agriculture

MDPI

Article

^{1}H-NMR Spectroscopy Coupled with Chemometrics to Classify Wines According to Different Grape Varieties and Different Terroirs

Paola Bambina [1], Alberto Spinella [2], Giuseppe Lo Papa [1], Delia Francesca Chillura Martino [3], Paolo Lo Meo [3], Luciano Cinquanta [1] and Pellegrino Conte [1,*]

[1] Department of Agricultural, Food and Forest Sciences, University of Palermo, V.le delle Scienze 13, 90128 Palermo, Italy; paola.bambina@unipa.it (P.B.); giuseppe.lopapa@unipa.it (G.L.P.); luciano.cinquanta@unipa.it (L.C.)
[2] Advanced Technologies Network Center (ATeN Center), University of Palermo, Via F. Marini 14, 90128 Palermo, Italy; alberto.spinella@unipa.it
[3] Department of Biological, Chemical and Pharmaceutical Sciences and Technologies, University of Palermo, V.le delle Scienze 16, 90128 Palermo, Italy; delia.chilluramartino@unipa.it (D.F.C.M.); paolo.lomeo@unipa.it (P.L.M.)
* Correspondence: pellegrino.conte@unipa.it

Abstract: In this study, ^{1}H-NMR spectroscopy coupled with chemometrics was applied to study the wine metabolome and to classify wines according to different grape varieties and different terroirs. By obtaining the metabolomic fingerprinting and profiling of the wines, it was possible to assess the metabolic biomarkers leading the classification (i.e., phenolic compounds, aroma compounds, amino acids, and organic acids). Moreover, information about the influence of the soil in shaping wine metabolome was obtained. For instance, the relationship between the soil texture and the content of amino acids and organic acids in wines was highlighted. The analysis conducted in this study allowed extraction of relevant spectral information not only from the most populated and concentrated spectral areas (e.g., aliphatic and carbinolic areas), but also from crowded spectral areas held by lowly concentrated compounds (i.e., polyphenols). This may be due to a successful combination between the parameters used for data reduction, preprocessing and elaboration. The metabolomic fingerprinting also allowed exploration of the H-bonds network inside the wines, which affects both gustatory and olfactory perceptions, by modulating the way how solutes interact with the human sensory receptors. These findings may have important implications in the context of food traceability and quality control, providing information about the chemical composition and biomolecular markers from a holistic point of view.

Keywords: ^{1}H-NMR spectroscopy; metabolomics; chemometrics; wine; terroir

Citation: Bambina, P.; Spinella, A.; Lo Papa, G.; Chillura Martino, D.F.; Lo Meo, P.; Cinquanta, L.; Conte, P. ^{1}H-NMR Spectroscopy Coupled with Chemometrics to Classify Wines According to Different Grape Varieties and Different Terroirs. *Agriculture* **2024**, *14*, 749. https://doi.org/10.3390/agriculture14050749

Academic Editor: Hongbin Pu

Received: 2 April 2024
Revised: 29 April 2024
Accepted: 9 May 2024
Published: 11 May 2024

1. Introduction

Food traceability refers to the ability to access information related to food products, ingredients, and raw materials through their entire life cycle [1]. It can be considered as a key tool to assess authenticity, and refers to the assurance that a food product is genuine, unadulterated, and accurately labeled with respect to its composition, origin, and quality attributes. Different approaches exist to authenticate food products that aim at analyzing the overall chemical composition and identifying biomolecular markers [2]. In this context, the so-called omics sciences have proved to be optimal tools at the service of traceability and authenticity of food and beverages [3]. Omics sciences encompass data-driven approaches that aim to comprehensively analyze biological systems by employing high-throughput technologies [4]. These sciences focus on the analysis of complete sets of biomolecules, including DNA, RNA, proteins, and other metabolites. The study of the ensemble of biomolecules provides a thorough overview of the molecular processes

occurring in a biological system. Among the major omics sciences, which include genomics, transcriptomics, proteomics, epigenomics, etc., metabolomics focuses on the metabolome, that is the ensemble of all the metabolites contained in a biological sample in a specific moment. Metabolomic science is defined as "the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification" [5]. It is a comprehensive approach that enables study of the metabolic state of a system, by exploring the dynamic changes in metabolite profiles as caused by genetic variations or changes in environmental conditions. One of the major analytical methods used to perform metabolomic analysis is Nuclear Magnetic Resonance (NMR) spectroscopy [6,7]. This technique is based on the absorption and re-emission of energy by targeted nuclei due to the application of an external magnetic field [8]. Depending on the type of nucleus being targeted by the applied magnetic field, different types of NMR analysis can be conducted, including ^{1}H-NMR, ^{13}C-NMR, ^{31}P-NMR, ^{19}F-NMR, etc. Among these, ^{1}H-NMR spectroscopy plays a major role because it provides structural and quantitative information on a wide range of metabolites simultaneously, based on the abundance of the hydrogen-nuclei, in a non-destructive and highly reproducible way [5]. Metabolomics is currently applied to a great variety of subjects, including agricultural, food and nutrition sciences [9–11]. Foods and beverages are evaluated by analyzing their macro- and micro-component composition with the aim of investigating sensory quality, nutritional content, safety, authenticity, and traceability [12,13]. With reference to wine science, metabolomic analysis has been employed to characterize wines derived from different grape varieties [14], geographical origins [15,16], vintages [17], and winemaking techniques [18]. Despite that these studies have provided valuable insights into the chemical composition, quality, and sensory characteristics of wines with respect to different environmental factors, some gaps and challenges still emerge, mainly related to the limited capability of compound identification (especially of low-abundant analytes) due to the complexity of the wine spectrum [19]. Moreover, the inherent heterogeneity of wine composition, due to different environmental factors, geographical origins, agronomical and oenological practices, introduces variability and batch effects into metabolomic data, making it harder to extract information about the origin of the products. It is well recognized that wines produced in different production areas develop different sensory characteristics [20]. This phenomenon, identified with the French term terroir, is due to the synergistic interactions among several factors, including grape variety, climate, soil, biodiversity features and human activities (e.g., viticultural and oenological techniques) [21]. Among these factors, the soil is a critical element for plant life, because it plays several ecological functions, such as carbon storage and water regulation, and it is responsible for the vehiculation of nutrients towards plant roots [21,22]. The role of the soil is complex and multifaceted, and several authors have highlighted the importance of some soil features (e.g., temperature, water supply, and mineral content, particularly nitrogen supply) in modulating vine vigor and grape composition [22,23]. However, the mechanisms regulating the effects of the soil on wine quality have been only moderately addressed [24,25]. In one of our previous studies [26], we employed ^{1}H-NMR-based metabolomics to classify Nero d'Avola red wines from different terroirs. In particular, we focused on the influence of the soil in shaping the micro-component composition of Nero d'Avola wines. We proposed a revised and boosted ^{1}H-NMR-based metabolomic approach with the aim of enhancing data quality and throughput performance. The analysis allowed us to point out that the soil strongly affected wine metabolome. However, we reported that the workflow for such analysis needed to be further optimized and verified. Therefore, in order to test the powerfulness (and weaknesses) of our analytical methodology, the primary goal of this new study was to evaluate the capability of our revised method to assess the chemical variability of wines obtained by different grape varieties. To attain this aim, the metabolomic analysis of wines from two *Vitis vinifera* L. grape varieties (e.g., Nero d'Avola red and Grillo white cultivars) was carried out. Moreover, in order to further verify the findings of our previous studies [26–28], the second goal was to investigate the effect of the soil on the

metabolome of white wines produced from *Vitis vinifera* L. cv Grillo grapes. The effects of the main soil chemical–physical parameters (namely texture, pH, total carbonates, cation exchange capacity, electric conductivity, organic matter, and mineral composition) on the determination of wine metabolome were investigated.

2. Materials and Methods

2.1. Wine Terroirs

The study was performed on red and white wines obtained from grapes grown in eight vineyards located in the hilly landscape nearby the city of Menfi (Agrigento), along the southwestern coast of Sicily (Southern Italy). The vineyards hosted the cultivars Nero d'Avola and Grillo of *Vitis vinifera* L. In order to investigate the effect of the soil on wine metabolome, the main soil chemical physical parameters were considered, namely texture, pH, total carbonates, organic matter, cation exchange capacity, electric conductivity, and mineral composition [27]. According to the results of the aforementioned analyses, the pedological classification was performed according to the Soil Taxonomy [29].

Soils were identified as Soil 1–8, where the soils from 1 to 4 hosted the Nero d'Avola vines and the soils from 5 to 8 hosted the Grillo vines.

Details about the chemical–physical parameters of the soils are reported in the Soil description section of the Supplementary Information and in Tables S1–S3 of the same Supplementary section.

2.2. Wine Elaboration

Grapes from each vineyard were manually harvested in September 2021, when about 200 g L^{-1} of fermentable sugars were accumulated in berries. The red wines, made from Nero d'Avola grapes, were obtained by applying the vinification process described in Bambina et al. [26]. Briefly, 150 kg of grapes for each trial were destemmed and softly crushed. Then, 5 g hL^{-1} of $K_2S_2O_5$ were added in order to inactivate the grape polyphenol-oxidase activity. The alcoholic fermentation was carried out by the inoculation of a *pied de cuvè* prepared with 20 g hL^{-1} of Lalvin EC 1118 *Saccharomyces cerevisiae* yeasts. 0.06 g hL^{-1} of thiamine and diammonium phosphate (DAP) were added to achieve 200 mg L^{-1} of yeast assimilable nitrogen (YAN). The fermentation temperature was 25 ± 1 °C. A punching down per day was carried out until the achievement of 6 % v v^{-1} of ethanol. Then, punching down was performed twice per day till the end of the alcoholic fermentation. Two pump-overs were performed: the first after 24 h from the beginning of the alcoholic fermentation and the second at the achievement of 6 % v v^{-1} of ethanol. The racking was carried out at the end of the alcoholic fermentation. Then, the malolactic fermentation was started by inoculation of selected lactic bacteria (Enartis-ML ONE). At the end of malolactic fermentation, wines were racked and $K_2S_2O_5$ was added until wines reached 25 mg L^{-1} of free SO_2.

