Progress Toward the Rapid Nondestructive Assessment of the Floral Origin of European Honey Using Dispersive Raman Spectroscopy

ROYSTON GOODACRE,* BRANKA S. RADOVIC, and ELKE ANKLAM

Institute of Biological Sciences, Cledwyn Building, The University of Wales, Aberystwyth, Ceredigion, SY23 3DD, UK (R.G.); and Food Products and Consumer Goods Unit, Institute for Health and Consumer Protection, Joint Research Centre Ispra, Commission of the European Union, I-21020 Ispra, Italy (B.S.R., E.A.)

Raman spectroscopy was investigated for its ability to discriminate between honey samples from different floral and geographical origins. The major vibrational modes in the Stokes Raman spectra were assigned and could be attributed to the four main sugars found in the honeys. The chemometric clustering method of discriminant function analysis indicated that the major differences between the honeys was due to their botanical origin rather than their country of origin, and this was confirmed by artificial neural network analyses. We consider the noninvasive nondestructive analysis of honey by Raman spectroscopy to be an alternative to the laborious and highly specialized mellisopalynology typing method currently used to identify the floral origin of honey.

Index Headings: **Raman spectroscopy; Chemometrics; Honey; Authenticity; Botanical origin.**

INTRODUCTION

The composition and the manufacture of honey are regulated by Community Directive 74/409/EEC (OJEC L 221, 12.8.1974). In order to harmonize the common European market, the European Commission has adopted a proposal to amend this Directive. According to this amendment, the name 'honey' has to be supplemented by information referring to the product's floral and geographical origin.

Traditionally, the determination of the botanical origin of honey has been achieved by analysis of the pollen (mellisopalynology) present in honey.[1,2] This method is based on the identification of pollen by microscopic examination, and so requires a very experienced analyst; it is thus very time consuming and dependent on the expert's ability and judgment.[3] The development of new methods that do not depend on expert analysis and potentially subjective opinion is therefore desirable.

Dispersive Raman spectroscopy is a physico-chemical method that measures the vibrations of bonds within functional groups by measuring the *exchange* of energy with EM radiation of a particular wavelength of light (e.g., a 780-nm near-infrared diode laser, as conducted here). This exchange of energy results in a measurable Raman shift in the wavelength of the incident laser light.[4–6] The Raman effect is, however, very weak because only 1 in every 10^8 photons exchange energy with a molecular bond vibration and the rest of the photons are Rayleigh scattered (that is to say, scattered with the same frequency as the incident monochromatic (ν_o) laser light). The Raman shift can result in two lines, $\nu_o - \nu_m$ and $\nu_o + \nu_m$, which are called Stokes and anti-Stokes lines, respectively. The Stokes Raman shift is considerably stronger than anti-Stokes Raman scattering, and thus, these are usually collected and can be used to construct a Raman 'fingerprint' of the sample. Because different bonds scatter different wavelengths of EM radiation, these Raman 'fingerprints' are made up of the vibrational features of all the sample components. Therefore, this method will give quantitative information about the total chemical composition of a honey sample, without its destruction (that is to say, it is totally ''noninvasive''), and produce 'fingerprints' that are reproducible and distinct for different materials.

Raman spectroscopy has only relatively recently been investigated as a potential tool for food quality control, for food compositional identification,[7] and for the detection of adulteration in foodstuffs,[8] as well as for basic research in the elucidation of structural or conformational changes that occur during processing of foods.[9] With reference to our own studies, we have found that dispersive Raman spectroscopy with laser excitation at 780 nm has been very useful for the classification of bacteria,[10] identification of cosmetics,[11] and the analysis of on-line fermentations.[12,13] The aim of the present study was to investigate dispersive Raman spectroscopy for the classification of honey according to its floral origin.

MATERIALS AND METHODS

Samples. Initially 80+ honey samples were obtained from various hive sites in seven different EU Member States. Standard pollen analysis was performed on all honey samples in order to confirm their floral authenticity. Some of the honey samples received contained multiple pollen types and so were not of unifloral origin. The honeys that could not be designated to a pure botanical origin were precluded from Raman spectroscopic analyses because we did not want to include erroneous honey assignments that would pollute the validation of this method. Thirteen confirmed unifloral types were thus provided by 51 samples. These were: acacia (7 samples), chestnut (9 samples), eucalyptus (4 samples), heather (10 samples), lime (4 samples), rape (5 samples), sunflower (4 samples), citrus (2 samples), lavender (2 samples), rosemary (1 sample), *Echium plantagineum* (1 sample), orange (1 sample), and fior di sulla (1 sample) (see Table I for full details).

Raman Spectroscopy. Spectra were collected using a

Received 6 June 2001; accepted 16 November 2001.
* Author to whom correspondence should be sent.

0003-7028/02/5604-0521$2.00/0

TABLE I. List of the honey samples analyzed.

