JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Application of Cryoprobe ^{1}H Nuclear Magnetic Resonance Spectroscopy and Multivariate Analysis for the Verification of Corsican Honey

JAMES A. DONARSKI, STEPHEN A. JONES, AND ADRIAN J. CHARLTON*

Department for Environment, Food and Rural Affairs, Central Science Laboratory, Sand Hutton, York YO41 1LZ, United Kingdom

Proton nuclear magnetic resonance spectroscopy (^{1}H NMR) and multivariate analysis techniques have been used to classify honey into two groups by geographical origin. Honey from Corsica (Miel de Corse) was used as an example of a protected designation of origin product. Mathematical models were constructed to determine the feasibility of distinguishing between honey from Corsica and that from other geographical locations in Europe, using ^{1}H NMR spectroscopy. Honey from 10 different regions within five countries was analyzed. ^{1}H NMR spectra were used as input variables for projection to latent structures (PLS) followed by linear discriminant analysis (LDA) and genetic programming (GP). Models were generated using three methods, PLS-LDA, two-stage GP, and a combination of PLS and GP (PLS-GP). The PLS-GP model used variables selected by PLS for subsequent GP calculations. All models were generated using Venetian blind cross-validation. Overall classification rates for the discrimination of Corsican and non-Corsican honey of 75.8, 94.5, and 96.2% were determined using PLS-LDA, two-stage GP, and PLS-GP, respectively. The variables utilized by PLS-GP were related to their ^{1}H NMR chemical shifts, and this led to the identification of trigonelline in honey for the first time.

KEYWORDS: Honey; NMR; metabolomics; genetic programming; PLS; geographical origin; food authenticity; chemometrics

INTRODUCTION

The country of origin of honey sold in the European Union (EU) must be clearly presented on the product label, as prescribed by EU legislation (Council Directive 2001/110/EC). This legislation is required because the geographical origin of the honey is linked to consumer perception of its quality with protected designation of origin (PDO) honeys often attracting a premium price. The verification of labeling claims relating to geographical origin utilizes diverse food traceability systems to address EU legislation (2002/178/EC). In support of this legislation, the EU-funded TRACE project was initiated (www.trace.eu.org). The TRACE project aims to "improve the health and well-being of European citizens by delivering improved traceability of food products". In accordance with the aims of the TRACE project, analytical fingerprinting techniques are being exploited to confirm the geographical origin of foodstuffs.

Recent reviews of current analytical techniques applied to the geographical origin determination of foodstuffs have been presented by Reid et al. (*1*) and Luyex et al. (*2*). The geographical origin of honey has previously been defined by a range of analytical parameters, including elemental composition, isotopic ratios (*3*), and complex chemical fingerprints (*4*). Honey is a complex matrix reported to consist of at least 200 compounds (*5*). These compounds are particularly diverse and include sugars (*6, 7*), flavonoids (*8*), enzymes (*5*), organic acids (*9*), amino acids (*10*), phenols, and polyphenols (*11*). The floral origin (*12–15*) and the species of bee that produces the honey (*16*) influence the range and distribution of the compounds that are present in honey. The determination of the concentration of these compounds often requires specific analytical methods for each compound type. Complex sample pretreatment is frequently required to concentrate the analytes of interest and to remove interfering compounds (*12–14*).

Here, the determination of the geographical origin of PDO honey from Corsica has been undertaken using ^{1}H NMR spectroscopy and multivariate analysis techniques. The approach taken requires minimal pretreatment of the honey, potentially expediting geographical origin confirmation. ^{1}H NMR spectroscopy facilitates the determination of highly specific and quantitative fingerprints relating to the composition of the product being analyzed. NMR spectroscopy also provides a range of measurements with which to resolve the chemical structure of compounds present in complex mixtures, and thus, NMR fingerprint data are highly interpretable.

NMR spectroscopic data from Corsican and non-Corsican honey were analyzed using three supervised statistical techniques: projection to latent structures−linear discriminant

* To whom correspondence should be addressed. Tel: +44(0)1904 462513. Fax: +44(0)1904 462133. E-mail: adrian.charlton@csl.gov.uk.

10.1021/jf072402x CCC: $40.75 Published 2008 by the American Chemical Society
Published on Web 06/20/2008

Table 1. Table Showing the Distribution of the Honey Samples

region	no. of samples	country	no. of samples
Marchfeld	8	Austria	18
Muehlviertel	10		
Carpentras	7	France	129
Corsica	111		
Limousin	11		
Bavaria	18	Germany	18
Galway	2	Ireland	2
Sicily	5	Italy	15
Trentino	5		
Tuscany	5		

Table 2. Specific Types and Sample Numbers of Corsican Honey[a]

Corsican honey type	no. of samples analyzed
Arbousier (strawberry tree)	1
Châtaigneraie (chestnut)	20
Maquis d'automne (autumn scrubland)	17
Miellats du maquis (honeydew scrubland)	19
Maquis de printemps (spring scrubland)	15
Maquis d'été (summer scrubland)	6
Printemps (nonspecific spring)	14
Non AOC (nonspecific)	15
Printemps—Clémentinier (spring clementine)	4

[a] The English translation of the honey type is presented in parentheses where appropriate.

analysis (PLS-LDA) (*17*), two-stage genetic programming (two-stage GP) (*18*), and a novel combination of PLS and GP (PLS-GP). The results from the PLS-GP model were used to determine the presence of resonance peaks that confirmed the geographical origin of Corsican honey. This study uses Corsican PDO honey to demonstrate the use of ^{1}H NMR spectroscopy and mathematical modeling methods for the verification of product origin.

