Malaysian Journal of Analytical Sciences, Vol 23 No 1 (2019): 1 - 13
DOI: https://doi.org/10.17576/mjas-2019-2301-01

MALAYSIAN JOURNAL OF ANALYTICAL SCIENCES
Published by The Malaysian Analytical Sciences Society
ISSN 1394 - 2506

UNIQUE SIGNATURES OF HONEYS AS A MEANS TO ESTABLISH PROVENANCE

(Tanda Kenalan Unik Madu Sebagai Satu Cara Menentukan Provenans)

Syazwani Dzolin[1], Wan Aini Wan Ibrahim[1, 2]*, Naji Arafat Mahat[1], Aemi Syazwani Abdul Keyon[1], Zuhaimy Ismail[3]

[1]*Department of Chemistry, Faculty of Science*
[2]*Centre for Sustainable Nanomaterials, Ibnu Sina Institute for Scientific and Industrial Research*
[3]*Department of Mathematical Sciences, Faculty of Science*
Universiti Teknologi Malaysia, 81310 UTM, Johor Bahru, Johor, Malaysia

**Corresponding author: waini@utm.my, wanaini@kimia.fs.utm.my*

Received: 8 August 2018; Accepted: 12 December 2018

Abstract

Global consumer demands for high quality genuine honey are increasing continuously. The ability to identify the geographical and botanical origins of honey would be of great importance to the improvement of quality control, as well as protection of reputation and confidence of the honey industry. The establishment of unique signatures of honey is one of the potential approaches for honey provenance. A great deal of literature has been published on honey authenticity and provenance. However, specific study focusing on the provenance of honey, especially those of the stingless bee species is lacking, especially in Malaysia. Consequently, this mini review is to highlight the presence of biomarkers in honeys for establishing their provenance *via* the application of spectroscopy and chemometric techniques.

Keywords: stingless bee, bee honey authenticity, provenance, spectroscopy, chemometric

Abstrak

Permintaan global pengguna terhadap madu tulen kualiti tinggi meningkat secara berterusan. Keupayaan untuk mengenal-pasti asal geografi dan botani madu akan menjadi amat penting kepada peningkatan kawalan kualiti, serta melindungi reputasi dan keyakinan industri madu. Kewujudan tanda kenalan unik madu adalah salah satu pendekatan yang berpotensi untuk asal madu. Banyak manuskrip telah diterbitkan untuk keaslian madu dan provenans. Walau bagaimanapun, kajian khusus yang memberi tumpuan kepada provenans madu, terutama bagi lebah tanpa sengat adalah sedikit terutamanya di Malaysia. Oleh itu, kajian tinjauan mini ini adalah untuk menyoroti kehadiran penanda bio dalam madu untuk menentukan provenansnya penggunaan teknik spektroskopi dan kemometrik.

Kata kunci: lebah tanpa sengat, keaslian madu lebah, provenans, spektroskopi, kemometrik

Introduction

Bee honey is appreciated worldwide by old folk and modern practitioners as a natural, nutritious product, mainly because of its therapeutic significance towards human health. Honey is produced by bees collecting nectar from flowers and converting it into a sweet natural product that is claimed to be healthier favourable nutritional option than plain sugar [1]. The properties of honey such as aroma, color, composition and flavour depend generally on the flowers, climate, geographical regions and honeybee species involved in its production [2, 3]. Honey is composed

mainly of fructose and glucose (65%), water (18%), minimal protein and lipid contents [4, 5]. It also consists of different quantities of minerals ranging between 0.02 g/100 g and 1.03 g/100 g, with potassium being the most abundant element comprising approximately one-third of the total mineral content [6-8]. Trace minerals such as potassium, calcium, sodium, iron, copper, zinc and manganese play a critical role in biological systems in which they maintain normal physiological systems and influence reproduction as catalyst for various biochemical reactions [9]. Because of the presence of various chemical compositions, honey has been found to significantly affect human nutrition, healing and preventing illness. In addition to its medicinal benefits, presence of contaminants such as pesticide residues and excessive amounts of minerals and heavy metals has been reported in honey, accentuating the needs for producing safe honey for human consumption [10].

Consumers have the rights to receive the truthfulness of information about the honey in market. The authenticity of honey is the major concern in order to prevent from the adulteration of honey and fraudulent practices during the manufacturing process. Depending on the common fraud, it also may represent a health risk to the consumers and therefore it is important to trace the authenticity of honey in the industry not only for economical but also for safety reasons. Generally, authenticity of honey means that correlation between declarations of honey label and the provenance of the honey itself. Over the past decade, honey is known to be the natural food product that possesses hundreds of benefits towards human health and becomes favourable by old folk and modern practitioners. Unfortunately, the authenticity of the honey decreases as the demand for it increases daily.

Normally, the composition of honey is closely related to its botanical origin and also the geographical area in which it originated. The characteristics of soil and climate will determine the floral that produce honey [11]. Contaminants are transferred from bees to ripened honey and alter its unique composition and the high quality [14]. The surrounding area of mines and steelworks, industrialized and urban areas or motorways with exhausts from vehicles has usually reflected in much higher concentrations of trace metals such as Al, Ba, Ca, Cd, Cr, Cu, Mg, Mn, Ni, Pb, Pd or Zn [12-16].

Provenance establishment of worldwide honey

In this day and age, the development of analytical approaches and advanced statistical techniques has gained interest for the determination of certain major and minor compounds in honey to serve as the biomarkers for its geographical or botanical origin. Thus far, several techniques have been used to analyze the compound composition of honeys. Generally, the most common methods are flame atomic absorption spectroscopy (FAAS) [17], electrothermal atomic absorption spectrometry (ETAAS) [18], high-performance ion chromatography (HPIC) [19], inductively coupled plasma-atomic emission spectroscopy (ICP-AES) [20, 21] and inductively coupled plasma-mass spectrometry (ICP-MS) [20, 22, 23].

