Talanta 128 (2014) 284–292

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Practical procedure for discriminating monofloral honey with a broad pollen profile variability using an electronic tongue

Mara E.B.C. Sousa [a], Luís G. Dias [a,*], Ana C.A. Veloso [b,c], Letícia Estevinho [a], António M. Peres [a,d], Adélio A.S.C. Machado [e]

[a] CIMO – Mountain Research Centre, Escola Superior Agrária, Instituto Politécnico de Bragança, Campus Santa Apolónia, Apartado 1172, 5301-855 Bragança, Portugal
[b] Instituto Politécnico de Coimbra, ISEC, DEQB, Rua Pedro Nunes, Quinta da Nora, 3030-199 Coimbra, Portugal
[c] CEB – Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal
[d] LSRE – Laboratory of Separation and Reaction Engineering – Associate Laboratory LSRE/LCM, Escola Superior Agrária, Instituto Politécnico de Bragança, Campus Santa Apolónia, Apartado 1172, 5301-855 Bragança, Portugal
[e] LAQUIPAI – Laboratório de Química Inorgânica Pura e de Aplicação Interdisciplinar, Departamento de Química, Faculdade de Ciências da Universidade do Porto, R. Campo Alegre, 687, 4169-007 Porto, Portugal

ARTICLE INFO

Article history:
Received 10 January 2014
Received in revised form
28 April 2014
Accepted 4 May 2014
Available online 10 May 2014

Keywords:
Electronic tongue
Monofloral honey
Honey colour
Linear discriminant analysis
Simulated annealing algorithm

ABSTRACT

Colour and floral origin are key parameters that may influence the honey market. Monofloral light honey are more demanded by consumers, mainly due to their flavour, being more valuable for producers due to their higher price when compared to darker honey. The latter usually have a high anti-oxidant content that increases their healthy potential. This work showed that it is possible to correctly classify monofloral honey with a high variability in floral origin with a potentiometric electronic tongue after making a preliminary selection of honey according their colours: white, amber and dark honey. The results showed that the device had a very satisfactory sensitivity towards floral origin (*Castanea* sp., *Echium* sp., *Erica* sp., *Lavandula* sp., *Prunus* sp. and *Rubus* sp.), allowing a leave-one-out cross validation correct classification of 100%. Therefore, the E-tongue shows potential to be used at analytical laboratory level for honey samples classification according to market and quality parameters, as a practical tool for ensuring monofloral honey authenticity.

1. Introduction

The term honey can be related to different food products taking into account the raw material and their floral origin, namely honeydew, monofloral or polyfloral honey. Also, honey is usually classified considering its colour. Colours of honey cover a continuous range from very pale yellow through ambers to nearly black, which may be attributed mainly to the plant source of the honey [1,2]. From a consumer's point of view, the colour is a key factor when purchasing, since honey colour is reported to be a factor in its grading and marketing, lighter colour honey being associated with more smooth flavour and in general more expensive [3–5]. On the other hand, monofloral honey is also much more valuable from a commercial point of view, due to their higher market prices, when compared namely to polyfloral honey [6–8]. Therefore, colour and floral classifications of honey samples are of major economic importance for both producers and consumers. Colour-Pfund classification is usually based on UV/vis spectrophotometric analysis, results of which are later related to a colour scale, considered a simple methodology. On the other hand, floral honey classification, is a time-consuming task that requires high-skilled technicians, since it is mainly based on melissopalynological analysis, which in some cases must be complemented by sensory analysis, as for some types of honey the interpretation of melissopalynological analysis can be complicated and sometimes imprecise and ambiguous [6]. Even so, Corbella and Cozzolino [9] reported the combination of multivariate techniques and pollen count analysis to classify honey samples from Uruguay according to botanical sources, although not in the industrial context. Therefore, novel, fast and low-cost practical tools that could be easily applied to the analysis of the majority of honey samples are needed and, in fact, even the EU commission encourages the development of new methods for honey authentication, as pointed out by some researchers [10].

In the last years several studies have dealt the honey authentication problematic by applying multivariate statistical techniques to different physicochemical data (such as colour, diastase activity, water content, ash, free amino acids, lactone and total acidity, pH,

* Corresponding author. Tel.: +351 273303323; fax: +351 273325405.
E-mail address: ldias@ipb.pt (L.G. Dias).

http://dx.doi.org/10.1016/j.talanta.2014.05.004
0039-9140/

electrical conductivity, viscosity, total antioxidant activity, flavonoids and phenolic acids) [7,11–14] or to honey volatile fraction measurements [15–17], among other techniques reported in literature for geographical and botanical classification of honey [8,18,19]. In general, all these approaches showed good discrimination capabilities, precision, accuracy and reliability, but they are in general destructive, time-consuming and expensive, being unsuitable for in situ monitoring [10]. To overcome these drawbacks other more simple and user-friendly methodologies have been proposed, namely the use of potentiometric [6,10,21,22], voltammetric [23,24] or impedance [25] electronic tongues (E-tongues). The results reported in these studies clearly show that all these electrochemical devices can be used as effective and practical tools to discriminate honey according to their botanical origin, allowing distinguishing among different monofloral and/or polyfloral samples and in some cases among different geographical origins.

With reference to the potentiometric devices, the first approach was published by Dias et al. [20], where an E-tongue, containing all-solid-state potentiometric sensors with polymeric membranes, together with multivariate chemometric tools (principal component analysis, PCA; and linear discriminant analysis, LDA), being developed and applied in the classification of 52 commercial Portuguese honey samples according to their botanical origin, namely as *Lavandula*, *Erica* or *Echium* monofloral honey (around 70% of correct classifications). Wei et al. [21] used a commercial potentiometric E-tongue with 7 sensors to classify 192 Chinese honey samples of different floral origin (*Acacia*, *Astragali*, *Buckwheat*, *Coptis*, *Data*, *Motherwort*, *Radix Changll* and *Vitex*) from a Chinese geographical area and 120 samples from five other Chinese geographical origins, harvested in 2008, using three-pattern recognitions techniques (PCA; Cluster Analysis, CA; and Artificial Neural Networks, ANN), with correct classification rates greater than 90%. Zakaria et al. [22] showed that an electronic nose used together with a chalcogenide-based potentiometric E-tongue containing ion-selective sensors, was the best solution for classifying honey of different floral origin. In this study, 80 honey samples from 10 monofloral honey from Malaysia and New Zealand and 4 brands of polyfloral honey from Malaysia were used. The sensor fusion methodology allowed a 100% correct classification for cross-validation procedure using a LDA approach and over than 90% using ANN for an external validation group. Major et al. [10] used a commercial potentiometric E-tongue, comprised of 7 sensors, for botanical classification and physicochemical characterisation of honey samples (acacia, chestnut and honeydew honey) from Croatia. Using ANN models, all training and testing samples were correctly classified according to their botanical origin. Escriche et al. [6] proposed a potentiometric E-tongue, made of metals and metallic compounds that, together with ANN, was able to correctly classify the botanical origin of 32 Spanish honey samples harvested in 2008 and 2009 with a success greater than 90% for three honey of floral origin (*citrus*, *rosemary* monofloral and polyfloral) and one honeydew (forest origin). Furthermore, a good correlation was observed between the E-tongue and colour-Pfund, luminosity and diastase activity. More recently, this same research team [26] used the same potentiometric device to satisfactorily classify honey samples (with a correct recognition greater than 75%) according to their botanical origin (citrus, rosemary and polyfloral) and physical treatment, usually applied to commercial honey (raw, liquation and pasteurisation), using algorithms based on Fuzzy ARTMAP simplified artificial neural networks.

