

# Black Pepper Authenticity Workflow Using the High-Resolution Agilent 7250 GC/Q-TOF

## Author

Sofia Nieto and  
Melissa Churley  
Agilent Technologies, Inc.

## Abstract

Black pepper is a highly valued commodity known to be subject to economically motivated adulteration.<sup>1</sup> This application note discusses a novel food authenticity workflow targeted for routine analysis that allows to distinguish black pepper samples from different geographic regions and detect adulteration. The workflow uses high-resolution mass spectrometry (HRMS) with Agilent Mass Profiler Professional (MPP) and Agilent MassHunter Classifier software to build and apply the classification model.

## Introduction

Food fraud is a rapidly growing and evolving business,<sup>2</sup> therefore several analytical methods have recently been developed to address the concern regarding food adulteration. Targeted methods, measuring the presence or absence of certain analytes are simpler to use but also easier to counteract by sophisticated fraudsters. The technique of HRMS, typically time-of-flight (TOF), combined with a nontargeted approach is a fast rising method for detecting food fraud and adulteration.<sup>3,4</sup> In this approach, an additional statistical software package is used to determine if a given sample belongs to a determined sample group. High resolution and sensitivity in full spectrum acquisition mode enable TOF analyzers to distinguish large numbers of features that can be used for sample classification. A nontargeted approach ensures that dozens of features are selected to differentiate an authentic sample from an adulterated one. The more features included in a class prediction model, the more effort it would take to generate an adulterated sample that resembles an authentic one, even with a low degree of adulteration.

This application note describes an authenticity workflow using the high-resolution Agilent 7250 GC/Q-TOF and a classification model to distinguish

black pepper adulteration with Szechuan pepper or papaya seeds. While this study is focused on the GC/Q-TOF workflow, a similar approach can be used with an LC/Q-TOF and is described elsewhere.<sup>5</sup>

## Experimental

### Sample preparation

Black pepper from two different geographical regions (Malabar, from India, and Phu Quoc, from Vietnam), as well as Szechuan pepper and papaya seeds (Figure 1) were ground. In separate sample groups for each adulterant, Szechuan pepper and papaya seeds were mixed in varying proportions to Malabar to mimic 5 to 50% adulteration. All samples (0.5 g each) were extracted sequentially using hexane and acetone (4 mL each), and extracts were combined. Method blank samples were prepared by extraction using the same solvents. The extracts were filtered through 0.45  $\mu\text{m}$  nylon filters.

### Data acquisition and processing

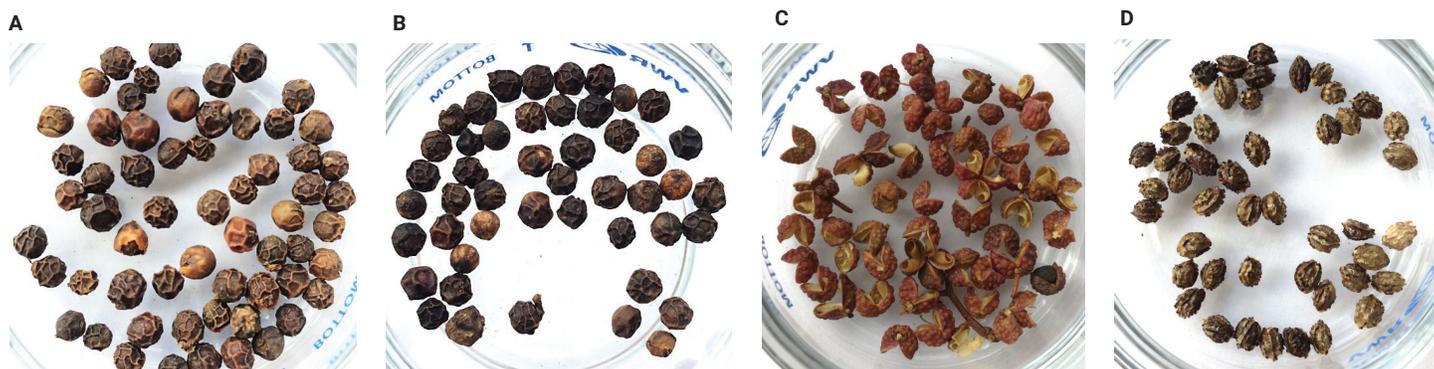
The samples were analyzed in random order using an Agilent 7890 GC coupled to a high-resolution 7250 Q-TOF MS in full spectrum acquisition mode. The data were acquired in standard electron ionization (EI) as well as low energy EI modes. Low energy EI mode, which helps to enhance the molecular ion in the spectrum, was used to confirm the identity of some of the

