Food Chemistry 328 (2020) 127164

Contents lists available at ScienceDirect

Food Chemistry

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

# Simple distinction of grapevine (*Vitis vinifera* L.) genotypes by direct ATR-FTIR

Ángela Álvarez[a], Jorge Yáñez[a,*], Yamil Neira[b], Rosario Castillo-Felices[b], Patricio Hinrichsen[c]

[a] Departamento de Química Analítica e Inorgánica, Laboratorio de Trazas Elementales y Especiación (LabTres), Facultad de Ciencias Químicas, Universidad de Concepción, POBox 160-C, Concepción, Chile

[b] Departamento de Análisis Instrumental, Facultad de Farmacia, Universidad de Concepción, POBox 160-C, Concepción, Chile

[c] Instituto de Investigaciones Agropecuarias, INIA La Platina. Santa Rosa 11, 610 Santiago, Chile

ARTICLE INFO

*Keywords:*
Leaves
Grapevine (*Vitis vinifera* L.)
Infrared spectroscopy (IR)
Non-destructive
Genotypes
Classification

ABSTRACT

The identification of grapevine (*Vitis vinifera* L.) genotypes is conventionally a laborious activity that must be carried out by specialized staff. In this work a novel and simple method for differentiation of grapevine genotypes is presented. Direct measurements of leaves by attenuated total reflectance Fourier-transform infrared spectroscopy (ATR-FTIR) combined with chemometric methods were used for classification of six genotypes (five varieties and a pair of clones), viz. Cinsault, Gewurztraminer (clone 643), Moscatel de Alejandría, País, Pinot Noir (French clone 777), Pinot Noir (local clone 'Valdivieso'). These were successfully classified and identified through supervised pattern recognition methods such as soft independent modeling of class analogy (SIMCA) and partial least square discriminant analysis (PLS-DA). The error rate for spectra classification of test sets by both models was 0.08. The results demonstrate the advantages of using ATR-FTIR as a rapid and non-destructive tool that achieves accurate grapevine genotype differentiation.

## 1. Introduction

Monitoring quality throughout the winemaking process (from soil to bottle) has great importance from both the viticulture and the enological point of view. Nowadays, methods for the identification of grapevine (*Vitis vinifera* L.) genotypes are visual ampelography (Galet, 1979) and DNA analysis (Borrego, De Andrés, Gómez, & Ibáñez, 2002). Nevertheless, these methods are time-consuming and involve labor-intensive intervention of experts. In this work, a novel screening method based on attenuated total reflectance (ATR) and Fourier transform infrared spectroscopy (FTIR) combined with advanced statistics for grapevine genotypes differentiation is presented. Advantages of ATR-FTIR are rapid data acquisition, user-friendliness, its relatively low cost and the fact that it is nondestructive. The vibrational spectrum is a molecular fingerprint for any type of biological/botanical samples (da Luz, 2006; Milosevic, 2004; Schmidtke, Smith, Müller, & Holzapfel, 2012; Shah, Cynkar, Smith, & Cozzolino, 2010).

In ATR-FTIR the IR spectrum is obtained from the surface of the sample. The sample is placed in contact with an internal reflection material (in this case Ge crystal) with a high refractive index, allowing total reflection of incident light multiple times. The surface of the sample interacts with the evanescent wave resulting in the absorption of radiation at each point of reflection (da Luz, 2006; Milosevic, 2004; Schmidtke et al., 2012).

Different studies in viticulture have used a number of analytical techniques such as gas chromatography-mass spectrometry (GC–MS), high performance liquid chromatography (HPLC), atomic absorption/emission spectrometry (AAS, AES), inductively coupled plasma optical emission spectrometry (ICP-OES), X-ray and voltammetry (Carrascon, Ontañón, Bueno, & Ferreira, 2017; Dalipi, Marguí, Borgese, Bilo, & Depero, 2016; Eftekhari et al., 2017; Kilmartin, 2016; Mierczynska-Vasilev & Smith, 2016; Mollo, Sixto, Falchi, Medina, & Knochen, 2017; Mutic et al., 2011; Ozbek & Akman, 2015; Shi et al., 2014). Vibrational spectroscopy has played a special role in viticulture, allowing to follow all the winemaking process analytically. Several studies have been focused on the evaluation of berry growing and maturation, monitoring wine fermentation, measuring quality parameters and sensory attributes, and even determining wine's geographic origin (Cozzolino, Cynkar, Dambergs, Shah, & Smith, 2013; Dambergs, Gishen, & Cozzolino, 2015; De Bei et al., 2011; Levasseur-garcia, Malaurie, & Mailhac, 2016; Martin et al., 2015; Nogales-Bueno et al., 2017; dos Santos, Claudia, Páscoa, Lopes, & Lopes, 2017; dos Santos et al., 2017; Tian et al., 2017). In addition, spectroscopic data is processed using chemometric methods that provide mathematical tools for processing

* Corresponding author.
*E-mail address:* jyanez@udec.cl (J. Yáñez).

https://doi.org/10.1016/j.foodchem.2020.127164
Received 17 September 2019; Received in revised form 14 May 2020; Accepted 25 May 2020
Available online 27 May 2020
0308-8146/ © 2020 Elsevier Ltd. All rights reserved.

large amounts of information and reveal correlations between spectroscopic patterns between samples. Classification methods such as soft independent modeling of class analogy (SIMCA) and partial least square discriminant analysis (PLS-DA) can be applied for sample identification. Classification and calibration models are developed and evaluated by appropriated validation methods (Otto, 2016).

The aim of this work is to differentiate and classify grapevine genotypes applying simple and direct screening techniques based on ATR-FTIR spectroscopy. These can be used as an alternative method to conventional genetic analysis used for variety identification, currently based on Simple Sequence Repeats (SSR) markers (This et al., 2004). This work is a proof of concept that demonstrates the tremendous potential of vibrational spectroscopy for direct and in-field classification technique of grapevine cultivars.

## 2. Materials and methods

### 2.1. Grapevine (Vitis vinifera L.) genotypes

The Región de Bío-Bío in Central-South Chile is a traditional wine production area where a number of wine grape varieties are used in small wineries. Therefore, there is a need for genotypes identity verification for these small wine producers in isolated regions, where conventional methods are scarcely accessible or not available at all.