EXPERIMENTAL PROCEDURES

Materials. All chemicals used were of analytical grade (≥99% purity). Deuterium oxide (^2H_2O) was supplied by Goss Scientific Instruments Ltd. (Cheshire, United Kingdom), 3-trimethylsilyl[2,2,3,3-$^2H_4$] propionic acid (TSP) was supplied by Avocado Research Chemicals Ltd. (Lancashire, United Kingdom), sodium azide was supplied by Sigma-Aldrich Co. (Dorset, United Kingdom), and dipotassium hydrogen phosphate and dihydrogen potassium phosphate were supplied by BDH Chemicals Ltd. (Dorset, United Kingdom). Ultrapure water was provided from an Elga Option 2 water purifier.

Samples. One hundred and eighty-two honey samples were collected from five countries, consisting of 10 different regions. The geographical origin and the number of samples collected in each region or country are given in **Table 1**. The Corsican honeys are representative of the main production areas and floral types within Corsica. The honey types and the number of samples analyzed are presented in **Table 2**.

Sample Preparation. Honey (50 g) that had previously been diluted with distilled water to 70° Brix was homogenized using a Turrax mixer (11000 RPM) for five 20 s periods. To prevent heating, a period of 5 min separated each mixing period. The honey was then placed onto a rolling mixer (35 RPM) for 16 h until homogeneous (data not shown).

The homogeneous honey samples were diluted to 15 ± 0.5° Brix (approximately 0.25 g of honey was dissolved into 1 mL of deionized water). The diluted samples were centrifuged to remove any remaining suspended material before filtration through a 0.2 μm PTFE syringe filter. To prepare the sample for NMR spectroscopy, 480 μL of filtered honey solution, 60 μL of sodium azide solution (10 mM dissolved in ^2H_2O), and 60 μL phosphate buffer solution (250 mM, pH 7.2 containing 10 mM TSP dissolved in ^2H_2O) were added to a labeled 5 mm diameter NMR tube. The tube contents were thoroughly mixed using a vortex mixer before ^{1}H NMR spectroscopy.

Spectral Acquisition. All experiments were carried out using a Bruker Avance 500 MHz NMR spectrometer equipped with a TCI cryoprobe. Spectra were acquired at a central frequency of 500.1323546 MHz using on-resonance presaturation to suppress the intensity of the water signal. A 60° observation pulse length of 5.1 μs and a delay between transients of 14 s were used. A total of 32768 complex data points were acquired with a spectral width of 14 ppm, giving an acquisition time of 4.6794 s. A recycle time of 18.7 s was determined experimentally to produce quantitative data with optimized sensitivity. Eight unrecorded (dummy) transients and 256 acquisition transients were used, giving a total experiment time of approximately 1 h and 20 min. One-dimensional (1D) ^{1}H NMR spectroscopic data were processed using FELIX software (Accelrys, San Diego, CA). A sine bell-shaped window function phase shifted by 90° was applied over all data points before Fourier transformation, phase, and baseline correction. The chemical shift of all data was referenced to the TSP resonance at 0 ppm. The area of this resonance was set to unity for all spectra acquired.

Two-dimensional (2D) ^{1}H−^{1}H total correlation spectroscopy (TOCSY) data were acquired using the standard spectrometer library pulse sequence (*19*) with 56 transients, 384 increments, a spectral width of 13.33 ppm, and a spin-lock time of 100 ms. In the F2 dimension, 4096 datapoints were collected, and the F1 dimension was zero-filled to give 2048 data points. Optimized pulse lengths were calibrated for individual experiments, and on-resonance presaturation was used to suppress the intensity of the water signal.

Statistical Analysis. Statistical analysis was completed using a custom written graphical user interface (GUI) for Matlab (The Mathworks, Natick, MA; Version 7.4.0.287 [R2007a]) known as Metabolab. Statistical analyses were completed using the full NMR spectrum excluding data from the residual water resonance (4.7−4.9 ppm).

PLS-LDA. PLS (*20*) is a supervised multivariate method for the determination of combinations of variables that result in optimal separation of specified experimental groupings. A linear distance metric applied following a PLS calculation constitutes PLS-LDA. PLS-LDA was performed using dummy *Y* variables that represent membership of either Corsican or non-Corsican groupings. Ten PLS components were calculated using the NIPALS algorithm (*21*). Preprocessing methods were evaluated, and the data were set to unit variance as this was found to produce the most robust predictive models. The Mahalanobis distance metric was used for LDA.