**Table 1.** GC/Q-TOF acquisition parameters.

GC and MS Conditions	Q-TOF (7250)
GC	Agilent 7890B
Column	Agilent DB-5MS UI, 30 m $\times$ 0.25 mm, 0.25 $\mu\text{m}$
Inlet	MMI, 4-mm UI liner single taper with wool
Injection Volume	1 $\mu\text{L}$
Injection Mode	Split, 10:1
Inlet Temperature	280 $^{\circ}\text{C}$
Oven Temperature Program	50 $^{\circ}\text{C}$ for 2 min; 10 $^{\circ}\text{C}/\text{min}$ to 300 $^{\circ}\text{C}$ ; 10 min hold
Carrier Gas	Helium
Column Flow	1.2 mL/min
Transfer Line Temperature	300 $^{\circ}\text{C}$
Quadrupole Temperature	150 $^{\circ}\text{C}$
Source Temperature	200 $^{\circ}\text{C}$
Electron Energy	70 eV
Emission Current	5 $\mu\text{A}$
Spectral Acquisition Rate	5 Hz
Mass Range	45 to 650 $m/z$

compounds of potential interest or to identify an unknown. The conditions are listed in Table 1. When using the low energy EI mode, the data acquisition was performed at 12 eV using 0.4  $\mu\text{A}$  emission current.

The chromatographic deconvolution and NIST17 library search were performed using the Agilent MassHunter Unknown Analysis tool of MassHunter Quantitative Analysis Software 10.1. The retention indices (RIs) for analyte peaks were calculated based on the alkane ladder to ensure correct compound identification.



**Figure 1.** Pure samples used in this study: (A) Malabar, (B) Phu Quoc, (C) Szechuan, and (D) papaya seeds.

The classification model based on the GC/Q-TOF data was built in Mass Profiler Professional 15.1. Classifier 1.1 was used to validate the model as well as to process "unknown" samples.

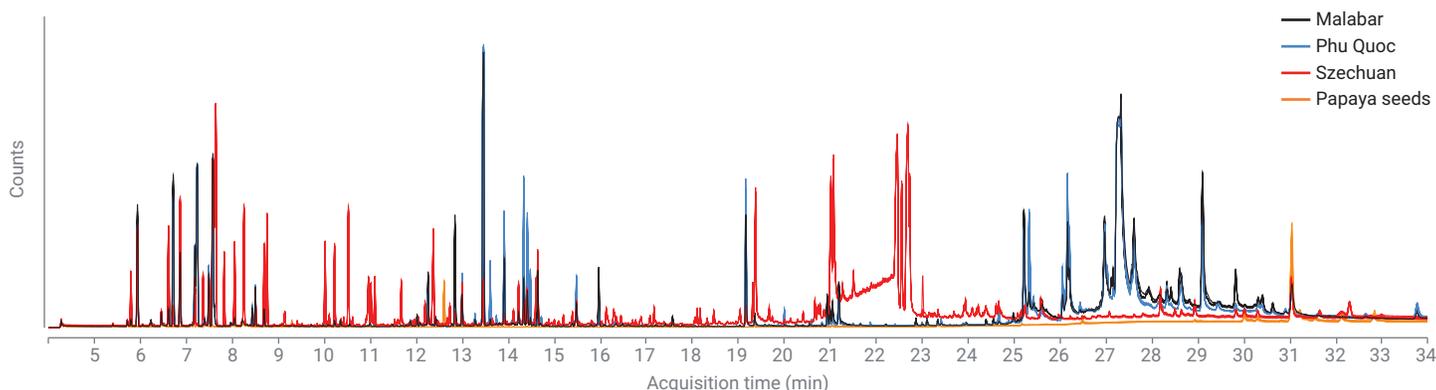
The data from pure samples of each type (black peppers Malabar and Phu Quoc, Szechuan pepper, and papaya seeds) were used to build the classification model. "Unknown" samples that consisted of Malabar black pepper mixed with either Szechuan or papaya seeds at various proportions, were analyzed in the same sequence together with positive controls (additional pure samples) and negative controls (additional samples of Malabar "adulterated" with either Szechuan or papaya seeds).

