Research Article

Shaochen Xu, Jiangyu Zhu, Qi Zhao, Jin Gao, Huining Zhang*, Boran Hu*

Quality evaluation of Cabernet Sauvignon wines in different vintages by ^1H nuclear magnetic resonance-based metabolomics

<https://doi.org/10.1515/chem-2020-0126>
received May 20, 2020; accepted March 2, 2021

Abstract: A proton nuclear magnetic resonance (NMR)-based metabolomic study was used to characterize 2009, 2010, 2011, and 2012 vintages of Cabernet Sauvignon wines from Ningxia, which were vinified using the same fermentation technique. The pattern recognition methods of principal component analysis (PCA), partial least squares discriminant analysis (PLS-DA), and orthogonal PLS-DA (OPLS-DA) clearly distinguished between the different vintages of wine driven by the following metabolites: valine, 2,3-butanediol, ethyl acetate, proline, succinic acid, lactic acid, acetic acid, glycerol, gallic acid, and choline. The PLS-DA loading plots also differentiated among the metabolites of different vintages. In the 2009 vintage wines, we found the highest levels of gallic acid, valine, proline, and 2,3-butanediol. The 2011 vintage wines contained the highest levels of lactic acid, and the highest levels of ethyl acetate,

succinic acid, glycerol, and choline were observed in the 2012 vintage wines. We selected eight metabolites from the ^1H NMR spectra that were quantified according to their peak areas, and the concentrations were in agreement with the results of PLS-DA and OPLS-DA analyses.

Keywords: Cabernet Sauvignon, vintage, metabolomics, NMR, pattern recognition

1 Introduction

Wine obtains several metabolites from grape berries during fermentation. Many factors, including the soil, climate, viticultural practices (soil tillage and covering), winemaking process, and vintage, contribute to the metabolite composition and content of wines [1–5]. So far, most studies about wines have focused on the characterization and evaluation of the biological activities of selected extractable components, whereas there is a comparative lack of research on the metabolites in wines. The common parameters used to evaluate the quality of wine are the total soluble solids, alcohol concentration, total acids, and total phenols. These basic parameters are significant, and the classical analytical methods can easily detect many other important compounds [1,6–9]. However, these parameters reflect only the health of the wine and cannot fully explain the quality of the wine. Therefore, for wine quality assessment, powerful advanced analysis methods are necessary to determine the metabolites in wines [6,10].

In proton nuclear magnetic resonance-based (^1H NMR-based) metabolomics, one pair of potential information extraction and classification of samples provide a new method for evaluating metabolic functions. ^1H NMR spectroscopy has been combined with principal component analysis (PCA), partial least squares discriminant analysis (PLS-DA), and orthogonal PLS-DA (OPLS-DA) to distinguish different wines obtained from the same variety grown in different geographical regions [1,6,11] and different varieties grown in the same geographical

* **Corresponding author: Huining Zhang**, Yangzhou Hospitality Institute, 500 Zhannan Road, Yangzhou 225002, Jiangsu, China, e-mail: zhanghuiningzh@sina.com

* **Corresponding author: Boran Hu**, The School of Tourism and Culinary Arts (The School of Food Science and Engineering), Yangzhou University, 196 Huayang West Road, Yangzhou 225127, Jiangsu, China, e-mail: huboran@yzu.edu.cn

Shaochen Xu: The Graduate School of East Asian Studies, Yamaguchi University, 1677-1 Yoshida, Yamaguchi-shi, Yamaguchi 753-8511, Japan; The School of Tourism and Culinary Arts (The School of Food Science and Engineering), Yangzhou University, 196 Huayang West Road, Yangzhou 225127, Jiangsu, China, e-mail: b002sn@yamaguchi-u.ac.jp

Jiangyu Zhu: The School of Tourism and Culinary Arts (The School of Food Science and Engineering), Yangzhou University, 196 Huayang West Road, Yangzhou 225127, Jiangsu, China, e-mail: zhu.jiangyu272@mail.kyutech.jp

Qi Zhao: The School of Tourism and Culinary Arts (The School of Food Science and Engineering), Yangzhou University, 196 Huayang West Road, Yangzhou 225127, Jiangsu, China, e-mail: 85425674@qq.com

Jin Gao: The School of Tourism and Culinary Arts (The School of Food Science and Engineering), Yangzhou University, 196 Huayang West Road, Yangzhou 225127, Jiangsu, China, e-mail: 2430449113@qq.com

region [3]. Its discriminant analysis counterpart (OPLS-DA) was demonstrated as a powerful tool for the analysis of qualitative data structures while the prediction results are equivalent to classification using standard PLS-DA [12].