Six grapevine genotypes (varieties and clones), among the most commonly planted in this region, were considered for this study. Adult plants were healthy, free of fungal diseases and of regular vigor. All samples were collected from the same vineyard of around 100 ha, where all grapevine genotypes are cultivated (36° 36′29.4″ S, 72° 39′42.1″ W, Altos de Guarilihue, Región de Bío-Bío, Chile) in December 2015. The genotypes analyzed were: Cinsault, Gewurztraminer (clone 643), Moscatel de Alejandría, País and Pinot Noir (French clone 777 and the local clone 'Valdivieso'). From these, Moscatel de Alejandría and País are patrimonial varieties with more than 300 years of local presence (Cantillana, 2017; Lacoste, Soto, & Pszczolkowski, 2015; Milla Tapia et al., 2007); Pinot Noir and Gewurztraminer were imported from France before the peak of the phylloxera crisis (second half of XIX$^{th}$ century), and Cinsault was introduced in the beginning of the XX$^{th}$ century (Cantillana, 2017).

Leaf samples were collected directly from 20 to 30 plants of each genotype, stored in plastic bags and then transported to the lab in refrigerated containers in the same day. Each leaf was cleaned with tissue paper (Kimberly-Clark Worldwide, Inc., Neenah, WI, USA) before the spectroscopic analysis to remove dirt.

Representative leaf samples of each variety and clone were separated in order to extract DNA for varietal verification using SSR markers, as has been described in the literature (Narváez, Castro, Valenzuela, & Hinrichsen, 2001). The result of these analyses is summarized in Table S1; the allelic patterns of each sample perfectly fitted Instituto de Investigaciones Agropecuarias (INIA) and Vitis International Variety Catalogue (VIVC) databases for each variety considered in this study.

### 2.2. Instrumentation

Sample spectra were record using a Ge crystal ATR sampling accessory mounted in a Nicolet iS10 FT-IR spectrometer (Thermo Fisher Scientific, Madison, WI, USA). The spectrum of each sample was collected in the range of 4000 to 690 cm$^{-1}$ with 25 scans at 8 cm$^{-1}$ resolution and measured after one reflection between the crystal and the sample. The spectrum of air (background) was collected before analyses of each sample. The spectral collection was operated using Omnic Specta software suite version 7.3 (Thermo Fisher Scientific, Madison, WI, USA).

### 2.3. Spectra pre-treatment

The spectra were saved as data files in .CSV format. First, the region at 2400–2280 cm$^{-1}$ corresponding to absorption band of atmospheric $CO_2$ was removed from data. Later, standard normal variate (SNV), first derivative and smoothing were tested as transformation techniques. SNV was used to remove multiplicative interferences of scatter in the spectral signal. First derivative and smoothing (Savitzky-Golay polynomial filter 25 points per window) were applied for minimizing baseline deviation and spectral noise, respectively (Gemperline, 2006; Otto, 2016; Zeaiter, Roger, & Bellon-Maurel, 2005). Finally, spectra were preprocessed by mean centering (MC) before calibration and classification.

### 2.4. Data sets

The abbreviations CT, GW, MA, PA, PN and PNV correspond to Cinsault, Gewurztraminer (clone 643), Moscatel de Alejandría, País, Pinot Noir (French clone 777) and Pinot Noir (local clone 'Valdivieso'), respectively. A total of 174 grapevine leaf samples of CT (29), GW (27), MA (30), PA (28), PN (30), PNV (30) were analyzed. Spectral data were divided by random selection in two sets (80:20): training set (138 samples) for obtaining the classification models and test set (36 samples) for external validation purpose. Each sample was measured twice and average spectra were considered for further analysis.

### 2.5. Chemometric analysis

For classification models, SIMCA and PLS-DA algorithms were utilized. Full spectra (4000 to 690 cm$^{-1}$) were considered for the models' development. Both models were applied over the training set. The designed classification was assessed by leave one out (LOO) cross-validation to select the optimal number of factors (Otto, 2016; Voet & Coenegracht, 1987).

The models were evaluated and compared using different parameters viz. error rate (ER) which is a measure of the number (percentage) of incorrectly classified samples; sensitivity (Sn) describes the model ability to correctly recognize objects belonging to the $g^{th}$ class; specificity (Sp) characterizes the capability of the $g^{th}$ class to reject objects of all other classes; and the precision (Pre) measures the ability of a classification model to exclude objects of other classes in the considered class (Wehrens, 2011). All these parameters were calculated on the training and test sets of samples. Root mean square error of calibration (RMSEC) and cross-validation (RMSECV) were reported for PLS-DA model. Chemometric analysis was performed by Pirouette 4.5 software (Infometrix Inc., Woodinville, WA, USA).

## 3. Results and discussion

### 3.1. ATR-FTIR and genetic analysis

Grapevine leaves are complex samples which include assemblages of organic compounds. The infrared-absorption bands can be attributed to biological polymeric and condensed compounds, such as pectin, lignin, polysaccharides, lipids, terpenoids and flavonoids (da Luz, 2006; Heredia-Guerrero et al., 2014; Nogales-Bueno et al., 2017). Fig. 1 shows the average spectra for the six classes of grapevine genotypes belonging to the training set. Previous studies determined the spectral zones; the main functional groups presented for these genotypes samples are described in Table 1. It is remarkable that this approach was able to differentiate the two clones of Pinot noir (clon 777 and clone Valdivieso), which is, up to now, not possible to afford by means of microsatellite/SSR or other available molecular analyses, with very specific exceptions (Moncada, Pelsy, Merdinoglu, & Hinrichsen, 2006; Pelsy et al., 2010). Even using whole genome sequencing for clones of particular varieties, differentiation of grapevine clones has been quite elusive (Vondras