PLS models were constructed using 90% of the NMR data. The omitted 10% of the data were used to determine the cross-validation classification rate. This process was repeated iteratively until all data had been used for cross-validation, resulting in 10 PLS models containing 1−10 PLS components. This approach is often referred to as Venetian blinds cross-validation. Cross-validation rates were the number of correct predictions made by the PLS-LDA model expressed as a percentage of the total number of predictions made per model and presented as the mean for the 10 PLS-LDA models used for cross-validation. Overfitting was determined to have occurred when the cross-validation rate began to decline when further PLS components were added to the calculation. The number of PLS components used was thus determined to be one fewer than the point at which this occurred.

Two-Stage GP (18). GP (*22*) is a computational learning technique based on Darwin's theory of evolution, the output of which is a classification tree. The GP tree is a symbolic expression that can easily be assessed to determine the variables from the input data that were utilized for classification. In two-stage GP, the first stage used all variables present in the data set. The second stage used only those variables within a specified tolerance range, used by the final trees of the first stage.

Each round of the two GP stages was completed using an island population of 300, five island populations, 10 generations between island migration, an island migration rate of 10%, a maximum GP tree size of eight, and a tree mutation rate of 50%. At each generation, trees were ordered by the value of their fitness function, which in this case was the classification rate. The top 10% of GP trees were kept, and the bottom 10% were removed. The stopping conditions enforced were either classification of 100% of the data or evolution for 200 generations. The first stage of GP used all of the variables within the ^{1}H NMR spectrum. The variables used in the second stage were constrained to

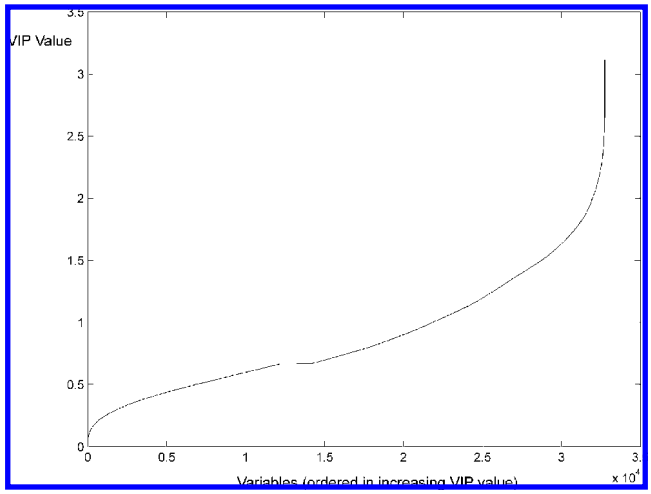

Figure 1. Plot of the VIP scores used in the PLS-LDA model for Corsican honey classification. Variables are arranged in order of increasing VIP value.

those that were used in the generation of the final GP trees of the first stage, with a tolerance on variable position of 10 units. Venetian blind cross-validation was used omitting 10% of the data from each calculation for subsequent validation.

PLS-GP. GP was performed using input variables determined from a PLS calculation that included all of the ^{1}H NMR data. To determine the input variables used in the PLS-GP model, a plot of the numerically ordered variable importance of projection (VIP) values from all variables was constructed (**Figure 1**). A threshold was determined by visual examination of **Figure 1**, and all variables with a VIP score of greater than or equal to 1.67 were used to constrain the variables used by GP. The conditions imposed on the GP algorithm were those used in the two-stage GP. Venetian blind cross-validation was used omitting 10% of the data from each calculation for subsequent validation.

RESULTS AND DISCUSSION

1D ^{1}H NMR Spectra. The 1D ^{1}H NMR spectrum of a Corsican honey is shown in **Figure 2A**. Glucose and fructose resonances are labeled. The resonance observed at 0.0 ppm is from the internal standard, TSP. Many other resonances are observed in the NMR spectrum of the honey, and these are largely shown in **Figure 2B**,**C**.

Statistical Results. The classification results obtained using PLS-LDA, two-stage GP, and PLS-GP are presented in **Table 3** as the percentage of the honeys that were correctly classified upon cross-validation. These data are presented according to sensitivity (Corsican samples correctly classified as Corsican samples), selectivity (non-Corsican samples correctly classified as non-Corsican samples), and overall correct classification (all samples correctly classified by the model). The classification rate that would be expected by chance is 50% for sensitivity, selectivity, and overall classification rate. Although the sample numbers were not equally balanced (61.0% of the samples were Corsican, and 39.0% of the samples were non-Corsican), the overall classification rates determined were significantly higher than those that would be expected by probability alone. Therefore, it is assumed that the imbalance in sample numbers did not significantly alter the efficacy of the models.