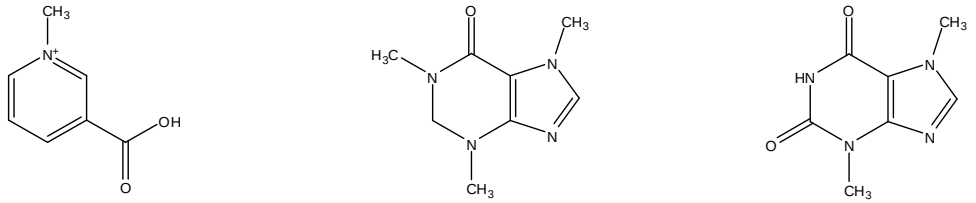
In recent years, authors have begun to apply chemometric techniques into spectroscopy methods in order to establish the provenance of honey. The chemometric evaluation is important to classify, categorize and discriminate the honey arising from different geographical areas [4, 24-26]. Chemometrics can be defined as a way to analyze complicated chemical raw data to more specific and precise data that can be seen and understand clearly. This process may take into account the extraction of information from raw data and ensuring the data selected consisted of maximum valuable information. Chemometrics can be grouped into three analyses such as exploratory analysis, classification or discriminant analysis and regression or prediction models [27]. Unsupervised approaches are usually used in exploratory analysis, which consist of algorithms that cluster metabolites into groups and visualize the data to show their similarities and differences. Principal component analysis (PCA) and hierarchical cluster analysis (HCA) are the commonly used of unsupervised methods. Other nested algorithmic techniques are soft independent modelling of class analogy (SIMCA) and k-nearest neighbours (kNN). For many years, few methods have been conducted in discriminant analysis such as linear discriminant analysis (LDA), partial least square discriminant analysis (PLS-DA), artificial neural networks (ANN), discriminant function analysis (DFA), canonical variant analysis (CVA) and support vector machine (SVM). In regression analysis, methods previously used were principle components regression (PCR), multiple linear regression (MLR), partial least square (PLS), orthogonal partial least square (OPLS), genetic programming/computing (GP/GC), genetic algorithm (GA), evolutionary programming (EP) and evolutionary algorithm (EA), classification and regression tree (CART) and multivariate adaptive regression splines (MARS).

This mini review focuses on the use of analytical approaches combined with chemometric techniques between year 2000 and 2017 for honey provenance (geographical and botanical origins) based on the presence of its biomarkers.

Nuclear magnetic resonance spectroscopy

Previously, many researches used nuclear magnetic resonance (NMR) to identify the origin of honeys as it possessed a powerful ability for compounds structures determination. This technique was frequently used as it provides broad understanding in complex structures especially in food like honey. Donarski et al. [28] studied on one hundred and eighty-two Corsican honey samples collected from five countries, namely Austria, France, Germany, Ireland and Italy that involved ten different regions. The results have demonstrated the viability of developing the accurate models to be used in identification of Corsican honey using proton nuclear magnetic resonance spectroscopy (^{1}H NMR). The produced NMR spectra were used as an input variable for PLS followed by LDA and GP. Overall, models were generated based on three methods; PLS-LDA, two-stage GP and combination of PLS-GP with classification rates for the discrimination of Corsican and non-Corsican honeys were 75.8, 94.5 and 96.2%, respectively. Venetian blind cross-validation was used to generate all models. The NMR spectroscopy was used to identify the molecular structural characteristics of present compounds that gave identity to variables that is most relevant to the classification of Corsican honey. Amazingly, for the first time, the alkaloid trigonelline have been reported to exist in Corsican honey (orange honey) which may prove to be an indicator of its geographical origin.

Schievano et al. [29] carried out a study using the ^{1}H NMR method. They identified and quantified trigonelline compound for the first time in Coffea honey from Colombia and Honduras with addition compounds, namely caffeine and theobromine which previously have not been reported by Donarski et al. [28]. Three marker compounds (Figure 1) with following range of amount have been proposed to be indicators of the botanical origin of the Coffea honeys; trigonelline (23 to 86 mg/kg), caffeine (15 to 98 mg/kg) and theobromine (25 to 160 mg/kg). The constraint in this study was the small number of samples because the Coffea honey is by no means a common product.



Trigonelline Caffeine Theobromine

Figure 1. Identified marker compounds in orange and *Coffea* honeys using NMR spectroscopy

In 2010, Donarski et al. [30] carried out another investigation to identify the botanical biomarkers from Corsican honey samples. The study was conducted for two years. In the first year, one hundred and eighty-two Corsican honey samples were collected from five countries; Austria, France, Germany, Ireland and Italy while in the second year, one hundred and ninety-two Corsican honey samples were collected from four countries namely Austria, France, Germany and Italy. Commercial chestnut honey, non-commercial honey and Corsican honey from the region Châtaigneraie were used as the outsource samples to compare and confirm the identity of the biomarkers present. Four biomarkers have been identified previously by Donarski et al. [28] during collection of the Corsican honey samples in the first year. The efficacy of GP was demonstrated during the determination of the biomarkers from the complexity of NMR spectroscopy data. They have elucidated the four biomarkers and claimed that the biomarkers were confirmed to be kynurenic acid and 2,5-dihydroxyphenylacetic acid. The kynurenic acid was found in sweet chestnut honey, whereas the 2, 5-dihydroxyphenylacetic acid was detected in strawberry-tree honey. Another two biomarkers have been characterized but unidentified. Therefore, they concluded that even honey

samples collected from similar geographical areas contained different compound composition basically according to their botanical origin.

Consonni and Cagliani [31] characterized geographically twenty-three polyflorals honey samples from Argentina, European Committee (EC) countries, Hungary and Italy and eighteen acacia honey from Italy and Hungary using NMR spectroscopy and chemometrics namely PCA and PLS-DA. The ^{1}H NMR spectra of the Acacia honey samples have similarity with the polyfloral honey samples. It means that, both showing the same water-soluble compounds content, even though they are different in concentrations. The compounds presence was grouped based on specific compound resonances comprised methyl amino acid region (0-2 ppm), sugars region (3-5.5 ppm) and aromatic region (6-10 ppm). Between the resonances, few compounds were readily identified, such as tyrosine, phenylalanine, formic acid, proline, alanine, threonine, acetate and lactic acid. The first two principal components (PCs) showed clear differentiation among samples based on different variety with 95.5% of the total variance. Corresponding to loading plot, acacia honey possessed high sugars content namely sucrose, turanose, α-glucose and β-glucose compared to the polyfloral honey. The PCA that was performed on a full spectrum did not provide a clear differentiation among the polyfloral honey samples according to the geographical origin. Following the unsuccessful separation of the polyfloral honey, hierarchical PLS-DA was conducted on thirteen polyfloral honey samples of certain origin and result showed a clear differentiation of samples based on their geographical origin with goodness fit, R^2 (91.9%) and prediction goodness parameter, Q^2 (72.7%). The unsupervised PCA technique demonstrated a clear geographical differentiation of the eighteen Acacia honey samples between Italian and Hungarian origins. Total six PCs exhibited 99.7% of total variance with $Q^2 = 98.1\%$.

Beretta et al. [32] conducted chemometric analysis to identify the botanical origin of commercial honey samples that were purchased from different locations in Milan, Italy in 2006 and 2007. Forty-four selected honey samples were from following floral sources; thyme, linden, orange, eucalyptus, multi-flora, honeydew, chestnut, almond, heather, rosa mosqueta, lavender, acacia, rosemary, rhododendron, cardoon, liquorice, sunflower, strawberry tree and French honeysuckle (*Hedysarum coronarium*). The 300 MHz ^{1}H NMR was used to construct the pattern of chemical compositions in honey samples and the results displayed the relationship between the compound composition and the botanical origin of honeys. Three botanical origin of honey samples, namely honeydew, chestnut and linden were chosen as suitable markers of origin based on their special characteristic that were not detectable in other honey origins. The honeydew honey showed the typical resonance of an aliphatic compound which claimed to derive from the plant phloem sap whereas in the chestnut honey, it displayed the typical signal of kynurenic acid. From the ^{1}H NMR profile in the linden honey, strong signals confirmed the presence of mono-terpene derivative cyclohexa-1,3-diene-1-carboxylic acid (CDCA) and 1-O-β-gentiobiosyl ester (CDCA-GBE). According to the specific markers obtained, the PCA multivariate technique had found different clusters of honeys. Therefore, this study strongly suggested that the proposed ^{1}H NMR method can distinguish the three honey botanical origins compared to other honey origins.

