The discrimination of honey origin using melissopalynology and Raman spectroscopy techniques coupled with multivariate analysis

The discrimination of honey origin using melissopalynology and Raman spectroscopy techniques coupled with multivariate analysis

Francesca Corvucci [a,*], Lara Nobili [b], Dora Melucci [b], Francesca-Vittoria Grillenzoni [a]

[a] CRA-API Agricultural Research Council, Honeybee and Silkworm Research Unit, Via di Saliceto, 80, 40128 Bologna, BO, Italy
[b] Department of Chemistry "Ciamician", University of Bologna, Via Selmi 2, I40126 Bologna, Italy

ARTICLE INFO

Article history:
Received 20 April 2014
Received in revised form 27 June 2014
Accepted 25 July 2014
Available online 6 August 2014

Keywords:
Honey
Melissopalynology
Pollen analysis
Chemometrics
PCA
FT-microRaman
Botanical origin
Geographical origin

ABSTRACT

Honey traceability to food quality is required by consumers and food control institutions. Melissopalynologists traditionally use percentages of nectariferous pollens to discriminate the botanical origin and the entire pollen spectrum (presence/absence, type and quantities and association of some pollen types) to determinate the geographical origin of honeys. To improve melissopalynological routine analysis, principal components analysis (PCA) was used. A remarkable and innovative result was that the most significant pollens for the traditional discrimination of the botanical and geographical origin of honeys were the same as those individuated with the chemometric model. The reliability of assignments of samples to honey classes was estimated through explained variance (85%). This confirms that the chemometric model properly describes the melissopalynological data. With the aim to improve honey discrimination, FT-microRaman spectrography and multivariate analysis were also applied. Well performing PCA models and good agreement with known classes were achieved. Encouraging results were obtained for botanical discrimination.

1. Introduction

Council Directive 2001/110 EC states that the term "honey" shall be applied only to the natural sweet substances produced by *Apis mellifera* bees from the nectar of plants, which the bees collect and transform by using specific substances to ripen and mature. Supplemental information concerning the floral or vegetal and the geographical origin can be given if samples come entirely from the indicated source.

Honey contains pollen grains and other microscopic particles handled by the bees from the targetted plants. Its composition reflects the vegetation types that form the honey matrix and is useful for the determination of its botanical and geographical origin. The number of pollen grains in honey is influenced by plant morphology and physiology (Todd & Vansell, 1942), the action of foraging bees (Molan, 1998), contamination in the hive and during the uncapping and processing. Honey analysis is a reasonable method for authenticating honey origins and characteristics; it also helps to determine and control the botanical origin of honey (Von der Ohe, Persano Oddo, Piana, Morlot, & Martin, 2004).

In order to establish representative ranges of some relevant parameters suitable for determining botanical and geographical origins of honey, an alternative analytical method is needed to evaluate the physicochemical components of honey, as described in the previous studies (Anupama, Bath, & Sapna, 2003; Conti, Stripeikis, Campanella, Cucina, & Tudino, 2007; Corbella & Cozzolino, 2006; Devillers, Morlot, Pham-Delègue, & Doré, 2004; Escriche, Kadar, Juan-Borras, & Domenech, 2014; Krauze & Zalewski, 1991; Latorre et al., 1999; Lopez et al., 1996; Persano Oddo, Piazza, Sabatini, & Accorti, 1995; Popek, 2002; Serrano, Villarejo, Espejo, & Jodral, 2004; Soria, Gonzales, De Lorenzo, Martinez-Castro, & Sanz, 2004); these involve phenolic acids and polyphenols (Andrade, Ferreres, Gil, & Tomas-Berberan, 1997; Ferreres, Andrade, & Tomas-Barberan, 1994; Tomas-Berberan, Martos, Ferreres, Radovic, & Anklam, 2001), volatiles (Guyot-Declerck, Resons, Bouseta, & Collin, 2002; Perez, Sanchez-Brunette, Calvo, & Tadeo, 2002; Soria et al., 2004) and amino acids and proteins (Cometto, Faye, Naranjo, Rubio, & Aldao, 2003; Davies & Harris, 1982; Pirini, Conte, Francioso, & Lecker, 1992; Siede, Schmidt, & Buchler, 2004). All of these methods involve a large number of parameters, and data treatments are performed by various statistical methods, such as: analysis of variance (Conti et al., 2007), canonical analysis (Gallego Pica & Fernandez Hernando, 2013, chap. 20), principal component analysis (PCA) and multivariate analysis (Devillers et al., 2004), with the purpose of selecting

* Corresponding author. Tel.: +39 051353103; fax: +39 051356361.
E-mail address: francesca.corvucci@entecra.it (F. Corvucci).

the factors for discriminating the studied varieties. The limitations of these methods are underlined by their inability to classify all honeys according to individual types and varieties. Therefore, the aim of this study was to discriminate the botanical origin of honeys, using fast and rapid analytical tool (Raman Spectroscopy) coupled with chemometric analysis.

2. Materials and methods

2.1. Sampling

The pollen spectra of 184 samples, acquired in 2013 from different botanical and geographical origins, were collected, including 35 multifloral (27 samples from Italy, 6 from east Europe and 2 from Argentina), 93 acacia (47 from Italy, 46 from east Europe), 31 citrus (27 from Italy, 4 from Spain) and 25 chestnut (all from Italy).

Then the Raman spectra of 308 unifloral honeys were collected (122 acacia samples, constituted of 91 samples from Italy and 31 from eastern Europe; 44 citrus samples containing 40 samples from Italy and 4 from Spain; 33 chestnut, 24 sunflower, 39 lime and 46 honeydew honey all from Italy).

Acquired spectra are shown in Fig. 1.

