

Chemical Profiling of Wines Produced in Bulgaria and Distinction from International Grape Varieties

Dessislava Gerginova and Svetlana Simova*

Cite This: *ACS Omega* 2023, 8, 18702–18713

Read Online

ACCESS |



Metrics & More

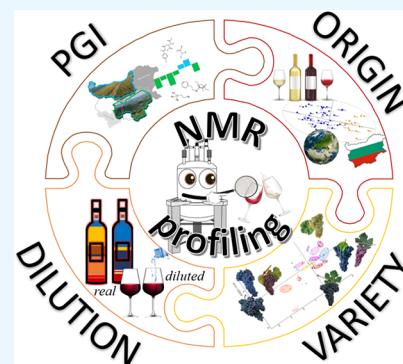


Article Recommendations



Supporting Information

ABSTRACT: Distinguishing the botanical and geographical origin of wine is important to prevent wine adulteration and to determine its quality. The combined use of ^1H NMR profiling and chemometrics allows the quantification of 31 common organic components in the NMR spectra of 70 wines from different sources. Using the NMR metabolomics approach, a successful differentiation of wines produced from Bulgarian and international grape varieties is achieved using linear discriminant analysis. Wines produced from typical local grape varieties contain higher average amounts of galacturonic, malic, tartaric, and succinic acid, alanine, choline, several alcohols, and saccharides arabinose, galactose, and sucrose than imported wine assortments. A practical decision tree is proposed for distinguishing 15 different grape varieties based on the amounts of the common wine components. An example of distinction of real from diluted wine via creation of a PLS-DA model is presented. Wines from the two subregions officially recognized by the EU at the Protected Geographical Indication (PGI) level are unequivocally recognized.



1. INTRODUCTION

Wine is a favorite alcoholic beverage and the oldest documented medicine. It contains a lot of compounds, but the composition depends on many factors, including grape variety, climate, relief, soil, production recipe, etc. Proving its authenticity requires multiple procedures, including identification and quantification of the substances contained in wine, development of a trustworthy database for distinguishing wine by grape variety, and possibilities to quickly distinguish real from diluted wine.¹ High demand, value, and variability of wine make it one of the frequently adulterated foods. Most often, fraud is related to declaration of false botanical and/or geographical origin or vintage. NMR spectroscopy is a robust and reproducible methodology, with ^1H NMR profiling often used for authentication of botanical origin and quality control of wine in European countries,² but not yet sufficiently utilized in the region of Southeast Europe.³ In Bulgaria, grape pips were found at Neolithic sites dating back to 6000 BC, and the Thracians, famous for their “sweet wine of Thrace”, inhabited the region from about 1000 BC.⁴ Bulgarian wine is famous for its aroma, taste, vinicultural traditions, and high quality, with antioxidant properties proved to be among the highest in European wines.⁵

The composition of grapes is of prime importance for the quality of the wine produced. Many compounds pass from the grape juice into the wine, and others participate in biochemical reactions with the formation of new characteristic substances. Grapes are one of the fruits whose composition is highly sensitive to the natural environment as drought and high temperatures can significantly affect the yield and final composition of primary and secondary metabolites. Light

radiation leads to an increase in the content of flavonols and anthocyanins. Apart from the grape variety and the climate in which the vines are grown, the acidity and type of the soil, the topography, the use of machine picking, the application of other techniques in growing grapes, the diseases, and the age of the vines influence the composition of the grapes (Figure 1).

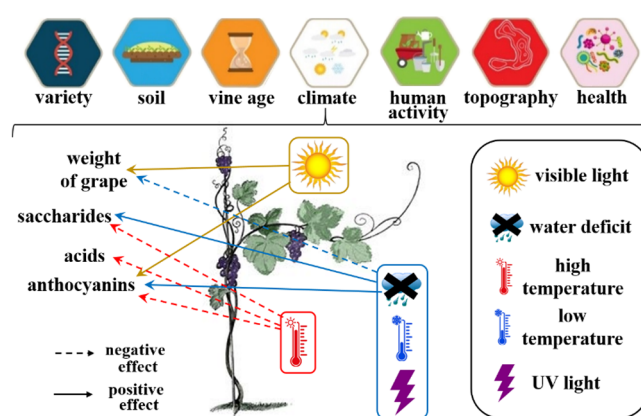
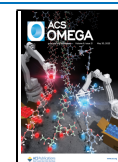


Figure 1. Main factors influencing the composition of grapes.

Received: January 31, 2023

Accepted: May 4, 2023

Published: May 16, 2023



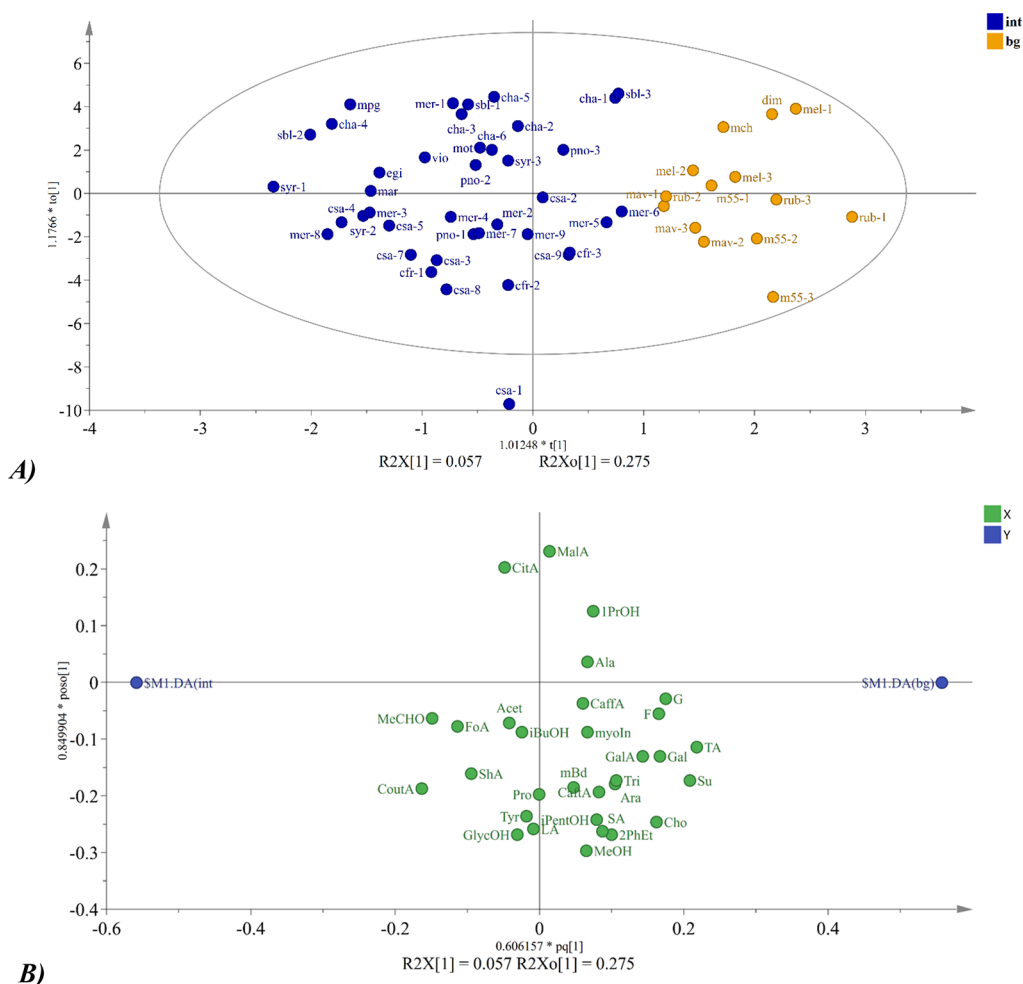


Figure 3. (A) Two-dimensional score plot of OPLS-DA analysis illustrating distinction of Bulgarian (bg) from international (int) wine grape varieties. (B) Loading plot of OPLS-DA analysis.

is highly effective for determining wine types but not yet for detection of wine fraud. Only recently, ^1H NMR spectroscopy is recognized as one of the official methods for qualitative and quantitative determination of glucose, acetic, sorbic, fumaric, shikimic, and malic acids in wine.¹⁴ Their quantification is performed using signal integration as proposed by Godelmann et al.¹⁵ Despite the lower sensitivity and higher cost, NMR has made considerable progress, particularly in large screening projects for authenticity and quality screening of food.¹⁶

Wine made from traditional Bulgarian varieties is valued by international experts,⁵ but little is known about its chemical composition. In the present work, we characterize quality and specificity of 70 wines using a metabolomics approach based on ^1H NMR profiling and chemometrics, reported in part as posters on regional conferences.¹⁷ We report a study on wines from Bulgarian and international grape varieties with the aim to differentiate them and to derive quantitative estimates to distinguish wines from 15 different grape varieties. Characterization of wines from the two officially recognized subregions by the EU at the Protected Geographical Indication (PGI) level in the country as well as differences between real and diluted wine is provided.

