Food Chemistry 131 (2012) 360–366

Contents lists available at SciVerse ScienceDirect

Food Chemistry

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

Analytical Methods

# Discrimination of five Tunisian cultivars by Mid InfraRed spectroscopy combined with chemometric analyses of olive *Olea europaea* leaves

Fathia Aouidi [a,c,*], Nathalie Dupuy [b], Jacques Artaud [b], Sevastianos Roussos [c], Monji Msallem [d], Isabelle Perraud-Gaime [c], Moktar Hamdi [a]

[a] *Laboratory of Microbial Ecology and Technology, Department of Biological and Chemical Engineering, National Institute of Applied Sciences and Technology (INSAT), Centre Urbain Nord, 2 Boulevard de la Terre, B.P. 676, 1080 Tunis, Tunisia*
[b] *ISM², UMR 6263, Equipe AD²EM, Groupe systèmes chimiques complexes, Faculté des Sciences et Technique Saint Jérôme – Université Paul Cézanne, 13397 Marseille Cedex 20, France*
[c] *IMEP UMR CNRS 6116/IRD UMR 193, Faculté des Sciences et Technique Saint Jérôme – Université Paul Cézanne, 13397 Marseille Cedex 20, France*
[d] *Institut de l'olivier, B.P. 208, 1082 Tunis, Tunisia*

ARTICLE INFO

*Article history:*
Received 13 December 2010
Received in revised form 7 March 2011
Accepted 17 August 2011
Available online 23 August 2011

*Keywords:*
Olive (*Olea europaea*) leaf
ATR-Mid InfraRed spectroscopy
Chemometric analysis
Cultivar
Discrimination
Chemical fingerprint

ABSTRACT

The high biodiversity of olive tree and the economic needs require tools for the correct classification and identification of the different cultivars. Simple and rapid methods are in increasing demand. In the present work, FT-MIR spectroscopy associated to chemometric treatment is proposed as a direct and rapid tool to discriminate cultivars according to their olive leaves, a persistent tissue the whole year. A set of 75 samples of olive leaves representative of five Tunisian cultivars (Chemlali, Sayali, Meski, Zarrazi and Chétoui) cultivated in the same geographical area was analysed. Discrimination between the five Tunisian cultivars was performed by the chemometric approach, principal component analysis (PCA), based on the FT-MIR spectral data provided by olive leaves. Furthermore, a correct classification (100%) of the five Tunisian cultivars was obtained by the Partial Least Square Discriminate Analysis (PLS-DA) method.



## 1. Introduction

It is well known that qualitative and quantitative aspects of the production of olive oils and table olives vary, among others, with the cultivars (Gómez-Rico, Fregapane, & Salvador, 2008; Hannachi et al., 2008; Pinheiro & Esteves da Silva, 2005). Nowadays, the identification of the olive cultivar become a topic of great economic relevance since the demand of table olives and olive oils is increasing and there is a growing commercial interest in high quality products. Furthermore, correct cultivar identification can represent a useful tool for nursery owners who need to certify and patent their plant material.

Until recently, olive cultivars are especially discriminated and classified on the base of morphological and agronomic characteristics (Pinheiro & Esteves da Silva, 2005; Trigui & Msallem, 2002). Identification based on the analysis of isoenzymes has also been cited (Trujillo, Rallo, & Arus, 1995). Although these markers provide useful tools for cultivar identification, their actual limits are the small number of detected polymorphisms and the influence of environment and of olive growing techniques. The recent development of DNA molecular markers has led to the emergence of new genetic markers for identifying olive cultivars. Molecular markers, such as Simple Sequence Repeat (SSR) (Alba, Montemurro, Sabetta, Pasqualone, & Blanco, 2009; Bracci et al., 2009; Doveri et al., 2008; Hannachi et al., 2008), Random Amplified Polymorphic DNA (RADP) (Zitoun et al., 2008) and Amplified Fragment Length Polymorphism (AFLP) (Belaj, Rallo, Trujillo, & Baldoni, 2004), are environment-independent and efficient to identify olive varieties. Recently, a chemometric approach based on the analytical data has been developed for discrimination and classification of olive cultivars. Analyses were carried out on a particular olive tissue such as leaves (Di Donna et al., 2010; Japón-Luján, Ruiz-Jimnez, & Luque de Castro, 2006) or olive fruits (Casale et al., 2010; Dupuy et al., 2010) or on olive oil (Casale, Sinelli, Oliveri, Di Egidio, & Lanteri, 2010; Dupuy, Galtier, Ollivier, Vanloot, & Artaud, 2010; Sinelli et al., 2010). The discrimination between varieties of olives trees cultivated in the same geographical area can be performed by chemotaxonomic markers, such as secondary metabolites of phenolic structure identified by high performance liquid chromatography/electrospray ionisation tandem mass spectroscopy (HPLC–ESI-MS) (Di Donna et al., 2010), peak areas from a high performance liquid chromatography-diode array (HPLC–DAD) analysis of biophenols (Japón-Luján et al.,

* Corresponding author at: Laboratory of Microbial Ecology and Technology, Department of Biological and Chemical Engineering, National Institute of Applied Sciences and Technology (INSAT), Centre Urbain Nord, 2 Boulevard de la Terre, University of Carthage, B.P. 676, 1080 Tunis, Tunisia. Tel.: +216 98 326 675.
*E-mail address:* aouidifathia@yahoo.fr (F. Aouidi).

0308-8146/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodchem.2011.08.041

2006), oleuropein content determined by high performance liquid chromatography (HPLC) and high resolution gas chromatography (HRGC) (Ranalli et al., 2006).

Most of analyses used for discriminating cultivars are time-consuming, expensive and involve a considerable amount of manual work. Very often, complex chemical treatment of the sample and the use of sophisticated instruments are required.