Therefore, considering: (i) the encouraging results achieved so far by different groups regarding honey botanical origin classification using potentiomettric E-tongues coupled with chemometric tools, and (ii) the commercial interest in classifying honey according to botanical origin, in this work a new classification strategy was evaluated envisaging the correct classification of samples that greatly vary in colour and pollen profile, which will be the normal practical case, using an E-tongue.

To the best of author's knowledge, an E-tongue has never been applied before to classify honey samples with the wide variability in colour and floral origin that was evidenced in the analysed samples. For that purpose a new all-solid-state potentiometric E-tongue device was constructed, using a print-screen technique, comprising cross-sensitivity sensors. The analysis carried out included a broader database than that previously used [20], namely, 65 monofloral Portuguese honey samples harvested in 3 consecutive years (from 2009 up to 2011) from all regions of beekeeping production of the mainland of Portugal, resulting in 6 different floral origins, according to pollinic analysis. E-tongue data were analysed using a two-step sequential procedure. First, a LDA model was established based on the potentiometric signal sensor profiles selected using a meta-heuristic variable selection algorithm, to show that it is possible to discriminate honey samples according to their colours (white, amber and dark honey, which were classified according to a colour-Pfund method). Finally, the main purpose, for each colour group selected, was to classify honey samples considering their floral origin by using a LDA based on the E-tongue data after variable selection.

2. Materials and methods

2.1. Honey samples

Sixty five monofloral Portuguese honey samples, harvested between 2009 and 2011, were kindly disposed by the Federação Nacional dos Apicultores de Portugal (FNAP), being a representative sampling of the most productive Portuguese honey regions, both in the continent (e.g., Trás-os-Montes e Alto Douro; Entre Douro e Minho; Beira Interior; Beira Litoral; Estremadura e Ribatejo; Alentejo; and Algarve regions) and Azores (Pico and São Miguel islands). Samples were stored at room temperature until analysis. Each sample was split for colour evaluation, floral origin classification based on melissopalynological assays and potentiometric signal profile analysis with an E-tongue device.

2.2. Honey colour classification

The colour of each honey sample was evaluated using a quantitative millimetre Pfund (mm Pfund) scale [27]. The mm Pfund values were calculated from the absorbance of diluted honey samples (5.0 g of honey in 10.0 mL of deionised water) recorded at 635 nm measured with an UV/vis spectrophotometer (Jenway, Genova model), according to [27]

$$\text{mm Pfund} = -3870 + 37,139 \times \text{Absorbance} \quad (1)$$

The quantitative mm Pfund scale was transformed into a qualitative colour classification according to the scale defined by the United States Department of Agriculture [28], which considers 7 levels of colour for honey: water white (≤ 8 mm Pfund), extra white ($8 <$ mm Pfund ≤ 17), white ($17 <$ mm Pfund ≤ 34), extra light amber ($34 <$ mm Pfund ≤ 50), light amber ($50 <$ mm Pfund ≤ 85), amber ($85 <$ mm Pfund ≤ 114) and dark amber (> 114 mm Pfund). Considering a broader colour classification scale, honey samples were split into only 3 main colour groups: white (mm Pfund ≤ 34), amber ($34 <$ mm Pfund ≤ 114) and dark (> 114 mm Pfund) colours. This colour grouping was chosen considering the dimension of database used and the main colour groups that are naturally identified in honey samples.

2.3. Melissopalynology analysis of honey

The honey pollen quantitative spectrum analysis was performed according to the method reported by Louveaux et al. [29]. For each analysis, 10 g of honey were diluted with 30 mL of distilled water and the sediment was concentrated by centrifugation at 1500 rpm during 30 min. To the recovered sediment an addition of 10.0 mL of anhydride acetic (Panreac) and sulphuric acid (M&B; 9:1, v/v) was made. After incubation in a water bath (100 °C during 3 min) with agitation, a new centrifugation was carried out and the solution was decanted. Then, 12.0 mL of acetic acid (Merck) were added to the sediment and, after agitation, a new centrifugation and decantation were made. The sediment was washed and re-suspended in 12.0 mL of distilled water, and centrifuged and decanted again. The final wash was made with 12.0 mL of KOH 7% (Merck) solution, and a repetition of the agitation, centrifugation and decantation steps was accomplished. Finally, the pollen grains were stained using a fuchsin solution (Merck) mixed with glycerine (Absolve).

Pollen identification and count were carried out using an optic microscope (Leitz Messtechnik GmbH, Wetzlar, Germany) with 400× and 1000× objectives (the last one was used when greater detail was required for pollen identification). For each honey sample, a minimum of 1000 grains of pollen was counted, and in case of doubt the analysis was repeated. Reference standards obtained from Portugal honey flora (available at Escola Superior Agrária – Instituto Politécnico de Bragança, Portugal) were used for grain pollen identification and the samples were classified based on their floral origin according to their found pollen morphology.

2.4. E-tongue analysis

2.4.1. Multisensor system

The multisensor system was printed in both sides of a PVC board using a print-screen technique, by applying an epoxy conductive silver paste (EPO-TK E4110, Epoxy Technology, Inc.) with low temperature curing, prepared by mixture of two reagents (paste and hardener). In each side of the system, 10 chemical sensors can be applied. The system was cured at 40 °C allowing obtaining a dried paste in 8 h. After cutting and cleaning the impressed circuit, the system was waterproofed using an acrylic resin (PLASTIK 70, da Kontakt Chemie). During the waterproofing, the contact spots where polymeric membranes would be applied as well as the connection section where the RS-232 pin 25 male plug will be connected, were protected. Fig. 1 shows an example of a multi-sensor system showing black covers protecting the spots where polymeric membranes will be applied. Before use, each system was tested with a multimeter (Digital Multimeter UniVolt DT-64) to verify the conduction of the electrical signal.

