Classification of Olive Oils through the use of High Resolution GC/MS

Jennifer Gushue, Stephan Baumann and Sofia Aronova : Agilent Technologies, Santa Clara, CA

PITTCON 2013
TUES 1510-24

Agilent Technologies

Overview

A model was constructed to predicts whether an olive oil would pass the extra virgin sensory test. Running a GC/Q-TOF MS in both electron impact (EI) and positive chemical ionization (PCI) modes identified a large number of compounds in olive oil. Multivariate software was then used to perform statistical analysis and construct a classification model utilizing the presence of five specific compounds to accurately predict whether an olive oil would fail the sensory test.

Introduction

The demand for olive oil is growing rapidly in the United States. The US market is expected to surpass $1.8 billion by 2013. The International Olive Council (IOC) and USDA have established standards for the classification of extra virgin olive oil (EVOO), including a sensory test conducted by a tasting panel, as well as chemical tests. However, recent studies have stated that imported olive oils, which account for 99% of the EVOO on the US market, often fail the sensory test for EVOO classification. Multivariate software can be used to model olive oil classification and a GC/Q-TOF can be used to determine the identity of markers that correlate with a failed sensory test.

In this experiment we demonstrated that untargeted compound analysis done on an accurate mass GC/Q-TOF could be used with an EI spectral library to help identify markers that correlate with a failed sensory test. We also showed that orthogonal techniques such as positive chemical ionization could be used to collaborate the EI identifications.

Experimental

Sample Preparation Procedure

Ten olive oil samples were obtained from the UC Davis Olive Center. All of these samples had been subjected to IOC sensory test using a panel sanctioned by the IOC to determine if they passed or failed the criteria for EVOO. The samples were stored in the dark at room temperature. The samples were diluted 1:10 in cyclohexane and injected into the GC/Q-TOF with a 1:10 split. The results were evaluated using multivariate analysis and a model was developed that accurately predicted if an EVOO would fail the sensory test. Representative samples were then analyzed in PCI and EI Product Ion Scan modes to aid in Structural Elucidation.

Experimental

Figure 1: 7890 GC / 7200 Q-TOF instrumentation.

GC Run Conditions	
Column	DB-5MS, 30 meter, 0.25 mm ID, 0.25 µm film (P/N 122-5532)
Injection volume	1 µL
MMI Injector	50 °C for 0.01 min 600 °C/min to 300°C
Purge to split vent	60 mL/min at 1 minute
Oven Temperature Program	45 °C for 4.25 min 5 °C/min to 75 °C, 0 min hold 10 °C/min to 320 °C, 10 min hold
Carrier gas	Helium at 1.3mL/min constant flow
Transfer line temperature	290 °C
MS Conditions	
Ionization mode	EI, positive CI (20 % methane flow)
Source temperature	230°C
Quadrupole temperature	150°C
m/z scan	40 to 800 m/z
Spectra acquisition rate	5 Hz, collecting both in scan and profile modes

Table 1: A cold splitless injection was used to minimize thermal decomposition.

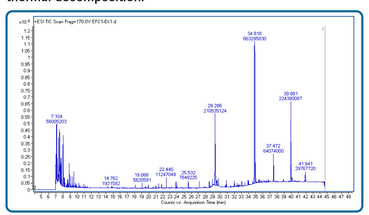

Figure 2: Typically, about 150 peaks are identified by chromatographic deconvolution with a relative area filter of 0.1% of largest peak.

Results and Discussion

Statistical Evaluation

Mass Profiler Professional (MPP) was used for statistical evaluation of the data. This software provides the possibility of using a guided workflow that helps the user evaluate the analytical data. The guided approach efficiently reduces the data set from 442 components to five statistically relevant compounds in a few steps. The steps were as follows:

1. Define the experiment type, workflow, and organism.
2. Select the data source as MassHunter Qual.
3. Import CEF files from MassHunter Qual.
4. Set abundance and model ion filters.
5. Define retention and match factor alignment parameters. (Total Compounds: 442)
6. Select internal, external, or no calibration.
7. Choose baseline correction.
8. Set condition filter flags. (Total Compounds: 91)
9. Filter based on Volcano Plot. This combines the Student's t-test for significance ($p < 0.05$) with fold change threshold = 4. (Total Compounds: 5)

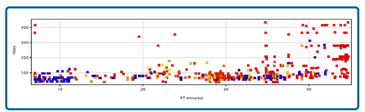

Figure 3: This mass vs. retention time plot shows that 442 unique compounds were distinguished by chromatographic deconvolution, most of which occurred only once or twice and were filtered out by MPP. The low frequency components are shown in red while the higher frequency components are shown in blue.

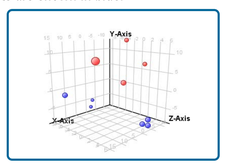

Figure 4: Principal Component Analysis (PCA) shows how data clusters. The samples that failed the sensory test are marked in red and the ones that passed are blue.

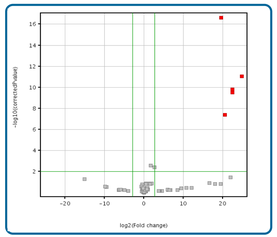

Figure 5: The Volcano Plot shows fold-change for each entity on the x-axis and significance on the y-axis. The five compounds in the upper right hand corner are accumulated in the samples that failed the sensory test.