JRC sample number	Botanical origin	Geographical origin	JRC sample number	Botanical origin	Geographical origin
1	Acacia	Germany	30	Rape	France
2	Acacia	Germany	31	Acacia	France
3	Lime	Germany	32	Lavender	France
4	Rape	Germany	33	Heather	France
5	Rape	Germany	34	Sunflower	France
6	Heather	Germany	35	Chestnut	France
7	Heather	Germany	36	Heather	France
8	Rape	Denmark	37	Heather	Netherlands
10	Chestnut	Italy	38	Lime	Netherlands
11	Acacia	Italy	39	Heather	Netherlands
12	Orange	Italy	40	Sunflower	France
13	Sunflower	Italy	41	Lime	Netherlands
14	Eucalyptus	Italy	42	Chestnut	Germany
15	Sunflower	Italy	43	Chestnut	Germany
16	Chestnut	Italy	44	Lime	Germany
17	Eucalyptus	Italy	45	Citrus	Italy
18	Fior di sulla	Italy	46	Rape	England
19	Acacia	Italy	47	Heather	England
20	Acacia	Italy	48	Heather	England
21	Chestnut	Italy	49	Heather	England
22	Acacia	Italy	50	Heather	England
23	Chestnut	Italy	51	*Echium plantagineum*	Portugal
24	Chestnut	Italy			
25	Eucalyptus	Spain	52	Eucalyptus	Portugal
27	Citrus	Spain	53	Lavender	Portugal
29	Chestnut	France	54	Rosemary	Portugal

Renishaw System 100 dispersive Raman spectrometer (Renishaw, UK), with a near-infrared 780-nm diode laser with the power at the sampling point typically at 80 mW.[14,15] The instrument grating was calibrated using neon lines[16] and was routinely checked with a silicon wafer centered at 520 nm and 100% ethanol for the C–C–O vibration at 880 cm^{-1}. A spectrum from each sample was collected for 60 s using the continuous extended scan (so that actual collection time was ~6 min), and the spectral resolution was 6 cm^{-1}. In order to reduce fluorescence, each honey was diluted with distilled water one tenth in a total volume of 4 mL. These were pipetted into a 4-mL Supelco vial (Supelco, PA); these were 10 mm in diameter and made of borosilicate glass. The vial was placed into a pre-fixed sample holder such that the laser was focused into the center of the vial (12 mm from the collection lens). Samples were analyzed in triplicate.

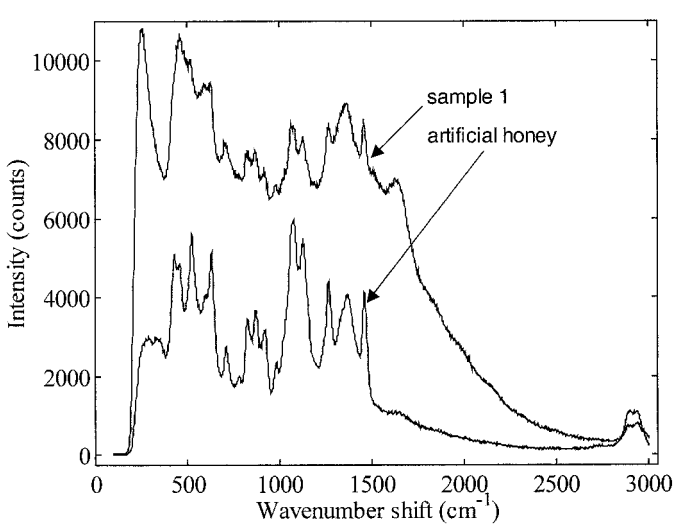

FIG. 1. Stokes Raman spectra of an acacia honey from Germany (sample 1) and artificial honey (see text for details).

The GRAMS WiRE software package (Galactic Industries Corporation, NH) running under Windows 95 was employed for instrument control and data capture. Spectra were collected over 100–3000 cm^{-1} wavenumber shifts with 1735 data points; therefore, the data binning was ~1.67 cm^{-1}. The data may be displayed as the intensity of Raman photon counts against Stokes Raman shift in wavenumbers (see Fig. 1 for a typical spectrum). Prior to chemometric analyses, ASCII data were exported from the GRAMS WiRE software used to control the Raman instrument. To account for photon count differences, the spectra were scaled such that the offset = 0 and the height of the first line (where the laser line is cut out by the holographic filter) at 250 cm^{-1} = 1.

Cluster Analyses. Multivariate data (such as that generated by Raman spectroscopy) consist of the results of observations of many different characters or variables (light frequency shifts) for a number of individuals or objects.[17] Each frequency shift (wavenumber) may be regarded as constituting a different dimension, such that if there are n variables (where n = 1735 measurements) each object may be said to reside at a unique position in an abstract entity referred to as n-dimensional hyperspace. This hyperspace is necessarily difficult to visualize, and an underlying theme of multivariate analysis is thus simplification[18,19] or dimensionality reduction, which usually means that we want to summarize a large body of data by means of relatively few parameters, preferably the two or three which lend themselves to graphical display, with minimal loss of information. Thus the initial

TABLE II. Mean outputs from five different 10-4-8 multilayer perceptrons.