In this study, PCA, PLS-DA, and OPLS-DA models were employed to distinguish the Cabernet Sauvignon wines from different vintages.

2 Methodology

2.1 Cabernet Sauvignon wines from different vintages

In this study, no specific permissions were required for the research activities, and the field studies did not involve protected species.

All the wine samples were vinified in Ningxia province of Northwest China. The grapes were grown nongrafted in a single vineyard with a uniform soil type (containing gravel mainly consisting of light sierozem and an organic matter content of 0.4–1.0%) in the Helan Mountains of Ningxia Province. This area is located in the warm temperate region of the northern hemisphere with a dry continental climate characterized by dry summers and severe winters. During the growth vintage (March–October) over the last 40 years, the average temperature was 15.24°C, the rainfall was 264.45 mm, and the evaporation was 1312.0 mm. Small changes in climate were registered from 2009 to 2012, and the climate information is shown in Table 1. The vineyard was planted in 1994 in north-south lines; the line spacing is 2.5 m and the spacing within the line is 1.2–1.5 m, using standardized management.

2.2 Sample origin

Samples were obtained from 2009, 2010, 2011, and 2012 vintages of Cabernet Sauvignon wines that were vinified

by the GUANG XIA (YINCHUAN) HELAN MOUNTAIN WINERY CO. LTD, which were named S1, S2, S3, and S4, respectively; six samples were tested for each year, and each sample had three parallel samples. The wines were vinified using the same fermentation technique and the same yeast (Lalvin CY 3079) without other chemical adjustments except for potassium metabisulfite (50 mg/L). The wines were not aged in oak barrels. After fermentation, the wines were stored in fermenting tanks (50 t).

We obtained three parallel samples of each wine from the sampling mouth. Every replicate sample was funneled into a brown glass bottle (750 mL) that was then sealed with a cork and transported to the laboratory storage (−4°C). The grapes of each vintage were harvested at similar concentrations of reducing sugar and titrable acidity (Table 2). The chemical and physical features of the wines met the China national test standard (GB/T 15038-2006), as shown in Table 3. In addition, there is a nice liner relationship ($R^2 = 0.958$) between the residue sugar content and the alcohol content illustrated in Figure 1, reflecting the process in which natural glucose is consumed and alcohol is produced during the fermentation.

2.3 NMR sample preparation

Ten milliliters of wine were centrifuged at 4,000 rpm for 20 min, and 3 mL supernatants were frozen at −70°C for 12 h and then lyophilized for 48 h. The lyophilized wine was dissolved in 400 μ L of oxalate buffer (pH = 4.0), mixed with 140 μ L of D₂O and 60 μ L of a 0.75% 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) in D₂O solution, and then centrifuged at 13,000 rpm for 20 min. Next 500 μ L supernatants were placed in 5 mm NMR tubes. The chemical shift of DSS provided reference ($\delta = 0$) and internal standard quantitative analyses.

All the chemical reagents were of analytical grade. D₂O (99.9%) and DSS were purchased from SIGMA-ALDRICH.

2.4 ¹H NMR spectroscopy

¹H NMR spectra were recorded on a Bruker AVANCE 600 spectrometer operating at a ¹H frequency of 600.13 MHz and a temperature of 298 K using a ¹H {¹³C/¹⁵N} probe. A NOESYPRESAT pulse sequence was used to suppress the residual water signal. A total of 256 transients were

Table 1: Climate information of the vineyard during the growth vintage (March–October) in 2009–2012

Vintages	Average temperature (°C)	Rainfall (mm)	Evaporation (mm)
2009	17.65	243.5	1562.1
2010	17.11	233.6	1398.7
2011	15.24	262.2	1423.6
2012	16.52	251.4	1266.7