Class Prediction

Through multivariate analysis we identified five compounds that correlated with a failed sensory test; the next step was to build a class prediction model that could be run independently of MPP. To do this we had to train the model with the data.

There are several classification models that can be applied however Partial Least Square (PLS) analysis is particularly adapted to situations where there are fewer observations (i.e. number of samples) than measured variables (e.g. detected entities, *m/z*). Its use has become very popular due to its ability to deal with many correlated and noisy variables. Therefore, PLSDA was used to construct the olive oil classification model.

Identifier	Training	Predicted(Training)	Confidence
CSCI-E1: Ig2	[F,Training]	[F,Training]	1.000
FSW2-E1: Ig2	[F,Training]	[F,Training]	1.000
ESCI-E1: Ig2	[P,Training]	[P,Training]	1.000
ESC2-E1: Ig2	[P,Training]	[P,Training]	1.000

Table 2: PLSD training set which contains representatives from each of the three clusters found in the PCA plot.

We can see in table three that a small data set such as this one is capable of identifying lower grade olive oil being misclassified as EVOO, a larger model could also determine if an olive oil has any of the characteristic compounds related to 'fusty', rancid, musty or vinegary flavor notes.

Results and Discussion

Identifier	Grade	Training	Predicted (Class Prediction)	Confidence
PAC1-E1: Ig2	F	None	[F,Training]	1.000
ESC2-E1: Ig2	P	Training	[P,Training]	1.000
ESCI-E1: Ig2	P	Training	[P,Training]	1.000
SAC1-E1: Ig2	F	None	[F,Training]	1.000
RFC2-E1: Ig2	P	None	[P,Training]	1.000
RSA2-E1: Ig2	P	None	[P,Training]	1.000
CSC1-E1: Ig2	F	Training	[F,Training]	1.000
RSA1-E1: Ig2	P	Training	[P,Training]	1.000
EFC1-E1: Ig2	P	None	[P,Training]	1.000
FSW2-E1: Ig2	F	Training	[F,Training]	1.000

Table 3: The model correctly predicted the pass or fail status of all samples, including those not used to construct the model. The samples that were not used for building the prediction model are listed with the Training variable set as 'None'.

Compound Identification/Structure Elucidation

While it is not necessary to know the identity of the compounds used in the classification model, identification could lead to an understanding of the mechanism by which those chemical components might adversely affect the sensory qualities of olive oil. The advantage of the 7200 GC/Q-TOF is that it can collect data in EI and CI, for both high resolution accurate mass MS and MS/MS. These orthogonal modes of operation aid confirmation. EI spectra allow library searching and provide fragmentation data, CI provides information about the empirical formula, and full spectrum MS/MS generates data for an accurate mass substructure search that can be applied to EI or CI generated ions.

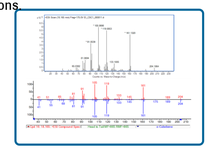

Figure 6: Commercial unit mass EI spectral libraries like Wiley and NIST can be searched using accurate mass EI GC/Q-TOF data to identify compounds.

As we see in Figure 7, the Molecular Structure Correlator (MSC) performs a substructure search of the ChemSpider database using MS/MS product ion data. MSC correlates the product ion results to all the possible structural isomers. Each individual fragment ion is ranked based on mass error corresponding to the proposed formula, along with a penalty based on how many bonds needed to be broken to generate that proposed formula. An isomer's individual compatibility score is a weighted average of the fragment ion scores, taking into account the intensity and the mass of each fragment ion.

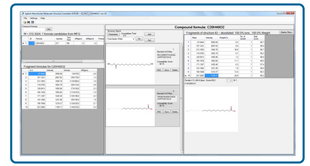

Figure 7: MSC performs a ChemSpider substructure search on the product ion fragments of ethyl octadecanoate's molecular ion peak and generates a compatibility score. Note that this substructure is equally compatible with eicosanoic acid as to ethyl octadecanoate. The EI fragmentation pattern is what allows us to distinguish between structural isomers.

Exact Mass

NIST ID	Formula	EI([M+]+ Measured	Mass Error (PPM)	PCI ([M+H]+ Measured	Mass Error (PPM)
n-Hexadecanoic acid	$C_{16}H_{32}O_2$	256.2395	4.7	257.2470	1.9
Octadecanoic acid, ethyl ester	$C_{20}H_{40}O_2$	312.3008	4.8	313.3091	3.2
Squalene	$C_{30}H_{50}$	410.39037	0.8	411.3987	0.5
α-Cubebene	$C_{15}H_{24}$	204.1883	4.9	205.1945	2.9

Table 4: PCI spectral data provided confirmatory accurate mass information for molecular ions of the four up-regulated compounds that had molecular ion peaks and were identified through EI library searching.

Conclusions

Using accurate mass EI and positive CI data generated by a GC/Q-TOF system, a model can be constructed that accurately predicts whether an olive oil will pass the EVOO sensory test. Although a study constructed using a very small sample set generates an incomplete model, it demonstrates the feasibility of the approach. A predictive model constructed using a significantly larger sample size would give olive oil producers a quick determination as to whether their oil would fail the sensory test and the defects present.

Structure elucidation of four of the model compounds was accomplished by EI library searching combined with Molecular Formula Generation, Isotopic Pattern Filtering, and Molecular Structure Correlation.