



# NMR spectroscopy evaluation of direct relationship between soils and molecular composition of red wines from Aglianico grapes

Pierluigi Mazzei<sup>a,b</sup>, Nicola Francesca<sup>c</sup>, Giancarlo Moschetti<sup>c</sup>, Alessandro Piccolo<sup>a,\*</sup>

<sup>a</sup> Centro Interdipartimentale per la Risonanza Magnetica Nucleare (CERMANU), Università 'di Napoli Federico II, Via Università' 100, 80055 Portici, Italy

<sup>b</sup> Istituto di Metodologie Chimiche, CNR, Via Salaria Km 29.300, 00016 Monterotondo Stazione, Italy

<sup>c</sup> Dipartimento di Scienze Entomologiche, Fitopatologiche, Microbiologiche Agrarie e Zootecniche, Università di Palermo, Viale delle Scienze, 90128 Palermo, Italy

## ARTICLE INFO

### Article history:

Received 11 March 2010

Received in revised form 25 May 2010

Accepted 2 June 2010

Available online 9 June 2010

### Keywords:

Nuclear Magnetic Resonance

Aglianico red wines

Multivariate statistical analysis

Terroir

## ABSTRACT

<sup>1</sup>H NMR spectroscopy was employed to investigate the molecular quality of Aglianico red wines from the Campania region of Italy. The wines were obtained from three different Aglianico vineyards characterized by different microclimatic and pedological properties. In order to reach an objective evaluation of “terroir” influence on wine quality, grapes were subjected to the same winemaking procedures. The careful subtraction of water and ethanol signals from NMR spectra allowed to statistically recognize the metabolites to be employed in multivariate statistical methods: Principal Component Analysis (PCA), Discriminant Analysis (DA) and Hierarchical Clustering Analysis (HCA). The three wines were differentiated from each other by six metabolites:  $\alpha$ -hydroxyisobutyrate, lactic acid, succinic acid, glycerol,  $\alpha$ -fructose and  $\beta$ -D-glucuronic acid. All multivariate analyses confirmed that the differentiation among the wines were related to micro-climate, and carbonate, clay, and organic matter content of soils. Additionally, the wine discrimination ability of NMR spectroscopy combined with chemometric methods, was proved when commercial Aglianico wines, deriving from different soils, were shown to be statistically different from the studied wines. Our findings indicate that multivariate statistical elaboration of NMR spectra of wines is a fast and accurate method to evaluate the molecular quality of wines, underlining the objective relation with *terroir*.

© 2010 Elsevier B.V. All rights reserved.

## 1. Introduction

Red wine is a complex molecular mixture resulting from fermentation of juice of *Vitis Vinifera* must [1]. The flavours and chemical composition of wine are usually related to the general environmental conditions insisting on vineyards from which wine is produced [2]. Many geographical factors such as climate, soil geology and composition, wild yeasts [3,4] and lactic acid bacteria [5], are able to synergistically influence wine molecular characteristics [6]. The relationship between wine quality and its specific site of production is commonly described by the French term “*terroir*” [7].

The analytical methods most used to investigate correlation between chemical composition of red wines and environmental factors related to winemaking are: High Performance Liquid Chromatography (HPLC) [8], Gas Chromatography–Mass Spectrometry (GC–MS) [9], Fourier Transform Ion Cyclotron Resonance Mass

Spectrometry (FT-ICR-MS) [10,11], High Performance Ion Exclusion Chromatography (HPICE) [12], and Nuclear Magnetic Resonance Spectrometry (NMR), used both with classical [12–14] and Site Natural Isotopic Fractionation (SNIF) NMR techniques [15].

Although NMR has poor sensitivity, liquid state NMR can rapidly provide spectra with highly reproducibility describing the main molecular profile of wines without laborious sample pre-treatments. Moreover, it has been shown [8,13,16] that NMR results, elaborated by multivariate statistical methods, can be used to identify and classify metabolites, common to many wines. The simplification of a multidimensional dataset obtained by liquid state <sup>1</sup>H NMR spectra, was successfully achieved with a number of multivariate statistical techniques, such as Principal Component Analysis (PCA), Discriminant Analysis (DA) and Hierarchical Clustering Analysis (HCA).

Recently, the combination of NMR spectroscopy with multivariate analysis has shown its high potential in wine quality assessment. Kosir et al. [15] attempted to classify Slovenian wines according to both geographical origins and beet sugar enrichment, by coupling the SNIF-NMR with some multivariate statistical methods. However, they were not able to significantly differentiate the wines produced in two adjacent Slovenian regions [15]. Brescia et al. [6], using the classical NMR spectroscopy with statistical analy-

\* Corresponding author at: Centro Interdipartimentale per la Risonanza Magnetica Nucleare (CERMANU), Università 'di Napoli Federico II, Via Università' 100, 80055 Portici (NA), Italy. Tel.: +39 081 788 5236; fax: +39 081 775 5130.