Lolli et al. [33] examined the data from seventy-one honey samples of five different floral origins, specifically chestnut, robinia, citrus, eucalyptus and polyfloral using high resolution nuclear magnetic resonance (HR-NMR). Both ^{1}H and ^{13}C NMR signals assigned the presence of α-glucopyranose, β-glucopyranose, α-fructofuranose, β-fructofuranose and β-fructopyranose as the marker compounds using deuterium oxide (D_2O) and (methyl sulfoxide)-$_{d6}$ solvents. Moreover, according to the overall data obtained, ^{1}H-^{13}C HMBC followed by chemometric analysis showed an effective technique to identify honey botanical origins. The chemometric techniques, PCA and general discriminant analysis (GDA) were used to analyze the spectral data. Between the two techniques, GDA showed higher efficiency compared to PCA in an effort to construct models in the prediction of honey origin.

Schievano and colleagues [34] carried out an investigation to identify the botanical origin of hundred and eighteen chloroform honey extracts obtained from Veneto beehives from four different botanical origins; acacia (Pseudo Acacia *robinia* L.), chestnut (*Castanea sativa*), linden (*Tilia* spp.) and polyfloral. A new simple and reproducible NMR approach combined with PCA and PLS-DA chemometric analysis was used to characterize the honey samples. Results through separate PCA modelling exhibited a specific characteristic (marker compounds) of each botanical origin based on the specific resonances obtained with 95% confidence intervals of the DmodX criterion and the critical value (D_{crit}). The identified marker compounds were 4-(1-hyfroxy-1-methylethyl)cyclohexa-1,3-

dienecarboxylic acid and 4-(1-methylethenyl)cyclohexa-1,3-dienecarboxylic acid from linden honey; chrysin from acacia honey; LACT-3-PKA from chestnut honey and hexanal from all botanical origin honeys. Data from PLS-DA found the coefficient of determination $R^2 = 0.67$ and validation correlation coefficient, $Q^2 = 0.77$.

Schievano et al. [35] performed a similar experiment as previously reported [33] to show that the botanical origin of three hundred and fifty-three chloroform extract of honey samples from seven different floral origins; six were monofloral (acacia, honeydew, orange, chestnut, eucalyptus, linden) and polyfloral. Specific markers of monofloral honeys have been identified using NMR-based metabolomic approach and associated with O2PLS-DA multivariate data analysis. The results exhibited high precision of classification that suggests the usefulness of this approach to discriminate the honey sample origin characteristic. Figure 2 displays the marker compounds detected from six monofloral honeys.The protons originated signals that extracted from the S-Plots were labeled in bold and marked with star.

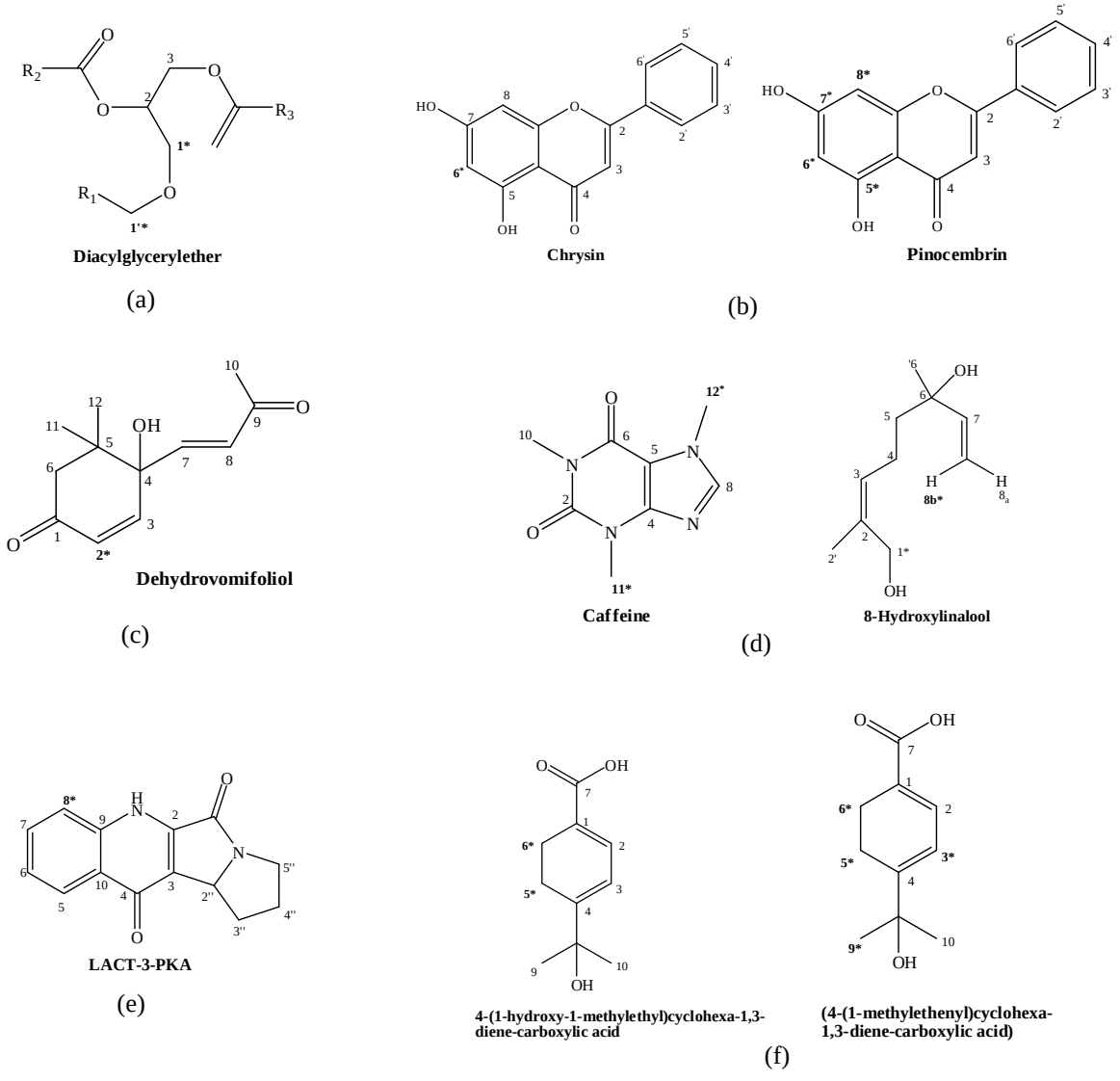


Figure 2. Identified marker compounds of monofloral honeys from (a) honeydew, (b) acacia, (c) eucalyptus, (d) orange, (e) chestnut and (f) linden