PLS-LDA. Classification rates for the sensitivity and selectivity were determined for the training and validation sets and from each model generated. The classification rate obtained from the validation data was highest when only two components were used; therefore, the model generated using two components was

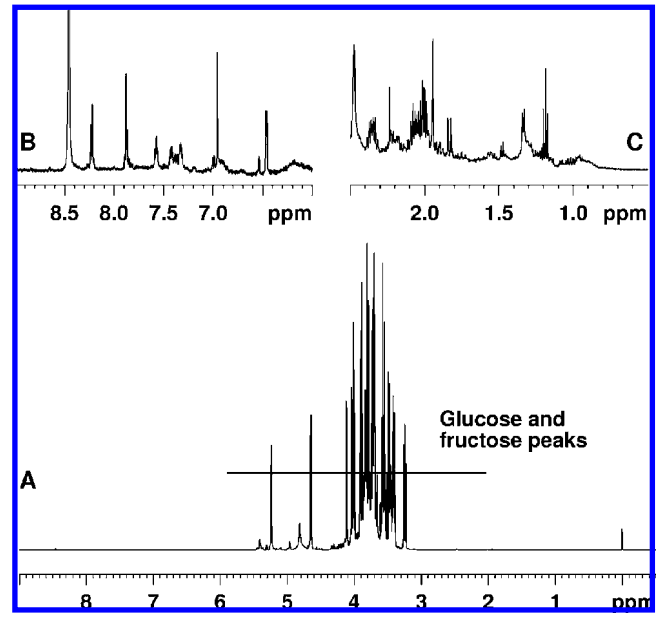

Figure 2. 1D ^{1}H NMR spectrum of a Corsican honey. (**A**) Whole NMR spectrum, all resonances above the horizontal line are from either glucose or fructose. Panels **B** and **C** are expansions of the spectral regions 6.0−9.0 and 0.5−2.5 ppm, respectively. These regions have been magnified along the vertical scale to show the presence of nonglucose or fructose resonances present in the honey spectrum.

Table 3. Overview of the Classification Rates upon Cross-Validation Determined for Corsican Honey Using PLS-LDA (Two Components), Two-Stage GP, and PLS-GP Models

model	sensitivity (%)	selectivity (%)	overall (%)
PLS-LDA	72.1	81.6	75.8
two-stage GP	96.4	91.5	94.5
PLS-GP	98.2	93.0	96.2

chosen as the best PLS-LDA model that was not over fitted. The first two components captured 76.6% of the variance in the data.

A PLS scores plot was generated using components 1 and 2 to visualize the separation of the data and is shown in **Figure 3**. The classification results obtained during cross-validation from the PLS-LDA model using two PLS components are presented in **Table 3**.

The overall classification rate that was obtained by PLS-LDA of 75.8% is significantly greater than would be expected by chance, and thus, the PLS calculation was able to identify traits in the NMR data that relate to the geographical origin of the honey. However, because of the relatively low classification rate obtained using PLS-LDA, when compared to two-stage GP and PLS-GP, the PLS-LDA models were not further interpreted.

Two-Stage GP. The classification results obtained by two-stage GP are presented in **Table 3**. To visualize the variables that were utilized in the generation of the GP trees, a histogram showing the frequency of occurrence of the variables used in the 10 GP trees, created during cross-validation, is shown in **Figure 4A**. These variables were compared to those used in the PLS-GP model. The overall classification rate that was obtained using two-stage GP was 94.5%, providing a strong indication that the NMR data can be used to verify the origin of Corsican honey.

PLS-GP. The classification results obtained using a combination of PLS and GP were higher than either method used alone and are presented in **Table 3**. The overall cross-validated

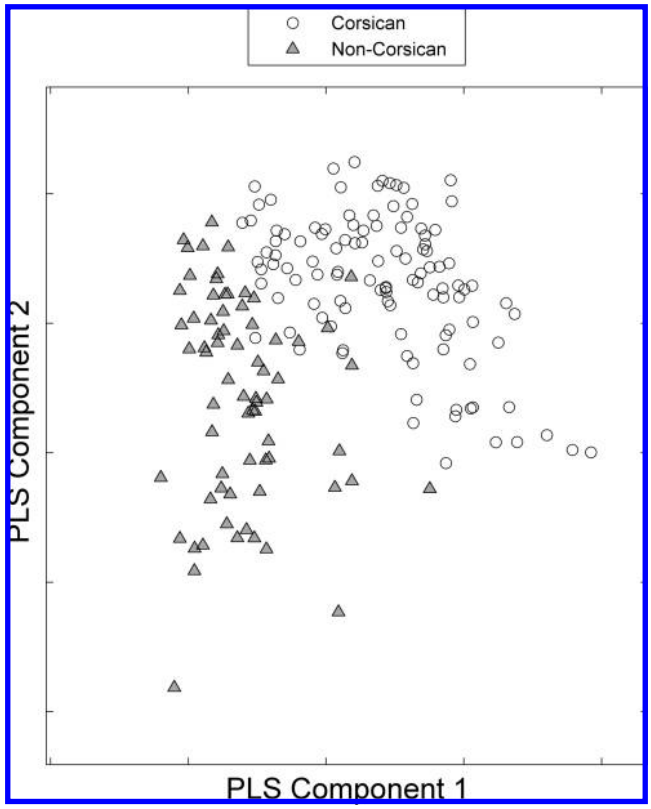

Figure 3. PLS-LDA scores plot of the Corsican honey classification model. The axes of the plot are PLS-LDA components 1 and 2.

