Metabolomic Profiling of Teas Using HPLC-QTOF/MS Coupled with Statistical Tools

ASMS 2010

Dr. Ying Wang, Agilent Technologies (China), Shanghai, China

Metabolomic Profiling of Teas Using HPLC-QTOF/MS Coupled with Statistical Tools

Introduction

Tea, as one of the most widely consumed and traded beverages in the world, has at least six varieties: white, yellow, green, oolong, black and pu'er, which are processed differently and the final content of each variety is therefore changed accordingly. Distinguishing and controlling the quality of teas with modern technologies has become of utmost importance. The methodology of Metabolomic profiling coupled with statistical analysis has been widely used in the area of food agriculture and manufacture, which involves finding the interesting metabolites with statistically significant variations in abundance within a set of samples.

We report here the results of an analysis of 13 tea samples of different varieties, places of origin and fermentation conditions. The 13 samples included green tea as Longjin from Zhejiang province harvested in 2008 and 2009, Laoshan tea from Shandong province in 2009; oolong tea as 5 Tieguanyins were semi-fermented in different levels, among which Dahongpao were fermented up to 70%, others were fermented from 20% to 50%, both Tieguanyins and black tea were from Fujian province and manufactured in 2009; Pu'er tea of both fermented and non-fermented were from Yunnan province, manufactured 3 years ago.

Experimental

Equipment

- Agilent 1200 Series Rapid Resolution LC system (RRLC): Agilent 1200 Series binary pump SL with degasser, Agilent 1200 Series high performance autosampler SL, Agilent 1200 Series thermostated column compartment
- Agilent 6530 Series Quadrupole Time-of-Flight mass spectrometer equipped with jet stream ion focus source, the internal reference compounds were introduced into the system at real time from a universal interface.
- Software: MassHunter Workstation for instrument control and data analysis of background subtraction and compound finding.
- Statistic analysis software: Mass Profiler Professional

Acquisition Method

- Solvent A: Water with 0.1% formic acid
- Solvent B: Methanol/solvent A (90/10, V/V)
- Column: Zorbax SB-Aq C18 2.1×100mm, 1.8μm particle size
- Flow rate: 0.3mL/min
- Gradient: 0-2min, 0%B; 2-20min, 0-40%B; 20-25min, 40-75%B; held for 5min, then stopped. Post time 10min
- Column temperature: 50℃
- QTOF MS: Negative polarity
  drying gas temperature 325℃, drying gas flow 7L/min,
  nebulizer 40psi
  sheath gas temperature 400℃, sheath gas flow 12L/min
  capillary voltage 3500V, nozzle voltage 0V, fragmentor 175V
  mass range m/z 50-1000 (MS), m/z 25-1000 (MSMS)
  scan rate: MS only mode, 2spectra/sec; Target MSMS mode, 5spectra/sec (MS), 2spectra/sec (MSMS)
  collision energy: 5V+3V/100Da
  reference mass on

The mass spectrometer was operated in MS only mode for the profiling, followed by Target MS/MS for the component founded by differential analysis processing.

Sample preparation

50mL of 80ºC water was poured over 3g of each sample, the first infusion was discarded after steeping for 3min, then another 50mL of 80ºC water was added into each residue. The tea soup was filtered after steeping for 10min. 5μL of filtrate was injected into the LC/QTOFMS system. At least 3 replicates were injected for each sample.

Results and Discussion

LCMS analysis

After optimization, the column packed with Zorbax SB-Aq particles was selected for the experiment due to its good performance to retain and separate the hydrophilic compounds in the tea soup.
During the MS data acquisition, reference ions were used for the real time calibration, thus maintaining mass accuracy with mass errors in MS and MS/MS modes below 2 ppm and 5 ppm, respectively. In order to obtain enough data points across the narrow peaks eluted from the column packed with STM packing material, the scan rate was set at 2spectra/sec.