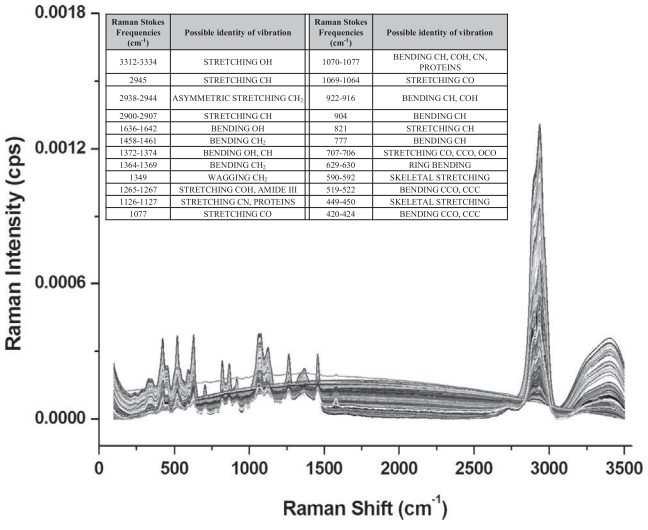
2.2. Traditional pollen analysis (melissopalynology)

Melissopalynology has been traditionally used for ascertaining the botanical and geographical origins of honey (Soria et al., 2004). Moreover, pollen analysis provides some important information about honey composition and its organoleptic properties (Von der Ohe et al., 2004).

Honey always includes pollen grains (PG), mainly from the plant species foraged by bees, and honeydew elements (HE) that provide a good fingerprint of the environment (where honey comes from).

Due to the high sugar content and viscosity of honey, pollens are extracted by centrifugation of a solution of honey and distilled water, as described in the harmonized methods of melissopalynology (Von der Ohe et al., 2004). In order to eliminate most of the sugars present in honey and to recover most of the pollen, samples were first diluted in 40 ml, then the sediment was stirred with 12 ml of distilled water and centrifuged for 15 min and 7 min, respectively.

The sediment was spread over a slide, drawing a square about 1 cm^2, and covered with an 18 × 18 mm coverslip.



Raman Stokes Frequencies (cm^{-1})	Possible identity of vibration	Raman Stokes Frequencies (cm^{-1})	Possible identity of vibration
3312-3334	STRETCHING OH	1070-1077	BENDING CH, COH, CN, PROTEINS
2945	STRETCHING CH	1069-1064	STRETCHING CO
2938-2944	ASYMMETRIC STRETCHING CH_2	922-916	BENDING CH, COH
2900-2907	STRETCHING CH	904	BENDING CH
1636-1642	BENDING OH	821	STRETCHING CH
1458-1461	BENDING CH_2	777	BENDING CH
1372-1374	BENDING OH, CH	707-706	STRETCHING CO, CCO, OCO
1364-1369	BENDING CH_2	629-630	RING BENDING
1349	WAGGING CH_2	590-592	SKELETAL STRETCHING
1265-1267	STRETCHING COH, AMIDE III	519-522	BENDING CCO, CCC
1126-1127	STRETCHING CN, PROTEINS	449-450	SKELETAL STRETCHING
1077	STRETCHING CO	420-424	BENDING CCO, CCC

Fig. 1. FT-microRaman spectra and assignment of principal frequencies.

Qualitative pollen analysis is based on the recognition and counting of PG and HE in honey samples. Pollens are identified and listed by the scientific name of the botanical group (*species*, *genus* or *family*), depending on how reliable the identification may be (Persano Oddo & Ricciardelli D'Albore, 1989); otherwise, a term such as *group*, *form* or *type,* to indicate a larger taxonomic group, is added to the scientific name.

The optical microscope used was a Zeiss, *Axiolab E re*, lens with magnification of 400× or 1000×. For every sample 500 (or 1000 in honey with overrepresented pollen, e.g. chestnut) nectariferous PG were counted, and the percentages of pollen types were determined. In the case of non nectariferous PG, they are just listed and considered as 1 (presence or absence).

The determination of the botanical origin is based on the relative frequencies of nectariferous species' pollen types. The frequency classes of pollen grains were given as predominant (>45%), secondary pollen (15–45%), important minor pollen (3–15%) and minor pollen (1–3%).

Generally a honey can be defined as unifloral if the "characteristic" pollen (e.g. *Brassica* in rape honey) exceeds 45%. It is considered honeydew if the ratio "HE/PG" exceeds 3. These are general guidelines but many pollen types are underrepresented (*Robinia pseudoacacia*, *Citrus* spp., *Tilia* spp.) or overrepresented (*Castanea sativa*, *Eucaliptus* spp.). For instance, to characterize acacia honey as unifloral, *R. pseudoacacia* pollen must be over 15%, citrus must have at least 10% of *Citrus* spp. pollen while, for chestnut honey, a content of 90% of *Castanea* pollen is required to classify honey as unifloral (Sabatini, Bortolotti, & Marcazan, 2007).

Particular caution in the interpretation of melissopalynological results is required due to the very different levels of abundance of a given pollen type in the nectar and also due to other sources of variability, such as secondary (inclusion of pollens inside the hive), tertiary (inclusion of pollens during the extraction process of the honey) and quaternary (aerial contamination) enrichment.

The determination of geographical origin is based on the entire pollen spectrum, which must be consistent with the flora of a particular region and/or with any reference pollen spectrum or description in the literature.

2.3. Fourier transform – microRaman spectroscopy

Raman spectroscopy has only recently been investigated as a potential tool for food quality control and compositional identification (Goodacre, Radovic, & Anklam, 2002). It is an analytical tool based on vibrational transitions within functional groups; the sample is illuminated by a particular wavelength of incident laser light, and the diffused light is analysed.

Spectra of the samples were obtained using a Thermo Scientific DXR Fourier Transform (FT) microRaman instrument equipped with one excitation laser: 532 nm; the spectrograph allows a resolution of 5 cm^{-1} (FWHM) using 900 lines/mm. Each sample was analysed, using a laser power at the sample of 10 mW for the green, an aperture of 25 μm pinhole and a 20× objective.

A drop of every sample was placed on a glass slide; the visual focus was correctly identified through the microscope, and then the spectrum was collected. A spectrum from each sample was collected for 5 min, using the continuous extended scan from 200 to 3500 cm^{-1}.

Possible vibrational modes (Fig. 1) were assigned to significant frequencies (Özbalci, Boyaci, Topcu, Kadilar, & Tamer, 2013; Fernández Pierna et al., 2010; Goodacre et al., 2002).