2. RESULTS AND DISCUSSION

2.1. Common Wine Components. Using the methodology of NMR metabolomics and the official method,⁹ we

determined the quantities of 31 common components in 70 wines—of different grape varieties and origin and three wines used as a test set (Table S1). The typical wine ^1H NMR spectrum is presented on Figure S2. Substances were from the following classes: alcohols (*meso*-2,3-butanediol, glycerol, isobutanol, isopentanol, myo-inositol, methanol, 1-propanol, and 2-phenylethanol), organic acids (tartaric, galacturonic, citric, lactic, formic, shikimic, malic, and succinic), amino acids (alanine, proline, and tyrosine), saccharides (arabinose, galactose, glucose, fructose, and sucrose), phenolic compounds (caffeic, caftaric, and coumaric acids), acetaldehyde, acetoin, choline, and trigonelline, determined in mg/L. The structural formulas of common wine components and their percentage content are visualized in Figure 2, in a periodic table analogous way.¹⁸ Due to signal overlap, the concentrations of ethanol, acetic, and gallic acids could not be reliably determined. Quercitol and valine were found in traces and not quantified. Sorbic acid and 5-hydroxymethylfurfural were not detected in the analyzed samples. The average values of all compounds determined from the 67 samples analyzed, excluding the three diluted, and the main substances are clearly distinguished from the diagram—glucose is the dominant saccharide followed by fructose, proline is the main amino acid in wine, and the higher quantities are caftaric and tartaric acids. The results are consistent with literature data, with glycerol being the second major alcohol after ethanol. The minimum, maximum, and

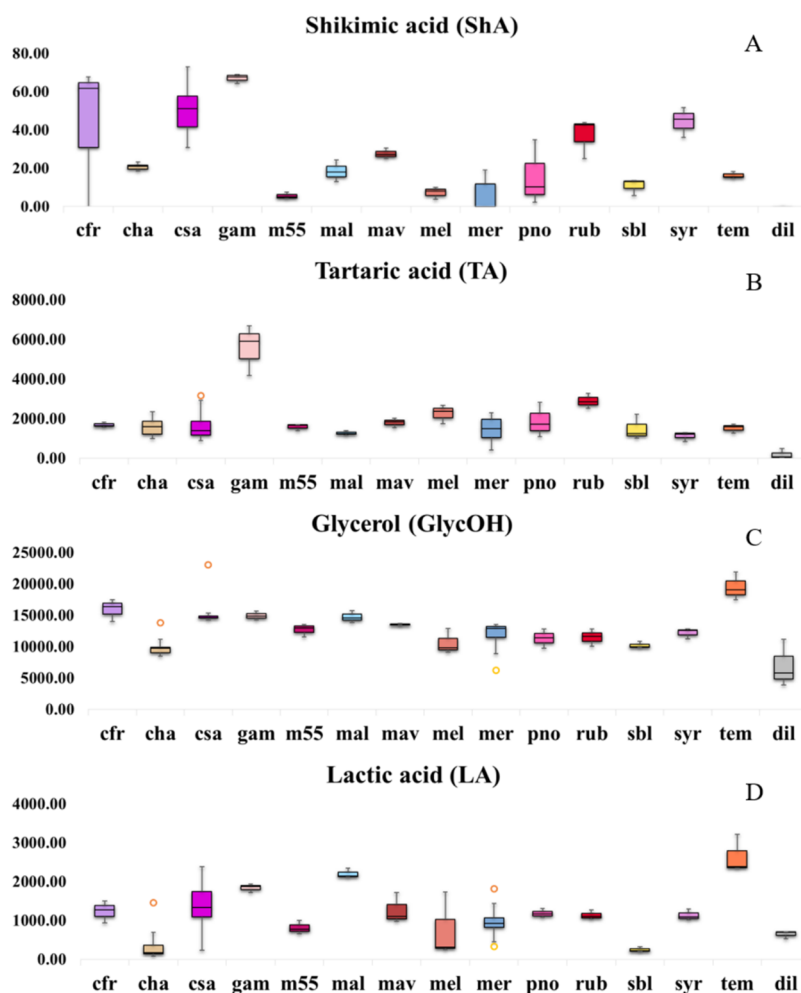


Figure 4. Box plot for (A) shikimic acid, (B) tartaric acid, (C) glycerol, and (D) lactic acid, most significant for distinguishing 15 wine types, $F_{\text{val}} > 7.00$, $p < 1 \times 10^{-7}$.

average values of 31 substances for 15 wine types, chemical shifts of the used signals, and Fisher values (F_{val}) that differentiate wines by grape variety are presented in Table S2.

Quantitative data for various wines does not differ sufficiently to allow detection of potential markers. Differences in the amounts of some components are characteristic for some wine varieties, e.g., lactic acid, higher in red wines, especially in Malbec and Tempranillo ($\text{LA} > 2117 \text{ mg/L}$), probably as a result of a more complete malolactic fermentation process. Malic acid is found in higher concentrations in white wines—Chardonnay and Sauvignon blanc—and in red wines of the variety Melnik ($\text{MaLA} > 341 \text{ mg/L}$). High levels of tartaric acid are observed in young wines of the variety Gamay from the Beaujolais region, France ($\text{TA} > 4177 \text{ mg/L}$). Comparison of our and literature quantitative data show similarities in the content of some substances, primarily alanine, 1-propanol, and 2-phenylethanol as well as differences in others, including glucose and fructose, which are climate-dependent. Variations in the component's distribution, mainly concentrations of the majority of acids (tartaric, citric, lactic, shikimic, malic, and succinic) and other substances, depend on the winemaking process, oak barrel aging, weather conditions and terrain, possible inaccuracies in the analysis method, and specific taste preferences of a nation. Chemometric techniques have to be used to identify the components

that can distinguish wines based on grape variety and geographical origin.

Comparing quantitative data from different wine types reveals that diluted wines have lower concentrations of many ingredients, including *meso*-2,3-butanediol, glycerol, isobutanol, isopentanol, methanol, myo-inositol, 1-propanol, tartaric, galacturonic, and succinic acids, proline, tyrosine, sucrose, trigonelline, and choline. Common wine components such as galactose, shikimic, caffeic, tartaric, and coumaric acids were not detected in any of the diluted samples.

2.2. Differentiation of Wines from Typical Bulgarian Grape Varieties. Chemometric methods described in the methods part, such as dispersion and discriminant analysis, were used to analyze quantitative data from the ^1H NMR spectra of wines of different origin. Distinction of Bulgarian wines from local and international grape varieties is possible by orthogonal partial least squares discriminant analysis (OPLS-DA) using quantitative data for all identified components in 54 white and red Bulgarian wine samples. The samples were split into two classes: 14 traditional wines from Mavrud, Melnik, Melnik-55, Rubin, Dimyat, and Misket cherven (bg) and 40 wines from Egiodola, Cabernet Sauvignon, Cabernet Franc, Marselan, Merlot, Pinot noir, Syrah, Viognier, Muscat Ottonel, Sauvignon blanc, Muscat blanc a Petits Grains, and Chardonnay (int). The differences of both wine types are visualized with a score and a loading plot in Figure 3.