## Results and discussion

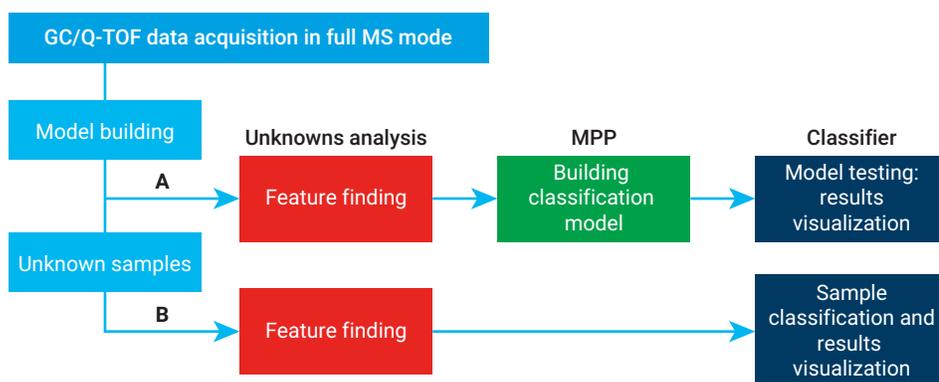
### Building the classification model

The data acquired from six replicate extracts of each type of pepper as well as papaya seeds were used to build the classification model. Figure 2 shows a chromatogram overlay of six replicates of each extract. Malabar and Phu Quoc black pepper extracts displayed similar chromatographic profiles, while Szechuan appeared quite distinct, with many unique and abundant peaks. The chromatographic profile of the papaya seeds extract was relatively simple compared to pepper extracts, which presented an additional challenge for detecting adulteration with this type of matrix.

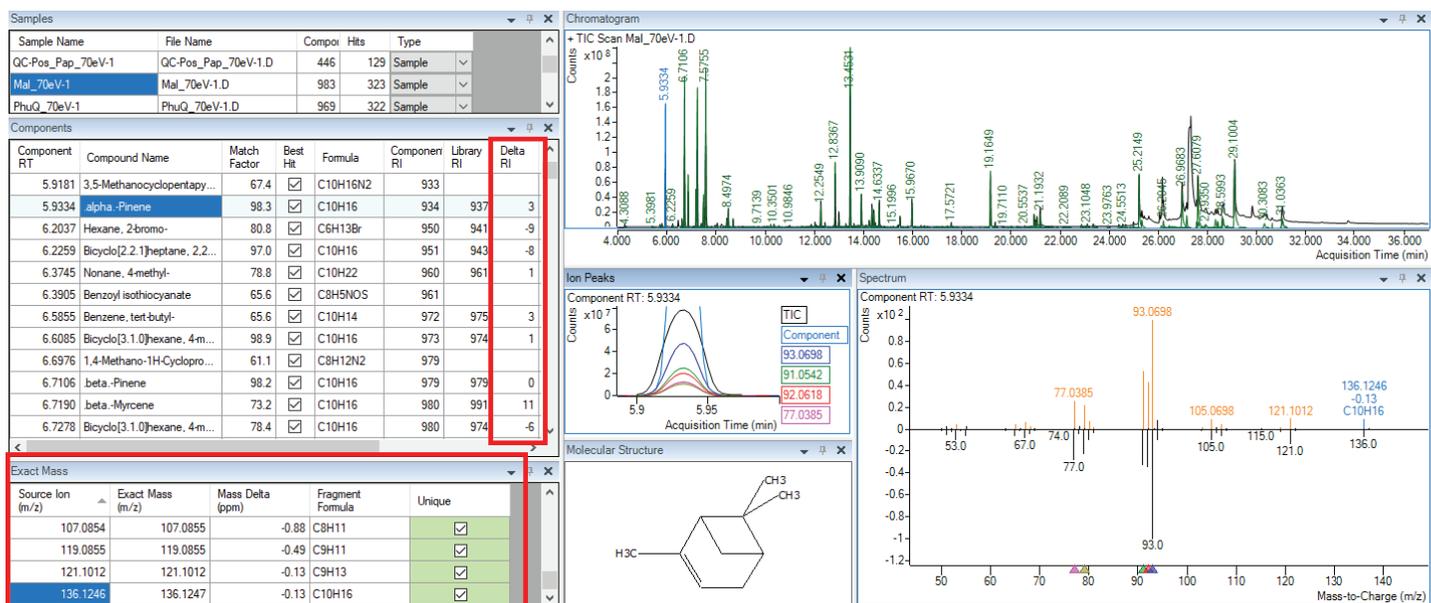
The general workflow for building a classification model and processing unknown samples is outlined in Figure 3. The first step involved data acquisition for the pure samples and negative controls for model validation. The accurate mass GC/Q-TOF data were processed using Unknowns Analysis to perform chromatographic deconvolution and NIST17.L library search (Figures 3 and 4). While it is not necessary for this type of workflow to identify compounds, the library search helped to ensure that important features were included into the classification model. To help correctly identify the compounds, the RI calculation was performed. Furthermore, confirmation of the compound identification based on accurate mass information was achieved using the ExactMass feature (Figure 4).