E-mail addresses: [Alessandro.piccolo@unina.it](mailto:Alessandro.piccolo@unina.it), [alpiccol@unina.it](mailto:alpiccol@unina.it) (A. Piccolo).

sis, were able to classify red wines produced in near by provinces of Apulia (Italy), but failed to distinguish wines obtained from different grape varieties [12]. Later, with a similar approach, they discriminated Apulian from Slovenian red wines confirming the efficacy of this method in assessing the geographical origin of wines [16]. Moreover, wines of Montepulciano d'Abruzzo (Italy) were satisfactorily depending on production years, yeast strains, and winemaking processes [8]. The applicability of a polyphasic approach was further confirmed when wines deriving from different cultivars, as grown on three different soils, were discriminated [7].

The aim of this work was to combine NMR spectroscopy with multivariate analysis in order to correlate wine quality to soil locations from which the cultivar originated. In order to reach an objective classification of the wine *terroir*, the grape variety "Aglianico", an ancient autoctonous cultivar of Campania (Italy), was collected from three near by soils with distinct pedological properties, and subjected to the same winemaking process. The wine samples were analyzed by  $^1\text{H}$  NMR spectroscopy and spectral data elaborated by PCA, DA and HCA statistical methods.

## 2. Experimental

### 2.1. Vineyards, soil sites and wine details

The vineyards under study are located in Taurasi (Avellino, Campania, Italy) ( $41^\circ 00' 35.39''\text{N}/14^\circ 57' 34.29''\text{E}$ ), but grown in three different soil sites: Case D'Alto [C], Coste Morante [S] and Macchia dei Goti [M]. The vineyards of "Aglianico di Taurasi" were planted in the same year (1982), on the same rootstock (1103 Paulsen) within a spurred cordon culture system. The plant density was 2000 stumps per hectare.

Soil C ( $41^\circ 00' 11.05''\text{N}/14^\circ 58' 19.00''\text{E}$ ) is located 400 m above sea level and developed from volcanic ashes and pomices on cinereous platforms. The soil is deep, mildly coarse textured, well drained and rich in nitrogen (N) and organic carbon (C). Soil S ( $41^\circ 01' 18.38''\text{N}/14^\circ 56' 50.51''\text{E}$ ) is situated 325 m above sea level. It is clayey and calcareous, slightly alkaline, with discrete exchangeable potassium and good C/N ratio. Soil M ( $40^\circ 59' 54.63''\text{N}/14^\circ 57' 42.81''\text{E}$ ) is located 340 m above sea level, and is calcareous clayey and rich in exchangeable calcium and potassium.

Samples from the C, S and M vineyard soils were collected in 2005. Soil samples were air-dried, sieved on 2 mm sieves, and analyzed for organic carbon, total nitrogen, C/N ratio, available phosphorous ( $\text{P}_2\text{O}_5$ ), electrical conductivity, cation exchange capacity (CEC), pH and total carbonates, as by the Official Italian Analytical Methods for Soil [17].

The 2005 vintage grapes were manually harvested from these vineyards and subjected to the same winemaking process as that carried out at the winery "Azienda Agricola Contrade di Taurasi di Enza Lonardo" located in Taurasi. Briefly, five quintals of harvested grapes were de-stemmed and crushed. Then, the three musts (about 3 hL) were transferred in different sanitised steel tanks, with the same inner volume (10 hL). The musts were treated with  $\text{K}_2\text{S}_2\text{O}_5$  (60 mg  $\text{kg}^{-1}$  of grapes). Fermentations took place at  $28^\circ\text{C}$  with indigenous yeast and caps were submerged twice a day. Macerations of pomaces lasted 12 days. The musts were pressed (about 8 bar) to obtain finished wines. About seven months after the end of malo-lactic fermentation, the wines were sieved through a rotary drum filter and finally bottled on December 2006.

A total number of 16 wine bottles were used for NMR analysis: five from Case d'Alto (C), six from Coste Morante (S), and five from Macchia dei Goti (M). Two more wine samples (F) were obtained from bottles of another Aglianico red wine commercially sold as

"Fidelis 2005", and originated from the Taburno area (Benevento, Italy) at a distance of about 40 km from Taurasi.

### 2.2. NMR experiments

Wine samples were used immediately after bottle uncorking and prepared by mixing 0.25 mL of wine with 0.75 mL of a deuterated water (99.8%  $\text{D}_2\text{O}/\text{H}_2\text{O}$ , Armar Chemicals) solution containing 2% (v/v) formic acid (98–100% RG, Merck Chemicals) as internal standard. Before analysis, each sample was stirred in a vortex, and transferred into a stoppered NMR tube (5 mm, 7", 507-HP-7, NORELL) and the remaining void volume gently degassed by a  $\text{N}_2$  flux.

A 400 MHz Bruker Avance spectrometer, equipped with a 5 mm BBI Bruker probe and working at the  $^1\text{H}$  frequency of 400.13 MHz, was employed to conduct all liquid state NMR measurements at a temperature of  $298 \pm 1\text{ K}$ . Mono-dimensional  $^1\text{H}$  spectra were acquired with 2 s of thermal equilibrium delay, a  $90^\circ$  pulse length ranging between 8.21 and 8.86  $\mu\text{s}$ , 200 transients, 32,768 time domain points, and 11.9692 ppm (4.7893 kHz) as spectral width. The free induction decay (FID) was multiplied by an exponential factor corresponding to 0.5 Hz. An excitation sculpting (ZGESGP) [18] and a modified T1 inversion-recovery (*mT1IR*) multi-suppression pulse sequences were both applied to suppress residual water signal and ethanol multiplets.