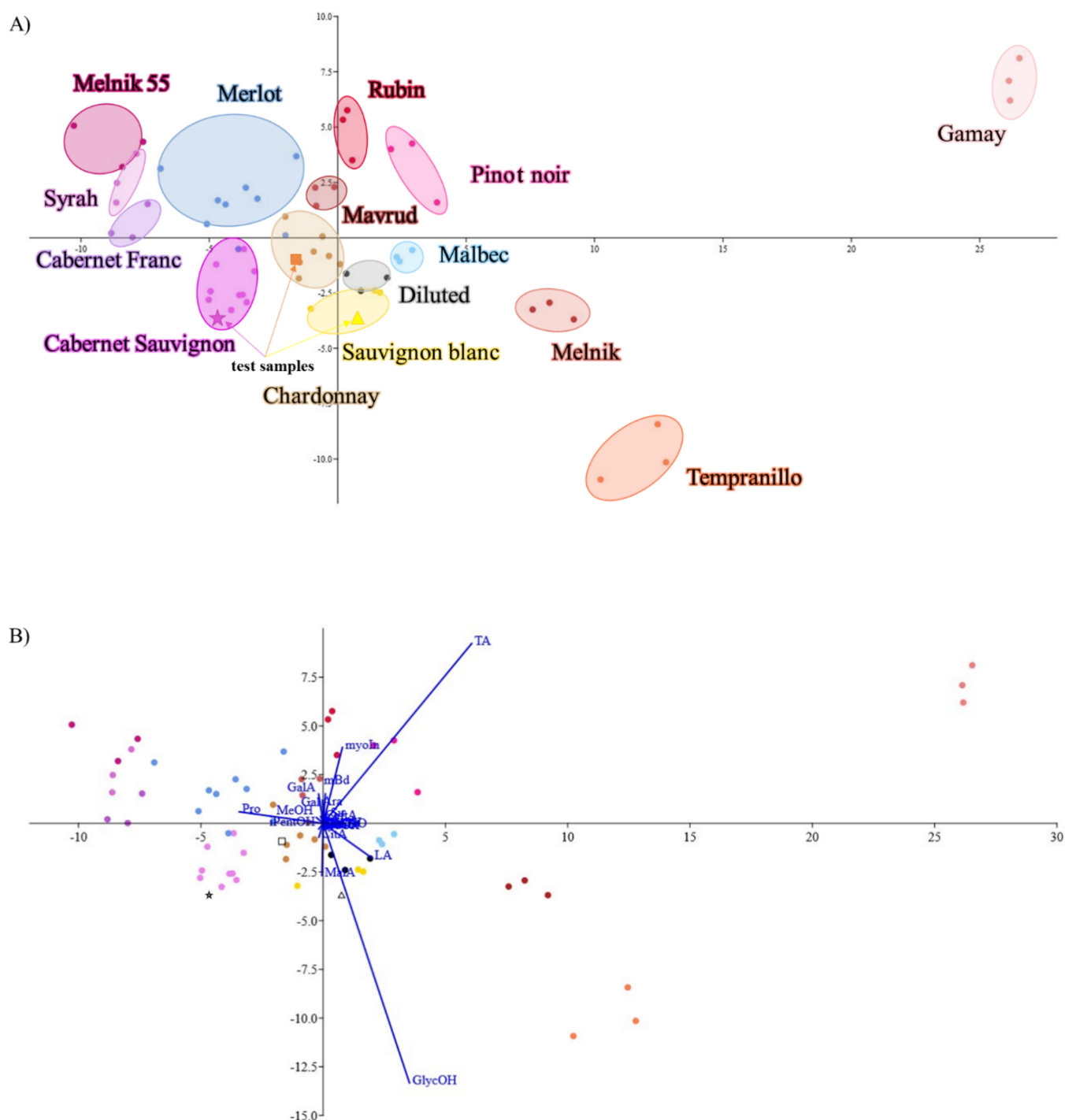


Figure 5. (A) Two-dimensional score plot of the applied LDA analysis for distinguishing 61 wine and 3 test samples according to the grape variety used. (B) Biplot derived from the LDA model.

The figure unequivocally illustrates the separation of the two groups, regardless of the wine's color or production technique. Nightingale's diagrams provide more detailed analysis presented in Figure S3 that illustrates separately the differences between white and red wines made in Bulgaria from local and foreign grape varieties.

Typical local Bulgarian wines contain more than 1.6 times more galacturonic and malic acid than imported wine varieties. Some alcohols (myo-inositol, 1-propanol, and 2-phenyl-ethanol), tartaric and succinic acids, alanine, choline, and the majority of saccharides, including arabinose, galactose, and

sucrose are also present in higher concentrations. White Bulgarian wines Dimyat and Misket cherven are characterized by the presence of large amounts of glycerol, isobutanol, methanol, citric acid, and phenolic compounds—caftaric and coumaric acids, while local red wines from Mavrud, Melnik, Melnik-55, and Rubin have higher concentrations of isopentanol, glucose and fructose, caffeic acid, and trigonelline. The quantities of some organic and amino acids such as lactic, formic and shikimic acids, proline and tyrosine, and acetaldehyde and acetoin are in lower concentrations in traditional Bulgarian white and red wines. These distinctive

features could be related to the taste preferences of Bulgarians for thick, full-bodied wines.

2.3. Differentiation of Wines by Grape Variety. One-factor analysis of variance (ANOVA) was applied to 61 of the measured samples, excluding the single types—cav, dim, egi, mar, mch, mot, mpg, san, and vio. Measured wines were divided into 15 classes, according to the grape variety used in their production, with one class for diluted wines. The influence of grape variety on the concentration of 31 substances was tested with ANOVA, using the confidence level $\alpha = 0.05$. The quantities of 26 identified components, excluding formic acid, glucose, fructose, caffeic acid, and acetoin, were found statistically significant to distinguish different wine types. Concentrations of four compounds—shikimic acid, tartaric acid, glycerol, and lactic acid with the highest F -values and p -values smaller than 1×10^{-7} —for the individual classes are visualized by box plots (Figure 4). Additional plots for the other less discriminating components are presented in Figure S4.

Box plot presentations of the quantitative data for the wine components visualize very well the essential differences between the individual wine types, e.g., it is easily seen that Gamay wines are rich in shikimic and tartaric acids, while Tempranillo contains larger amounts of glycerol and lactic acid. However, they cannot effectively distinguish between different grape varieties without applying additional chemometric techniques.

The literature provides scarce information regarding the quantification of local wine components, while a number of literature sources using mainly HPLC, suitable for phenolic compounds and certain organic acids or GC–MS used for volatile substances, including alcohols have been published. A detailed comparison of the values for the components in the measured wine grape varieties from different countries taken from the literature and measured in this study are presented in Table S3. The similarities between the data for local wines and literature data are summarized below.

Local Chardonnay varieties have close average values for lactic/galacturonic acids and glucose/fructose to Italian wines,¹⁹ probably caused by close climatic conditions in both countries. Bulgarian Chardonnay is also similar to Australian Chardonnay²⁰ in terms of malic acid and isopentanol. Methanol, 2,3-butanediol, and most higher alcohols contained in Californian wines analogously match local Chardonnay. Turkish wines²¹ show similar caffeic acid amounts. The concentrations of galactose/arabinose in French Chardonnay wines²² fall within the quantitative range of the saccharides found in the Bulgarian samples.

Sauvignon blanc wines produced in Brazil,²³ Slovenia,²⁴ and Bulgaria contain similar quantities of succinic acid. The shikimic acid concentration in Italian²⁵ and isopentanol and 2-phenylethanol in New Zealand wines²⁶ are close to the quantitative data for Bulgarian wines.

Literature data on malic, tartaric, lactic, and citric acids in wines from Chile,²⁷ caffeic and caftaric acids in Hungarian wines,²⁸ and isobutanol, isopentanol, and 2-phenylethanol in wine from Australia²⁹ match the data for local Pinot noir.