Table 2: Grape composition at harvest

Harvest date	Cultivar	Reducing sugar (g/L)	Titrate acidity (g/L)	pH
August 15, 2009	Cabernet Sauvignon	232.7 ± 1.1	7.67 ± 0.21	3.73 ± 0.01
August 16, 2010		228.8 ± 1.5	6.98 ± 0.30	3.88 ± 0.01
August 22, 2011		220.2 ± 0.9	8.32 ± 0.24	3.23 ± 0.02
August 18, 2012		225.6 ± 1.3	7.72 ± 0.20	3.54 ± 0.01

Table 3: Physical and chemical features of the wines

Index*	Vintages			
	2009	2010	2011	2012
Alcohol content (% Vol)	13.2 ± 0.0	12.9 ± 0.1	12.4 ± 0.1	12.8 ± 0.0
Residual sugar (g/L)	2.20 ± 0.09	2.55 ± 0.17	3.10 ± 0.12	2.50 ± 0.07
Total acid (g/L)	6.4 ± 0.0	6.7 ± 0.0	6.1 ± 0.0	5.8 ± 0.0
Volatile acid (g/L)	0.42 ± 0.02	0.45 ± 0.00	0.46 ± 0.01	0.43 ± 0.01
Dry extract (g/L)	27.9 ± 0.1	28.9 ± 0.5	27.6 ± 0.1	29.1 ± 0.3
pH	3.47 ± 0.02	3.22 ± 0.01	3.56 ± 0.03	3.73 ± 0.00
Total SO ₂ (mg/L)	86 ± 1	88 ± 2	82 ± 1	85 ± 0
Free SO ₂ (mg/L)	31 ± 0	28 ± 1	32 ± 1	33 ± 0
Methanol (mg/L)	205 ± 3	220 ± 1	214 ± 4	206 ± 2
Fe ³⁺ (mg/L)	2.2 ± 0.1	1.9 ± 0.0	2.1 ± 0.0	2.0 ± 0.0
Cu ²⁺ (mg/L)	0.055 ± 0.003	0.053 ± 0.001	0.065 ± 0.005	0.059 ± 0.002
K ⁺ (mg/L)	946 ± 7	936 ± 8	957 ± 4	955 ± 11
Ca ²⁺ (mg/L)	103 ± 3	97 ± 1	99 ± 1	102 ± 2
Tartaric acid (g/L)	2.64 ± 0.13	2.28 ± 0.05	2.32 ± 0.08	2.44 ± 0.11
Citric acid (g/L)	0.31 ± 0.00	0.28 ± 0.02	0.29 ± 0.01	0.26 ± 0.03
Lactic acid (g/L)	2.66 ± 0.04	2.64 ± 0.02	2.73 ± 0.05	2.53 ± 0.01
Color tone	12.5 ± 0.1	12.8 ± 0.1	12.3 ± 0.1	12.7 ± 0.2
Color tint	0.83 ± 0.00	0.82 ± 0.01	0.80 ± 0.00	0.81 ± 0.00

*The methods used to determine the physical and chemical features met the China National Test Standard GB/T15038–2006.

collected into 32,000 complex data points with a spectral width of 7183.9 Hz, an acquisition time of 2.3 s, a mixing time of 100 ms, and a relaxation delay of 2 s. The NMR spectra were processed with a line-broadening factor of 0.3 Hz prior to Fourier transformation.

2.5 NMR data reduction

The NMR spectral data were reduced into 0.005 ppm spectral buckets. The regions corresponding to water (4.6–4.8 ppm), incompletely removed DSS (−0.5–0.5 ppm, 1.74–1.84 ppm, and 2.90–2.95 ppm), and ethanol (1.18–1.22 ppm and 3.57–3.72 ppm) were removed by AMIX software. The data set was then imported into SIMCA-P 12.0 for multivariate statistical analysis.

Figure 1: Linear relationship between residual sugar content (g/L) and alcohol content (% vol).

2.6 Pattern recognition

We used PCA and PLS-DA to check the intrinsic variability of the data set and to separate out the different vintages of the wine, respectively. PCA was employed to examine the intrinsic variation in the data set. Applying an orthogonal signal correction (OSC) followed by PLS-DA analysis, OPLS-DA can eliminate the information that did not contribute to the discrimination. PLS-DA score plots from the ^1H NMR spectra of different vintage wines were generated in pairwise comparisons, then analyzed with OPLS-DA [13–15]. OPLS-DA is an improvement in the PLS-DA method to discriminate two groups using multivariate data [16], combine OSC and PLS-DA analysis and filter data. The advantage of OPLS-DA compared to PLS-DA is that OPLS-DA uses a single component as the predictor of groups, while the other components describe the variation in orthogonal to the first predictive component [17].