Feature extraction

The compound features in 45 raw data files acquired with the mode of MS Only, including 39 sample files and 6 blank runs, were extracted using a proprietary algorithm called Molecular Feature Extraction (MFE) in MassHunter Qualitative Analysis software. Ion responses whose individual abundances rose and fell together over time, indicated that they were from the same compound, and these responses, including isotopes, dimers, different adduct ions (here are $[M-H]^-$, $[M+Cl]^-$ and $[M+NO_3]^-$) were grouped as one compound. A compound list was generated for each data file. Background ion responses, occurred in the blank runs, were filtered out of the samples data at same time. Finally, extracted features for each sample data file were exported as XML-based Compound Exchange Format (.cef) files for chemometric analysis.

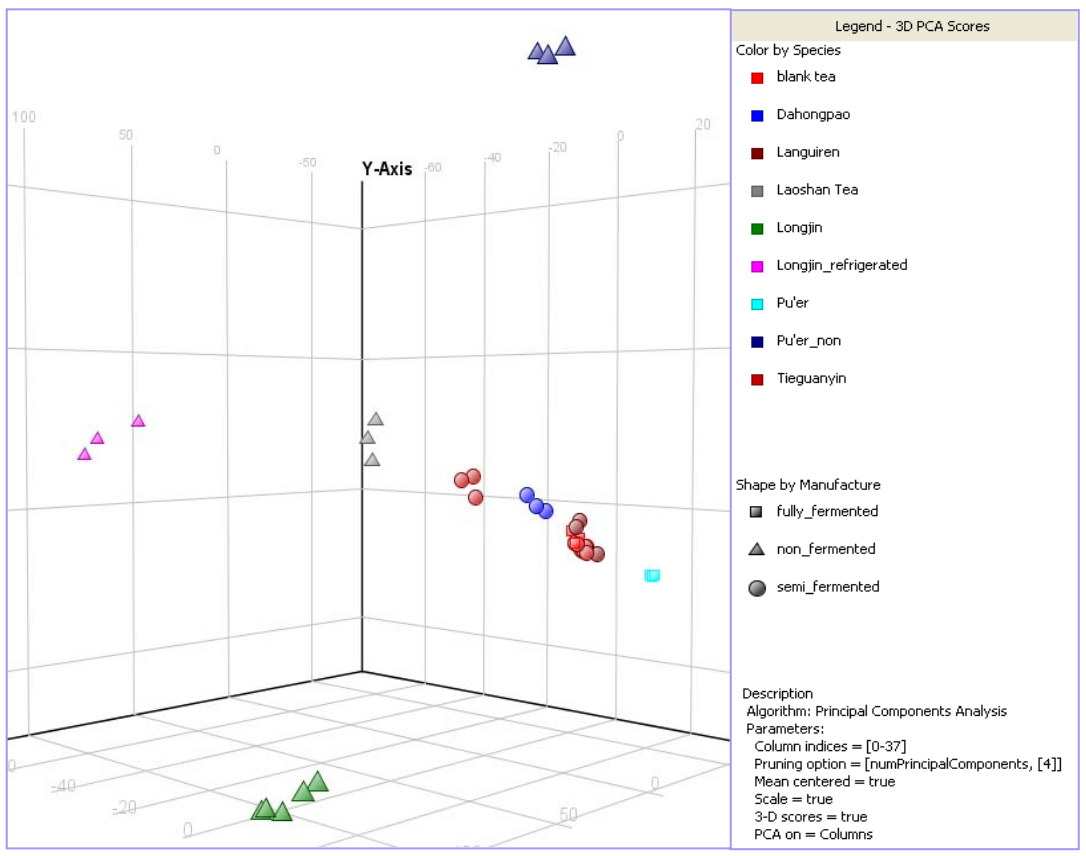


Figure 1. PCA score plot on Tea Species

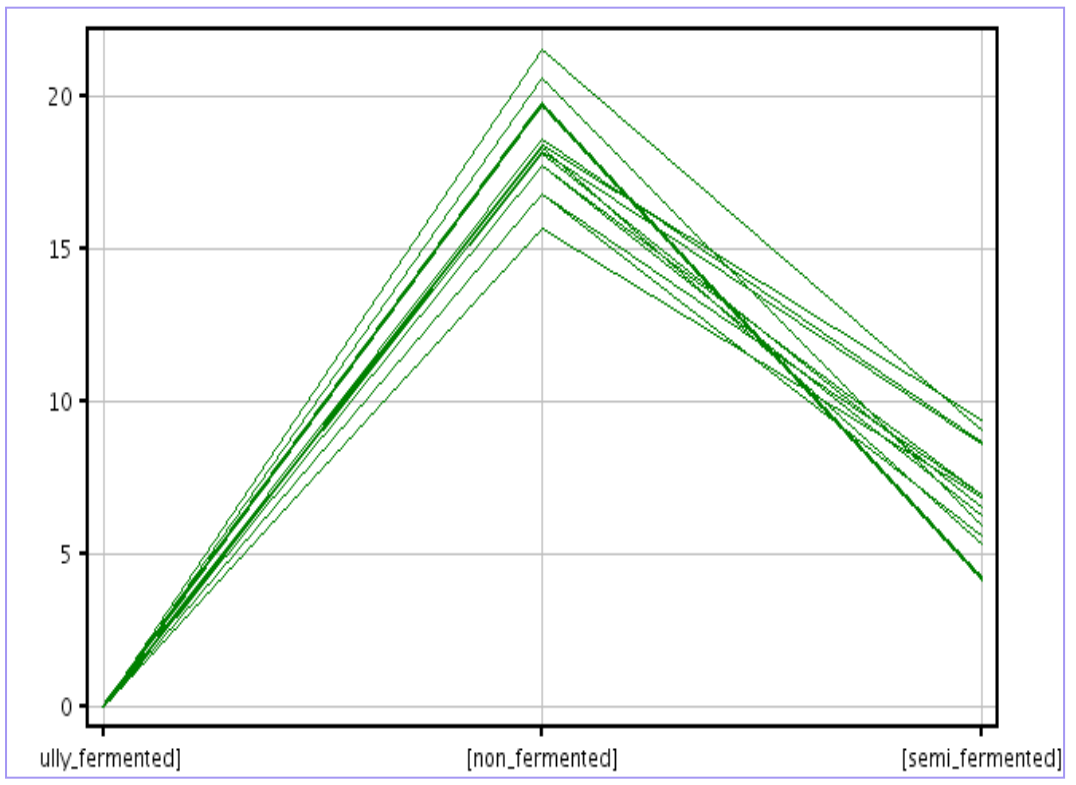
Statistic analysis

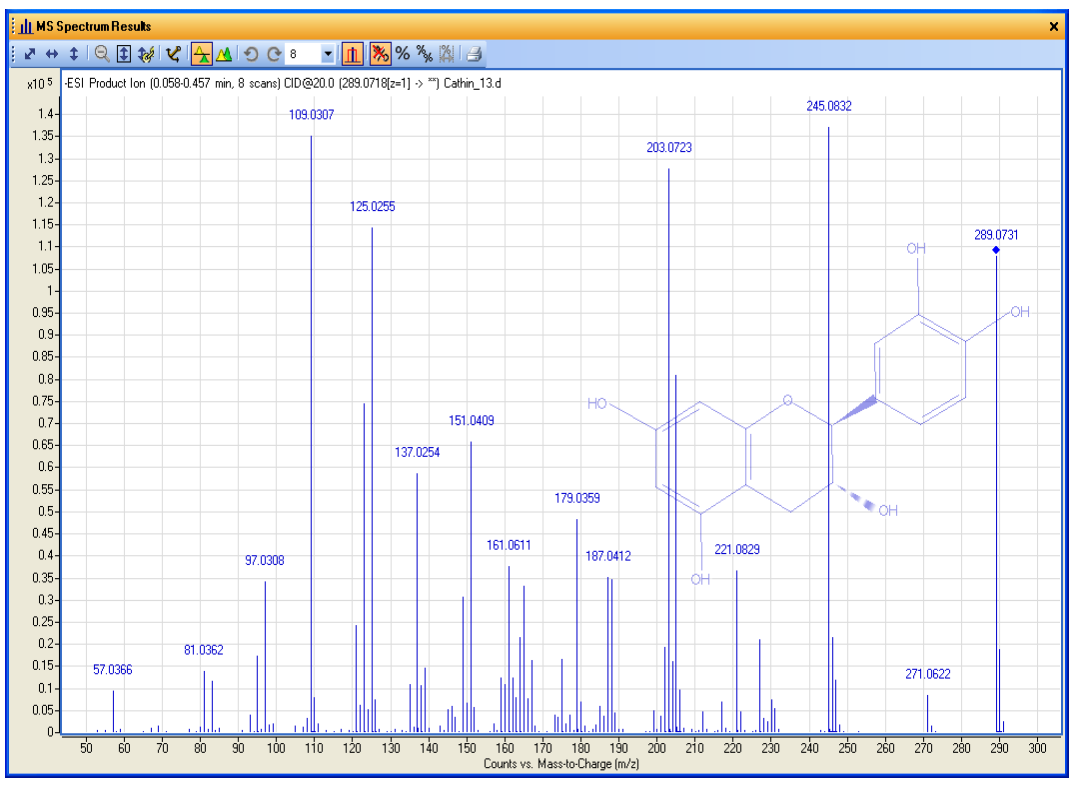
.cef files of all of 39 samples were imported into the statistic software, Mass Profiler Professional (MPP). The same compounds identified between each sample set were aligned for comparison. All data files were divided into groups, based on the parameters of species and/or manufacture processing. The features that appear in all of samples in at least one condition and with a coefficient of variation of less than 5% were retained for the statistic calculation.