The values for caffeic and caftaric acids in local Syrah and Cabernet Sauvignon wines are similar to the literature data for Turkish,²¹ Hungarian,²⁸ and Brazilian wines,³⁰ while Chilean wines from the above two varieties have values of shikimic acid close to ours.³¹ Like quantities of alanine are found in Greek Syrah and Cabernet Sauvignon.³² Citric acid concentrations in

French, Chilean, South African, and Bulgarian Syrah wines are comparable. Cabernet Sauvignon wines from Chile,²⁷ France,³³ China,³⁴ and Bulgaria show analogous content of tartaric and lactic acids. Chinese Cabernet Sauvignon wines show similarities with Bulgarian wines in alanine/proline, glucose/fructose, 1-propanol/isobutanol, and succinic/malic acids.³⁵

The caffeic acid content of Bulgarian and Hungarian Cabernet Franc wines is similar,²⁸ and China wine similarity is noticed in the quantity of acetoin and 1-propanol.³⁶

Australian, Chinese, French, Chilean, and South African Merlot wines²⁷ show similar concentrations of tartaric, citric, lactic, and malic acids with local Merlot wines and analogous concentrations of alanine in Greek wines.³² The amount of caffeic acid in the analyzed Merlot samples is close to wines from Montenegro,³⁸ Turkey,²¹ and Hungary.²⁸ Merlot wines from Romania³⁷ and Bulgaria have close fructose content, likely due to similar temperature amplitudes.

To the best of our knowledge, only few articles are available on the composition of wine from the typical Bulgarian variety Rubin and no on Mavrud, Melnik-55, and Shiroka Melnishka loza. The values that we obtained differ significantly from the data for 1-propanol³⁹ and for tartaric and malic acid⁴⁰ in Rubin.

Quantitative composition studies of compounds in wine, except on phenolic substances and volatile components, are scarce, even though this study indicates similarities in the content of several substances, mainly alanine, 1-propanol, and 2-phenylethanol. Some components, among which glucose and fructose, are climate-dependent. This study confirms the possibility to characterize and distinguish wine produced from a specific grape variety. Differences in concentrations of most acids, tartaric, citric, lactic, shikimic, malic, succinic, and other components depend on the method of wine production, on different aging, on weather conditions and relief, on possible inaccuracy in the method of analysis, and by specific taste requirements from a given nation.

One of our goals was to differentiate the wine samples based on the grape variety from wine production. Linear discriminant analysis (LDA) provides a two-dimensional score plot, shown in Figure 5, using a 4-component model from 26 statistically significant substances with F_{val} higher F_{crit} (1.91). To validate the classification performance of the model, a test set with three samples (Chardonnay, Sauvignon blanc, and Cabernet Sauvignon) was used. The first three LDA components describe 76.42% of the variation.

The first two components are used to present the score plot, containing data only for the grape varieties with three or more samples. Figure 5A provides a clear distinction between different wine types with good sample grouping. The wine color does not affect the grouping as rosé and blue wines are classified based on their respective varieties. Although no clear distinction of all types of white from red wines is observed, the biplot shows that most wines from red grape varieties produced in Bulgaria are characterized by a higher content of proline, methanol, and gallic acid. All test samples represented by star, square, and triangle symbols were correctly grouped as Cabernet, Chardonnay, and Sauvignon. The corresponding plot for all samples is presented in Figure S5.

Despite the limited number of wines used, we created an indicative heatmap that provides information on the substances specific to individual wine types, presented in Figure 6. Red and orange colors indicate large amounts of the

component in each class, while blue colors indicate small or absent amounts.

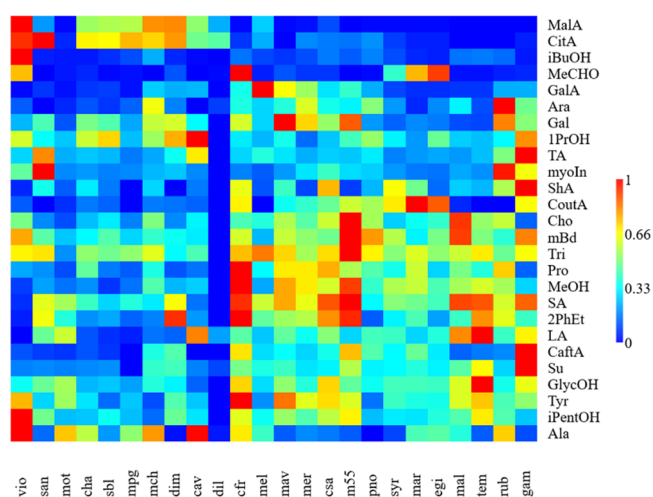


Figure 6. Heatmap of 24 different wine types, indicating quantitative information for common wine components.

A detailed characterization for each type of wine is presented below that, due to the limited sample number, is only indicative. Viognier wine differs from all others due to a very high isobutanol content of 373 mg/L, which is more than 7 times higher than the average value of 51 mg/L. White grape wines have high levels of malic (over 564 mg/L) and citric acid (over 232 mg/L). Citric acid is more than malic only in wines from Chardonnay and Muscat blanc a Petits Grains. Both wines differ in the concentration of 1-propanol, which is higher in Chardonnay. 1-Propanol is also present in large quantities in Sauvignon Blanc, Dimyat, Misket cherven, Viognier, and Cava. High levels of lactic acid (2214 mg/L) and tartaric acid (3923 mg/L) are characteristic for Cava wine, while higher amounts of higher alcohols (mBd, iBuOH, iPentOH, and 1PrOH), sugar alcohol (myoIn), and acetaldehyde are characteristic of

Viognier. The profile of white wines from traditional Bulgarian grape varieties Dimyat and Misket cherven is distinguished by significant concentrations of saccharides (Su and Gal), caftaric acid, and trigonelline. The presence of malic acid above 340 mg/L is typical for two red wines, Melnik (Shiroka Melnishka loza) and Sangiovese, as well as for Cabernet Sauvignon rosé. The Gamay variety contains high concentrations of citric, tartaric, phenolic acids (caftaric and coutaric), and shikimic acid. The latter are present in significant quantities in several red wines—Mavrud, Syrah, Cabernet Sauvignon, Cabernet Franc, and Marselan. Methanol and acetaldehyde are both present in large amounts in samples Cabernet Franc, Syrah, and Marselan. Amino acids alanine and proline are present in high concentrations in Ruben and Cabernet Franc. The composition of Cabernet Franc and Cabernet Sauvignon is alike, with exception of *meso*-2,3-butanediol, alanine, acetaldehyde, and sucrose, which are higher in Cabernet Franc, while choline and myo-inositol are lower. Ruben contains significant amounts of choline and myo-inositol. A high content of galactose is measured in wines from typical Bulgarian red grape varieties Ruben, Melnik-55, and Mavrud as well as in Merlot. Mavrud wines have higher levels of tyrosine and shikimic acid than Merlot and Melnik-55, which are distinguished by the concentration of arabinose in Merlot, and of *meso*-2,3-butanediol and choline in Melnik-55. Malbec contains higher levels of choline, *meso*-2,3-butanediol, and lactic acid (>2117 mg/L), which are also typical for Tempranillo wines. Tempranillo and Pinot noir wines demonstrate high contents of isobutanol, choline, and trigonelline. In addition, more 1-propanol and coutaric acid are found in Pinot noir. Diluted wines have a profile similar to that of white wines—they are rich in malic and citric acid. Accuracy of the misclassification matrix of the model used, presented in Table S4, is 98.44%.

For practical distinction of the measured wines, the information collected so far allows to derive a decision tree using the CHAID algorithm. Samples were divided into 15 classes according to the measured wines with at least three samples. The visualization is presented in Figure 7. Ten

