

# Origin Identification and Quantitative Analysis of Honey by Nuclear Magnetic Resonance and Chemometric Techniques

Xin Zheng<sup>1</sup> · Yanrong Zhao<sup>1</sup> · Huifeng Wu<sup>2</sup> · Jiyang Dong<sup>1</sup> · Jianghua Feng<sup>1</sup>

Received: 2 July 2015 / Accepted: 21 September 2015 / Published online: 3 October 2015  
© Springer Science+Business Media New York 2015

**Abstract** The combination of <sup>1</sup>H NMR spectroscopy and multivariate statistical analysis has become a promising method for the discrimination of food origins. In this paper, this method has been successfully employed to analyze 70 Chinese honey samples from eight botanic origins, three geographical origins, and five production dates. Thirty-three components in honey samples were detected and identified from their <sup>1</sup>H NMR spectra, and 20 of them were accurately quantified by comparing their integral area with that of internal standards with relaxation time correction. Nontargeted principal component analysis (PCA) has been applied to distinguish the honeys from different botanical and geographical origins. The variations of components in the honeys, including saccharides and all kind of amino and organic carboxylic acids, confirmed their clustering according to their origins in PCA scores plots. Orthogonal partial least squares discriminant analysis (OPLS-DA) based on the NMR data for the different pairwise honey samples allows to identify the compositional variations contributed to geographical discrimination and storage time. Hence, NMR spectroscopy coupled with chemometric techniques offers an efficient tool for quality control of honey, and it could further serve to the classification, qualitative and quantitative control of other foods.

**Keywords** Honey · Quality control · Nuclear magnetic resonance · Chemometric analysis · Origin identification

## Introduction

Honey is a natural sweet product collected by honeybees from the nectar of flowers (Aparna and Rajalakshmi 1999), and it is very popular all over the world for its nutrition value. Honeys from different origins have the similar relative sweetness to granulated sugar as they have a large amount of fructose and glucose (Cavia et al. 2002). However, there are also some minor components like organic acids, amino acids, and polysaccharides, which make a certain contribution to the organoleptic character of honeys (Janiszewska et al. 2012). Each honey from different botanical and geographical origins gives a unique and individual organoleptic character on the basis of its chemistry, amount, and combination of the various components (Juan-Borras et al. 2014; Karabagias et al. 2014).

Over the years, China has been the largest exporter of honey in the world with around 100,000 t per year which takes up a quarter of total exports. Because of the vast territory, complicated terrain, and climate diversity, China has various honey origins. Differences in quality and price exist among the different parts of China. The taste, aroma, and combination of honeys are tightly associated with their botanical and geographical origins, as well as seasonal conditions. The control and characterization of quality and origins of honeys are of great importance and interest in apiculture.

Traditionally, pollen analysis (melissopalynological analysis) and the evaluation of organoleptic characteristics are the most important methods to determine the botanical and geographical origins with great accuracy. However, these traditional methods are often time-consuming and operator-dependent. (Conte et al. 1998; Yang et al. 2012).

---

✉ Jianghua Feng  
jianghua.feng@xmu.edu.cn

<sup>1</sup> Fujian Provincial Key Laboratory of Plasma and Magnetic Resonance, Department of Electronic Science, Xiamen University, Xiamen 361005, China

<sup>2</sup> Key Laboratory of Coastal Zone Environment Processes, CAS, Shandong Provincial Key Laboratory of Coastal Zone Environment Processes, Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai 264003, China

The liquid-state  $^1\text{H}$  high-resolution NMR spectroscopy has been widely used in food analysis in particular for quality control of food with general acceptance due to its nondestructive characteristics, simultaneous detection of various components, and high reproducibility (Mannina et al. 2012). The coupling of chemometric analysis with  $^1\text{H}$  NMR spectroscopy gives much richer and interesting results for food analysis and provides in-depth understanding for the detailed information (Monakhova et al. 2013). Several papers have demonstrated that combination of NMR spectroscopy with multivariate statistical analysis is indeed a promising method to determine honey origin, and also a high-efficient quantitative tool (Boffo et al. 2012; Consonni and Cagliani 2008; Donarski et al. 2010; Ohmenhaeuser et al. 2013).

In this study, various honeys from different botanical or geographical origins and production dates were investigated by  $^1\text{H}$  NMR spectroscopy combined with multivariate statistical analysis. The principal component analysis (PCA) was used to classify the honey samples according to different botanical and geographical origins. Furthermore, the compositional variations of honeys with geographical origins and storage time were identified by the orthogonal partial least squares discriminant analysis (OPLS-DA). The potential of high-resolution  $^1\text{H}$  NMR was also explored to quantify components in honeys and to confirm their botanical origin statistically.

## Materials and Methods

### Honey Samples and Preparation

A total of 70 Chinese honey samples obtained from different botanic and geographical origins are provided by the local supermarkets. Among the honey samples, 25 were multi-floral (Floral) samples in 5 various production dates over a 14-month period from a same brand, 15 were *Flos sophorae* (FS) samples from three geographical origins (Fujian, Zhejiang, and Shanghai), and the rest were from six honey botanical origins including Chinese jujube (CJ), *Lonicera japonica* (LJ), *Flos chrysanthemi indicis* (FCI), *Schefflera actinophylla* (SA), *Dimocarpus longan* (DL), and *Eriobotrya japonica* (EJ) from the same producer. All samples were collected in the years from 2012 to 2013.

All above honeys were kept under the condition of 17–25 °C in case of crystallization. They were stored in the dark to keep out of the sun so that no significant change would occur (Venir et al. 2010). Potential of hydrogen (pH) was measured for each honey sample on a Mettler Toledo pH meter. As a result, the measured pH ranged from 3.4 to 5.2. Honey samples were prepared by dissolving 100 mg of honey in 500  $\mu\text{L}$  of 600 mM of deuterated phosphate buffer (pH 4.92) containing 0.05 % of sodium 3-(trimethylsilyl)-2,2,3,3- $^2\text{H}_4$  propionate (TSP). The honey–buffer mixture was

left to stand for 5 min at room temperature and then centrifuged at 10,000 rpm for 10 min. The supernatant solution (500  $\mu\text{L}$ ) was then transferred into a 5-mm NMR tube for measurement.