The white wines, made by Grillo grapes, were vinified according to Bambina et al. [27]. A total of 150 kg of grapes for each trial were stored in a refrigerated room (8–10 °C) for 24 h. Then, the grapes were destemmed, softly crushed, and pressed. The obtained musts were added with 5 g hL^{-1} of $K_2S_2O_5$ and clarified by static settling at 5 °C for 24 h, with the addition of 3 g hL^{-1} of pectolytic enzymes (Endozym ICS 10 Èclairs, Pascal Biotech, Negeri Sembilan, Malaysia). After racking, the clear musts were inoculated by a *pied de cuvè* prepared with 20 g hL^{-1} of Lalvin QA23 *Saccharomyces cerevisiae* yeasts. 0.06 g hL^{-1} of thiamine and 20 g hL^{-1} of DAP were added to the musts in order to achieve 200 mg L^{-1} of YAN. The fermentation temperature was 16 ± 1 °C. Oxygenations of 1/5 of each must were performed at the achievement of 5–6% and 8–9% v v^{-1} of ethanol. At the end of the alcoholic fermentation, wines were racked and $K_2S_2O_5$ was added until wines reached 25 mg L^{-1} of free SO_2. All wines were bottled and stored at 10 °C until the analyses.

In order to consider and reduce the natural variability derived from the vinification process, three separate trials were vinified for each vineyard, obtaining a total of 24 wines (3 trials × 8 vineyards). The results of the analyses are reported throughout the text as

Mean ± Standard Deviation of the three trials for each vineyard. The main chemical–physical parameters of the wines, measured by means of a Winescan^{TM} instrument (FOSS, Hilleroed, Denmark), are reported in Table S4 of the Supplementary Information.

2.3. ^{1}H-NMR-Based Metabolomic Analyses of Wines

To ensure samples' integrity and to minimize the variability associated with the sample processing, the preparation step only involved the addiction of 0.1 mL of D_2O to 0.5 mL of wine. The diluted wines were then analysed by means of a Bruker Avance II 400 spectrometer operating at a proton Larmor frequency of 400.15 MHz. The ^{1}H-NMR spectra were acquired by applying the NOESYGPPS1D pulse sequence, which allowed suppression of water and ethanol signals. The suppressed signals included the singlet produced by water (at 4.81 ppm), the quartet produced by ethanol methylene protons (at 3.60–3.68 ppm) and the triplet produced by ethanol methyl protons (at 1.15–1.20 ppm). The acquiring temperature was 25 ± 1 °C. Free induction decays (FID) were collected with a 64k time domain, with a spectral width of 8012.82 Hz, a relaxation delay of 4 s and an acquisition time of 4 s. In total, 128 scans, 4 dummy scans, a shaped pulse of 0.08 mW and a mixing time of 0.01 s were used. A line broadening of 0.3 Hz was applied to the exponential function prior to the Fourier transformation. D_2O was used to optimize the field frequency lock. No quantitative internal standard was used. No artificial pH adjustment of wines was carried out with the aim to avoid any kind of modification of the matrix. As a matter of fact, wine pH could affect the chemical interactions among molecules (e.g., H bonding). Therefore, its correction could affect the results of the study. Spectra were manually phased and the baselines were manually corrected by means of the Whittaker smoother method. These operations were performed by using MNova 14.2.3 software (Mestrelab Research, Santiago de Compostela, Spain). To correct vertical scale errors deriving from the residual water and ethanol signals, spectra were normalized to the total spectral area after having removed the spectral regions containing the water protons signal (at 4.81 ppm), the ethanol methyl protons signal (at 1.10–1.20 ppm) and the ^{13}C satellites of ethanol (at 0.97–1.05 ppm and 1.28–1.36 ppm). The spectral region containing the ethanol methylene protons signal (at 3.60–3.68 ppm) was not removed because this region hosts signals produced by other wine metabolites and its removal obstructs the identification of some compounds.

Wine samples were identified as Wine 1–8, where wines from 1 to 4 were Nero d'Avola red wines and wines from 5 to 8 were Grillo white wines. The numbering of each wine corresponds to the vineyard numbering reported above.

Untargeted and Targeted Metabolomic Approaches

In the elaboration of ^{1}H-NMR spectroscopic data, two different approaches were applied, namely the untargeted and the targeted approach. The untargeted analysis gave the metabolic fingerprinting by processing the entire ^{1}H-NMR spectra. The targeted analysis provided the wine profiling by identifying and quantifying an ensemble of metabolites. Both type of analyses have been already described in Bambina et al. [26]. The input variables for the untargeted analysis were generated via bucketing the spectra. The bucketing, performed by means of MNova software, allowed division of each spectrum into segments of constant width, called buckets. This operation enabled reduction of the large number of original data points (about 64,000 data points for each spectrum) into a smaller number of input variables. The bucketing operation was performed in the spectral range 0.50–9.70 ppm by using the width of 0.01 ppm. This operation yielded about 900 spectral buckets, which were used as input variables for subsequent chemometric analysis.

To perform the targeted analysis, the signals assignment and the consequent metabolites identification were carried out by comparison with the spectral information (i.e., chemical shift, fine structure, and coupling constant) of the spectra of pure compounds. The comparison was semi-automatically performed by using the Simple Mixture Analysis (SMA) plug-in of the MNova software. This identifies single compounds inside complex mixtures by building specific libraries with the spectra of the pure compounds.

The library for wine metabolomic analysis was developed by collecting the spectra of about 60 pure compounds downloaded by the Natural Products Magnetic Resonance Database (NP-MRD) [5]. The spectral information used for the identification of compounds is listed in Table S5 of the Supplementary Information. A chemical shift range, rather than a single specific value, was used for compound identification to account for the possible chemical shift dispersion deriving from different wines' pH and from different interactions among wine molecules. A centroid tolerance of 0.10 ppm was applied.

The identification of minor compounds (such as polyphenols) was carried out by applying the global spectral deconvolution method (GSD) to deconvolute overlapping signals in crowded spectral regions.

Identified metabolites were semi-quantified as percentage of the assigned peaks on the total spectral area, by using the qNMR plug-in of MNova software.

By means of the described protocol, up to 58 different metabolites were identified in Nero d'Avola and Grillo wines, far more than the number of metabolites simultaneously identified in other ^{1}H-NMR-based metabolomic studies of wines.

2.4. Chemometrics

In order to extract meaningful information and to identify patterns inside the datasets, Hierarchical Clustering Analysis (HCA), Principal Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA) were performed, which enable the exploration, the visualization, and the interpretation of the multivariate data.

Hierarchical Clustering Analysis (HCA) is an unsupervised pattern recognition technique that aims at identifying metabolite patterns and to group similar samples into clusters based on their chemical similarity, by means of a hierarchical tree-like structure (called a dendrogram) [30]. It also allows visualization of the relative variation of the concentration of the metabolites. The sample clusters were obtained by applying the Ward's minimum variance method, which works according to the minimum within-cluster variance.

PCA is an unsupervised pattern recognition method primarily used for dimensionality reduction, data visualization and feature extraction. This method transforms the original variables into a new set of uncorrelated variables, called principal components (PCs), that capture the maximum variance of the dataset [31].

PLS–DA is a supervised method designed for classification and discrimination tasks. It focuses on finding latent variables (LVs) that maximize the covariance between predictor variables and class labels. The quality and the robustness of a PLS–DA model is tested by the cross validation (CV) test, the number of misclassifications (NMC) and the Area Under the Receiver Operating Characteristics (AUROC). The CV test provides the values of accuracy, R^2 and Q^2. The accuracy indicates the overall correctness of the predictions, the R^2 (also known as coefficient of determination) indicates the goodness of the fitting regression model, and the Q^2 is an estimate of the predictive ability of the model. The NMC measures the number of misclassified samples by constructing a confusion matrix. The AUROC measures the overall discriminative ability, or predictive performance, of a classification model across different decision thresholds. The ROC curve plots the sensitivity, defined as the number of true positives measured as percentage of all positives, against the 1-specificity, i.e., the number of false positives expressed as percentage of all negatives. An AUROC value near to 1 indicates a great discrimination performance, while an AUROC value of =0.5 indicates the absence of discriminative power [32].

The statistical significance of the model is measured by the Permutation test [33]. In this study, the PLS-DA was applied in a double cross validation scheme. This means that the diagnostic tools were used not only to determine the quality of the discrimination model but also to optimize the hyperparameters of the model, i.e., the optimal number of LVs. The number of LVs was chosen in order to preserve the best balance between the predictive performance and the interpretability of the model.

In order to ensure that the analyses were not biased by differences in the scales of the concentrations of different metabolites, the chemometric elaborations were preceded by the

preprocessing step, consisting of data normalization and scaling. While the normalization turns the skewed distributions into Gaussian-like curves, the scaling reduces the relative importance of the large fold changes on small fold changes, keeping the data structure almost unchanged [34]. The type of data preprocessing was chosen after the visual evaluation of the data distribution before and after normalization and scaling. All the statistical elaborations were performed by means of the MetaboAnalyst 6.0 web-based tool.

3. Results and Discussion

3.1. ^{1}H-NMR-Based Metabolomics to Classify Wines According to Different Grape Varieties

3.1.1. Wines Fingerprinting

The untargeted analysis led to the fingerprinting of Grillo and Nero d'Avola wines by processing the entire ^{1}H NMR spectra. The stacked spectra of the analysed wines are reported in Figure 1. The spectra are divided into different regions. Namely, between 0 and 2 ppm, the proton Larmor frequencies of simple aliphatic protons are observed; between 2 and 3 ppm, the resonance of protons adjacent to unsaturated groups occurs; between 3 and 6 ppm, protons adjacent to electronegative atoms and vinyl protons resonate; between 6 and 9 ppm, the resonance of the aromatic protons is observed and, finally, between 9 and 10 ppm, the resonance of the aldehydic protons occurs.

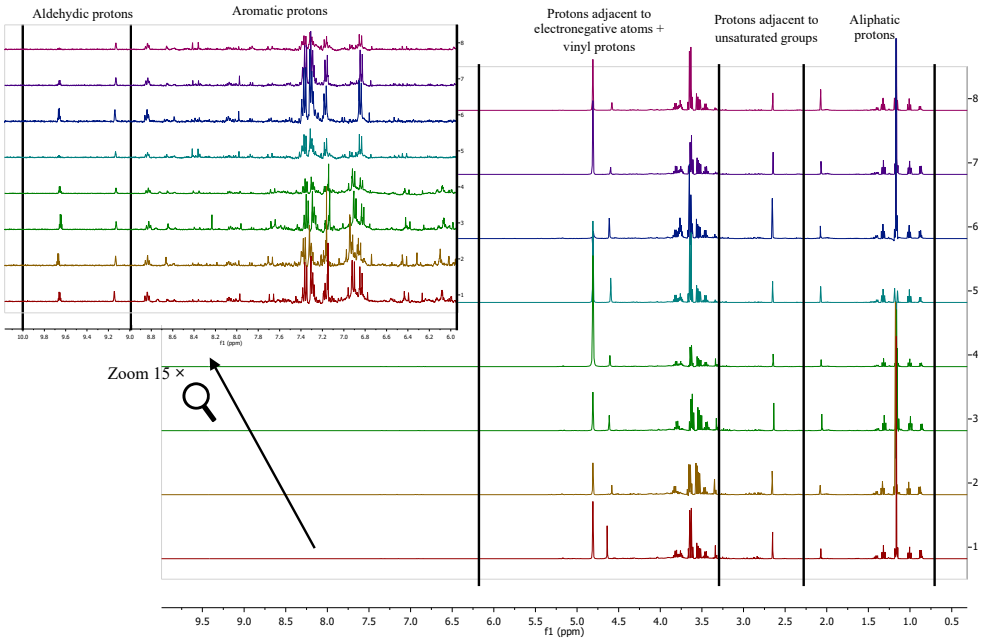

Figure 1. Stacked ^{1}H-NMR spectra of the analyzed wines. In the figure, the ranges of Larmor frequencies of chemically different protons are indicated. The numbering at the right of the plot indicates the wine samples as explained in Materials and Methods.