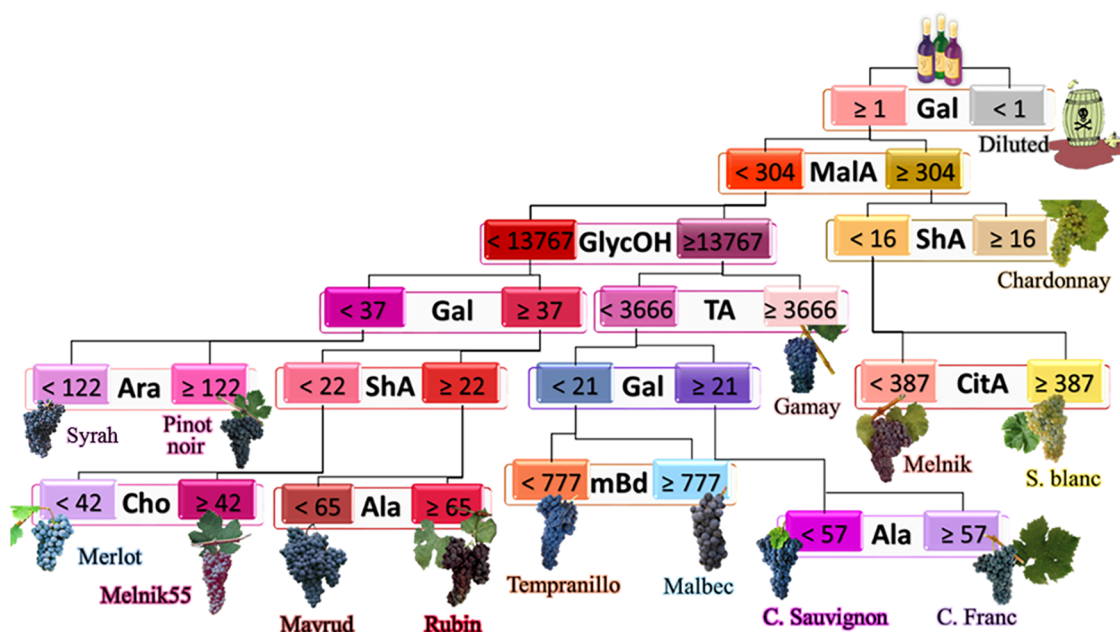


Figure 7. Information from a decision tree distinguishing wines from 15 different grape types.

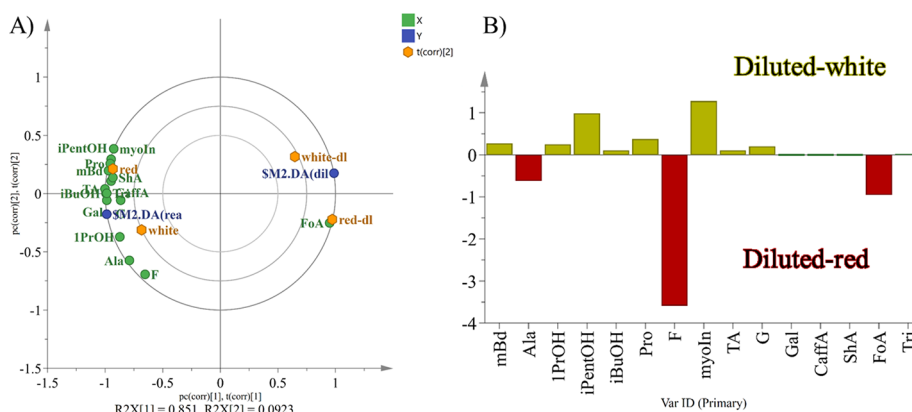


Figure 8. (A) Biplot of the created PLS-DA model, demonstrating clear discrimination between real and diluted wines. (B) Contribution plots for diluted white and red wines.

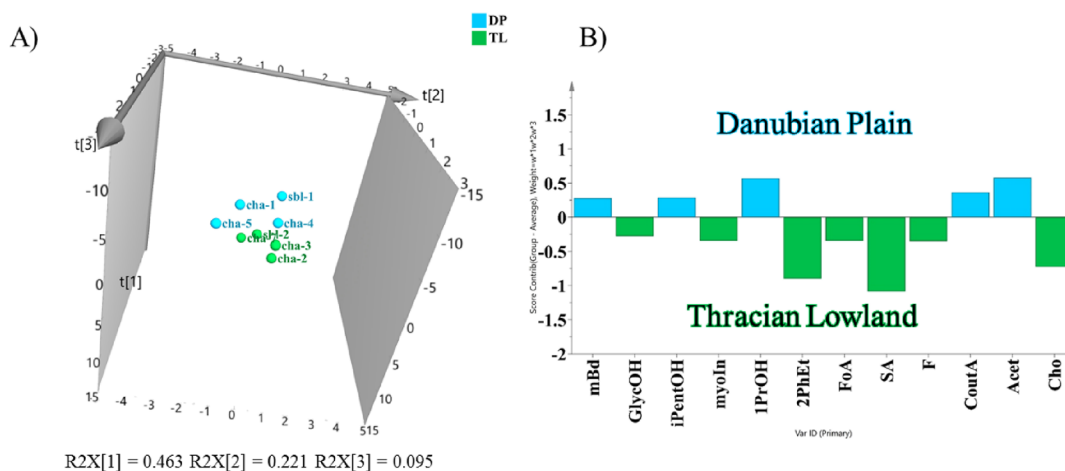


Figure 9. (A) Three-dimensional graph of results and (B) contribution of individual components for white wines produced in the Danubian Plain (DP) and the Thracian Lowland (TL).

components from different classes—monosaccharides (Gal and Ara), organic acids (TA, MalA, CitA, and ShA), alcohols (mBd and GlycOH), alanine, and choline can be determined as useful for differentiation of grape varieties used in winemaking and characterized in this study. The tree confirms that white wine and Shiroka Melnishka loza have a higher concentration of malic acid. Wines from Cabernet, Tempranillo, Malbec, and Gamay varieties contain higher levels of glycerol, while Merlot and typical Bulgarian wines contain more galactose. The decision tree allows accurate identification of 15 wine types. The only incorrectly classified sample is csa-6, predicted as Chardonnay. Despite the effectiveness in classifying the 15 wine varieties studied, the created decision tree is not suitable for assessing other types of wines and those made from multiple grape varieties. It can provide preliminary information about the type, but further statistical analysis is necessary to confirm the specific varietal.

Given the limited sample size for some wine varieties in this study, conducting future research with a larger sample size is necessary to improve reliability and substantiate the conclusions of our findings.

2.4. Distinguishing Real from Diluted Wine. One example of distinction of real from diluted wine was performed by partial least squares discriminant analysis. Average values of red, white, and diluted red wines were used for analysis, and data were grouped into two classes—real and diluted. The

model's VIP parameter identified 15 substances crucial for differentiating between the two wine types—mBd, iBuOH, iPentOH, myoIn, iPrOH, TA, FoA, ShA, Ala, Pro, Gal, G, F, CaffA, and Tri. Figure 8A illustrates the biplot of the created PLS-DA model, demonstrating clear discrimination between real and diluted wines.

The biplot suggests that most substances in diluted wines are present either in much lower concentrations or entirely absent except formic acid. Contribution plots depicted on Figure 8B demonstrate the significant variation in the quantities of 11 compounds between diluted red and white wines. Fructose is found in higher abundance in diluted red wines, while glucose and tartaric acids are more prominent in white wines. Some higher alcohols—*isopentanol*, *meso*-2,3-butanediol, and *myo*-inositol—have slightly higher concentrations in white than in red adulterated wines, indicating differences in dilution. The process used to make diluted wines is also influenced by the desired product's color. Figure S6 shows the percentage content of the main components (G, TA, myoIn, F, Pro, iPentOH, and mBd) calculated in terms of the sum of the amounts of the fifteen substances according to the color and quality of the wine.

Distinction of white and blue wine, red and rosé, and red wines produced with and without maceration is provided in the Supporting Information (Figures S7–S9).

2.5. Differentiation of Wine by Geographical Origin.

Climate and topography affect the wine composition in addition to the grape variety used and the production method. We compared limited series of eight analyzed white wines, six Chardonnay and two Sauvignon Blanc samples, and eight red wines, two Cabernet Sauvignon, two Cabernet Franc, two Merlot, and two Syrah samples, produced in both PGIs in Bulgaria to prove the potential of the methodology. White and red wines were considered separately to limit the influence of the grape variety. The samples of each wine color were divided into two groups, containing an equal number of wines of the same grape type and different geographical origin. PLS-DA was applied to determine the components specific for distinction of white wines from the Danubian Plain (DP) and from the Thracian Lowland (TL). Using VIP characteristics, we were able to identify 12 substances—coutaric, formic and succinic acids, fructose, *meso*-2,3-butanediol, 1-propanol, isopentanol, glycerol, myo-inositol, 2-phenylethanol, acetoin, and choline—whose amounts allow us to distinguish the white wines produced in the two regions. The new PLS-DA model using these substances provides good separation characteristics for the two PGI regions, visualized in Figure 9 with a three-dimensional score plot and the contribution of the individual components for both classes.

