

# Sample Profiler – A Cloud-Based Application for Confirming Food Authenticity

Srividya Kailasam, Ralf Tautenhahn, and Tim Stratton, Thermo Fisher Scientific, ITPL Road, Bangalore, Karnataka, India, 560048; 355 River Oaks Parkway, San Jose, CA, 95134 USA

## ABSTRACT

**Purpose:** Establishing food authenticity involves analyzing the HRMS data of a large number of samples of known quality to determine their chemical composition. The chemical profile of genuine samples can be used for comparing against suspect samples. In this study, soy sauces with different characteristics, such as samples manufactured in different regions, or those packaged in different types of containers, or even samples manufactured using different processes, were analyzed and their chemical compositions determined.

**Methods:** HRMS data was collected using a benchtop Orbitrap™ mass spectrometer post reverse phase separation using a simple linear gradient. The data was analyzed using a cloud-based software application, Sample Profiler.

**Results:** Chemical databases or 'profiles' of samples having different characteristics were easily determined, and comparing samples against these profiles yielded valuable information regarding their similarities and differences.

## INTRODUCTION

Confirming food authenticity is important not only for consumers and the food industry but also for government regulators. There is often the need to detect trace levels of contaminants in complex matrices, which requires the use of state-of-the-art mass spectrometers. Sophisticated software tools are needed to quickly process large data sets generated in these studies.

Thermo Scientific™ Sample Profiler is a new cloud-based software application that can be applied in diverse fields such as food authenticity, environmental monitoring, metabolomics, and detection of new designer drugs. Using this application, large data sets can be processed to detect and integrate compound peaks. Further, compound identification is automated by comparing the fragmentation spectra of compounds against mzCloud™, a web-based spectral database.

Using the Sample Profiler application, chemical profiles of samples of known quality and authenticity can be created within minutes. Individual samples of suspect/questionable quality can be readily compared against these profiles. Statistical evaluation of these comparisons helps users identify certain compounds of interest or 'marker compounds' that are present in significantly different concentration levels in the two sets – authentic and suspect samples. Thus, the presence or absence of marker compounds helps with the quick verification of authenticity and/or characterization of challenge samples when compared against profiles of samples of known/desired characteristics.

In this study, we used Sample Profiler to distinguish between soy sauce samples that were brewed during manufacturing from those that were not.

## MATERIALS AND METHODS

### Sample Preparation

Different brands of soy sauces were purchased locally. Samples were prepared in triplicate by diluting aliquots of soy sauce with water (3:1). Pooled quality control (QC) samples were prepared by combining a fixed amount of each soy sauce into a pool and preparing a set of dilution samples as normal.

### Test Method

Analysis was performed by LC-MS with separation achieved by reversed-phase linear gradient (ACN/water, 0.1% formic acid) with a run time of 10 min. Mass spectrometric analysis was performed on a Thermo Scientific™ Q Exactive™ HF hybrid quadrupole-Orbitrap mass spectrometer operated in positive mode with electrospray ionization. Full scan MS1 (120,000 resolution FWHM at  $m/z$  200) and precursor selected MS/MS (30,000 resolution) data was acquired. Sample injection order was randomized with QC samples injected every 10 injections.

### Data Analysis

Mass spectral data for different types of soy sauces, such as those manufactured in different regions, stored in different types of containers, and even processed differently (by brewing or by not brewing), were collected. Data for blanks and QCs were also obtained. Raw data files were uploaded to Thermo Fisher Cloud and imported into the Sample Profiler application for further processing. The components were detected in samples after the application aligned all the retention times in all samples. Compound identification was done by matching fragmentation (MS2) spectra against the mzCloud spectral database. Sample Profiler aggregated the name, predicted formulae,  $m/z$ , RT, and area counts of each compound from all the samples.

The Sample Profiler application allows the user to visualize the total ion current chromatograms, extracted ion chromatograms, trendlines and distribution plots for all the detected compounds. In addition to creating a database of all compounds detected in all the samples, the user can also easily create desired subsets or 'profiles' from this master-database. Profiles were created based on the region of manufacturing and manufacturing process itself. Further, Sample Profiler allows the user to set up comparative analysis of user-selected samples against specific profiles to determine differences between the two.

