Research Article

Received: 29 September 2020 Revised: 26 October 2020 Accepted article published: 23 November 2020 Published online in Wiley Online Library: 7 December 2020

(wileyonlinelibrary.com) DOI 10.1002/jsfa.10961

Botanical origin discrimination of Greek honeys: physicochemical parameters *versus* Raman spectroscopy

Marinos Xagoraris,[a] Elisavet Lazarou,[a] Eleftheria H Kaparakou,[a] Eleftherios Alissandrakis,[b] Petros A Tarantilis[a] and Christos S Pappas[a]*

Abstract

BACKGROUND: The authenticity of honey is of high importance since it affects its commercial value. The discrimination of the origin of honey is of prime importance to reinforce consumer trust. In this study, four chemometric models were developed based on the physicochemical parameters according to European and Greek legislation and one using Raman spectroscopy to discriminate Greek honey samples from three commercial monofloral botanical sources.

RESULTS: The results of physicochemical (glucose, fructose, electrical activity) parameters chemometric models showed that the percentage of correct recognition fluctuated from 92.2% to 93.8% with cross-validation 90.6–92.2%, and the placement of test set was 79.0–84.3% successful. The addition of maltose content in the previous discrimination models did not significantly improve the discrimination. The corresponding percentages of the Raman chemometric model were 95.3%, 90.6%, and 84.3%.

CONCLUSION: The five chemometric models developed presented similar and very satisfactory results. Given that the recording of Raman spectra is simple, fast, a minimal amount of sample is needed for the analysis, no solvent (environmentally friendly) is used, and no specialized personnel are required, we conclude that the chemometric model based on Raman spectroscopy is an efficient tool to discriminate the botanical origin of fir, pine, and thyme honey varieties.

Supporting information may be found in the online version of this article.

Keywords: fir honey; pine honey; thyme honey; botanical origin; Raman spectroscopy; discrimination

INTRODUCTION

According to Codex Alimentarius,[1] 'honey is the natural sweet substance, produced by honeybees from the nectar of plants or from secretions of living parts of plants, or excretions of plant-sucking insects on the living parts of plants, which the bees collect, transform by combining with specific substances of their own, deposit, dehydrate, store and leave in honeycombs to ripen and mature'. Nowadays, honey is the most important primary product of beekeeping, from both quantitative and economic points of view. The most recently published statistics estimate an annual Greek production for 2017 of 21.939 t with total consumption 1.7 kg per person.[2] Domestic production covers about 90% of consumption.[3] Beekeeping is widespread throughout the country, but there are areas of greater interest in beekeeping.

Among Greek unifloral honeys, pine, fir, and thyme varieties are commercially available. Pine honey comes from honeydew secretions of *Marchallina hellenica* (Gennadius), a hemipteran insect species parasitizing living parts of various *Pinus* species. The main pine species in the Greek region are *Pinus halepensis* Ten. and *Pinus brutia* Mill., which belong to the Pinaceae family. Pine honey is mainly collected during August to October, and it constitutes approximately 65% of total Greek annual honey production.[4] Fir honey is also a honeydew honey, produced mainly from *Abies cephalonica* Loudon and *Abies alba* Mill., also of the Pinaceae family.[5] It is estimated that about 5% of the annual production in Greece is fir honey. Thyme honey is a blossom variety derived from *Thymbra capitata* L. of the Lamiaceae family and amounts to about 10% of the total annual production. We mainly find it in the Greek islands. The commercial value of these honeys is high for Greece; therefore, standardization acquires increasingly further interest.

According to European legislation, the honey botanical origin is determined by physicochemical characteristics such as sugar content, moisture, water-insoluble fraction, electrical conductivity,

* Correspondence to: CS Pappas, Department of Food Science and Human Nutrition, Agricultural University of Athens, 75 Iera Odos, 11855 Athens, Greece, E-mail: *chrispap@aua.gr*

a *Laboratory of Chemistry, Department of Food Science and Human Nutrition, Agricultural University of Athens, Athens, Greece*

b *Laboratory of Quality and Safety of Agricultural Products, Landscape and Environment, Department of Agriculture, Hellenic Mediterranean University, Crete, Greece*

free acid, diastase activity, and hydroxymethylfurfural.[6] Greece also uses melissopalynological analysis.[7]

Many studies have been published regarding honey botanical origin based on physicochemical parameters in combination with chemometric techniques.[8–13] However, a wide dispersion of these parameters for the honey type was found, which is associated with the natural heterogeneity of honey as a product. It produced overlapping of these variables, which reduces their usefulness in honey source classification.[14]

Raman spectroscopy is a simple spectroscopic technique used in condensed matter physics and chemistry to study vibrational, rotational, and other low-frequency modes in a system.[15,16] Raman spectra are not affected by the presence of water, and the effect of fluorescence is minimal.[17] Nowadays, Raman spectroscopy is increasingly used as an analytical technique for the evaluation of food safety and quality. In recent years, several studies on honey were performed, proving the potential of Raman spectroscopy as a feasible alternative for honey authentication.[16,18–20]

Multivariate methods have been recently applied in the field of honey authentication to different extents, starting from principal component analysis (PCA), and clustering methods.[18,19] Supervised multivariate classification models have also been applied, the most exploited tools being: linear discriminant analysis (LDA),[15,20] PCA–LDA,[21] factorial discriminant analysis[22] and *K* nearest neighbors.[23] Other approaches are based on artificial neural networks[23,24] and support vector machines.[25]

The aim of this work is to investigate the potential of chemometric models developed based on (i) physicochemical parameters according to European and Greek legislation and (ii) Raman spectroscopy for the purpose the discrimination of botanical origin of three Greek commercial honeys. An additional goal is to compare these chemometric models with each other.