**Figure 2.** Chromatogram overlay of six replicates of each extract.

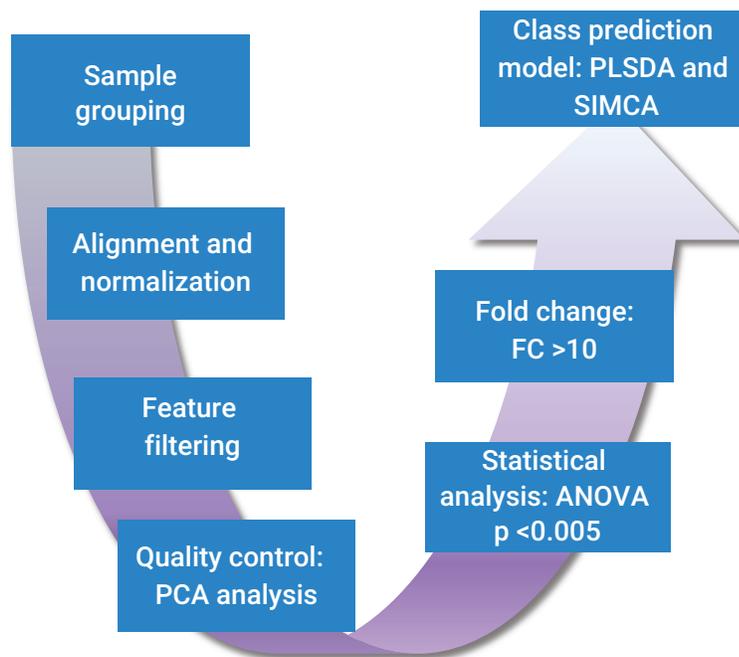


**Figure 3.** Overview of food authenticity workflow using GC/Q-TOF. (A) Building classification model. (B) Analysis of unknown samples.



**Figure 4.** Feature finding in Unknowns Analysis. Highlighted in red boxes: RI calibration function helps confirm compound ID. ExactMass feature provides additional confirmation of compound identification using accurate mass spectral information.

MPP was then used to build, test, and validate the classification model. An example of the workflow in MPP used in this study is outlined in Figure 5. After the sample grouping and alignment steps, the normalization was performed using an internal standard (triphenyl phosphate). Feature filtering was then applied to eliminate poorly reproducible features within a sample group. Next, the separation between the sample groups was evaluated using a principle component analysis (PCA) plot. In the following steps, statistical analysis (ANOVA) and the fold change (FC) analysis were performed using the strict cutoff of  $p < 0.005$  and  $FC > 10$ , respectively. While the number of features gets significantly reduced during this process, this assists in creating a robust classification model.



**Figure 5.** An example of MPP workflow for building a classification model.

After the classification model was built, the data were exported to the Classifier software, which enables convenient results visualization and reporting to facilitate model testing. When the model was finalized, the unknown samples were processed in Unknowns Analysis and Classifier, bypassing data processing in MPP (Figure 3B).