## RESULTS

### Background

Soy sauce is a popular condiment manufactured and consumed in many Asian countries. Traditionally, soy sauces are prepared by two methods: a) by naturally brewing fermented soya beans with wheat and brine, which can take several months, or b) by using hydrolyzed vegetable protein (HVP), which takes only days. Soy sauces manufactured by the two processes differ in their physical characteristics such as aroma, flavor, and texture. In addition, soy sauces made by acid-hydrolysis of soy proteins have a longer shelf life.

### Data Processing in the Cloud

Large data sets uploaded to the Data Manager in the Thermo Fisher Cloud can be easily added to projects created in Sample Profiler and processed using suitable processing methods. The application mines the raw data for accurate mass and retention times of the compounds along with their area counts in all the samples where they are found. From the accurate mass, formulae are predicted and by comparing fragmentation spectra and predicted formulae against mzCloud and/or ChemSpider databases, compounds are identified. As can be seen in Figure 1, the distribution of intensities for each compound in all the samples where it is present is also shown in the Compounds table.

### Creating Profiles

After all samples including QCs and blanks are processed, two databases of compounds or profiles, a) All Samples and b) All QCs samples, are automatically created by the application. The user can also create additional profiles or subsets of the All Samples profile by partitioning it based on study factors or criteria such as manufacturing region or packaging material used to characterize the samples. For example, in the present study, a database or a profile of compounds present in soy samples that were brewed during manufacture was created.

Figure 1, which is a snapshot of the profile of "Brewed" soy sauces, shows details such as compound name, predicted formula, MW, RT, and highest intensity value from all samples for 11,572 compounds found in this subset of soy sauces. The distribution plots shown in the last column give an idea of the spread in the intensity values found across all samples. This table can be filtered to show the compound details in any of the profiles created in the project.

Figure 1. Profile of soy samples brewed during manufacturing – A snapshot.

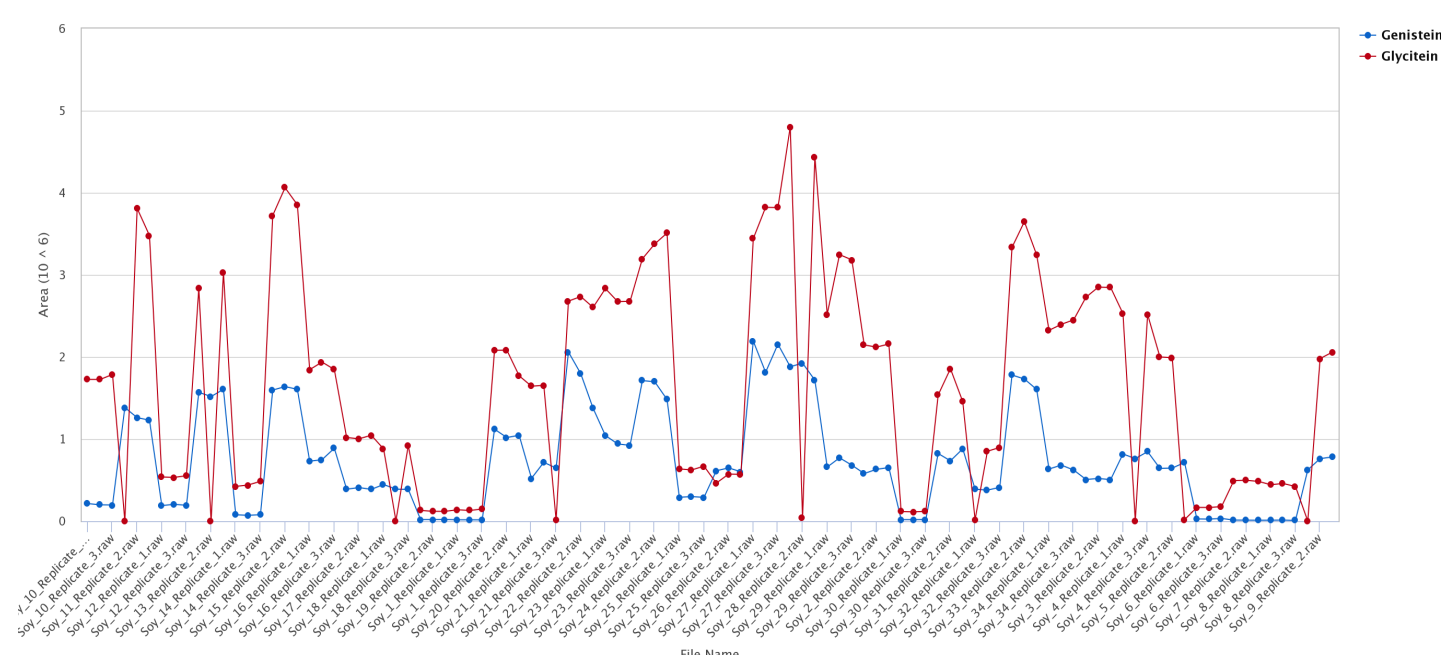
Marker	Name	Predicted Formula	Molecular Weight	RT (min)	Max Area	Distribution
1	C15H12N2	C15H12N2	228.19238	4.856	3.11e+6	
2	ethyl 4-hydroxy-2-phenylpyridine-5-carboxylate	C13H12N2O3	246.09815	1.966	3.05e+6	
3	Glycine	C2H4N2O2	294.06795	5.621	3.00e+6	
4	2816-N	C18H18N	119.07343	4.074	3.59e+6	
5	N-(3-benzoyloxycarbonyl-L-threonine)	C12H15NO5	253.09403	2.410	3.59e+6	
6	2816-N	C18H18N	119.07343	5.621	3.59e+6	
7	3-tert-butyl-1-methyl-1H-imidazo-5-carboxamide	C19H24N2O2	102.10543	1.854	3.59e+6	
8	methyl (R)-2-carboxylate-3-carboxylate	C15H13NO2	226.07402	3.245	3.58e+6	
9	2,5-Dioxo-1-oxindole-N-(3-benzoyloxycarbonyl-L-threonine)	C17H13NO6	348.13155	3.959	3.58e+6	
10	2716-N	C17H19NO3S	291.12570	0.899	3.57e+6	