Ethical approval: The conducted research is not related to either human or animal use.

3 Results

3.1 Metabolite differences in wines of different vintages

The PCA score plot shows a clear differentiation among the Cabernet Sauvignon wines of different vintages. The models show good adaptability and high predictability

with high statistical values of R^2X (0.867) and Q^2 (0.789) (Figure 2).

The PLS-DA and OPLS-DA models were used to compare the different vintages of wine. As shown in Figure 3, the PLS-DA score plots derived from the ^1H NMR spectra of the 2009 and 2010 vintages of Cabernet Sauvignon wine had the highest values for the pairwise comparison of R^2X and Q^2 , and these indicate a clear separation between the 2009 and 2010 vintages of wine. As shown in Figure 4, the OPLS-DA score plots derived from the ^1H NMR spectra of the 2009 and 2010 vintages of Cabernet Sauvignon wine had the highest values for the pairwise comparison of R^2X and R^2Y . Figure 4 shows a clearer separation between the 2009 and 2010 vintages of wine than Figure 3.

The complementary load plot gives the contribution of the metabolite differentiation (Figure 5). The loading plot shows that the metabolites of the 2009 vintages are higher than those of the 2010 vintages. The loading plot shows high levels of valine, glycerol, 2,3-butanediol, α -glucose, acetic acid, proline, succinic acid, sucrose, tartaric acid, gallic acid, and tyrosine in the 2009 vintages, while ethyl acetate, lactic acid, choline, β -glucose, and α -D-glucuronic acid were at relatively low levels in the 2010 vintages.

Both the PLS-DA and OPLS-DA score plots showed a clear discrimination between the 2009 and 2011 vintages of the Cabernet Sauvignon wine (Figures 6 and 7), and the loading plot provides the metabolites that contributed to this discrimination (Figure 8). Higher levels of 2,3-butanediol, ethyl acetate, proline, succinic acid, glycerol, α -glucose, tartaric acid, choline, and sucrose and

Figure 2: PCA score plot of Cabernet Sauvignon wines of different vintages.

Figure 3: PLS-DA score plot chart from ¹H NMR spectra of 2009 and 2010 vintage Cabernet Sauvignon wines.

Figure 4: OPLS-DA score plot chart from ¹H NMR spectra of 2009 and 2010 vintage Cabernet Sauvignon wines.

Figure 5: PLS-DA loading plot chart from ¹H NMR spectra of 2009 and 2010 vintage Cabernet Sauvignon wines.

Figure 6: PLS-DA score plot chart from ¹H NMR spectra of 2009 and 2011 vintage Cabernet Sauvignon wines.

Figure 7: OPLS-DA score plot chart from ¹H NMR spectra of 2009 and 2011 vintage Cabernet Sauvignon wines.

Figure 8: PLS-DA loading plot chart from ¹H NMR spectra of 2009 and 2011 vintage Cabernet Sauvignon wines.

Figure 9: PLS-DA score plot chart from ^1H NMR spectra of 2009 and 2012 vintage Cabernet Sauvignon wines.

Figure 10: OPLS-DA score plot chart from ^1H NMR spectra of 2009 and 2012 vintage Cabernet Sauvignon wines.

lower levels of lactate and α -D-glucuronic acid were detected in the 2009 vintages compared to the 2011 vintages.

The PCA and OPLS-DA score plots of the 2009 and 2012 vintage Cabernet Sauvignon wines also showed clear separation (Figures 9 and 10) identified by higher levels of valine, 2,3-butanediol, proline, succinic acid, D-sucrose, tartaric acid, gallic acid, α -glucose, and β -glucose and lower levels of lactate, ethyl acetate, acetic acid, glycerol, α -D-glucuronic acid, and choline in the 2009 vintages (Figure 11).

The PCA and OPLS-DA score plots of the 2010 and 2011 vintage Cabernet Sauvignon wines also showed significant separation (Figures 12 and 13). The loading plot illustrates higher levels of choline, proline, and 2,3-butanediol and lower levels of valine, lactic acid, succinic acid, and glycerol in the 2010 vintages compared to those in the 2011 vintages (Figure 14).