White wines from the Danubian Plain contain larger amounts of higher alcohols (mBd, iPentOH, and 1PrOH), coutaric acid, and acetoin, while white wines from the Thracian Lowland are richer in acids (FoA and SA), sugars (F), sugar alcohols (myoIn), choline, glycerol, and 2-phenylethanol.

Analogous chemometric analysis was performed to determine the significant components for distinguishing red wines from protected geographical indications in the country (Figure S10). The VIP parameter defines 12 substances allowing the wines from the two regions to be distinguished, six of which (1PrOH, FoA, F, CoutA, Acet, and Cho) are identical to those found in white wines. Red and white wines from the Thracian Lowlands contain more sugars (F and Su) and choline. Red wines from the Danubian Plain (cfr-2, csa-4, mer-3, and syr-3), like white wines from the same area, contain more alcohols (MeOH and 1PrOH) and amino acids (Ala and Pro), caffeic acid, and trigonelline. Red wines from areas south of the Balkan Mountains (cfr-1, csa-2, mer-9, and syr-1) typically contain more acetoin and coutaric acid as white wines from the Danubian Plain.

Grounds for distinction are the temperature amplitude and the different distribution of precipitation by seasons. In addition to the climate, different soil types and a lower slope of the relief in the Thracian lowland could be important. The observed differences show that climate and topography affect differently some components of white and red grapes and the wines that they produce.

Although the classification matrices used to validate both PLS-DA models (Table S5) do not indicate that any of the tested samples were misclassified, further research involving more objects is needed for more robust conclusions.

3. MATERIALS AND METHODS

3.1. Wine Samples. In the present work, 70 wine and 3 test samples have been investigated, and 30 of them (cfr-1, cfr-2, csa-2–csa-4, m55-1, mar, mel-1, mer-2, mer-3, mer-9, pno-2, pno-3, syr-1–syr-3, egi, cha-1–cha-5, mot, mpg, sbl-1, sbl-2, vio, and tes-1–tes 3) were obtained from colleagues at the University of Food Technologies - Plovdiv (project “Tradi-

tional Bulgarian Wines - Characteristics and Identification” No. FNI B 02/217, 2014–2016). They were produced from 50 kg of an average sample of technologically ripe grapes of the given variety and region, according to Table S1. The grapes were purchased from licensed grape growers, according to the requirements of the Executive Agency for the Vine and Wine and were processed in industrial conditions according to the traditional technology in the five wine cellars until receiving wine. Production places were in village Starosel (Plovdiv region), Chateau “Burgozone” (Vratsa region), “Logodazh” (Blagoevgrad region), and training and experimental base in Brestnik (Plovdiv region). Samples were taken from each young wine after completion of the alcoholic and malolactic fermentations, which were stored in conditions of the cellars for 6 months for self-clarification. During this period, they were cared for in accordance with the technological instructions in the wine industry. Samples mch and mel-2 and three diluted wine samples dil-1–3 were provided from local winemakers. International wines were bought from groceries in Bulgaria and abroad.

3.2. Sample Preparation. Wines were analyzed by NMR spectroscopy immediately after delivery/purchase, using the methodology adopted by the OIV in 2020.⁹ Samples were prepared according to the method described by Godelman¹⁵ and Fresenius⁸ and accepted by the OIV, which involves dissolving 495 μ L of wine in 55 μ L of deuterated 0.1 M phosphate buffer solution, NaH_2PO_4 and H_3PO_4 , pH = 2.90, containing 0.1% TSP and 0.05% NaN_3 . A small amount of 0.1 M H_3PO_4 or NaOH was added to the 70 wine samples after careful stirring to adjust the pH of each sample to 3.10 ± 0.02 if necessary. The solutions were transferred to 5 mm NMR tubes and homogenized in an ultrasonic bath for 1 min.

3.3. NMR Spectroscopy. The NMR spectra were recorded at 300.0 ± 0.1 K on a Bruker Avance II⁺ 600 spectrometer (Biospin GmbH, Germany) equipped with a PA BBO 600S3 BB-H-D-05 Z probe equipped with a gradient coil. For the analysis of all 70 wine samples, ¹H spectra with suppression of the water signal were acquired. The following parameters were applied: *zg0pr*, 45° pulse, 16 dummy and 256 accumulations, spectral width 13.6 ppm, 64k data points, acquisition time/relaxation delay 4.00/4.00 s. It should be noted that under these condition pulse sequences *zgcppr*, *noesypr1d*, and *zg0pr* provided completely identical spectra. Exponential multiplication with a line width of 0.3 Hz and FT were applied before referencing to the signal of TSP at 0 ppm.

Wine presents a complex mixture; that is why for unambiguous assignment of the signals, additional ¹³C and two-dimensional experiments (JRES, TOCSY, HSQC) were recorded on 11 of the samples. Carbon spectra (*zgdc30*) were recorded with the following parameters: SW = 238 ppm, NS = 20k, TD = 32k, AQ = 0.45 s, d1 = 1.5 s. JRES and TOCSY experiments obtained by *jresprqf* and *zdipsi2gpphpr* pulse sequences were registered. The following parameters were applied: spectral width 10.6 ppm/66 Hz, NS = 16/8, DS = 16, registration 8k \times 64/2k \times 256 data points, relaxation time 2 s, mixing time 0.15 s. HSQC spectra (*hsqcedetgppsp.3*) were acquired accordingly 10.6/175 ppm, 16/16 scans, 2k \times 400, 1.5 s.

Selected non-overlapped signals of 31 identified substances were integrated for quantitative analysis (Table S6). Only one signal was used for each of the identified compounds, and the TSP signal was used as a standard of known quantity. Quantification was performed using the following formula:¹⁵

$$m_x = \frac{N_{\text{TSP}} \times \text{Int}_x \times M_x \times m_{\text{TSP}}}{N_x \times \text{Int}_{\text{TSP}} \times M_{\text{TSP}}}$$

where N_{TSP} represents the number of protons from the TSP signal ($N_{\text{TSP}} = 9$), N_x is the number of chemically equivalent protons responsible for the signal of the investigated compound x ; Int_x represents the integral intensity for the signal of the analyte x ; Int_{TSP} is the integral intensity for the TSP signal ($\text{Int}_{\text{TSP}} = 1$); M_x is the molar mass of the analyzed compound; M_{TSP} is the molar mass of the standard ($M_{\text{TSP}} = 172.27$ g/mol); and m_{TSP} is the amount of TSP in each sample (mg/L).

3.4. Chemometrics: Methods and Software Used.

Excel software (Microsoft Office Standard 2019) was used for preparation of boxplots, one-factor analysis of variance (ANOVA), and Nightingale's diagrams. Supervised chemometrics methods (partial least squares linear discriminant analysis (PLS-DA) and orthogonal partial least squares discriminant analysis (OPLS-DA)) and unsupervised cluster analysis with Euclidean distance were created using SIMCA 17.0.2 software (Sartorius Stedim Data Analytics AB, Umetrics). Linear discriminant analysis (LDA) and heatmap were performed by software package Past 4.06 (manufacturer Oyvind Hammer), while the decision tree (C&R Tree) was created using a CHAID algorithm and Sipina 3.12 software package (producer Ricco Rakotomalala).

The chemometric techniques were used to distinguish wines according to the grape variety and the geographical area where they are produced and real from diluted wine and to determine the differences between wines of Bulgarian and international varieties.

3.4.1. Differentiation of Wine from Typical Bulgarian Grape Varieties. To differentiate typical Bulgarian from international grape varieties grown in Bulgaria, the data from Table S2 for 40 international (int) and 14 Bulgarian (bg) samples was analyzed using OPLS-DA analysis, composed of two classes. The model was built using one predictive and four orthogonal components ($R_2X(\text{cum}) = 0.541$; $R_2Y(\text{cum}) = 0.689$; $Q_2(\text{cum}) = 0.0109$), and for the prepared graph in Figure 4, the first orthogonal and the predictive component were used. A misclassification matrix was used to determine the classification accuracy. Visualization of the considerable differences between white and red wines of typical Bulgarian and international varieties was carried out by Nightingale's diagrams, presented in Figure S3.