### Trendline Plots

Sample Profiler also allows for visualization of the trends in compound areas. When trendlines for compound areas across all samples in the profile are plotted, those samples that may have significantly higher or lower concentrations of one or more compounds can be easily identified. These samples can then be investigated further to get a better understanding of the products' chemical composition.

Soy products are known to contain isoflavones like glycitein and genistein. When the area trendlines of these two compounds across all samples in the "Brewed" profile were plotted as shown in Figure 2, they exhibit similar variations across samples.

Figure 2. Trendlines showing the variation in compound intensities across all soy sauces.



### Comparing Different Types of Samples

Using Sample Profiler, comparisons between different types of soy sauces can be easily set up and data processed at different fold changes and adjusted p-values. In the present study, a profile of brewed soy sauces was created and used for comparing against not-brewed samples. Data analysis performed by setting the fold change value to 8 and adjusted p-value to > 0.001. From Figure 3 it can be seen that of a total of 11,803 compounds, the not-brewed samples were found to contain 267 compounds in significantly higher abundance, and 1791 compounds in significantly lower abundance than in brewed samples.

Figure 3. Comparison of not-brewed soy samples to a profile of brewed soy samples.

Mr.	Name	Predicted Formula	Molecular weight	RT (min)	Profile Median	Compared Sa.	Match S.	p-Val.	Adjusted p-Value (0.001 and Log2 Fold Change > 3 (FC) > 8)	Ratio	Log2 F.	Distribution
3	N-(4,5-Tetrahydroxybenzo-)	C12H15NO6	268.08660	2.542	2.05e+3	1.90e+5	Additional	1.9e-54	3.09e+2	91.002	6.51	
4	4-methylpiperazino-1-	C10H15NO2	208.11629	1.102	1.05e+3	1.41e+5	Additional	3.9e-54	2.46e+2	72.334	6.18	
5	2-Methyl-2-propenyl[2,2-	C14H21NO7	315.15118	4.137	1.53e+5	3.36e+3	Match	2.91e-54	1.52e+2	0.222	-5.51	
6	C18H19NO2P S	348.10213	6.159	2.01e+3	6.36e+4	Additional	6.52e-54	1.01e+2	3.522	-4.93		
7	C18H15NO5	335.17257	6.388	7.34e+4	2.79e+3	Match	9.96e-54	1.02e+2	0.036	-4.81		
8	Genistein	C15H10O5	270.05228	6.266	7.34e+5	1.73e+4	Match	7.41e-54	1.29e+2	0.204	-5.41	
9	C13H19NO3	237.13004	3.542	2.23e+4	9.82e+4	Match	7.52e-54	1.29e+2	0.944	-4.91		
10	4-Oxopiperazine	C12H14N2O3	234.10019	3.894	2.07e+3	1.17e+5	Additional	1.9e-54	1.49e+2	66.241	6.05	