The PCA and OPLS-DA score plots also showed significant differentiation between the Cabernet Sauvignon wines vinified in 2010 and 2012 (Figures 15 and 16). Relatively higher levels of valine and 2,3-butanediol and lower levels

Figure 11: PLS-DA loading plot chart from ^1H NMR spectra of 2009 and 2012 vintage Cabernet Sauvignon wines.

Figure 12: PLS-DA score plot chart from ^1H NMR spectra of 2010 and 2011 vintage Cabernet Sauvignon wines.

of lactic acid, proline, acetic acid, succinic acid, choline, glycerol, and ethyl acetate were found in the Cabernet Sauvignon wines vinified in 2010 compared to those vinified in 2012, as shown in the PLS-DA loading plot (Figure 17).

The PLS-DA and OPLS-DA score plots showed clear separation between the 2011 and 2012 vintage Cabernet Sauvignon wines based on the first component (Figures 18 and 19). The corresponding loading plot showed relatively high load levels of valine, lactic acid, and succinic acid and low levels of 2,3-butanediol, proline, acetic acid, choline, glycerol, D-sucrose, acetate, α-glucose, gallic acid, and tyrosine in the 2011 vintages compared with the 2012 vintages (Figure 20).

3.2 Quantitative analysis

The PLS-DA analysis and OPLS-DA analysis revealed small differences in the metabolite compositions and large differences in the metabolite concentrations in the Cabernet Sauvignon wines vinified from 2009 to 2012. The abbreviations S1–S4 represent the 2009–2012 vintages, respectively. The concentration of valine in the four vintages in order from high to low was $S1 > S3 > S2 > S4$, the concentration of 2,3-butanediol in order from high to low was $S1 > S2 > S4 > S3$, the concentration of glycerol in order from high to low was $S4 > S1 > S3 > S2$, the concentration of ethyl acetate in order from high to low was

Figure 13: OPLS-DA score plot chart from ^1H NMR spectra of 2010 and 2011 vintage Cabernet Sauvignon wines.

Figure 14: PLS-DA loading plot chart from ^1H NMR spectra of 2010 and 2011 vintage Cabernet Sauvignon wines.

$S4 > S2 > S1 > S3$, the concentration of succinic acid in order from high to low was $S4 > S3 > S1 > S2$, the concentration of lactate in order from high to low was $S3 > S4 > S1 > S2$, the concentration of choline in order from high to low was $S4 > S2 > S1 > S3$, and the concentration of gallic acid in order from high to low was $S1 > S4 > S3 > S2$.

From the ^1H NMR spectra, eight metabolites were selected, and their concentrations were calculated according to their peak areas. The results of this quantitative analysis (Figure 21) agree with that of PLS-DA and OPLS-DA.

4 Discussion

4.1 Polyols and ethyl acetate

In the present study, ethanol composes a large proportion of the wines, and the difference in the ethanol content of each wine is quite small. The ethanol signals of the samples were so intense in the spectra that they cover the signals of the other less abundant components. Therefore, ethanol was not a major discriminating compound.

Figure 15: PLS-DA score plot chart from ^1H NMR spectra of 2010 and 2012 vintage Cabernet Sauvignon wines.

Figure 16: OPLS-DA score plot chart from ^1H NMR spectra of 2010 and 2012 vintage Cabernet Sauvignon wines.

The 2,3-butanediol is a by-product of fermentation in wine, probably from the reduction of acetoin or pyruvic acid [1,18]. Because the taste threshold of 2,3-butanediol is 150 mg/L, it does not usually affect the flavor. However, the average content of 2,3-butanediol in each wine was approximately 0.243 g/L, which will make the wine slightly bitter with a sticky texture. In our study, the 2011 vintage wines contained the highest levels of 2,3-butanediol.

Glycerol is formed as a by-product of alcohol fermentation. The pH, sulfite concentration, grape variety, fermentation temperature, yeast, and nitrogen composition of the wine influence the level of glycerol [1,19,20]. In our

study, the winemaking conditions, such as the sulfite concentration, fermentation temperature, and yeast were approximately the same. Therefore, the glycerol contents may have resulted from the sugar contents in the grape berries.