Figure 1, the PCA score plot based on condition of species, shows that the Pu'er teas of both fermented and non-fermented are quite different from others, because the tea-leaf were collected from a special tea plant where the difference between the Pu'ers is due to manufacture processing. The refrigerated Longjin tea is distinguished from those stored in ambient. The difference with the Laoshan tea is due to the place of origin. the Blank tea is quite similar to some of Tieguanyins, but Dahongpao and one of Tieguanyin are distinguished from others due to the different fermentation level.
Table 1, A class prediction model was built via Partial Least Squares Discrimination (PLSD) to classify the candidate samples, the overall accuracy of prediction was over 95%. Data dispersion showed some overlap between the Dahongpao and Tieguanyin tea . The model can be used to predict the tea samples.
Figure 2, one of the cluster based on the K-means analysis, indicates a trend of the concentration of some compounds which decreased when the tea was fermented further. These compounds are mainly catechins, including catechin gallate, gallocatechin and gallocatechin gallate, et al. Catechins were highest in concentration in green teas that are non-fermented, while in the fully-fermented teas, such as black tea and Pu'er had substantially fewer due to their oxidative preparation. All the compounds included in the cluster were exported in a inclusion list for acquisition of TargetMSMS.
Figure 3 shows the accurate mass library search result for one of MS^2 spectrum, catechin was matched at score of 86.6.

Metabolomic Profiling of Teas Using HPLC-QTOF/MS Coupled with Statistical Tools

Results and Discussion

	[black tea] (Predicted)	[Dahongpao] (Predicted)	[Laoshan] (Predicted)	[Longjin] (Predicted)	[Longjin_refrigerated] (Predicted)	[Pu'er] (Predicted)	[Pu'er_non] (Predicted)	[Tieguanyin] (Predicted)	Accuracy
(True)[black tea]	3	0	0	0	0	0	0	0	100
(True)[Dahongpao]	0	2	0	0	0	0	0	1	66.67
(True)[Laoshan]	0	0	3	0	0	0	0	0	100
(True)[Longjin]	0	0	0	5	0	0	0	0	100
(True)[Longjin_refrigerated]	0	0	0	0	3	0	0	0	100
(True)[Pu'er]	0	0	0	0	0	3	0	0	100
(True)[Pu'er_non]	0	0	0	0	0	0	3	0	100
(True)[Tieguanyin]	0	0	0	0	0	0	0	12	100
Overall Accuracy									95.8





Conclusions

- The methodology of metabolomic profiling coupled with statistic analysis can be used to determine differential analysis among teas
- Data reduction via MFE has been achieved effectively; with filtering by frequency and variability, the data quality was improved further.
- Both PCA on condition and class prediction via PLSD can be used to classify the samples with higher accuracy.
- Clustering analysis showed the trend of compounds under different conditions.
- Accurate MS^2 and library searching were used to identify the interesting metabolites with statistically significant variations in abundance variation across samples.