### $^1\text{H}$ NMR Spectroscopy and Data Processing

$^1\text{H}$  NMR spectrum was recorded for each honey sample on a 500 MHz Varian NMR spectrometer (Agilent Technologies, Santa Clara, CA, USA) equipped with a 5-mm indirect detection probe operating at 499.74 MHz and under the room temperature (295 K) without sample rotation. All the  $^1\text{H}$  NMR spectra were acquired using water suppressed NOESYPR1D pulse sequence (recycle delay- $90^\circ$ - $t_1$ - $90^\circ$ - $t_m$ - $90^\circ$ -acquisition). Thirty-two K data points of the NMR signal were acquired with a 6000-Hz spectral width, an acquisition time of 2.5 s, and a recycle delay time of 3.0 s using 32 scans and 4 prior dummy scans. The water suppression was achieved with the selective irradiation on the water peak during a fixed interval  $t_1$  of 4  $\mu\text{s}$  and the mixing time  $t_m$  of 100 ms.

All free induction decays (FIDs) were weighted by an exponential function with a 1.0-Hz line-broadening factor before Fourier transformation. All spectra were manually phased and baseline-corrected by using the Bernstein polynomial fit method with a respective polynomial order for the purpose of best baseline correction effect by using MestReNova (Version 7.0.14).  $^1\text{H}$  NMR spectra of honey samples were referenced to TSP at 0.0 ppm. The regions of water resonance ( $\delta$  4.76–4.86) were removed to eliminate baseline effects of imperfect water signal. The spectral region of  $\delta$  0.50–9.50 was automatically integrated with the integral width of 0.004 ppm and normalized to TSP area ( $\delta$  -0.10 to 0.10). Then, the NMR data were converted into ASCII format for further multivariate statistical analysis.

### Quantitative Analyses of Components in Honey Samples

For the quantification of the components, an integral width of 0.002 ppm was used in order to give more accurate quantitative results. Spectral region of  $\delta$  -0.10 to 9.50 was automatically integrated and normalized to TSP (internal standard) area ( $\delta$  -0.10 to 0.10). The integration of the selected signals and the comparison with TSP area permitted the quantitative determination of the corresponding compounds in the honey samples. As the integral area of corresponding signal is proportional to proton concentration of a compound, its quantification from an NMR spectrum requires knowledge of the number of protons contributing to the signals of this compound and the internal standard.

For accurate quantitation purposes, spin–lattice relaxation time ( $T_1$ ) of each resonance in the NMR spectra was measured using the classical inversion recovery sequence with water suppression during both recycle delay (RD) and relaxation

delay  $\tau$ . Sixteen  $\tau$  values of 0.05–16 s were employed with the total repeat time of 22.9 s (including both recycle delay and acquisition time) and 32 K data points of the NMR signal were acquired using 32 scans. The areas from the top of peaks were used to calculate the  $T_1$  values by fitting experimental data into a single exponential relaxation process.

Because the integral area of a compound in an NMR spectrum is proportional to its proton concentration in the corresponding solution, the concentration of a compound and the concentration of internal standard (here is TSP) are then related by

$$\frac{C_x}{MW_x/N_x} \times \frac{1}{A_x} = \frac{C_{TSP}}{MW_{TSP}/N_{TSP}} \times \frac{1}{A_{TSP}} \quad (1)$$

in a completely relaxed  $^1\text{H}$  NMR spectrum.

Thus, the concentration of any a component  $x$  in honey samples could be calculated by an equation as below:

$$C_x = \frac{A_x/N_x}{A_{TSP}/N_{TSP}} \times \frac{MW_x}{MW_{TSP}} \times C_{TSP} \quad (2)$$

where  $C_x$ ,  $A_x$ ,  $N_x$ , and  $MW_x$  are the concentration in milligrams per gram, the integral area in a completely relaxed  $^1\text{H}$  NMR spectrum, the number of hydrogens contributing to the signals, and molecular weight of any a component  $x$  in honey samples, respectively, and  $C_{TSP}$ ,  $A_{TSP}$ ,  $N_{TSP}$ , and  $MW_{TSP}$  are the concentration, the integral area in a completely relaxed  $^1\text{H}$  NMR spectrum, the number of hydrogens contributing to the signals, and molecular weight of TSP, respectively.

### Chemometric analyses

Multivariate data analysis was performed on SIMCA-P+ 14 software (Umetrics, Umea, Sweden). Principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA) were performed with “mean centering” (Ctr) and “Pareto scaling” (Pareto) (van den Berg et al. 2006), respectively. PCA was applied for visualization and classifying the samples from different botanical and geographical origins and production dates. OPLS-DA was performed to improve the separation between pairwise groups and production date groups respectively. OPLS is an extension to the supervised PLS regression method (Trygg et al. 2007). It can fulfill the task of discrimination between groups and interpretation of group differences in meaningful ways.  $R^2$  and  $Q^2$  values were used to assess the amount of variation represented by the PCs and robustness of the model, respectively. All OPLS-DA models were cross-validated by permutation tests (permutation numbers=200) and CV-ANOVA. The coefficients plots were then back-transformed from the coefficients incorporating the weight of the variables with MATLAB

scripts with some in-house modifications. In our study, correlation coefficient of  $|r| > 0.811$  (degree of freedom=4) was used as the cutoff value for statistical significance based on the discrimination significance at the level of  $p=0.05$ . The color-coded plots can enhance interpretability of the models: The intensity corresponds to the mean-centered model (variance) and the color-scale derives from the Pareto-scaled model (correlation) in the coefficient plots.