Ohmenhaeuser et al. [36] studied the qualitative and quantitative analysis of three hundred and twenty-eight honey samples to observe their botanical origin using ^{1}H and ^{13}C of 400 MHz NMR spectroscopy. Data analyzed by PCA showed clusters of honeys that came from the same botanical origin and the chemical shifts confirmed the presence of glucose and fructose to be one of the reasons of this differentiation and both compounds exhibited much higher peak intensity compared to other compounds using the NMR method. Additionally, the large number of overlapped signals caused other compounds to obscure. Further analysis showed 95-100% of accuracy has been determined using SIMCA and PLS-DA to classify spectra based on the honey botanical origins.

Gas chromatography-mass spectrometry

Gas chromatography (GC) has been widely used for classification of the geographical and botanical origins. When the technique was coupled with mass spectrometry (MS), this selective and sensitive detector allows structural elucidation of the analyzed compounds. A recent study by Azevedo et al. [45] conducted in Santa Catarina, Brazil demonstrated the geographical origin of honeydew honey. A total of twenty-one samples of bracatinga honeydew honey (Bhh) were collected from five mountainous regions namely Bocaino do Sul, Bom Retiro, Lages, Urubici and Urupema between February to June 2014. Free amino acids (FAA) have been identified to be marker compounds using GC-MS and from the data obtained, further classification by PCA technique have been demonstrated for these regions. In this study, PCA technique able to discriminate the honey samples based on their geographical origin with 82% of total variance (sum of two principal components). The authors claimed that FAA profile can be used in honey authenticity determination as it is a reliable analytical method.

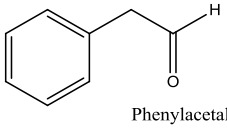
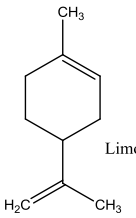
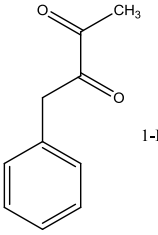
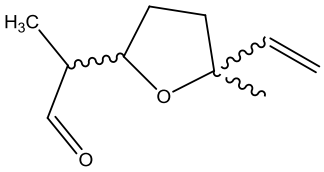
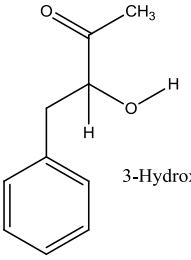
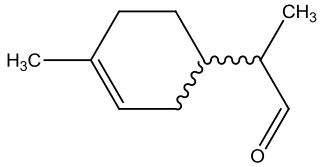
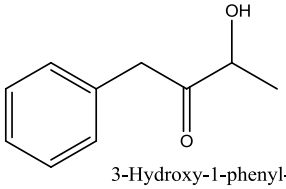
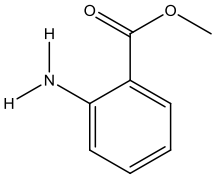
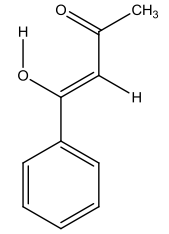
Castro-Vázquez [38] studied a total of forty-nine of Spanish monofloral honey sample from eucalyptus, citrus, thyme, rosemary, heather and lavender botanical origins to undergo analysis to isolate the marker volatile compounds using GC-MS and sensory analysis. HCA was first conducted and presented by the dendrogram to show the distribution of botanical origins of honey samples. The honey samples from the same botanical origins were assigned in the same cluster and this highlighted their similarity. Some similarities of different botanical origins have been detected by successively higher-level clusters, however in order to obtain more specific information, PCA analysis has been carried out. Therefore, several volatile compounds have been detected in higher composition such as in citrus honey contain linalool derivatives for examples sinensal isomers, limonyl alcohol α-4-dimethyl-3-cyclohexene-1-actaldehyde and citrus aroma; eucalyptus honey possessed hydroxyketones, (acetoin, 5-hydroxy-2,7-dimethyl-4-octanone), p-cymene derivatives, 3-caren-2-ol and spathulenol and cheese aroma; lavender honey profiled hexanal, nerolidol oxide, coumarin, hexanol, hotrienol and aromatic herb aromas; heather honey presented benzene, phenolic compounds, ripe fruit and spicy aromas. The compounds that were found in a specific honey sample can be used as markers of their botanical origins.

An isolated aroma compounds from a set of seventy-seven of unifloral honey samples from various botanical origins of Greece such as thyme (*Corydothymus capitatus*), pine (*Pinus* spp.), citrus (*Citrus* spp.), chestnut (*Castanea sativa*), fir (*Abies* spp.), cotton (*Gossypium hirsutum*) and heather (*Erica manipuliflora*) were obtained from headspace solid-phase micro-extraction (HS-SPME), followed by the analysis using GC-MS fingerprinting. For the discrimination of different types of honey, several biomarker groups have been detected, namely phenolics, aliphatics, aldehydes, terpenoids and alcohols. Specifically, for the thyme and citrus honeys, the compounds that act as biomarkers for their discrimination are shown in Table 1. According to the results from combined MS spectra of seven botanical origins, orthogonal partial least squares-discriminant analysis (OPLSTM-DA) exhibited a strong discriminant model. Moreover, the OPLSTM-hierarchical cluster analysis (OPLSTM-HCA) showed higher than 98% of percentage value of correct classification of honey samples. From the results obtained, it clearly demonstrated the potential of the use of combined HS-SPME-GC-MS and chemometric analysis (OPLSTM-HCA) in order to discriminate and classify honey samples based on their botanical origin [37].