classification rate using this combined approach was 96.2%. A histogram showing the frequency of occurrence of the variables used in the 10 GP trees, created during cross-validation, is shown in **Figure 4B**. The variables chosen by the PLS-GP model are similar to those chosen by the two-stage GP, although the PLS-GP model used a reduced number in the predictive trees, aiding the identification of the most pertinent variables. Analysis of the 10 individual GP trees generated by this model was completed, and the tree with the best classification rate was determined. A graphical representation of this tree, which correctly classified the honey data set with classification rates of 97.3, 97.2, and 97.3% for the sensitivity, selectivity, and overall rate, respectively, is shown in **Figure 5**.

Corsican Honey Biomarkers. The variables used by the most successful PLS-GP model were converted to their corresponding 1D ^{1}H NMR chemical shifts. The typical peak width at half-height for a singlet resonance present in the honey ^{1}H NMR data was approximately 1.1 Hz. The spectral resolution was 0.21 Hz; therefore, each NMR resonance corresponded to a minimum of 10 variables. Variables corresponding to the same NMR resonance were grouped, and their combined frequency of occurrence within the cross-validation GP trees was calculated. The honey NMR spectra were examined to determine the significance of these resonances in the determination of the geographical origin of the honey. A description of the significance of the highest occurring variables is presented in **Table 4**.

The identification of compounds giving rise to resonances present in the honey NMR spectra was undertaken using both 1D and 2D NMR techniques. Where standard compounds were available, spiking experiments were used to confirm tentative NMR assignments made by comparison to spectral databases and by data interpretation. 2D ^{1}H−^{1}H TOCSY NMR spectra were acquired of honeys containing elevated amounts of each

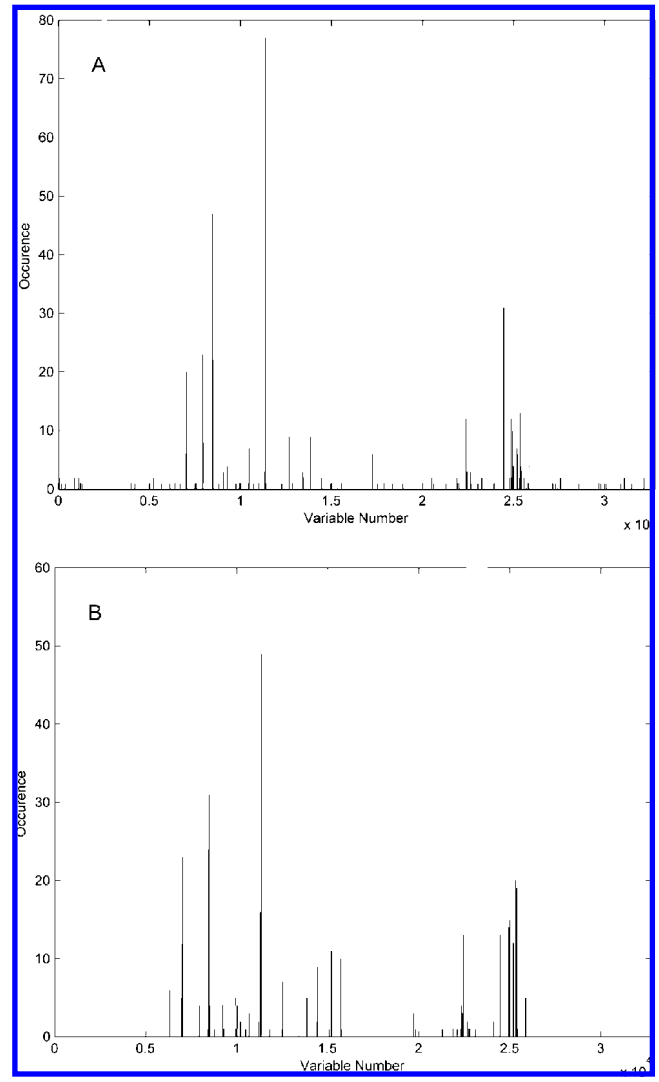

Figure 4. Plot of the number of times each variable was used in the final Corsican honey classification models generated by (**A**) two-stage GP and (**B**) PLS-GP.

variable listed in **Table 4**. Where TOCSY correlations were observed, the chemical shifts of the coupled protons are presented.

The chemical environment was determined from chemical shift information (*23*). Where available, the relative integrals, peak multiplicity (d = doublet, dd = double doublet, m = multiplet, s = singlet, and t = triplet) and coupling constants are presented in parentheses. Data assessment focused on those variables listed in **Table 4** and is presented below in the order of the frequency of occurrence of the variables in the PLS-GP trees.