MATERIALS AND METHODS

Samples

Eighty-three monofloral honey samples from three botanical sources (27 pine, 19 fir, and 37 thyme) were purchased directly from Greek beekeepers from 2018 to 2019 harvest years. To ensure that all samples could be classified as monofloral, mellisopalynological analysis was performed. The water content of samples ranged between 13.70 and 18.40% (w/w). Samples were delivered to the laboratory and kept in the dark at 25 °C until further analysis. Honey samples were liquefied in a water bath at 55 °C for 20 min and their Raman spectra were recorded.[20]

Raman spectroscopy

A DeltaNu Advantage 785 visible–infrared Raman spectrometer (DeltaNu Inc., Laramie, WY, USA) equipped with a 785 nm diode laser for excitation with a maximum output power of 71.6 mW was used to record the honey's spectra. Each spectrum was a 10 s acquisition over the spectral range of 2000–200 cm^{-1} using a resolution of 8 cm^{-1}. The spectrometer was accompanied by NuSpec software. A small amount of each sample of honey was placed in Wilmad NMR sample tubes (40 x 8.2mm) (Warminster, USA), , and remained at room temperature before analysis to remove the bubbles. Nine spectra were obtained for each honey sample.

Raman spectra were smoothed using the Savitsky–Golay algorithm and their baselines were corrected. These pretreatments were performed with 'automatic smoothing' (five-point moving second-degree polynomial) and 'baseline correction' (second-degree polynomial, 20 iterations) functions. Finally, using the Statistical Spectra function, the mean of nine spectra for each sample was taken from the nine initial spectra and the mean spectrum was normalized (absorbance maximum value of 1). Spectrum processing was performed using OMNIC v.9.1 software (Thermo Fisher Scientific Inc., Waltham, MA, USA).

Physicochemical and melissopalynological analysis

The determination of the honey sugars was performed by high-performance liquid chromatography using a Shimadzu CTO-10A column oven and a Shimadzu RID-20A detector (Duisburg, Germany).[26] Fructose, glucose, maltose, and sucrose content were determined. Furthermore, the sum of fructose and glucose was calculated. Electrical conductivity (μS cm^{-1}) was performed with a Consort C3010 multi-parameter analyzer (Turnhout, Belgium) according to the International Honey Commission.[27] Finally, moisture (%, w/w) was measured using a 38-01 OPTi refractometer (Bellingham & Stanley, Tunbridge Wells, UK) and Edmund Buhler water bath (Labortechnik Biotechnologie Materialtechnik Unwelttechnik, Wasserburg, Germany) according to the International Honey Commission.[27]

The melissopalynological analysis was performed with a Krüss microscope (A. Krüss Optronic GmbH, Hamburg, Germany).[28]

Statistical analysis

Chemometric models were developed for the discrimination of the samples according to the botanical origin. The chemometric models were based on physicochemical variables and the LDA statistical technique, one according to European legislation and one other according to Greek legislation.[6,7] In these two chemometric models, maltose content was added and they were recalculated and compared. Then, a chemometric model based on Raman spectroscopy and a stepwise-LDA statistical technique was developed.

For the development of chemometric models, the 83 samples were randomly allocated into two groups. The first group of 64 samples (termed 'standards'; St1–St64) was used as a calibration set and the second set of 19 samples (termed 'unknown'; T1–T19) was used as the test set. The normal distribution ($P > 0.05$) of variables of each group was exanimate by the Kolmogorov–Smirnov and Shapiro–Wilk normality tests. The correlation coefficient between variables was checked ($P < 0.05$). Each chemometric model was checked by cross-validation. The statistical analysis was performed using SPSS v.25 software (IBM Corp., Armonk, NY, USA).

RESULTS AND DISCUSSION

Physicochemical and melissopalynological analysis

Table 1 shows the results of physicochemical analysis. Fructose and glucose are the main sugars in honey, and their actual proportion depends largely on the source of the nectar.[29] The sum of fructose and glucose for honeydew and blossom honeys must be not less than 45% and 60% (w/w) respectively.[6] In our case, fir and pine honeys had a minimum content of 46.30% and 46.40% (w/w) respectively, and thyme honeys had 60.10% (w/w). It has been reported that honeydew honey presented a lower mean content of this parameter than blossom honey.[9,30] This can be confirmed by our results, as shown in Table 1. The maltose content is correlated with the botanical origin,[31] and honeydew honeys usually have a higher maltose content than blossom honeys,[32] yet our results showed similar maltose content in all three botanical origins (2.37–2.50% w/w). The sucrose content presented differences, with a higher mean value found in thyme honey samples (0.63% w/w).

Table 1. Results of physicochemical analysis[a]

Botanical origin	Aggregate functions	Fructose (%)	Glucose (%)[b]	Maltose (%)[c]	Sucrose (%)[d]	Fructose + glucose[e]	Electrical conductivity (μS cm^{-1})[f]	Moisture (%)[g]
Fir honeys	Min.	25.50	20.80	0.00	0.00	46.30	1241	13.70
	Max.	35.50	32.70	7.10	0.10	64.50	2000	18.40
	Average	30.16	24.53	2.37	0.07	54.69	1562	15.37
Pine honeys	Min.	26.20	22.60	0.00	0.00	46.40	901	14.40
	Max.	39.20	42.20	6.10	1.00	77.20	1431	17.90
	Average	31.41	28.39	2.50	0.13	59.43	1114	16.02
Thyme honeys	Min.	31.30	25.50	0.00	0.00	60.10	273	14.30
	Max.	43.70	45.40	6.30	2.30	86.40	588	17.90
	Average	36.82	31.08	2.50	0.63	67.88	443	15.93

[a] All percentages w/w.
[b] Glucose (%w/w).
[c] Maltose (%w/w).
[d] Sucrose (%w/w).
[e] Sugars (fructose, glucose) (%w/w).
[f] Electrical conductivity (μS cm^{-1}).
[g] Moisture (%w/w).