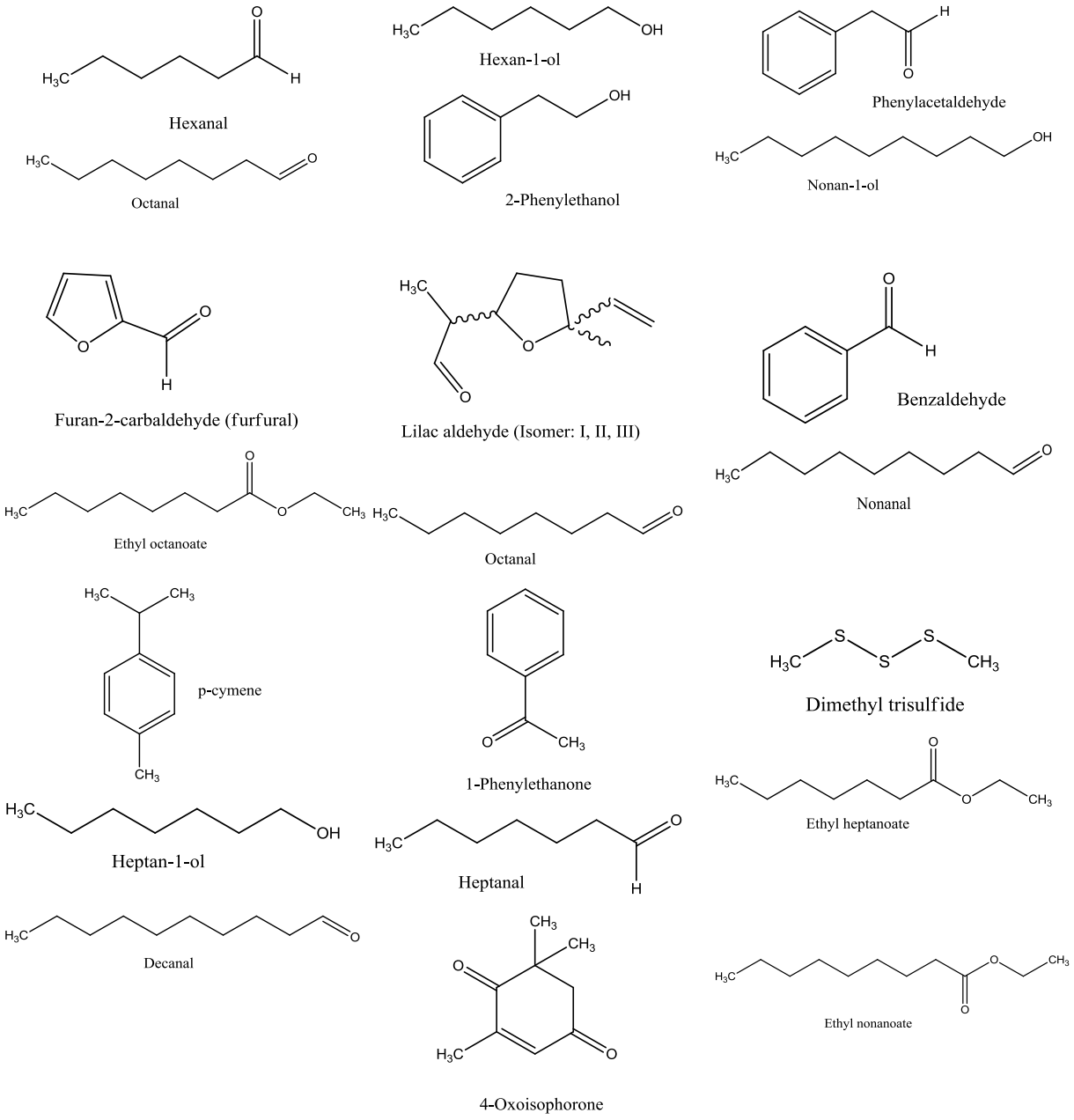
A study under Europe's largest research project on food traceability (EU TRACE) have been conducted by Cajka and coworkers [39] on three hundred and seventy-four honey samples consequently to trace honey origin based on volatile pattern processing using artificial neural network (ANN). The honey samples have been directly provided by producers and claimed their authenticity. The harvesting process took two years (2006 and 2007). In the first harvest period (2006), six types of honey samples were collected namely Corsican, non-Corsican-French, Italian, Austrian, Irish and German. For the second harvest process in 2007, only five types of honey samples were

collected; Corsican, non-Corsican-French, Italian, Austrian and German. In the preliminary data analysis, PCA analysis performed clustering among the samples and divided them into two groups, labelled as 'Corsica' and 'non-Corsica'. The honey samples from various geographical and floral origins have been analyzed using head-space solid-phase micro-extraction (HS-SPME) coupled with 2D-gas chromatography-time-of-flight mass spectrometry (GC-GC-TOF-MS). Twenty-six aroma compound markers (Figure 3) were selected based on the observed peaks that varied significantly in their intensities. The selected marker compounds were used to test the chemometric model, ANN with multilayer perceptron (MLP). The results obtained showed the best prediction (94.5%) and classification (96.5%) capabilities of ANN-MLP model to improve the model performance of total 2 years harvest compared to separate year harvests.

Table 1. Identified marker compounds of thyme and citrus honeys

Thyme honey	**Citrus honey**
 Phenylacetaldehyde	 Limonene
 1-Phenyl-2,3-butanedione	 Lilac aldehyde (Isomer: I, II, III)
 3-Hydroxy-4-phenyl-2-butanone	 1-p-menthen-9-al isomers
 3-Hydroxy-1-phenyl-2-butanone	 Methyl anthranilate
 3-Hydroxy-4-phenyl-3-buten-2-one	

A continuity from an earlier study on three hundred and seventy-four honey samples have been done by Stanimirova et al. [40] to trace the geographical origin of honeys based on volatile compound profiles using four pattern recognition techniques. In order to investigate the influence of the year of sampling on the discrimination, LDA, discriminant partial least squares (DPLS), SIMCA and SVM have been used to construct models for the discrimination of Corsican and non-Corsican-French honeys for both years using the selected marker aroma compounds as identified by Cajka et al. [39] (Figure 3). The results obtained from four chemometric models demonstrated that the SVM combined with Pearson VII universal kemel (PUK) gave the best performance (91.5%) even though the DPLS (87.4%) and the LDA (85.2%) showed high sensitivities and specificities. The combination of GC-GC-TOF-MS with LDA, DPLS and SVM models presented a successful application to detect mislabeling of Corsican honey.



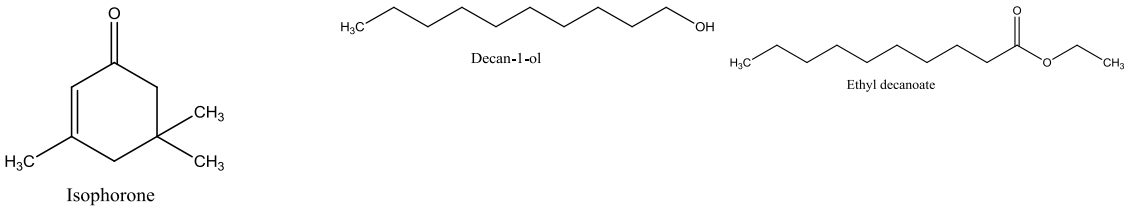


Figure 3. Identified marker aroma compounds of Corsican and non-Corsican-French honeys

Selected ion flow tube mass spectrometry

Selected ion flow tube mass spectrometry (SIFT-MS) has been used for trace gas analysis. The process involved the chemical ionization of volatile compounds in samples. Agila and Barringer [41] tested on unifloral American honeys from Ohio and Indiana, namely clover (*Trifolium* spp.), wildflower, star thistle (*Centaurea americana*), cranberry (*Vaccinium* spp.), blueberry (*Vaccinium* spp.) and unknown source to study the effect between the botanical origin and volatile compound profile of honey samples. These honey samples were analyzed using SIFT-MS coupled with SIMCA. Many volatile compounds were detected form all honey sample, however only several exhibited at high level in the samples such as 1-octen-3-ol, furfural, acetic acid, acetone, benzaldehyde and ethyl acetate. The honey samples from Ohio and Indiana with same botanical origins were totally having different volatile compound compositions. Indiana wildflower honey possessed the highest volatile compound composition, then Ohio wildflower honey. The same trend also can be observed in Indiana clover honey with higher volatile compounds concentration compared to Ohio clover honey. However, in this study, the highest concentration of most of the volatile compounds has been indicated in the unknown source of honey.