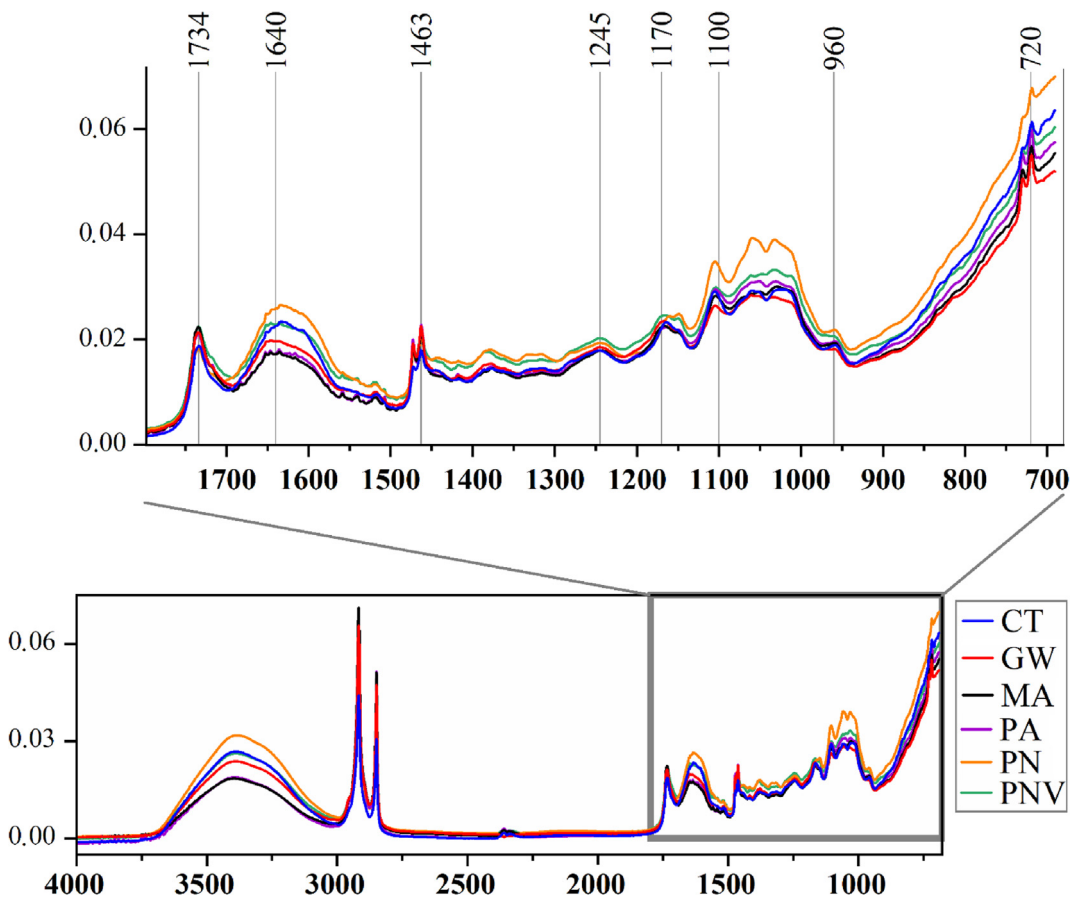


**Fig. 1.** ATR-FTIR average spectra for each grapevine *(Vitis vinifera L.)* genotypes.

**Table 1**
Peak assignment for the ATR-FTIR spectra shown in Fig. 1.

| Assignment | Absorption bands ($cm^{-1}$) | Component related |
|---|---|---|
| $\nu$(O—H) | 3400 | Polysaccharides, lignins |
| $\nu_a$, $\nu_s$ ($CH_2$) | 2916, 2850 | Lignins, lipids, terpenoids |
| $\nu$(C=O) | 1734 | Polyesters, pectins, lignins |
| $\nu_a(COO^-)$ and $\nu$ $(C{=}C)_{aromatic}$ | 1640 | Pectins, phenolics |
| $\delta_a(CH_2)$ | 1472,1463 | Lipids |
| $\nu$(C—O), $\delta$(O—H) | 1245 | Polysaccharides, pectins, lignins |
| $\nu$(C—O—C) | 1170 | Polysaccharides, pectins |
| $\nu$(C—O) and $\nu$(C—C) | 1100–1023 | Polysaccharides, pectins |
| $\delta$ $(C{-}H)_{out\ of\ plane}$ | 960, 730 | Phenolics |
| $\delta$ $(CH_2)_{rocking}$ | 720 | Phenolics |

Assignment: $\nu$ (stretching), $\delta$ (bending); s (symmetric), a (asymmetric)

et al., 2019). The evaluation of larger sets of clones belonging to different varieties would be an evident target for future applications of this IR platform.

### 3.2. Supervised analysis

SIMCA and PLS-DA supervised classification methods were chosen for building differentiation models based on training datasets, which were obtained from spectral data of grapevine leaves. The classification models were externally validated to achieve prediction success (correct class assignment) using test data, which were omitted during the training process. Table S2 details the parameters optimized for the best classification performance of the models.

After testing different spectral transformations, SNV and first derivative/smooth showed the best results for pretreatment of ATR-FTIR raw data. Along with the application of the SIMCA model, both transformed techniques showed a successful classification. SNV was selected because higher interclass distance values were obtained (Table S3) in spite of a slightly greater number of components being required. Higher interclass distance is important to produce reliable predictions. Interclass distances show how far the two classes are separated, which is calculated as the geometric distance between individual clusters for a chosen confidence interval (95% in this case).

SIMCA requires the development of individual principal component (PC) based on models for each class. The number of PCs selected in each PCA model was optimized using the cross-validation procedure. Eight PCs were selected for SIMCA classification. The cumulative variance for each class were CT (99.3%), GW (99.6%), MA (99.5%), PA (99.1%), PN (99.3%) and PNV (99.6%). Fig. 2 shows the discriminating ability of the variables for the set of grapevine genotypes. The discrimination power plot shows which wavenumbers mostly contribute to the separation of grapevine genotypes. The higher the value, the more influential the wavenumber is in the discrimination between class (Infometrix, 2011). It can be seen that variables corresponding at bands C—H stretching ($\nu_a$2956 and $\nu_s$2873 $cm^{-1}$) and C—H deformation (1400 $cm^{-1}$) provide high discriminating power; these regions can be related with component of polysaccharides, lignins and pectins. Likewise, C=O stretching at ~1748 $cm^{-1}$ was important to discriminate. Around ~1696 $cm^{-1}$ were noted the influence of the bands $COO^-$ asymmetric stretching and C=C stretching associate at pectins and phenolics. Finally, absorptions of C—O—C glycosidic at 1113 $cm^{-1}$ showed also significant discriminating power.

