure-based MS<sup>n</sup> approach

enabled more data points for precise statistical analysis. As shown from the heat map (Hierarchical Cluster Analysis) of detected flavonoids (Figure 16), more high abundance flavonoids were detected from Kale Blazer and Berries Gomega juice samples.

On another hand, most flavonoids detected from the Red Rhapsody juice sample were less abundant. Figure 17 depicts an extracted ion chromatogram for the quantitation of detected neoeriocitrin. Figure 18 shows that the three juice samples are clearly differentiated by principal component analysis.

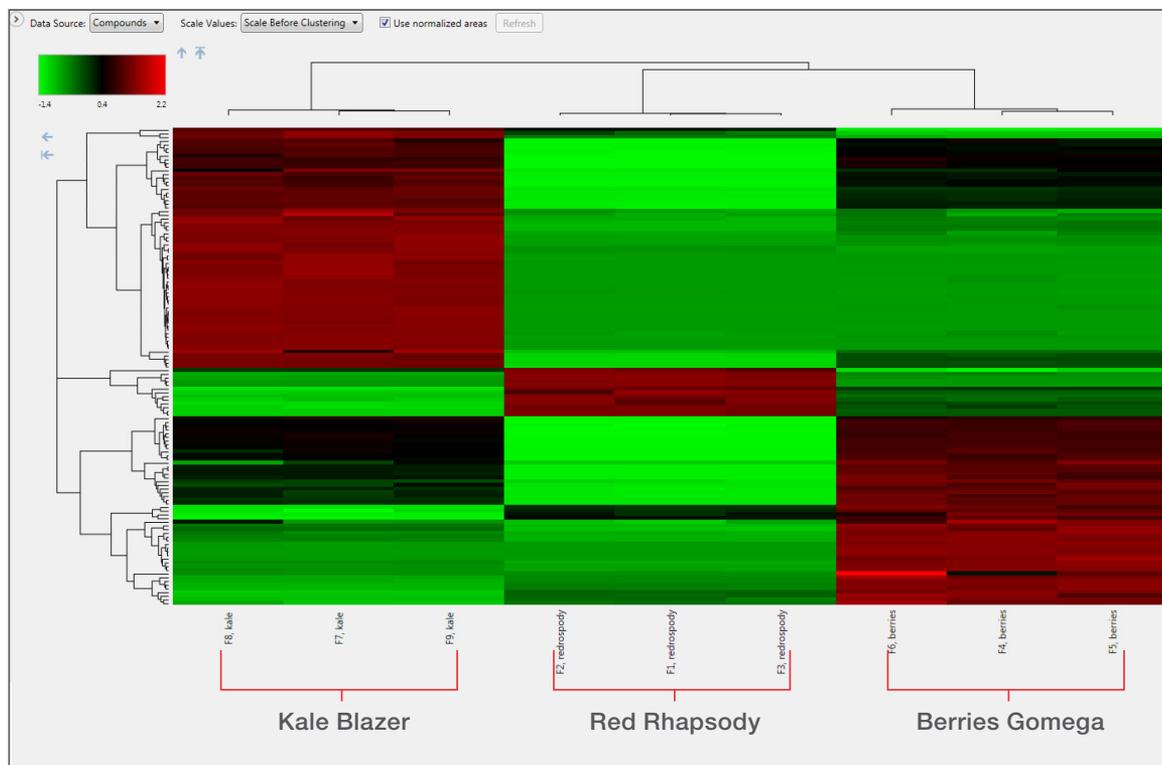


Figure 16. Hierarchical cluster analysis.

## Extracted ion chromatogram

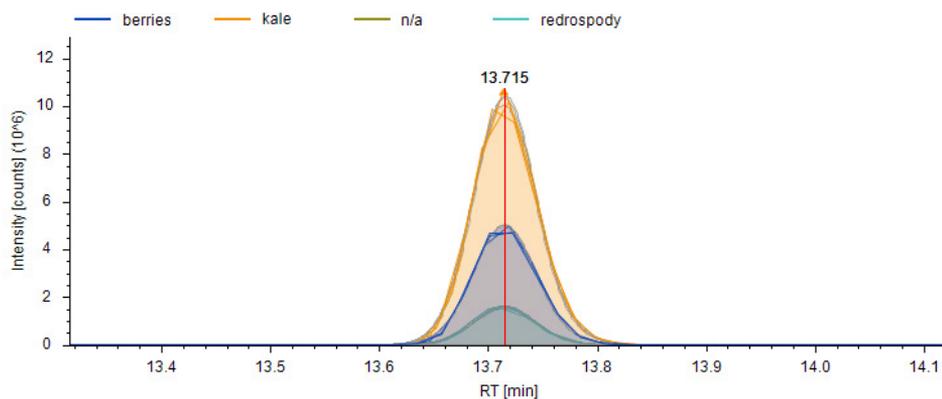
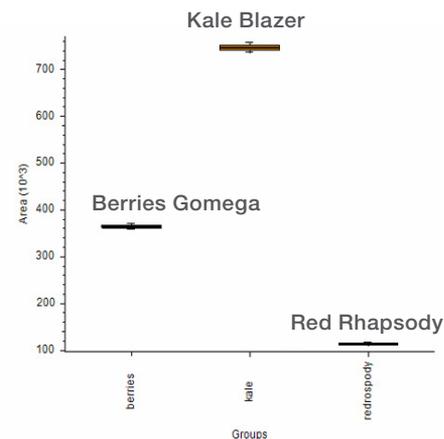