Recently, Fourier transform infrared (FT-IR) spectroscopy has become an emerging well-accepted analytical technique, due to its simplicity with advantages in terms of cost per sample. It achieves high analysis speed and requires little or no sample preparation.

FT-IR spectroscopy has been widely used as an analytical tool in various laboratories and industrial sectors such as food agricultural (Galtier et al., 2008; Hennessy, Downey, & O'Donnell, 2009), petrochemical (Abbas, Rebufa, Dupuy, Permanyer, & Kister, 2008; Roman & Ravilya, 2008), textile (Durand, Devos, Ruckebusch, & Huvenne, 2007; Langeron, Doussot, Hewson, & Duchêne, 2007) and pharmaceutical (Blanco, Cueva-Mestanza, & Peguero, 2010; Wu, Sun, Zhou, & Leung, 2008).

Up to date, a lot of studies have been published on the utilisation of Near and Mid FT-IR for authentication, identification or classification of many agro-foods, notably olive oils (Casale, Sinelli, et al., 2010; Dupuy, Galtier, olliver, et al., 2010; Galtier et al., 2008; Sinelli et al., 2010) and table olives (Casale et al., 2010; Dupuy et al., 2010) by multivariate statistical analysis of spectral data. Near FT-IR (FT-NIR) has been applied on olive leaves to discriminate between the juvenile and adult leaves (Leon & Downey, 2006) and for prediction of nutritive composition (Fernandez-Cabanas, Garrido-Varo, Delgado-Pertinez, & Gomez-Cabrera, 2008). To our knowledge, no work has been conducted on olive leaves by application of Mid FT-IR (FT-MIR) spectroscopy in association with chemometric treatments.

The aim of this study was to develop by FT-MIR spectroscopy associated to chemometric treatment, a direct and rapid tool to discriminate five Tunisian cultivars according to their olive leaves.

## 2. Materials and methods

### 2.1. Leaf samples

Experiments were carried out on 75 samples of olive leaves from five Tunisian cultivars (Chemlali (A), Sayali (B), Meski (C), Zarrazi (D) and Chétoui (E)). Well-expanded leaves, from the current season, were collected in 19th October 2009 from 15 olive trees. All trees were cultivated in the same experimental field and cultural conditions of the collection of olive trees varieties (RESGE CFC/COI/03) – Oued Souhil – Nabeul 2004 (Tunisia). For each of the five cultivars, the samples were obtained from three different trees (coded (1), (2) and (3)). From each tree, leaves were harvested from the five orientations of tree (Centre (C), North (N), South (S), East (E) and West (W)). In all cases, every sample contains about 100 healthy leaves. A morphological characteristic of five cultivar leaves is presented in Table 1.

**Table 1**
Morphological characteristics of five cultivar leaves.

| Cultivar leaves | Length (mm) | Width (mm) | Form = length/width |
|---|---|---|---|
| Chemlali | 52,93 ± 8,84$^{a}$ | 10,70 ± 1,76$^{a}$ | 5,03 ± 1,06$^{bc}$ |
| Chétoui | 54,97 ± 5,34$^{a}$ | 12,37 ± 1,45$^{b}$ | 4,49 ± 0,57$^{a}$ |
| Meski | 62,07 ± 6,68$^{b}$ | 11,63 ± 1,40$^{b}$ | 5,38 ± 0,65$^{c}$ |
| Sayali | 79,60 ± 8,81$^{c}$ | 11,87 ± 1,50$^{b}$ | 6,79 ± 1,00$^{d}$ |
| Zarrazi | 65,97 ± 10,34$^{b}$ | 13,52 ± 1,88$^{c}$ | 4,89 ± 0,48$^{b}$ |

Results were expressed as mean ± standard deviation.
Values in each column with different letters differs significantly (*p-Values* < 0.05).

All samples were dried on site in a microwave two times for 2 min at maximum power 800 W (2450 MHz). Dried leaves were powdered in a blender IKA Labortechnik A10 and stored at 4 °C in the dark until use.

### 2.2. Fourier transform infra-red (FTIR) analysis

#### 2.2.1. Instrumentation

Mid-Infrared spectra of samples were obtained using a Thermo Nicolet AVATAR 370 FT-IR spectrometer equipped with a deuterotriglycine sulphate (DTGS) detector, an Ever-Glo Mid-Infrared source and a KBr/germanium beam splitter. The InfraRed spectrometer was situated in an air-conditioned room (21 °C).

#### 2.2.2. Analysis conditions

Dried powdered olive leaf samples were analysed by Mid-Infra-Red spectroscopy. Around 30–50 mg of powder is sufficient to perform Mid-Infrared spectral analysis.

Samples were deposited on attenuated total reflection (ATR) cell equipped with a diamond crystal prism (monoreflexion). Mid-Infrared spectra were recorded between 4000 and 700 $cm^{-1}$. The nominal resolution was 4 $cm^{-1}$, and 100 scans were co-added. Air was taken as reference for the background spectrum before collection of each sample spectrum. Between each spectrum, the ATR plate was cleaned in situ by scrubbing with ethanol solution, which made it possible to dry the ATR. Cleanliness was verified by collecting a background spectrum and comparing it to the previous background spectrum.

Each sample was scanned with three replicates. The scans of each sample were examined visually for consistence, and the average spectrum of each sample was used for further analyses.

### 2.3. Chemometric treatments of FT-MIR spectra

The FT-MIR data were transformed with standard normal variant (SNV) to remove slope variation and to correct for scatter effects. The spectral variation was analysed by principal component analysis (PCA) on the average of the spectral data for each sample in order to discriminate olives leaf samples on the basis of their cultivars. Partial Least Square Discriminate Analysis (PLS-DA) method was performed with the purpose of olive cultivar identification.