2.4. Chemometrics

To process melissopalynological data and Raman spectra, PCA was applied (Vandeginste et al., 1998; Wold, Esbensen, & Geladi,

1987). Principal component analysis has been chosen for its capability to explore and model experimental data, evaluating variables relevance and correlations, and producing easy-to-read bidimensional or 3D-plots in which samples (also indicated as objects) and variables are visualized (scores plot and loadings plot, respectively). So, even starting from matrices of experimental data which may count tens or hundreds of rows and columns, it is possible to observe samples projected in a particular plane or 3D-space: that plane or that 3D-space in which the quantity of useful analytical information is maximized. This new space is obtained by simply rotating the original space, looking for the directions of maximum variance. Variance is an estimation of quantity of information. The scores are the coordinates of objects in this new, special space and the principal components (PCs) are the eigenvectors of that space; the relevant eigenvalues are proportional to the percentage of variance (Explained Variance, EV). The compared analysis of scores plots and loadings plots allows us to study similarities among objects, relevance of variables (which is proportional to loadings), relative relevance of variables (loadings close to each other are an indication of similar relevance), links between objects and those variables which influence their properties.

Since PCA is based on evaluation and comparison of variances, it is important to verify that variance values are comparable. Eventual data-pretreatment should be considered. Moreover, the analyst should choose (with great care) the standard samples in order to be highly confident of the correspondence between declared standard sample properties and their actual properties. Otherwise the model achieved will not describe the real system.

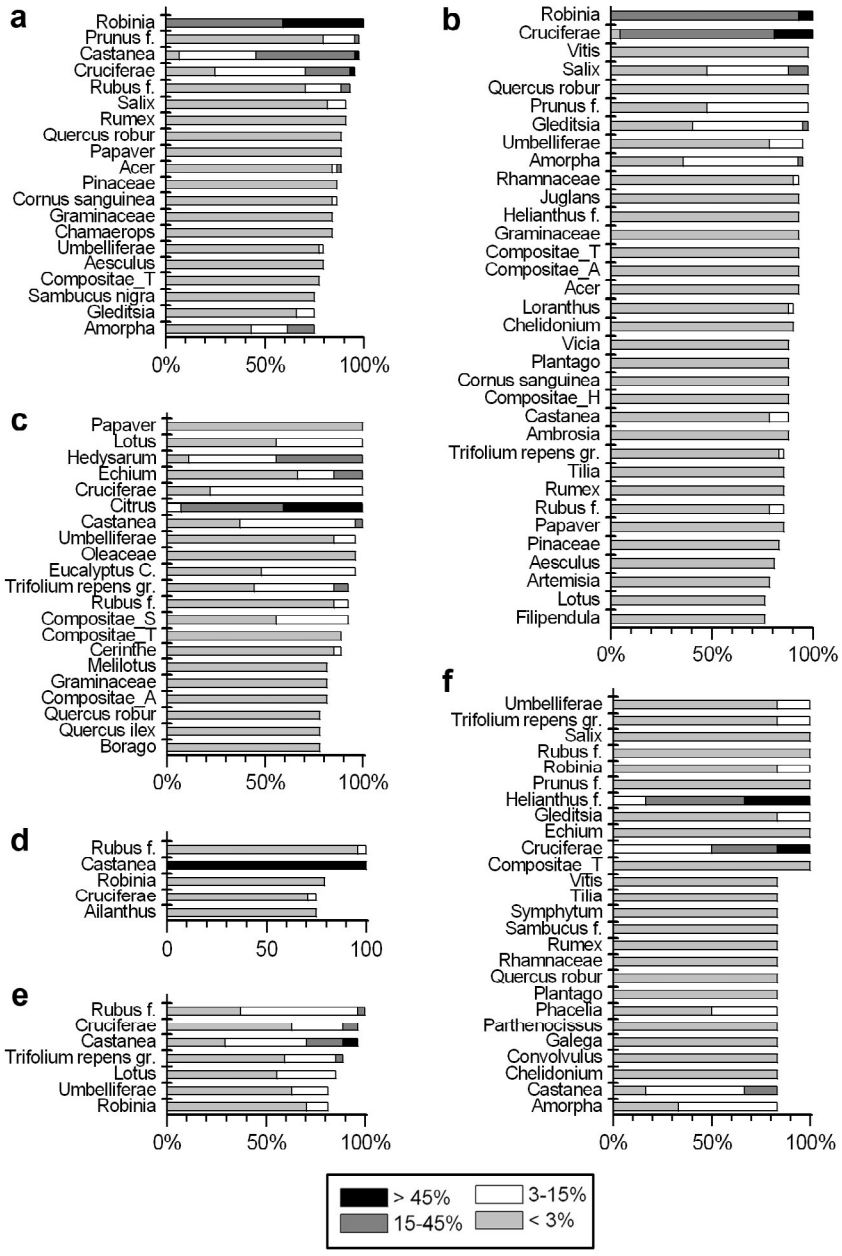


Fig. 2. Pollen spectra of Italian acacia honeys (a), east European acacia honeys (b), Italian citrus honeys (c), Italian chestnut honeys (d), Italian multifloral honeys (e) and east European multifloral honeys (f). Only pollen taxa found in more than 75% of the honey samples are reported. Data show absolute frequency (percent of analyzed honeys in which each pollen type was found) distributed into four classes, indicating the relative quantity of each pollen type.

Since explained variance is necessarily lower than 100% in order to reduce the complexity of the problem, a consequent drawback of PCA lies in the necessity to lose part of the original information.

In the case of melissopalynological data, no scaling was necessary. In the case of the Raman spectra here processed, autoscaling and area normalisation were applied, respectively, to account for the possible different orders of magnitude of measured variables and to avoid strong peaks covering the weak ones, while baseline correction was applied to correct the fluorescence effect and to reduce spectral noise and background effects.

To create PCA models, all objects in datasets were used as a training set. Models were validated by full cross-validation (leave-one-out) to evaluate the performance and robustness of the developed models.

Multivariate data processing was performed by the software The Unscrambler® v. 9.8 (CAMO Software, Oslo, Norway).

3. Results and discussion

3.1. Traditional pollen analysis

3.1.1. Acacia honeys

The characterizing pollen is *R. pseudoacacia.* Seven samples were excluded because the percentage of *R. pseudoacacia* pollen was under 15% (Beckh & Camps, 2009); the mean value of this pollen in selected samples (n = 86) was 35% (min 15%–max 88%). This value was higher than that reported in 2004 which was 28.1% (min 7.0% max 59.5%) (Persano Oddo & Piro, 2004).

The samples from Italy were 44 and 42 were from east Europe. In Italian honey samples 148 pollen types were found. *R. Pseudoacacia* was found in all the samples, in 41% as predominant pollen (>45%) with *Castanea* and Cruciferae (2% of samples), and in 59% as secondary pollen. The most frequent pollens (found in over 75% of samples) are represented in Fig. 2a.