For excitation sculpting experiments, a specific shape pulse was built by modifying the basic shape pulse *SQUA100.1000* reported in the Bruker Topspin database. The most efficient signals multi-suppression occurred when the shape pulse length (P12) and power (SP1) were adequately calibrated. These parameters were optimized, respectively, at 7000  $\mu\text{s}$  for all samples, and from 40.01 to 41.4 dB, depending strictly on sample. The T1 inversion-recovery sequence was adjusted to provide simultaneous suppression of both water and ethanol signals by, respectively, water pre-saturation and selection of appropriate inversion-recovery time-length.

Structural identification of wine metabolites was achieved by 2D NMR experiments: homo-nuclear  $^1\text{H}$ - $^1\text{H}$  COSY (CORrelation SpectroscopY) and TOCSY (Total CORrelation SpectroscopY), and hetero-nuclear  $^1\text{H}$ - $^{13}\text{C}$  HSQC (Hetero-nuclear Single-Quantum CORrelation). In both 2D homo-nuclear experiments excitation sculpting was applied by employing two Bruker pulse sequences *COSYDFESGPPH* and *MLEVESGPPH* respectively. 2D spectra were acquired with 96 total transients, a time domain of 2k points (F2) and 256 experiments (F1), 16 dummy scans, and 11.9692 ppm as spectral width. Moreover, TOCSY experiments were conducted with a mixing time of 80 ms and trim pulse length of 2500 ms. The hetero-nuclear HSQC experiments were acquired with 80 total transients, 16 dummy scans, 0.5 ms of trim pulse length, a time domain of 2k points and 256 experiments, and 300 ppm (120.039 kHz) as  $^{13}\text{C}$  spectral width. A baseline correction was applied to all mono- and bi-dimensional spectra and the proton frequency axis was calibrated by associating the formic acid signal at 8.226 ppm [19]. Spectra were processed by using both Bruker Topspin Software (v.1.3) and MestReC NMR Processing Software (v.4.9.9.9).

### 2.3. Multivariate data analyses

Statistical data elaboration was achieved by the XLStat software v.7.5.2 (Addinsoft). A number (119) of  $^1\text{H}$  NMR variable buckets were screened, each bucket width being adequately chosen to include individual peaks or distinct multiplets, when possible. These buckets were equally adopted for all spectra. Prior to multivariate analysis, each spectrum was normalized by dividing the single bucket area by the sum of each signal integrated area. The

**Table 1**  
<sup>1</sup>H chemical shift and assignment of signals which mostly differentiated wine samples.

Peak identifier	Buckets interval (ppm)	Assignments	$\delta$ (ppm)	Multiplicity and group
1	1.299–1.248	$\alpha$ -Hydroxyisobutyrate	1.267	s, CH <sub>3</sub>
2	1.438–1.373	Lactic acid	1.410	d, 7 $\beta$ -CH <sub>3</sub>
3	2.688–2.638	Succinic acid	2.665	s, CH <sub>3</sub>
4	3.577–3.510	Glycerol	3.542	q, CH <sub>2</sub>
5	3.782–3.743	$\alpha$ -Fructose	3.766	t, C2H, ring
6	4.624–4.586	$\beta$ -D-Glucuronic acid	4.605	d, C1H, ring

s: singlet; d: doublet; t: triplet; q: quartet.

normalized values were further mean centred for PCA (Principal Component Analysis) and DA (Discriminant Analysis). The validation of the DA method validation consisted in building a validation set, achieved by dividing all samples in a training set and in a test set. The former set employed a discriminant model formed by 10 of the 16 samples, whereas the latter set was used to test the model and comprised 6 of the remaining samples. A similar partition, containing 37.5% of samples in the test set, allowed a sufficient number of samples in the training set as well as a representative number of samples in the test set [16]. This validation was repeated 5 times by changing, each time randomly, the samples included in the training and test set.

A HCA (Hierarchical Component Analysis) was developed by applying a hierarchical ascendant cluster analysis. By this technique, the extent of similarity among samples was measured by Euclidean distances, while cluster aggregation was based on the average linkage method [20].