	Botanical origin of honey							
Sample (origin)	Acacia	Chestnut	Eucalyptus	Heather	Lime	Rape	Sunflower	Other[a]
1 (Acacia)	**0.3**	−0.1	−0.1	−0.1	0.0	0.0	−0.1	0.2
22 (Acacia)	**0.9**	−0.1	−0.1	−0.1	0.1	0.0	−0.1	0.2
24 (Chestnut)	−0.1	**1.0**	0.1	0.0	0.1	−0.1	−0.1	0.1
35 (Chestnut)	−0.1	**0.8**	0.1	0.0	0.0	−0.1	−0.1	0.3
43 (Chestnut)	−0.1	**1.1**	0.1	−0.1	0.2	−0.1	−0.1	0.1
17 (Eucalyptus)	−0.1	−0.1	**0.5**	0.2	−0.1	0.0	0.1	0.3
7 (Heather)	−0.1	0.4	0.0	**0.8**	0.3	−0.1	−0.1	0.1
36 (Heather)	−0.1	0.5	0.2	**0.6**	0.0	−0.1	−0.1	0.1
39 (Heather)	−0.1	0.5	**0.7**	0.1	−0.1	0.1	−0.1	0.4
50 (Heather)	−0.1	0.0	0.1	**1.0**	0.1	−0.1	0.0	0.0
41 (Lime)	0.0	−0.1	−0.1	−0.1	**0.8**	−0.1	−0.1	0.0
5 (Rape)	0.0	−0.1	−0.1	−0.1	−0.1	**0.8**	0.0	0.4
8 (Rape)	−0.1	−0.1	−0.1	−0.1	0.0	**0.6**	−0.1	0.5
15 (Sunflower)	−0.1	−0.1	0.1	0.0	−0.1	0.2	**1.1**	0.0

[a] Unifloral types belonging to citrus, lavender, rosemary, *Echium plantagineum*, orange, and fior di sulla honeys. This dummy node has been used previously for analyzing other spectroscopic data (Ref. 41).
Bold = winning node.
Underlined = correct identity.

stage of the chemometric analyses involved the reduction of the multidimensional Raman data by principal components analysis (PCA).[18,20] PCA is a well known technique for reducing the dimensionality of multivariate data while preserving most of the variance, and Matlab was employed to perform PCA according to the NIPALS algorithm.[21] Discriminant function analysis (DFA; also known as canonical variates analysis (CVA)) then discriminated between groups on the basis of the retained principal components (PCs) and the *a priori* knowledge of which spectra were replicates, and thus, this process does not bias the analysis in any way.[22] These types of analysis fall into the category of ''unsupervised learning'', in which the relevant multivariate algorithms seek ''clusters'' in the data,[23] thus allowing the investigator to group objects together on the basis of their perceived closeness in the *n*-dimensional hyperspace referred to above. These methods were implemented using Matlab version 5 (The Math Works, Inc., MA), which runs under Microsoft Windows NT on an IBM compatible PC.

Common Supervised Analysis Methods. When the desired responses (targets) associated with each of the inputs (spectra) are known then the system may be ''supervised''. The goal of supervised learning is to find a model that will correctly associate the inputs with the targets; this is usually achieved by minimizing the error between the target and the model's response (output).[24] A popular method for achieving this is the multilayer perceptron (MLP) using log sigmoidals as the transfer functions and standard back-propagation.[25–27] All the ANNs were carried out with a user-friendly neural network simulation program, NeuFrame version 3,0,0,0 (Neural Computer Sciences, Southampton, Hants), which runs under Microsoft Windows NT on an IBM compatible PC.

To attempt to predict the botanical origin for those honeys, only those honeys that contained enough (>3) samples were used to classify to a unifloral variety of honey; these were acacia (7 samples), chestnut (9), eucalyptus (4), heather (10), lime (4), rape (5), and sunflower (4). One third (14 honeys) of these were chosen randomly as a test set (JRC sample numbers shown); 1, 22 (acacia); 24, 35, 43 (chestnut); 17 (eucalyptus); 7, 36, 39, 50 (heather); 41 (lime); 5, 8 (rape); and 15 (sunflower). The 29 other honey samples, including those honeys that contained <3 samples, were used as a training set. For the latter, containing citrus, lavender, rosemary, *Echium plantagineum,* orange, and fior di sulla, these were encoded in a single node called 'other floral origin' honey.

Using the full original Raman spectra the number of inputs would be 1735 Raman scatters; because this is so large with respect to the number of training examples (36 × 3 = 108), in order to obey the parsimony principle,[27–29] the number of inputs was reduced by using the first 10 principal components, a method we and others have found to be useful as a preprocessing step to ANNs.[30–33] PCA was performed on both the training and test sets, and the total percentage explained variance was 99.8%. As 7 unifloral botanical origins plus one mixed botanical origin were to be assessed, the output was binary encoded in 8 nodes (see Table II for details). Various MLP architectures (n_{inputs}-n_{hidden}-$n_{\text{output nodes}}$) were employed that differed in the number of hidden nodes: 10-4-8, 10-7-8, 10-10-8. It was found that in training each MLP to 0.15% RMSEC (root mean squared error of calibration), all ANNs gave very similar results; therefore, the 10-4-8 MLP was used, as it was the most parsimonious.