The number of different pollen types found in eastern Europe honey samples was 148. *R. Pseudoacacia* was found as predominant pollen in 7%, and as secondary pollen in 92%: these values are lower than those in the Italian acacia honeys.

East European acacia honeys are characterized by Cruciferae, *Crataegus*, *Amorpha*, *Symphytum*, *Gleditsia*, *Chelidonium*, *Vicia*, *Cornus sanguinea*, *Trifolium* spp., *Helianthus*, *Phacelia*, *Loranthus*, *Centaurea cyanus*, *Vitis* (Persano Oddo, Piana, & Ricciardelli D'Albore, 2007).

The most frequent pollens (over 75% of samples) found in the samples are shown in Fig. 2b. Cruciferae and *Salix* were present in higher percentage than in Italian honeys while *Castanea* was present in lower percentage than in Italian honeys.

Amorpha was found in higher percentage than in Italian honeys. Cruciferae, *Gleditsia*, *Helianthus*, *Loranthus*, *Cornus sanguinea*, *Vicia, Vitis*, *Chelidonium* and *Amorpha* were present in more than 75% of samples; they were considered as markers of geographical origin, as reported in the literature (Persano Oddo et al., 2007). These markers of geographical origin, generally, are pollens, found in a

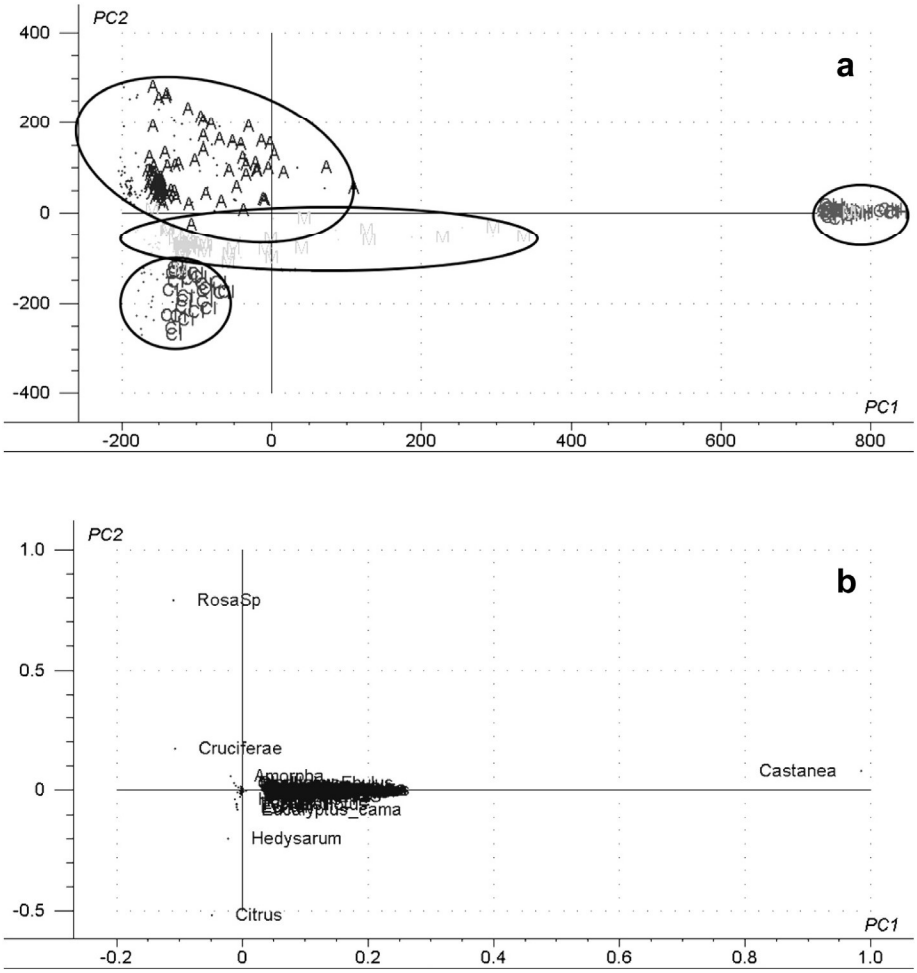


Fig. 3. Scores (a) and loadings (b) plots of the PCA model achieved on 175 pollen spectra of honeys with different botanical origins. The scores are plotted in the first two principal components (total explained variance 85%) and labelled according to their botanical origin. The codes are: A – Acacia, CH – chestnut, CI – citrus, M – multifloral.

lower frequency or percentage, because the most frequent pollens are often widespread.

Moreover, *Hedysarum* (marker pollen of Italian origin) was found in only 54% of Italian samples (15% as minor pollen, 5% as important minor pollen, 3% as secondary, and never as predominant). Thus, the absence of *Hedysarum* allows *Chamaerops* or *Castanea* to be selected as suitable for determining the Italian origin.

Also, for east European honeys, the typical association (*Phacelia*, *Loranthus*, and *Symphytum*) was not found in all samples but the presence of *Eleagnus* or *Centaurea cyanus,* in lower percentage, is considered as a marker of origin, as well.

3.1.2. Citrus honeys

The characterizing pollen is *Citrus* spp. All citrus samples ($n = 31$) were characterized as unifloral. The mean value of *Citrus* spp. was 38% (range from 14% to 76%), and this is higher than the value reported in 2004 which was 18.6% (min 2.3% max 42.2%) (Persano Oddo & Piro, 2004).

The Italian citrus honeys were 27 and 4 samples were from Spain. In Italian honey samples 124 pollen types were found. *Citrus* spp. pollen was found to be predominant in 40% of samples, higher than values found in 2002, as secondary pollen in 51% of samples, similar to the values found in 2002, and an important minor pollen in 7% of samples, but lower than the values found in 2002 (Grillenzoni, Anfossi, Marcazzan, & Sabatini, 2002).

The most frequent pollens (Fig. 2c) *Echium*, *Hedysarum*, *Lotus*, Compositae S, *Rubus* f., *Trifolium repens* gr., *Papaver*, Oleaceae were present; they are markers of geographical origin, as reported in the literature (Grillenzoni et al., 2002). Also *Eucalyptus camaldulensis, Cerithe*, *Borago*, and the non-nectariferous *Quercus ilex* pollens were present as markers of Mediterranean origin (Grillenzoni et al., 2002).