PLS-DA was carried out using an exclusive binary coding scheme with one bit per cultivar, providing a quintuplet {a; b; c; d; e} since the aim of this work was to discriminate between five cultivar. Each number represents a "membership value" for each cultivar, e.g., a response encoded {0; 1; 0; 0; 0} means that the sample belongs to cultivar B (Sayali). During the calibration process, the PLS-DA method was trained to compute the five "membership values", one for each cultivar; the sample was then assigned to the cultivar showing the highest membership value.

The whole spectral collection included 75 spectra, each corresponding to one sample. In order to carry out a validation test, the data set was divided into two groups: two third of the samples (50 observations) were used for calibration and one third (25 observations) was used for validation.

The model was built by the block cross validation method during the calibration developments. Each block contained 10 samples. The evaluation of the errors in the calibration was carried out by computing the standard error of calibration (SEC) after comparing the real "membership value" with the computed one for each cultivar. The formula for the standard error of calibration is:

$$SEC = \sqrt{\left(\frac{\sum_{i=1}^{N}(C_i - C_i')^2}{N - 1 - p}\right)} \quad (1)$$

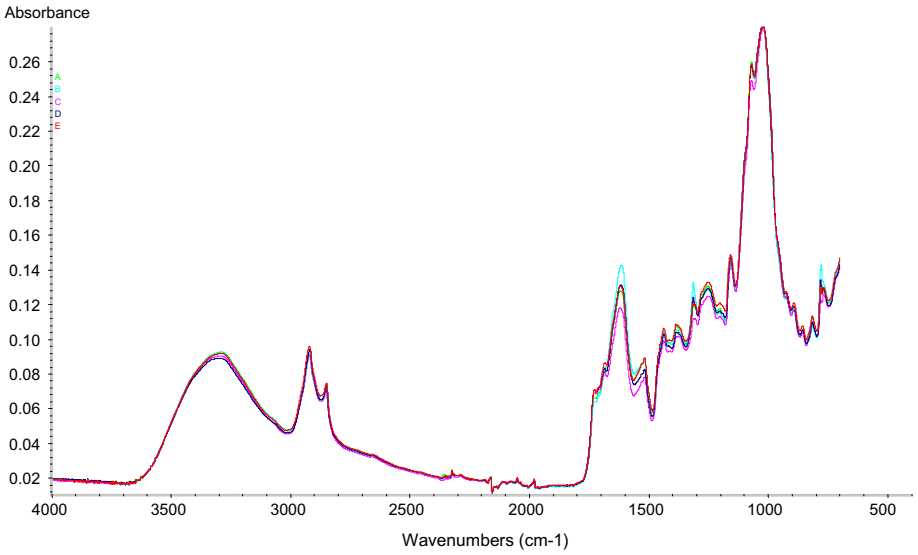


**Fig. 1.** The typical Mid FTIR-ATR spectra of powdered olive leaf samples recorded for the five different cultivars.

where $C_i$ is the known value, $C_i'$ is the value calculated by the calibration equation, $N$ is the number of samples and $p$ is the number of independent variables in the regression optimised by cross-validation. The standard error of prediction (SEP) gives an estimation of the prediction performance during the step of validation of the calibration equation:

$$\mathrm{SEP} = \sqrt{\left(\frac{\sum_{i=1}^{M}(C_i - C_i')^2}{M}\right)} \qquad (2)$$

where $M$ is the number of samples in the prediction set. Acceptable models should have low SEC and SEP, and high Coefficient of correlation $R^2$ between predicted and reference value.

### 2.4. Software for FTIR data treatments

FTIR-ATR spectra have been recorded by the instrument software OMNIC 4.1b (Thermo Nicolet). UNSCRAMBLER software the version 9.6 from CAMO (Computer Aided Modelling, Trondheim, Norway) has been used for chemometric treatment of FTIR-ATR data spectra.

## 3. Results and discussion

### 3.1. Typical FTIR spectra of olive leaves

The Mid-FTIR spectra obtained for all samples of dried powdered olive leaves were similar by visual inspection. Typical average spectra from each cultivar were presented in Fig 1. All spectra are characterised by common absorption bands. Many overlapped bands appear at the spectra region between 700 and 1750 $cm^{-1}$. A large band was observed at 3000–3600 $cm^{-1}$. An intense absorbance was recorded for two bands at 2800–3000 $cm^{-1}$.

Despite the similarity of the general pattern, some differences in the relative intensities of spectral absorbance were observed between samples from various cultivars. Direct molecular interpretation of the spectra is very difficult for complex sample matrices especially for samples such as olive leaves.