Near infrared spectroscopy

Near infrared (NIR) spectroscopy represents an emerging technique in the study of compounds containing O-H, N-H and C-H bonds but seldom been used by researchers. Woodcock and coworkers [42] investigated on the same number of honey samples as reported by Stanimirova et al. [40] and Cajka et al. [39] under a EU-funded TRACE project. In order to confirm the claimed by protected designation of origin (PDO) regions that is specified in European legislation, NIR spectroscopy have been used to construct a specific spectral fingerprinting for Corsican honey. In the preliminary screening, no separation was observed between the Corsican and non-Corsican honeys. Analysis using PCA gave some separation spectra between harvest 1 and harvest 2. However, there was a large amount of overlap between Corsican and non-Corsican honeys. By using cross-validation, a variable selection algorithm and a second derivative data pre-treatment, DPLS models have been developed. The results exhibited correct classification for Corsican (90.0%) and non-Corsican (90.3%) honey samples. Over again, using a variable selection procedure, highest correct classification percentage has been achieved, 90.4% and 86.3% for Corsican and non-Corsican honeys, respectively. This study is much on screening method, but in future it can be improved using advanced method to confirm the presence of honey biomarkers in order to identify its geographical origin.

Fourier transform-Raman spectroscopy

Fourier Transform (FT)-Raman Spectroscopy was not a popular technique to be used to profile honey compounds among honey researchers. However, Piernna and colleagues [43] investigated the geographical origin of a total of three hundred and seventy-four Corsican honey harvested in the year 2006 and 2007 from numerous countries in the Mediterranean region. The study focused on the discrimination of the honeys from different geographical origins namely France, Germany, Italy, Ireland and Austria in order to observe the potential of fingerprinting and profiling using FT-Raman spectroscopy. According to FT-Raman band assignments, the chemical information of honey may be due to the saccharides as shown in the region of 200 and 600 cm^{-1} of scattering bands associated the skeletal vibration motions with major contributions from the deformation of C-C-C, C-C-O, C-C and C-O groups of saccharides (Table 2). In addition, by the application of Raman spectroscopy, the minor contributions of other honey constituents such as proteins, unknown carbohydrates or organic acids can be observed in the region of specific wavenumbers (351; 424; 1,077; 1,126; 1,266 and 1,460 cm^{-1}). The date was further analyzed by PCA technique and result obtained did not show a clear pattern of discrimination or any outliers. Other exploratory methods used were PLS-DA and SVM where these models showed correct grouping between 85% and 90%. The

authors promised that the results from combination of Raman spectroscopy and chemometric techniques exhibited non-expensive discrimination of the honey origins. Unfortunately, in this study, they have not identified and characterized the actual compounds presence in the honey samples to be served as biomarker.

Table 2. Scattering bands of honey and their respective functional groups

Wavenumber (cm^{-1})	**Intensity**	**Type of Vibration**
3312-3334	medium	Stretching of O-H
2938-2944	very strong	Asymmetric stretching of CH_2
2900-2907	shoulder	Stretching of C-H
1636-1642	weak	Deformation of O-H of water
1458-1461	strong	Symmetric deformation in the plane of CH_2
1364-1369	medium	Asymmetric deformation in the plane of CH_2
1265-1267	strong	Deformation of C-C-H, O-C-H and C-O-H; vibration of peptide bond
1126-1127	strong	Deformation of C-O-H, vibration of C-N (protein or amino acid)
1077	strong	Stretching of C-O
1064-1069	strong	Stretching of C-O
979	very weak	Unknown
916-918	weak	Deformation of C-H and C-O-H
904	shoulder	Deformation of C-H
865-871, 822	weak	
777	very weak	Deformation of C-H
708-710	weak	
629-630	strong	Ring deformation
590-592	shoulder	Skeletal vibration
519-522	strong	Deformation of C-C-O and C-C-C
449-450	shoulder	Skeletal vibration
420-424	strong	Deformation of C-C-O and C-C-C
351	very weak	Unknown carbohydrates and proteins

Inductively coupled plasma

Inductively coupled plasma (ICP) technique is becoming widely accepted in food analysis for determination of major and trace elements. Chudzinska and Baralkiewicz [46] studied fifteen elements (Al, B, Ba, Ca, Cd, Cr, Cu, K, Mg, Mn, Na, Ni, Pb, Sr and Zn) in each of hundred and forty honey samples of three different types namely honeydew, buckwheat and rape honeys. These honey samples were collected in different areas of Poland between May to November in the years 2006-2007. The data obtained was subjected to LDA technique and revealed 100% discrimination between honey samples. Furthermore, the authors claimed that the values of K and Mn elements are important during the classification prediction and believed they could be markers for honey botanical origin.

A study by Chua et al. [20] constructed the elemental profiles of six honey from Malaysia, namely Tualang (*Koompassia excels*) from Tasik Pedu, Kedah, Gelam (*Malaleuca cajuputi*) from Gubir, Kedah, acacia from Kota Tinggi, Johor and other three forest honeys labeled as forest NS from Kampung Jerjak Seberang, Negeri Sembilan, forest M1 from Kampung Kemuning, Melaka and forest M2 from Kampung Tanjung Rimau Dalam, Melaka using the data obtained from both ICP-AES and ICP-MS. Results showed that the honeys possessed abundant of

potassium and sodium in the range of 69.3-78.6% and 14.1-28.7%, respectively. According to the PCA data obtained, these two elements were claimed to be mineral markers to distinguish the geographical origins of selected honeys. However, no explanation on provenance has been reported clearly as the results exhibited no correlation between the origin of honeys and the type of elements.