### 3. Results and discussion

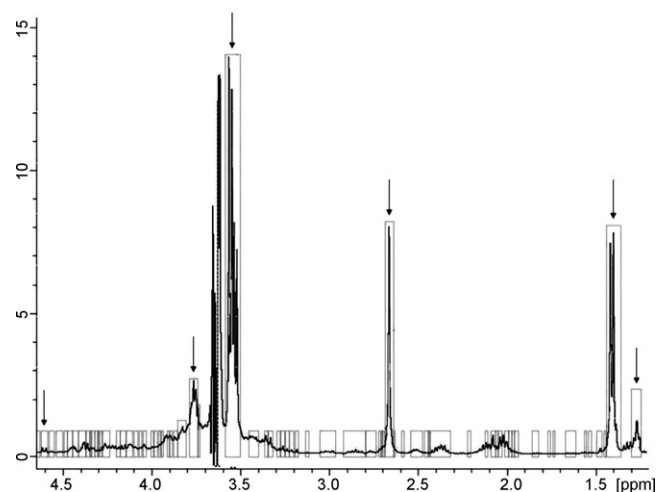
The <sup>1</sup>H NMR spectra of red wines (Supplementary Fig. S1a) showed that most intense signals were those of ethanol (a triplet at 1.204 ppm and a quartet at 3.674 ppm) and water (a singlet at 4.804 ppm). These signals interfere with spectral elucidation since their large intensities not only depress other meaningful signals but also overlap other molecular peaks resonating at near by frequencies. Excitation sculpting (ZGESGP) and a modified T1 inversion-recovery (*mT1IR*) pulse sequences were applied to suppress these interfering water and ethanol signals.

The *mT1IR* pulse sequence is related to <sup>1</sup>H spin-lattice relaxation times of all visible signals and showed that spins of ethanol signals relaxed almost completely before all other molecular spins began to significantly return to equilibrium. In fact, it was observed that application of a 3 s delay between initial 180° and second 90° hard pulse, efficiently suppressed the ethanol multiplets and acetic acid singlet at 2.09 ppm [13,21], while it allowed recovering all other signals without significant intensity loss. In addition, this pulse sequence included a water signal pre-saturation, that resulted in a concomitant removal of both undesired water and ethanol signals (Supplementary Fig. S1c). Even though both excitation sculpting and modified T1 inversion-recovery techniques were efficient in suppressing the interfering signals, the *mT1IR* pulse sequence was more successful in showing a cleaner <sup>1</sup>H spectrum of red wine (see comparison in Supplementary Fig. S1b and c). Consequently, the *mT1IR* sequence was applied to all mono-dimensional spectra of wine samples. Moreover, *mT1IR* spectra of wine samples from the same soil did not show significant differences, whereas variations were noticeable in wines from different soils. Therefore, these spectra could be integrated to evaluate, in details, the effect of soil site on wines molecular composition. An example of selected buckets with variable amplitudes within the 4.55–1.15 ppm spectral region is shown in Fig. 1.

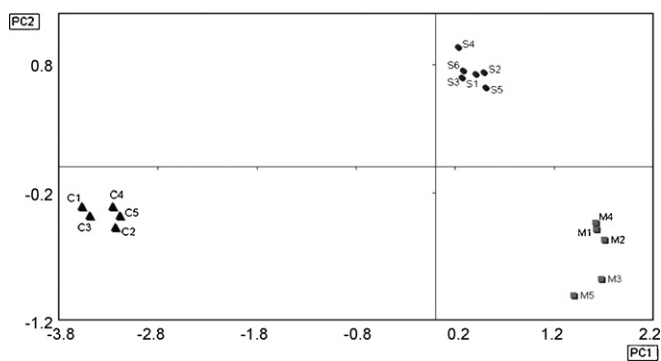
A preliminary Analysis of Variance (ANOVA) was applied to a data matrix constituted by 119 variables and 16 samples. It was observed that only 6 out of the 119 employed variables con-

tributed to wine differentiation within a 95% confidence interval (Fisher test) [22]. The buckets capable of statistical differentiation is indicated by arrows in Fig. 1, while precise bucket intervals with structural assignment of related compound variables are reported in Table 1. These compounds were  $\alpha$ -hydroxyisobutyrate [13], lactic acid, succinic acid, glycerol,  $\alpha$ -fructose, and  $\beta$ -D-glucuronic acid [13,23]. A spectral expansion for these metabolites is reported in Supplementary Fig. S2. Molecular attribution to signals in respective buckets was achieved by combining spectral interpretation of mono- and double-dimensional experiments (Supplementary Figs. 3–5). Signals identification was further supported by molecular attribution reported in literature for wine <sup>1</sup>H NMR spectra [13,14,16,21,24]. Most of the above compounds are mainly of microbial origin: lactic acid is the metabolic product of malo-lactic fermentation performed by malo-lactic bacteria (MLB) [25]; succinic acid is the major organic acid produced by yeast metabolism [26]; glycerol is the final product of glyceropiruvic fermentation by yeast and its formation accompanies ethanol production in sugar fermentation [27]. Although fructose is not a product of microbial origin, its concentration, related to glucose content, may influence yeast fermentation, the glycerol concentration depending on glucose/fructose content in must [28]. Moreover, the composition of yeasts wild strains and MLB were found slightly different in the three musts and related wines (data not shown).

The multivariate Principal Component Analysis (PCA) calculates linear combinations of a starting set of variables on the basis of their maximum variance [16]. It reduces the dimensions multiplicity of the original data matrix, while retaining the maximum amount of variability, as well as the original information contained in the data set [29]. PCA decomposes a data matrix with *m* rows (wine samples) and *n* columns (NMR spectral areas) into the product obtained by



**Fig. 1.** <sup>1</sup>H spectrum of M1 sample. The superimposed rectangles contain the selected buckets within the 4.55–1.15 ppm region, which were equally employed to integrate all wine spectra. Arrows indicate peaks which statistically contributed to discriminate among C, S and M wine samples [Case d'Alto (C), Coste Morante (S) and Macchia dei Goti (M)].