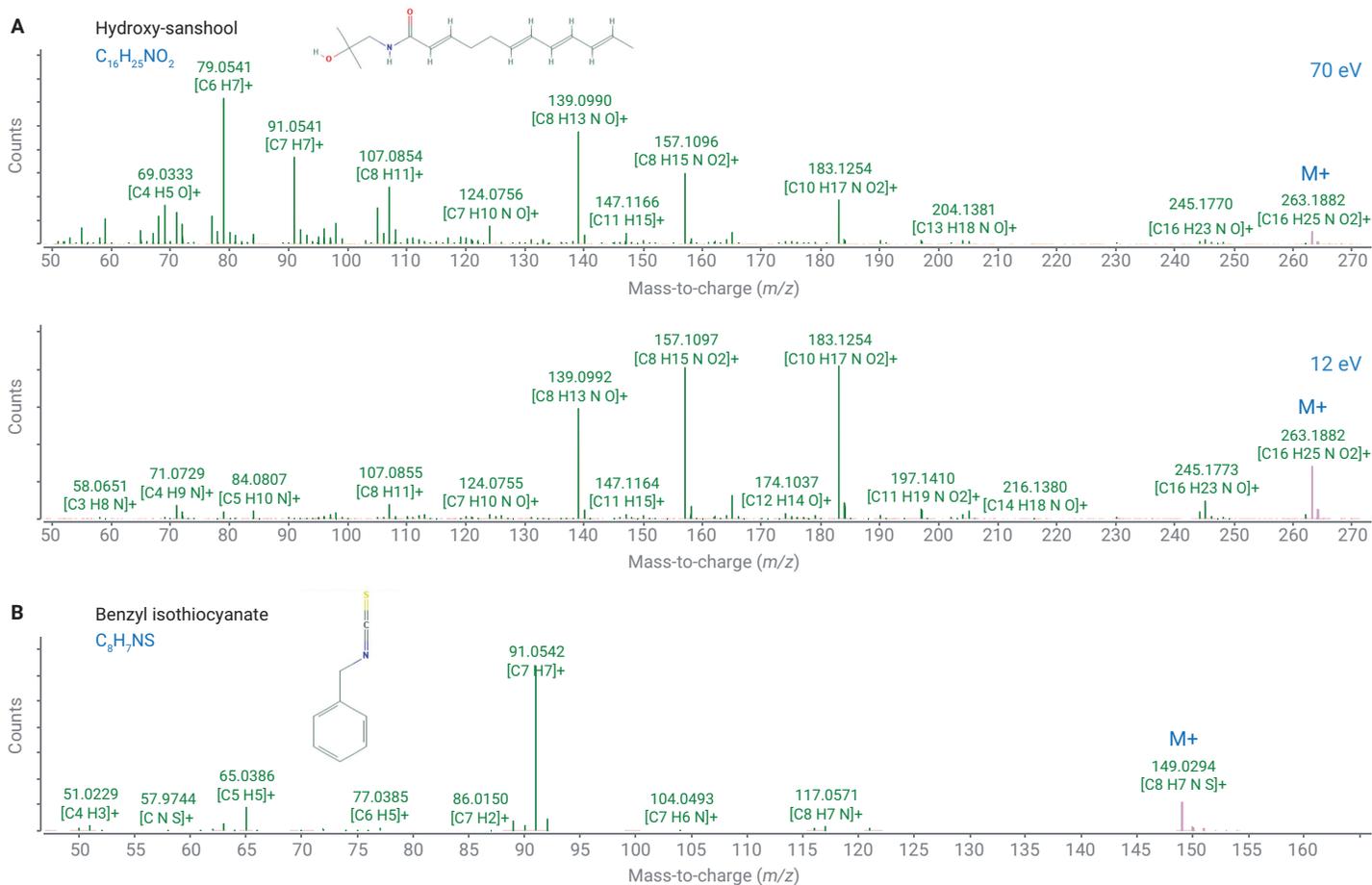
The classification models were built using two different algorithms: partial least squares discriminant analysis (PLSDA) and soft independent modeling of class analogy (SIMCA). Note that the PLSDA model reports a confidence score, while the SIMCA model uses a distance score.

### Overview of the compounds included in the class prediction model

Table 2 lists some of the most abundant compounds from each type of pure sample included in the classification model. Notably, for black peppers Malabar and Phu Quoc, these included the primary marker compound of black pepper: piperine, as well as some of the secondary marker compounds, which were terpenes  $\beta$ -pinene, 3-carene, d-limonene, and caryophyllene. For Szechuan, some of its characteristic compounds such as eucalyptol and hydroxy-sanshool isomers were also among the compounds included in the model. Finally, the papaya seeds extracts were rich in nitrogen-containing compounds such as benzyl isothiocyanate, as well as sterols and fatty acids.

**Table 2.** Major compounds included in the classification model. Relative amounts in each of the extracts (% base peak on the chromatogram) are shown. Percent of base peak was calculated after averaging across all the replicates from each group (<1% if not indicated).