2.4.2. Chemical sensors

The two multi-sensor arrays constructed used different cross-sensitivity membranes as chemical sensors, with different pre-established mass combinations of 4 lipidic additives (octadecylamine, oleyl alcohol, methyltrioctylammonium chloride and oleic acid from Fluka; between 2.8 and 3.2%), 5 plasticizers (bis (1-butylpentyl) adipate, dibutyl sebacate, 2-nitrophenyl-octylether, tris(2-ethylhexyl)phosphate and dioctyl phenylphosphonate, from Fluka; between 64.7% and 65.2%) and PVC high molecular weight polymer (poly(vinyl chloride) polymer, between 31.9% and 32.3%), as shown in Tables 1 and 2, identified with a code with a letter S (for sensor) followed by the number of the array (1 or 2) followed by the number of the membrane (1–20, corresponding to different combinations of plasticizer and additive used). The membrane additives and plasticizers used in the polymeric membranes preparation were selected taking into account the sensor performance, especially the signal stability and repeatability in time of the sensor responses towards basic standard taste substances (salty, sweet, bitter, acid and umami), as previously shown by Dias et al. [30]. These sensor arrays were homemade and their cost only concerns the price of reagents and materials. Furthermore, these devices may be re-used several times, markedly amortising the low initial investment.

Each mixture was prepared by weighting pre-established masses of each one of the above-mentioned 3 products, which were diluted in tetrahydrofurane (from Sigma) in order to obtain a viscous and homogeneous solution. In the multi-sensor system, after removing the protective caps with the aid of a sharp cutting edge, each membrane was prepared by using a drop-by-drop technique (with a 3–5 min interval to ensure complete solvent evaporation) until a transparent crystalline membrane was obtained.

2.4.3. E-tongue assays

The assays with the E-tongue device were carried out at room temperature using aqueous solutions of honey (10.00 g of honey diluted in 50.00 g of deionized water). Before each analysis, the analytical system (two sensor arrays) and the Ag/AgCl reference electrode (Crison, 5241) were carefully washed with deionised water and cleaned with an adsorbent paper and then submersed into the stirred aqueous sample solution and allowed to stabilise during 7 min time-period. Finally the potentiometric signals, varying from −2.0 V to +2.0 V, of 20 different sensor membranes, used in duplicate (S1:1–S1:20 and S2:1–S2:20), were recorded for each sample.

2.5. Statistical analysis

This work is focused on the classification of honey samples using linear discriminant analysis (LDA, a supervised multivariate statistical method for classification) according to the colour and floral origin of honey. The mathematical models established are linear combinations of the independent variables (E-tongue sensor signals) that allow the best separation of the different groups of honey (dependent variable). The proposed methodology involved two steps: first, the selection of the most informative independent variables (sensors) by using a Simulated Annealing (SA)

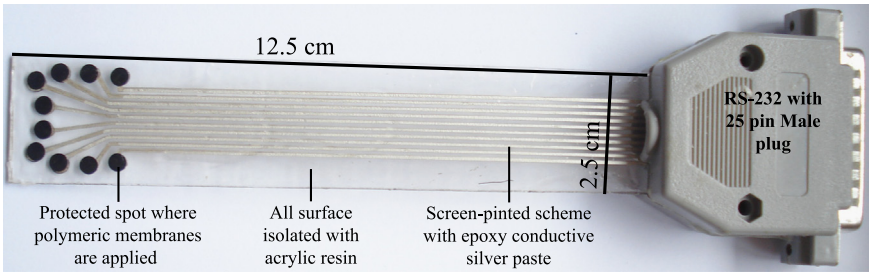

Fig. 1. Multi-sensor system built to use 20 polymeric membranes for potentiometric analysis.

Table 1
Polymeric membrane compositions of the sensors applied into the E-tongue multi-sensor arrays.

Plasticizer	Additive	ID no.[a]	Additive (%)	Plasticizer (%)	PVC [b] (%)
Bis(1-butylpentyl) adipate	Octadecylamine	S1:1 or S2:1	2.99	65.02	31.99
	Oleyl alcohol	S1:2 or S2:2	2.96	65.00	32.04
	Methyltrioctylammonium chloride	S1:3 or S2:3	3.00	65.01	31.99
	Oleic acid	S1:4 or S2:4	3.00	65.04	31.96
Dibutyl sebacate	Octadecylamine	S1:5 or S2:5	3.02	65.11	31.87
	Oleyl alcohol	S1:6 or S2:6	2.96	65.02	32.02
	Methyltrioctylammonium chloride	S1:7 or S2:7	3.00	64.91	32.10
	Oleic acid	S1:8 or S2:8	2.99	64.93	32.09
2-Nitrophenyl-octylether	Octadecylamine	S1:9 or S2:9	2.98	64.68	32.34
	Oleyl alcohol	S1:10 or S2:10	3.00	65.10	31.90
	Methyltrioctylammonium chloride	S1:11 or S2:11	2.99	65.00	32.00
	Oleic acid	S1:12 or S2:12	3.01	65.04	31.95
Tris(2-ethylhexyl)phosphate	Octadecylamine	S1:13 or S2:13	2.96	64.99	32.05
	Oleyl alcohol	S1:14 or S2:14	2.97	65.04	32.00
	Methyltrioctylammonium chloride	S1:15 or S2:15	3.00	64.99	32.01
	Oleic acid	S1:16 or S2:16	2.98	65.04	31.98
Dioctyl phenylphosphonate	Octadecylamine	S1:17 or S2:17	3.12	64.85	32.03
	Oleyl alcohol	S1:18 or S2:18	2.95	65.04	32.01
	Methyltrioctylammonium chloride	S1:19 or S2:19	2.99	64.69	32.32
	Oleic acid	S1:20 or S2:20	3.01	65.11	31.88

[a] Identification number.
[b] Polyvinyl chloride.

Table 2
Floral and colour commercial honey samples classification according to reference melissopalynological and Pfund-colour group analysis.