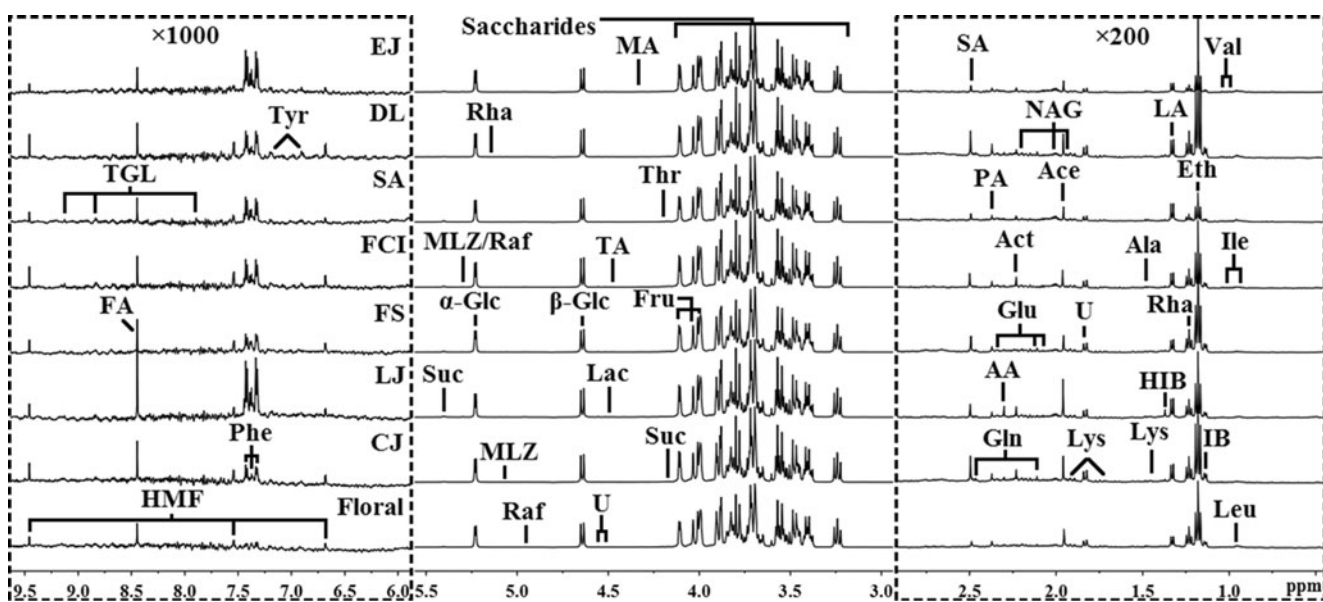
## Results and Discussion

### $^1\text{H}$ NMR Spectral Profiles of Honeys from Different Botanical Origins

$^1\text{H}$  NMR spectrum can provide global information about complex samples and maintain the natural ratio of the components (Cazor et al. 2006). Figure 1 represents the typical  $^1\text{H}$  NMR spectra of the honey samples in water solution from different botanical origins including multi-floral, Chinese jujube, *L. japonica*, *F. sophorae*, *F. chrysanthemi indici*, *S. actinophylla*, *D. longan*, and *E. japonica*. These spectra show the dominant resonances of main components content. Resonance assignments were performed according to the existing literature (Boffo et al. 2012; Consonni et al. 2012; Ohmenhaeuser et al. 2013) and confirmed by public Human Metabolome Database (HMDB) and in-house NMR database.

It can be seen that the NMR spectra are dominated with sugar regions ( $\delta$  3.0–5.5), in which the very intensive resonances are from the major monosaccharides ( $\alpha$ - and  $\beta$ -glucose and fructose). These signals are mainly equal to all honey samples in our experiment with very little variations in the intensity (Mazzoni et al. 1997). Some other saccharides including mono- (rhamnose), di- (lactose and sucrose), and tri-saccharides (melezitose and raffinose), with very low proton signal intensity, are also detected and identified in this region.

Other less intensive resonances of the minor components can be observed and recognized in the aliphatic regions ( $\delta$  0.5–3.0) and aromatic regions ( $\delta$  6.0–9.5) in the NMR spectra. These resonances from organic carboxylic acids (acetic, acetoacetic, N-acetylglutamic, formic,  $\alpha$ -hydroxyisobutyric, isobutyric, lactic, malic, pyruvic, succinic and tartaric acids), amino acids (alanine, glutamic acid, glutamine, isoleucine, leucine, lysine, phenylalanine, threonine, tyrosine, and valine), and some other compounds (acetone, ethanol, hydroxymethylfurfural, and trigonelline) show some obvious variations in the intensity among the honey samples with the different botanical origins. For example, a big variation in the intensity is displayed in the resonances from ethanol, formic acid, succinic acid, and phenylalanine among the eight botanical honey samples in Fig. 1. However, visual comparison of the spectra is impossible for complete analysis of the large number of samples due to the



**Fig. 1** A 500-MHz  $^1\text{H}$  NMR of honeys from various botanical origins. Floral (multi-floral); CJ (Chinese jujube); LJ (*Lonicera japonica*); FS (*Flos Sophorae*); FCI (*Flos chrysanthemi indicici*); SA (Schefflera actinophylla); DL (*Dimocarpus longan*); EJ (*Eriobotrya japonica*). The regions of  $\delta$  0.5–3.0 and  $\delta$  6.0–9.5 were respectively vertically magnified 200 and 1000 times compared with corresponding region of  $\delta$  3.0–5.5 for the purpose of clarity. Keys: Ace, acetic acid; Act, acetone; Ala, alanine; AA, acetoacetic acid; Eth, ethanol; FA, formic acid; Fru, fructose; Gln,

glutamine; Glu, glutamic acid; HIB,  $\alpha$ -hydroxyisobutyric acid; HMF, hydroxymethylfurfural; Ile, isoleucine; IB, isobutyric acid; LA, lactic acid; Lac, lactose; Leu, leucine; Lys, lysine; MA, malic acid; MLZ, melezitose; NAG, N-acetyl glycoprotein signals; PA, pyruvic acid; Phe, phenylalanine; Rha, rhamnose; Raf, raffinose; SA, succinic acid; Suc, sucrose; TA, tartaric acid; TGL, trigonelline; Thr, threonine; Tyr, tyrosine; Val, valine;  $\alpha$ -Glc,  $\alpha$ -glucose;  $\beta$ -Glc,  $\beta$ -glucose

complexity of honey composition. Multivariate data analysis was employed to analyze and identify such variations.

### Botanical Origins Discrimination of Honey with PCA and Quantitative Analysis

PCA was performed on the 40 honey samples from eight botanical origins. The first three PCs explaining 94.6 % of the total variance lead to a clear differentiation among the honey samples in the 3D PCA scores plot (Fig. 2a). The scores plot suggests that the honey samples can be separated into three groups, SA, Floral, and FS, and other five classes of honeys as CJ, LJ, DL, FCI, and EJ. Furthermore, clusters from those five botanical origins (CJ, LJ, DL, FCI, and EJ) were further separated from each other in the 3D PCA sub-scores plot (Fig. 2b), among which PCA scores of CJ, LJ, and DL were mixed with each other, while minor differences from FCI and EJ can be traced. It could be noticed that PCA score of EJ is close to that of SA. Although the corresponding loading plots (data not shown) allow to specify the variables (chemical shifts, i.e., components) which are responsible for the observed clustering for the botanical origins, only the factors from the primary components such as fructose and glucose can be obviously observed and identified, while the factors from minor components will be easily ignored though they also play a certain role. Therefore, it would be necessary to establish a quantitative NMR method for the detectable

components in the honey for the identification of characterized compounds of different botanical origins.