Classification methods as SIMCA and PLS-DA not only determine whether a sample belongs to one of the chosen classes but can also indicate if it does not belong to any of them (Gemperline, 2006). The

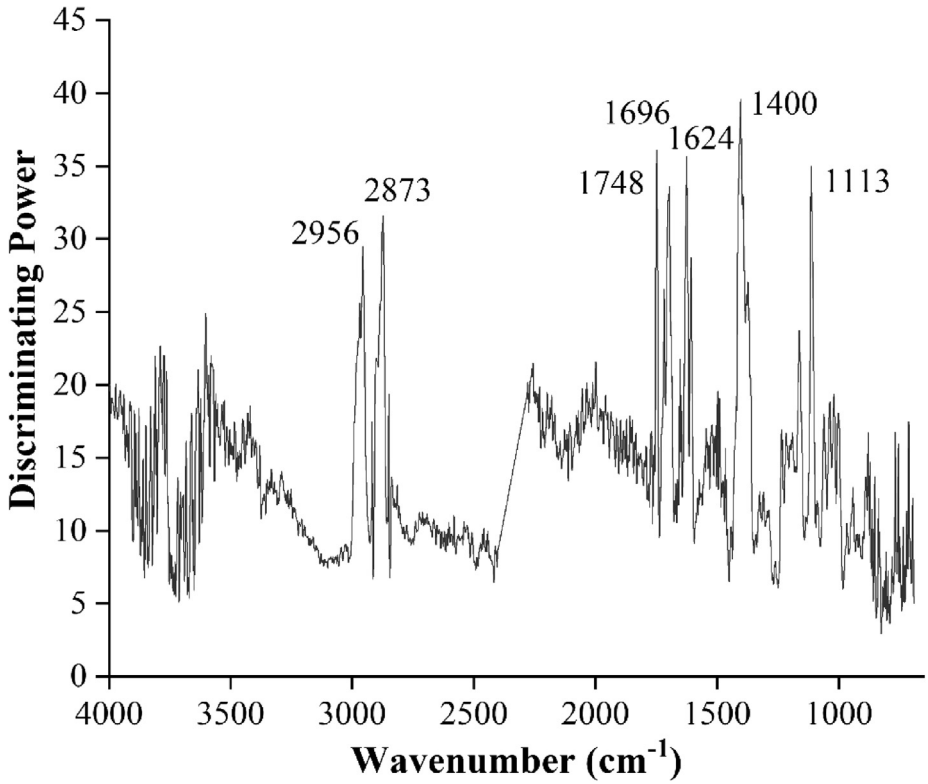


**Fig. 2.** Discriminating power in the SIMCA model built.

**Table 2**
SIMCA and (b) PLS-DA confusion matrix for the grapevine (*Vitis vinifera* L.) genotypes.

| **(a) SIMCA model** | | Error rate | Predicted class | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | CT | GW | MA | PA | PN | PNV | No match |
| Real class | CT | *Training* (0) | **23** | | | | | | |
| | | *Test* (0.17) | **5** | | | | 1 | | |
| | GW | *Training* (0) | | **21** | | | | | |
| | | *Test* (0) | | **6** | | | | | |
| | MA | *Training* (0) | | | **24** | | | | |
| | | *Test* (0) | | | **6** | | | | |
| | PA | *Training* (0) | | | | **22** | | | |
| | | *Test* (0.33) | | 1 | | **4** | | 1 | |
| | PN | *Training* (0) | | | | | **24** | | |
| | | *Test* (0) | | | | | **6** | | |
| | PNV | *Training* (0) | | | | | | **24** | |
| | | *Test* (0) | | | | | | **6** | |

| **(b) PLS-DA model** | | Error rate | Predicted class | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | CT | GW | MA | PA | PN | PNV | No match |
| Real class | CT | *Training* (0.09) | **21** | | | | | | 2 |
| | | *Test* (0.33) | **4** | | | | | | 2 |
| | GW | *Training* (0.24) | | **16** | | | | 1 | 4 |
| | | *Test* (0) | | **6** | | | | | |
| | MA | *Training* (0.12) | | 1 | **21** | | | | 2 |
| | | *Test* (0) | | | **6** | | | | |
| | PA | *Training* (0.33) | | | | **17** | 1 | | 4 |
| | | *Test* (0) | | | | **6** | | | |
| | PN | *Training* (0.17) | | | | | **20** | | 4 |
| | | *Test* (0.17) | | | | | **5** | | 1 |
| | PNV | *Training* (0) | | | | | | **24** | |
| | | *Test* (0) | | | | | | **6** | |

**Table 3**
The quality of the built models evaluated in terms of the prediction sensitivity (Sn, %), specificity (Sp, %) and precision (Pre, %).

| | | SIMCA model | | | PLS-DA model | | |
|---|---|---|---|---|---|---|---|
| | | Sn | Sp | Pre | Sn | Sp | Pre |
| CT | *Training* | 100 | 100 | 100 | 100 | 100 | 100 |
| | *Test* | 83 | 100 | 100 | 100 | 100 | 100 |
| GW | *Training* | 100 | 100 | 100 | 94 | 99 | 94 |
| | *Test* | 100 | 96 | 86 | 100 | 100 | 100 |
| MA | *Training* | 100 | 100 | 100 | 96 | 100 | 100 |
| | *Test* | 100 | 100 | 100 | 100 | 100 | 100 |
| PA | *Training* | 100 | 100 | 100 | 94 | 100 | 100 |
| | *Test* | 67 | 100 | 100 | 100 | 100 | 100 |
| PN | *Training* | 100 | 100 | 100 | 100 | 99 | 95 |
| | *Test* | 100 | 96 | 86 | 100 | 100 | 100 |
| PNV | *Training* | 100 | 100 | 100 | 100 | 99 | 96 |
| | *Test* | 100 | 96 | 86 | 100 | 100 | 100 |

classification results are shown in a confusion matrix. The confusion matrix indicates two possible outcomes including *(i)* the sample was classified into some of the six classes (CT, GW, MA, PA, PN and PNV) or *(ii)* the sample was not assigned to any class. The results of SIMCA classification by internal and external validation are shown in Table 2a. For LOO cross validation all training sets were correctly assigned with 100% precision, sensitivity and specificity. Whereas the error rates for CT and PA were 0.17 and 0.33, respectively. The values of sensitivity for CT (Sn: 83%), PA (Sn: 67%) and precision for GW (Pre: 86%), PN (Pre: 86%), PNV (Pre: 86%) were acceptable (Table 3). Fig. 3a illustrates the SIMCA classification for external validation of the unknown grapevine genotypes testing set. Y-axis values correspond to predicted "grapevine genotype". A value of 1, 2, 3, 4, 5 and 6 indicates that a sample belongs to CT, GW, MA, PA, PN, and PNV respectively. While 0 corresponds to "unmatched" samples. It can be noticed that only few CT and PA testing samples were wrongly assigned.