The typical Italian association of geographical origin is *Citrus* spp. and *Hedysarum*, which distinguishes Italian citrus honey from those of other countries.

In Spanish honey samples, 188 pollen types were found. *Citrus* spp. pollen was never found to be predominant, while it was present as secondary pollen in 4% of samples. Since Spanish apicultural techniques are different, the citrus honeys are quantitatively and qualitatively richer than Italian ones. Since the number of Spanish samples was very low ($n = 4$), it is meaningless to list the most frequent pollens.

3.1.3. Chestnut honeys

The characterizing pollen is *Castanea.* Two samples were excluded because the percentage of *Castanea* pollen was under 90% (Sabatini et al., 2007). The mean value of this pollen in selected samples ($n = 23$) was 96% (min 90%–max 100%). This value is higher than that reported in 2004, which was 94.5% (min 85.6%–max 100%) (Persano Oddo & Piro, 2004).

All samples ($n = 23$) were from Italy; 112 pollen types were found. *Castanea* was always found in all the samples as predominant pollen (>45%).

The most frequent pollens (found in over 75% of samples) are shown in Fig. 2d. They are *Rubus* f., *Robinia*, *Ailanthus*, Cruciferae, as reported in the literature (Grillenzoni & Ferro, 2008).

3.1.4. Multifloral honeys

Multifloral honeys do not have a typical pollen. When it is not possible to define the sample as unifloral because the honey does

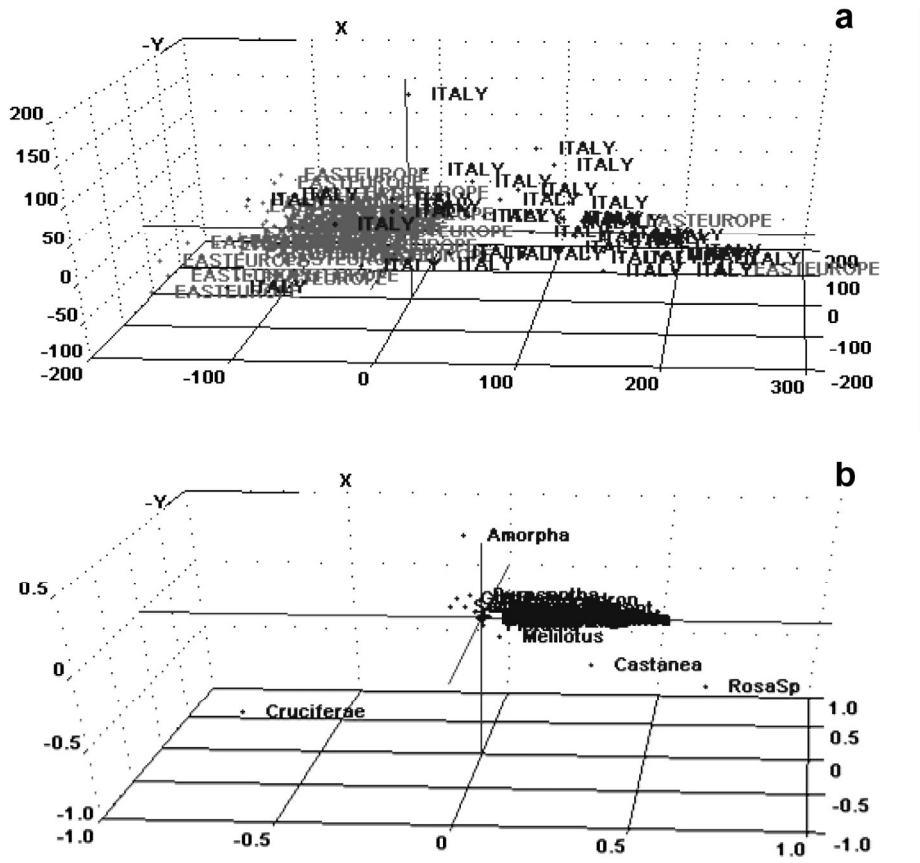


Fig. 4. Scores (a) and loadings (b) plots of the PCA model achieved on the acacia honeys. The scores are plotted in the first three principal components (total explained variance is 84%) and labelled according to their geographical origin.

not come either mainly from a single nectar source or from honeydew, honey is labelled as multifloral.

One sample was excluded because it was honeydew.

In Italian honeys, very common pollens are *Castanea*, Cruciferae, Graminaceae, *Rubus* f., *Salix*, *Trifolium pratense* s.l., *Trifolium repens*, *Malus*, *Pyrus* f., *Prunus* f., Compositae forma T, *Lotus*, *Papaver* and *Plantago* (Persano Oddo et al., 2007).

In Italian honey samples, 145 pollen types were found. Fig. 2e shows that *Rubus* f. was found in all the samples, the most frequent nectariferous pollens (found in over 75% of samples) were *Castanea* (as predominant pollen in 7%), Cruciferae, *Trifolium repens* gr., *Lotus*, according to the literature (Persano Oddo et al., 2007). *Hedysarum*, the marker of Italian honeys, was found in 7% as predominant and in 7% as secondary pollen.

In east European honey samples, 118 pollen types were found.

East European honeys are characterized by Robinia, Cruciferae, Crataegus, Amorpha, Symphytum, Gleditsia, Chelidonium, Vicia, Cornus sanguinea, Trifolium spp., Helianthus, Phacelia, Loranthus, Centaurea cyanus and Vitis (Ricciardelli D'Albore, 1997).

In all samples, Cruciferae (as predominant in 17%), *Gleditsia*, *Helianthus* (as predominant in 33%), *Robinia*, *Trifolium repens* gr. were considered as markers of geographical origin, as reported in the literature (Ricciardelli D'Albore, 1997) but many other pollens were found in all samples, suggesting further characterization for east European honeys.

The most frequent pollens (found in over 75% of samples) are shown in Fig. 2f. *Phacelia, Loranthus, Symphytum* and *Chelidonium* were considered as markers of geographical origin, as reported by Piana (Piana, 1997), often in association with *Tilia*.

In Argentine honey samples, 75 pollen types were found. It is meaningless to list the most frequent pollens, since the number of Argentine samples was very low ($n = 2$).

3.2. Traditional pollen analysis integrated by multivariate analysis

To create the dataset for multivariate analysis, 175 samples were selected among melissopalynological ones, based on ascertained compliance with botanical origin. The number of pollens belonging to the 203 different pollen types and honeydew elements, starch granule, amorphous and crystalline materials were the variables.