A recent study in Malaysia by Shadan et al. [44] exhibited the provenance of stingless bee honey collected from four different geographical regions of Johor, Malaysia (North, South, East and West) represented by five districts namely Segamat, Johor Bahru, Muar, Batu Pahat and Kota Tinggi. The collection period was during less rainy season in May, June and July. Multi-element distribution patterns were established using Inductively Coupled Plasma Optical Emission Spectrometry (ICP-OES) in combination with PCA and LDA techniques. The PCA results demonstrated 87.0% correct classification rate and improved to 96.2% with the use of LDA. The results indicated possible discrimination for the different geographical regions. Until now, this is the first documented data on the provenance establishment of stingless bee honey in Johor, Malaysia. Therefore, our aim is to further studies on the same aspect but in larger scale of the stingless bee honey geographical origins in Malaysia covering four regions namely Kedah, Johor, Selangor and Pahang which are represented by sixteen districts and covering different seasons within a year. Eighteen multi-trace metal ions (Pb, Cu, Mn, Zn, Ni, Cr, Cd, Al, Ba, Se, Ag, As, Co, Be, V, Tl, Th and U) were selected to construct their fingerprinting in identifying the authenticity with the assist of chemometric techniques (PCA and LDA).

Conclusion

Honey is well known as a complex matrix that is very difficult to analyze and characterize. Often the determination of their provenance is complicated because of the incomplete correlation between analytical parameters and honey provenance marker compounds. With globalization technologies, numerous studies have attempted to show the important used of the spectroscopy coupled with the chemometric techniques to exhibit the biomarkers of honeys according to their geographical and botanical origins. For now, it shows a great advantage to solve several problems in this field. Creation of honey databases can be much useful to help identify and quantify the upcoming data in order to monitor unexpected changes or fraud performers based on the established honey fingerprinting. However, issue dealing with the food policies, including legal requirements and standard need to be taken into consideration. Therefore, this field of study needs more time and efforts to improve the capability of technologies and develop databases for the determination of the honey authentication and the provenance according to its unique signatures. In addition, from the review, it can be concluded that the significant to highlight the type of bee species when studying about provenance in order to compare and compile the results from worldwide honey. To date, no comprehensive study in Malaysia has been conducted to establish provenance of bee honeys according to their marker compounds especially from stingless bee species. To the best of our knowledge, only a study by Shadan et al. [44] had documented the provenance establishment of stingless bee honey using multi-element analysis combined with chemometric techniques. Thus, our on-going study will further investigate the geographical origin of stingless bee honey samples specifically from *Heterotrigona itama* species collected from the four regions of Peninsular Malaysia. Concurrently, our study will classify honey samples and find markers of the honey authenticity according to their multi-trace metals concentrations.

Acknowledgment

We are thankful for the advice and assistance provided by Department of Chemistry, Malaysia. We are grateful for the financial support of the work by Universiti Teknologi Malaysia (UTM) (Research University Grant - Grant number QJ130000.2526.12H57).

References

1. Vanhanen, A. Emmertz, L. P. and Savage, P. (2011). Mineral analysis of mono-floral New Zealand honey. *Food Chemistry*, 128(1): 236–240.
2. Escuredo, O., Dobre, I., Fernández-González, M. and Seijo, M. C. (2014). Contribution of botanical origin and sugar composition of honeys on the crystallization phenomenon. *Food Chemistry*, (149): 84–90.
3. Tornuk, F., Karaman, S., Ozturk, I., Toker, O. S., Tastemur, B. and Sagdic, O. (2013). Quality characterization of artisanal and retail turkish blossom honeys: Determination of physicochemical, microbiological, bioactive properties and aroma profile. *Industrial Crops and Products*, (46):124–131.

4. Silva, L. R., Videira, R., Monteiro, A. P., Valentão, P. and Andrade, P. B. (2009). Honey from Luso region (Portugal): Physicochemical characteristics and mineral contents. *Microchemistry Journal*, (93): 73–77.
5. Khalil, M. I., Motallib, M. A., Anisuzzaman, A., Sathi, Z. S., Hye, M. and Shahjahan, M. (2001). Biochemical analysis of different brands of unifloral honey available at the northern region of Bangladesh. *Journal of Medical Science*, (1): 385–388.
6. Chakir, A., Romane, A., Barbagianni, N., Bartoli, D. and Ferrazzi, P. (2011). Major and trace elements in different types of Moroccan honeys. *Australian Journal of Basic and Applied Sciences*, 5(4):223–231.
7. Bogdanov, S., Haldimann, M., Luginbuhl, W. and Gallmann, P. (2007). Minerals in Honey: Environmental, geographical and botanical aspects. *Journal of Apicultural Research*, (46): 269–275.
8. de Ferrer, B. S. G., de Rodriguez, O., Pena, J., Martinez, J., and Moran, M. (2004). Mineral content of the honey produced in Zulia state, Venezuela. *Archivos Latinoamericanos de Nutricion*, 54(3): 346–348.
9. Staniškien, B., Matusevičius, P. and Budreckien, R. (2006). Honey as an indicator of environmental pollution. *Environmental Research, Engineering and Management*, 2(36): 53–58.
10. Mckee, B. (2003). Prevention of residues in honey: A future perspective. *Apiacta*, (38): 173–177.
11. Rashed, M. N. and Soltan, M. E. (2004). Major and trace elements in in different types of Egyptian mono-floral and non-floral bee honeys. *Journal of Food Composition and Analysis*, (17): 725-735.
12. Achudume, A. C. and Nwafor, B. N. (2010). The ecological assessment of metals in local brands of honey in Southwest Nigeria. *African Journal of Agricultural Research*, 5(18):2608-2610.
13. Stankovska, E. Stafilov, T. and Šajn, R. (2008). Monitoring of trace elements in honey from the Republic of Macedonia by atomic absorption spectrometry. *Environmental Monitoring Assessment*, (142): 117-126.
14. Tuzen, M. Silici, S. Mendil, D. and Soylak, M. (2007). Trace element levels in honeys from different regions of Turkey. *Food Chemistry*, 103(2): 325-330.
15. Floryszak-Wieczorek, F. Milczarek, G. Arasimowicz, M. and Ciszewski, A. (2006). Do nitric oxide donors mimic endogenous NO-related response in plants? *Planta*, (224): 1363-1372.
16. Bratu, I. and Georgescu, C. (2005). Chemical contamination of bee honey–identifying sensor of the environment pollution. *Journal of Central European Agriculture*, (6):95–98.
17. Yilmaz, H., and Yavuz, O. (1999). Content of some trace metals in honey from South-Eastern Anatolia. *Food Chemistry*, 4(65):475–476.
18. Przybylowski, P. and Wilczyńska, A. (2001). Honey as an environmental marker. *Food Chemistry*. (74): 3289–3291.
19. Buldini, P. L., Cavalli, S., Mevoli, A. and Sharma, J. L., (2001). Ion chromatographic and voltammetric determination of heavy and transition metals in honey. *Food Chemistry*, (73):487–495.
20. Chua, L. S., Abdul-Rahaman, N. L., Sarmidi, M. R. and Aziz, R. (2012). Multi-elemental composition and physical properties of honey samples from Malaysia. *Food Chemistry*, 135(3): 880–887.
21. Devillers, J., Dore, J., Marenco, M., Poirier-Duchene, F., Galand, N. and Viel, C. (2002). Chemometrical analysis of 18 metallic and non-metallic elements found in honeys sold in France. *Journal of Agricultural Food Chemistry*, (50): 5998–6007.
22. Chudzinska, M. and Barałkiewicz, D. (2011). Application of ICP-MS method of determination of 15 elements in honey with chemometric approach for the verification of their authenticity. *Food and Chemical Toxicology*, 49: 2741–2749.
23. Chudzinska, M., and Barałkiewicz, D. (2010). Estimation of honey authenticity by multi-element characteristics inductively coupled plasma-mass spectrometry (ICP-MS) combined with chemometrics. *Food and Chemical Toxicology*. 48: 284–290.
24. Latorre, M. J. Peña, R. Pita, C. Botana, A. García, S. and Herrero, C. (1999). Chemometric classification of honeys according to their type II. Metal content data. *Food Chemistry*, (66): 263–268.
25. González Paramás, A. M. Bárez, J. A. G. Garcia-Villanova, R. J. Palá, T. R. Albajar, R. A. and Sánchez, J. S. (2000). Geographical discrimination of honeys by using mineral composition and common chemical quality parameters. *Journal of the Science of Food and Agriculture*, (80): 157–165.
26. Baroni, M. V. Arrua, C. Nores, M. L. Fayé, P. del Pilar Díaz, M. Chiabrando, G. A. and Wunderlin, D. A. (2009). Composition of honey from Córdoba (Argentina): Assessment of North/South provenance by chemometrics. *Food Chemistry*, (114): 727–723.
27. Goodacre, R. Broadhurst, D. Smilde, A. Kristal, B. Baker, J. and Beger, R. (2007). Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics*. 3(3): 231–241.