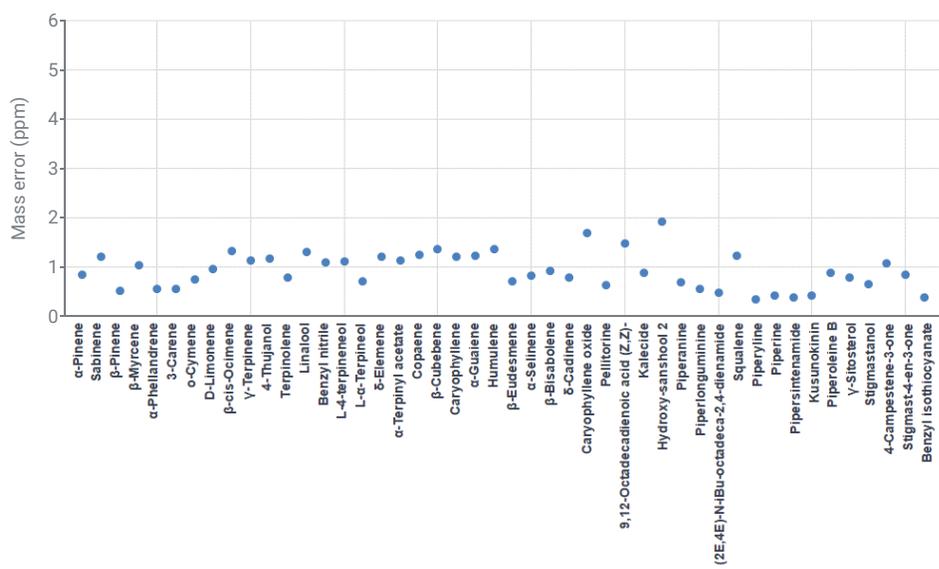
	RT	Malabar	Phu Quoc	Szechuan	Papaya Seeds
		% Base Peak			
$\alpha$ -Pinene	5.93	21.7	16.9	8.4	
Sabinene	6.61			16.7	
$\beta$ -Pinene	6.71	37.0	31.3		
$\beta$ -Myrcene	6.86	1.3	2.0	19.4	
$\alpha$ -Phellandrene	7.18	4.5	10.6		
3-Carene	7.23	64.5	86.6		
o-Cymene	7.49	2.1	1.7		
D-Limonene	7.58	23.9	21.4	22.1	
Eucalyptol	7.64	23.7	1.6	100	
$\beta$ -cis-Ocimene	7.82			3.2	
$\gamma$ -Terpinene	8.04			4.5	
4-Thujanol	8.25			4.0	
Terpinolene	8.49		1.0		
Linalool	8.69			2.1	
Benzyl nitrile	9.30				2.8
L-4-terpineneol	10.01			2.8	
L- $\alpha$ -Terpineol	10.22			2.5	
$\delta$ -Elemene	12.26		1.1		
$\alpha$ -Terpinyl acetate	12.37			10.3	
Benzyl isothiocyanate	12.61				80.7
Copaene	12.84	4.1			
$\beta$ -Cubebene	12.99		1.4		
Caryophyllene	13.45	24.9	59.6		
$\alpha$ -Guaiene	13.61		1.6		
Humulene	13.91	1.1	3.4		
$\beta$ -Eudesmene	14.33		4.4		
$\alpha$ -Selinene	14.40		3.3		
$\beta$ -Bisabolene	14.47		1.5		
$\delta$ -Cadinene	14.64			2.2	
Caryophyllene oxide	15.48		1.7		
Pellitorine	19.16	5.4	4.8		
9,12-Octadecadienoic acid (Z,Z)-	21.02			5.3	
Kalecide	21.17		1.1		
Hydroxy-sanshool 2	22.55			4.3	
Piperanine	25.23	18.6	3.7		
Piperlonguminine	25.29		3.8		
(2E,4E)-N-Isobutyloctadeca-2,4-dienamide	26.22		2.9		
Squalene	26.48				2.4
Piperlyline	26.96	4.4	3.7		
Piperine	27.21	100	100		
Pipersintenamide	27.61	2.2	11.0		
Kusunokinin	28.64		2.3		
Piperoleine B	29.01	8.3	5.1		
$\gamma$ -Sitosterol	31.03	1.6	1.5	1.6	100
Stigmastanol	31.22				8.1
4-Campestene-3-one	31.54				2.9
Sitostenone (Stigmast-4-en-3-one)	32.79				7.8



**Figure 6.** Example of spectra for two unique and predominant compounds identified in “adulterants” in (A) Szechuan pepper, at 70 and 12 eV and (B) in papaya seeds. Structures are from the PubChem database.

Example spectra for two unique and predominant compounds identified in the “adulterants,” namely Szechuan and papaya seeds, are shown in Figure 6. The Fragment Formula Annotation tool of MassHunter Qualitative Analysis 10 was used to confirm the compound identity. Note that hydroxy-sanshool does not have an EI spectrum in NIST17. However, a tentative ID can still be assigned based on published data,<sup>6</sup> low energy EI data (for molecular ion confirmation) (Figure 6A, lower panel), and results from the fragment formula annotation.

Excellent mass accuracy (mass error within 2 ppm) was observed for all the major compounds included in the model (Figure 7).