RESULTS AND DISCUSSION

A typical Raman spectrum from one of the German honeys of acacia botanical origin is shown in Fig. 1. The most prominent peaks that can be observed are of carbohydrate origin,[34] and this was perhaps not surprising because honey consists of ~80 g/100 g of mono- and disaccharides (OJEC L 221, 12.8.1974). When artificial honey was made in distilled H_2O, comprising (/100 g) fructose (38.5 g), glucose (31.0 g), maltose (7.2 g), sucrose (1.5 g), and H_2O (21.8 g),[35] and analyzed by dispersive Raman spectroscopy (Fig. 1), the majority of the bands observed in the real honey were seen to be attributable to just the sugar composition found in the honey. Close visible inspection of the spectrum in Fig. 1 and the others collected showed very few, if any, prominent extra

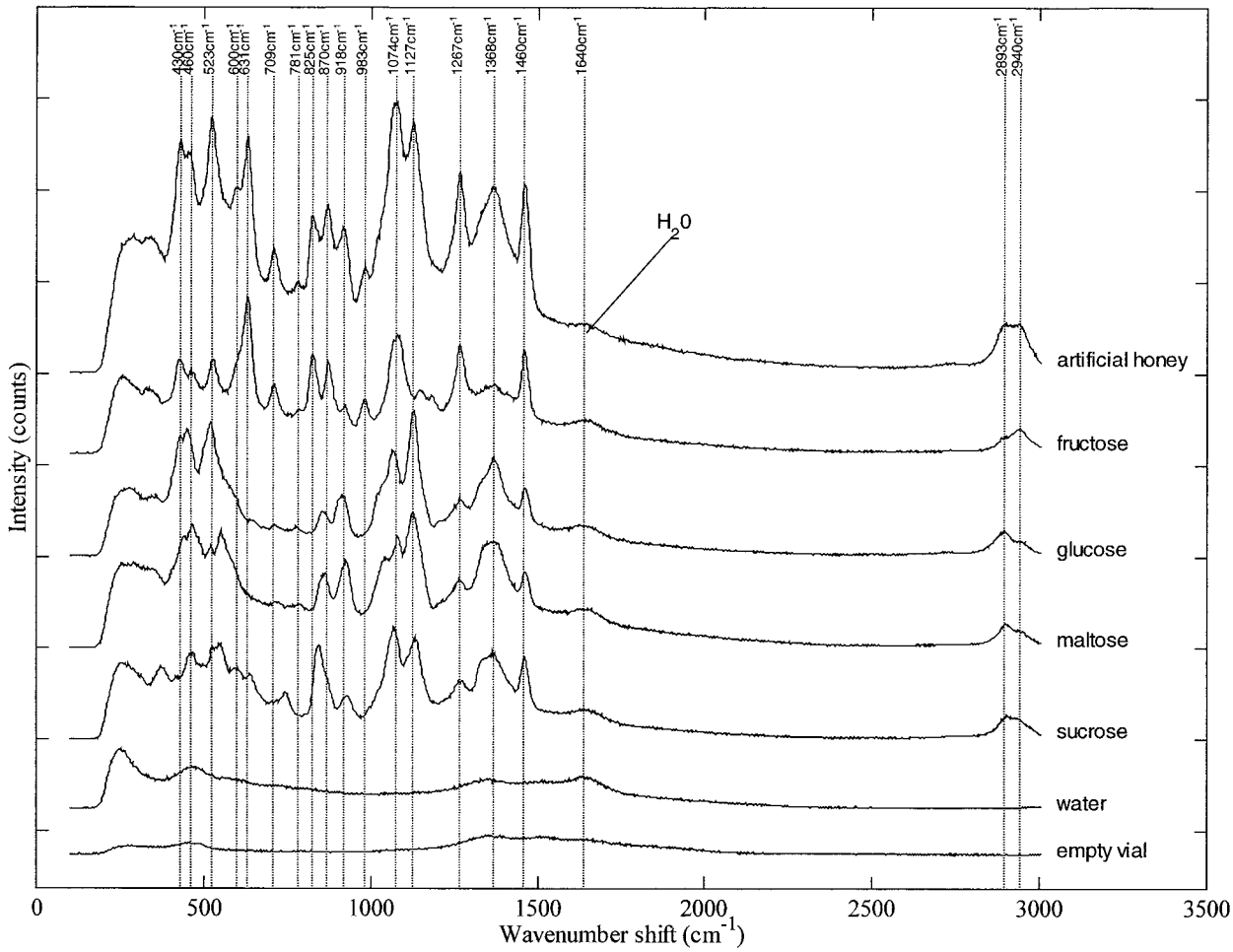

FIG. 2. Stokes Raman spectra of artificial honey (80 g/100 g), and the four common sugars found in honey (40 g/100 g), water, and an empty glass sample carrier.

bands, although some minor additional bands were observed that could arise from the small floral contribution (pollen, proteins, and higher sugars, as well as bees' saliva) also found in honey. Moreover, these may be coincident with the sugar vibrations and thus hard to see.

Because there are approximately 20 Raman Stokes bands that can be attributed to the sugars it is important to try to allocate each frequency to a particular sugar(s) and to assign the Raman frequencies to the specific Raman vibration modes. Therefore, the four pure sugars were dissolved in distilled water at 40 g/100 g and analyzed along with the suitable controls of pure H_2O and an empty glass sample carrier vial. The resulting spectra, along with that from the artificial honey, are shown in Fig. 2; also shown are the wavenumber shifts of the 20 most prominent bands. The first 'hump' at around 250 cm^{-1} is ignored because this is due to the filter cut-off from the laser, and the 'peak' at ~340 cm^{-1} due to fluorescence from impurities in the glass vials is also ignored. Note, of course, from Figs. 1 and 2 that very little fluorescence is seen in the Raman spectra from the honeys, and that the contributions from the glass vial and, indeed, the water are very small.