According to the present legislation, there are two areas in Bulgaria⁴¹ where PGI wines are produced—the Danubian Plain (DP) and the Thracian Lowland (TL). A three-component PLS-DA analysis was applied to distinguish wines from the two areas using eight white wines divided into two groups (TL ($n = 4$) and DP ($n = 4$)) so that both classes contain the same types of wines. The PLS-DA model ($R_2X(\text{cum}) = 0.779$; $R_2Y(\text{cum}) = 0.783$; $Q_2(\text{cum}) = 0.419$) provides 12 significant substances with a VIP value above 1.00 that allowed to distinguish both areas. Visualization of the results is presented in the tridimensional score plot in Figure 9. A completely identical approach was used to distinguish eight red wines produced in the two regions (TL ($n = 4$) and DP ($n = 4$)). For the preparation of this PLS-DA model, 12 substances with VIP values above 0.93 were used. The characteristics of the red wine model are as follows: $R_2X(\text{cum}) = 0.672$; $R_2Y(\text{cum}) = 0.920$; $Q_2(\text{cum}) = 0.412$, visualized in

Figure S10. Classification matrixes for both cases are presented in Table S5.

3.4.2. Discrimination of Botanical Origin. Botanical origin of wine was analyzed using the 61 samples grouped into 15 classes: cha, Chardonnay ($n = 7$); cfr, Cabernet Franc ($n = 3$); csa, Cabernet Sauvignon ($n = 9$); egi, Egiodola ($n = 1$); gam, Gamay ($n = 3$); m5S, Melnik 55 ($n = 3$); mal, Malbec ($n = 3$); mav, Mavrud ($n = 3$); mel, Melnik, Shiroka Melnishka loza ($n = 3$); mer, Merlot ($n = 9$); pno, Pinot noir ($n = 3$); rub, Rubin ($n = 3$); sbl, Sauvignon blanc ($n = 3$); syr, Syrah ($n = 3$); tem, Tempranillo ($n = 3$); dil, diluted ($n = 3$); and tree test samples. To determine the substances differing in the individual wine types, the ANOVA method was applied at a statistical certainty of α 0.05 to the data of 15 classes, excluding the test samples. Quantitative data for four components (shikimic acid, tartaric acid, glycerol, and lactic acid) in each group are visualized by boxplots (Figure 4). LDA and PLS-DA techniques using three components were performed to differentiate all 15 classes and to determine the characteristic compounds for each class. Both methods provided different results. While LDA described more than 75% of the variables, PLS-DA described only 50%. That is why, the LDA model provides a convincing graph for differentiation of the grape varieties using the first two components presented in Figure 5 as well as a classification matrix (Table S4). A heatmap was used to determine the differences in each class. It was created based on quantitative data of all significant substances after their normalization and is shown in Figure 6. An easy and reliable overview that allows determination of the variety in new wine samples provides the prepared decision tree, shown in Figure 7. It uses the quantities of the substances described in Table S2 for the main 15 wine classes. An LDA model was created for all 73 samples (70 from the training set and 3 from the test set). The score plot and biplot of the LDA analysis can be seen in Figure S5.

3.4.3. Distinction of Real from Diluted Wine. Easy distinction of diluted from real wine could be made both by cluster analysis and PLS-DA. To prepare the analyses, average values for the compounds in Table S2 of the analyzed white and red wines and of the two diluted red wines were used. PLS-DA analysis with two distinct classes (diluted ($n = 2$) and real ($n = 2$)) provides the results. The PLS-DA model with two components ($R_2X(\text{cum}) = 0.924$; $R_2Y(\text{cum}) = 1$; $Q_2(\text{cum}) = 0.981$) provided 15 substances with VIP values above 1.00, significant for distinguishing the four groups. The model created with data for the 15 substances was characterized by 2 components. The variance explained by them related to the class information (R_2X) was 94.4% and the prediction coefficient (Q_2) was 99.3%.

Nightingale's diagrams (Figures S7 and S8) are applied to visualize differences in the amounts of 31 components of blue and white wines and red and rosé produced from the same grape variety (Chardonnay and Cabernet Sauvignon). Average values for the substances of a given type and color of wine, as well as the values for blue wine or rosé, were used to create the charts. The amounts of compounds in Pinot noir wine prepared with and without maceration are analogously compared in Figure S9.

4. CONCLUSIONS

In this work, we show the detailed studies of wine and quantitation of its common components that can be used to characterize safety, quality, and authentication of wine from local and international producers. The used NMR method-

ology and chemometric tools are very powerful and allow to check subtle differences in wines as grape variety, country of origin, color, production methods, and others. A decision tree has been proposed to easily predict the origin of new wines from 15 grape varieties. This study provides evidence that this methodology is quite adequate to be used for food control issues and for preparation of databases for detailed characterization of wine from individual grape varieties.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsomega.3c00636>.

Data for the 70 measured wine samples and comparison with literature data, misclassification matrixes, typical ^1H NMR spectrum and $^1\text{H}/^{13}\text{C}$ chemical shifts of components in wine, comparison of different wines via Nightingale's diagrams, 2D/3D score plots, biplot and contribution plots (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Svetlana Simova – Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria; orcid.org/0000-0002-9565-3850; Email: Svetlana.Simova@orgchm.bas.bg

Author

Dessislava Gerginova – Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acsomega.3c00636>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Financial support of the National Science Fund (Projects UNA-17/2005, DRNF-02-13/2009), the Ministry of Education and Science-Bulgaria (INFRAMAT project), and IOCCP-BAS are gratefully acknowledged for building and maintaining the NMR infrastructure.

■ REFERENCES

- (1) Johnson, R. *Food fraud and economically motivated adulteration of food and food ingredients*. Congressional Research Service Washington DC: 2014.
- (2) Godelmann, R.; Fang, F.; Humpfer, E.; Schütz, B.; Bansbach, M.; Schäfer, H.; Spraul, M. Targeted and nontargeted wine analysis by ^1H 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.
- (3) Anastasiadi, M.; Zira, A.; Magiatis, P.; Haroutounian, S. A.; Skaltsounis, A. L.; Mikros, E. ^1H NMR-based metabonomics for the classification of Greek wines according to variety, region, and vintage. Comparison with HPLC data. *J. Agric. Food Chem.* **2009**, *57*, 11067–11074.
- (4) Gilby, C., *The Wines of Bulgaria, Romania and Moldova*; Infinite Ideas: 2018.
- (5) Burns, J.; Gardner, P. T.; O'Neil, J.; Crawford, S.; Morecroft, I.; McPhail, D. B.; Lister, C.; Matthews, D.; MacLean, M. R.; Lean, M. E.; Duthie, G. G.; Crozier, A. Relationship among antioxidant activity,

vasodilation capacity, and phenolic content of red wines. *J. Agric. Food Chem.* **2000**, *48*, 220–230.

(6) International Organization of Vine and Wine, International Code of Oenological Practices, 2021.

(7) Waterhouse, A. L.; Sacks, G. L.; Jeffery, D. W., *Understanding wine chemistry*. John Wiley & Sons: 2016.

(8) Borgmann, E.; Fresenius, T. W., *Anleitung zur chemischen Analyse des Weines*. CW Kreidel: 1898.

(9) OIV, *Compendium of International methods of wine and must analysis*. 2022.

(10) Ogrinc, N.; Košir, I. J.; Kocjančič, M.; Kidrič, J. Determination of authenticity, regional origin, and vintage of Slovenian wines using a combination of IRMS and SNIF-NMR analyses. *J. Agric. Food Chem.* **2001**, *49*, 1432–1440.

(11) Ranaweera, R. K.; Capone, D. L.; Bastian, S. E.; Cozzolino, D.; Jeffery, D. W. A review of wine authentication using spectroscopic approaches in combination with chemometrics. *Molecules* **2021**, *26*, 4334.

(12) Popirdă, A.; Luchian, C. E.; Cotea, V. V.; Colibaba, L. C.; Scutarășu, E. C.; Toader, A. M. A review of representative methods used in wine authentication. *Agriculture* **2021**, *11*, 225.

(13) Viskič, M.; Bandič, L. M.; Korenika, A.-M. J.; Jeromel, A. NMR in the service of wine differentiation. *Foods* **2021**, *10*, 120.

(14) *Resolution OIV-oen 618–2020*, 2020, <https://www.oiv.int/public/medias/7590/oiv-oen-618-2020-en.pdf> (accessed 2023-04-06).