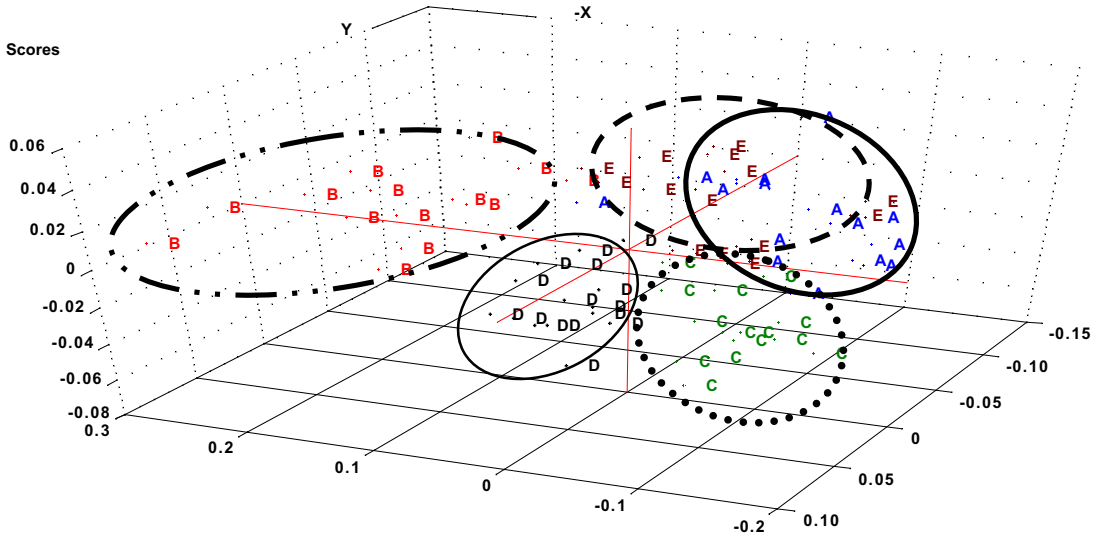


**Fig. 2.** Discrimination of five Tunisian cultivars (Chemlali (A), Sayali (B), Meski (C), Zarrazi (D) and Chétoui (E)): Three dimensional plots of powdered olive leaf samples investigated in principal component analysis using the Mid-Infrared spectra of samples analysed in triplicate (74% of variance explained).

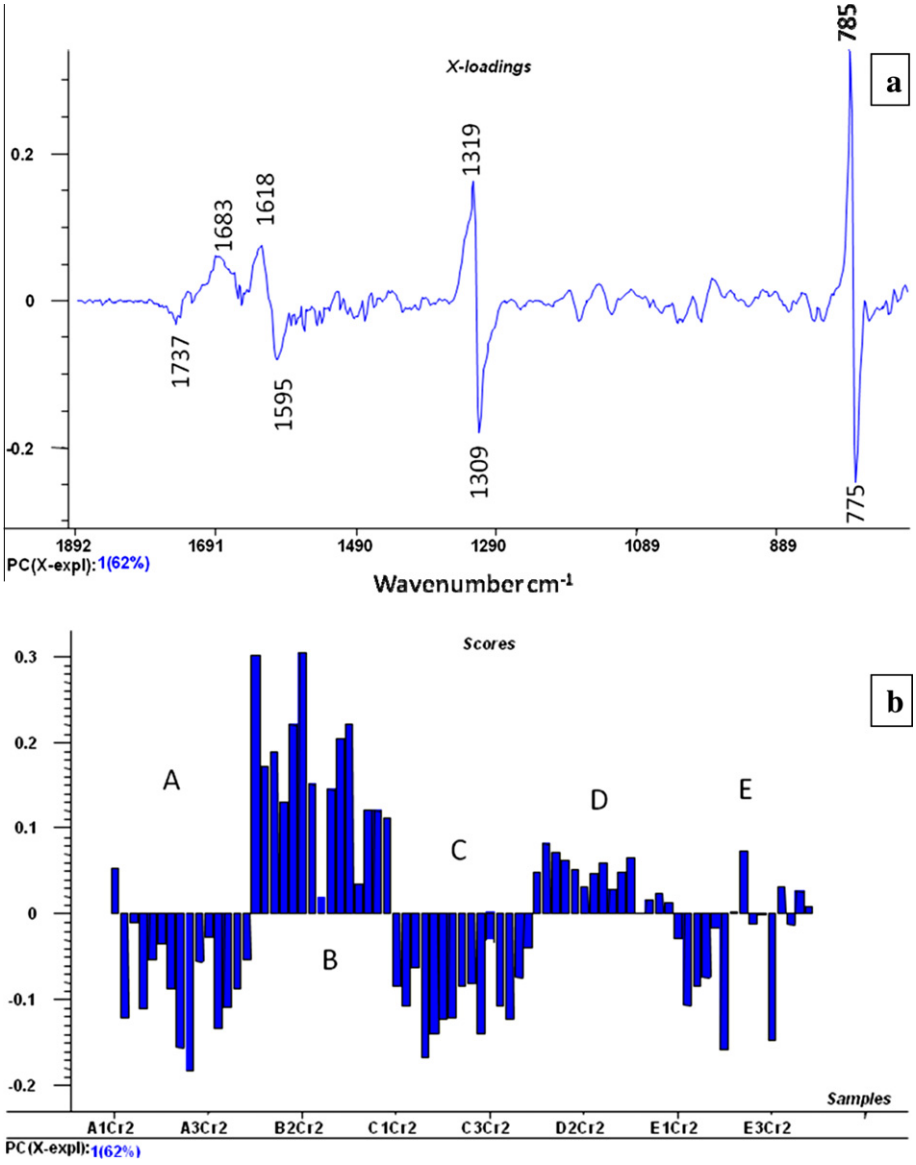


**Fig. 3.** (a) First loading obtained in principal component analysis. (b) Scores obtained on the first loading.

Because each chemical compound in the olive leaf sample contributes to the absorbance spectrum, the MIR spectrum of olive leaf sample contains information for the compounds which are present. Thus, the olive leaf spectrum may be considered to represent its overall chemical composition and therefore has the potential to characterise leaves and acts as a fingerprint.

The visual examination of the spectral variations did not permit to apprehend clearly the difference of chemical structure and chemical species concentration in olive leaves depending on cultivar. Chemometric treatments are, often, applied in order to extract information from the spectral data set.

### *3.2. Discrimination and classification of olive leaf samples according to their cultivar*

Today, multivariate analysis is an essential chemometric tool to study data came from many observations made on several variables. Its aim is to resume information contained in data with a reduced number of dimensions to characterise as well as possible the differences or similarities between observations and variables.

Olive leaf samples were powder's solids in the experiments. Effect on scatter of MIR light existed when passing powder sample. Therefore, SNV preprocessing method was applied on all spectra in order to remove physical spectral information (due to particle size), so that discrimination and predicted models will be performed based on mainly chemical spectral information.