Table III contains details of the 20 Raman bands seen and the occurrence and strength of each of these in fructose, glucose, maltose, and sucrose. The mono- and disaccharides, although biologically relatively simple, are chemically quite complex molecules and have many Raman active bonds.[34] However, remembering that glucose is a 6-membered ring and fructose a 5-membered ring, while the disaccharide maltose contains two glucose subunits and sucrose comprises glucose and fructose, allows, in consultation with the relevant literature,[36–38] assignment of the Raman bands to specific vibrations from the honeys. Full details of these are found in Table III.

Despite these assignments, the complexity and similarity of all 51 spectra was such that the classification (or clustering) of these spectra would not be possible by simple visual inspection, and this readily illustrates the need to employ chemometric techniques for the cluster analysis of Raman data. The next stage was therefore to use discriminant function analyses (DFA) to observe the relationships between the honey samples as judged from their Raman spectra. Because triplicate measurements for each honey had been made, the 153 spectra that had been recorded were coded so as to give 51 groups, one for each honey (see Table I), and the data were analyzed by DFA as detailed above. The resulting ordination plots of all 51 honeys (see Table I for identifiers) are shown in Fig. 3A. It is clear from this figure that some structure can be seen in the data, but what this relates to can only be seen by plotting the discriminant functions for each

Table III. Proposed identities and occurrence of the Raman bands.

Raman band	Possible identities of the vibration[a]	Found in[b]			
		Fructose	Glucose	Maltose	Sucrose
430 cm^{-1}	skeletal vibration	++	+	−	−
460 cm^{-1}	skeletal vibration	+	−	++	++
523 cm^{-1}	skeletal vibration	+	++	+	+
600 cm^{-1}	skeletal vibration	−	−	−	+
631 cm^{-1}	ring deformation	++	−	−	+
709 cm^{-1}	skeletal vibration	++	−	−	−
781 cm^{-1}	ring vibration	+	+	+	−
825 cm^{-1}	C–OH stretch	++	−	−	−
870 cm^{-1}	C–O–C cyclic alkyl ethers	++	−	−	−
918 cm^{-1}	CH, COH bend	+	++	++	+
983 cm^{-1}	ring ''breathing''	+	−	−	−
1074 cm^{-1}	C–O–C cyclic alkyl ethers	++	+	+	+
1127 cm^{-1}	C–OH deformation	−	++	++	++
1267 cm^{-1}	C–O–C cyclic alkyl ethers	++	+	+	+
1368 cm^{-1}	CH bend + OH bend	−	++	++	++
1460 cm^{-1}	CH_2 bend	++	+	+	+
1640 cm^{-1}	O–H bend from H_2O	+	+	+	+
2893 cm^{-1}	CH bend	−	+	+	+
2940 cm^{-1}	CH_2 bend	+	−	−	+

[a] From Refs. 34, 36–38.
[b] Key: − absent, + medium strength vibration, ++ strong vibration.

of the 51 group means and coding according to floral (Fig. 3B) or geographical (Fig. 3C) origins.

The DFA plot labeled with details of where the honey was produced (Fig. 3C) shows no clustering according to country of origin. Nor was there any evidence of clustering when lower DFs were plotted (data not shown). A possible reason that it was not possible to detect the geographical origin was that the number of representative samples from each country was too small. For example, while 15 honeys were supplied from Italy, six different floral origins of honey were represented: 5 chestnut, 4 acacia, 2 sunflower, 2 eucalyptus, 1 orange, and 1 fior di sulla. It is likely that having this very large (bio)chemical difference within regions will necessarily mean that it will be more difficult to separate samples between regions, a phenomenon observed when using pyrolysis mass spectrometry to investigate the geographical origin of olive oils[39] and honeys.[40]

An obvious question to be asked is ''Is the biochemical signature similar for honeys produced by bees collecting nectar from the same flower?'' Figure 3B shows the floral origin of the honeys and it is clear from this plot that some evidence of botanical origin of the honeys is present. Seven clusters can be seen, which are highlighted in the figure. However, this does require the 'eye-of-faith,' as knowledge of which honey is from which floral origin is needed before the clusters become evident. Moreover, the clusters do overlap and in some cases not all botanical origins cluster together; for example, only three of the five honeys of rape floral origin cluster together (in particular, sample 46 is very different), and one of the acacia honeys (sample 2) is very different from the others. This necessarily means that using simple 'average' Raman spectra to discriminate between the different honeys would likely be unsuccessful.