Electrical conductivity is a good indicator of the botanical origin of honey.[33] Overall, honeydew honeys are generally characterized by higher values of electrical conductivity than blossom honeys are.[9,34,35] The European legislation states that blossom honeys must have values <800 μS cm^{-1}, whereas honeydew honey and mixtures of honeydew and blossom honeys must have values of >800 μS cm^{-1}.[6] Simultaneously, Greek legislation states that fir honey must have values >1000 μS cm^{-1}, pine honey >900 μS cm^{-1}, and thyme honey <600 μS cm^{-1}.[7] The average electrical conductivity we found for the fir honey samples was 1562 μS cm^{-1}, for pine honey it was 1114 μS cm^{-1}, and for thyme honey it was 443 μS cm^{-1}.

Moisture content is an important quality parameter that influences the shelf life of honey.[36] This depends on various factors, including the harvesting season, the degree of maturity reached in the hive, and climate factors.[37] Moisture content cannot be higher than 20% (w/w)[6] and no more than 18.5% (w/w) for fir honeys.[7] The moisture content of all samples ranged between 13.70% and 18.40% w/w.

The melissopalynological analysis results are shown in Supporting Information Table S1. Thyme honey samples have ≥18% thyme pollen grains, and at the same time the percentage of pollen grains of other plant species does not exceed 45%. Melissopalynological analysis is not a safe criterion for botanical origin of pine and fir honeys.[38] The existence of fungal spores, mold hyphae, microscopic algae, trichomes, and pollen from nectarless and anemophilous plants, stated as honeydew elements (HDEs), contributes to the botanical recognition of these honeys. Pine honey samples have important presence of HDEs, and fir honey samples have a small presence of HDEs.[38]

Spectroscopic analysis

Representative spectra from each botanical origin are presented in Fig. 1. The assignments of the major peaks are shown in Table 2. It was observed that the spectra showed significant similarities. The most important area of the spectrum is between 1700 and700 cm^{-1}, where the most characteristic groups and sugars absorb.

LDA based on European legislation physicochemical analysis

Hydroxymethylfurfural and diastase activity are usually used as a measure of honey freshness.[47] The pH value is related to the stability, the shelf life of honey, and as an indicator for possible microbial contamination.[48,49] Therefore, the previous parameters

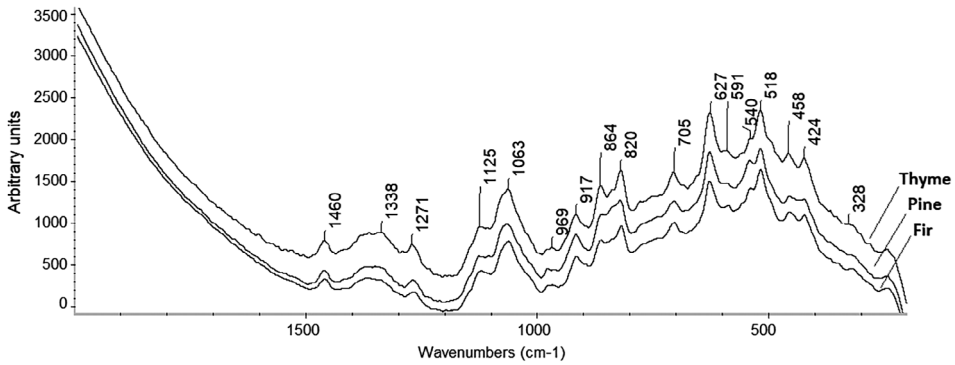

Figure 1. Mean Raman spectra derived from thyme, pine, and fir honey samples.

Table 2. Main peaks of the Raman spectrum derived from honey samples

Wavenumber (cm^{-1})	Functional group	Peak performance	Assignment	Reference
~325–328	C—C—C	Glucose and fructose	Stretching	20
~423	C—C—O and C—C—C	Glucose and fructose	Bending	20
~455–458	Skeletal	Sugars	Stretching	20
~517–518	C—C—O and C—C—C	Glucose and fructose	Bending	39,40
~539–541	Ring	Fructose	Stretching	39,40
~588–590	Ring	Fructose	Deformation	39,40
~627	Skeletal	Sugars	Stretching	19,41
~705	C—O, C—C—O, and O—C—O	Sugars	Stretching and Bending	20,40
~819–821	C—H	Sugars	Stretching	20,42
~860–864	C—H, C—O—H, and CH_2	Sugars	Stretching	20,42
~916–917	C—H and C—O—H	Sugars	Bending	20,40
~978–980	C—C—H	Fructose	Bending	43
~1063–1065	C—O	Sugars	Stretching	20,44
~1121–1125	C—N	Proteins – amino acids	Stretching	40
~1266–1271	C—OH	Protein – amide (III)	Stretching	25,45
~1363	CH_2	Sugars	Deformation	25,45
~1460–1461	CH_2 and COO^-	Flavanols and organic acids	Bending and stretching	44–46

are not related to the botanical origin. For this reason, they were not included in the chemometric models based on the physicochemical characteristics.

It emerged from the Kolmogorov–Smirnov and Shapiro–Wilk tests that the results of some variables did not confirm the normality hypothesis, since the values of the statistical *P*-value were lower than the significance level of 5% ($P < 0.05$; Table 3). Then, the correlation coefficient between variables was checked with Spearman's rho test. The results showed that the variables are somewhat correlated.