Variables 8425−8507. This corresponds to a resonance at δ = 8.221 (1H, d, 8.3 Hz). This resonance was coupled to resonances at 7.873, 7.865 (overlapping resonances 2H, m), and 7.571 (1H, m) ppm. Another resonance, δ = 6.952 (1H, s) ppm, was also from the same molecule. It was concluded that this molecule was a fused aromatic ring system, possibly a coumarin derivative substituted at either the 5- or the 6-position within the fused ring structure. Coumarin has previously been found in French lavender honey and was proposed as a marker for honey age (*24*). Acquisition of NMR spectroscopic data from a standard determined that these resonances were not from coumarin.

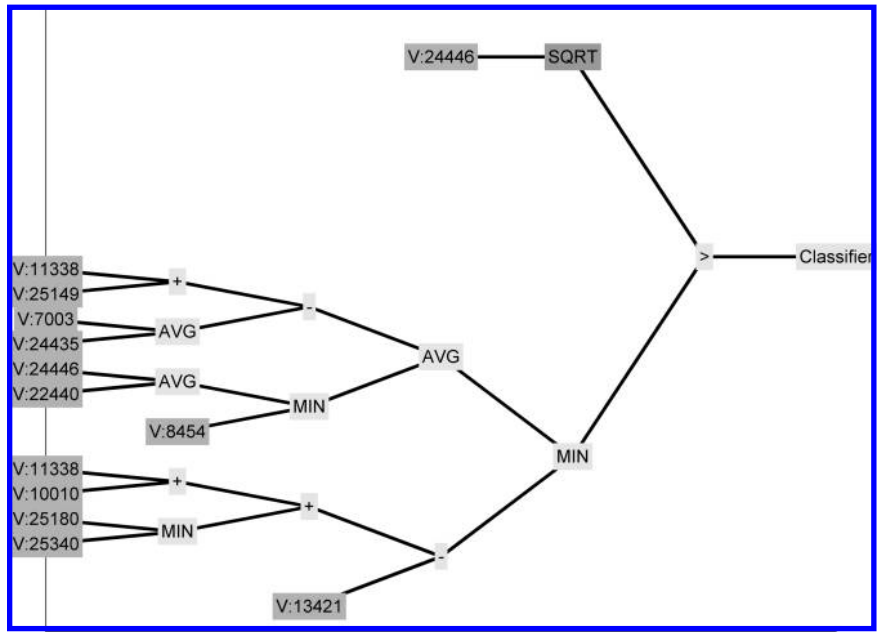

Figure 5. GP tree for the classification of Corsican honey using PLS-GP. The overall classification rate obtained by this tree when used to classify the honey as Corsican or non-Corsican is 97.3%. V, variable; SQRT, square root; +, summation; −, subtraction; min, minimum; and AVG, average.

Table 4. Highest Occurring Variables Used in the PLS-GP Model and a Brief Description of Their Significance for the Classification of Corsican Honey

variable	^{1}H chemical shift range (ppm)	occurrence in PLS-GP	comment[a]
8425–8507	8.1989–8.234	233	elevated Corsican marker
0–5	11.832–11.834	208	no resonances, baseline value
11299–11343	6.987–7.006	173	elevated Corsican marker
25307–25432	0.967–1.020	115	elevated Corsican marker
24435–24474	1.376–1.393	80	elevated non-Corsican marker
25148–25189	1.070–1.088	60	elevated Corsican marker
7003–7027	8.831–8.842	57	elevated Non-Corsican marker—mainly Austria/Germany
22316–22448	2.244–2.298	51	decreased in non-Corsican samples—Austria

[a] Comments are given to describe the significance of each variable range. No comment is applicable to all samples but describes a significant subset of each region or country.

Variables 0−5. NMR resonances were not observed in any of the honey spectra at this chemical shift range. The intensities at these positions represent typical baseline values.

Variables 11299−11343. This variable range corresponds to a resonance at $\delta = 6.997$ (1H, d, 8.7 Hz) ppm that is coupled to a resonance at $\delta = 7.350$ (1H, d, 8.7 Hz) ppm. It was concluded that these resonances were from either a 1,4- or a 1,2,3,4-substituted aromatic molecule. Aromatic compounds matching this description that have been demonstrated to be present in honey include *p*-hydroxybenzoic acid and *p*-coumaric acid (*25*). Acquisition of data from standards of these compounds determined that neither gave rise to these variables.

Variables 25307−25432, 24435−24474, and 25148−25189. These variable ranges correspond to resonances at $\delta = 1.018$ (1H, s), $\delta = 1.376$ (1H, s), and $\delta = 1.080$ (1H, s) ppm, respectively. The chemical shift ranges of these resonances suggest that they are from methyl groups.

Variables 7003−7027. This variable range corresponds to a resonance at $\delta = 9.131$ (1H, s), which was coupled to resonances at $\delta = 8.848$, 8.834 (overlapping resonances, 2H, m), 8.088 (1H, m), and 4.440 (3H, s) ppm. It was hypothesized that these resonances were from trigonelline. A spiking experiment was performed (**Figure 6**), and the identification of this biomarker as trigonelline was confirmed. The samples containing trigonelline were from Germany, Austria, and France (Limousin), where nine, six, and one sample contained trigonelline, respectively. Trigonelline is a plant hormone generally present in herbaceous species of saline and dry habitats (*26*) and is known to accumulate in salt-stressed plants (*27*).

