Food Chemistry 118 (2010) 78–83

Contents lists available at ScienceDirect

Food Chemistry

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

Authentication of extra virgin olive oils by Fourier-transform infrared spectroscopy

M.J. Lerma-García *, G. Ramis-Ramos, J.M. Herrero-Martínez, E.F. Simó-Alfonso

Department of Analytical Chemistry, Faculty of Chemistry, University of Valencia, 46100 Burjassot, Valencia, Spain

ARTICLE INFO

Article history:
Received 30 April 2008
Received in revised form 11 March 2009
Accepted 21 April 2009

Keywords:
Botanical origin
Extra virgin olive oil adulteration
Fourier-transform infrared spectroscopy
Linear discriminant analysis
Multiple linear regression

ABSTRACT

Fourier-transform infrared spectroscopy (FTIR), followed by multivariate treatment of the spectral data, was used to classify vegetable oils according to their botanical origin, and also to establish the composition of binary mixtures of extra virgin olive oil (EVOO) with other low cost edible oils. Oil samples corresponding to five different botanical origins (EVOO, sunflower, corn, soybean and hazelnut) were used. The wavelength scale of the FTIR spectra of the oils was divided in 26 regions. The normalized absorbance peak areas within these regions were used as predictors. Classification of the oil samples according to their botanical origin was achieved by linear discriminant analysis (LDA). An excellent resolution among all categories was achieved using an LDA model constructed with eight predictors. In addition, multiple linear regression models were used to predict the composition of binary mixtures of EVOO with sunflower, corn, soybean and hazelnut oils. For all the binary mixtures, models capable of detecting a low cost oil content in EVOO as low as 5% were obtained.

1. Introduction

Oil genuineness is a very important aspect of quality edible oils. Extra virgin olive oil (EVOO) has unique nutritional and sensory characteristics, being also a basic component of the Mediterranean diet. The importance of EVOO is mainly attributed to its high content of oleic acid and its richness in phenolic compounds, which act as natural antioxidants (Bendini et al., 2007). On the other hand, EVOO is expensive owing to the hard and time-consuming tasks involved in the cultivation of olive trees, the harvesting of the fruits, and the extraction of the oil. For these reasons, adulterations of EVOO with olive oils of lower quality, or with oils of a different botanical origin are occasionally detected (Catharino et al., 2005; Chiavaro, Vittadini, Rodriguez-Estrada, Cerretani, & Bendini, 2008; Marcos-Lorenzo, Pérez-Pavón, Fernández-Laespada, García-Pinto, & Moreno-Cordero, 2002; Mariani, Bellan, Lestini, & Aparicio, 2006; Poulli, Mousdis, & Georgiou, 2006; Tay, Singh, Krishnan, & Gore, 2002; Vlachos et al., 2006). For this reason, European Mediterranean countries, which are major suppliers of olive oil in the world market, have adopted common regulations to protect growers and consumers from fraud (European Union Commission, 1991).

To establish the authenticity of edible oils, a number of chromatographic (Brodnjak-Voncina, Kodba, & Novic, 2005; Marcos-Lorenzo et al., 2002; Mariani et al., 2006), thermal (Chiavaro et al., 2008) and spectroscopic methods, including fluorescence (Poulli et al., 2006; Sikorska, Górecki, Khmelinskii, Sikorski, & Kozioł, 2005), NIR (Christy, Kasemsumran, Du, & Ozaki, 2004; Downey, McIntyre, & Davies, 2002; Kasemsumran & Kang, 2005; Sato, 1994; Wesley, Pacheco, & McGill, 1996; Yang, Irudayaraj, & Paradkar, 2005), FTIR (Baeten et al., 2005; Dupuy, Duponchel, Huvenne, Sombret, & Legrand, 1995; Lai, Kemsley, & Wilson, 1999; Ozen & Mauer, 2002; Tay et al., 2002; Vlachos et al., 2006; Yang et al., 2005), FT-Raman (Baeten & Meurens, 1996; López-Díez, Bianchi, & Goodacre, 2003; Yang et al., 2005), NMR (Dais & Spyros, 2007; García-González, Mannina, D'Imperio, Segre, & Aparicio, 2004; Vigli, Philippidis, Spyros, & Dais, 2003) and MS (Catharino et al., 2005; Lay, Liyanage, Durham, & Brooks, 2006; Lerma-García, Ramis-Ramos, Herrero-Martínez, & Simó-Alfonso, 2007, 2008; Lerma-García, Simó-Alfonso, Ramis-Ramos, & Herrero-Martínez, 2007; Marcos-Lorenzo et al., 2002), followed by multivariate statistical analysis of the data, have been described. For this purpose, the contents of fatty acids (Brodnjak-Voncina et al., 2005), tocopherols (Lerma-García, Simó-Alfonso, et al., 2007; Sikorska et al., 2005), volatile compounds (Marcos-Lorenzo et al., 2002), amino acids (Lerma-García, Ramis-Ramos, et al., 2007) and sterols (Lerma-García et al., 2008; Mariani et al., 2006), have been used.