Sample	Harvest year	Pollen quantitative spectrum[a] (abundance, %)			mm Pfund	Classification	
		1st pollen	2nd pollen	3rd pollen		Floral origin	Colour group
1	2011	Ech 53%	Pru 22%	Foe 16%	32	Ech	White
2	2011	Ech 71%	Pru 15%	Lav 10%	16	Ech	Extra white
3	2011	Ech 73%	Tri 10%	Lav 5%	27	Ech	White
4	2011	Ech 79%	Lav 12%	Rub 8%	26	Ech	White
5	2011	Ech 60%	Lav 17%	Tri 14%	32	Lav	White
6	2010	Ech 34%	Rub 32%	Lav 19%	26	Lav	White
7	2010	Ech 38%	Rub 25%	Lav 19%	33	Lav	White
8	2009	Rub 39%	Lav 19%	Tri 11%	20	Lav	White
9	2010	Ech 39%	Lav 21%	Rub 20%	21	Lav	White
10	2010	Ech 39%	Rub 26%	Lav 21%	33	Lav	White
11	2011	Ech 36%	Lav 31%	Tri 21%	33	Lav	White
12	2010	Lav 31%	Ech 29%	Pru 22%	28	Lav	White
13	2009	Lav 32%	Cas 19%	Rub 13%	32	Lav	White
14	2011	Lav 39%	Ech 39%	Euc 7%	27	Lav	White
15	2009	Lav 45%	Ech 25%	Rub 20%	27	Lav	White
16	2011	Lav 45%	Ech 31%	Tri 7%	16	Lav	Extra white
17	2009	Lav 47%	Cas 17%	Ech 8%	28	Lav	White
18	2009	Lav 49%	Rub 20%	Ech 10%	27	Lav	White
19	2011	Lav 56%	Leo 20%	Rub 15%	22	Lav	White
20	2011	Lav 67%	Eri 12%	Rub 12%	31	Lav	White
21	2010	Ech 53%	Rub 12%	Cas 10%	39	Ech	Amber
22	2009	Ech 53%	Cas 16%	Lav 8%	40	Ech	Amber
23	2010	Ech 58%	Cas 7%	Rub 7%	45	Ech	Amber
24	2011	Ech 70%	Pru 13%	Cas 10%	47	Ech	Amber
25	2009	Ech 70%	Lav 10%	Pru 8%	49	Ech	Amber
26	2010	Ech 49%	Rub 16%	Cas 13%	66	Ech	Amber
27	2010	Ech 53%	Rub 28%	Euc 5%	67	Ech	Amber
28	2011	Ech 54%	Pru 24%	Lav 10%	78	Ech	Amber
29	2011	Ech 63%	Pru 26%	Tri 5%	78	Ech	Amber
30	2011	Ech 54%	Pru 16%	Cas 10%	100	Ech	Amber
31	2011	Ech 58%	Cas 26%	Tri 13%	103	Ech	Amber
32	2011	Lav 38%	Ech 30%	Rub 18%	35	Lav	Amber
33	2009	Lav 33%	Rub 20%	Thy 18%	36	Lav	Amber
34	2011	Lav 46%	Pru 40%	Tri 8%	42	Lav	Amber
35	2009	Lav 31%	Pru 29%	Ech 20%	51	Lav	Amber
36	2011	Pru 32%	Lav 27%	Ech 26%	104	Lav	Amber
37	2011	Pru 64%	Ech 15%	Aca 13%	61	Pru	Amber
38	2011	Pru 50%	Cas 38%	Rub 3%	63	Pru	Amber
39	2011	Pru 80%	Cas 6%	Euc 6%	69	Pru	Amber
40	2009	Rub 58%	Aca 22%	Pru 7%	40	Rub	Amber
41	2011	Rub 47%	Euc 21%	Pru 19%	59	Rub	Amber

Table 2 (*continued*)

Sample	Harvest year	Pollen quantitative spectrum[a] (abundance, %)			mm Pfund	Classification	
		1st pollen	2nd pollen	3rd pollen		Floral origin	Colour group
42	2010	Rub 62%	Ech 14%	Aca 13%	60	Rub	Amber
43	2010	Rub 49%	Cas 17%	Gen 11%	66	Rub	Amber
44	2009	Rub 52%	Eri 12%	Pru 11%	74	Rub	Amber
45	2009	Rub 50%	Cas 27%	Tri 8%	76	Rub	Amber
46	2009	Rub 69%	Pru 7%	Eri 3%	81	Rub	Amber
47	2009	Rub 49%	Tri 17%	Cas 14%	101	Rub	Amber
48	2010	Rub 57%	Tri 22%	Eri 10%	111	Rub	Amber
49	2010	Rub 50%	Lav 14%	Cas 12%	113	Rub	Amber
50	2009	Rub 51%	Cas 11%	Lav 6%	114	Rub	Amber
51	2009	Cas 92%	Eri 1%	Pru 1%	119	Cas	Dark
52	2011	Cas 95%	Rub 5%	–	134	Cas	Dark
53	2011	Cas 91%	Rub 5%	Tri 4%	147	Cas	Dark
54	2011	Cas 90%	Euc 6%	Ech 3%	176	Cas	Dark
55	2009	Eri 50%	Cas 28%	Pru 8%	127	Eri	Dark
56	2009	Eri 53%	Cas 15%	Rub 13%	129	Eri	Dark
57	2009	Eri 61%	Tri 17%	Pru 12%	141	Eri	Dark
58	2011	Eri 56%	Tri 23%	Pru 8%	171	Eri	Dark
59	2010	Eri 54%	Cas 19%	Rub 8%	172	Eri	Dark
60	2009	Eri 82%	Cas 12%	Ech 6%	193	Eri	Dark
61	2011	Eri 54%	Cas 26%	Ech 7%	199	Eri	Dark
62	2011	Eri 63%	Rub 19%	Cas 15%	201	Eri	Dark
63	2010	Rub 50%	Ech 17%	Cas 16%	142	Rub	Dark
64	2011	Rub 61%	Ech 14%	Eri 13%	147	Rub	Dark
65	2009	Rub 49%	Cas 22%	Lav 10%	196	Rub	Dark

[a] Aca – *Acacia* sp.; Cas – *Castanea* sp.; Ech – *Echium* sp.; Eri – *Erica* sp.; Euc – *Eucalyptus* sp.; Foe – *Foeniculum* sp.; Gen – *Genista sp.*; Lav – *Lavandula* sp.; Pru – *Prunus* sp.; Rub – *Rubus* sp.; Thy – *Thymus* sp.; and Tri – *Trifolium* sp.

meta-heuristic variable selection algorithm, which is a key task considering that signal profiles recorded by E-tongue devices usually show a high multicollinearity degree; then, the performance evaluation of the selected models regarding sample classification was carried out using a leave-one-out (LOO) cross-validation technique to avoid overoptimistic correct classification results. This methodology enabled to evaluate the model's prediction performance by removing a set of n samples from the database and then predict their rank with the LDA model obtained with the remaining $n-1$ samples. This process was repeated n times, to obtain the classification errors of all the samples, and the overall sum of errors for each test divided by n.

The subsets of independent variables selected by applying the SA algorithm allowed obtaining LDA models with fewer sensors, eliminating redundant sensor that had a similar contribution to the differences between groups, increasing the accuracy of the prediction. The final model selected will be simpler and easier to interpret, allowing the best prediction performance with the minimum number of sensors. The algorithm was programmed to give the best model for each subset, using 2–20 sensors selected among the 40 recorded potentiometric signals (which resulted in 19 models) and after 10,000 attempts. The subset range was set between 2 and 20 with the purpose of achieving the simplest model (with the minimum number of sensors) and also considering the number of honey samples belonging to each colour group. The number of attempts was set equal to 10,000 as this value enabled to reach always the same best solution (type of sensors included in the model) for subsets with higher number of variables. The quality criterion ccr12 (Roy's first root statistical coefficient) was used to assess the goodness of fitting between the dependent variable (colour or types of monofloral honey) and each one of the 19 subsets of sensors chosen. Maximising this criterion is equivalent to the maximisation of the first root Roy, which is the ratio between the unexplained and the explained variance for the first discriminant function (conceptually equivalent to the value F ratio in the analysis of variance) [31,32]. All statistical analysis was performed using Subselect [31,32] and MASS [33] packages of the open source statistical program R (version 2.15.1).