When PLS-DA model was applied, first derivative/smoothing were

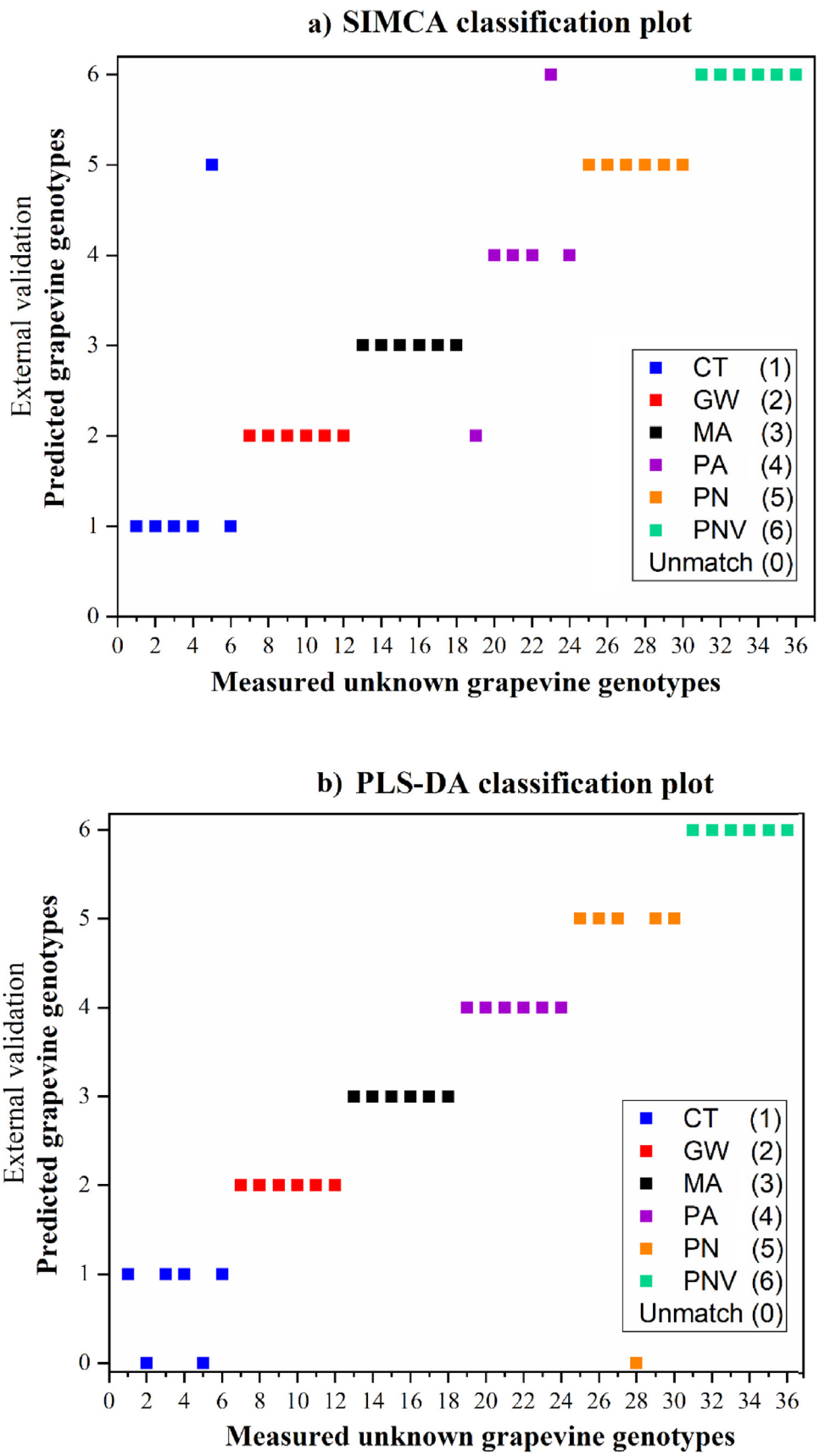


**Fig. 3.** (a) SIMCA and (b) PLS-DA classification plot for external validation from the unknown grapevine genotypes (testing set).

selected as preprocessing technique since they showed the best prediction rate. The latent variables used and the cumulative variance obtained for each class are shown in Table S2. Regression vectors obtained from PLS-DA analysis of the six grapevine genotypes are shown in Fig. S1. It can be noted that for Gewurztraminer and País the C—H stretching at 3000–2800 $cm^{-1}$ were relevant bands for classification purposes. These are related with components of polysaccharides, lignins and pectins. Also C═O stretching at ~1747 $cm^{-1}$ was observed in CT and MA spectra. In particular, the band at ~1637 $cm^{-1}$ belonging $COO^-$ asymmetric stretching and C═C stretching was present in CT and can be assigned to pectins and phenolics components. Specifically, a band at ~1463 $cm^{-1}$ corresponding at C—H deformation was significant for GW and PA. All classes excepting PNV showed high contribution by the region at 1100–1000 $cm^{-1}$ associated to the absorption of the C—O stretching from polysaccharides and pectins. For all classes, the influence of the bands at ~820–700 $cm^{-1}$, corresponding to C—H out of plane deformation and rocking, was noted.

Table 3 shows the results of PLS-DA classification for training and test set. The error rate was 0.14 for the training set. The values of sensitivity for GW (Sn: 94%), MA (Sn: 96%), PA (Sn: 94%) and precision for GW (Pre: 94%), PN (Pre: 95%), PNV (Pre: 96%) were acceptable (Table 3). Additionally, Table S5 summarizes the figure of merit of PLS-DA classification model for the grapevine genotypes. It is reported according to the number of components in each brand. The error rates for the test set for CT and PN were 0.33 and 0.17, respectively. The statistics (precision, sensitivity and specificity) for test set from the six classes are shown in Table 2b. Besides, Fig. 3b shows the PLS-DA classification for external validation from the unknown grapevine genotypes testing set. It can be seen that only CT and PN samples were unmatched.