So, the next step was the use of PCA. However, it was necessary to check the suitability of the variables for PCA (varimax rotation). The value of Kaiser–Meyer–Olkin measure of sampling adequacy was found to be 0.466. The Bartlett test of sphericity was $\chi^2 = 691.516$ ($P < 0.05$). Finally, the values of sampling adequacy from the anti-image correlation table were <0.5. The

Table 3. Correlation between physicochemical variables

Physicochemical variables	Aggregate functions	Fructose (%)[a]	Glucose (%)[b]	Maltose (%)[c]	Sucrose (%)[d]	Fructose + glucose (%)[e]	EC (μS cm^{-1})[f]	Moisture (%)[g]
Fructose	CC[h]							
	P-value							
Glucose	CC	0.55**						
	P-value	0.00						
Maltose	CC	0.11	0.05					
	P-value	0.37	0.69					
Sucrose	CC	0.20	0.01	0.40**				
	P-value	0.14	0.93	0.00				
Fru-Glu	CC	0.88**	0.85**	0.11	0.08			
	P-value	0.00	0.00	0.37	0.53			
EC	CC	−0.63**	−0.41**	−0.15	−0.58**	−0.55**		
	P-value	0.00	0.00	0.22	0.00	0.00		
Moisture	CC	0.10	0.39**	−0.02	−0.06	0.29*	−0.17	
	P-value	0.44	0.00	0.90	0.65	0.02	0.17	

*$P < 0.05$;
** $P < 0.01$.
[a] Fructose (%w/w).
[b] Glucose (%w/w).
[c] Maltose (%w/w).
[d] Sucrose (%w/w).
[e] Sugars (fructose, glucose) (%w/w).
[f] Electrical conductivity (μS cm^{-1}).
[g] Moisture (%w/w).
[h] Correlation Coefficient.
$N = 64$. All percentages w/w. CC: correlation coefficient; EC: electrical conductivity.

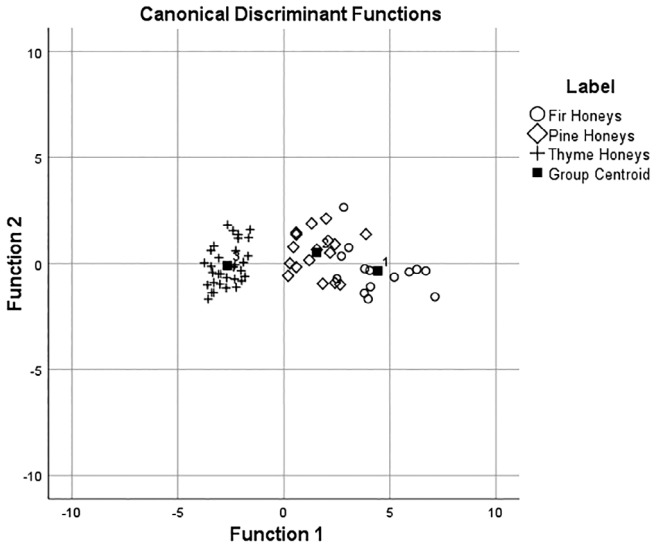

Figure 2. Discrimination results based on European legislation physicochemical analysis.

aforemetioned results indicate that the variables were not suitable for PCA.

Moreover, the LDA model can be used excluding some variables. Sucrose was the first variable to be excluded, because it did not confirm the normality hypothesis and had outliers according to the Kolmogorov–Smirnov and Shapiro–Wilk tests. The sum of fructose and glucose was excluded as it had a high correlation with other variables. Finally, electrical conductivity and moisture were used for the LDA model, based on European legislation.

Box's *M* test ($P > 0.001$) confirmed that we could continue with LDA.[50] Figure 2 shows the results of LDA. The group centroid values, which represent the unstandardized canonical discriminant functions evaluated at group means, are also plotted. Each centroid gives information about the coordinates (discriminant functions) of the group means in the polyparametric space. Specifically, the percentage of samples that were classified correctly was 92.2%, whereas using the cross-validation method resulted in 90.6% (Table 4).

According to the Wilks' Lambda (Λ) statistical test, the percentage that cannot be explained by the variability variable is very small ($\Lambda = 0.090$ with $P < 0.05$ for the first and $\Lambda = 0.904$ with $P < 0.05$ for the second canonical discriminant function).

With regard to the eigenvalues, the first discriminant function recorded the higher eigenvalue (9.025) and the second a much lower one (0.106). The canonical correlation for the first discriminant function was estimated at 94.9%, which explains 98.8% of total variance. The corresponding values for the second discriminant function were found to be 31.0% and 1.2%. This means that the first function contributes significantly to the separation of regions. Discriminant functions accounted for 100% of total variance.

Of the 'unknown' fir honey samples, 80% were correctly classified and the other 20% were in the pine honey group. Of the 'unknown' pine honey samples, 70% were correctly classified and 30% were in the fir honey group. Every 'unknown' thyme honey sample was correctly classified. In total, from the 19 'unknown' honey samples, 15 (79.0%) were correctly classified and four (21.0%) misclassified. Samples that were not classified correctly could in some cases be justified. Pine and fir honey are honeydew honeys, with several similarities in physicochemical characteristics (sugars, electrical conductivity, moisture), so false prediction is explicable. Also, thyme honey often has a variable contribution of pine honeydew.[47]

Then, maltose was added and the model was recalculated (Supporting Information Table S2). It should be mentioned that maltose confirmed the normality hypothesis and did not correlate with the other variables. The percentage of samples that were classified correctly was 92.2%, and the cross-validation method result was 92.2% too. Compared with the European chemometric model, the percentage when using the cross-validation method was a little bit higher. Also, the prediction set had better classification results, with 16 (84.3%) samples that were correctly classified and three (15.7%) misclassified.

Table 4. Classification results[a,b] based on European legislation physicochemical analysis

			Predicted group membership			
		Label	Fir honeys	Pine honeys	Thyme honeys	Total
Original	Count	Fir honeys	10	4	0	14
		Pine honeys	1	16	0	17
		Thyme honeys	0	0	33	33
	%	Fir honeys	71.4	28.6	0.0	100.0
		Pine honeys	5.9	94.1	0.0	100.0
		Thyme honeys	0.0	0.0	100.0	100.0
Cross-validated[c]	Count	Fir honeys	10	4	0	14
		Pine honeys	2	15	0	17
		Thyme honeys	0	0	33	33
	%	Fir honeys	71.4	28.6	0.0	100.0
		Pine honeys	11.8	88.2	0.0	100.0
		Thyme honeys	0.0	0.0	100.0	100.0

[a] 92.2% of original grouped cases correctly classified.
[b] 90.6% of cross-validated grouped cases correctly classified.
[c] Cross-validation is done only for those cases in the analysis. In cross-validation, each case is classified by the functions derived from all cases other than that case.