Of the 11,803 compounds found in all the samples in this comparison, 7041 compounds with match status 'Match' were found to be present in both, 4531 compounds with match status 'Missing' were found only in brewed samples, and 231 compounds with match status of 'Additional' were found only in not-brewed soy samples. It can be seen that the application also determines the distribution of compound intensities in samples present in both groups – the profile and compared samples at the defined statistical thresholds.

Brewed soy sauces have lower abundances of sorbic acid and higher abundances of azelaic acid than not-brewed soy sauces. The comparative trends of compound intensities in the two groups can be visualized in Figures 4 and 5.

Figure 4. Trendline showing the relative abundances of azelaic acid in brewed and not-brewed soy sauces.

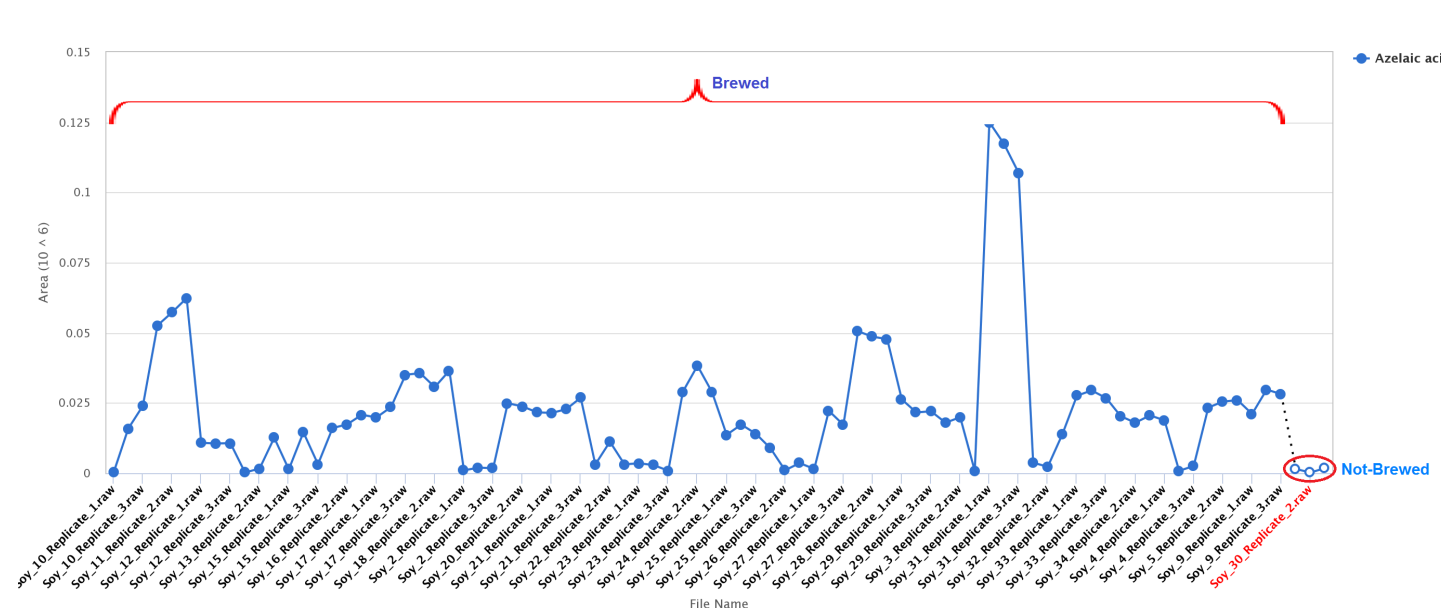
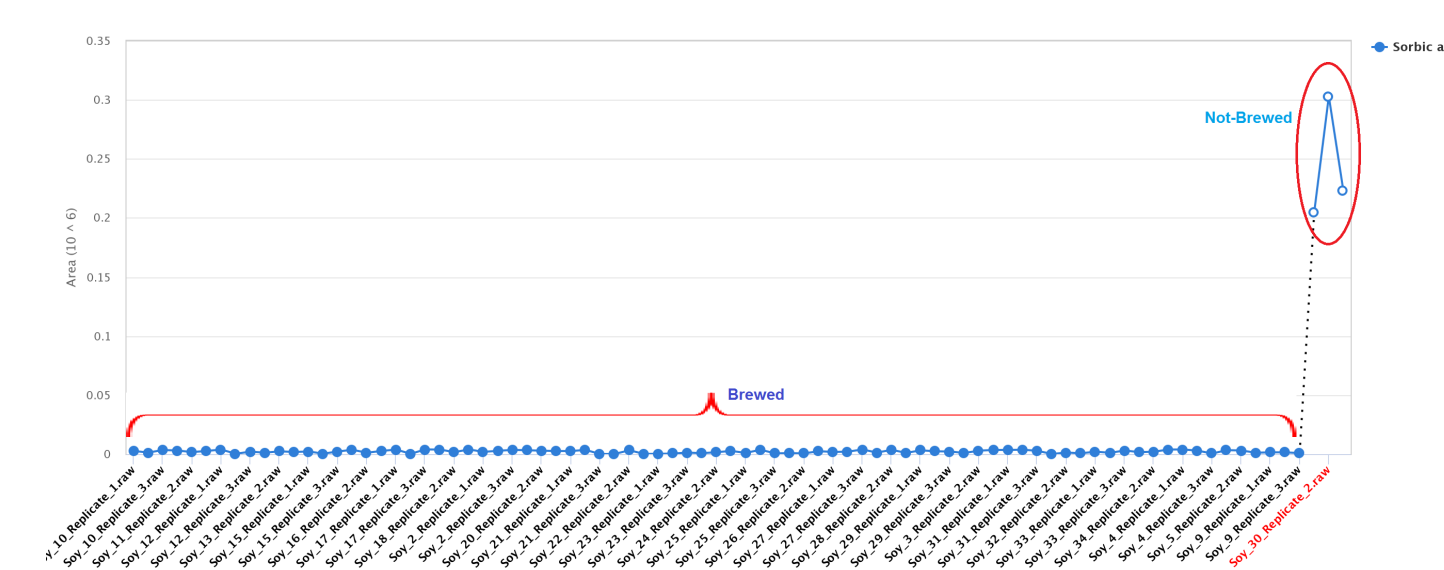


Figure 5. Trendline showing the relative abundances of sorbic acid in brewed and not-brewed soy sauces.



### Marker Compounds

A key feature of this application is the ability to search for the presence or absence of compounds of interest or 'marker compounds' in all the processed samples and comparisons by specifying either one or more of the following four parameters: compound name, formula, MW, or RT. After a marker has been defined, the table is automatically updated with relevant information such as Found In X of Y samples (Figure 6).

In this study, Sample Profiler was used to search for genistein and glycitein specified by name. Post search, the table was updated formulae, MW, and RTs of the two compounds along with the total number of samples and QCs in which they were found.