Basically, the concentration of any component in honey samples could be calculated according to Eq. 2 in the section “Materials and Methods.” However, in our experiments, NMR spectra were acquired in an incomplete relaxation; thus, the integral area for a given proton resonance should be corrected with relaxation time in order to get an accurate quantification as the following equation:

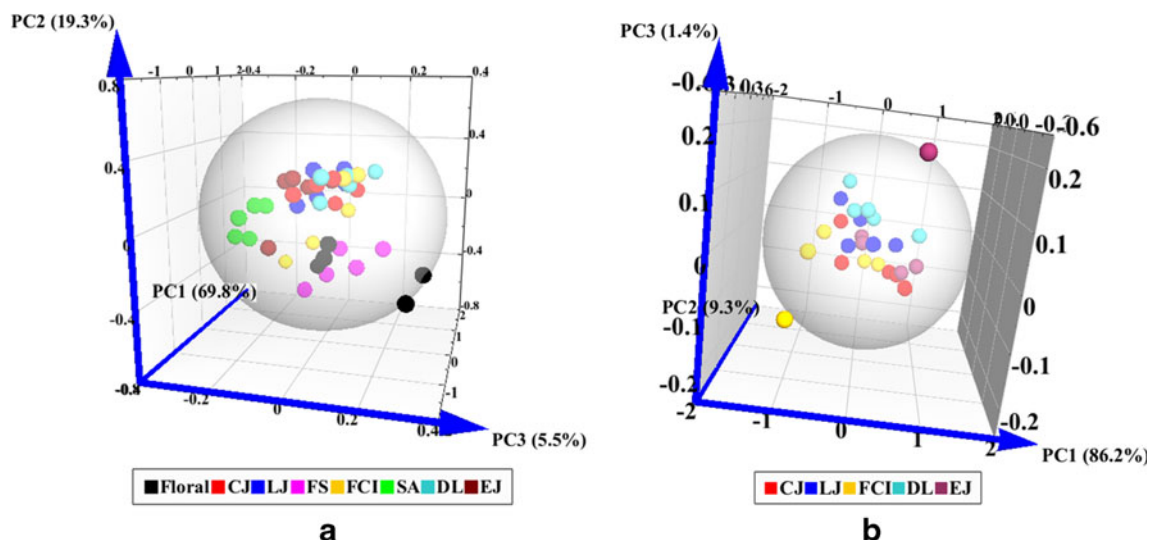
$$A'_x = A_x \left[ 1 - \exp\left(-\frac{t}{T_x}\right) \right] \quad (3)$$

where  $A'_x$  and  $A_x$  are integral areas of a given proton  $x$  with a spin–lattice relaxation time of  $T_x$  in a given spectrum and the completely relaxed spectrum, respectively, and  $t$  is the real relaxation time in our experiment (i.e., recycle delay time plus acquisition time).

With the known concentration of TSP, the concentration of a given component  $x$  could be derived from the following relationship:

$$C_x = \frac{A'_x}{A_{TSP}} \times \frac{\left[ 1 - \exp\left(-\frac{t}{T_x}\right) \right]}{\left[ 1 - \exp\left(-\frac{t}{T_{TSP}}\right) \right]} \times \frac{N_{TSP}}{N_x} \times \frac{MW_x}{MW_{TSP}} \times C_{TSP} \quad (4)$$

where  $A_{TSP}$  and  $A'_{TSP}$  are the integral area for TSP in the completely and incompletely relaxed spectra, respectively,



**Fig. 2** 3D PCA scores plots derived from the  $^1\text{H}$  NMR spectra of different botanical resources of honeys. Floral, multi-floral; CJ, Chinese jujube; LJ, *Lonicera japonica*; FS, *Flos sophorae*; FCI, *Flos chrysanthemi indicii*; SA, *Schefflera actinophylla*; DL, *Dimocarpus longan*; EJ, *Eriobotrya japonica*

and  $T_{TSP}$  is the spin–lattice relaxation time for methyl protons of TSP. In our study, thirty-three compounds were detected and identified from NMR spectra for most of the honeys. Basically, a direct quantification of any component may be obtained by comparing all of its integration areas with that of TSP with relaxation time correction as described above (Eq. 4). However, very low signal/noise (S/N) ratios of some peaks will forbid the accurate quantification of their corresponding components; therefore,  $S/N=10$  was defined as the cutoff level of quantification in our study. In addition, it is almost impossible to accurately quantify for some compounds due to a number of appreciably overlapped signals with each other in our case, especially in the sugar region ( $\delta$  3.0–5.5). Thus, it is valid to try to select the characteristic and well-resolved resonances corresponding to each component in order to get the exact quantitative results. Finally, 20 components in these eight botanical-origin honey samples were simultaneously quantified, and their signals information selected for the quantification and the quantitative results were tabulated in Table 1. It is especially remarkable that the number of hydrogens contributing to the signals is complicated to count in a few cases. For example, the chosen peaks ( $\delta$  4.11) from fructose for quantification correspond to two hydrogens; however, the two hydrogens also contribute to other signals ( $\delta$  3.84 and  $\delta$  3.90). Hence, the corresponding integral areas were calculated for each signals of fructose according to NMR spectra of the standard solution. As a result, the signal at  $\delta$  4.11 corresponds to 0.6 hydrogen contribution (Table 1).

According to the comparison of quantitative results (Table 1) and PCA analysis (Fig. 2), it is obvious that the clustering of different botanic-origin honey in the PCA scores plots was a direct consequence of the variations in their chemical composition. Not only the concentrations of primary

components but also those of minor components play important roles in the differentiation of honey. The higher amounts of sucrose and some amino acids such as lysine and glutamine and the lower amounts of fructose,  $\alpha$ - and  $\beta$ -glucose, isobutyric acid (IB), and ethanol in SA honey than in the others led to a separate place of SA honey from others in PCA scores plot (Fig. 2a). The grouping of CJ, LJ, DL, FCI, and EJ honeys in PCA scores plot (Fig. 2a) is in agreement with their similar compositions (Table 1). In particular, the visual approaching of EJ honey to SA honey was related to their lower quantity of fructose. A separation of FCI honey from the sub-cluster of CJ, LJ, and DL honeys may be due to the higher concentrations of fructose and glucose and the lower  $\alpha$ -hydroxyisobutyrate (HIB) in FCI honey. The higher quantity of glucose and the medium level of lysine induced the separation of Floral and FS honeys from others.