Both PLS and SIMCA are well established and easily manageable methods. When compared, the PLS-DA algorithm yields better classification rates than SIMCA. Nevertheless, the interpretation of PLS-DA tends to be more complicated when the number of classes rises. In this work, both methods showed acceptable classification rates for classes belonging to the calibration set. However, SIMCA showed better results regarding the prediction rate of the training set in cross validation. In this sense, a crucial comparison point is the ability of the models to obtain good results for the training set. Taking this feature into account, the SIMCA model could be considered as the best option for the classification of our samples, nevertheless, the capacity of classification of both models should be reassessed when new samples and genotypes of different vineyards are included in future works.

In comparison with our results, other studies of grapevine classification by vibrational spectroscopy have been reported. NIR hyperspectral image of three grapevine varieties (Tempranillo, Grenache and Cabernet Sauvignon) coupled with PLS analysis were evaluated. Similarly to our results, the model proposed reached a 92% satisfactory classification (Diago, Fernandes, Millan, Tardaguila, & Melo-pinto, 2013). Furthermore, when classification of four clones of Cabernet Sauvignon was studied by NIR hyperspectral image and PLS analysis, a correct prediction rate of 97.8% was achieved (Fernandes, Melo-Pinto, Millan, Tardaguila, & Diago, 2014). The results reported in this work using ATR-FTIR show the predictive ability in order to get fast, reliable and accurate classification of new samples. The proposed methodology requires to increase the number of *Vitis* spp. genotypes (possibly applicable in rootstock hybrids identification) in order to attest the versatility and predictive ability of this technique. Besides, further studies could be conducted exploring additional prediction methods. Artificial Neural Networks (ANNs), Support Vector Machine (SVM), and Random Forest (RF) are computational methods widely used for food authentication (Esteki, Regueiro, & Simal-Gándara, 2019; Esteki, Shahsavari, & Simal-Gandara, 2018, 2020, 2019; Gao et al., 2019; Gonzalez-Fernandez et al., 2019), and winemaking (Astray et al., 2019; Moldes, Mejuto, Rial-Otero, & Simal-Gandara, 2017; Wu et al., 2019). Future opportunities must include the assessment of food quality, safety, geographical origin and authenticity (e.g. grapes and wines). In fact, blockchain for food traceability analysis is being used for tracking and authenticating the food supply chain (Galvez, Mejuto, & Simal-Gandara, 2018).

## 4. Conclusions

The ability of ATR-FTIR spectroscopy coupled with chemometric analysis was successfully evaluated as a simple and fast screening method for the differentiation and identification of grapevine genotypes. The analysis was developed and validated with leaves of six genotypes. Differentiation and classification of these genotypes was achieved by using both SIMCA and PLS-DA models. All spectra were classified correctly during the LOO-cross validation applying SIMCA, while the error rate for PLS-DA model was 0.14. Parameters such as sensitivity, specificity and precision were evaluated. The ability of the models was tested by correct assignment of external data. The classification error rates for the external validation of SIMCA and PLS-DA models built were the same (0.08).

The main advantage of the proposed method is its ability to provide rapid and non-destructive analysis. In particular, its environmental friendliness should be highlighted, as it provides an alternative to other analytical platforms for grapevine varietal identification, such as the classical ampelography or the nowadays most popular SSR-based molecular screening. It is worth to mention that a larger number of grapevine genotypes must be included through more experimentation before proposing spectral-based methodologies as the recommended method for grapevine fingerprinting. Additionally, the reproducibility of the IR-based methods presented here must be evaluated on leaf samples of the same genotype (variety, clone) obtained from different environments and with different managements and trellis systems as well as on samples having different physiological conditions, sanitary conditions and developmental stages, in order to fully establish its confidence, technical strengths and weaknesses.

## CRediT authorship contribution statement

**Ángela Álvarez:** Investigation, Conceptualization, Methodology, Software, Validation, Writing - original draft. **Jorge Yáñez:** Conceptualization, Methodology, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Yamil Neira:** Resources, Writing - review & editing. **Rosario Castillo-Felices:** Formal analysis. **Patricio Hinrichsen:** Writing - review & editing, Resources.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

The authors thank the National Commission for Scientific and Technological Research (CONICYT) project FONDECYT N° 1151296, FONDECYT N° 1191572 and PIA ANILLO ACM 170008. Also appreciate the CONICYT doctoral scholarship of Ángela Álvarez N° 21161209.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2020.127164.

## References


Astray, G., Mejuto, J. C., Martínez-Martínez, V., Nevares, I., Alamo-Sanza, M., & Simal-Gandara, J. (2019). Prediction models to control aging time in red wine. *Molecules, 24*(5), 1–11. https://doi.org/10.3390/molecules24050826.

Borrego, J., De Andrés, M. T., Gómez, J. L., & Ibáñez, J. (2002). Genetic study of Malvasia and Torrontes groups through molecular markers. *American Journal of Enology and Viticulture, 53*(2), 125–130.

Cantillana, J. C. (2017). Persistencia de cepas tradicionales en el secano interior del Maule sur: Rescate de elementos patrimoniales. *Trilogía, 29*(40), 60–75.

Carrascon, V., Ontañón, I., Bueno, M., & Ferreira, V. (2017). Gas chromatography-mass spectrometry strategies for the accurate and sensitive speciation of sulfur dioxide in wine. *Journal of Chromatography A, 1504*, 27–34. https://doi.org/10.1016/j.chroma.2017.05.012.

Cozzolino, D., Cynkar, W. U., Dambergs, R. G., Shah, N., & Smith, P. (2013). In situ measurement of soil chemical composition by near-infrared spectroscopy: A tool toward sustainable vineyard management in situ measurement of soil chemical composition by near-infrared spectroscopy. *Communications in Soil Science and Plant Analysis, 44*, 1610–1619. https://doi.org/10.1080/00103624.2013.768263.

da Luz, R. B. (2006). Attenuated total reflectance spectroscopy of plant leaves: A tool for ecological and botanical studies. *New Phytologist, 172*, 305–318. https://doi.org/10.1111/j.1469-8137.2006.01823.x.

Dalipi, R., Marguí, E., Borgese, L., Bilo, F., & Depero, L. E. (2016). Analytical performance of benchtop total reflection X-ray fluorescence instrumentation for multielemental analysis of wine samples. *Spectrochimica Acta – Part B Atomic Spectroscopy, 120*, 37–43. https://doi.org/10.1016/j.sab.2016.04.001.

Dambergs, R., Gishen, M., & Cozzolino, D. (2015). A review of the state of the art, limitations, and perspectives of infrared spectroscopy for the analysis of wine grapes, must, and grapevine tissue. *Applied Spectroscopy Reviews, 50*, 261–278. https://doi.org/10.1080/05704928.2014.966380.