FTIR is a rapid and non-destructive powerful analytical tool for the study of edible oils and fats, requiring minimum sample preparation. FTIR is also an excellent tool for quantitative analysis, since the intensities of the spectral bands are proportional to concentration. For this reason, FTIR has been used to distinguish oils from different botanical origins using non-supervised classificatory techniques (Dupuy et al., 1995; Lai et al., 1999; Rusak, Brown, & Martin, 2003). FTIR has been also used to distinguish EVOOs from different geographical origins (Bendini, Cerretani, et al., 2007; Galtier et al., 2007; Tapp, Defernez, & Kemsley, 2003) and different genetic varieties (Gurdeniz, Tokatli, & Ozen, 2007). FTIR

* Corresponding author. Tel.: +34 96 354 43 34; fax: +34 96 354 44 36.
E-mail address: m.jesus.lerma@uv.es (M.J. Lerma-García).

0308-8146/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodchem.2009.04.092

applications addressed to detect olive oil adulteration with low cost edible oils (Baeten et al., 2005; Ozen and Mauer, 2002; Tay et al., 2002; Vlachos et al., 2006), to evaluate olive oil freshness (Sinelli, Cosio, Gigliotti, & Casiraghi, 2007), to study changes produced by frying (Valdes & Garcia, 2006) and to assess oil oxidation (Guillén & Cabo, 2002; Muik, Lendl, Molina-Diaz, Valcarcel, & Ayora-Canada, 2007; Vlachos et al., 2006) have been also described.

In this work, FTIR followed by linear discriminant analysis (LDA) of the spectral data was used to classify vegetable oils according to their botanical origin. Also, data treatment by multiple linear regression (MLR) was used to detect and quantify EVOO adulteration with other low cost edible oils, including sunflower, corn, soybean and hazelnut oils.

2. Experimental

2.1. Oil samples and mixtures

The vegetable oils employed in this study (Table 1) were either purchased at the local market or kindly donated by the manufacturers. The botanical origin and quality grade of all the samples were guaranteed by the suppliers. As indicated in Table 1, four samples of each botanical origin were used for training purposes in classification studies, being the other two samples of each category used to evaluate the prediction capability of the classification models. To estimate the adulteration of EVOO with low cost oils by using regression models, binary mixtures containing EVOO and increasing percentages of low cost oil (sunflower, corn, soybean or hazelnut) were prepared. To improve robustness of MLR models, the objects of the calibration matrix were prepared using EVOOs and low cost oils from different geographical origins. For instance, for the sunflower-EVOO pair, oils from different geographical origins were selected to prepare a total of seven mixtures containing 0%, 5%, 10%, 30%, 50%, 75% and 100% sunflower oil. Sets of mixtures containing the same percentages of low cost oil were also prepared for the corn-, soybean- and hazelnut-EVOO pairs. The resulting 28 mixtures were used as calibration set to construct regression models. Additional mixtures of the sunflower-, corn-, soybean- and hazelnut-EVOO pairs, also using oils from different geographical origins, and containing 5%, 50% and 80% low cost oil were prepared. These 12 additional binary mixtures were used to validate the prediction performance of the regression models.