**Fig. 2.** Scatter-plot of the scores generated from the two principal components, PC1 and PC2, obtained using analytical data from  $^1\text{H}$  NMR red wine spectra.

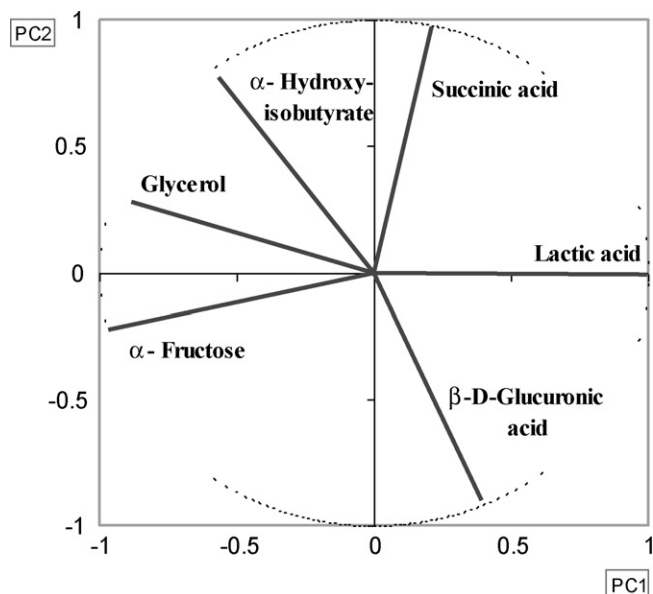
multiplying a matrix of scores ( $m$  rows for wine samples and  $P < n$  columns for Principal Components, PCs), with a matrix of loadings ( $P < m$  rows for PCs and  $n$  columns for NMR spectral areas). Each combination of two resulting PCs, creates a bi-dimensional space, constituted by PCs as axis. Loading vectors are associated to each of such factorial planes, correlated to the original variables, and orientated toward the direction in which the maximum variance of variables is expressed. This representation projects the original observations in a new reference system, whose coordinates are regulated by the variance, and provides a final bi-dimensional score-plot, that may be interpreted with loading-plot information [30]. This means that score-plots can indirectly highlight possible differences or similarities among samples, while loading-plots reveal the responsible spectral components (variables).

The PCA analysis was conducted on a data matrix containing only the six variables capable of sample differentiation (Table 1) [20]. In PCA, the first two Principal Components explained 86.0 and 11.2% of total variance, respectively (Fig. 2). The three different wine samples were projected to three separated score-plot regions, whose differences resulted statistically significant ( $p < 0.0001$ ) by the Barlett's sphericity test [31]. In fact, it was found (Fig. 2) that PCA well discriminated the C wine from S and M along the PC1 direction, being the former positioned in the negative PC1 score. On the other hand, M and S wines were mutually well separated along the PC2 axis. Furthermore, this PCA distribution showed that wine replicates deriving from the same soils were associated with close score-plot domains, while groupings of wine replicates were spatially well separated according to soil differences (Fig. 2). This finding testified that the molecular characteristics of the three wines were sufficiently varied to show PCA distinctions.

The three variables which mostly contributed to differences by PC1 were provided by lactic acid (74.0%),  $\alpha$ -fructose (12.7%), and glycerol (12.2%) (Table 2). The PCA linear orientations of variance vectors for these variables are shown in the loading-plot (Fig. 3), that correlates score-plot domains with variables. The large distance among coordinate values along PC1 between the C and the S and M domains (Supplementary Table S1) may then be attributed

**Table 2**  
Principal Component Analysis eigenvalues and variable contributions (%).

Eigenvalues and variables	PC1	PC2
Eigenvalue	3.472	0.453
Variance (%)	85.967	11.220
Cumulate variance (%)	85.967	97.187
$\beta$ -D-Glucuronic acid	0.033	1.334
$\alpha$ -Fructose	12.732	5.224
Glycerol	12.203	9.228
Succinic acid	0.464	76.362
Lactic acid	74.025	0.020
$\alpha$ -Hydroxyisobutyrate	0.543	7.831



**Fig. 3.** Loading-plot of the variables associated with the first two principal components of PCA, calculated using analytical data.

to larger lactic acid in S and M wines than in C wine. Such score-plot distinction is additionally provided by the greater content of  $\alpha$ -fructose and glycerol found for C than for S and M wines. However, the more positive PC1 scoring for M than for S wines is explained by a greater glycerol content in the latter. As for PC2, the main difference among samples was provided by succinic acid (76.362%), whose content was largest in S wine. Moreover, M samples showed both greater and smaller content of  $\beta$ -D-glucuronic acid and  $\alpha$ -hydroxyisobutyrate, respectively, than C and S samples. It is furthermore noted that within the M statistical domain, M3 and M5 samples showed more negative PC2 scorings than the remaining M samples, due to concomitant larger  $\beta$ -D-glucuronic and lower glycerol amounts.