3. Results and discussion

3.1. Honey colour and floral origin classifications

When the 65 honey samples were analysed by UV–vis spectrophotometry and the absorbances of aqueous diluted samples transformed into mm Pfund according to Eq. (1), values from 15.9 up to 204.2 mm Pfund were determined, enabling honey to be classified according to the honey colour scale [27,28]. The results showed that samples had colours between extra white and dark amber: 2 extra white, 18 white, 9 extra light amber, 14 light amber, 7 amber and 15 dark amber honey samples.

Using the colour classification methodology proposed in this work (Section 2.1), samples were split into 3 main groups as follows: 20 white samples with mm Pfund lower than 34 (between 15.9 and 33.3 mm Pfund); 30 amber samples with mm Pfund ranging from 34.8 to 113.6 mm Pfund; and 15 dark samples with mm Pfund greater than 118.8 mm Pfund. This procedure is needed as a first data treatment step for honey analysis with the potentiometric E-tongue due to the wide colour variability observed in honey samples, which is mainly influenced by the honey floral origin.

Concerning the floral origin of each honey sample, it was evaluated using pollen grain identification and count. Globally, for the 65 honey samples analysed, it was possible to identify 23 different kinds of pollen. By descending order considering their presence in the overall samples, these were: *Rubus* sp., *Lavandula* sp., *Prunus* sp., *Echium* sp., *Castanea* sp., *Trifolium* sp. and *Erica* sp. (present in more than 44% of the samples), followed by *Eucalyptus* sp., *Leontodeon* sp., *Thymus* sp., *Cytisus* sp., *Acacia* sp. and *Pinus* sp. (detected in 12–34% of the samples) and finally the other type of pollens, *Sandix* sp., *Foeniculum* sp., *Helianthus* sp., *Genista* sp., *Tilia sp.*, *Mentha* sp., *Persea* sp., *Medicago sp.*, *Crepis* sp. and *Mimosaceae*

sp. (identified in less than 10% of the samples). These different identified pollens were already reported for other Portuguese honey [34]. The pollinic profiles determined showed that in 61 honey samples it was possible to identify between 4 and 9 different pollens, in 3 honey samples 3 different pollens were detected and, finally, in only 1 sample 2 different pollens were found. Independently of their colours and based on their pollinic profiles and on the relative abundance of each pollen [34,35], the 65 honey samples were classified as *Castanea*, *Echium*, *Erica*, *Lavandula*, *Prunus* and *Rubus* monofloral honey and, further split according to the 3 main colour groups previously defined: white, amber and dark. Table 2 presents information concerning harvest year, colour classification and pollinic profile (considering only the 3 pollens most predominante) of the honey samples analysed, which enables a detailed overview of the wide variability found for these two important honey characteristics.

These results, as expected, also confirm the high floral origin variability of the honey samples analysed in this work. Only the *Castanea* honey showed lower variability in the pollinic profile, since the main pollen (*Castanea* sp.) represented 90–95% of prevalence. For the other monofloral honey, the main pollen varied over a wider range of percentages: *Echium* between 49% and 79%; *Erica* between 50% and 82%; *Lavandula* between 17% and 67%; *Prunus* between 49% and 80%; and *Rubus* between 47% and 69%.

3.2. E-tongue results

3.2.1. E-tongue signals profiles of honey samples

In the total, 65 assays were carried out, each providing 40 potentiometric signals (20 different sensor membranes used in duplicate: S1:1–S1:20 and S2:1–S2:20). Fig. 2 shows the potentiometric signal's box-plots for each sensor and monofloral honey samples grouped according to colour (white, amber and dark). The potentiometric signals varied from +0.09 V to +0.23 V for all sensors included in the E-tongue, avoiding the need of data scaling. Fig. 2 shows that slight differences in signal intensities occur for some sets of sensors, implying the need of a variable

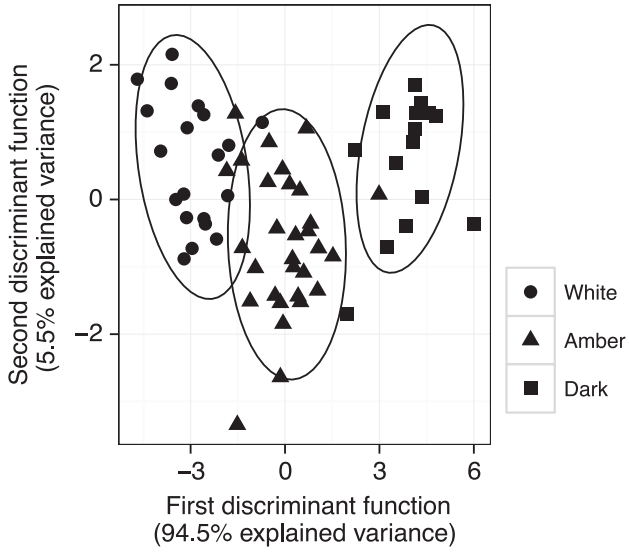

Fig. 3. Linear discriminant analysis between the three colour groups associated with the primary selection of samples. The ellipses in each group are associated with the confidence interval of 90%, assuming a multivariate normal distribution.