In general, the clustering of different botanic-origin honeys in PCA scores plot suggested their variations in chemical composition, and their compositional differences confirmed the classification of honeys in PCA scores plot.

### Compositional Variations of Honey from Different Geographical Origins

Geographical origin also plays a key role on the composition of honey. PCA was performed on the FS honey samples from Fujian, Shanghai, and Zhejiang in order to understand the geographical factors. PCA scores plot (Fig. 3a) displayed a clear sample separation according to their geographical origin, which revealed some subtle compositional differences in honey with same botanic origin but different geographical origin.

To get an insight into the types of components responsible for the separation between different geographical-origin

**Table 1** The componential quantification in honeys from different botanical origins (values are given in milligrams per gram of honey)

Component and characteristic signals	Honey sample							
	Floral	CJ	LJ	FS	FCI	SA	DL	EJ
<b>Saccharides</b>								
Rhamnose: 1.23(m, 3H) <sup>a</sup> , 5.10(d, 0.5H)	0.19±0.04 <sup>b</sup>	0.19±0.06	0.22±0.03	0.16±0.03	0.18±0.03	0.18±0.04	0.17±0.002	0.14±0.001
Fructose: 4.11(m, 0.6H)	417.1±16.0	429.7±5.2	437.7±2.2	416.1±5.6	444.6±4.8	390.9±3.6	422.9±3.2	393.1±4.0
Sucrose: 4.21(d, 1H), 5.40(d, 1H)	11.3±1.0	11.0±0.8	13.6±0.4	11.1±0.5	13.0±0.4	15.3±0.5	6.7±0.3	9.6±0.3
Lactose: 4.47(d, 1H)	3.8±0.3	3.2±0.7	3.4±0.2	3.4±0.4	3.6±0.2	2.7±0.1	2.8±0.2	2.6±0.1
β-Glucose: 4.64(d, 1H)	190.5±12.4	174.8±2.9	172.8±1.8	178.5±5.9	186.9±2.7	131.1±5.2	171.4±5.5	180.9±3.1
Melezitose: 5.06(d, 1H), 5.49(d, 1H)	1.8±0.1	1.7±0.2	1.7±0.2	1.5±0.2	1.8±0.2	1.6±0.1	1.2±0.1	1.4±0.2
α-Glucose: 5.23(d, 1H)	164.9±11.5	156.6±3.0	154.2±1.4	164.9±4.2	165.6±3.5	131.8±3.02	154.5±1.3	165.8±3.7
<b>Amino acids</b>								
Lysine: 1.73(m, 1H)	0.1±0.06	nd	nd	0.08±0.04	nd	0.1±0.1	nd	nd
Glutamine: 2.14(m, 2H)	0.06±0.02	0.06±0.07	nd	0.08±0.07	nd	0.09±0.04	0.05±0.06	0.06±0.04
Glutamic acid: 2.34(m, 2H)	0.06±0.02	0.07±0.04	0.05±0.03	0.1±0.08	0.1±0.07	0.08±0.04	0.07±0.07	0.07±0.08
Threonine: 4.24(m, 1H)	3.8±0.2	3.2±0.5	3.7±0.2	3.4±0.2	3.9±0.2	3.5±0.2	2.6±0.2	2.9±0.06
Tyrosine: 6.90(d, 2H), 7.19(d, 2H)	nd	0.3±0.2	0.3±0.04	0.1±0.02	0.2±0.1	0.2±0.1	0.3±0.1	0.2±0.1
Phenylalanine: 7.33(d, 2H), 7.37(t, 1H), 7.42(dd, 2H)	nd	0.1±0.008	0.2±0.02	0.1±0.03	0.2±0.03	0.1±0.009	0.1±0.04	0.2±0.006
<b>Organic acids</b>								
Isobutyric acid: 1.14(d, 6H)	0.03±0.001	0.04±0.003	nd	0.04±0.003	nd	nd	0.03±0.002	nd
Lactic acid: 1.33(d, 3H)	0.08±0.007	0.09±0.009	0.09±0.006	0.09±0.003	0.09±0.02	0.09±0.01	0.07±0.008	0.05±0.02
Malic acid: 4.31(dd, 1H)	4.3±0.3	3.8±0.7	5.2±0.3	4.0±0.3	4.3±0.2	3.9±0.2	3.2±0.2	3.5±0.08
Tartaric acid: 4.41(s, 2H)	0.4±0.04	0.4±0.1	0.5±0.03	0.4±0.05	0.4±0.03	0.4±0.03	0.2±0.03	0.2±0.02
Formic acid: 8.45(s, 1H)	nd	0.03±0.004	0.06±0.005	0.03±0.005	nd	nd	nd	nd
<b>Other compounds</b>								
Ethanol: 1.18(t, 3H)	0.2±0.01	0.3±0.01	0.2±0.002	0.3±0.01	0.2±0.01	0.06±0.006	0.3±0.005	0.07±0.009
Hydroxymethylfurfural: 6.68(d, 1H), 7.54(d, 1H), 9.46(s, 1H)	nd	nd	nd	nd	0.03±0.004	nd	0.03±0.002	nd

<sup>a</sup> Chemical shift, multiplicity, and the corresponding proton number. Multiplicity: s, singlet; d, doublet; t, triplet; dd, doublet of doublets; m, multiplet

<sup>b</sup> The values are shown as mean±standard deviation, which were obtained from five samples in each honey group

<sup>c</sup> nd: not detectable or below NMR detection limit

honeys, the scores and loadings plots with correlation coefficients were obtained by using orthogonal partial least squares discriminant analysis (OPLS-DA). Clear differentiation of *F. sophorae* honeys from different origin was achieved in the geographical models (Fig. 4a, b). An overall quality of fit ( $R^2Y$ ) of 0.991 and 0.997 and an overall cross-validation coefficient ( $Q^2$ ) of 0.851 and 0.943 were obtained for the pairwise group from Fujian and Shanghai and Fujian and Zhejiang, respectively.