Prior to chemometric analysis, the fingerprinting data were auto-scaled (mean-centered and divided by the standard deviation of each variable). Then, the PCA was carried out, which yielded two clearly distinguishable groups. As can be observed in the 2D PCA Scores Plot of Figure 2A, Nero d'Avola red wines and Grillo white wines are grouped separately. The spectral fragments responsible for the separation are listed in Table 1. The table reports the contribution of each original variable (spectral fragment) to the PCs, helping the interpretation of the spatial separation among wines.

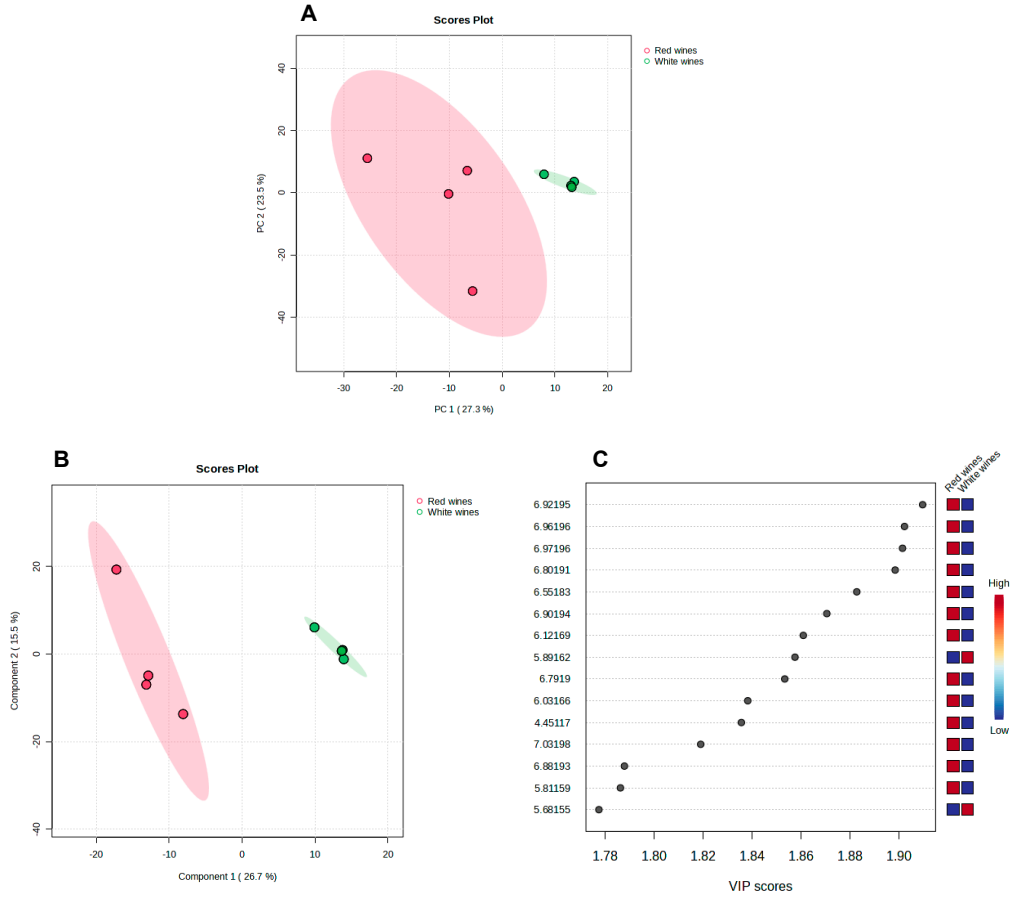

Figure 2. (**A**) 2D PCA Scores Plot showing the spatial separation among Nero d'Avola and Grillo wines as based on untargeted analysis. The explained variance is shown in brackets. (**B**) 2D PLS-DA showing the spatial separation among Nero d'Avola and Grillo wines. (**C**) Variable Importance in Projection scores (VIPs) providing the measure of the importance of each variable (in this case, of each spectral fragment) in the discrimination among wine groups.

The PC1 was mostly driven by simple aliphatic protons (which can be attributed to aroma compounds and amino acids), protons adjacent to unsaturated groups (belonging to organic acids), protons adjacent to electronegative atoms and vinyl groups (attributable to carbohydrates and polyols), and aromatic protons (proper to polyphenols). The PC2 was mostly driven by protons adjacent to electronegative atoms and vinyl groups, and aromatic protons. The separation and classification of wines obtained with the PCA was mostly due to the protons from the most concentrated area (e.g., aliphatic area). Some authors have already reported wine classification models based on differences on amino acids, organic acids, and polyols [14–16]. However, the fact that amino acids, organic acids, and polyols (especially glycerol) are some of the most concentrated and simply identifiable compounds via ^{1}H-NMR suggests that these results can be due to the approach used for data extrapolation, preprocessing and statistical elaboration. In fact, it often occurs that the relative variations of highly concentrated compounds overshadow small variations of lowly concentrated compounds, possibly carrying the most interesting metabolomic information. As a matter of fact, it is well recognized that some of the most important wine sensory properties, such as color and mouthfeel (i.e., viscosity, astringency, bitterness, etc.) are particularly due to low molecular weight polyphenolic compounds, that are present in much lower concentration with respect to other compounds (e.g., glycerol, polysaccharides, ethanol, organic acids, etc.). Therefore, the result of the PCA suggests that this statistical model (PCA) may not be the best solution for the extraction of significant features and the selection of possible biomarkers, despite yielding a good separation among wine types.

Table 1. Spectral fragments mostly contributing to each PC for Nero d'Avola and Grillo wine fingerprinting. The values indicating the contribution of each spectral fragment to the PCs represent the strength and direction of the relationship between the original variable and the PC.

Spectral Fragment	Contribution to PC1 (Positive Side)	Spectral Fragment	Contribution to PC2 (Positive Side)
0.89	0.06	3.43	0.06
2.08	0.06	3.45	0.06
2.37	0.06	3.52	0.05
3.67	0.06	3.61	0.05
3.75	0.06	3.70	0.05
3.76	0.06	3.71	0.06
5.68	0.06	5.42	0.07
5.89	0.06	5.93	0.05
6.52	0.06	6.17	0.06
7.18	0.05	6.18	0.05
7.23	0.06	7.12	0.05
7.24	0.05	7.26	0.06
7.39	0.05	7.28	0.06
7.88	0.06	7.35	0.06
7.93	0.06	8.58	0.06
7.95	0.06	8.94	0.07
8.29	0.06	9.26	0.07
8.38	0.06	9.28	0.05
8.39	0.06		
Spectral Fragment	**Contribution to PC1 (Negative Side)**	**Spectral Fragment**	**Contribution to PC2 (Negative Side)**
3.89	−0.07	2.67	−0.08
3.90	−0.07	2.92	−0.08
3.98	−0.07	3.35	−0.08
4.42	−0.07	3.56	−0.08
4.43	−0.07	3.58	−0.08
4.44	−0.07	3.59	−0.07
5.15	−0.07	3.93	−0.07
5.20	−0.07	5.23	−0.08
6.03	−0.07	5.24	−0.08
6.04	−0.07	5.25	−0.08
6.05	−0.07	5.26	−0.08
6.07	−0.07	5.31	−0.08
6.12	−0.07	5.44	−0.08
6.79	−0.07	5.45	−0.07
6.80	−0.07	6.15	−0.08
6.82	−0.07	6.32	−0.08
6.89	−0.07	6.64	−0.08
6.91	−0.07	7.74	−0.08
7.61	−0.07	8.19	−0.08
8.63	−0.07	8.67	−0.07

The supervised chemometric method, namely the PLS–DA, provided a clear separation among red and white wines (Figure 2B). The permutation test was performed by using n = 1000 permutations and yielded a p value < 0.001, ensuring the significance of the test. The CV test was performed by using the Leave-One-Out Cross Validation procedure (LOOCV) (Table S6a of the Supplementary Materials). By means of the CV test, 2 LVs were selected, which explained the 43% of the total variance of the matrix. For the 2 LVs model, the accuracy was 1, the R^2 was 0.992, and the Q^2 was 0.967, indicating the correctness of the predictions, the goodness of the fitting model and the predictive ability of the model. The NMC analysis (Figure S1A in the Supplementary Materials) ensured an optimal classification performance of the model, yielding 0 misclassifications. The AUROC

analysis (Figure S1B in the Supplementary Materials) provided an AUC (area under the curve) = 1, that means a great discriminative performance.

Given the statistically significant discrimination between the two types of wine, the Variable Importance in Projection scores (VIPs) shown in Figure 2C provides indicative information about the metabolic biomarkers responsible for wines differentiation based on different vine cultivars. The VIPs provide a measure of the importance of each variable (in this case, of each spectral fragment) in discriminating among the groups. This discrimination was essentially led by aromatic protons, which were more abundant in Nero d'Avola red wines with respect to Grillo white wines. This is a remarkable result, since the major differences between a red and white wine are known to lie in their polyphenolic profile (e.g., the lack of anthocyanins and the lower concentration of flavanols in white wines), which cause important implications into wines sensory characteristics, such as color and mouthfeel. This result highlighted that this analytical methodology, consisting of the combination between ^{1}H-NMR spectroscopy and supervised chemometrics, was able to detect significant variations in crowded spectral areas held by lowly concentrated compounds (namely between 6.00 and 9.00 ppm). By using the traditional metabolomic approaches [18,26,28], only big variations of highly concentrated compounds are usually considered and evaluated, whereas the variations in lowly concentrated compounds (often carrying the most important information) are left unnoticed. Thus, it can be pointed out that the strong point of this method is the ideal combination between the parameters used for data reduction, data preprocessing, and chemometric models.