(15) Godelmann, R.; Kost, C.; Patz, C. D.; Ristow, R.; Wachter, H. Quantitation of Compounds in Wine Using ^1H NMR Spectroscopy: Description of the Method and Collaborative Study. *J. AOAC Int.* **2016**, *99*, 1295–1304.

(16) Schwarzinger, S. Echtheits- und Qualitätsscreening von Lebensmitteln. NMR, oder was sonst? *GIT Labor-Fachzeitschrift* **2023**, *67*, 15.

(17) Conference abstracts: (a)) Gerginova, D.; Mitrev, Y.; Simova, S. HNMR chemical profile of several wines produced in Bulgaria. In *Book of Abstracts Ampere NMR School*, Zakopane, 2018. (b)) Gerginova, D.; Mitrev, Y.; Simova, S.; Karadjova, I. NMR spectroscopy of Bulgarian wine – an alcoholic beverage rich in bio-antioxidants. In *Book of Abstracts First International Conference on Bio-antioxidants: Natural Bio-antioxidants – as a base for new synthetic drugs and food additives/supplements*, Sofia, Bulgaria, 2017 (c)) Gerginova, D.; Mitrev, Y.; Simova, S. Wines produced in Bulgaria - chemical profile via ^1H NMR. In *Book of Abstracts XXIIInd National Conference with International Participation New Cryogenic and Isotope Technologies for Energy and Environment*. Băile Govora, Romania, 2018.

(18) Gerginova, D.; Mitrev, Y.; Simova, S. Periodic table of compounds in wine using ^1H NMR spectroscopy and chemometrics, poster at 47th World Chemistry Congress of IUPAC, 07–12.07.2019.

(19) Boroski, M.; Crupi, P.; Tamborra, P.; Antonacci, D.; Toci, A. T. Influence of winemaking techniques with low sulphur dioxide on wine varieties Chardonnay, Pinot and Montepulciano. *J. Food Nutr. Res.* **2017**, *56*, 326–334.

(20) Simpson, R.; Miller, G. Aroma composition of Chardonnay wine. *Vitis* **1984**, 143–158.

(21) Sen, I.; Tokatli, F. Authenticity of wines made with economically important grape varieties grown in Anatolia by their phenolic profiles. *Food Control* **2014**, *46*, 446–454.

(22) Jégou, S.; Hoang, D. A.; Salmon, T.; Williams, P.; Oluwa, S.; Vrigneau, C.; Doco, T.; Marchal, R. Effect of grape juice press fractioning on polysaccharide and oligosaccharide compositions of Pinot meunier and Chardonnay Champagne base wines. *Food Chem.* **2017**, *232*, 49–59.

(23) Gomes, T. M.; Mazon, L. F.; Panceri, C. P.; Machado, B. D.; Brighenti, A.; Burin, V. M.; Bordignon-Luiz, M. T. Changes in vineyard productive attributes and phytochemical composition of sauvignon blanc grape and wine induced by the application of silicon and calcium. *J. Sci. Food Agric.* **2020**, *100*, 1547–1557.

(24) Kosir, I.; Kocjancic, M.; Kidric, J. Wine analysis by 1D and 2D NMR spectroscopy. *Analisis* **1998**, *26*, 97–101.

- (25) Roman, T.; Nicolini, G.; Barp, L.; Malacarne, M.; Tait, F.; Larcher, R. Shikimic acid concentration in white wines produced with different processing protocols from fungus-resistant grapes growing in the Alps. *Vitis* **2018**, *57*, 41–46.
- (26) Parish-Virtue, K.; Herbst-Johnstone, M.; Bouda, F.; Fedrizzi, B.; Deed, R. C.; Kilmartin, P. A. Aroma and Sensory Profiles of Sauvignon Blanc Wines from Commercially Produced Free Run and Pressed Juices. *Beverages* **2021**, *7*, 29.
- (27) Huang, X.-Y.; Jiang, Z.-T.; Tan, J.; Li, R. Geographical Origin Traceability of Red Wines Based on Chemometric Classification via Organic Acid Profiles. *Journal of Food Quality* **2017**, *2017*, 1.
- (28) Avar, P.; Nikfardjam, M. S.; Kunsági-Máté, S.; Montskó, G.; Szabó, Z.; Böddi, K.; Ohmacht, R.; Márk, L. Investigation of phenolic components of Hungarian wines. *Int. J. Mol. Sci.* **2007**, *8*, 1028–1038.
- (29) Longo, R.; Pearson, W.; Merry, A.; Solomon, M.; Nicolotti, L.; Westmore, H.; Damberg, R.; Kerslake, F. Preliminary study of Australian Pinot noir wines by colour and volatile analyses, and the Pivot Profile method using wine professionals. *Foods* **2020**, *9*, 1142.
- (30) Dias, F. D. S.; David, J. M.; David, J. P. Determination of phenolic acids and quercetin in Brazilian red wines from Vale do São Francisco region using liquid-liquid ultrasound-assisted extraction and HPLC-DAD-MS. *J. Braz. Chem. Soc.* **2016**, *27*, 1055–1059.
- (31) Mardones, C.; Hitschfeld, A.; Contreras, A.; Lepe, K.; Gutiérrez, L.; von Baer, D. Comparison of shikimic acid determination by capillary zone electrophoresis with direct and indirect detection with liquid chromatography for varietal differentiation of red wines. *J. Chromatogr. A* **2005**, *1085*, 285–292.
- (32) Bouloumpasi, E.; Soufleros, E.; Tsarchopoulos, C.; Biliaderis, C. Primary amino acid composition and its use in discrimination of Greek red wines with regard to variety and cultivation region. *VITIS-GEILWEILERHOF* **2002**, *41*, 195–202.
- (33) Fowles, G. W. A. Acids in grapes and wines: a review. *J. Wine Res.* **1992**, *3*, 25–41.
- (34) Hu, B.; Yue, Y.; Zhu, Y.; Wen, W.; Zhang, F.; Hardie, J. W. Proton nuclear magnetic resonance-spectroscopic discrimination of wines reflects genetic homology of several different grape (*V. vinifera* L.) cultivars. *PLoS One* **2015**, *10*, No. e0142840.
- (35) Hu, B.; Zhao, Q.; Yue, Y.; Zhu, J.; Lu, G.; Li, H.; Zhang, F.; Hardie, W. *1H Nuclear magnetic resonance-based metabolomic study for Cabernet Sauvignon wines in different vintages*; *PeerJ Preprints* **2016**, e2332v1.
- (36) Jiang, B.; Zhang, Z. W. A Preliminary Study of Aroma Composition and Impact Odorants of Cabernet Franc Wines under Different Terrain Conditions of the Loess Plateau Region (China). *Molecules* **2018**, *23*, 1096.
- (37) Niculaua, M.; Tudose-Sandu-Ville, S.; Cotea, V. V.; Luchian, C. E.; Tudose-Sandu-Ville, O.-F. Phenolic Compounds Content in Merlot Wines Obtained through Different Thermomaceration Techniques. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* **2017**, *45*, 548–552.
- (38) Đorđević, N. O.; Pejin, B.; Novaković, M. M.; Stanković, D. M.; Mutić, J. J.; Pajović, S. B.; Tešević, V. V. Some chemical characteristics and antioxidant capacity of novel Merlot wine clones developed in Montenegro. *Scientia Horticulturae* **2017**, *225*, 505–511.
- (39) Dimitrov, D.; Haygarov, V.; Yoncheva, T.; Nedelkovski, D. Aromatic profile of red wines from grapevine varieties grown in the conditions of Northern Bulgaria. *Journal of Mountain Agriculture on the Balkans* **2017**, *20*, 187–200.
- (40) Haygarov, V.; Yoncheva, T.; Dimitrov, D. Study of trans-resveratrol content in grapes and wine of the varieties Storgozia, Kaylashki Rubin, Trapezitsa, Rubin, Bouquet and Pinot Noir. *Rastenievadni nauki - Bulgarian Journal of Crop Science* **2020**, *57*, 50–58.
- (41) *eAmbrosia - the EU geographical indications register*, <https://ec.europa.eu/info/food-farming-fisheries/food-safety-and-quality/certification/quality-labels/geographical-indications-register/> (accessed 2023-04-06).