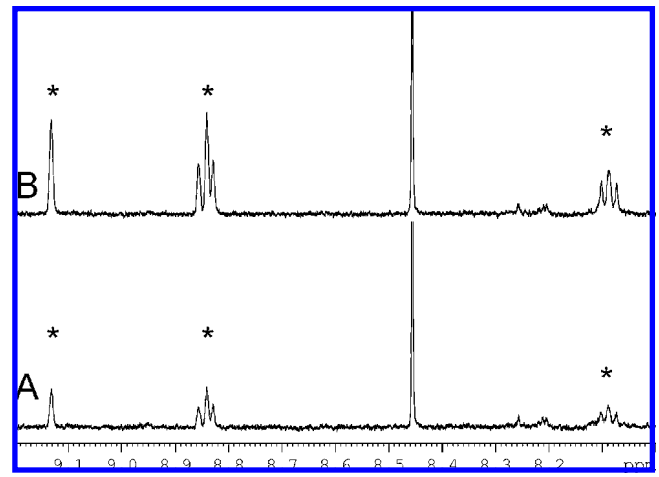

Figure 6. 1D ^{1}H NMR spectrum of the region 8.0−9.2 ppm of (**A**) Bavarian honey and (**B**) Bavarian honey spiked with trigonelline. Resonances observed from trigonelline are labeled*.

Variables 22316−22448. Analysis of this region determined that it contained several resonances and that the profile of these resonances varied between samples. Therefore, a specific resonance assignment was not possible.

This study has demonstrated the feasibility of developing accurate models that can be used for the identification of Corsican honey using ^{1}H NMR spectroscopy. The most accurate of these models was produced using a combination of PLS and GP with PLS used as a variable selection step to determine input variables for GP. This approach resulted in an overall correct classification rate for the verification of Corsican honey of 96.2% upon cross-validation, with individual GP trees classifying with an overall rate of up to 97.3%. The PLS-GP model was easily

interpretable and was used to identify the variables (and the resulting NMR spectroscopic resonances) used for classification. NMR spectroscopy has been used to determine the molecular structural characteristics of those compounds that gave rise to the variables that were most pertinent to the classification of Corsican honey. Trigonelline is reported for the first time in honey. Trigonelline may prove to be indicative of geographical origin (saline habitat) or growth conditions (dry habitat). These results have also identified the presence of four biomarkers of Corsican honey that will be further interrogated in future studies to fully elucidate their structure.

The methods that have been presented here clearly provide a basis for the use of molecular fingerprinting for the determination of the origin of food. Often the determination of geographical origin can be complicated by the incomplete correlation between analytical parameters and geographical boundaries. Here, the data generated have been interpreted in the context of honey derived from the island of Corsica. However, the approach taken may be more appropriate for the determination of botanical and processing factors that are often regional in nature and have a highly specific impact on the overall molecular composition of foodstuffs.

ACKNOWLEDGMENT

The honey samples utilized in this research were collected as part of the TRACE project by the Czech Bee Research Institute. The samples were subsampled and distributed by The Irish Agriculture and Food Development Authority.

LITERATURE CITED

(1) Reid, L. M.; O'Donnell, C. P.; Downey, G. Recent technological advances for the determination of food authenticity. *Trends Food Sci. Technol.* **2006**, *17*, 344–353.

(2) Luykx, D. M. A. M.; van Ruth, S. M. An overview of analytical methods for determining the geographical origin of food products. *Food Chem.* **2008**, *107*, 897–911.

(3) Kelly, S.; Heaton, K.; Hoogewerff, J. Tracing the geographical origin of food: The application of multi-element and multi-isotope analysis. *Trends Food Sci. Technol.* **2005**, *16*, 555–567.

(4) Woodcock, T.; Downey, G.; Kelly, J. D.; O'Donnell, C. Geographical classification of honey samples by near-infrared spectroscopy: A feasibility study. *J. Agric. Food Chem.* **2007**, *55*, 9128–9134.

(5) White, J. W. *Honey, A Comprehensive Survey*; Crane, Russak & Company: New York, 1979.

(6) Cotte, J. F.; Casabianca, H.; Chardon, S.; Lheritier, J.; Grenier-Loustalot, M. F. Application of carbohydrate analysis to verify honey authenticity. *J. Chromatogr. A* **2003**, *1021*, 145–155.

(7) Mazzoni, V.; Bradesi, P.; Tomi, F.; Casanova, J. Direct qualitative and quantitative analysis of carbohydrate mixtures using C-13 NMR spectroscopy: Application to honey. *Magn. Reson. Chem.* **1997**, *35*, S81−S90.