The Discriminant Analysis (DA) is a supervised statistical method to classify samples. DA extracts discriminant functions from a data matrix composed by independent variables in order to maximize inter-class variance and minimize intra-class variance. This criterium can be supported by a validation test in order to verify the statistical confidence with which the a priori classification of some observations (test set) coincides with the a posteriori DA prediction as elaborated by information provided by the remaining samples (training set). One of the five applied discriminant models, which were obtained by a validation test is shown (Supplementary Fig. 6). In this model, C3, C4, S3, S6, M3 and M5 samples were used as test-set samples. Since PCA indicated that M3 and M5 samples were slightly differentiated from other M samples, and may thus be hardly recognized by DA, we chose M3 and M5 samples as M unknown objects to verify the reliability of DA analysis. This DA exercise proved that all wine observations were correctly associated to respective group with 100% of success (observation coordinates and square distances for this model are reported in Supplementary Table S2). Moreover, in all five developed discriminant models, each wine sample was correctly associated to the respective group with 100% of success and a significance of 0.05.

The observed PCA differentiation among wine samples is reflected by different soil properties. In fact, while S and M soils are definitely carbonatic soils, as shown by pH value and total carbonates content, the volcanic C soil is less carbonatic (Table 3). Moreover, the S and M soils have a larger clay character and a greater organic carbon and total nitrogen content than the C soil. The diverse soil properties were confirmed by PCA differentiation

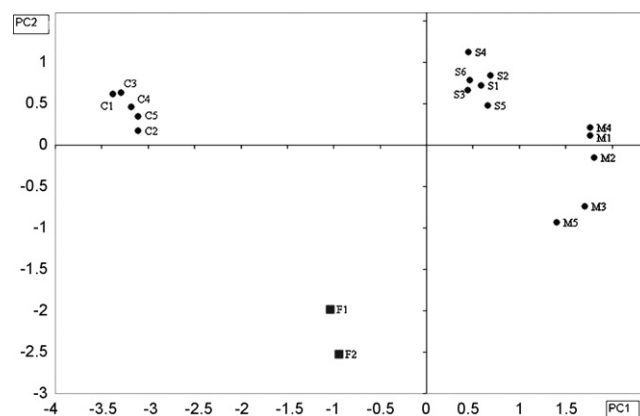
**Table 3**  
Properties of soils under vineyards.

Properties	C	S	M
Gross sand (%)	43.8	13.8	16.7
Fine sand (%)	41.6	33.1	34.7
Silt (%)	12.1	22.8	21
Clay (%)	2.5	30.3	27.6
Organic carbon (g kg <sup>-1</sup> )	27.1	13.5	12.2
N total (g kg <sup>-1</sup> )	2.45	1.28	1.05
C/N	11.1	10.6	11.6
P <sub>2</sub> O <sub>5</sub> (mg kg <sup>-1</sup> )	24.3	43.6	16.9
Electric conductivity (dS m <sup>-1</sup> )	0.099	0.145	0.126
C.E.C. (mequiv./100 g)	26.3	29.3	24.4
pH (H <sub>2</sub> O)	6.58	8.15	8.26
Total carbonates (g kg <sup>-1</sup> )	10.5	60.6	109

among wine samples and this suggested that soil characteristics had a bearing on the molecular quality of wines obtained from the soils.

The PCA was modified (PCAe) by adding to the data matrix two more samples of a commercial Aglianico wine (F) from Campania region (supplementary Table S3). Since PCA score-plots arrange similar samples in near by zones, while placing dissimilar samples in distant zones, addition of different unrelated samples to the dataset would presumably project their scoring in a distinct domain of the score-plot. Fig. 4 shows the new PCAe score-plot where the separation between F (negative scoring on both principal functions) and previous C, S, and M samples is significant. Such separation of F samples was attributed to much lower content of glycerol and succinic acid for PC2, and to both smaller and larger amount of lactic acid and  $\alpha$ -fructose, respectively, in PC1, than for C, S, and M samples (Fig. 4).

HCA is an unsupervised method that recognizes and distributes data grouping, according to their affinity, in clusters of progressive dissimilarity, as described in a dendrogram. It is assumed that the closer the objects in a space defined by variables, the more similar are their properties. HCA was thus conducted to verify the classification of C, S and M wines according to their mutual dissimilarity and their relation to F wines. The resulting dendrogram and main descriptive values are shown in Fig. 5 (relative matrix with quantitative dissimilarities in Supplementary Table S4). Knot 35 in the dendrogram shows that the most dissimilar clusters are those