Then, the spectral data set of olive leaf samples was subjected to the basic tool for data analysis: principal component analysis (PCA). PCA is very important especially in the preliminary steps of a multivariate analysis to perform an exploratory analysis in order to have an overview of data. It allows describing data set without a priori knowledge of the data structure.

PCA has been carried out on Mid-FTIR spectra in order to assay to discriminate olive leaf samples according to their cultivar. It was performed on the 75 spectra: 15 average spectra recorded in triplicate for each one of the five cultivars.

Fig 2 shows the 3D plot of principal component scores. This plot represents a 74% of total variance in the spectral data. The first principal component accounts for 62% of the variance. It is clear that there was information related to the varietal origin in the MIR spectra of olive leaf samples but there is an overlay between Chemlali leaves (A) and Chétoui one (E). The examination of the first loading (Fig. 3a) and the associated scores (Fig 3b) shown the wavenumbers highly correlated with the species. The Fig 3b

**Table 2**
Determination of calibration and prediction parameters of cultivars by chemometric treatment of MIR spectra.

| Cultivars | Calibration | | | | Prediction | | | |
|---|---|---|---|---|---|---|---|---|
| | Samples Numbers | $R^2$ | SEC | Bias | Samples Numbers | $Q^2$ | SEP | Bias |
| Chemlali | 50 | 0.98 | 0.054 | 0.000 | 25 | 0.88 | 0.193 | 0.004 |
| Chétoui | 50 | 0.98 | 0.047 | −0.000 | 25 | 0.88 | 0.193 | −0.045 |
| Meski | 50 | 0.97 | 0.064 | −0.001 | 25 | 0.89 | 0.182 | −0.010 |
| Sayali | 50 | 0.98 | 0.053 | 0.000 | 25 | 0.85 | 0.216 | −0.019 |
| Zarrazi | 50 | 0.96 | 0.073 | 0.001 | 25 | 0.85 | 0.209 | 0.083 |

Spectral rang ($cm^{-1}$): 4000–700; $R^2$: correlation coefficient between calibrated and reference value; SEC: standard error of calibration; $Q^2$: correlation coefficient between predicted and reference value; SEP: standard error of prediction; bias: difference between an estimator's expectation and the true value of the parameter being estimated.

**Table 3**
Prediction of leaf cultivars by chemometric analysis of MIR spectra.

| Cultivars | Samples | Chemlali | | Chétoui | | Meski | | Sayali | | Zarrazi | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Pred. | Ref. | Pred. | Ref. | Pred. | Ref. | Pred. | Ref. | Pred. | Ref. |
| Chemlali | A1N | 0.965 | 1 | −0.180 | 0 | −0.047 | 0 | −0.155 | 0 | **0.477** | 0 |
| | A2C | 0.587 | 1 | 0.160 | 0 | −0.074 | 0 | −0.074 | 0 | 0.140 | 0 |
| | A2S | 0.911 | 1 | −0.133 | 0 | −0.074 | 0 | −0.152 | 0 | 0.186 | 0 |
| | A3E | 1.047 | 1 | −0.280 | 0 | −0.163 | 0 | 0.184 | 0 | 0.174 | 0 |
| | A3W | 0.981 | 1 | −0.210 | 0 | −0.029 | 0 | 0.081 | 0 | 0.043 | 0 |
| Chétoui | E1N | 0.043 | 0 | 0.882 | 1 | 0.186 | 0 | −0.099 | 0 | 0.124 | 0 |
| | E2C | 0.164 | 0 | 0.916 | 1 | −0.035 | 0 | −0.042 | 0 | 0.032 | 0 |
| | E2S | −0.165 | 0 | 0.825 | 1 | 0.183 | 0 | 0.118 | 0 | −0.025 | 0 |
| | E3E | 0.228 | 0 | 0.780 | 1 | −0.106 | 0 | 0.196 | 0 | 0.171 | 0 |
| | E3W | 0.314 | 0 | 0.702 | 1 | 0.085 | 0 | 0.007 | 0 | −0.052 | 0 |
| Meski | C1N | −0.142 | 0 | 0.035 | 0 | 0.690 | 1 | 0.033 | 0 | 0.419 | 0 |
| | C2C | 0.211 | 0 | −0.122 | 0 | 1.078 | 1 | −0.131 | 0 | 0.072 | 0 |
| | C2S | −0.144 | 0 | 0.226 | 0 | 0.722 | 1 | 0.128 | 0 | 0.067 | 0 |
| | C3E | 0.111 | 0 | −0.019 | 0 | 0.617 | 1 | 0.172 | 0 | 0.070 | 0 |
| | C3W | 0.230 | 0 | −0.409 | 0 | 0.798 | 1 | 0.032 | 0 | 0.256 | 0 |
| Sayali | B1N | 0.367 | 0 | −0.293 | 0 | −0.001 | 0 | 0.702 | 1 | **0.450** | 0 |
| | B2C | 0.094 | 0 | −0.069 | 0 | 0.071 | 0 | 0.608 | 1 | **0.464** | 0 |
| | B2S | −0.138 | 0 | 0.151 | 0 | 0.062 | 0 | 0.862 | 1 | −0.039 | 0 |
| | B3E | 0.064 | 0 | 0.272 | 0 | 0.061 | 0 | 0.984 | 1 | −0.172 | 0 |
| | B3W | −0.300 | 0 | 0.003 | 0 | **0.469** | 0 | 1.122 | 1 | −0.141 | 0 |
| Zarrazi | D1N | 0.097 | 0 | 0.158 | 0 | 0.170 | 0 | 0.054 | 0 | 0.775 | 1 |
| | D2C | −0.273 | 0 | 0.166 | 0 | 0.036 | 0 | −0.037 | 0 | 1.011 | 1 |
| | D2S | −0.125 | 0 | −0.093 | 0 | −0.103 | 0 | 0.072 | 0 | 0.995 | 1 |
| | D3E | −0.051 | 0 | 0.192 | 0 | −0.186 | 0 | 0.065 | 0 | 0.750 | 1 |
| | D3W | −0.189 | 0 | 0.200 | 0 | 0.023 | 0 | 0.107 | 0 | 0.836 | 1 |