At the beginning, a PCA model was created, taking all objects as the training set. From the scores plot of this model, one outlier sample could be identified (east European acacia honey) and excluded from the elaboration. This outlier was characterized by a very high quantity of spores.

A new PCA model was created and the scores plot relevant to the first two PCs is shown in Fig. 3a. This figure exhibits the data-forming clusters corresponding to botanical classes, with explained variance of 85% (76% on the PC1, and 9% on the PC2). The same scores plot on form 3D, relevant to the first three PCs, indicated an explained variance of 90% (76% on the PC1, 9% on the PC2, and 5% on the PC3). The graphical model displayed in Fig. 3a allowed us to observe all the studied samples simultaneously, and clusters correspond to possible classes. In particular, it is possible to visually discriminate between unifloral and multifloral samples and to quantify the reliability of conclusions by

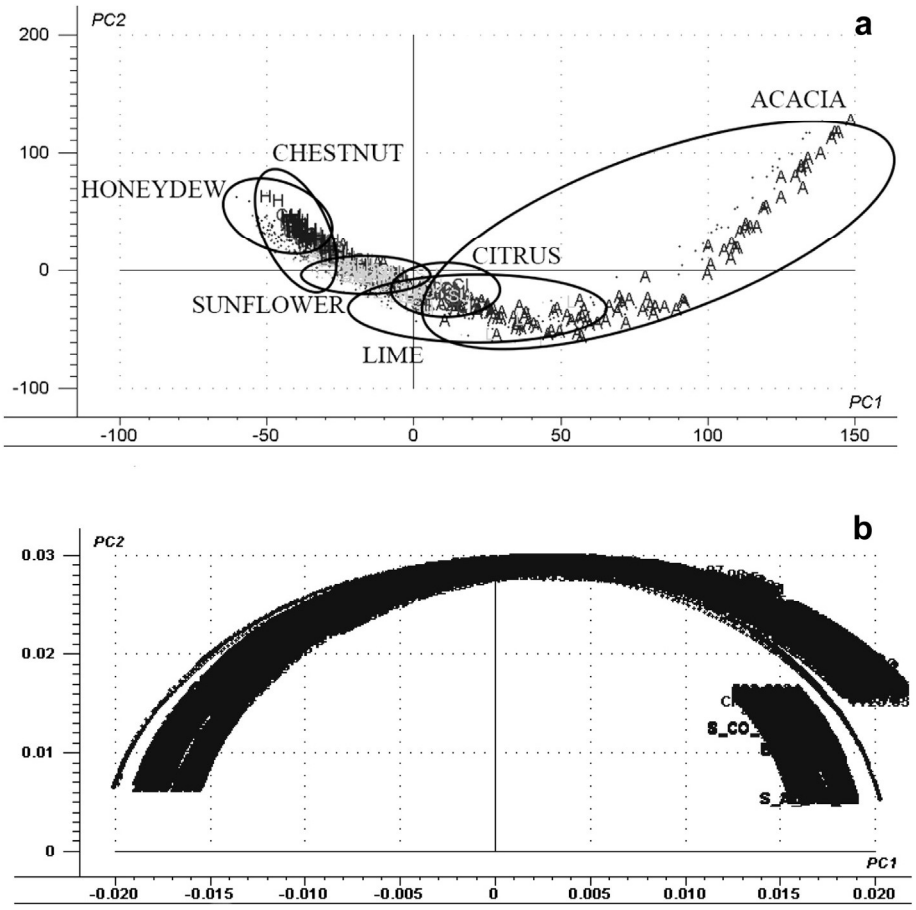


Fig. 5. Scores (a) and loadings (b) plots of the PCA model achieved on 308 FT-microRaman spectra of honeys with different botanical origins. The scores are plotted in the first two principal components (total explained variance 98%) and labelled according to their botanical origins. The codes are: A – Acacia, CH – chestnut, CI – citrus, H – honeydew, L – lime, SU – sunflower.

explained variance. This discrimination is the main issue in the authentication of honey.

The loadings plot corresponding to Fig. 3a is shown in Fig. 3b and allowed us to identify the most significant variables of the systems, including *Castanea, Citrus,* Cruciferae, *Rosa* spp. It helped to quantify and compare their relevance by the numerical values of loadings. The perfect correspondence between classes and characterising pollens was confirmed. For instance, chestnut samples and the relevant *Castanea* pollen were found at high positive PC1 values; the citrus samples and the relevant *Citrus* pollen were found at high negative PC2 values; the acacia samples and the relevant Cruciferae and *Rosa* spp. pollens were found at high positive PC2 values; multifloral samples did not show any specific pollen type, but they were very well separated with respect to all the other classes. In general, high graphical selectivity was shown in the separation between classes.

As a second step, individual PCA models relevant to botanical categories were created. These individual models allowed us to also achieve a geographical discrimination of samples and find significant variables.

The acacia model (Fig. 4a and b) discriminated between Italian and east European samples, particularly *Castanea* (for Italian honeys) and Cruciferae (for east European honeys) pollens were found to be the most significant ones (explained variance was 84%; 61% on the PC1, 17% on the PC2, and 6% on the PC3).

The citrus model discriminated between Italian and Spanish samples, particularly *Citrus* and *Hedysarum* pollens were found to be the most significant for Italian honeys (explained variance was 78%; 50% on the PC1, 19% on the PC2, and 9% on the PC3).

The multifloral model discriminated between Italian, east European and Argentine honeys. In particular, *Hedysarum* and *Castanea* pollens, for Italian honeys, and Cruciferae and *Helianthus* pollens, for east European honeys, were found to be the most significant ones (explained variance was 73%; 58% on the PC1, 9% on the PC2, and 6% on the PC3).

3.3. FT-microRaman spectroscopy

For multivariate analysis, a data matrix of FT-Raman Stokes spectra was prepared.

The 308 samples were considered as objects for the training set.

Before the reduction of the multidimensional Raman data by chemometric elaboration, all spectra were pre-treated by area normalization and baseline correction. These transformations are required to correct for fluorescence effect and to reduce spectral noise and background effects. Moreover, all samples were scaled with the function autoscaling to avoid strong peaks covering the weak ones.