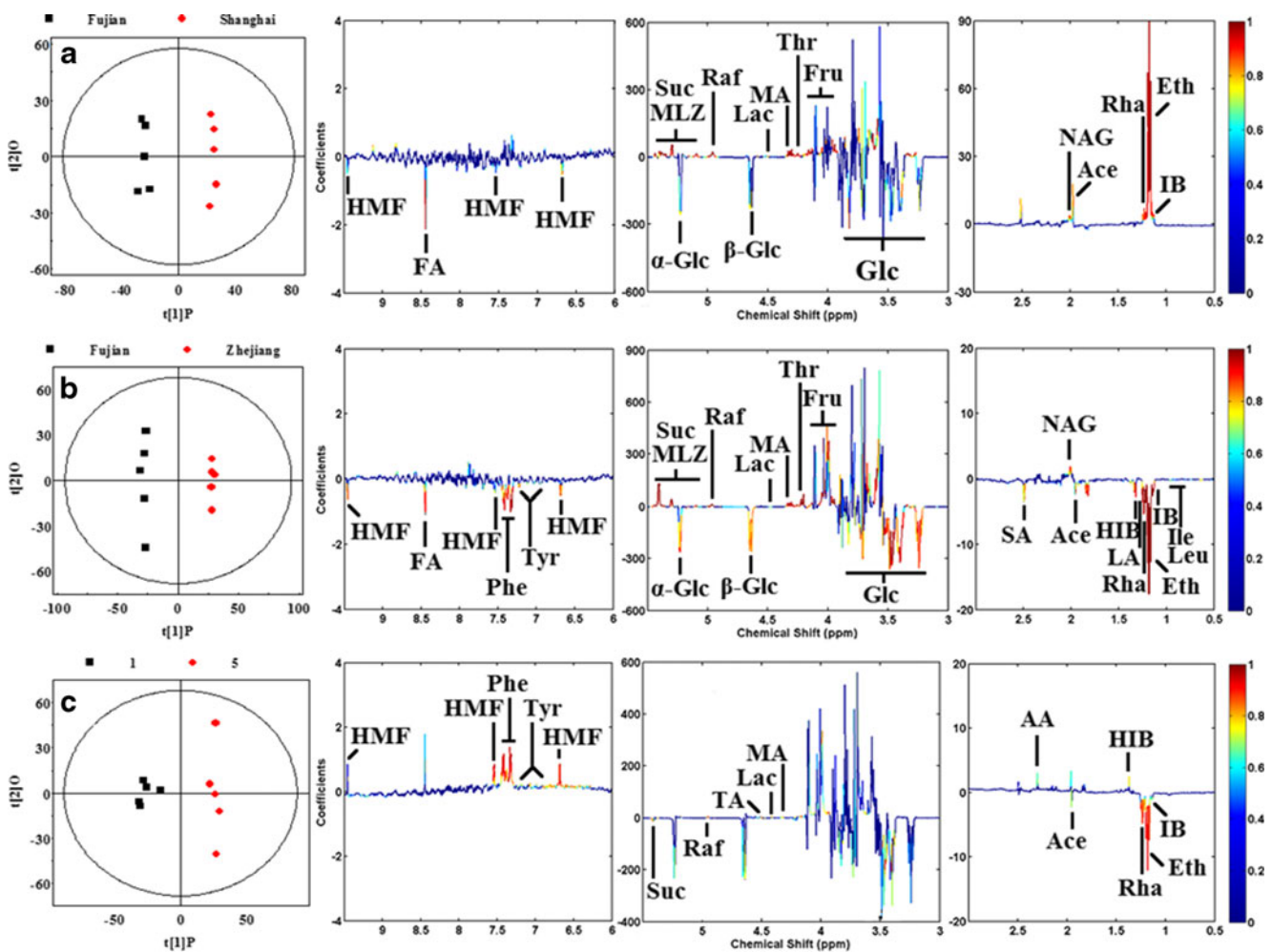
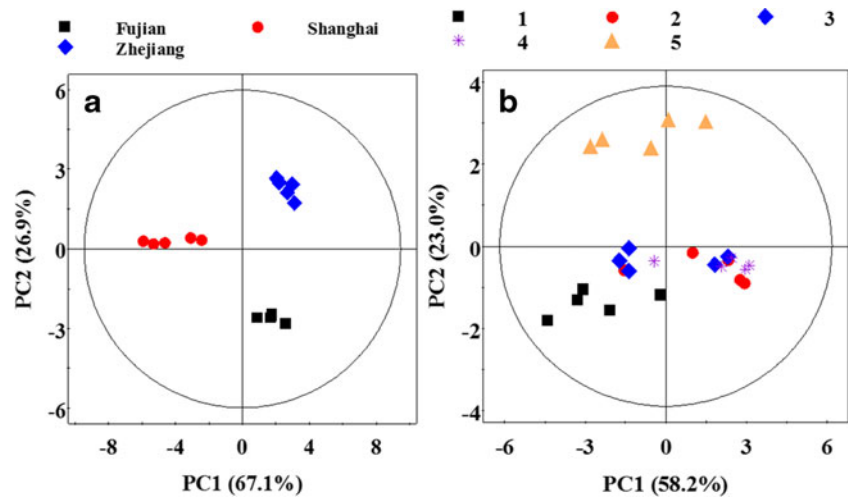
The relative contribution of the components in the well-separated Fujian and Shanghai FS honeys can be obtained from the loadings plot (Fig. 4a and Table 2): high concentrations of acetic acid, ethanol, IB, malic acid, NAA, threonine, and some saccharides including fructose, lactose, melezitose, raffinose, rhamnose, and sucrose are characteristics of FS honeys from Shanghai, while α- and β-glucose, formic acid, and hydroxymethylfurfural (HMF) are characteristics of FS

honeys from Fujian. Similarly, the compositional differences can also be derived from the geographical model of Fujian and Zhejiang FS honeys. Compared with Fujian FS honey, higher levels of NAA, threonine, and some saccharides including fructose, lactose, melezitose, raffinose, and sucrose existed in Zhejiang FS honey. More amino acids including isoleucine, leucine, phenylalanine, and tyrosine, glucose, and organic acids including acetic acid, HIB, formic acid, IB, lactic acid, malic acid, succinic acid, and tartaric acid were detected in Fujian FS honey than Zhejiang FS honey.