Ethyl acetate in wine is the major ester produced by yeast, which can result in the sensory perception of volatile acidity, and has a certain odor of nail polish remover. When the content of ethyl acetate in wine is low, it can contribute fruity aroma properties, thereby increasing the complexity of the aroma and taste of the wine. Generally, the content of ethyl acetate is affected by yeast strains,

Figure 17: PLS-DA loading plot chart from ^1H NMR spectra of 2010 and 2012 vintage Cabernet Sauvignon wines.

Figure 18: PLS-DA score plot chart from ^1H NMR spectra of 2011 and 2012 vintage Cabernet Sauvignon wines.

fermentation temperature, amino acid content, and sulfur dioxide content in juice [21]. In this study, the 2012 Cabernet Sauvignon wine had the highest ethyl acetate content. Since the fermentation conditions and production process in different vintages were the same, the difference in ethyl acetate content may be due to the amino acid content in the juice.

4.2 Organic acids

Tartaric acid, malic acid, and citric acid in wine mostly derive from the grape berries. The concentration of tartaric

acid in grape berries usually remains stable despite increases in berry volume during maturation. Precipitation is related to the brewing conditions, including fermentation temperature, pH, and concentrations of potassium and calcium [1,22]. Therefore, tartaric acid in wines cannot be used as a biomarker for describing the characteristics of wines.

Higher lactate contents in wines show that malolactic fermentation has occurred, in which bacteria completely transform into lactic acid, citric acid, and malic acid [23,24]. Therefore, we cannot detect malic acid or citric acid in dry red wine.

Succinic acid is the main nonvolatile organic acid present during alcoholic fermentation and MLF [1]. As

Figure 19: OPLS-DA score plot chart from ^1H NMR spectra of 2011 and 2012 vintage Cabernet Sauvignon wines.

Figure 20: PLS-DA loading plot chart from ^1H NMR spectra of 2011 and 2012 vintage Cabernet Sauvignon wines.

one of the major metabolic products, succinic acid is very stable and does not change with age.

4.3 Amino acids

The wine amino acids have different origins. Some are released from dead yeast or at the end of fermentation, whereas some are indigenous to the grape and can be partially or fully metabolized by yeast; others are vinified by protein enzymatic degradation [25]. Classically, alanine is used in the growth of yeast in wine, so little is

detected in the finished wine product. Proline is not a nutrient used by yeast and can therefore be used as a biological marker of wine. Lee *et al.* [26] states that the proline content in wine depends on the environmental factors and grape varieties. Among the four different vintages of Cabernet Sauvignon wine tested, the 2009 vintage had the highest proline content, and the 2011 vintage had the lowest level of proline. This pattern may have resulted from the greater sunshine and lower rainfall in 2009.

Valine, another amino acid biomarker, was also revealed by the PLS-DA and OPLS-DA analyses. Valine

Figure 21: Comparison of the main metabolite concentrations in the 2009–2012 vintage wines. *The error bars indicate the standard deviations. S1, S2, S3, and S4 represent 2009, 2010, 2011, and 2012 vintages of Cabernet Sauvignon wines, respectively. (a): Valine content; (b): 2,3-butanediol; (c): ethyl acetate; (d): proline; (e): succinic acid; (f): choline; (g): glycerol; (h): gallic acid.

is used by yeast during fermentation and appears with yeast autolysis.

4.4 Choline

Choline is a precursor of glycine betaine, and betaine is related to homocysteine. The average level of choline in wines is 5.6 mg/100 g [27,28]. In our study, the 2012 vintage Cabernet Sauvignon wines had the highest levels of choline whereas the 2011 vintage had the lowest levels.

4.5 Carbohydrates

Glucose and fructose are the main sugars in grapes. When grape maturity begins, the glucose content in the grape is higher than the fructose content; both contents become nearly equal by harvest time. Dry wine refers to wine with a sugar level of less than or equal to 4.0 g/L. We detected sucrose, α -glucose, and β -glucose, and the differences in the concentrations of these three sugars were small. Therefore, we cannot use the carbohydrate in these wines as characteristic metabolites.