Figure 4 shows the baseline-corrected Raman spectra (using the multipoint linear baseline correction routine in the GRAMS WiRE software) of two acacia and two rape honeys; these samples have been chosen because of their difference in DFA (Fig. 3B, and above text). Since there was no appreciable background variation or spurious resonance-enhanced bands in these spectra (data not shown) baseline correction was used to attempt to highlight any differences in the key carbohydrate bands. It can be seen (Fig. 4A) that JRC sample 2 has an enhancement in the bands at 460 and 523 cm^{-1} and these can be assigned to skeletal vibrations in maltose/sucrose and glucose respectively (Table III). For the honey samples of rape floral origin, JRC sample 46 has reduced band intensities at 870 and 983 cm^{-1}, which can be assigned to C–O–C cyclic alkyl ethers and ring ''breathing'' from fructose, respectively (Table III), thus indicating that this sample might have a lower content of fructose than the other rape honeys.

Because the interpretation, in terms of the botanical origin of the honey, of the unsupervised cluster analysis method of DFA (unsupervised because the class structure in the DFA were replicates and not origin of honey), used the knowledge of which plant the honey was made from, it seems logical to use this *a priori* knowledge to our advantage before doing the analysis. Initially, experiments using DFA on a subset of the honeys (see Materials and Methods section for details of training and test sets) were calibrated with the *a priori* knowledge of the floral origin of the honeys. However, while the separation of the botanical variety of the honey was successful for a training set (as one would expect for the *calibration* data), projection of the test set into this space (as detailed in Ref. 40) was unsuccessful. Therefore, supervised learning by neural network analysis was conducted as detailed in the Materials and Methods section. Briefly, (1) only those honeys that contained greater than three samples were used as a unifloral output node (seven honeys in total; acacia, chestnut, eucalyptus, heather, lime, and rape or sunflower), while an eighth output node was used to classify honeys of other botanical varieties; (2) because the full original Raman spectra contained 1735 Raman scatters, the number of inputs was reduced by PCA and PCs 1–10, which explained 99.8% of the total variance, were employed; (3) the optimal number of hid-

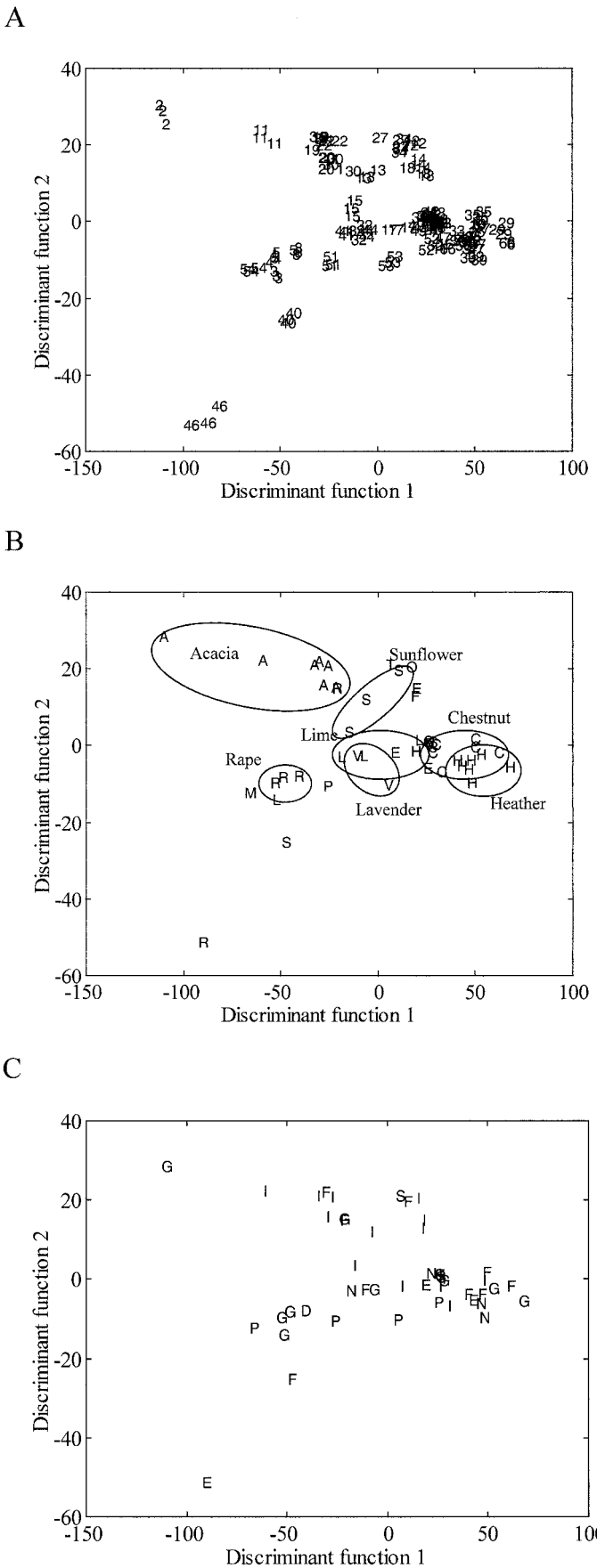

FIG. 3. Results of discriminant function analysis on all 51 honeys; (**A**) coding according to JRC sample number (Table I) with triplicate points shown; the other plots are the means of these DF scores labeled according to botanical (**B**) and geographical (**C**) origins. For (**B**) the codes are: acacia (***A***), chestnut (***C***), citrus (***T***), *Echium plantagineum* (***P***), eucalyptus (***E***), fior di sulla (***F***), heather (***H***), lavender (***V***), lime (***L***), orange (***O***), rape (***R***), rosemary (***M***), and sunflower (***S***). For (**C**) the codes are: Denmark (***D***), England (***E***), France (***F***), Germany (***G***), Italy (***I***), Netherlands (***N***), Portugal (***P***), and Spain (***S***).