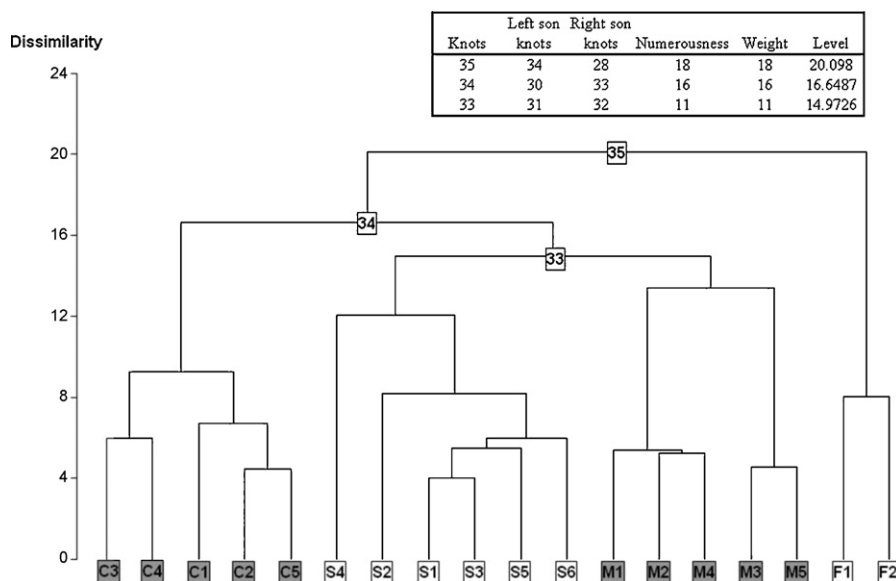


**Fig. 4.** Scatter-plot of PCAe. The scores generated from the two principal components, PC1 and PC2, obtained using analytical data from <sup>1</sup>H NMR red wine spectra. F is two samples of an Aglianico different from C, S, and M. The PC1 and PC2 explained, respectively, 70.48% and 23.068% of the total variance.

between F wines and the rest of Aglianico samples. This indicates that this HCA approach visibly discriminates wines of different origin and quality. Furthermore, knot 34 reveals that the C cluster is recognized by HCA as different from the macro-cluster containing S and M wines. This is in accordance with results from both the PCA score-plot and its relation to soil properties. HCA provides further insight by closely classifying M and S clusters (knot 33) and confirming PCA results, which indicated that M3 and M5 were slightly different, within the M cluster, from M1, M2 and M4 samples.

#### 4. Conclusions

The study shows that <sup>1</sup>H NMR spectroscopy combined with multivariate statistical analyses can rapidly and efficiently discriminate three red wines derived from the same Aglianico grape variety grown on different soils. Multi-suppression of both water and ethanol signals in wine samples, by a specifically modified inversion-recovery pulse technique, offered highly reproducible NMR spectra without overlapping of meaningful peaks. Although this approach also removed the acetic acid signal, the discrimination among the wines were anyhow assured.



**Fig. 5.** Dendrogram of red wines obtained using analytical data. Euclidean distance was employed to find the dissimilarity among the samples, and the average correlation as aggregation method.

Despite the wines were produced employing the same grape variety and the same winemaking procedures, the PCA elaboration of their  $^1\text{H}$  NMR spectra was able to significantly recognize dissimilarities among their molecular properties, as shown by the content of lactic acid,  $\alpha$ -fructose, glycerol, succinic acid,  $\beta$ -D-glucuronic acid and  $\alpha$ -hydroxyisobutyrate. Statistical variables of PCA were also employed to perform a discriminant analysis, whose five validation tests provided a 100% correct prediction of differences among wines.

Since the diversity of the three wines could be only reconciled to differences in soils, these results suggest that bio-pedo-climatic factors determined the ultimate wine quality. However, some metabolites (lactic and succinic acid, and glycerol) may have derived from must and wine microbial activity, thus suggesting that indigenous yeasts and MLB may also contribute to the expression of terroir. Multivariate statistical analysis of  $^1\text{H}$  NMR spectra indicated a marked difference between one of the Aglianico wines and the other two. Since the vineyards from which the wines originated were closely located (lower than 3 km apart), wine quality appears to be conditioned less by macro-climate than by micro-climate and intrinsic soil properties, which in turn influence the must microflora. In fact, the differentiated wine originated from a soil of higher altitude than for the other two wines, with consequent differences in temperature and humidity. Moreover, soil characteristics, such as texture and carbonate as well as organic carbon content, varied between one soil and the other two ones. The peculiarity of this one different soil is reflected by PCA, DA and HCA results from  $^1\text{H}$  NMR spectra of wines, thereby sustaining the assumption that soil physical and chemical properties are able to determine the overall wine quality.

A further evidence of the efficacy of combining careful NMR spectroscopy with multivariate statistics to assess wine quality and its *terroir*, was shown by comparing results of the three wines under study with those of a commercially available Aglianico wine from another soil site. Both  $^1\text{H}$  NMR spectroscopy and multivariate elaboration clearly indicated that the latter wine had a significant different quality than the formers. Thus, the findings of this work are promising in indicating  $^1\text{H}$  NMR spectroscopy as a rapid and objectively sound technique for a geographical identification of wine quality.

#### Acknowledgements

We thank Alessandro Lonardo, owner of “Azienda Agricola Contrade di Taurasi” (Taurasi, AV, Italy). We are grateful to Maurizio De Simone for his oenological support in the cellar.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.aca.2010.06.003](https://doi.org/10.1016/j.aca.2010.06.003).