4.6 Cause of differences in the metabolites of Cabernet Sauvignon wines from different vintages

^1H NMR-based metabolomics were used to study the metabolite differences in various vintages of Cabernet Sauvignon wines. Pattern recognition showed clear differentiation between the wines vinified in 2009, 2010, 2011, and 2012. The metabolites used for the differentiation were 2,3-butanediol, ethyl acetate, valine, proline, succinic acid, lactate, acetic acid, glycerol, gallic acid, and choline. Wines were vinified using the same fermentation technique, yeast, and grape varieties. Therefore, climatic factors such as average temperature, rainfall, and evaporation were the main reasons for the differences in the wine metabolites of different vintages. Most likely, the higher the average temperature, higher the evaporation; and lower rainfall in 2009 increased the sugar content of the grapes and allowed the grapes to reach optimum ripeness. Therefore, the 2009 vintage wines have the highest levels of valine, 2,3-butanediol, gallic acid, and proline. Due to the lower average

temperature, higher rainfall, and higher evaporation in 2011 and 2012, these grapes experienced a long, slow ripening season. The 2011 vintage wines contained the highest level of lactic acid, and the highest levels of ethyl acetate, succinic acid, glycerol, and choline were detected in the 2012 vintage wines. Some metabolites were selected from the ^1H NMR spectra and quantified according to their peak areas. The results of the quantitative analysis agreed with the PLS-DA results.

5 Conclusion

This study shows that the NMR-based metabolomics approach can effectively classify wine. Certification of a vintage's geographical indications as well as adulteration and quality monitoring, provides the theoretical basis and technical support for this method.

Abbreviations

NMR	nuclear magnetic resonance
PCA	principal component analysis
PLS-DA	partial least squares discriminant analysis
OPLS-DA	orthogonal partial least squares discriminant analysis
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid

Acknowledgments: We gratefully acknowledge the support of the Great Wall Wines, COFCO Group, P. R. China, for providing the wine samples. The authors would like to thank Dr. William James Hardie, Dr. Guangnian Lu, Dr. Hua Li, and Mr. Yaqing Yue for their kind help in this work. We also would like to direct our deepest gratitude to the comments and suggestions by anonymous journal reviewers.

Research Funding: We acknowledge financial support by the National Natural Science Foundation of China (Project No. 31271857).

Author contributions: S. X. was in charge of data curation, formal analysis, software, validation, visualization, writing, reviewing, and editing; J. Z. contributed to data curation, formal analysis, and software; Q. Z. was involved in formal analysis and software; J. G. contributed to formal analysis and software; H. Z. was in charge of resources; B. H. was involved in funding acquisition,

investigation, methodology, project administration, data curation, supervision, validation, and writing of the original draft.

Conflict of interest: We confirm that none of the authors have any competing interests in the manuscript.

Data availability statement: The data sets generated during and/or analyzed during the current study are available with the corresponding author on reasonable request.