3.1.2. Wine Profiling

Once the fingerprinting of the wines was obtained and the entire ^{1}H-NMR spectrum was explored, the next step in wine metabolomic characterization is the peaks annotation and the identification of metabolites. Figure 3 shows the peaks annotation in the spectrum of one of the Grillo wines. The Nero d'Avola spectrum with compounds identification has been already reported in our previous study [26].

By applying the workflow described in Materials and Methods, up to 58 different metabolites were identified in Nero d'Avola and Grillo wines. These belong to different chemical classes, including aroma compounds, organic acids, amino acids, carbohydrates, polyphenols, etc. The complete list of the identified compounds, together with the relative concentration among the different wines are shown in Figure 4. This figure represents the Hierarchical Clustering Analysis. The dendrogram at the top of figure clearly grouped the wine samples based on different grapevine varieties, according to their chemical similarities. Nero d'Avola wines showed higher concentrations of polyphenolic compounds (e.g., anthocyanins and flavanols), acetaldehyde, some amino acids (e.g., proline, threonine, arginine, valine), and some organic acids (shikimic acid, syringic acid, citric acid tartaric acid, malic acid). Grillo wines showed higher concentrations of most of the identified aroma compounds, together with a number of organic acids, including acetic acid, sinapic acid, succinic acid and caffeic acid.

The PCA yielded two clearly distinguishable groups of wines, the first one for the red and the second one for the white wines (Figure 5A). The metabolites mostly responsible for the separation are reported in Figure 5B. The positive side of PC1 was essentially driven by amino acids, aroma compounds and organic acids, while the negative side of PC1 was mostly led by phenolic compounds. The PC2 was mostly driven by organic acids, phenols, and aroma compounds. Even the PLS-DA supplied a clear discrimination among the two kinds of wine (Figure 5C). The permutation test, performed by using n = 1000 permutations, yielded a p value < 0.001, meaning that the classification model was statistically significant. The CV test, carried out by means of the LOOCV algorithm, measured that the first 2 LVs (which explained the 63% of the total variance of the dataset) showed accuracy = 1, R^2 = 0.994 and Q^2 = 0.986 (Table 2). These values ensured the exactness of the classification, the goodness of fit and a great prediction ability. The NMC analysis highlighted that no samples were misclassified in the analysis (Figure S2A in the Supplementary Materials).

The AUROC = 1 highlighted the perfect discrimination power of the model (Figure S2B in the Supplementary Materials).

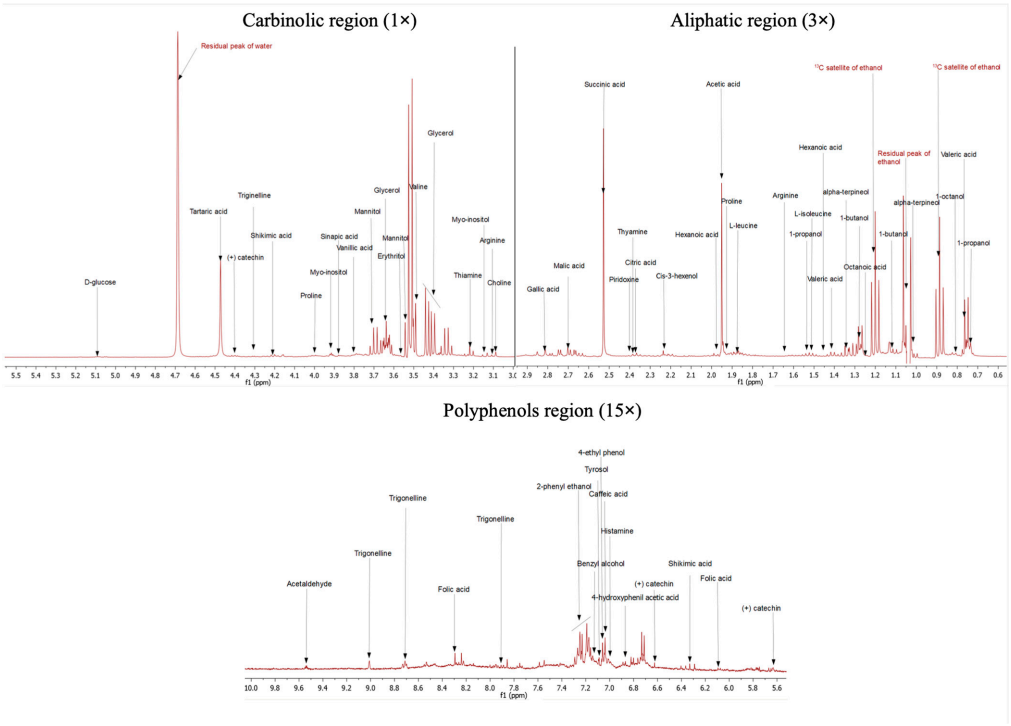

Figure 3. Peaks annotation and metabolites identification in the spectrum of one of the analysed Grillo wines.

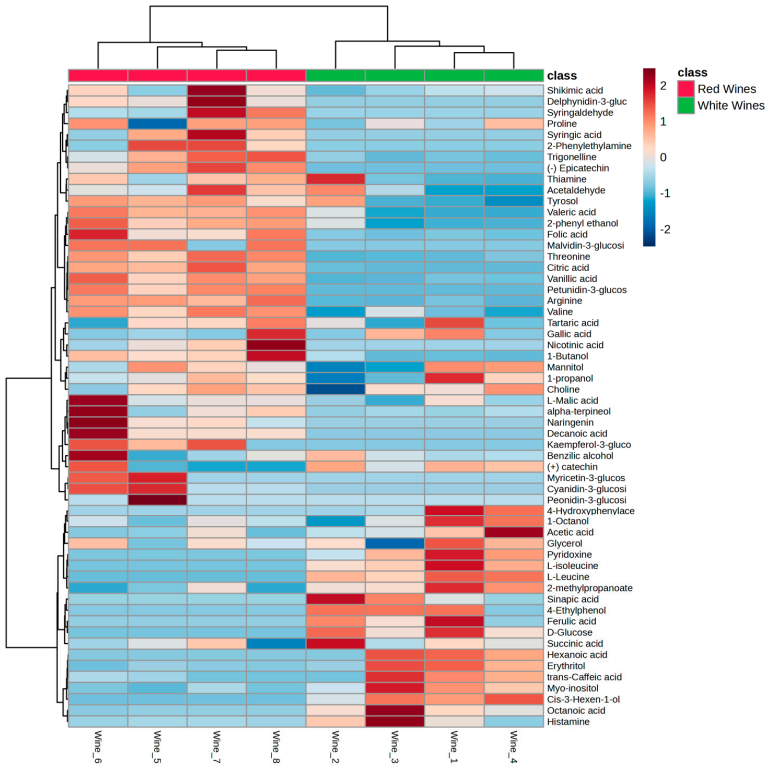

Figure 4. Hierarchical Clustering Heatmap providing the visualization of the relative variation of the concentration of the identified compounds among the different wines. Each colored cell on the map

corresponds to a concentration value and it can be used to quickly identify similarity/differences among wine samples. The dendrogram at the top of figure clearly grouped the wine samples based on different grapevine varieties, according to their chemical similarities.

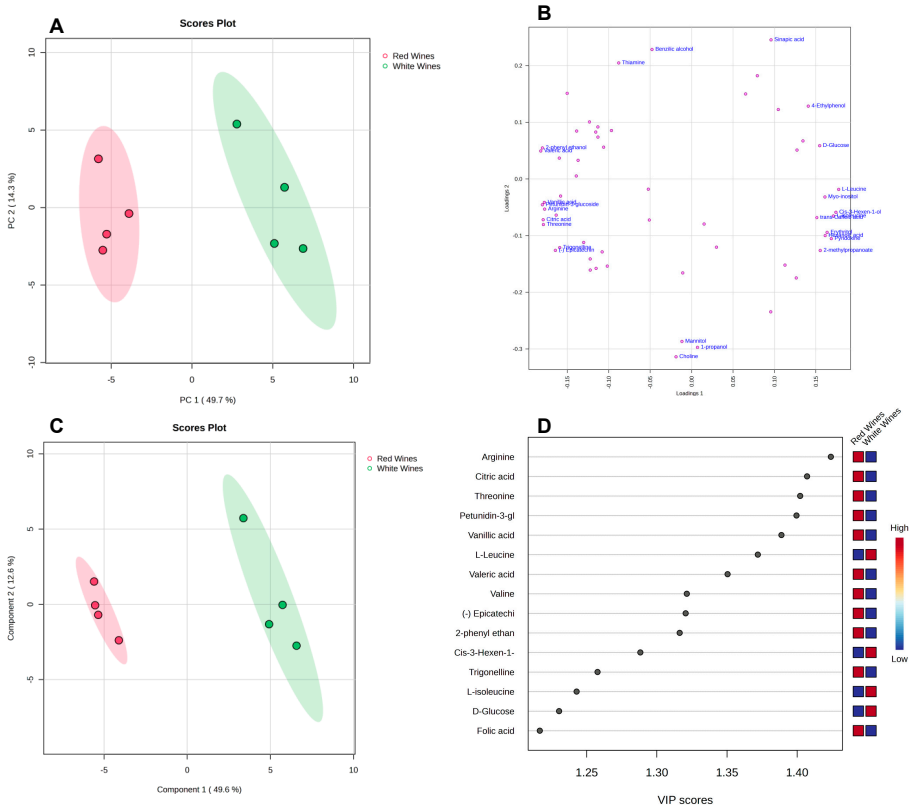

Figure 5. (**A**) 2D PCA Scores Plot showing the spatial separation among Nero d'Avola and Grillo wines as based on the targeted analysis. The explained variance is shown in brackets. (**B**) 2D PCA Loadings Plot showing the metabolites mostly responsible for wines' separation. (**C**) 2D PLS-DA showing the spatial separation among Nero d'Avola and Grillo wines. (**D**) Variable Importance in Projection scores (VIPs) providing the measure of the importance of the metabolites responsible for wines' discrimination.

The VIPs scores plot of Figure 5D highlighted that amino acids, organic acids and phenols are the compounds mostly responsible for the separation and can be considered as metabolic biomarkers.

Despite the good results obtained by both the unsupervised and supervised chemometric models in separating samples according to metabolite concentrations and in extracting the significant features, it must be noticed that the identification of compounds in a ^{1}H-NMR spectrum inevitably leads to a loss of spectral information. This is essentially because the targeted analysis involves the selective detection and quantification of signals corresponding to specific metabolites, which may lead to the neglecting of not assigned signals.

However, the great number of metabolites identified with our method with respect to the other studies reported in the literature [14,15] represents a step forward for further improvements in the targeted analysis.

Table 2. Spectral fragments mostly contributing to each PC for Grillo wines fingerprinting. The values indicating the contribution of each spectral fragment to the PCs represent the strength and direction of the relationship between the original variable and the PC.