#### References

- [1] A. Ramos, H. Santos, in: G.A. Webb (Ed.), *Annual Reports on NMR Spectroscopy*, vol. 37, Academic Press, New York, 1998, pp. 179–199.
- [2] G.V. Jones, M.A. White, O.R. Cooper, K. Storchmann, *Clim. Change* 73 (2005) 319–343.
- [3] N. Francesca, M. Chiurazzi, R. Romano, M. Aponte, L. Settanni, G. Moschetti, *World J. Microbiol. Biotechnol.* 26 (2010) 337–351.
- [4] E. Nikolaou, E.H. Soufleros, E. Bouloumpasi, N. Tzanetakakis, *Food Microbiol.* 23 (2006) 205–211.
- [5] E. Boido, K. Medina, L. Farina, F. Carrau, G. Versini, E. Dellacassa, *J. Agric. Food Chem.* 57 (2009) 3278–6271.
- [6] G.V. Jones, in: R.W. Macqueen, L.D. Meinert (Eds.), *Geoscience Canada Reprint Series Number 9*, Geological Association of Canada, Newfoundland, 2006, pp. 1–14.
- [7] G.E. Pereira, J.P. Gaudillere, C.V. Leeuwen, G. Hilbert, M. Maucourt, C. Deborde, A. Moing, D. Rolin, VI International Terroir Congress, Montpellier, 2006.
- [8] M. Giaccio, A. Del Signore, *J. Sci. Food Agric.* 84 (2004) 164–172.
- [9] M. Ugliano, L. Moio, *J. Sci. Food Agric.* 86 (2006) 2468–2476.
- [10] R.D. Gougeon, M. Lucio, M. Frommberger, D. Peyron, D. Chassagne, H. Alexandre, F. Feuillat, A. Voilley, P. Cayot, I. Gebefugi, N. Hertkorn, P. Schmitt-kopplin, *PNAS* 106 (2009) 9174–9179.
- [11] A. Cuadros-Inostroza, P. Giavalisco, J. Hummel, A. Eckardt, L. Willmitzer, H. Pena-cortes, *Anal. Chem.* 82 (2010) 3573–3580.
- [12] M.A. Brescia, V. Caldarola, A. De Giglio, D. Benedetti, F.P. Panizzi, A. Sacco, *Anal. Chim. Acta* 458 (2002) 177–186.
- [13] Y.Y. Du, G.Y. Bai, X. Zhang, M.L. Liu, *Chin. J. Chem.* 25 (2007) 930–936.
- [14] S. Clark, N.W. Barnett, M. Adams, I.B. Cook, G.A. Dyson, G. Johnston, *Anal. Chim. Acta* 563 (2006) 339–345.
- [15] I.J. Kosir, M. Kocjancic, N. Ogrinc, J. Kidric, *Anal. Chim. Acta* 429 (2001) 195–206.
- [16] M.A. Brescia, I.J. Kosir, V. Caldarola, J. Kidric, A. Sacco, *J. Agric. Food Chem.* 51 (2003) 21–26.
- [17] Ministerial Decree of Italian Republic, Decreto Ministeriale delle Politiche Agricole e Forestali del 13 Settembre 1999, Published on Official Gazette n° 248, 21st October 1999.
- [18] T. Parella, P. Adell, F. Sanchez-Ferrando, A. Virgili, *Magn. Reson. Chem.* 36 (1998) 245–249.
- [19] L.A. Cardoza, B.J. Cutak, J. Ketter, C.K. Larive, *J. Chromatogr. A* 1022 (2004) 131–137.
- [20] R. Todeschini, *Introduzione alla Chemiometria*, EdiSES s.r.l., Naples, 1998, pp. 37–79.
- [21] E.F. Boffo, L.A. Tavares, M.M.C. Ferreira, A.G. Ferreira, *Food Sci. Technol.* 42 (2009) 1455–1460.
- [22] P. Dagnelie, *Théorie et méthodes statistiques*, vol. 2, Les Presses Agronomiques de Gembloux, Gembloux, 1986.
- [23] A. Caridi, *Int. J. Food Microbiol.* 120 (2007) 167–172.
- [24] I.J. Kosir, J. Kidric, *Anal. Chim. Acta* 458 (2002) 77–84.
- [25] C. Davis, D. Wibowo, R.E. Eschenbruch, T.H. Lee, G.H. Fleet, *Am. J. Enol. Vitic.* 36 (1985) 290–301.
- [26] T. Magarifuchi, K. Goto, Y. Iimura, M. Tadenuma, G. Tamura, *J. Ferment. Bioeng.* 80 (1995) 335–361.
- [27] M.J. Taherzadeh, L. Adler, G. Liden, *Enzyme Microb. Technol.* 31 (2002) 53–66.
- [28] S.K. Yalcin, Z.Y. Ozbas, *World J. Microbiol. Biotechnol.* 21 (2005) 1303–1310.
- [29] D. Smejkalova, R. Spaccini, A. Piccolo, *Eur. J. Soil Sci.* 59 (2008) 496–504.
- [30] R.G. Brereton, in: J. Wiley, sons (Eds.), *Chemometrics—Data analysis for the Laboratory and Chemical Plant*, Wiley, England, 2003, pp. 184–249.
- [31] W.R. Dillon, M. Goldstein, *Multivariate Analysis. Methods and Applications*, John Wiley & Sons, New York, 1984, pp. 44–47.