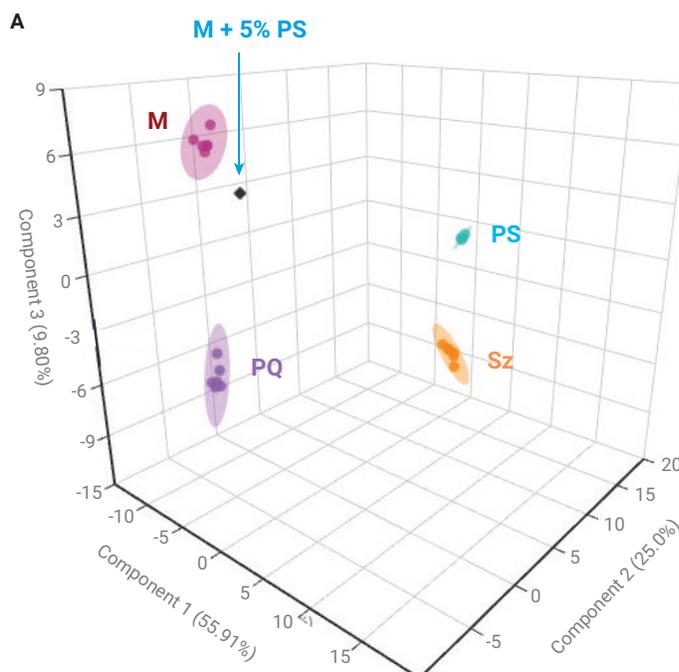


**Figure 7.** Mass accuracy of the most prominent model compounds.

## Analysis of unknown samples in Classifier

The PLS-DA and SIMCA class prediction models were exported from MPP to the Classifier software for further validation. Two models were created for each class prediction algorithm: one that included papaya seeds and one that did not. Since it is not always feasible to predict all the possible adulterants, papaya was omitted from some of the models to mimic this situation. The models were evaluated using the samples adulterated with 5 to 50% of either papaya seeds or Szechuan pepper.

Figure 8 shows visualization of some of the results for the SIMCA model on a 3D PCA plot. Extracts of Malabar, Phu Quoc, and Szechuan were all well clustered. Malabar sample adulterated with 5% papaya seeds was well separated from the other sample groups and can easily be distinguished (Figure 8A). The results displayed in the compound table (Figure 8B) for this sample indicated that a few compounds were out of the model range. The identities for the model compounds are not displayed in the Classifier software, so that the analyst performing the routine analysis would not know which compounds are used in the classification model.