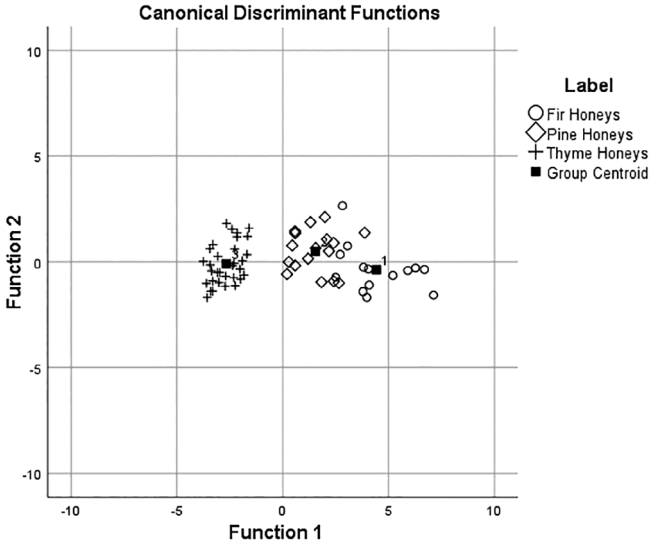

Figure 3. Discrimination results based on Greek legislation physicochemical analysis.

LDA based on Greek legislation physicochemical analysis

Greek legislation includes the variables of both European legislation and melissopalynological analysis. The results of discrimination analysis are shown in Fig. 3. Some 93.8% of honey samples were classified correctly, whereas the corresponding percentage using the cross-validation technique was found to be 92.2% (Table 5). Wilks' Lambda values were found to be very small, at 0.051 and 0.666 ($P < 0.05$) for the first and second canonical discriminant functions, respectively. The corresponding eigenvalues were 12.061 and 0.502. The canonical correlation for the first and second discriminant functions was estimated at 96.1% and 57.8%, which explain 96.0% and 4.0% of total variance, respectively. This shows that the first function contributes significantly to the separation of regions.

Ther were 19 'unknown' honey samples (84.3%) correctly classified and three (15.7%) misclassified. In particular, 100% of the fir honey samples were correctly classified. For the pine honey samples, 70% were correctly classified and the other 30% were in the fir honey group. Every 'unknown' for thyme honey was correctly classified.

A second chemometric model based on Greek legislation plus the maltose content was developed (Supporting Information Table S3). The percentage of samples that were classified correctly was 93.8%, whereas when using the cross-validation method it was 92.2%. The prediction set had the same classification results as the Greek legislation model.

From the previous analysis, we conclude that the four proposed chemometric models can discriminate the honey samples satisfactorily. The addition of maltose content in the chemometric model increases the discrimination percentage. This confirms its association with the botanical origin of honey.

Stepwise-LDA of Raman spectra

The statistically significant spectral regions (Fig. 4) that the statistical model was based on for the discrimination of the three varieties (fir, pine, and thyme) were 600–680, 890–950, 955–1000, and 1000–1100 cm^{-1}. The first spectral region, 600–680 cm^{-1}, was related to deformation of the fructose ring. The second spectral region, 890–950 cm^{-1}, was related to the bending vibration of C—H and C—O—H. The third spectral region, 955–1000 cm^{-1}, was related to the C—C—H bending of the fructose. Finally, the fourth spectral region, 1000–1100 cm^{-1}, was related to the vibration of C—O, and the region from 1070 to 1077 cm^{-1} was mainly due to vibration of C—H and C—O—H of sugars and a small contribution from the vibration of the C—N bond of proteins and amino acids.

We based the creation of the calibration model on the spectral differences in the 1280–600 cm^{-1} spectral region. The equivalence between the groups was checked with Box's *M* test ($P > 0.001$), because the number of samples per group is not the same. The results showed that six steps were formed. The Wilks' Lambda value of the first step was found to be 0.829 with $P < 0.05$, the second was 0.220 with $P < 0.05$, the third was

Table 5. Classification results[a,b] based on Greek legislation physicochemical analysis

			Predicted group membership			
		Label	Fir honeys	Pine honeys	Thyme honeys	Total
Original	Count	Fir honeys	12	2	0	14
		Pine honeys	2	15	0	17
		Thyme honeys	0	0	33	33
	%	Fir honeys	85.7	14.3	0.0	100.0
		Pine honeys	11.8	88.2	0.0	100.0
		Thyme honeys	0.0	0.0	100.0	100.0
Cross-validated[c]	Count	Fir honeys	11	3	0	14
		Pine honeys	2	15	0	17
		Thyme honeys	0	0	33	33
	%	Fir honeys	78.6	21.4	0.0	100.0
		Pine honeys	11.8	88.2	0.0	100.0
		Thyme honeys	0.0	0.0	100.0	100.0

[a] 93.8% of original grouped cases correctly classified.
[b] 92.2% of cross-validated grouped cases correctly classified.
[c] Cross-validation is done only for those cases in the analysis. In cross-validation, each case is classified by the functions derived from all cases other than that case.