Pred.: predicted; Ref.: reference; **0.000:** suspicious results; spectral rang ($cm^{-1}$): 4000–700.

shows that the species B and D are positively correlated to the first loading positive part (1683, 1618, 1319, and 785 $cm^{-1}$) and negatively correlated to the negative part. The species A and C are negatively correlated to the first loading positive part and positively correlated to the negative part (1737, 1595, 1309 and 775 $cm^{-1}$). The species E do not present correlation on this loading. The chemical information associated is difficult to interpret because olive leaves are very complex systems, even if the difference in spectral data may be attributed to the difference in chemical composition, with qualitative and quantitative aspects.

Olive tree cultivars have been previously separated according to their chemical composition, such as DNA (Bracci et al., 2009; Zitoun et al., 2008), isoenzyme (Trujillo et al., 1995), phenolic compounds (Di Donna et al., 2010; Japón-Luján et al., 2006).

In the same framework, Casale et al. and Dupuy et al. have investigated IR spectroscopy as a tool for olive cultivars discrimination by analysing table olive fruits (Casale et al., 2010; Dupuy et al., 2010) and extra virgin olive oil (Casale et al., 2010; Dupuy, Galtier, Olliver, et al., 2010; Sinelli et al., 2010).

Therefore, five classification models (Chemlali, Sayali, Meski, Zarrazi and Chétoui) have been built, using a PLS-DA method as a chemometric treatment of spectra. The performance of models was evaluated in term of correlation coefficient ($R^2$), standard error of calibration (SEC), standard error of prediction (SEP) and bias. Results of calibration and prediction models were presented in Table 2. The five models are acceptable. In fact, they have low SEC (0.047–0.073), low SEP (0.182–0.216), high $R^2$ (0.96–0.98), low bias (−0.019–0.083) between predicted and reference spectrum.

Results obtained for all samples used in prediction set of the five cultivars are presented in Table 3. The predicted cultivars are never given by 0 or 1 results because the chemical composition in the samples varies according to the olive leaf cultivars. As a matter of fact, there is a natural variation of the chemical composition which can notably be a function of tree age, olive leaf age.

Considering the difficulty to calibrate and predict the cultivar with binary variables, it was necessary to discriminate the results between the initial values 0 or 1. We considered that all the values between negative value and 0.49 conduce to a non recognised sample and the ones between 0.51 and high positive one to a recognised sample. With those limits, 100% of the samples from the Tunisian cultivars were well predicted. The results between 0.45 and 0.55 were suspicious because they were very close to the threshold 0.5. The 0.45 and 0.55 limits have been chosen because they express 10% of error in the results. With those limits four samples have been detected as having suspicious origin that represents 16% of all samples.

100% of the samples from "Chétoui", "Zarrazi" and "Meski" cultivars were well predicted. But, only 80% and 40%, respectively of