Spectral Fragment	Contribution to PC1 (Positive Side)	Spectral Fragment	Contribution to PC2 (Positive Side)
5.35	0.09	2.25	0.12
5.48	0.10	2.27	0.12
5.49	0.09	2.28	0.12
6.50	0.09	2.29	0.12
6.76	0.10	2.56	0.11
6.77	0.10	2.57	0.12
7.09	0.07	2.58	0.12
7.20	0.07	2.59	0.12
7.49	0.10	2.60	0.12
7.55	0.07	2.61	0.12
7.58	0.08	2.71	0.12
7.69	0.07	2.72	0.11
7.73	0.08	2.73	0.12
7.91	0.09	3.00	0.12
7.94	0.09	3.01	0.11
7.96	0.09	3.17	0.11
8.43	0.09	3.18	0.12
9.05	0.08	3.19	0.12
9.07	0.08	5.46	0.11
9.08	0.09	6.75	0.11
Spectral Fragment	**Contribution to PC1 (Negative Side)**	**Spectral Fragment**	**Contribution to PC2 (Negative Side)**
1.37	−0.11	5.44	−0.11
1.38	−0.12	5.82	−0.10
1.45	−0.12	5.85	−0.10
1.47	−0.12	6.46	−0.10
1.49	−0.11	6.74	−0.11
1.50	−0.11	6.97	−0.10
1.51	−0.11	7.00	−0.11
1.52	−0.11	7.02	−0.10
1.54	−0.11	7.83	−0.09
1.56	−0.11	7.84	−0.11
1.57	−0.11	8.19	−0.11
1.58	−0.11	8.23	−0.10
1.59	−0.11	8.24	−0.10
1.60	−0.11	8.31	−0.10
1.61	−0.11	8.46	−0.11
1.63	−0.11	8.57	−0.09
1.68	−0.11	8.61	−0.11
1.71	−0.11	8.68	−0.09
1.73	−0.11	8.69	−0.10
2.74	−0.11	8.79	−0.11

The Study of the H-Bond Network

The overall wine metabolic composition is very complex. For this reason, when the sensory properties of wine are considered, the interactions between the different wine components must be accounted for. The role of the interactions among compounds influencing food sensory properties has been already reported in some studies [20,35]. However, only the interactions among a few components were simultaneously studied. This may lead to the underestimation of important information.

As already introduced in our previous paper [26], the study of the ^{1}H-NMR chemical shift dispersion allows the obtaining of information about the H-bond network where the resonant protons are involved. In particular, protons involved into strong hydrogen bonds

typically exhibit higher chemical shift values due to a major de-shielding effect. In fact, when a proton participates in a strong hydrogen bond, its electron density decreases as it is shifted towards the more electronegative atom [36].

Figure 6 shows the chemical shift dispersion is some spectral areas, such as the alkyl zone (A), the unsaturated groups area (B) and the phenolic area (C). From the three sections mentioned above, it can be observed that the spectral peaks differ not only in terms of intensity (that reflects the abundance of resonant protons and, therefore, the concentration of the molecule generating the signal), but also in terms of chemical shift values. This leads to the conclusion that the H-bond network in which the showed protons are involved is different.

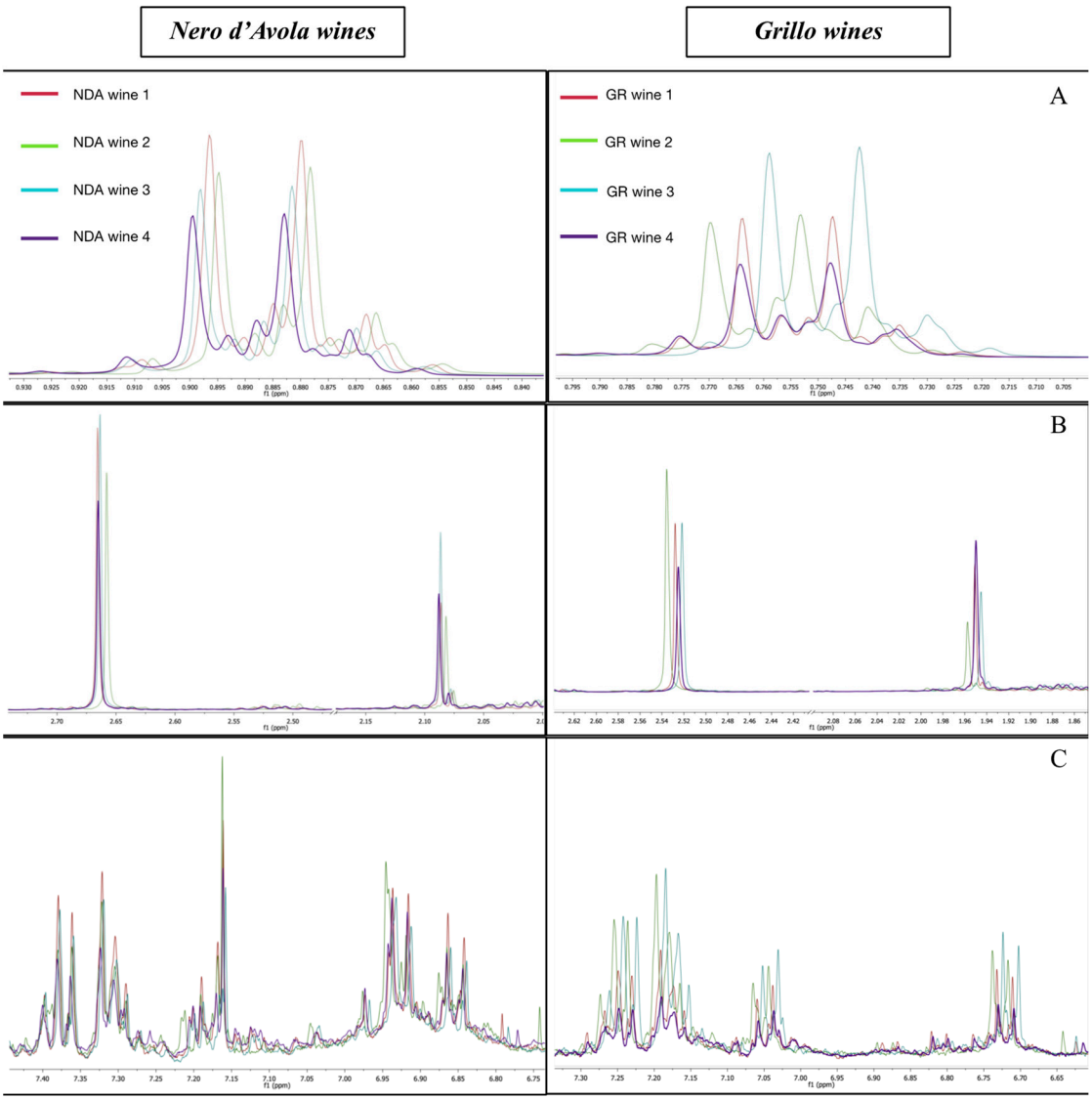

Figure 6. Chemical shift dispersion in the aliphatic zone (**A**), in the unsaturated groups area (**B**) and the phenolic area (**C**) in Nero d'Avola (left) and Grillo (right) wines' spectra. The spectral peaks differ not only in terms of intensity but also in terms of chemical shift values.

For example, the signals of Nero d'Avola 4 in the alkyl zone (A) are shifted towards higher chemical shift values with respect to other wines, while the signals of Grillo 3 are shifted towards higher chemical shift values in the three zones showed in figure. This suggested that the compounds in both Nero d'Avola 2 and Grillo 3 are involved in a stronger H-bond network with respect to the other wines.

This has important implications into wine sensory properties. Interactions, such as H-bonds, may occur between volatile aroma compounds and other non-volatile systems, including carbohydrates, organic acids, and polyphenols. Due to this, aroma compounds may become less volatile and, hence, less readily released into the gaseous phase. As a result, the concentration of aroma compounds in the wine headspace is lower, thereby leading to a reduced perceived aroma intensity. Moreover, the complexation of aroma compounds with other wine components can alter their sensory properties, masking or modifying their characteristic aromas [37,38].

3.2. ^{1}H-NMR-Based Metabolomics to Discriminate Wines from Different Terroirs

As already introduced above, the second part of this study deals with the possible capability of ^{1}H-NMR-based metabolomic analysis to explore how the metabolome of Grillo wines can be affected by different terroirs and, in particular, by the different soil types in which the grapes have grown.

The PCA analysis applied to the spectral fragments obtained by the untargeted analysis reduced the number of original variables into 2 PCs that accounted for 80% of the total variance of the data set. Figure 7A shows the 2D PCA Scores Plot, reporting a great separation among Grillo wines derived from grapes grown on different soils. The spectral fragments leading the two selected PCs together with their loading's values (i.e., the measure of the contribution of each spectral fragment to the selected principal components) are listed in Table 2.

The spectral fragments leading the positive side of PC1 are essentially aromatic protons (i.e., polyphenols), while the negative side of PC1 is mainly led by aliphatic protons (i.e., aroma compounds and amino acids); the spectral fragments leading the positive side of PC2 are protons adjacent to unsaturated groups, protons adjacent to electronegative atoms and vinyl groups (i.e., carbohydrates and polyols); the spectral fragments leading the negative side of PC2 are aromatic protons.

Even the PLS-DA provided a great separation among Grillo wines (Figure 7B). The empirical *p*-value obtained from the Permutation test (n = 1000 permutations) was< 0.001 (Table S6c in the Supplementary Materials). The CV test suggested that the optimal number of LVs was 3, showing accuracy = 1, R^2 = 1, and Q^2 = 1. The 3 LVs model explained the 100% of variance of the dataset. Given that both NMC and AUROC analyses are optimized to evaluate the performance of binary classification models, the One-vs-All (OvA) approach was applied in order to extend the analyses to the multi-classes classification models. Therefore, one random wine class was treated against the other classes (considered as a single class). The NMC analysis highlighted the absence of misclassifications of the samples, while the AUROC value was 0.876 (Figure S3A,B of the Supplementary Materials). These results indicated the correctness of the classification model and its good predictive ability.

Figure 7C reports the Variable Importance in Projections scores (VIPs), indicating the most important variables leading the separation among wines. These were mainly represented by aromatic protons (i.e., polyphenols) and protons adjacent to unsaturated groups (i.e., organic acids).

With the protocol described for compounds identification and quantification, up to 44 compounds were identified in Grillo wines.

Chemometric elaborations of the dataset obtained by the target analysis (namely PCA and PLS-DA) revealed that the metabolites mostly responsible for the separation included organic acids, amino acids, and aroma compounds (Figure 8B,D), thus confirming the results obtained by the elaboration of the untargeted data. The results of the diagnostic tests are visible in Table S6d and in Figure S4 of the Supplementary Materials.