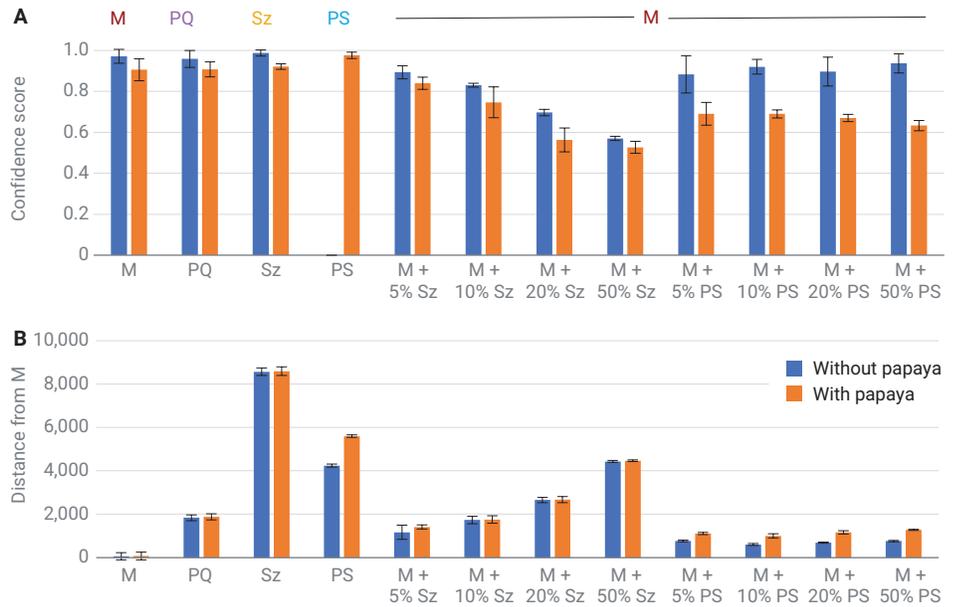
## B Compound Table: M

Mass	RT	Abundance	Flag	Profile
262.1440	20.864	-22.453	●	
257.2475	24.709	-22.453	●	
257.2265	20.137	-22.453	●	
256.0975	23.264	-22.453	●	
248.2012	33.760	3.587	○	
235.2058	24.665	0.626	○	
233.0812	23.220	-22.453	●	
218.2031	31.645	-22.453	●	
218.2028	32.301	-22.453	●	
215.1794	31.203	-1.694	○	
212.0704	23.602	-22.453	●	
212.0680	17.384	-22.453	●	
207.1744	22.996	-2.558	○	
205.0494	26.570	0.186	●	
204.1872	13.272	-0.530	●	
201.0547	27.270	-22.453	○	
201.0546	26.968	4.968	●	
192.1386	21.187	3.122	●	
192.1385	23.102	0.633	●	
191.1433	14.393	-22.453	●	
189.1638	14.411	2.273	●	
181.0495	17.385	-22.453	○	
179.1431	21.181	-22.453	●	
177.0546	30.896	-22.453	●	
176.0466	16.336	0.032	●	
175.1482	10.983	-0.394	○	
175.0388	24.548	0.344	○	
173.0596	20.551	-0.767	●	

**Figure 8.** Results visualization in Classifier that can be used for class prediction model evaluation or classification of unknown samples. (A) 3D PCA plot showing clustered samples of Malabar (M), Phu Quoc (PQ), Szechuan (Sz), and papaya seeds (PS). The arrow pointing to the Malabar sample adulterated with 5% papaya seeds. (B) Compound table shown for the same sample: Malabar adulterated with 5% papaya seeds. Compounds flagged green fall into the model range for Malabar.

The results of the PLSDA and SIMCA model comparison are shown in Figure 9. For PLSDA, when papaya was not included in the model, adulteration of black pepper with papaya seeds cannot be detected reliably (Figure 9A). With papaya included in the classification model, the confidence score for adulterated samples decreased, making it possible to better distinguish the adulterated samples.

The SIMCA results are reported as a distance from a certain sample class (Figure 9B displays the distance from Malabar). Compared to PLSDA, SIMCA was able to detect lower levels of adulteration with Szechuan as well as papaya seeds, even when papaya was not included in the classification model. When papaya was included, the distance from Malabar increased further, making this distinction easier.



**Figure 9.** Results of PLSDA versus SIMCA model comparison, both with and without papaya seeds extracts included in the model. (A) PLSDA, classification category for the confidence score given on the Y-axis is annotated above the bars. (B) SIMCA.

## Conclusion

In this application note, the adulteration of the black pepper Malabar was studied as an example to demonstrate a comprehensive approach that can be used for the routine analysis of food adulteration with GC/Q-TOF.

The model was able to better distinguish between pure and adulterated black pepper when adulterant was included in the model. In addition, the SIMCA model was able to detect as low as 5% adulteration with either papaya seeds or Szechuan pepper.

## References

1. Lafeuille, J.-L. *et al.* A Rapid Non-Targeted Method for Detecting the Adulteration of Black Pepper With a Broad Range of Endogenous and Exogenous Material at Economically Motivating Levels Using Micro-ATR-FT-MIR Imaging. *J. Agric. Food Chem.* **2020**, *68*(1), 390–401.
2. Medina, S. *et al.* Food Fingerprints – a Valuable Tool to Monitor Food Authenticity and Safety. *Food Chemistry* **2019**, *278*, 144–162.
3. Popping, B.; Everstine, K. The Food Fraud Combat Triumvirate: Vulnerability Management, Market Intelligence, and Detection Methods. *Food Quality Magazine* **2016**, *3*, 5–12.
4. Hong, E. *et al.* Modern Analytical Methods for the Detection of Food Fraud and Adulteration by Food Category. *J. Sci. Food Agric.* **2017**, *97*, 3877–3896.
5. Yannell, K. E.; Cuthbertson, D. Food Authenticity Testing with the Agilent 6546 LC/Q-TOF and MassHunter Classifier. *Agilent Technologies application note*, publication number 5994-0694EN, **2019**.
6. Ji, Y. *et al.* Chemical composition, sensory properties and application of Sichuan pepper (*Zanthoxylum* genus). *Food Science and Human Wellness* **2019**, *8*, 115–125.

[www.agilent.com/chem](http://www.agilent.com/chem)

DE.6819097222

This information is subject to change without notice.

© Agilent Technologies, Inc. 2020  
Printed in the USA, August 6, 2020  
5994-2264EN