(8) Tomas-Barberan, F. A.; Martos, I.; Ferreres, F.; Radovic, B. S.; Anklam, E. HPLC flavonoid profiles as markers for the botanical origin of European unifloral honeys. *J. Sci. Food Agric.* **2001**, *81*, 485–496.

(9) Cherchi, A.; Spanedda, L.; Tuberoso, C.; Cabras, P. Solid-phase extraction and high-performance liquid-chromatographic determination of organic-acids in honey. *J. Chromatogr. A* **1994**, *669*, 59–64.

(10) Paramas, A. M. G.; Barez, J. A. G.; Marcos, C. C.; Garcia-Villanova, R. J.; Sanchez, J. S. HPLC-fluorimetric method for analysis of amino acids in products of the hive (honey and bee-pollen). *Food Chem.* **2006**, *95*, 148–156.

(11) Andrade, P.; Ferreres, F.; Amaral, M. T. Analysis of honey phenolic acids by HPLC, its application to honey botanical characterization. *J. Liq. Chromatogr. Relat. Technol.* **1997**, *20*, 2281–2288.

(12) Andrade, P.; Ferreres, F.; Gil, M. I.; Tomas-Barberan, F. A. Determination of phenolic compounds in honeys with different floral origin by capillary zone electrophoresis. *Food Chem.* **1997**, *60*, 79–84.

(13) Dimitrova, B.; Gevrenova, R.; Anklam, E. Analysis of phenolic acids in honeys of different floral origin by solid-phase extraction and high-performance liquid chromatography. *Phytochem. Anal.* **2007**, *18*, 24–32.

(14) Foldhazi, G. Analysis and quantitation of sugars in honey of different botanical origin using high-performance liquid-chromatography. *Acta Aliment.* **1994**, *23*, 299–311.

(15) Gil, M. I.; Ferreres, F.; Ortiz, A.; Subra, E.; Tomas-Barberan, F. A. Plant phenolic metabolites and floral origin of rosemary honey. *J. Agric. Food Chem.* **1995**, *43*, 2833–2838.

(16) Souza, B.; Roubik, D.; Barth, O.; Heard, T.; Enriquez, E.; Carvalho, C.; Villas-Boas, J.; Marchini, L.; Locatelli, J.; Persano-Oddo, L.; Meida-Muradian, L.; Bogdanov, S.; Vit, P. Composition of stingless bee honey: Setting quality standards. *Interciencia* **2006**, *31*, 867–875.

(17) Al-Jowder, O.; Kemsley, E. K.; Wilson, R. H. Detection of adulteration in cooked meat products by mid-infrared spectroscopy. *J. Agric. Food Chem.* **2002**, *50*, 1325–1329.

(18) Davis, R. A.; Charlton, A. J.; Oehlschlager, S.; Wilson, J. C. Novel feature selection method for genetic programming using metabolomic H-1 NMR data. *Chemom. Intell. Lab.* **2006**, *81*, 50–59.

(19) Bax, A.; Davis, D. G. Mlev-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson.* **1985**, *65*, 355–360.

(20) Sjostrom, M.; Wold, S.; Lindberg, W.; Persson, J. A.; Martens, H. A multivariate calibration problem in analytical chemistry solved by partial least squares models in latent variables. *Anal. Chim. Acta* **1983**, *150*, 61–70.

(21) Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab.* **1987**, *2*, 37–52.

(22) Koza, J. R. Genetic programming as a means for programming computers by natural selection. *Stat. Comput.* **1994**, *4*, 87–112.

(23) Abraham, R. J.; Fisher, U.; Loftus, P. The chemical shift. In *Introduction to NMR Spectroscopy*, 2nd ed.; John Wiley & Sons: New York, 1991; pp 13−29.

(24) Guyot-Declerck, C.; Renson, S.; Bouseta, A.; Collin, S. Floral quality and discrimination of *Lavandula stoechas*, *Lavandula angustifolia*, and *Lavandula angustifoliaxlatifolia* honeys. *Food Chem.* **2002**, *79*, 453–459.

(25) Gheldof, N.; Wang, X. H.; Engeseth, N. J. Identification and quantification of antioxidant components of honeys from various floral sources. *J. Agric. Food Chem.* **2002**, *50*, 5870–5877.

(26) Dini, I.; Tenore, G. C.; Trimarco, E.; Dini, A. Two novel betaine derivatives from Kancolla seeds (Chenopodiaceae). *Food Chem.* **2006**, *98*, 209–213.

(27) Tramontano, W. A.; Jouve, D. Trigonelline accumulation in salt-stressed legumes and the role of other osmoregulators as cell cycle control agents. *Phytochemistry* **1997**, *44*, 1037–1040.

Received for review August 10, 2007. Revised manuscript received April 28, 2008. Accepted April 29, 2008. We thank the European Commission, through the Sixth Framework Programme under the Food Quality and Safety Priority as part of the TRACE project (www.trace.eu.org), and the UK Food Standards Agency, for funding this work. The information contained in this article reflects the authors' views; the European Commission is not liable for any use of the information contained therein.

JF072402X