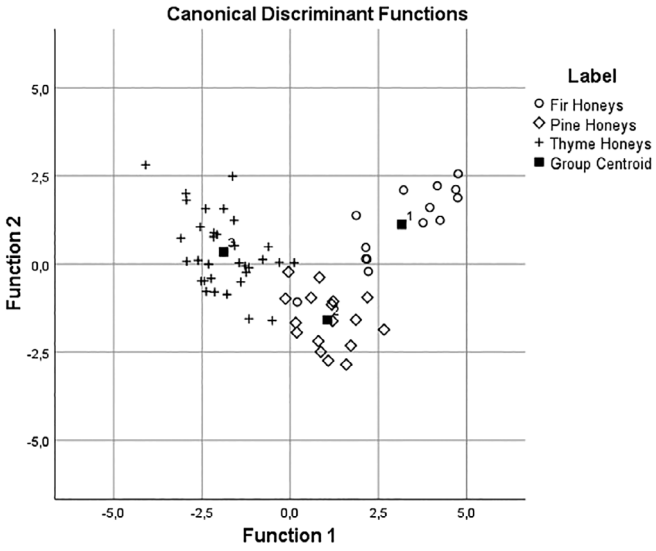

Figure 4. The statistically significant spectral regions that the statistical region was based on for the differentiation of the three honey varieties.

0.159 with $P < 0.05$, the fourth was 0.123 with $P < 0.05$, the fifth was 0.101 with $P < 0.05$, and the sixth was 0.088 with $P < 0.05$.

Then, stepwise-LDA was performed in honey samples to set the calibration model. According to the results, a satisfactory discrimination was observed (Fig. 5). Honeys were differentiated according to botanical origin. The group centroid values, which represent the unstandardized canonical discriminant functions evaluated at group means, are also plotted. Each centroid gives information about the coordinates (discriminant functions) of the group means in the polyparametric space. More specifically, the percentage of samples that were classified correctly was 95.3%, whereas with the method of cross-validation it was 90.6% (Table 6).

According to the Wilks' Lambda statistical test, the percentage that cannot be explained by the variability variable is very small ($\Lambda = 0.088$ with $P < 0.05$ for the first and $\Lambda = 0.487$ with $P < 0.05$ for the second canonical discriminant function), which shows satisfactory discriminating ability of the chemometric model.

The calibration model was also confirmed by eigenvalues. The first discriminant function recorded the higher eigenvalue (4.535) and the second a much lower eigenvalue (1.053). The

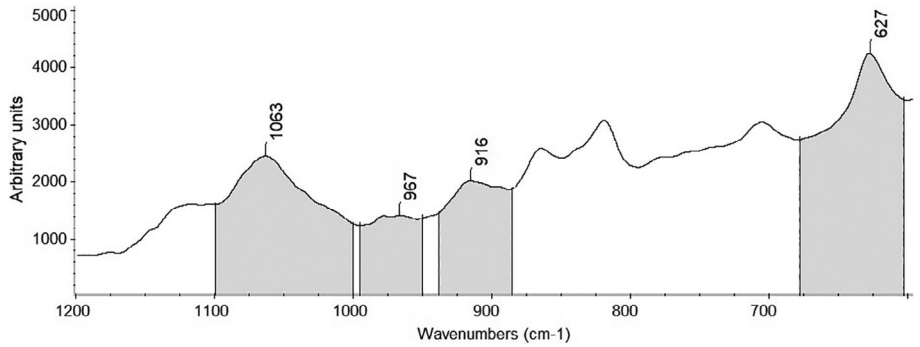

Figure 5. Discrimination results based on Raman spectra analysis.

Table 6. Classification[a,b] results based on Raman spectra

			Predicted group membership			
		Label	Fir honeys	Pine honeys	Thyme honeys	Total
Original	Count	Fir honeys	13	1	0	14
		Pine honeys	0	16	1	17
		Thyme honeys	0	1	32	33
	%	Fir honeys	92.9	7.1	0.0	100.0
		Pine honeys	0.0	94.1	5.9	100.0
		Thyme honeys	0.0	3.0	97.0	100.0
Cross-validated[c]	Count	Fir honeys	12	2	0	14
		Pine honeys	0	16	1	17
		Thyme honeys	0	3	30	33
	%	Fir honeys	85.7	14.3	0.0	100.0
		Pine honeys	0.0	94.1	5.9	100.0
		Thyme honeys	0.0	9.1	90.9	100.0

[a] 95.3% of original grouped cases correctly classified.
[b] 90.6% of cross-validated grouped cases correctly classified.
[c] Cross-validation is done only for those cases in the analysis. In cross-validation, each case is classified by the functions derived from all cases other than that case.

canonical correlation for the first discriminant function was estimated at 90.5%, which explains 81.2% of total variance. The second discriminant function was estimated at 71.6%, which explains 18.8% of total variance. This confirms that the first function contributes significantly to the separation of regions. Discriminant functions accounted for 100% of total variance.

In total, from the 19 'unknown' honey samples, 16 (84.3%) were correctly classified and three (15.7%) misclassified. Of the 'unknown' fir honey samples, 60% were correctly classified. The other 40% were classified in the pine honey group. The corresponding percentage for the 'unknown' pine honey samples (90%) was correctly classified in the pine honey group. The other 10% were classified in the fir honey group. Of the 'unknown' thyme honey samples, 100% were correctly classified in the thyme honey group. Pine and fir honeys are honeydew varieties, with several similarities in physicochemical characteristics (sugars, electrical conductivity, moisture), so false prediction is explicable. Also, thyme honey often has some contribution of pine honey.

For the estimation of these results, partial least-squares regression models were performed between the Raman spectra and the parameters that were considered as significant for the chemometric models.[45] The results showed the parameters evaluated had a high correlation coefficient (Supporting Information Figs S1–S3). Good calibration models were obtained between Raman spectra and electrical conductivity ($R^2 = 0.967$), moisture ($R^2 = 0.877$), and maltose ($R^2 = 0.760$). These results confirm the hypothesis that Raman spectroscopy is a useful technique with acceptable accuracy in determining the botanical origin of monofloral honeys.