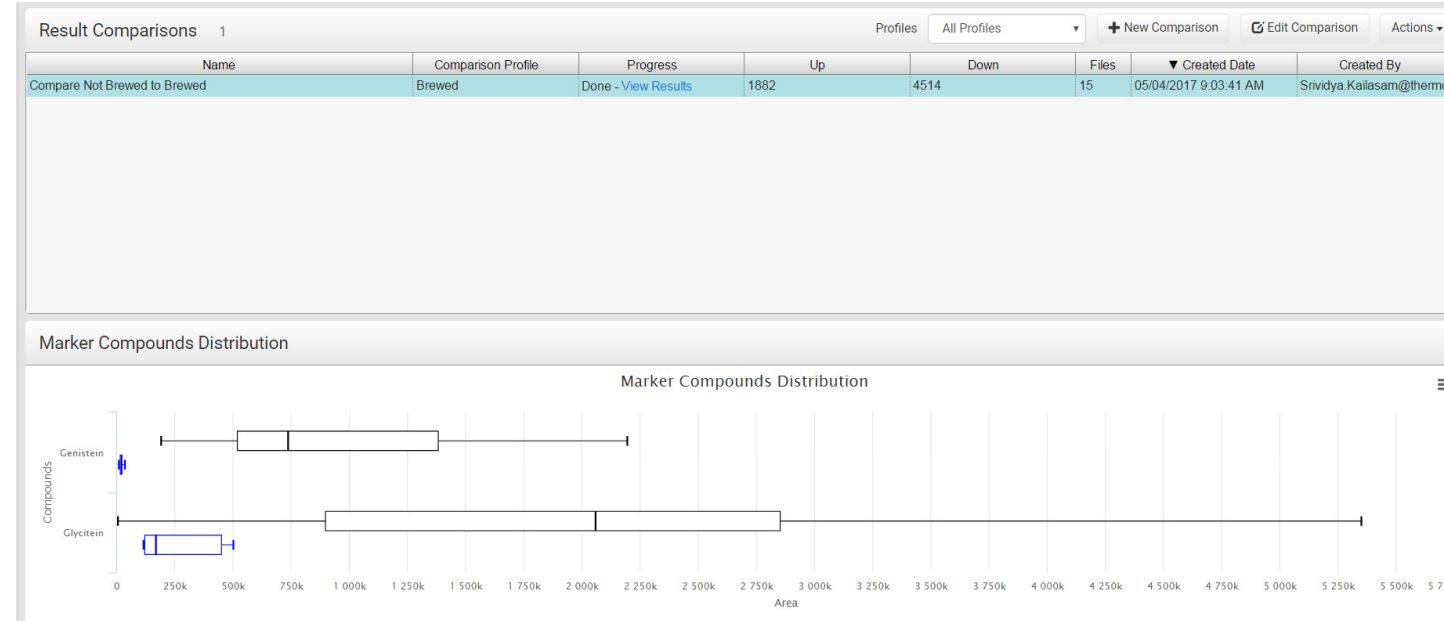
Figure 6. Marker compounds expected to be found in brewed soy sauces.

Assign N.	Match By	Compound Name	Formula	MW	MW Tolerance (ppm)	RT (Observed (min))	RT Tolerance (min)	Found In
1	RT	Genistein	C15H10O5	270.05228	5	6.266	0.006	102 of 114 samples
2	RT	Glycitein	C13H12O2	246.06798	5	5.628	0.006	109 of 114 samples

After the compounds are added to the table, Sample Profiler will automatically search for their presence in any comparison set up in the project.

In Figure 7, the relative distributions of the two marker compounds in the comparison between brewed and not-brewed soy sauces are shown. For both compounds, the distributions shown in blue are for abundances in not-brewed samples and the distributions shown in black are for abundances in brewed soy samples. From these results, brewed soy sauces are understood to have higher abundances of these two marker compounds.