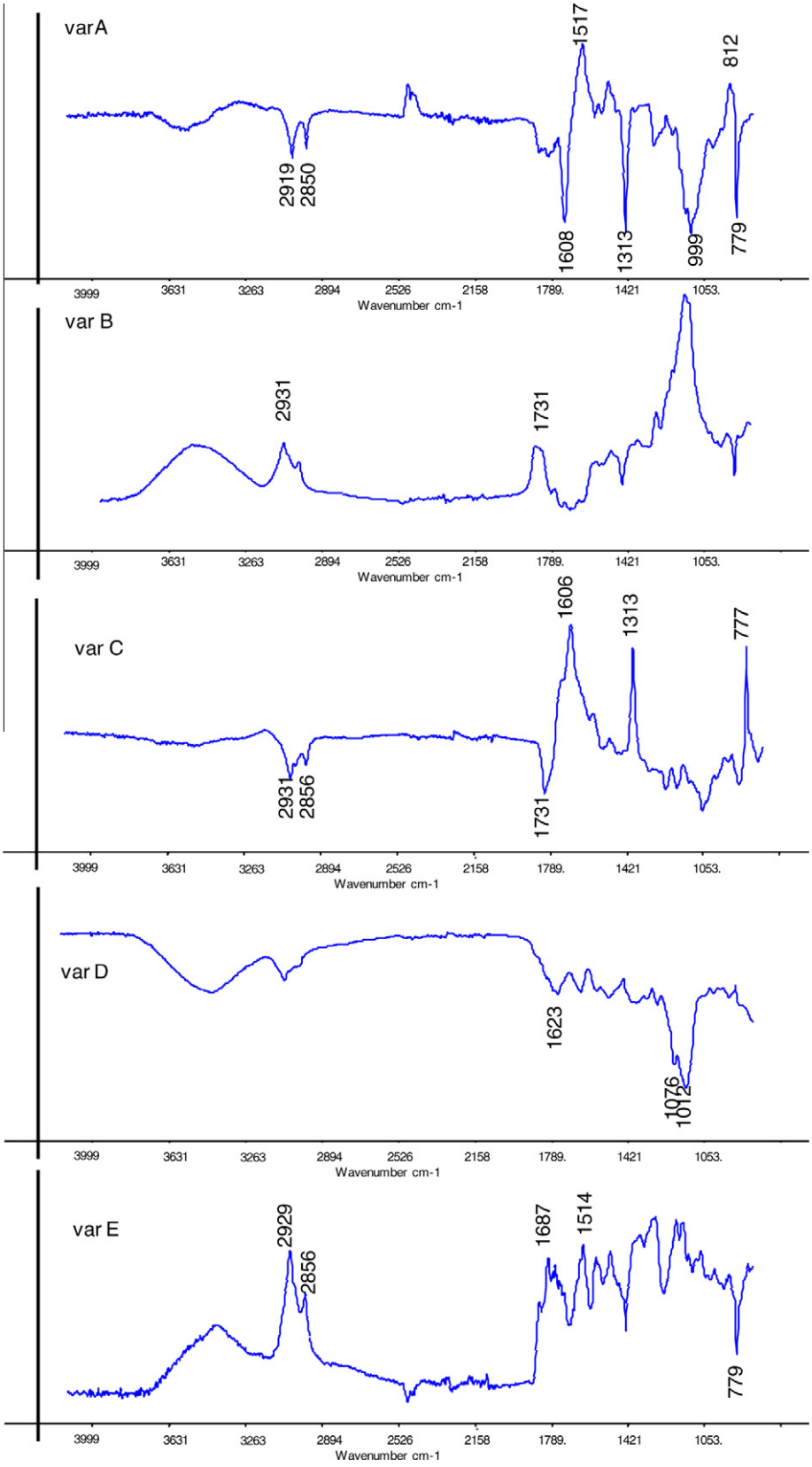


**Fig. 4.** First regression coefficient obtained for the five determinations of cultivars (Chemlali (A), Sayali (B), Meski (C), Zarrazi (D) and Chétoui (E)).

the samples from "Chemlali" and "Sayali" cultivars were well predicted. Some samples were not reliable to their cultivars: one sample for "Chemlali" cultivar (A1N) has some characteristics with "Zarrazi" cultivar; three samples for "Sayali" cultivar (B1N, B2C and B3W) are identified as "Sayali" cultivar but they have some characteristics with "Meski" and "Zarrazi" cultivar. As it was

possible to classify samples as a function of their origin, it is interesting to understand how this classification was carried out on the basis of MIR spectra. In the case of PLS regression between spectra and compounds, it is well known that the first loading (noted "B") was a good approximation of the pure compound spectrum (Haaland & Thomas, 1988). Thus, loadings obtained for all the origins provide an approximation of the original feature of olive leaves (Fig. 4). The A variety is correlated at 1517 and 812 $cm^{-1}$, the B one to 2931, 1731 and 1016 $cm^{-1}$, the C one to 1606, 1313 and 777 $cm^{-1}$, the D one is anti correlated to 1623, 1076 and 1012 $cm^{-1}$ and the E one is correlated to 2929, 2856, 1687 and 1514 $cm^{-1}$. Each variety is identified on different spectral information.

The rapidity of analyses, the availability of leaves all the year and the feasibility of the technique, among other, allow the chemometric treatment of FT-MIR spectra of olive leaves to be a promising approach for olive tree cultivar discrimination and identification.

## 4. Conclusion

The discrimination between five Tunisian cultivars of olive trees cultivated in the same geographical area can be performed by a chemometric approach based on the FT-MIR spectral data provided by olive leaves, a persistent tissue the whole year. Furthermore, MIR spectroscopy presents high potential for cultivar differentiation and prediction by PLS-DA method. Therefore, a simple, rapid and reliable overall characterisation of olive cultivar may be obtained at a low cost. It might be an application for rapidly classification of olive cultivar.

However, this study presents some limitation and further work is necessary to obtain more robust classification rules, able to account in a better way both for the regional and for the time variability. It should analyse a larger number of samples representative of wide variability (geographical origin, harvesting period and years).

## Acknowledgements

Fathia Aouidi is thankful to the Tunisian Ministry of Scientific Research, Technology and Competences Development for providing grant. We also thank Dr. Ali Ben Diah from Olive Institute for technical assistance in sample collection.

## References

Abbas, O., Rebufa, C., Dupuy, N., Permanyer, A., & Kister, J. (2008). Assessing petroleum oils biodegradation by chemometric analysis of spectroscopic data. *Talanta, 75*, 857–871.

Alba, V., Montemurro, C., Sabetta, W., Pasqualone, A., & Blanco, A. (2009). SSR-based identification key of cultivars of *Olea europaea* L. diffused in Southern-Italy. *Scientia Horticulturae, 123*, 11–16.

Belaj, A., Rallo, L., Trujillo, I., & Baldoni, L. (2004). Using RAPD and AFLP markers to distinguish individuals obtained by clonal selection of Arbequina and Manzanilla de Sevilla olive. *HortScience, 39*, 1566–1570.

Blanco, M., Cueva-Mestanza, R., & Peguero, A. (2010). Controlling individual steps in the production process of paracetamol tablets by use of NIR spectroscopy. *Journal of Pharmaceutical and Biomedical Analysis, 51*(4), 797–804.

Bracci, T., Sebastiani, L., Busconi, M., Fogher, C., Belaj, A., & Trujillo, I. (2009). SSR markers reveal the uniqueness of olive cultivars from the Italian region of Liguria. *Scientia Horticulturae, 122*, 209–215.

Casale, M., Sinelli, N., Oliveri, P., Di Egidio, V., & Lanteri, S. (2010). Chemometrical strategies for feature selection and data compression applied to NIR and MIR spectra of extra virgin olive oils for cultivar identification. *Talanta, 80*, 1832–1837.

Casale, M., Zunin, P., Cosulich, M.-E., Pistarino, E., Perego, P., & Lanteri, S. (2010). Characterisation of table olive cultivar by NIR spectroscopy. *Food Chemistry, 122*(4), 1261–1265.

Di Donna, L., Mazzotti, F., Naccarato, A., Salerno, R., Tagarelli, A., Taverna, D., et al. (2010). Secondary metabolites of *Olea europaea* leaves as markers for the discrimination of cultivars and cultivation zones by multivariate analysis. *Food Chemistry, 121*, 492–496.