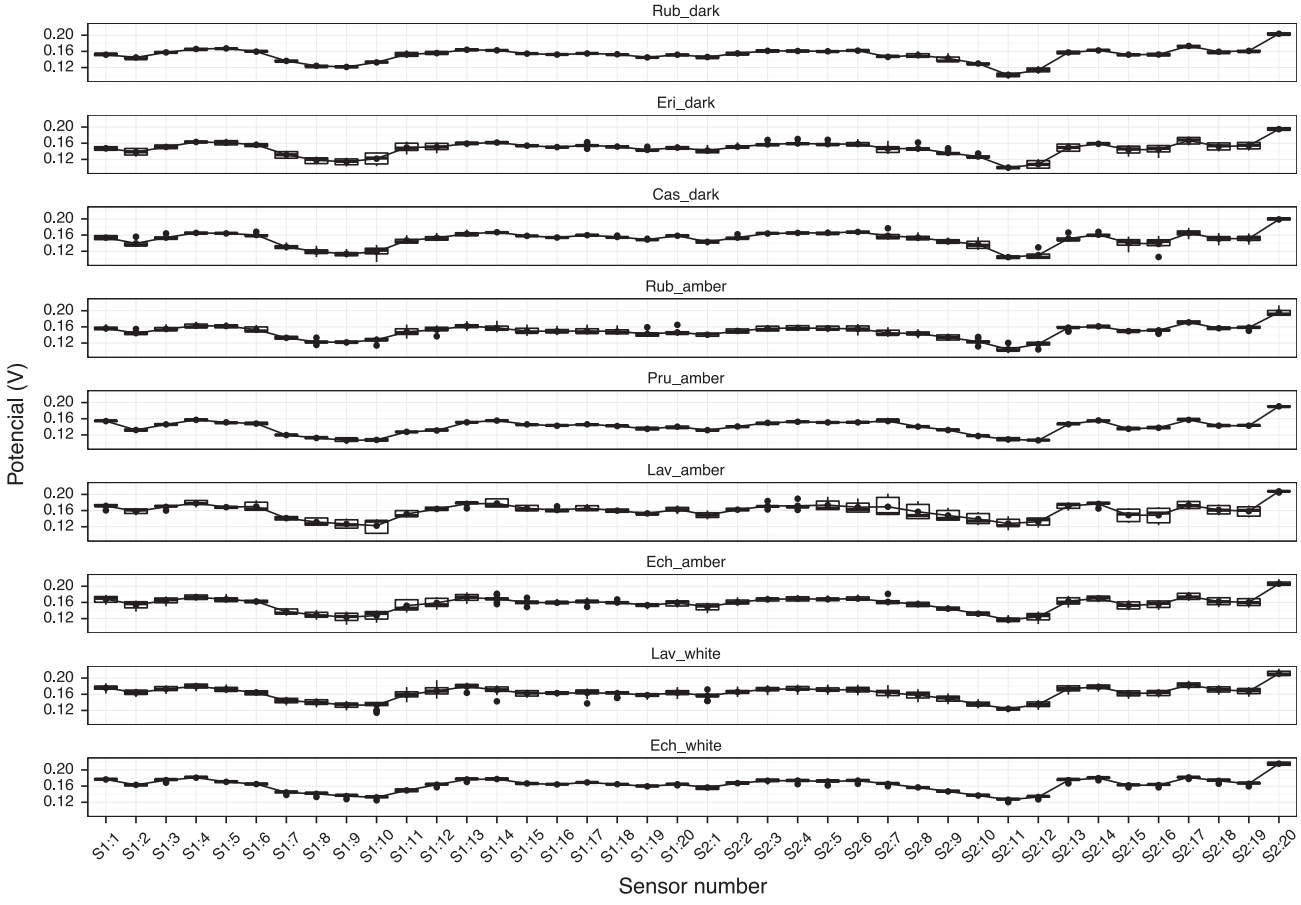

Fig. 2. Potentiometric signal's box-plots for each sensor and monofloral honey grouped according to colour (white, amber and dark).

selection procedure to choose the most relevant signals for establishing the best predictive LDA model.

3.2.2. E-tongue analysis

The classification performance of the E-tongue was evaluated considering colour and floral origin factors. A two-step procedure was adopted, using the potentiometric signal data and LDA, with the sensors subsets selected by the SA algorithm. First honey samples were classified according to colour, independently of their floral origin; and after, for each of the previous groups, samples were discriminated considering their floral origin. This procedure was adopted because it was not possible (data not shown) to establish a predictive LDA model that would allow a leave-one-out cross-validation satisfactory classification ($\leq$ 45% of correct classifications) when using simultaneously the 65 honey samples of the 6 monofloral honeys.

The results showed that, the potentiometric E-tongue developed allowed a satisfactory discrimination of honey samples into the 3 main colour groups established (white, amber and dark honey) using LDA based on a subset of 13 sensors selected by the SA algorithm (S1:3; S1:5; S1:7; S1:14; S1:16; S1:20; S2:2; S2:11; S2:14; S2:16; S2:17; S2:18 and S2:19) (Fig. 3), which may confirm that there is similarity among the potentiometric signal profiles of the samples of each group. Although colour is a visual characteristic, the capability of the E-tongue to distinguish honey samples according to their colours may be due to the different responses of the sensors device towards the different physicochemical and matrix composition of the different monofloral honey samples studied [36–40].

For that, two discriminant functions were established explaining 100% of the total original data variance (94.5% and 5.5%). The first function enables the discrimination of the three colour groups in an expected lighter-to-darkness tone sequence, being white honey placed in the negative region, amber honey in the middle zone (corresponding to scores near to zero) and dark honey in the positive region. The sensors selected covered all the plasticizers and additives used in the membrane composition. Both original data and LOO cross-validation (Table 3) classifications allowed achieving 91% of the honey samples correctly classified (only 6 samples of the 65 honey samples were misclassified), being white and dark honey classifications the most accurate. These results are quite satisfactory, especially if it is taken into account that 6 different floral origins were identified in the honey samples studied and also that they were collected during a 3-year time period.

The E-tongue gave very satisfactory results regarding the discrimination of monofloral honey according to their floral origin, within each honey colour group previously defined (Figs. 4–6), without showing any correlation with their harvest years.

Indeed, for white, amber and dark colour groups the E-tongue enabled 100% correct classifications with a LDA leave-one-out cross-validation procedure. For the white honey group, containing *Lavandula* and *Echium* monofloral honey (16.3–33.3 mm Pfund and 15.9 up to 32 mm Pfund, respectively), one linear discriminant function (explaining 100% of the data variance) was established

Table 3
LDA contingency matrix for the monofloral honey samples colour classification based on the E-tongue signals recorded (results from leave-one-out cross-validation procedure).

Actual honey colour group	Predicted honey colour group			Sensitivity (%)
	White	Amber	Dark	
White	19	1	0	95
Amber	3	26	1	87
Dark	0	1	14	93

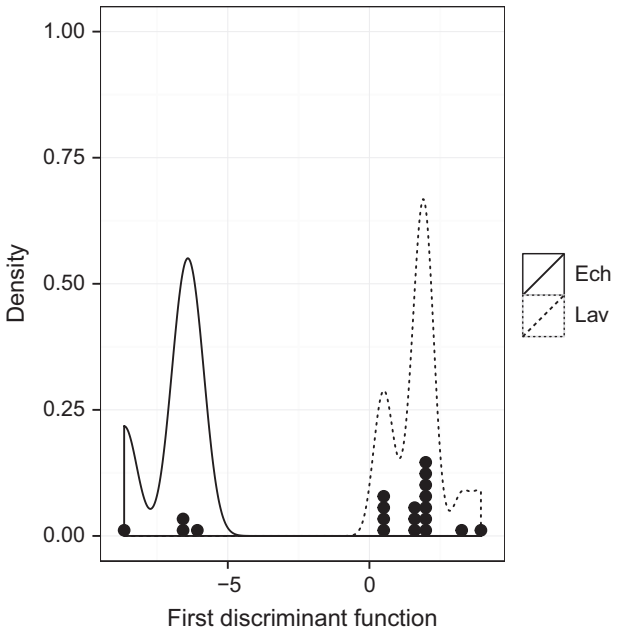

Fig. 4. Frequency distribution of the data referring to the first discriminant function classification of two groups of white-coloured monofloral honey: *Echium* sp. (Ech) and *Lavandula* sp. (Lav).