CONCLUSIONS

Two chemometric models were developed, employing data from physicochemical parameters according to European and Greek legislation, and two other models according to previous parameters plus maltose content, for the purpose of botanical discrimination of three commercial Greek honeys (fir, pine, thyme). According to the results, the percentage of correct standard recognition ranged from 92.2% to 93.8%, with cross-validation values of 90.6–92.2%, whereas the percentage of correct placement of the 'unknown' samples of the test set was 79.0–84.3%. The addition of maltose content to the chemometric models did not significantly affect the results of the discrimination.

In addition, a chemometric model of discrimination was developed based on Raman spectroscopy combined with stepwise-LDA. The results were similar to those of previous chemometric models (recognition of standards: 95.3%; cross-validation: 90.6%; test set: 84.3%).

Considering that Raman spectroscopy is simple, is not time-consuming, is non-destructive, requires small amounts of samples, is environmentally friendly (does not use solvents), does not require specialized personnel, and, at the same time, is just as accurate, we conclude that it can be used for botanical origin discrimination of the aforementioned honey samples, in combination with the stepwise-LDA statistical technique.

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Codex Alimentarius Commission, *Revised Codex Standard for Honey*. Codex Alimentarius Commission, Rome, Italy, pp. 12–1981 (2001) Codex STAN 12-1981.
2 Annual Agricultural Survey for Livestock Products, Greek Statistical Authority (2020). *Agriculture, Livestock*. [Online]. Available: https://www.statistics.gr/el/statistics/agr [15 July 2020].
3 Information and Statistics for Beekeeping, Greek Ministry of Rural Development and Food (2018) [Online]. Available: http://www.minagric.gr/index.php/el/for-farmer-2/animal-production/bees-honey [15 July 2020].
4 Tananaki C, Thrasyvoulou A, Giraudel JL and Montury M, Determination of volatile characteristics of Greek and Turkish pine honey samples and their classification by using Kohonen self organising maps. *Food Chem* **101**:1687–1693 (2007).
5 Karabagias IK, Halatsi EZ, Karabournioti S, Kontakos S and Kontominas MG, Impact of physicochemical parameters, pollen grains, and phenolic compounds on the correct geographical differentiation of fir honeys produced in Greece as assessed by multivariate analyses. *Int J Food Prop* **20**:S520–S533 (2017).
6 EU, Council Directive 2001/110/EC of 20 December 2001 relating to honey. *Off J Eur Communities Legis* **10**:47–52 (2002) https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:010:0047:0052:EN:PDF.
7 Government Gazette B-239/23–2-2005; Annex II, article 67 of Greek Food Code. [Online]. Available: http://www.minagric.gr/images/stories/docs/agrotis/MeliMelissokomia/KYA_Taytopoiisi_.pdf.
8 Karabagias IK, Badeka AV, Kontakos S, Karabournioti S and Kontominas MG, Botanical discrimination of Greek unifloral honeys with physicochemical and chemometric analyses. *Food Chem* **165**:181–190 (2014).
9 Bentabol-Manzanares A, Hernandez-Garcia Z, Rodriguez-Galdon B, Rodriguez- Rodriguez E and Diaz-Romero C, Differentiation of blossom and honeydew honeys using multivariate analysis on the physicochemical parameters and sugar composition. *Food Chem* **126**:664–672 (2011).
10 Leon-Ruiz V, Vera S, Gonzalez-Porto AV and Andres MPS, Vitamin C and sugar levels as simple markers for discriminating Spanish honey sources. *J Food Sci* **76**:C356–C361 (2011).
11 Senyuva HZ, Gilbert J, Silici S, Charlton A, Dal C, Gürel N *et al*., Profiling Turkish honeys to determine authenticity using physical and chemical characteristics. *J Agric Food Chem* **57**:3911–3919 (2009).
12 Marini F, Magrì AL, Balestrieri F, Fabretti F and Marini D, Supervised pattern recognition applied to the discrimination of the floral origin of six types of Italian honey samples. *Anal Chim Acta* **515**:117–125 (2004).
13 Terrab A, Dıez MJ and Heredia FJ, Characterisation of Moroccan unifloral honeys by their physicochemical characteristics. *Food Chem* **79**:373–379 (2002).
14 Soria AC, Gonzalez M, De Lorenzo C, Martínez-Castro I and Sanz J, Estimation of the honeydew ratio in honey samples from their physicochemical data and from their volatile composition obtained by SPME and GC–MS. *J Sci Food Agric* **85**:817–824 (2005).
15 Oroian M, Ropciuc S and Paduret S, Honey adulteration detection using Raman spectroscopy. *Food Anal Methods* **11**:959–968 (2018).
16 Goodacre R, Radovic BS and Anklam E, Progress toward the rapid nondestructive assessment of the floral origin of European honey using dispersive Raman spectroscopy. *Appl Spectrosc* **56**:521–527 (2002).
17 Owen CA, Notingher I, Hill R, Stevens M and Hench LL, Progress in Raman spectroscopy in the fields of tissue engineering, diagnostics and toxicological testing. *J Mater Sci Mater Med* **17**:1019–1023 (2006).
18 Corvucci F, Nobili L, Melucci D and Grillenzoni FV, The discrimination of honey origin using melissopalynology and Raman spectroscopy techniques coupled with multivariate analysis. *Food Chem* **169**:297–304 (2015).
19 Frausto-Reyes C, Casillas-Penuelas R, Quintanar-Stephano JL, Macias-Lopez E, Bujdud-Perez JM and Medina-Ramirez I, Spectroscopic study of honey from *Apis millifera* from different regions in Mexico. *Spectrochim Acta Part A* **178**:212–217 (2017).
20 Oroian M and Ropciuc S, Botanical authentication of honeys based on Raman spectra. *J Food Meas Charact* **12**:545–554 (2018).
21 Ruoff K, Luginbuhl W, Bogdanov S, Bosset JO, Estermann B, Ziolko T *et al*., Authentication of the botanical origin of honey by near-infrared spectroscopy. *J Agric Food Chem* **54**:6867–6872 (2006).
22 Hennessy S, Downey G and O'Donnell CP, Attempted confirmation of the provenance of Corsican PDO honey using FT-IR spectroscopy