At the beginning, the PCA model for all pollen samples was created. In the scores plot (Fig. 5a) of this model, one outlier could be identified (Italian lime honey) and excluded from further

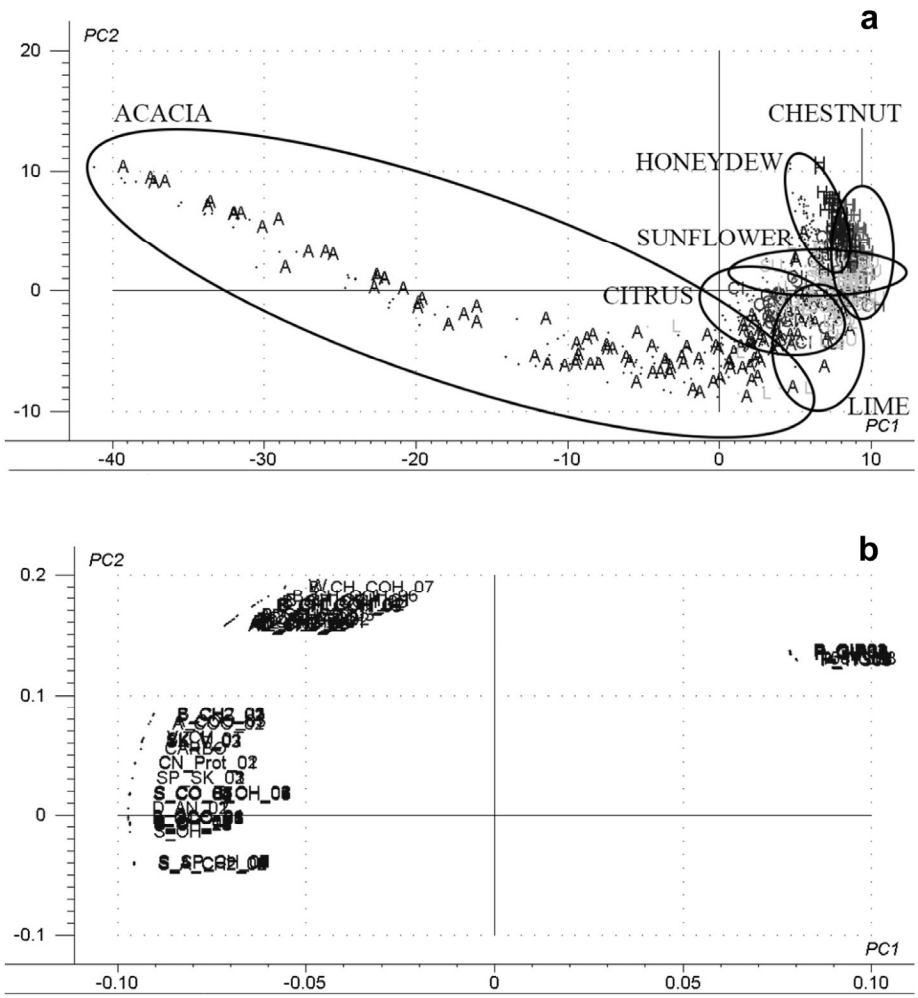


Fig. 6. Scores (a) and Loadings (b) plots of the PCA model obtained by the variables reduction on the FT-Raman spectra. The scores are plotted in the first two principal components (total explained variance 99%) and labelled according to their botanical origins. The codes are: A – acacia, CH – chestnut, CI – citrus, H – honeydew, L – lime, SU – sunflower.

elaboration (explained variance was 99%; 67% on the PC1, and 32% on the PC2). In the scores plot, data were structured according to their botanical origin, but the classes were found to be less separated than those of models reported in Fig. 3a. This result gives us encouragement to proceed with chemometric modelling.

The high value of explained variance on the first PC and the structure of the loadings plot (Fig. 5b) underline the useless information contained in these spectroscopic data. For that reason, the PCA analysis was repeated on a reduced variable set. The reduction was made by considering only the variables with a loading higher than 0.02 (absolute value), and then picking only the most chemically significant frequencies reported. This operation gave a better separation (Fig. 6a and b) among the botanical classes (explained variance was 100%; 85% on the PC1, and 15% on the PC2).

4. Conclusion

The melissopalynologic chemometric model provides an important tool of comparison to test the rapid and non-invasive FT-microRaman technique. The obtained results demonstrated that Raman data, in combination with proper PCA models, could be successfully adopted to identify the botanical origins.

To improve the discrimination, higher numbers of known samples should be collected. To improve the quality of the signal, and then the discrimination among different samples, it is better to register the FT-Raman spectra with the fluorescence correction.

The application of the Raman technique to honey origin discrimination is important because it is a simple, rapid, and non-destructive method.

Acknowledgments

Investigation was supported by Institutional Funds for Selected Research Topics (University of Bologna and Agricultural Research Council, Italy).

References

Andrade, P., Ferreres, F., Gil, M. I., & Tomas-Berberan, F. A. (1997). Determination of phenolic compounds in honey with different floral origin by capillary zone electrophoresis. *Food Chemistry, 60*, 79–84.

Anupama, D., Bath, K. K., & Sapna, V. K. (2003). Sensory and physic-chemical properties of commercial samples of honey. *Food Research International, 36*, 183–191.

Beckh, G., & Camps, G. (2009). Neue Spezifikationen für Trachthonige. *Deutsche Lebensmittel – Rundschau, 105*(2), 105–110.

Cometto, P. M., Faye, P. F., Naranjo, R. D. D., Rubio, M. A., & Aldao, M. A. J. (2003). Comparison of free amino acids profile in honey from three Argentinian regions. *Journal of Agricultural and Food Chemistry, 51*, 5079–5087.

Conti, M. E., Stripeikis, J., Campanella, L., Cucina, D., & Tudino, M. B. (2007). Characterization of Italian honeys (Marche Region) on the basis of their mineral content and some typical quality parameters. *Chemistry Central Journal, 1*, 14.

Corbella, E., & Cozzolino, D. (2006). Classification of the floral origin of Uruguayan honeys by chemical and physical characteristics combined with chemometrics. *LWT – Food Science and Technology, 39*, 534–539.

Davies, A. M. C., & Harris, R. G. (1982). Free amino acid analysis of honeys from England and Wales: Application to the determination of geographical origin of honeys. *Journal of Apicultural Research, 21*, 168–173.

Devillers, J., Morlot, M., Pham-Delègue, M. H., & Doré, J. C. (2004). Classification of monofloral honeys based on their quality control data. *Food Chemistry, 86*, 305–312.