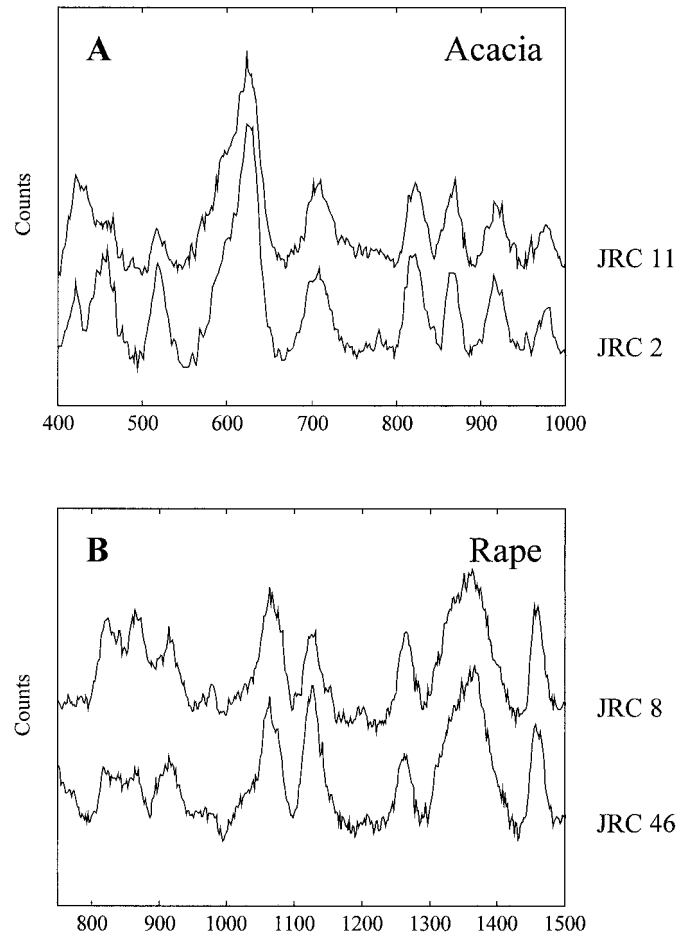

FIG. 4. Baseline-corrected Stokes Raman spectra of (**A**) acacia honeys, samples 2 and 11, and (**B**) rape honeys, samples 8 and 46.

den nodes was determined to be 4, because it was the most parsimonious; and (4) training was conducted five times to 0.15% RMSEC. This process took typically 2–3 $\times$ 10^3 epochs and in real time took only ~2 min to train. Table II shows the average of five different 10-4-8 MLPs. The correct identity was taken to be the winning output node that was given the highest score. As can be seen from this table, 13 of the 14 honeys were classified correctly and only one of the heather honeys was misidentified. Therefore, we believe that honeys of the same botanical origin have a similar biochemical composition and that Raman spectroscopy can be used to identify which floral type the honey comes from.

In conclusion, this study shows that Raman spectroscopy is a very useful tool for the rapid, noninvasive analysis of honey samples. The major vibrational modes in the Stokes Raman spectra were assigned and could be attributed to the four main sugars found in the honeys. Cluster analysis of the spectra with only the knowledge

of which spectra were replicates indicated that the major differences between the honeys were due to their botanical origin rather than their country of origin. This was confirmed by neural network-based analyses, which correctly classified 13 of the 14 honeys in an independent, randomly chosen test set. Finally, we believe that Raman spectroscopy has great potential as a physico-chemical method for the noninvasive nondestructive objective analysis of honey and would be an ideal alternative to the laborious and subjective mellisopalynology typing method currently used to identify the floral origin of honey.

ACKNOWLEDGMENTS

R.G. is indebted to the Engineering and Biological Systems Committee of the UK BBSRC for financial support. We also want to thank Mr. Harald Russmann (Wierz–Eggert–Joerissen GmbH, Hamburg, Germany), who performed the pollen analysis.