Table 1
Botanical origin, number of samples, brand and use during LDA model construction of the oil samples.

Origin	No. of samples	Brand	LDA set
Hazelnut	2	Guinama	Training
	2	Percheron	Training
	2	Flumen	Evaluation
Sunflower	2	Koipesol	Training
	2	Hacendado	Training
	1	Capicua	Evaluation
	1	Coosol	Evaluation
Corn	1	Guinama	Training
	1	Asua	Training
	1	Artua	Evaluation
	1	Mazola	Evaluation
Corn germ	1	Guinama	Training
	1	Hacendado	Training
Extra virgin olive	1	Carbonell	Training
	1	Grupo Hojiblanca	Training
	1	Borges	Training
	1	Torrereal	Training
	1	Coosur	Evaluation
	1	Hacendado	Evaluation
Soybean	2	Guinama	Training
	2	Biolasi	Training
	2	Sojola	Evaluation

2.2. FTIR spectra

FTIR spectra were obtained using a Nicolet Nexus FTIR spectrophotometer (Thermo Electron Corporation, Waltham, MA, USA) with a resolution of 4 cm^{-1} at 32 scans. A small quantity of the oil samples (≈2 μL) was directly deposited between two well-polished KBr disks, creating a thin film. Duplicated spectra were recorded for all the oil samples and binary mixtures, except the 12 mixtures used as validation set in regression studies which were recorded three times each. Spectra were scanned in the absorbance mode from 4000 to 500 cm^{-1} and the data were handled with the EZ OMNIC 7.3 software (Thermo Electron Corporation).

2.3. Data treatment and construction of matrices

FTIR spectra were divided in the 26 wavelength regions described in Table 2. Each selected spectral region corresponds to a peak or a shoulder, representing structural or functional group information, either about the lipids or minor components of the oil samples (see Table 2). For each region, the peak/shoulder area was measured. In order to reduce the variability associated to the total amount of oil sample used, and to minimize other sources of variance also affecting the intensity of all the peaks, such as the thickness of the sample and radiation source intensity, normalized rather than absolute areas were used. Two normalization procedures were tried. In procedure A, for each spectrum, the area of each region was divided by the sum of the areas of the 26 regions. In procedure B, the area of each region was divided by each one of the areas of the other 25 regions; in this way, and since any pair of areas should be considered only once, $(26 \times 25)/2 = 325$ normalized variables were obtained.

For classification studies, two matrices containing 20 objects each, which corresponded to the averages of the duplicated spectra of the training samples of Table 1, were constructed. These matrices had either 26 or 325 predictors, according to normalization procedures A and B, respectively. A response column, containing the categories corresponding to the five botanical origins of the oils (corn and corn germ were considered as a single category), was added to the training matrices. In order to reduce the internal dispersion of the categories, which was important to reduce the number of variables selected during model construction, the means of the two spectra of each sample, instead of the individual spectra, were used. For evaluation purposes, two more matrices containing 10 objects each, which corresponded to the averages of the duplicated spectra of each evaluation sample of Table 1, were constructed. These matrices also had either 26 or 325 predictors, according to normalization procedures A and B, respectively.

Concerning to the regression studies, and for the sunflower-EVOO pair, two calibration matrices containing seven objects each, which corresponded to the averages of the duplicated spectra of the calibration mixtures, were constructed. According to normalization procedures A and B, the number of predictors was either 26 or 325, respectively. A response column, containing the low cost oil percentages of the mixtures, was added to these matrices. For validation of the prediction performance, two more matrices containing three objects each, which corresponded to the averages of triplicate spectra of validation mixtures, were constructed. The number of predictors was either 26 or 325 predictors, as indicated. Analogously, matrices for the corn-, soybean- and hazelnut-EVOO pairs were also constructed. Statistical treatment of the data was performed using SPSS (v. 12.0.1, Statistical Package for the Social Sciences, Chicago, IL, USA).