Escriche, I., Kadar, M., Juan-Borras, M., & Domenech, E. (2014). Suitability of antioxidant capacity, flavonoids and phenolic acids for floral authentication of honey. Impact of industrial thermal treatment. *Food Chemistry, 142*, 135–143.

European Commission (2002). Council Directive 2001/110/CE concerning honey, Official J. Eur. Communities Jan 12th 202, L10, 47–52.

Ferreres, F., Andrade, P., & Tomas-Barberan, F. A. (1994). Flavonoids from Portuguese heather honey. *Zeitschrift für Lebensmittel-Untersuchung und - Forschung, 1991*, 32–37.

Gallego Pica, A., & Fernandez Hernando, P. (2013). *Food protected designation of origin* (1st ed.). Oxford: Elsevier.

Goodacre, R., Radovic, B. S., & Anklam, E. (2002). Progress toward the rapid nondestructive assessment of the floral origin of European honey using dispersive Raman spectroscopy. *Applied Spectroscopy, 56*(4), 521–527.

Grillenzoni, F. V., Anfossi, V., Marcazzan, G. L., & Sabatini, A. G. (2002). Una produzione italiana: il miele di agrumi. *L'ape nostra amica, 6*, 36–41.

Grillenzoni, F. V., & Ferro, A. (2008). Valutazione delle caratteristiche del miele di castagno italiano su base pluriennale e territoriale. *Apoidea, 5*(1), 27–33.

Guyot-Declerck, C., Resons, S., Bouseta, A., & Collin, S. (2002). Floral quality and discrimination of Lavandula stoechas, Lavandula angustifolia, and Lavandula angustifolia x latifolia honeys. *Food Chemistry, 79*, 453–459.

Krauze, A., & Zalewski, R. I. (1991). Classification of honey by principal component analysis on the basis of chemical and physical parameters. *Zeitschrift für Lebensmittel-Untersuchung und – Forschung, 192*, 19–23.

Latorre, M. J., Peña, R., Pita, C., Botana, A., Garcia, S., & Herrero, C. (1999). Chemometric classification of honey according to their type. Metal content data. *Food Chemistry, 66*, 263–268.

Lopez, B., Latorre, M. J., Fernandez, M. I., Garcia, M. A., Garcia, S., & Herrer, C. (1996). Chemometric classification of honeys according to their type based on quality control data. *Food Chemistry, 55*(3), 281–287.

Molan, P. C. (1998). The limitations of the methods of identifying the floral source of Honeys. *Bee World, 79*, 59–68.

Von der Ohe, W., Persano Oddo, L., Piana, M. L., Morlot, M., & Martin, P. (2004). Harmonized Methods of melissopalynology. *Apidologie, 35*(suppl. 1), 18–25.

Özbalci, B., Boyaci, İ. H., Topcu, A., Kadilar, C., & Tamer, U. (2013). Rapid analysis of sugars in honey by processing Raman spectrum using chemometric methods and artificial neural networks. *Food Chemistry, 136*, 1444–1452.

Persano Oddo, L., & Ricciardelli D'Albore, G. (1989). Nomenclatura melissopalinologica. *Apicoltura, 5*, 63–72.

Persano Oddo, L., Piazza, M. G., Sabatini, A. G., & Accorti, M. (1995). Characterization of unifloral honeys. *Apidologie, 26*, 453–465.

Persano Oddo, L., & Piro, R. (2004). Main European unifloral honeys: Descriptive sheets. *Apidologie, 35*, 38–81.

Persano Oddo, L., Piana, M. L., & Ricciardelli D'Albore, G. (2007). *I mieli regionali italiani, caratterizzazione melissopalinologica*. Viterbo. Ministero delle Politiche Agricole e Forestali and CRA – Agricoltural Research Conucil.

Perez, R. A., Sanchez-Brunette, C., Calvo, R. M., & Tadeo, J. L. (2002). Analysis of volatiles from Spanish honeys by solid-phase microextraction and gas chromatography–mass spectrometry. *Journal of Agricultural and Food Chemistry, 50*, 2633–2637.

Piana, M. L. (1997). La determinazione dell'origine geografica nel miele e le frodi collegate. *Lapis, 5*, 8–17.

Pirini, A., Conte, L., Francioso, O., & Lecker, G. (1992). Capillary gas chromatographic determination of free amino acids in honey as means of discrimination between botanical sources. *Journal of High Research of Chromatography, 15*, 165–170.

Popek, S. (2002). A procedure to identify a honey type. *Food Chemistry, 79*, 401–406.

Ricciardelli D'Albore, G. (1997). *Textbook of melissopalynology* (1st ed.). Bucharest: Apimondia publishing House.

Sabatini, A. G., Bortolotti, L., & Marcazan, G. L. (2007). *Conoscere il miele* (2nd ed.). Bologna: Avenue Media.

Serrano, S., Villarejo, M., Espejo, R., & Jodral, M. (2004). Chemical and physical parameters of Andalusian honey: Classification of Citrus and Eucalyptus honeys by discriminant analysis. *Food Chemistry, 84*, 619–625.

Siede, R., Schmidt, C., & Buchler, R. (2004). A PCR based apple detection methods as a complementary instrument for the honey quality assessment. *Deutsche Lebensmittel-Rundschau, 100*, 381–384.

Soria, A. C., Gonzales, M., De Lorenzo, C., Martinez-Castro, I., & Sanz, J. (2004). Characterization of artisanal honeys from Madrid (Central Spain) on the basis of their melissopalynological, physicochemical and volatile composition data. *Food Chemistry, 85*, 121–130.

Todd, F. E., & Vansell, G. H. (1942). Pollen grains in nectar and honey. *The Journal of Economic Entomology, 35*, 728–731.

Tomas-Berberan, F. A., Martos, I., Ferreres, F., Radovic, B. S., & Anklam, E. (2001). HPLC flavonoidi profiles as marker for the botanical origin of European unifloral Honeys. *Journal of the Science of Food and Agricolture, 81*, 485–496.

Vandeginste, B. G. M., Massart, D. L., Buyden, L. M. C., De Jong, S., Lewi, P. J., & Smeyers-Verbeke, J. (1998). *Handobook of chemometrics and qualimetric* (1st ed.). Amsterdam: Elsevier.

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems, 2*, 37–52.