### Effects of Storage Time on the Compositional Changes

Quality and composition of honey are also affected by storage time and conditions (Adriana et al. 1999; Anupama et al. 2003). In order to understand the effect of storage time on the compositional changes of honey, 25 floral honey samples

**Fig. 3** PCA scores plots derived from the  $^1\text{H}$  NMR spectra of *Flos sophorae* honeys from three different geographical origins (a) and from five different production dates (b); 1 to 5: from the latest to the earliest date



**Fig. 4** OPLS-DA score plots (left panel) and corresponding loading plots (right panel) derived from  $^1\text{H}$  NMR spectra of *Flos sophorae* honeys from different groups to show the differences in geographical origins and production dates. The color maps show the significance of

compositional variations between the pairwise groups. Peaks in the positive direction indicate that those components are more abundant in groups in the positive direction of first principal component and vice versa. Keys of the assignments are given in Fig. 1

**Table 2** Componential variations of *Flos sophorae* honeys with geographical origins and production dates

Component and spectral assignments	$r^a$		
	Fujian–Shanghai $R^2X=50.3\%$ , $R^2Y=0.991$ , $Q^2=0.851^c$ $p<0.001^d$	Fujian–Zhejiang $R^2X=62.4\%$ , $R^2Y=0.997$ , $Q^2=0.943$ $p<0.001$	1–5 <sup>b</sup> $R^2X=60.8\%$ , $R^2Y=0.973$ , $Q^2=0.866$ $p<0.005$
Acetic acid: 1.96(s <sup>e</sup> )	0.923	−0.961	−0.827
Acetoacetic acid: 2.30(s)	–	–	0.952
$\alpha$ -Glucose: 3.42(t), 3.54(dd), 3.71(t), 3.73(m), 3.84(m), 5.23(d)	−0.951	−0.983	−0.881
$\alpha$ -Hydroxyisobutyric acid: 1.37(s)	–	−0.943	0.948
$\beta$ -Glucose: 3.25(dd), 3.41(t), 3.46(m), 3.49(t), 3.90(dd), 4.64(d)	−0.922	−0.969	−0.912
Ethanol: 1.18(t)	0.996	−0.998	−0.986
Formic acid: 8.45(s)	−0.971	−0.973	–
Fructose: 3.56(m), 3.71(m), 3.81(m), 3.89(dd), 4.00(m), 4.04(d), 4.11(m)	0.989	0.997	–
Hydroxymethylfurfural: 6.68(d), 7.54(d), 9.46(s)	−0.980	−0.987	0.996
Isobutyric acid: 1.14(d)	0.979	−0.996	−0.905
Isoleucine: 0.94(t), 1.01(d)	–	−0.822	–
Lactic acid: 1.33(d)	–	−0.992	–
Lactose: 3.94(m), 4.47(d)	0.996	0.999	0.968
Leucine: 0.96(t)	–	−0.906	–
Malic acid: 4.31(dd)	0.998	0.999	0.960
Melezitose: 5.06(d), 5.29(d), 5.49(d)	0.998	0.997	–
N-Acetylglutamic acid: 1.90(m), 2.00(m), 2.23(t)	0.992	0.977	–
Phenylalanine: 7.33(d), 7.37(t), 7.42(dd)	–	−0.995	0.988
Raffinose: 4.95(d), 5.29(d), 5.43(d)	0.999	0.999	−0.957
Rhamnose: 1.23(m), 5.10(d)	0.995	−0.998	−0.976
Succinic acid: 2.50(s)	–	−0.982	–
Sucrose: 3.56(m), 4.21(d), 5.40(d)	0.998	0.999	−0.972
Tartaric acid: 4.41(s)	–	−0.831	0.966
Threonine: 4.24(m)	0.996	0.999	–
Tyrosine: 6.90(d), 7.19(d)	–	−0.861	0.961

An en dash “–” means the correlation coefficient  $|r|$  is less than 0.811

<sup>a</sup> Correlation coefficients, positive and negative signs indicate positive and negative correlation in the concentrations, respectively. The correlation coefficient of  $|r| > 0.811$  was used as the cutoff value for the statistical significance based on the discrimination significance at the level of  $p=0.05$  and  $df$  (degree of freedom)=4

<sup>b</sup> 1, the latest production date; 5, the earliest production date

<sup>c</sup> The cross-validation parameters of OPLS-DA models corresponding to the pairwise groups

<sup>d</sup> The results from CV-ANOVA of OPLS-DA models corresponding to the pairwise groups

<sup>e</sup> Multiplicity: s, singlet; d, doublet; t, triplet; dd, doublet of doublets; m, multiplet

in five production dates (five samples each group) were performed by PCA. The scores plot (Fig. 3b) shows that the honey samples were regularly classified by production dates, in which the earliest honey samples were separated clearly from others, and the samples with the similar production date (2, 3, and 4 in our study) were sub-clustered in PCA scores plot. More in-depth analyses were carried out for the same botanical-origin honeys to find the variation for the honey samples from the earliest to the latest

production dates by applying OPLS-DA models (Fig. 4c). The model for production dates offers an overall quality of fit ( $R^2Y$ ) of 0.973 and an overall cross-validation coefficient ( $Q^2$ ) of 0.866. The loadings plot derived from the OPLS-DA model applied to honeys of earliest and latest production dates (with 14-month interval) indicated that the levels of sucrose,  $\alpha$ - and  $\beta$ -glucose, raffinose, rhamnose, IB, acetic acid, and ethanol decreased with time whereas those of lactose, HMF, phenylalanine, tyrosine, and some organic



acids including acetoacetic acid, HIB, malic acid, and tartaric acid increased over the same period.

HMF, a sugar breakdown product (Zappala et al. 2005), is an important indicator for evaluation of storage time and honey freshness (Morales et al. 2009). As reported previously (Jimenez et al. 1994; Turhan et al. 2008), the increased level of HMF was observed in honey with the time in our study, which is negatively correlated to the levels of sucrose, glucose, raffinose, and rhamnose during the storage period. The contents of free amino acids, phenylalanine and tyrosine, in this study were found to be reduced with the storage time, suggesting the possible association with the Maillard reaction (Iglesias et al. 2006). The increased levels of total acidity and free acids in honey are the most pronounced effect in storage duration (Bath and Singh 2000). The obvious augment of organic acids including acetoacetic acid, HIB, malic acid, and tartaric acid in our study is correlated to storage time, and the reduced level of ethanol and acetic acid is also correlated to storage time as volatile fraction of honey.

## Conclusions

This investigation has suggested that NMR spectra can provide accurately qualitative information of 20 components in the honeys including saccharide and some carboxylic organic and amino acids. <sup>1</sup>H NMR spectroscopy combined with multivariate statistical analysis could be successfully used in the identification and differentiation of botanic and geographical origins of honeys. Further, OPLS-DA demonstrated that the detailed compositional variations are associated with the storage time and geographical origin. Our results and technologies will be also helpful for classification of other foods and their qualitative and quantitative controls.