References

- [1] Son HS, Hwang GS, Ahn HJ, Park WM, Lee CH, Hong YS. Characterization of wines from grape varieties through multivariate statistical analysis of ^1H NMR spectroscopic data. *Food Res Int.* 2009;42:1483–91.
- [2] De Pascali S, Coletta A, Del Coco L, Basile T, Gambacorta G, Fanizzi FP. Viticultural practice and winemaking effects on metabolic profile of Negroamaro. *Food Chem.* 2009;161:112–9.
- [3] Rochfort S, Ezernieks V, Bastian SEP, Downey MO. Sensory attributes of wine influenced by variety and berry shading discriminated by NMR metabolomics. *Food Chem.* 2010;121:1296–304.
- [4] Forvefle L, Vercauteren J, Rutledge DN. Multivariate statistical analysis of two-dimensional NMR data to differentiate grapevine cultivars and clones. *Food Chem.* 1996;57:441–50.
- [5] Hu B, Yue Y, Zhu Y, Wen W, Zhang F, Hardie JW. Proton nuclear magnetic resonance-Spectroscopic discrimination of wines reflects genetic homology of several different grape (*V. vinifera* L.) cultivars. *PLoS One.* 2010;10:e0142840.
- [6] Son HS, Kim KM, van den Berg F, Hwang GS, Park WM, Lee CH, et al. ^1H nuclear magnetic resonance-based metabolomic characterization of wines by grape varieties and production areas. *J Agric Food Chem.* 2008;56:8007–16.
- [7] Pereira GE, Gaudillere JP, Van Leeuwen C, Hilbert G, Lavielle O, Maucourt M, et al. and chemometrics to characterize mature grape berries in four wine-growing areas in Bordeaux, France. *J Agric Food Chem.* 2005;53:6382–9.
- [8] Amaral FM, Caro MSB. Investigation of different pre-concentration methods for NMR analyses of Brazilian white wine. *Food Chem.* 2005;93:507–10.
- [9] Zhu J, Hu B, Lu J, Xu S. Analysis of metabolites in Cabernet Sauvignon and shiraz dry red wines from Shanxi by ^1H NMR spectroscopy combined with pattern recognition analysis. *Open Chem.* 2018;16:446–52.
- [10] Sun SY, Che CY, Sun TF, Lv ZZ, He SX, Gu HN, et al. Evaluation of sequential inoculation of *Saccharomyces cerevisiae* and *Oenococcus oeni* strains on the chemical and aromatic profiles of cherry wines. *Food Chem.* 2013;138:2233–41.
- [11] Papotti G, Bertelli D, Graziosi R, Silvestri M, Bertacchini L, Durante C, et al. Application of one- and two-dimensional NMR spectroscopy for the characterization of protected designation of origin Lambrusco wines of Modena. *J Agric Food Chem.* 2013;61:1741–6.
- [12] Bylesjö M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J Chemom.* 2006;341–51.
- [13] Nicholson JK, Lindon JC, Holmes E. “Metabonomics”: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica.* 1999;29:1181–9.
- [14] Anastasiadi M, Zira A, Magiatis P, Haroutounian SA, Skaltsounis AL, 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;57:11067–74.
- [15] Lee JE, Hong YS, Lee CH. Characterization of fermentative behaviors of lactic acid bacteria in grape wines through ^1H NMR- and GC-based metabolic profiling. *J Agric Food Chem.* 2009;57:4810–7.
- [16] Trygg J, Wold S. Orthogonal projections to latent structures (O-PLS). *J Chemom.* 2002;16(3):119–28.
- [17] Trygg J, Holmes E, Lundstedt T. Chemometrics in metabolomics. *J Proteome Res.* 2007;6:469–79.
- [18] Romano P, Fiore C, Paraggio M, Caruso M, Capece A. Function of yeast species and strains in wine flavor. *Int J Food Microbiol.* 2003;86:169–80.
- [19] Radler F, Schütz H. Glycerol production of various strains of *Saccharomyces*. *Am J Enol Vitic.* 1982;33:36–40.
- [20] Gardner N, Rodrigue N, Champagne CP. Combined effects of sulfites, temperature, and agitation time on production of glycerol in grape juice by *Saccharomyces cerevisiae*. *Appl Env Microbiol.* 1993;59:2022–8.
- [21] Rojas V, Gil JV, Piñaga F, Manzanares P. Acetate ester formation in wine by mixed cultures in laboratory fermentations. *Int J Food Microbiol.* 2003;86:181–8.
- [22] Viggiani L, Morelli MA. Characterization of wines by nuclear magnetic resonance: a work study on wines from the Basilicata region in Italy. *J Agric Food Chem.* 2008;56:8273–9.
- [23] Avenoza A, Busto JH, Canal N, Peregrina JM. Time course of the evolution of malic and lactic acids in the alcoholic and malolactic fermentation of grape must by quantitative ^1H NMR (qHNMR) spectroscopy. *J Agric Food Chem.* 2006;54:4715–20.
- [24] Larsen FH, van den Berg F, Engelsens SB. An exploratory chemometric study of ^1H NMR spectra of table wines. *J Chemom.* 2006;20:198–208.
- [25] Košir IJ, Kidrič J. Use of modern nuclear magnetic resonance spectroscopy in wine analysis: determination of minor compounds. *Anal Chim Acta.* 2002;458:77–84.
- [26] Lee JE, Hwang GS, Van Den Berg F, Lee CH, Hong YS. Evidence of vintage effects on grape wines using ^1H NMR-based metabolomic study. *Anal Chim Acta.* 2009;648:7–6.
- [27] Zeisel SH, Da Costa KA, Franklin PD, Alexander EA, Lamont JT, Sheard NF, et al. Choline, an essential nutrient for humans. *FASEB J.* 1991;5:2093–8.
- [28] Mickelbart MV, Chapman P, Collier-Christian L. Endogenous levels and exogenous application of glycinebetaine to grapevines. *Sci Hortic.* 2006;111:7–16.