De Bei, R., Cozzolino, D., Sullivan, W., Cynkar, W., Fuentes, S., Dambergs, R., ... Tyerman, S. (2011). Non-destructive measurement of grapevine water potential using near infrared spectroscopy. *Australian Journal of Grape and Wine Research, 17*, 62–71. https://doi.org/10.1111/j.1755-0238.2010.00117.x.

Diago, M. P., Fernandes, A. M., Millan, B., Tardaguila, J., & Melo-pinto, P. (2013). Identification of grapevine varieties using leaf spectroscopy and partial least squares. *Computers and Electronics in Agriculture, 99*, 7–13. https://doi.org/10.1016/j.compag.2013.08.021.

Eftekhari, M., Yadollahi, A., Ford, C. M., Shojaeiyan, A., Ayyari, M., & Hokmabadi, H. (2017). Chemodiversity evaluation of grape (Vitis vinifera) vegetative parts during summer and early fall. *Industrial Crops and Products, 108*, 267–277. https://doi.org/10.1016/j.indcrop.2017.05.057.

Esteki, M., Regueiro, J., & Simal-Gándara, J. (2019). Tackling fraudsters with global strategies to expose fraud in the food chain. *Comprehensive Reviews in Food Science and Food Safety, 18*(2), 425–440. https://doi.org/10.1111/1541-4337.12419.

Esteki, M., Shahsavari, Z., & Simal-Gandara, J. (2018). Use of spectroscopic methods in combination with linear discriminant analysis for authentication of food products. *Food Control, 91*, 100–112. https://doi.org/10.1016/j.foodcont.2018.03.031.

Esteki, M., Shahsavari, Z., & Simal-Gandara, J. (2019). Food identification by high performance liquid chromatography fingerprinting and mathematical processing. *Food Research International, 122*(March), 303–317. https://doi.org/10.1016/j.foodres.2019.04.025.

Esteki, M., Shahsavari, Z., & Simal-Gandara, J. (2020). Gas chromatographic fingerprinting coupled to chemometrics for food authentication. *Food Reviews International, 36*(4), 384–427. https://doi.org/10.1080/87559129.2019.1649691.

Fernandes, A. M., Melo-Pinto, P., Millan, B., Tardaguila, J., & Diago, M. (2014). Automatic discrimination of grapevine (Vitis vinifera L .) clones using leaf hyperspectral imaging and partial least squares. *Journal of Agricultural Science, 153*(3), 1–11. https://doi.org/10.1017/S0021859614000252.

Galet, P. (1979). A practical ampelography. Comstock Pub. Associates.

Galvez, J. F., Mejuto, J. C., & Simal-Gandara, J. (2018). Future challenges on the use of blockchain for food traceability analysis. *TrAC - Trends in Analytical Chemistry, 107*, 222–232. https://doi.org/10.1016/j.trac.2018.08.011.

Gao, B., Holroyd, S. E., Moore, J. C., Laurvick, K., Gendel, S. M., & Xie, Z. (2019). Opportunities and challenges using non-targeted methods for food fraud detection. *Journal of Agricultural and Food Chemistry, 67*(31), 8425–8430. https://doi.org/10.1021/acs.jafc.9b03085.

Gemperline, P. (2006). *Practical Guide to Chemometrics* (2nd ed.). Boca Raton: CRC Taylor & Francis https://doi.org/10.2307/1269627.

Gonzalez-Fernandez, I., Iglesias-Otero, M. A., Esteki, M., Moldes, O. A., Mejuto, J. C., & Simal-Gandara, J. (2019). A critical review on the use of artificial neural networks in olive oil production, characterization and authentication. *Critical Reviews in Food Science and Nutrition, 59*(12), 1913–1926. https://doi.org/10.1080/10408398.2018.1433628.

Heredia-Guerrero, J. A., Benítez, J. J., Domínguez, E., Bayer, I. S., Cingolani, R., Athanassiou, A., & Heredia, A. (2014). Infrared and Raman spectroscopic features of plant cuticles : A review. *Frontier in Plant Science, 5*(305), 1–14. https://doi.org/10.3389/fpls.2014.00305.

Infometrix, I. (2011). Pirouette Multivariate Data Analysis Software. Bothell, WA. Retrieved from http://www.infometrix.com/.

Kilmartin, P. A. (2016). Electrochemistry applied to the analysis of wine: A mini-review. *Electrochemistry Communications, 67*, 39–42. https://doi.org/10.1016/j.elecom.2016.03.011.

Lacoste, P., Soto, N., & Pszczolkowski, P. (2015). Moscatel de Alejandría en Chile y Argentina: Origen y relación con el pisco. *IDESIA (Chile), 33*(3), 79–86.

Levasseur-garcia, C., Malaurie, H., & Mailhac, N. (2016). An infrared diagnostic system to detect causal agents of grapevine trunk diseases. *Journal of Microbiological Methods, 131*, 1–6. https://doi.org/10.1016/j.mimet.2016.09.022.

Martin, C., Bruneel, J., Guyon, F., Médina, B., Jourdes, M., Teissedre, P., & Guillaume, F. (2015). Raman spectroscopy of white wines. *Food Chemistry, 181*, 235–240. https://doi.org/10.1016/j.foodchem.2015.02.076.

Mierczynska-Vasilev, A., & Smith, P. A. (2016). Surface modification influencing adsorption of red wine constituents: The role of functional groups. *Applied Surface Science, 386*, 14–23. https://doi.org/10.1016/j.apsusc.2016.06.016.

Milla Tapia, A., Cabezas, J. A., Cabello, F., Lacombe, T., Martínez-Zapater, J. M., Hinrichsen, P., & Cervera, M. T. (2007). Determining the spanish origin of representative ancient american grapevine varieties. *American Journal of Enology and Viticulture, 58*(2), 242–251.

Milosevic, M. (2004). Internal reflection and ATR spectroscopy. *Applied Spectroscopy Reviews, 39*(3), 365–384. https://doi.org/10.1081/ASR-200030195.

Moldes, O. A., Mejuto, J. C., Rial-Otero, R., & Simal-Gandara, J. (2017). A critical review on the applications of artificial neural networks in winemaking technology. *Critical Reviews in Food Science and Nutrition, 57*(13), 2896–2908. https://doi.org/10.1080/10408398.2015.1078277.