Doveri, S., Gil, F. S., Diaz, A., Reale, S., Busconi, M., Camara Machado, A., et al. (2008). Standardization of a set of microsatellite markers for use in cultivar identification studies in olive (*Olea europaea* L.). *Scientia Horticulturae, 116*, 367–373.

Dupuy, N., Galtier, O., Le Dréau, Y., Pinatel, C., Kister, J., & Artaud, J. (2010). Chemometric analysis of combined NIR and MIR spectra to characterize French olives. *European Journal of Lipid Science and Technology, 112*, 463–475.

Dupuy, N., Galtier, O., Ollivier, D., Vanloot, P., & Artaud, J. (2010). Comparison between NIR, MIR, Concatenated NIR and MIR analysis and hierarchical PLS model. Application to virgin olive oils analysis. *Analytica Chimica Acta, 666*, 23–31.

Durand, A., Devos, O., Ruckebusch, C., & Huvenne, J. P. (2007). Genetic algorithm optimisation combined with partial least squares regression and mutual information variable selection procedures in near-infrared quantitative analysis of cotton–viscose textiles. *Analytica Chimica Acta, 595*, 72–79.

Fernandez-Cabanas, V. M., Garrido-Varo, A., Delgado-Pertinez, M., & Gomez-Cabrera, A. (2008). Nutritive evaluation of olive tree leaves by near-infrared spectroscopy: Effect of soil contamination and correction with spectral pretreatments. *Applied Spectroscopy, 62*(1), 51–58.

Galtier, O., Le Dréau, Y., Ollivier, D., Kister, J., Artaud, J., & Dupuy, N. (2008). Lipid compositions and French Registered Designations of origins of virgin olive oils predicted by chemometric analysis of Mid-infrared spectra. *Applied Spectroscopy, 62*(5), 583–590.

Gómez-Rico, A., Fregapane, G., & Salvador, M. D. (2008). Effect of cultivar and ripening on minor components in Spanish olive fruits and their corresponding virgin olive oils. *Food Research International, 41*, 433–440.

Haaland, D. M., & Thomas, E. V. (1988). Partial least-squares methods for spectral analysis. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry, 60*, 1193–1202.

Hannachi, H., Breton, C., Msallem, M., Ben El Hadj, S., El Gazzah, M., & Bervillé, A. (2008). Differences between native and introduced olive cultivars as revealed by morphology of drupes, oil composition and SSR polymorphisms: A case study in Tunisia. *Scientia Horticulturae, 116*, 280–290.

Hennessy, S., Downey, G., & O'Donnell, C. P. (2009). Confirmation of food origin claims by fourier transform infrared spectroscopy and chemometrics: Extra virgin olive oil from Liguria. *Journal of Agriculture and Food Chemistry, 57*(5), 1735–1741.

Japón-Luján, R., Ruiz-Jimnez, J., & Luque de Castro, M. D. (2006). Discrimination and classification of olive tree varieties and cultivation zones by biophenol contents. *Journal of Agricultural and Food Chemistry, 54*(26), 9706–9712.

Langeron, Y., Doussot, M., Hewson, D. J., & Duchêne, J. (2007). Classifying NIR spectra of textile products with kernel methods. *Engineering Applications of Artificial Intelligence, 20*, 415–427.

Leon, L., & Downey, G. (2006). Preliminary studies by visible and near-infrared reflectance spectroscopy of juvenile and adult olive (*Olea europaea* L.) leaves. *Journal of the Science of Food and Agriculture, 86*, 999–1004.

Pinheiro, P. B. M., & Esteves da Silva, J. C. G. (2005). Chemometric classification of olives from three Portuguese cultivars of *Olea europaea* L. *Analytica Chimica Acta, 544*, 229–235.

Ranalli, A., Contento, S., Lucera, L., Di Febo, M., Marchegiani, D., & Di Fonzo, V. (2006). Factors affecting the contents of iridoid oleuropein in olive leaves (*Olea europaea* L.). *Journal of Agricultural and Food Chemistry, 54*(2), 434–440.

Roman, M. B., & Ravilya, Z. S. (2008). Gasoline classification by source and type based on near infrared (NIR) spectroscopy data. *Fuel, 87*, 1096–1101.

Sinelli, N., Casale, M., Di Egidio, V., Oliveri, P., Bassi, D., Tura, D., et al. (2010). Varietal discrimination of extra virgin olive oils by near and mid infrared spectroscopy. *Food Research International, 43*, 2126–2131.

Trigui, A., & Msallem, M. (2002). Oliviers de Tunisie, In *Catalogue des variétés autochtones et types locaux, identification variétale et caractérisation morpho-pomologique des ressources génétiques oléicoles de Tunisie* (Vol. 1). (Fr) Ministère de l'Agriculture, IRESA, Institut de l'Olivier, Tunisia.

Trujillo, I., Rallo, L., & Arus, P. (1995). Identifying olive cultivars by isoenzyme analysis. *Journal of the American Society for Horticultural Science, 120*, 318–324.

Wu, Y. W., Sun, S. Q., Zhou, Q., & Leung, H. W. (2008). Fourier transform mid-infrared (MIR) and near-infrared (NIR) spectroscopy for rapid quality assessment of Chinese medicine preparation Honghua Oil. *Journal of Pharmaceutical and Biomedical Analysis, 46*(3), 498–504.

Zitoun, B., Bronzini de Caraffa, V., Giannettini, J., Breton, C., Trigui, T., Maury, J., et al. (2008). Genetic diversity in Tunisian olive accessions and their relatedness with other Mediterranean olive genotypes. *Scientia Horticulturae, 115*, 416–419.