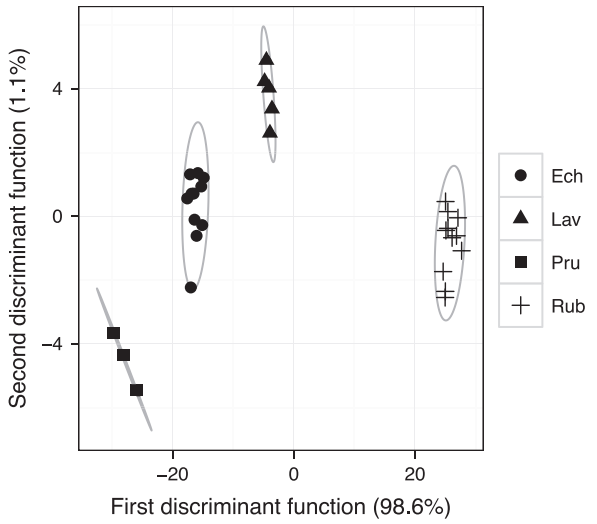

Fig. 5. Linear discriminant analysis of four groups of amber-coloured monofloral honey: *Echium* sp. (Ech), *Lavandula* sp. (Lav), *Prunus* sp. (Pru) and *Rubus* sp. (Rub). The ellipses in each group are associated with the confidence interval of 95%, assuming a multivariate normal distribution.

based on the potentiometric signal data recorded by a set of 6 sensors (S1:1, S1:3, S1:10, S2:2, S2:3, S2:16), which were selected using the SA algorithm. Regarding the amber honey group, which included *Echium*, *Lavandula*, *Prunus* and *Rubus* monofloral honey (38.9–103.2 mm Pfund; 34.8–103.5 mm Pfund; 60.8–69.0 mm Pfund and 40.0 up to 113.6 mm Pfund, respectively), three linear discrminant functions, with 16 sensors (S1:6, S1:7, S1:8, S1:10, S1:12, S1:14, S1:17, S2:1, S2:4, S2:8, S2:12, S2:14, S2:15, S2:16, S2:18, S2:19), were needed (explaining 98.6%, 1,1% and 0.3% of the original data total data variance, respectively). Finally, for the dark honey group, containing *Castanea*, *Erica* and *Rubus* monofloral honey (118.8–175.6 mm Pfund; 126.6–201.2 mm Pfund; and 142.2 up to 196.0 mm Pfund, respectively), two linear discriminant functions, based on the signals of 7 sensors (S1:12, S1:15, S1:18, S2:7, S2:9, S2:13, S2:19), were established

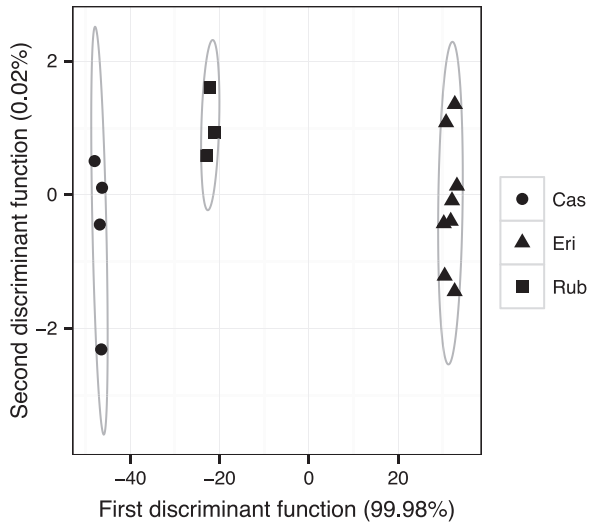

Fig. 6. Linear discriminant analysis of three groups of dark-coloured monofloral honey: *Castanea* sp. (Cas), *Erica* sp. (Eri) and *Rubus* sp. (Rub). The ellipses in each group are associated with the confidence interval of 95%, assuming a multivariate normal distribution.

(accounting for 99.98% and 0.02% of the original data total variance, respectively).

The analysis of the type of sensors selected for each colour group, considering the plasticizer plus additive combination, shows that, in general, all plasticizers and additives were used with the same frequency, which means that all the combinations used were adequate for extracting the most relevant information from the honey samples. Furthermore, for white and amber groups, the variable selection algorithm enabled the inclusion of repeated sensors, with the same membrane composition, in the discrimination models, each from one of the two sensor arrays used, which is in accordance with the findings previous reported [41], showing that the inclusion of repeated sensors in multivariate analysis can improve model performance. Indeed, from a modelling point of view, duplicate potentiometric sensors can be seen as independent variables considering that slight variations of the membrane composition and physical properties (transparency and porosity) may occur when a drop-by-drop technique is used for membrane preparation, resulting in slightly different signal profiles recorded by two sensor replicas. Also, some the sensors selected for two colour groups, by applying the SA algorithm, were equal, namely for white and amber groups (membrane nos. 1, 10 and 16) and amber and dark groups (membrane nos. 7, 12, 15, 18 and 19). The selection of these same sensors may be due to the presence of honey samples with the same floral origin belonging to different honey colour groups. The use of the meta-heuristic SA algorithm for variable selection enabled the identification and selection of a minimum set of sensors required to fully discriminate monofloral honey samples according to their floral origin, after colour honey classification. Moreover, compared with previous reported applications of E-tongue for floral origin classification of honey, the performance of the proposed potentiometric E-tongue is similar [10,22] or quite superior [6,20,21,26].

From a global point of view, it is interesting to note that the inclusion of more sensors (variables) in the linear discriminant functions previously described would result in a lower classification performance of the E-tongue, when a LOO cross-validation procedure was applied (data not shown). This fact corroborates the importance of using an adequate variable selection algorithm that could deal simultaneously with co-linearity issues between potentiometric signal profiles and the selection of the data that provide a truly chemical fingerprint of the samples. It is known that the choice of sensors is a key principle in designing E-tongues [42]. This potential may be enhanced if a robust variable selection algorithm, such as the SA meta-heuristic algorithm, is used together with traditional LDA.

Finally, it should be noted that this novel approach (sample split according to honey colour coupled with cross-sensitivity potentiometric E-tongue) showed a similar [10,22] or better [6,20,21,26] performance concerning floral origin honey sample classification when compared with previous works. Moreover, the proposed methodology is successfully applied to honey samples that, although being classified as monofloral, had a confirmed broader pollinic composition and have been collected during a 3-year period, which compared with the previous works represents a sampling procedure with greater intrinsic variability.

On the whole, these results show the usefulness of the device proposed in this work for potential use at analytical laboratory level for floral origin honey classification.

4. Conclusion

The combined strategy adopted in this work, coupling a prior split step of honey sample by colour determined by spectrophotometry analysis together with a potentiometric E-tongue enabled monofloral honey discrimination according to floral origin (100% of correct classifications for LOO cross-validation), for samples with high colour and pollen profile variabilities. The quality of the results achieved with the E-tongue designed and built in this work, with cross-sensitivity lipid membranes, shows that the selection and incorporation of cross-sensitivity lipidic membranes into the sensor-array provided a useful and informative chemical fingerprint from the monofloral honey samples, enabling their floral origin discrimination.