Mollo, A., Sixto, A., Falchi, L., Medina, M., & Knochen, M. (2017). Zinc determination in Tannat wine by direct injection onto graphite tube: Electrothermal AAS as an alternative to flame AAS. *Microchemical Journal, 135*, 239–244. https://doi.org/10.1016/j.microc.2017.09.008.

Moncada, X., Pelsy, F., Merdinoglu, D., & Hinrichsen, P. (2006). Genetic diversity and geographical dispersal in grapevine clones revealed by microsatellite markers. *Genome, 49*, 1459–1472. https://doi.org/10.1139/g06-102.

Mutic, J., Manojlovic, D., Kovacevic, R., Trifunovic, J., Amaizah, N. R., & Ignjatovic, L. (2011). Feasibility of the internal standardization in direct determination of arsenic in wine samples by inductively coupled plasma atomic emission spectrometry. *Microchemical Journal, 98*, 11–14. https://doi.org/10.1016/j.microc.2010.10.004.

Narváez, H. C., Castro, P. H. M., Valenzuela, B. J., & Hinrichsen, R. P. (2001). Patrones genéticos de los cultivares de vides de vinificación más comúnmente usados en Chile basados en marcadores de microsatélites. *Agricultura Técnica (Chile), 61*(3), 249–261. https://doi.org/10.4067/S0365-28072001000300001.

Nogales-Bueno, J., Baca-Bocanegra, B., Rooney, A., Hernández-Hierro, J. M., Byrne, H. J., & Heredia, F. J. (2017). Study of phenolic extractability in grape seeds by means of ATR-FTIR and Raman spectroscopy. *Food Chemistry, 232*, 602–609. https://doi.org/10.1016/j.foodchem.2017.04.049.

Otto, M. (2016). *Chemometrics: Statistics and Computer Application in Analytical Chemistry. Chemometrics* (3rd ed.). Weinheim: Wiley-VCH https://doi.org/10.1002/9783527699377.

Ozbek, N., & Akman, S. (2015). Determination of boron in Turkish wines by microwave plasma atomic emission spectrometry. *LWT – Food Science and Technology, 61*, 532–535. https://doi.org/10.1016/j.lwt.2014.11.047.

Pelsy, F., Hocquigny, S., Moncada, X., Barbeau, G., Forget, D., Hinrichsen, P., & Merdinoglu, D. (2010). An extensive study of the genetic diversity within seven French wine grape variety collections. *Theoretical and Applied Genetics, 120*, 1219–1231. https://doi.org/10.1007/s00122-009-1250-8.

Schmidtke, L. M., Smith, J. P., Müller, M. C., & Holzapfel, B. P. (2012). Rapid monitoring of grapevine reserves using ATR-FT-IR and chemometrics. *Analytica Chimica Acta, 732*, 16–25. https://doi.org/10.1016/j.aca.2011.10.055.

Shah, N., Cynkar, W., Smith, P., & Cozzolino, D. (2010). Use of attenuated total reflectance midinfrared for rapid and real-time analysis of compositional parameters in commercial white grape juice. *Journal of Agricultural and Food Chemistry, 58*, 3279–3283. https://doi.org/10.1021/jf100420z.

Shi, J., He, M., Cao, J., Wang, H., Ding, J., Jiao, Y., ... Wang, Y. (2014). The comparative analysis of the potential relationship between resveratrol and stilbene synthase gene family in the development stages of grapes (Vitis quinquangularis and Vitis vinifera). *Plant Physiology and Biochemistry, 74*, 24–32. https://doi.org/10.1016/j.plaphy.2013.10.021.

dos Santos, Teixeira, Claudia, A., Páscoa, R. N. M. J., Lopes, A., & Lopes, João A. (2017). A review on the application of vibrational spectroscopy in the wine industry : From soil to bottle. *Trends in Analytical Chemistry, 88*, 100–118. https://doi.org/10.1016/j.trac.2016.12.012.

Teixeira dos Santos, Cláudia A, Páscoa, R. N. M. J., Cruz Sarraguça, M., Porto, P. A. L. S., Cerdeira, A. L., González Sáiz, J. M., … Lopes, J. A. (2017). Merging vibrational spectroscopic data for wine classi fi cation according to the geographic origin. Food Research International Journal, 102, 504–510. https://doi.org/10.1016/j.foodres.2017.09.018.

This, P., Jung, A., Boccacci, P., Borrego, J., Botta, R., Costantini, L., ... Maul, E. (2004). Development of a standard set of microsatellite reference alleles for identification of grape cultivars. *Theoretical and Applied Genetics, 109*, 1448–1458. https://doi.org/10.1007/s00122-004-1760-3.

Tian, Y., Yan, C., Zhang, T., Tang, H., Li, H., Yu, J., ... Yu, J. (2017). Classification of wines according to their production regions with the contained trace elements using laser-induced breakdown spectroscopy ☆. *Spectrochimica Acta Part B: Atomic Spectroscopy, 135*, 91–101. https://doi.org/10.1016/j.sab.2017.07.003.

Voet, H. Van Der, & Coenegracht, P. M. J. (1987). New probabilistic versions of the SIMCA and classy classification methods. Part 1. Theorical Description. *Analytica Chimica Acta, 192*, 63–75. https://doi.org/10.1016/S0003-2670(00)85688-3.

Vondras, A. M., Minio, A., Blanco-Ulate, B., Figueroa-Balderas, R., Penn, M. A., Yongfeng, Z., … Cantu, D. (2019). The genomic diversification of clonally propagated grapevines. BioRixv. https://doi.org/http://dx.doi.org/10.1101/585869.

Wehrens, R. (2011). Chemometrics with R: Multivariate Data Analysis in the Natural Sciences and Life Sciences (1st ed.). New York: Springer Science & Business Media. https://doi.org/10.1007/978-0-387-78171-6.

Wu, H., Tian, L., Chen, B., Jin, B., Tian, B., Xie, L., ... Lin, G. (2019). Verification of imported red wine origin into China using multi isotope and elemental analyses. *Food Chemistry, 301*(July), https://doi.org/10.1016/j.foodchem.2019.125137.

Zeaiter, M., Roger, J.-M., & Bellon-Maurel, V. (2005). Robustness of models developed by multivariate calibration. Part II : The influence of pre-processing methods. *Trends in Analytical Chemistry, 24*(5), 437–445. https://doi.org/10.1016/j.trac.2004.11.023.