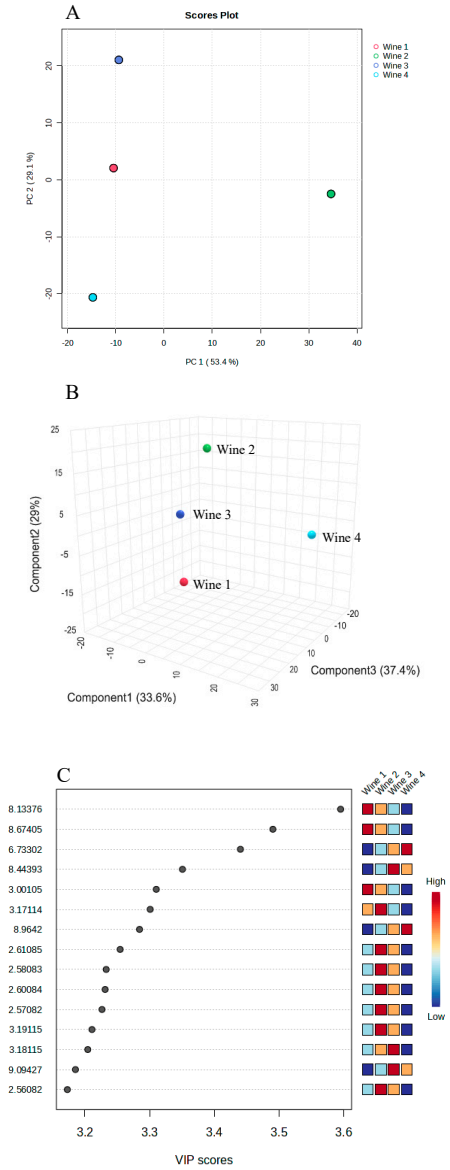

Figure 7. (**A**,**B**) 2D PCA Scores Plot and PLS-DA Scores Plot, respectively, highlighting the discrimination among Grillo wines derived from different soils obtained by the untargeted analysis. (**C**) Variables Importance in Projection Scores (VIPs).

Experiences of how different soil managements affect the volatile [39,40] and the non-volatile fraction of wines [41,42] are already reported in the literature. Mackenzie et al. [43] found correlations between several plant-available trace elements in soil and sugar content and titratable acidity in grapes, while Blotevogel et al. [25] observed that the soil type, together with the climatic conditions, systematically influenced the elemental composition of wines. Van Leeuwen et al. [21] assessed that the soil modulates vine development and grape ripening through the soil temperature, and the water and nutrient supply. While soil temperature seemed to have an effect on vine phenology, the water and nutrient supply appeared to affect the accumulation of secondary metabolites, such as polyphenols and aroma compounds.

The following section reports a discussion on the relationships between the soil features and the metabolic profiles of Grillo wines observed in this study.

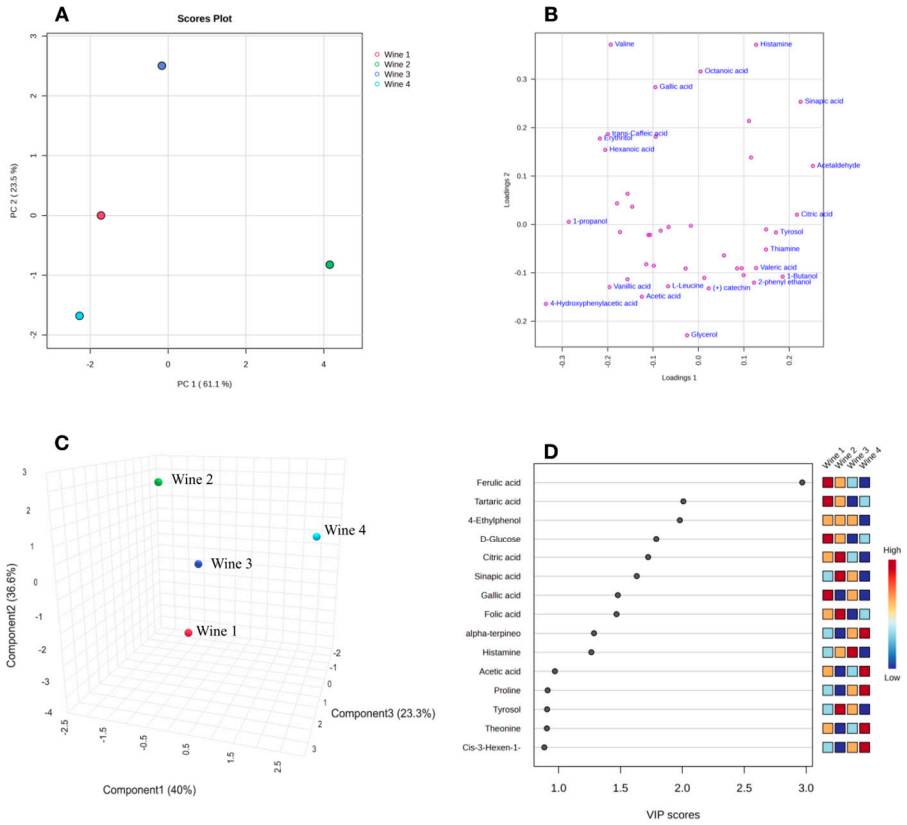

Figure 8. (**A**,**B**) 2D PCA Scores Plot and Loading Plot for Grillo wines targeted analysis. (**C**,**D**) PLS-DA Scores Plot and Variables Importance in Projection Scores (VIPs) for Grillo wines targeted analysis.

3.3. Correlations between Soil Features and Grillo Wines' Metabolic Profile

To analyze the effect of the soil on the determination of Grillo wines' metabolome, the relative concentrations of the metabolites obtained via the ^{1}H-NMR targeted analysis and the chemical–physical parameters of the soils were subjected to correlation analysis. The Pearson's correlation coefficient, also indicated as r, measures the strength and direction of a linear relationship between two continuous variables. The r coefficient ranges from −1 to +1, where +1 indicates the perfect positive linear relationship, −1 indicates the perfect negative linear relationship, and 0 indicates the absence of linear relationship. From a practical point of view, a positive linear relationship means that, as one soil parameter increases, the metabolite concentration in wine also tends to increase; on the contrary, a negative linear correlation means that as one soil parameter increases, the metabolite concentration in wine decreases.

Figure 9 reports the correlation heatmap graphically representing the correlation matrix between soil and wine features. Each cell in the heatmap represents the correlation between two combinations of variables. Positive correlations are indicated in red, while negative ones are reported in blue.

The color intensity of the cell reflects the strength and direction of the correlation. Only the significant correlations are reported (Pearson's r correlation coefficient > |0.90|, p value < 0.05). It can be observed that the clay and silt contents in soils established negative correlations with D-glucose, tartaric acid, 1-propanol, and L-isoleucine. Conversely, sand content established positive correlations with L-isoleucine, tartaric acid, and malic acid.

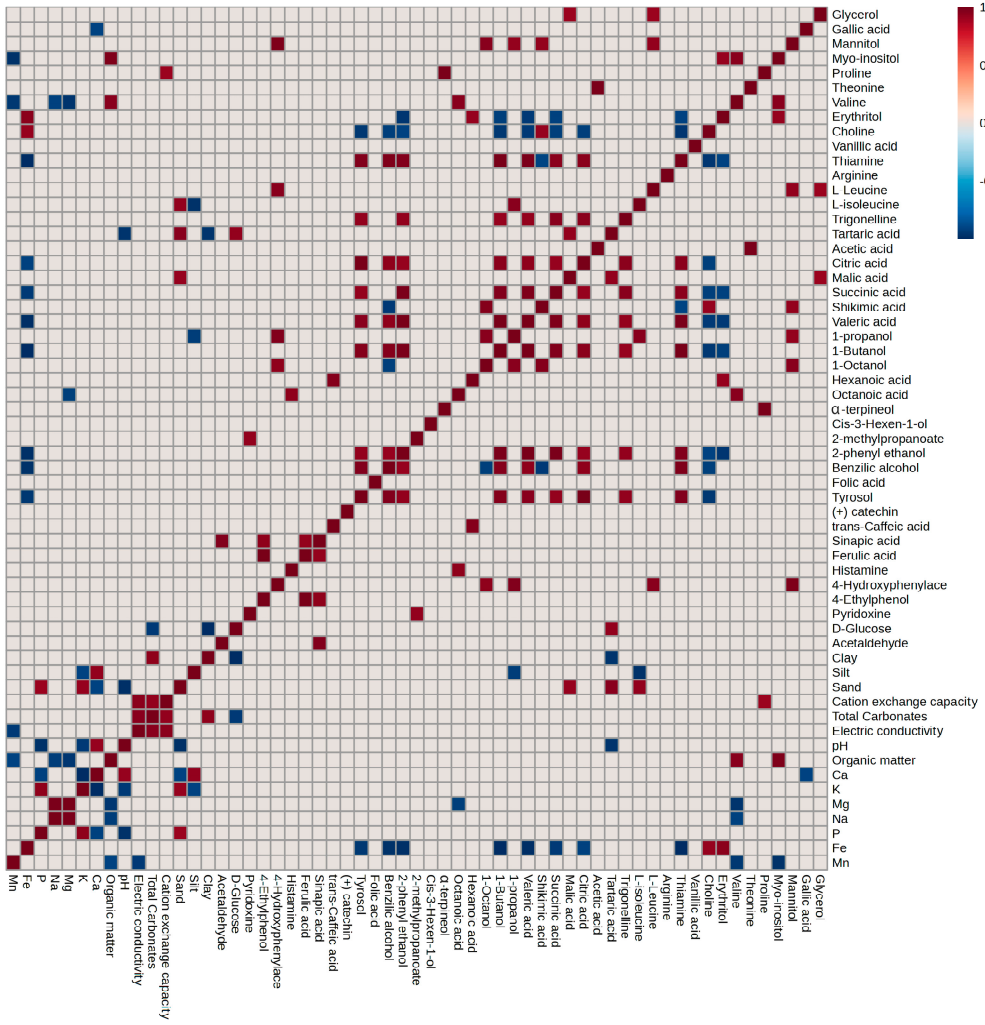

Figure 9. Correlation heatmaps reporting correlations between soil chemical physical parameters and wines' metabolomic composition. Only correlations with Pearson r correlation coefficient > |0.90| are highlighted in the figure. Positive correlations are indicated in red, negative correlations are indicated in blue.