and multivariate data analysis. *J Agric Food Chem* **58**:9401–9406 (2010).
23 Latorre CH, Crecente RMP, Martin SG and Garcia JB, A fast chemometric procedure based on NIR data for authentication of honey with protected geographical indication. *Food Chem* **141**:3559–3565 (2013).
24 Chen LZ, Wang JH, Ye ZH, Zhao J, Xue XF, Vander Heyden Y *et al*., Classification of Chinese honeys according to their floral origin by near infrared spectroscopy. *Food Chem* **135**:338–342 (2012).
25 Fernández Pierna JA, Abbas O, Dardenne P and Baeten V, Discrimination of Corsican honey by FT-Raman spectroscopy and chemometrics. *Biotechnol Agron Soc Environ* **15**:75–84 (2011).
26 Helrich K, *Official Methods of Analysis of Association of Official Analytical Chemists*, 15th edn. Association of Official Analytical Chemist, Arlington, VA, USA (1990).
27 IHC, Harmonised Methods of the International Honey Commission, [Online]. (2009). Available: https://www.ihc-platform.net/ihcmethods2009.pdf
28 Louveaux J, Maurizio A and Vorwohl G, Methods of melissopalynology. *Bee World* **51**:125–138 (1970).
29 Anklam E, A review of the analytical methods to determine the geographical and botanical origin of honey. *Food Chem* **63**:549–562 (1998).
30 Soria A, Gonzalez M, De Lorenzo C, Martınez-Castro I and Sanz J, Characterization of artisanal honeys from Madrid (central Spain) on the basis of their melissopalynological, physicochemical and volatile composition data. *Food Chem* **85**:121–130 (2004).
31 Cotte JF, Casabianca H, Chardon S, Lheritier J and Grenier-Loustalot MF, Chromatographic analysis of sugars applied to the characterisation of monofloral honey. *Anal Bioanal Chem* **380**: 698–705 (2004).
32 Mateo R and Bosch-Reig F, Sugar profiles of Spanish unifloral honeys. *Food Chem* **60**:33–41 (1997).
33 Bogdanov S, Lullmann C, Martin P, von der Ohe W, Russmann H, Vorwohl G *et al*., Honey quality and international regulatory standards: review by the International Honey Commission. *Bee World* **80**:61–69 (1999).
34 Thrasyvoulou A and Manikis J, Some physicochemical and microscopic characteristics of Greek unifloral honeys. *Apidologie* **26**:441–452 (1995).
35 Diez MJ, Andres C and Terrab A, Physicochemical parameters and pollen analysis of Moroccan honeydew honeys. *Int J Food Sci Technol* **39**:167–176 (2004).
36 Bogdanov S, Ruoff K and Persano Oddo L, Physicochemical methods for the characterisation of unofloral honeys: a review. *Apidologie* **35**:S4–S17 (2004).
37 Finola MS, Lasagno MC and Marioli JM, Microbiological and chemical characterization of honeys from central Argentina. *Food Chem* **100**: 1649–1653 (2007).
38 Alvarez-Suarez JM ed, *Bee Products – Chemical and Biological Properties*. Springer International Publishing AG, Cham, Switzerland (2017).
39 Batsoulis AN, Siatis NG, Kimbaris AC, Alissandrakis EK, Pappas CS, Tarantilis PA *et al*., FT-Raman spectroscopic simultaneous determination of fructose and glucose in honey. *J Agric Food Chem* **53**: 207–210 (2005).
40 Tahir HE, Xiaobo Z, Zhihua L, Jiyong S, Zhai X, Wang S *et al*., Rapid prediction of phenolic compounds and antioxidant activity of Sudanese honey using Raman and Fourier transform infrared (FT-IR) spectroscopy. *Food Chem* **226**:202–211 (2017).
41 Anjos O, Campos MG, Ruiz PC and Antunes P, Application of FTIR-ATR spectroscopy to the quantification of sugar in honey. *Food Chem* **169**:218–223 (2015).
42 Corbett EC and Vichy V, Fourier transform Raman studies of materials and compounds of biological importance – II. The effect of moisture on the molecular structure of the alpha and beta anomers of D-glucose. *Spectrochim Acta Part A* **47A**:1399–1411 (1991).
43 Jandrić Z, Haughey SA, Frew RD, McComb K, Galvin-King P, Elliott CT *et al*., Discrimination of honey of different floral origins by a combination of various chemical parameters. *Food Chem* **189**:52–59 (2015).
44 Paradkar MM and Irudayaraj J, Discrimination and classification of beet and cane inverts in honey by FT-Raman spectroscopy. *Food Chem* **76**:231–239 (2001).
45 Anjos O, Santos AJA, Paixão V and Estevinho LM, Physicochemical characterization of *Lavandula* spp. honey with FT-Raman spectroscopy. *Talanta* **178**:43–48 (2018).
46 Nickless EM, Holroyd SE, Stephens JM, Gordon KC and Wargent JJ, Analytical FT-Raman spectroscopy to chemotype *Leptospermum scoparium* and generate predictive models for screening for dihydroxyacetone levels in floral nectar. *J Raman Spectrosc* **45**: 890–894 (2014).
47 Calvo CP and Vazquez M, Differences between honeydew and blossom honeys: a review. *Trends Food Sci Technol* **59**:79–87 (2017).
48 Terrab A, Recamales AF, Hernanz D and Heredia FJ, Characterisation of Spanish thyme honeys by their physicochemical characteristics and mineral contents. *Food Chem* **88**:537–542 (2004).
49 Conti ME, Lazio region (central Italy) honeys: a survey of mineral content and typical quality parameters. *Food Control* **11**:459–463 (2000).
50 Hahs-Vaughn D, *Applied Multivariate Statistical Concepts*. Routledge, New York, NY, USA (2016).