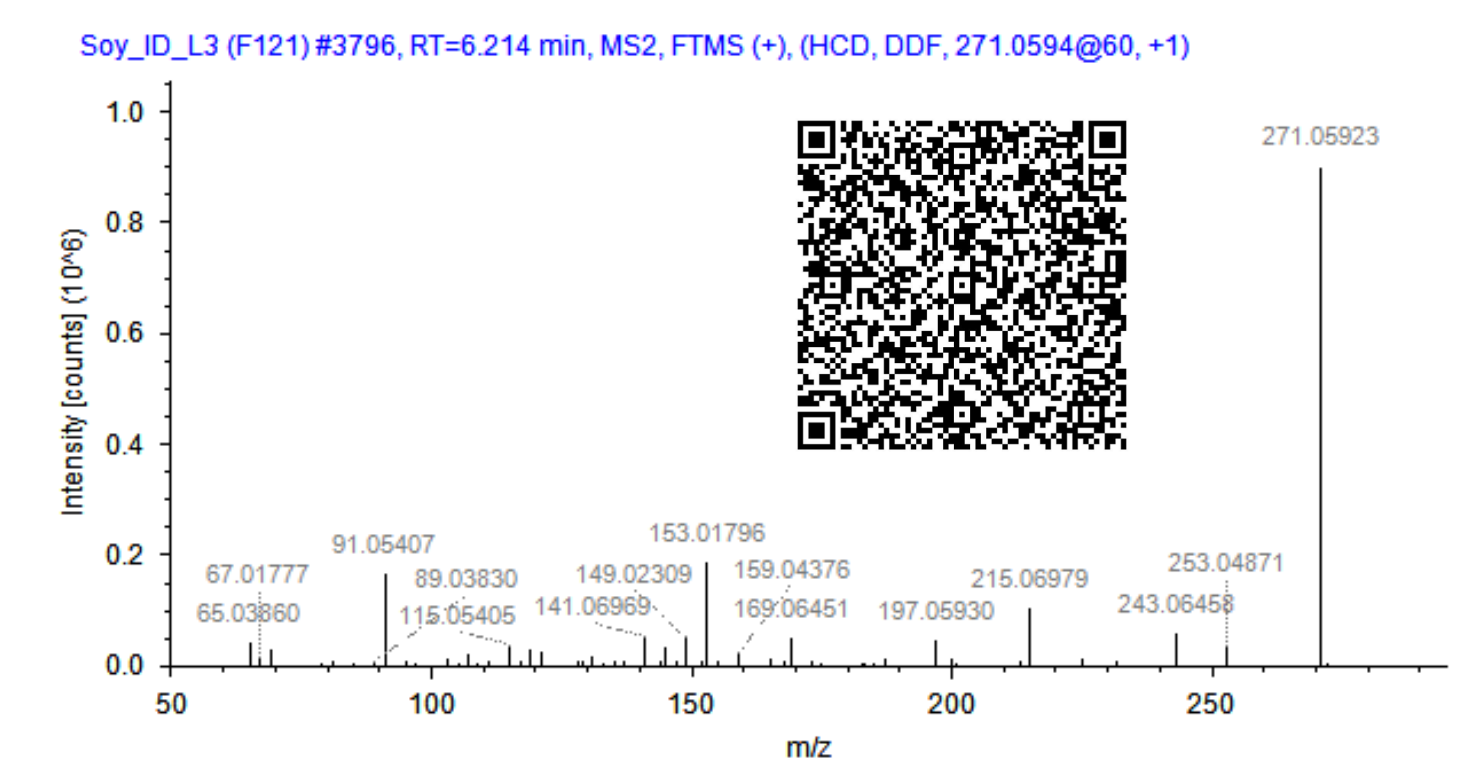
Figure 7. Relative distributions of marker compounds in the highlighted comparison.



This knowledge can be applied to quickly verify the authenticity of a challenge sample(s) by setting up appropriate comparisons between them and known profiles and studying the relative distributions of marker compounds.

mzCloud is a web-based mass spectral database that is extensively used in the Sample Profiler application for compound identification by comparing fragment spectra against the vast and growing library. Figure 8 shows the MS2 spectrum of genistein along with the SpectraQR.

Figure 8. Fragmentation spectrum of genistein.



Note: The displayed SpectraQR can be searched against mzCloud using the free mzCloud app (for iPhone® or Android™ devices). Scan the QR code with the mzCloud app and click "Search → Compound ID".

## CONCLUSIONS

- Compounds such as glycitein and genistein were found to be present in higher abundance in brewed soy sauces as compared to not-brewed samples.
- Soy sauces manufactured in different regions or packaged differently can be easily differentiated by the Sample Profiler application.
- Sample Profiler allows for visualization of trends in the data in addition to TIC and XIC traces.
- Marker compounds help with easy verification or characterization of suspect/challenge.

## TRADEMARKS/LICENSING

© 2017 Thermo Fisher Scientific Inc. All rights reserved. mzCloud is a trademark of HighChem, Ltd. iPhone is a registered trademark of Apple Inc. Android is a trademark of Android Google Inc. All other trademarks are the property of Thermo Fisher Scientific and its subsidiaries. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.