Soil texture, i.e., the relative amounts of sand, silt, and clay particles in soil, plays a crucial role in influencing water dynamics, particularly affecting water retention, drainage, and the movement of water within the soil profile. Soils with high clay content have high surface area and small pores, also called residual pores (diameter $\leq$ 0.5 µm) [44]. Water molecules are adsorbed on the negatively charged surface of clay particles and are held against the force of gravity. Moreover, due to electrostatic forces, water molecules move slowly through the soil profile, contributing to better moisture availability for plants over time. However, excessive clay content can lead to poor drainage and potential waterlogging issues.

Sandy soils are characterized by larger particles, lower surface area, and bigger pores as compared to prevalently silty and clayey soils. The large pores in sandy soils are also indicated as transmission pores (diameter $\geq$ 50 µm) [44]. Here, water molecules move rapidly. Therefore, water tends to drain towards deeper soil layers, thus quickly drying the soil. For this reason, sandy soils have lower water retention capacity and good drainage. However, this also leads to faster nutrient leaching. Finally, silty soils have intermediate particles size (as compared to the aforementioned soils) and moderate ability to retain water. They do not retain water molecules as strongly as clayey soils but hold water better than sandy soils. The packing of silty particles generates pores that are adequate for water

retention. They are also referred to as storage pores (diameter 0.5–50 μm) and make water available for plant uptake.

In light of the above, it appears evident that soil texture can significantly influence vine growth, grape quality and, ultimately, wine production. In fact, water is the medium for nutrient transportation towards plant roots. In particular, insufficient water availability limits nutrient uptake, thus affecting plant growth. Differences in plant vegetative development may modify the intensity of the sunlight irradiation in the cluster area. Consequently, biological processes involving sunlight-sensible metabolites (such as sugar accumulation and organic acids degradation) are greatly affected. This can explain the aforementioned correlations. Moreover, the observed results accord with what has been reported by some authors [40,45]. Namely, moderately reduced water regimes induce an increase of the accumulation of several metabolites in grape berries.

High CEC was associated with high concentration of proline. High total carbonates seemed to decrease the content of pyridoxine. Slightly basic pH determined low concentration of tartaric acid in wines. Finally, high contents of organic matter in soils seemed to produce high concentrations of valine and myo-inositol. All the aforementioned soil parameters are related to the amount and availability of nutrients. As a matter of fact, CEC is the capability of a soil to retain positively charged nutrients by electrostatic forces, while EC is the measure of the total ion amount dissolved in the soil solutions. The pH strongly affects many soil biogeochemical processes, such as the modulation of nutrient availability, the mineralization of organic matter, ammonia volatilization, dissolution and precipitation of organic matter and metals, nitrification, and denitrification. Finally, organic matter affects soil structure, soil water retention capacity and nutrient mobilization, playing an important role in role in soil health and fertility. The influence of the nutrient dynamics is further confirmed by the correlation patterns observed between soil nutrients and wine chemical composition. Ca was negatively correlated with gallic acid, while Mg, Na and Mn were negatively correlated with valine. However, the major correlations were observed with soil Fe. In particular, Fe was negatively correlated with some organic acids (such as citric, succinic and valeric acids) and some aroma compounds (such as 1-butanol, 2-phenylethanol and benzylic alcohol). Finally, it positively correlated with erythritol and choline. Fe is an essential mineral for plant nutrition and plays a fundamental role in many biological processes, including photosynthesis, respiration, nitrogen fixation and assimilation, and DNA synthesis. Fe is also involved in the biosynthesis of plant hormones, which play specific roles in plant development and adaptative responses to environmental conditions [46]. It seems that high Fe availability shifts the equilibrium towards plant vegetative growth at the expense of metabolites accumulation in grape berries. Despite some authors [47] stating that it is difficult to establish correlations between wine quality and soil nutrients, the correlation pattern observed in this study suggested otherwise, namely that soil mineral composition, as well as soil chemical–physical parameters modulating nutrients dynamics, play a fundamental role in determining wine quality.

Finally, it is important to mention that the correlations patterns observed between soil features and Grillo wines were quite different from those observed for Nero d'Avola wines in our previous study [26]. This suggests that other site-specific factors, such as grape varieties, climate, topography, etc., also play a role in how soil characteristics interact with vine growth. This underlines the importance of the vine-environment combination in the expression of wine terroir, justifying why oenological products show an extraordinary diversity through space and time.

4. Conclusions

In this study, ^{1}H-NMR spectroscopy coupled with chemometrics was successfully applied to classify wines according to different grapes varieties and different terroirs. In particular, the metabolomic fingerprinting and profiling of wines associated to the PLS-DA technique provided the most indicative information about the metabolic biomarkers responsible for wines differentiation. The metabolic biomarkers leading the differentia-

tion between vine cultivars and terroirs were mainly phenols, aroma compounds, amino acids, and organic acids. Therefore, our revised approach allowed us to extract relevant spectral information even from the crowded spectral areas held by lowly concentrated compounds (e.g., polyphenols). This is a remarkable result given that, to date, most wine metabolomic studies considered only variations of highly concentrated compounds, while overlooking lowly concentrated compounds. It can be assessed that the strong point of our method is the combination of the parameters used for data reduction, preprocessing, and chemometric models. Moreover, the method used for peaks annotation and metabolites identification allowed identification of about 60 compounds, far more than the number of metabolites identified with other approaches. The greater number of metabolites identified with our method represents a step forward for further improvements of the targeted analysis. Finally, the study of the H-bonds network inside the wines provided information about the strength of the interactions between the different wine components. This knowledge is useful because the H-bond network in wine affect wine sensory properties by modulating the way how the solutes interact with human sensory receptors. The study of the H-bond network in wine opens new issues in the comprehension of the chemical mechanisms involved in gustatory and olfactory perceptions. Further investigations on the relationships between the H-bond structure and wine sensory properties are the subject of ongoing studies.

The findings of this study are very useful in the context of traceability and authenticity of food and beverages, providing a suitable methodology for assessment of the complete metabolic composition of the wine and the biomolecular markers responsible for the differentiation among different grape varieties and terroirs.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agriculture14050749/s1, Section S1: Soils description; Table S1: chemical–physical parameters of the soils of Nero d'Avola vineyards; Table S2: chemical–physical parameters of the soils of Grillo vineyards; Table S3: mineral composition of the soils of the Nero d'Avola and Grillo vineyards; Table S4: main chemical–physical parameters measured for each wine; Table S5: chemical shift ranges, fine structure and coupling constants used for the identification of compounds; Table S6: results of Cross Validation (i.e., accuracy, R^2 and Q^2) and Permutation tests for the PLS-DA performed in this study; Figures S1–S4: Plots of predicted class probabilities (NMC) and ROC curve for the PLS-DA models of this study.

Author Contributions: Conceptualization, P.B., P.C. and P.L.M.; Methodology: P.B., P.C. and D.F.C.M.; Investigation, P.B., G.L.P. and P.C.; Formal analysis, P.B. and A.S.; Data Curation, P.B., P.C., P.L.M. and L.C.; Writing—Original Draft, P.B.; Writing—Review and Editing, P.C., L.C., G.L.P., D.F.C.M. and P.L.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Dataset available on request from the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Jiménez-Carvelo, A.M.; Martín-Torres, S.; Ortega-Gavilán, F.; Camacho, J. PLS-DA vs. Sparse PLS-DA in Food Traceability. A Case Study: Authentication of Avocado Samples. *Talanta* **2021**, *224*, 121904. [CrossRef] [PubMed]
2. Medina, S.; Pereira, J.A.; Silva, P.; Perestrelo, R.; Câmara, J.S. Food Fingerprints—A Valuable Tool to Monitor Food Authenticity and Safety. *Food Chem.* **2019**, *278*, 144–162. [CrossRef] [PubMed]
3. Ellis, D.I.; Muhamadali, H.; Allen, D.P.; Elliott, C.T.; Goodacre, R. A Flavour of Omics Approaches for the Detection of Food Fraud. *Curr. Opin. Food Sci.* **2016**, *10*, 7–15. [CrossRef]
4. Fiehn, O. Metabolomics—The Link between Genotypes and Phenotypes. In *Functional Genomics*; Town, C., Ed.; Springer: Dordrecht, The Netherlands, 2002; pp. 155–171. ISBN 978-94-010-3903-1.
5. Wishart, D.S.; Sayeeda, Z.; Budinski, Z.; Guo, A.; Lee, B.L.; Berjanskii, M.; Rout, M.; Peters, H.; Dizon, R.; Mah, R.; et al. NP-MRD: The Natural Products Magnetic Resonance Database. *Nucleic Acids Res.* **2022**, *50*, D665–D677. [CrossRef]
6. Fuhrer, T.; Zamboni, N. High-Throughput Discovery Metabolomics. *Curr. Opin. Biotechnol.* **2015**, *31*, 73–78. [CrossRef]