References

[1] J.W. White Jr., L.W. Doner, Honey Composition and Properties, Beekeeping in the United States, Agric Handbook, 198082–91.
[2] M.J. Diez, C. Andres, A. Terrab, Int. J. Food Sci. Technol. 39 (2004) 167–176.
[3] M.L. Al, D. Daniel, A. Moise, O. Bobis, L. Laslo, S. Bogdanov, Food Chem. 112 (2009) 863–867.
[4] J.W. White, Assoc. Off. Anal. Chem. J. 67 (1984) 1129–1131.
[5] M. Viuda-Martos, Y. Ruiz-Navajas, J.M. Zaldivar-Cruz, V. Kuri, J. Fernández-López, Á.A. Carbonell-Barrachina, J.Á. Pérez-Álvarez, Int. J. Food Sci. Technol. 45 (2010) 1111–1118.
[6] I. Escriche, M. Kadar, E. Domenech, L. Gil-Sánchez, J. Food Eng. 109 (2012) 449–456.
[7] I. Escriche, M. Kadar, M. Juan-Borrás, E. Domenech, Food Chem. 142 (2014) 135–143.
[8] G. Scandurra, G. Tripodi, A. Verzera, J. Food Eng. 119 (2013) 738–743.
[9] E. Corbella, D. Cozzolino, Chil. J. Agric. Res. 68 (2008) 102–107.
[10] N. Major, K. Marković, M. Krpan, G. Šarić, M. Hruškar, N. Vahčić, Talanta 85 (2011) 569–574.
[11] C. Acquarone, P. Buera, B. Elizalde, Food Chem. 101 (2007) 695–703.
[12] G. Beretta, P. Granata, M. Ferrero, M. Orioli, R.M. Facino, Anal. Chim. Acta 533 (2005) 185–191.
[13] N. Gheldof, X.H. Wang, N.J. Engeseth, J. Agric. Food Chem. 50 (2002) 5870–5877.
[14] Z. Wei, J. Wang, W. Liao, J. Food Eng. 96 (2010) 469–479.
[15] L. Castro-Vázquez, M.C. Díaz-Maroto, M.A. González-Viñas, M.S. Pérez-Coello, Food Chem. 112 (2009) 1022–1030.
[16] L.F. Cuevas-Glory, J.A. Pino, L.S. Santiago, E. Sauri-Duch, Food Chem. 103 (2007) 1032–1043.
[17] I. Escriche, M. Kadar, M. Juan-Borrás, E. Domenech, Food Res. Int. 44 (2011) 1504–1513.
[18] J.M. Camiña, R.G. Pellerano, E.J. Marchevsky, Curr. Anal. Chem. 8 (2012) 408–425.
[19] C.H. Latorre, R.M.P. Crecente, S.G. Martín, J.B. García, Food Chem. 141 (2013) 3559–3565.
[20] L.A. Dias, A.M. Peres, M. Vilas-Boas, M.A. Rocha, L. Estevinho, A.A.S.C. Machado, Microchim. Acta 163 (2008) 97–102.
[21] Z. Wei, J. Wang, W. Liao, J. Food Eng. 94 (2009) 260–266.
[22] A. Zakaria, A.Y. Md Shakaff, M.J. Masnan, M.N. Ahmad, A.H. Adom, M.N. Jaafar, S.A. Ghani, A.H. Abdullah, A.H.A. Aziz, L.M. Kamarudin, N. Subari, N.A. Fikri, Sensors 11 (2011) 7799–7822.

[23] K. Tiwari, B. Tudu, R. Bandyopadhyay, A. Chatterjee, J. Food Eng. 117 (2013) 205–210.
[24] Z. Wei, J. Wang, Electrochim. Acta 56 (2011) 4907–4915.
[25] P.A. Ulloa, R. Guerra, A.M. Cavaco, A.M.R. Costa, A.C. Figueira, A.F. Brigas, Comput. Electron. Agric. 94 (2013) 1–11.
[26] E. Garcia-Breijo, J. Garrigues, L.G. Sanchez, N. Laguarda-Miro, Sensors 13 (2013) 10418–10429.
[27] I.C.F.R. Ferreira, E. Aires, J.C.M. Barreira, L.M. Estevinho, Food Chem. 114 (2009) 1438–1443.
[28] USDA Agricultural Marketing Service, United States Standards for Grades of Extracted Honey, USDA, Washington DC, May 23, 1985.
[29] J. Louveaux, A. Maurizio, G. Vorwohl, Bee World 59 (1978) 139–157.
[30] L.A. Dias, A.M. Peres, A.C.A. Veloso, F.S. Reis, M. Vilas-Boas, A.A.S.C. Machado, Sens. Actuat. B 136 (2009) 209–217.
[31] J. Cadima, J.O. Cerdeira, M. Minhoto, Comput. Stat. Data Anal. 47 (2004) 225–236.
[32] J. Cadima, J.O. Cerdeira, P.D. Silva, M. Minhoto, The subselect R package, 2012. ⟨http://cran.r-project.org/web/packages/subselect/vignettes/subselect.pdf⟩.
[33] W.N. Venables, B.D. Ripley, Modern Applied Statistics with S, fourth ed., Springer, New York, ISBN 0-387-95457-0, 2002.
[34] S.M.A. Pires, T. Rodrigues, A. Rocha, A. Pajuelo, O. Pereira, R. Port. Zoot. 1 (2005) 87–99.
[35] W. Von Der Ohe, L.P. Oddo, M.L. Piana, M. Morlot, P. Martin, Apidologie 35 (2004) S18–S25.
[36] V. Baltrusaityte, P. Petras, V. Ceksteryte, Food Chem. 101 (2007) 502–514.
[37] E.F. Boffo, L.A. Tavares, A.C.T. Tobias, M.M.C. Ferreira, A.G. Ferreira, LWT – Food Sci. Technol. 49 (2012) 55–63.
[38] I. Jasicka-Misiak, A. Poliwoda, M. DereÅ, P. Kafarski, Food Chem. 131 (2012) 1149–1156.
[39] J.J.J. Lacerda, J.S. Santos, S.A. Santos, G.B. Rodrigues, M.L.P. Santos, Quím. Nova 33 (2010) 1022–1026.
[40] Md. I. Khalil, M. Moniruzzaman, L. Boukraâ, M. Benhanifia, Md. A. Islam, Md. N. Islam, S.A. Sulaiman, S.H. Gan, Molecules 17 (2012) 11199–11215.
[41] D.P.A. Correia, J.M.C.S. Magalhães, A.A.S.C. Machado, Talanta 67 (2005) 773–782.
[42] J.M. Gutiérrez, Z. Haddi, A. Amari, B. Bouchikhi, A. Mimendia, X. Cetó, M. del Valle, Sens. Actuat. B. 177 (2013) 989–996.