28. Donarski, J. A., Jones, S. A., and Charlton, A. J. (2008). Application of cryoprobe ^{1}H nuclear magnetic resonance spectroscopy and multivariate analysis for the verification of corsican honey. *Journal of Agricultural and Food Chemistry*, 56(14): 5451–5456.
29. Schievano, E., Finotello, C., Mammi, S., Belci, A. I., Colomban, S. and Navarini, L. (2015). Preliminary characterization of monofloral *Coffea* spp. honey: Correlation between potential biomarkers and pollen content. *Journal of Agricultural and Food Chemistry*, (63): 5858–5863.
30. Donarski, J. A., Jones, S. A., Harrison, M., Driffield, M. and Charlton, A. J. (2010). Identification of botanical biomarkers found in corsican honey. *Food Chemistry*, 118(4):987–994.
31. Consonni, R. and Cagliani, L. R. (2008). Geographical characterization of polyfloral and acacia honeys by nuclear magnetic resonance and chemometrics. *Journal of Agricultural and Food Chemistry*, 56(16): 6873–6880.
32. Beretta, G., Caneva, E., Regazzoni, L., Bakhtyari, N. G. and Maffei F. R. (2008). A solid-phase extraction procedure coupled to ^{1}H NMR, with chemometric analysis, to seek reliable markers of the botanical origin of honey. *Analytica Chimica Acta*, 620(1–2): 176–182.
33. Lolli, M. Bertelli, D. Plessi, M. Sabatini, A.G. and Restani, C. (2008). Classification of Italian honeys by 2D HR-NMR. *Journal of Agricultural and Food Chemistry*, 56(4): 1298–1304.
34. Schievano, E., Peggion, E. and Mammi, S. (2010). ^{1}H Nuclear magnetic resonance spectra of chloroform extracts of honey for chemometric determination of its botanical origin. *Journal of Agricultural and Food Chemistry*, 58(1):57–65.
35. Schievano, E., Stocchero, M., Morelato, E., Facchin, C. and Mammi, S. (2012). An NMR-based metabolomic approach to identify the botanical origin of honey. *Metabolomics*, 8(4):679–690.
36. Ohmenhaeuser, M., Monakhova, Y. B., Kuballa, T. and Lachenmeier, D. W. (2013). Qualitative and quantitative control of honeys using NMR spectroscopy and chemometrics. *ISRN Analytical Chemistry*, 2013: 1-9.
37. Aliferis, K. A., Tarantilis, P. A., Harizanis, P. C. and Alissandrakis, E. (2010). Botanical discrimination and classification of honey samples applying gas chromatography/mass spectrometry fingerprinting of headspace volatile compounds. *Food Chemistry*, 121(3): 856–862.
38. Castro-Vázquez, L., Díaz-Maroto, M. C., González-Viñas, M. A. and Pérez-Coello, M. S. (2009). Differentiation of monofloral citrus, rosemary, eucalyptus, lavender, thyme and heather honeys based on volatile composition and sensory descriptive analysis. *Food Chemistry*, 112(4):1022–1030.
39. Cajka, T., Hajslova, J., Pudil, F. and Riddellova, K. (2009). Traceability of honey origin based on volatiles pattern processing by artificial neural networks. *Journal of Chromatography A*, 1216(9):1458–1462.
40. Stanimirova, I., Üstünb, B., Cajka, T., Riddelova, K., Hajslova, J., and Buydens, L. M. C. (2010). Tracing the geographical origin of honeys based on volatile compounds profiles assessment using pattern recognition techniques. *Food Chemistry*, 118(1):171–176.
41. Agila, A. and Barringer, S. (2012). Application of selected ion flow tube mass spectrometry coupled with chemometrics to study the effect of location and botanical origin on volatile profile of unifloral american honeys. *Journal of Food Science*, 77(10):1103–1108.
42. Woodcock, T., Downey, G. and O'Donnell, C. P. (2009). Near infrared spectral fingerprinting for confirmation of claimed PDO provenance of honey. *Food Chemistry*, 114(2): 742–746.
43. Piernna, J. A. Abbas, O. Dardenne, P. and Baeten, V. (2011). Discrimination of corsican honey by FT-raman spectroscopy and chemometrics. *Biotechnology, Agronomy, Society and Environment*, 15(1): 75–84.
44. Shadan, A. F. Mahat, N. A., Wan Ibrahim, W. A., Ariffin, Z. and Ismail, D. (2017). Provenance establishment of stingless bee honey using multi-element analysis in combination with chemometrics techniques. *Journal of Forensic Sciences*, 63(1): 80–85.
45. Azevedo, M. S., Seraglio, S. K. T., Rocha, G., Balders, C. B., Piovezan, M., Gonzaga, L. V., Falkenberg, D. B., Fett, R., and Costa, A. C. O. (2017). Free amino acid determination by GC-MS combined with a chemometric approach for geographical classification of bracatinga honeydew honey (*Mimosa scabrella* Bentham). *Food Control*, 78: 383–392.
46. Chudzinska, M. and Baralkiewicz, D. (2011). Application of ICP-MS method of determination of 15 elements in honey with chemometric approach for the verification of their authenticity. *Food and Chemical Toxicology*, 49(11): 2741–2749.