It should be pointed out that the possible adulteration in the honeys (although we chose the well-respected supermarkets to get the authentic honeys as we can) will cause some variations in the final quantified results and multivariate statistical models; however, such method should be helpful to the identification of honey adulteration in the next step. And, added sample size can improve the accuracy, reliability, and transferability of the statistical models.

## Compliance with Ethical Standards

**Funding** This study was funded by the Key Project of Science and Technology Commission of Fujian Province (2014H0045) and the Fundamental Research Funds for the Central Universities (2013121007) and NCET.

**Conflict of Interest** All authors declare no conflict of interest.

**Ethical approval** This article does not contain any study with human participants or animals performed by any of the authors.

**Informed consent** Not applicable.

## References

- Adriana P, Leila B, Maria PB (1999) Color changes during storage of honeys in relation to their composition and initial color. *Food Res Int* 32:185–191
- Anupama D, Bhat KK, Sapna VK (2003) Sensory and physico-chemical properties of commercial samples of honey. *Food Res Int* 36:183–191
- Aparna AR, Rajalakshmi D (1999) Honey—its characteristics, sensory aspects, and applications. *Food Rev Int* 15(4):455–471
- Bath PK, Singh N (2000) A research note chemical changes in *Helianthus annuus* and *Eucalyptus lanceolatus* honey during storage. *J Food Qual* 23(4):443–451
- Boffo EF, Tavares LA, Tobias ACT, Ferreira MMC, Ferreira AG (2012) Identification of components of Brazilian honey by <sup>1</sup>H NMR and classification of its botanical origin by chemometric methods. *LWT Food Sci Technol* 49(1):55–63
- Cavia MM, Fernandez-Muino MA, Gomez-Alonso E, Montes-Perez MJ, Huidobro JF, Sancho MT (2002) Evolution of fructose and glucose in honey over one year: influence of induced granulation. *Food Chem* 78(2):157–161
- Cazor A, Deborde C, Moing A, Rolin D, This H (2006) Sucrose, glucose, and fructose extraction in aqueous carrot root extracts prepared at different temperatures by means of direct NMR measurements. *J Agric Food Chem* 54(13):4681–4686
- Consonni R, Cagliani LR (2008) Geographical characterization of polyfloral and acacia honeys by nuclear magnetic resonance and chemometrics. *J Agric Food Chem* 56(16):6873–6880
- Consonni R, Cagliani LR, Cogliati C (2012) NMR characterization of saccharides in Italian honeys of different floral sources. *J Agric Food Chem* 60(18):4526–4534
- Conte LS, Miorini M, Giomo A, Bertacco G, Zironi R (1998) Evaluation of some fixed components for unifloral honey characterization. *J Agric Food Chem* 46(5):1844–1849
- Donarski JA, Roberts DPT, Charlton AJ (2010) Quantitative NMR spectroscopy for the rapid measurement of methylglyoxal in manuka honey. *Anal Methods* 2(10):1479–1483
- Iglesias MT, Martin-Alvarez PJ, Polo MC, de Lorenzo C, Gonzalez M, Pueyo E (2006) Changes in the free amino acid contents of honeys during storage at ambient temperature. *J Agric Food Chem* 54(24):9099–9104
- Janiszewska K, Aniolowska M, Nowakowski P (2012) Free amino acids content of honeys from Poland. *Polish J Food Nutri Sci* 62(2):85–89
- Jimenez M, Mateo JJ, Huerta T, Mateo R (1994) Influence of the storage-conditions on some physicochemical and mycological parameters of honey. *J Sci Food Agric* 64(1):67–74
- Juan-Borras M, Domenech E, Hellebrandova M, Escriche I (2014) Effect of country origin on physicochemical, sugar and volatile composition of acacia, sunflower and tilia honeys. *Food Res Int* 60:86–94
- Karabagias IK, Badeka A, Kontakos S, Karabournioti S, Kontominas MG (2014) Characterization and classification of *Thymus capitatus* (L.) honey according to geographical origin based on volatile compounds, physicochemical parameters and chemometrics. *Food Res Int* 55:363–372
- Mannina L, Sobolev AP, Viel S (2012) Liquid state <sup>1</sup>H high field NMR in food analysis. *Prog Nucl Magn Reson Spectrosc* 66:1–39
- Mazzoni V, Bradesi P, Tomi F, Casanova J (1997) Direct qualitative and quantitative analysis of carbohydrate mixtures using C-13 NMR spectroscopy: application to honey. *Magn Reson Chem* 35:S81–S90

- Monakhova YB, Kuballa T, Lachenmeier DW (2013) Chemometric methods in NMR spectroscopic analysis of food products. *J Anal Chem* 68(9):755–766
- Morales V, Sanz ML, Martin-Alvarez PJ, Corzo N (2009) Combined use of HMF and furosine to assess fresh honey quality. *J Sci Food Agric* 89(8):1332–1338
- Ohmenhaeuser, M., Monakhova, Y. B., Kuballa, T., Lachenmeier, D. W. (2013). Qualitative and Quantitative Control of Honeys Using NMR Spectroscopy and Chemometrics. *ISRN Analytical Chemistry*, 2013: Article ID 825318
- Trygg J, Holmes E, Lundstedt T (2007) Chemometrics in metabolomics. *J Proteome Res* 6(2):469–479
- Turhan I, Tetik N, Karhan M, Gurel F, Tavukcuoglu HR (2008) Quality of honeys influenced by thermal treatment. *Lwt-Food Sci Technol* 41(8):1396–1399
- van den Berg RA, Hoefsloot HCJ, Westerhuis JA, Smilde AK, van der Werf MJ (2006) Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* 7
- Venir E, Spaziani M, Maltini E (2010) Crystallization in “Tarassaco” Italian honey studied by DSC. *Food Chem* 122(2):410–415
- Yang Y, Battesti M-J, Paolini J, Muselli A, Tomi P, Costa J (2012) Melissopalynological origin determination and volatile composition analysis of Corsican “Erica arborea spring maquis” honeys. *Food Chem* 134(1):37–47
- Zappala A, Fallico B, Arena E, Verzera A (2005) Methods for the determination of HMF in honey: a comparison. *Food Control* 16(3):273–277