1. R. W. Sawyer, J. Assoc. Pub. Analysts **13,** 64 (1975).
2. A. Maurizio, *Honey: a comprehensive survey. Vol. 2,* E. Crane, Ed. (Heinmann, London, 1979), p. 240.
3. V. W. Howells, J. Assoc. Pub. Analysts **7,** 88 (1969).
4. N. B. Colthup, L. H. Daly, and S. E. Wiberly, *Introduction to infrared and Raman spectroscopy* (Academic Press, New York, 1990).
5. J. R. Ferraro and K. Nakamoto, *Introductory Raman Spectroscopy* (Academic Press, London, 1994).
6. B. Schrader, *Infrared and Raman spectroscopy: methods and applications* (Verlag Chemie, Weinheim, 1995).
7. D. D. Archibald, S. E. Kays, D. S. Himmelsbach, and F. E. Barton, Appl. Spectrosc. **52,** 22 (1998).
8. V. Baeten, M. Meurens, M. T. Morales, and R. Aparicio, J. Agric. Food Chem. **44,** 2225 (1996).
9. E. C. Y. LiChan, Trends Food Sci. Technol. **7,** 361 (1996).
10. R. Goodacre, É. M. Timmins, R. Burton, N. Kaderbhai, A. M. Woodward, D. B. Kell, and P. J. Rooney, Microbiology **144,** 1157 (1998).
11. R. Goodacre, A. C. McGovern, N. Kaderbhai, and E. A. Goodacre, *Kohonen Maps,* E. Oja and S. Kaski, Eds. (Elsevier, Amsterdam, 1999), p. 335.
12. A. D. Shaw, N. Kaderbhai, A. Jones, A. M. Woodward, R. Goodacre, J. J. Rowland, and D. B. Kell, Appl. Spectrosc. **53,** 1419 (1999).
13. A. C. McGovern, D. Broadhurst, J. Taylor, N. Kaderbhai, M. K. Winson, D. A. P. Small, J. J. Rowland, D. B. Kell, and R. Goodacre, Biotechnol. Bioeng., paper in press (2002).
14. K. P. J. Williams, G. D. Pitt, B. J. E. Smith, A. Whitley, D. N. Batchelder, and I. P. Hayward, J. Raman Spectrosc. **25,** 131 (1994).
15. K. P. J. Williams, G. D. Pitt, D. N. Batchelder, and B. J. Kip, Appl. Spectrosc. **48,** 232 (1994).
16. C. H. Tseng, J. F. Ford, C. K. Mann, and T. J. Vickers, Appl. Spectrosc. **47,** 1808 (1993).
17. H. Martens and T. Næs, *Multivariate Calibration* (John Wiley, Chichester, 1989).
18. I. T. Jolliffe, *Principal Component Analysis* (Springer–Verlag, New York, 1986).
19. C. Chatfield and A. J. Collins, *Introduction to Multivariate Analysis* (Chapman and Hall, London, 1980).
20. D. R. Causton, *A Biologist's Advanced Mathematics* (Allen and Unwin, London, 1987).
21. H. Wold, *Multivariate Analysis,* K. R. Krishnaiah, Ed. (Academic Press, New York, 1966), p. 391.
22. B. F. J. Manly, *Multivariate Statistical Methods: A Primer* (Chapman and Hall, London, 1994).
23. B. S. Everitt, *Cluster Analysis* (Edward Arnold, London, 1993).
24. D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. de Jong, P. J. Lewi, and J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics: Part A* (Elsevier, Amsterdam, 1997).
25. D. E. Rumelhart, J. L. McClelland, and The PDP Research Group, *Parallel Distributed Processing, Experiments in the Microstructure of Cognition,* vol. I and II (MIT Press, Cambridge, MA, 1986).
26. P. D. Wasserman, *Neural Computing: Theory and Practice* (Van Nostrand Reinhold, New York, 1989).
27. C. M. Bishop, *Neural networks for pattern recognition* (Clarendon Press, Oxford, 1995).
28. M. B. Seasholtz and B. Kowalski, Anal. Chim. Acta **277,** 165 (1993).
29. D. B. Kell and B. Sonnleitner, Trends Biotechnol. **13,** 481 (1995).
30. P. J. Gemperline, J. R. Long, and V. G. Gregoriou, Anal. Chem. **63,** 2313 (1991).
31. M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, and M. Redon, Anal. Chem. **67,** 4477 (1995).
32. R. Goodacre, É. M. Timmins, P. J. Rooney, J. J. Rowland, and D. B. Kell, FEMS Microbiol. Lett. **140,** 233 (1996).
33. R. Goodacre, D. Hammond, and D. B. Kell, J. Anal. Appl. Pyrolysis **40/41,** 135 (1997).
34. B. Schrader, *Raman/Infrared Atlas of Organic Compounds* (Verlag Chemie, New York, 1989).
35. J. W. White, Jr., M. L. Reithof, M. H. Subers, and I. Kushnir, U. S. Dept. of Agriculture Technical Bulletin **1261** (1962).
36. F. R. Dollish, W. G. Fately, and F. F. Bently, *Characteristic Raman frequency of organic compounds* (John Wiley, New York, 1974).
37. D. Lin-Vien, N. B. Colthup, W. G. Fateley, and J. G. Grasselli, *The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules* (Academic Press, Boston, 1991).
38. I. A. Degen, *Tables of chracteristic group frequencies for the interpretation of infrared and Raman spectra* (Acolyte Publications, Harrow, UK, 1997).
39. G. J. Salter, M. Lazzari, L. Giansante, R. Goodacre, A. Jones, G. Surricchio, D. B. Kell, and G. Bianchi, J. Anal. Appl. Pyrolysis **40/41,** 159 (1997).
40. B. S. Radovic, R. Goodacre, and E. Anklam, J. Anal. Appl. Pyrolysis **60,** 79 (2001).
41. R. Goodacre, M. J. Neal, D. B. Kell, L. W. Greenham, W. C. Noble, and R. G. Harvey, J. Appl. Bacteriology **76,** 124 (1994).