Figure 17. Extracted ion chromatograms and trend chart of neoeriocitrin.

## Trend Chart



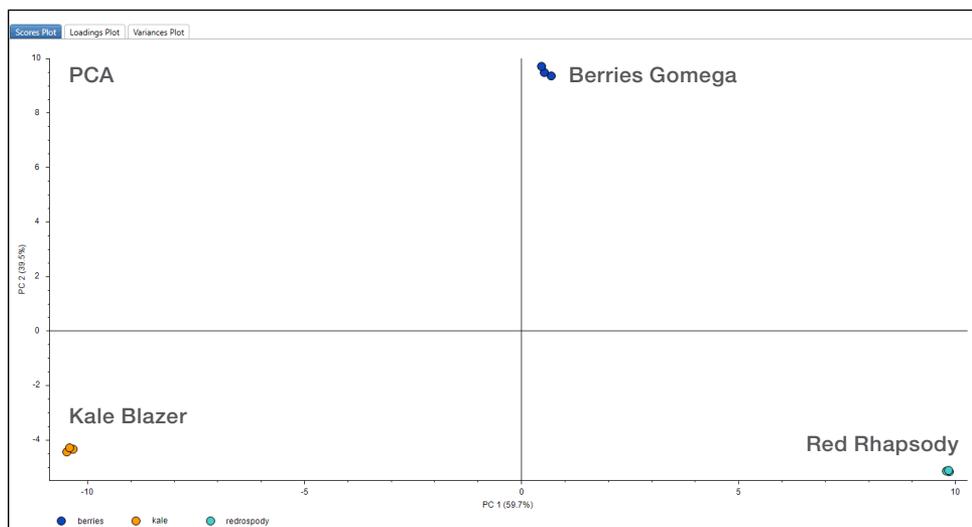


Figure 18. PCA of flavonoid compounds identified from the three juice samples.

## Conclusions

We have developed a new structure-based MS<sup>n</sup> approach to address the challenges of unknown flavonoid annotation. This new approach enables rapid flavonoid class compound detection and further structure annotation without the need to have expertise in flavonoid chemical structure and fragmentation rules. Two times more flavonoids were annotated from the three fruits and vegetable juice samples using this structure-based MS<sup>n</sup> approach compared to MS<sup>2</sup> only approach. This new structure-based MS<sup>n</sup> approach also allowed quantitation of annotated flavonoids simultaneously in a single LC-MS run. In addition to the class of flavonoid compounds, the concept of this structure-based MS<sup>n</sup> approach can be applied to other classes of compounds, such as steroids and endocannabinoids.

## References

1. <http://metabolomics.jp/wiki/Category:FL>
2. P. Kachlicki, A. Piasecka, M. Stobiecki, L. Marczak. Structural Characterization of Flavonoid Glycoconjugates and Their Derivatives with Mass Spectrometric Techniques, *Molecules* **2016**, 21, 1494.
3. D. Tsimogiannis, M. Samiotaki, G. Panayotou, V. Oreopoulou. Characterization of Flavonoid Subgroups and Hydroxy Substitution by HPLC-MS/MS, *Molecules* **2007**, 12, 593–606.
4. J.J.J. van der Hooft, J. Vervoort, R. J. Bino, J. Beekwilder, R.C.H. de Vos. Polyphenol Identification Based on Systematic and Robust High-Resolution Accurate Mass Spectrometry Fragmentation, *Anal. Chem.* **2011**, 83, 409–416.
5. M. Arita, K. Suwa. Search extension transforms Wiki into a relational system: A case for flavonoid metabolite database. *BioData Mining* **2008** 1:7.

Find out more at [www.thermofisher.com/orbitrapID-X](http://www.thermofisher.com/orbitrapID-X)