7. Gika, H.; Virgiliou, C.; Theodoridis, G.; Plumb, R.S.; Wilson, I.D. Untargeted LC/MS-Based Metabolic Phenotyping (Metabonomics/Metabolomics): The State of the Art. *J. Chromatogr. B* **2019**, *1117*, 136–147. [CrossRef] [PubMed]
8. Bothwell, J.H.F.; Griffin, J.L. An Introduction to Biological Nuclear Magnetic Resonance Spectroscopy. *Biol. Rev.* **2011**, *86*, 493–510. [CrossRef] [PubMed]
9. Beckner Whitener, M.E.; Stanstrup, J.; Panzeri, V.; Carlin, S.; Divol, B.; Du Toit, M.; Vrhovsek, U. Untangling the Wine Metabolome by Combining Untargeted SPME–GCxGC-TOF-MS and Sensory Analysis to Profile Sauvignon Blanc Co-Fermented with Seven Different Yeasts. *Metabolomics* **2016**, *12*, 53. [CrossRef]
10. Tabago, M.K.A.G.; Calingacion, M.N.; Garcia, J. Recent Advances in NMR-Based Metabolomics of Alcoholic Beverages. *Food Chem. Mol. Sci.* **2021**, *2*, 100009. [CrossRef] [PubMed]
11. Martins, C.; Brandão, T.; Almeida, A.; Rocha, S.M. Metabolomics Strategy for the Mapping of Volatile Exometabolome from *Saccharomyces* spp. Widely Used in the Food Industry Based on Comprehensive Two-dimensional Gas Chromatography. *J. Sep. Sci.* **2017**, *40*, 2228–2237. [CrossRef] [PubMed]
12. Amargianitaki, M.; Spyros, A. NMR-Based Metabolomics in Wine Quality Control and Authentication. *Chem. Biol. Technol. Agric.* **2017**, *4*, 9. [CrossRef]
13. Alañón, M.E.; Pérez-Coello, M.S.; Marina, M.L. Wine Science in the Metabolomics Era. *TrAC Trends Anal. Chem.* **2015**, *74*, 1–20. [CrossRef]
14. Son, H.-S.; Kim, K.M.; Van Den Berg, F.; Hwang, G.-S.; Park, W.-M.; Lee, C.-H.; Hong, Y.-S. 1 H Nuclear Magnetic Resonance-Based Metabolomic Characterization of Wines by Grape Varieties and Production Areas. *J. Agric. Food Chem.* **2008**, *56*, 8007–8016. [CrossRef]
15. Gougeon, L.; Da Costa, G.; Le Mao, I.; Ma, W.; Teissedre, P.-L.; Guyon, F.; Richard, T. Wine Analysis and Authenticity Using ^{1}H-NMR Metabolomics Data: Application to Chinese Wines. *Food Anal. Methods* **2018**, *11*, 3425–3434. [CrossRef]
16. Godelmann, R.; Fang, F.; Humpfer, E.; Schütz, B.; Bansbach, M.; Schäfer, H.; Spraul, M. Targeted and Nontargeted Wine Analysis by 1 H NMR Spectroscopy Combined with Multivariate Statistical Analysis. Differentiation of Important Parameters: Grape Variety, Geographical Origin, Year of Vintage. *J. Agric. Food Chem.* **2013**, *61*, 5610–5619. [CrossRef]
17. Lee, J.-E.; Hwang, G.-S.; Van Den Berg, F.; Lee, C.-H.; Hong, Y.-S. Evidence of Vintage Effects on Grape Wines Using ^{1}H NMR-Based Metabolomic Study. *Anal. Chim. Acta* **2009**, *648*, 71–76. [CrossRef]
18. López-Rituerto, E.; Savorani, F.; Avenoza, A.; Busto, J.H.; Peregrina, J.M.; Engelsen, S.B. Investigations of La Rioja Terroir for Wine Production Using 1 H NMR Metabolomics. *J. Agric. Food Chem.* **2012**, *60*, 3452–3461. [CrossRef] [PubMed]
19. Solovyev, P.A.; Fauhl-Hassek, C.; Riedl, J.; Esslinger, S.; Bontempo, L.; Camin, F. NMR Spectroscopy in Wine Authentication: An Official Control Perspective. *Comp. Rev. Food Sci. Food Safe* **2021**, *20*, 2040–2062. [CrossRef] [PubMed]
20. Jones, P.R.; Gawel, R.; Francis, I.L.; Waters, E.J. The Influence of Interactions between Major White Wine Components on the Aroma, Flavour and Texture of Model White Wine. *Food Qual. Prefer.* **2008**, *19*, 596–607. [CrossRef]
21. Van Leeuwen, C.; De Rességuier, L. Major Soil-Related Factors in Terroir Expression and Vineyard Siting. *Elements* **2018**, *14*, 159–165. [CrossRef]
22. White, R.E. The Value of Soil Knowledge in Understanding Wine Terroir. *Front. Environ. Sci.* **2020**, *8*, 12. [CrossRef]
23. Coipel, J.; Lovelle, B.R.; Sipp, C.; Leeuwen, C.V. «Terroir» Effect, as a Result of Environmental Stress, Depends More on Soil Depth than on Soil Type (*Vitis vinifera* L. Cv. *Grenache noir*, Côtes Du Rhône, France, 2000). *J. Int. Sci. Vigne Vin* **2006**, *40*, 177–185.
24. de Andrés-de Prado, R.; Yuste-Rojas, M.; Sort, X.; Andrés-Lacueva, C.; Torres, M.; Lamuela-Raventós, R.M. Effect of Soil Type on Wines Produced from *Vitis vinifera* L. Cv. Grenache in Commercial Vineyards. *J. Agric. Food Chem.* **2007**, *55*, 779–786. [CrossRef] [PubMed]
25. Blotevogel, S.; Schreck, E.; Laplanche, C.; Besson, P.; Saurin, N.; Audry, S.; Viers, J.; Oliva, P. Soil Chemistry and Meteorological Conditions Influence the Elemental Profiles of West European Wines. *Food Chem.* **2019**, *298*, 125033. [CrossRef] [PubMed]
26. Bambina, P.; Spinella, A.; Lo Papa, G.; Chillura Martino, D.F.; Lo Meo, P.; Corona, O.; Cinquanta, L.; Conte, P. ^{1}H NMR-Based Metabolomics to Assess the Impact of Soil Type on the Chemical Composition of Nero d'Avola Red Wines. *J. Agric. Food Chem.* **2023**, *71*, 5823–5835. [CrossRef] [PubMed]
27. Bambina, P.; Gancel, A.-L.; Corona, O.; Jourdes, M.; Teissedre, P.-L. Soil Effect on Proanthocyanidins Composition of Red and White Wines Obtained from Nero d'Avola and Grillo *Vitis vinifera* L. Cultivars. *Food Chem.* **2024**, *443*, 138521. [CrossRef]
28. Bambina, P.; Pollon, M.; Vitaggio, C.; Papa, G.L.; Conte, P.; Cinquanta, L.; Corona, O. Effect of Soil Type on Some Composition Parameters of *Vitis vinifera* L. Cv. Nero d'Avola Grapes at Different Stages of Ripening. *Int. J. Food Sci. Technol.* **2024**, *59*, 2361–2374. [CrossRef]
29. *Keys to Soil Taxonomy*, 13th ed.; U.S. Department of Agriculture: Washington, DC, USA, 2022.
30. Rahman, M.d.A.; Rahman, M.d.M.; Asma-Ul-Husna; Haque Mollah, M.d.N. Robust Hierarchical Clustering for Metabolomics Data Analysis in Presence of Cell-Wise and Case-Wise Outliers. In Proceedings of the 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 8–9 February 2018; IEEE: New York, NY, USA, 2018; pp. 1–4.
31. Kurita, T. Principal Component Analysis (PCA). In *Computer Vision*; Springer International Publishing: Cham, Switzerland, 2020; pp. 1–4. ISBN 978-3-030-03243-2.
32. Szymańska, E.; Saccenti, E.; Smilde, A.K.; Westerhuis, J.A. Double-Check: Validation of Diagnostic Statistics for PLS-DA Models in Metabolomics Studies. *Metabolomics* **2012**, *8*, 3–16. [CrossRef] [PubMed]

33. Lee, L.C.; Liong, C.-Y.; Jemain, A.A. Partial Least Squares-Discriminant Analysis (PLS-DA) for Classification of High-Dimensional (HD) Data: A Review of Contemporary Practice Strategies and Knowledge Gaps. *Analyst* **2018**, *143*, 3526–3539. [CrossRef] [PubMed]
34. Van Den Berg, R.A.; Hoefsloot, H.C.; Westerhuis, J.A.; Smilde, A.K.; Van Der Werf, M.J. Centering, Scaling, and Transformations: Improving the Biological Information Content of Metabolomics Data. *BMC Genom.* **2006**, *7*, 142. [CrossRef] [PubMed]
35. Keast, R.S.J. A Psychophysical Investigation of Binary Bitter-Compound Interactions. *Chem. Senses* **2003**, *28*, 301–313. [CrossRef] [PubMed]
36. Pacios, L.F.; Gómez, P.C. Dependence of Calculated NMR Proton Chemical Shifts on Electron Density Properties in Proton-Transfer Processes on Short Strong Hydrogen Bonds. *J. Phys. Chem. A* **2004**, *108*, 11783–11792. [CrossRef]
37. Villamor, R.R.; Ross, C.F. Wine Matrix Compounds Affect Perception of Wine Aromas. *Annu. Rev. Food Sci. Technol.* **2013**, *4*, 1–20. [CrossRef] [PubMed]
38. Pittari, E.; Moio, L.; Piombino, P. Interactions between Polyphenols and Volatile Compounds in Wine: A Literature Review on Physicochemical and Sensory Insights. *Appl. Sci.* **2021**, *11*, 1157. [CrossRef]
39. Monteiro, A.; Lopes, C.M. Influence of Cover Crop on Water Use and Performance of Vineyard in Mediterranean Portugal. *Agric. Ecosyst. Environ.* **2007**, *121*, 336–342. [CrossRef]
40. Koundouras, S.; Marinos, V.; Gkoulioti, A.; Kotseridis, Y.; Van Leeuwen, C. Influence of Vineyard Location and Vine Water Status on Fruit Maturation of Nonirrigated Cv. Agiorgitiko (*Vitis vinifera* L.). Effects on Wine Phenolic and Aroma Components. *J. Agric. Food Chem.* **2006**, *54*, 5077–5086. [CrossRef] [PubMed]
41. Bouzas-Cid, Y.; Trigo-Córdoba, E.; Falqué, E.; Orriols, I.; Mirás-Avalos, J.M. Influence of Supplementary Irrigation on the Amino Acid and Volatile Composition of Godello Wines from the Ribeiro Designation of Origin. *Food Res. Int.* **2018**, *111*, 715–723. [CrossRef] [PubMed]
42. Wang, R.; Sun, Q.; Chang, Q. Soil Types Effect on Grape and Wine Composition in Helan Mountain Area of Ningxia. *PLoS ONE* **2015**, *10*, e0116690. [CrossRef] [PubMed]
43. Mackenzie, D.E.; Christy, A.G. The Role of Soil Chemistry in Wine Grape Quality and Sustainable Soil Management in Vineyards. *Water Sci. Technol.* **2005**, *51*, 27–37. [CrossRef] [PubMed]
44. Conte, P.; Nestle, N. Water Dynamics in Different Biochar Fractions. *Magn. Reson. Chem.* **2015**, *53*, 726–734. [CrossRef] [PubMed]
45. Esteban, M.A.; Villanueva, M.J.; Lissarrague, J.R. Effect of Irrigation on Changes in the Anthocyanin Composition of the Skin of Cv Tempranillo (*Vitis vinifera* L) Grape Berries during Ripening. *J. Sci. Food Agric.* **2001**, *81*, 409–420. [CrossRef]
46. Zhang, X.; Zhang, D.; Sun, W.; Wang, T. The Adaptive Mechanism of Plants to Iron Deficiency via Iron Uptake, Transport, and Homeostasis. *Int. J. Mol. Sci.* **2019**, *20*, 2424. [CrossRef] [PubMed]
47. Pérez-Álvarez, E.P.; Martínez-Vidaurre, J.M.; Martín, I.; García-Escudero, E.; Peregrina, F. Relationships among Soil Nitrate Nitrogen and Nitrogen Nutritional Status, Yield Components, and Must Quality in Semi-Arid Vineyards from Rioja AOC, Spain. *Commun. Soil Sci. Plant Anal.* **2013**, *44*, 232–242. [CrossRef]