Research Article Open Access

Jiangyu Zhu, Boran Hu*, Jie Lu, Shaochen Xu

Analysis of Metabolites in Cabernet Sauvignon and Shiraz Dry Red Wines from Shanxi by ^{1}H NMR Spectroscopy Combined with Pattern Recognition Analysis

https://doi.org/10.1515/chem-2018-0052
received January 12, 2018; accepted March 20, 2018.

Abstract: Metabolomics technology based on proton nuclear magnetic resonance (^{1}H NMR) spectroscopy combined with pattern recognition analysis was used to characterize the Cabernet Sauvignon and Shiraz dry red wines vinified in the Linfen of Shanxi Province, China, in 2016. The results showed that there was a very significant difference between the metabolites of Cabernet Sauvignon and Shiraz dry red wines from the area of Linfen. Compared with Shiraz dry red wines, Cabernet Sauvignon dry red wines contained higher levels of proline, valine, tartaric acid, citric acid, malic acid, gallic acid, β-glucose and ethyl acetate, whereas 2,3-butanediol, lactic acid, choline, glycerol, α-D-glucuronic acid, succinic acid and alanine were present in lower levels. Application of NMR spectroscopy combined with pattern recognition analysis showed the discriminative power between wine varietals from the same production area. The loading plot from partial least squares discriminant analysis (PLS-DA) indicated that the key biomarkers for this differentiation were proline, tartaric acid, glycerol, lactic acid, choline, succinic acid and gallic acid, which was consistent with the result of quantitative analysis.

Keywords: nuclear magnetic resonance (NMR); pattern recognition analysis; metabolites; Cabernet Sauvignon; Shiraz.

***Corresponding author: Boran Hu,** School of Food Science and Engineering, Yangzhou University, Yangzhou city, Jiangsu Province 225127, China, E-mail: huboran@yzu.edu.cn
Jiangyu Zhu, Jie Lu, Shaochen Xu: School of Food Science and Engineering, Yangzhou University, Yangzhou city, Jiangsu Province 225127, China

1 Introduction

Wine is one of the most popular beverages throughout the world. In recent years, with the globalization of viticulture and wine trade, vineyard growth has gradually increased in emergent viticultural countries. According to the data of 2013 form the International Organization of the Vine and Wine (OIV), wine production in China reached a high level to 2.1 Mhl [1]. Tremendous wealth is generated in wine sector, and the requirements of consumers to wine quality, security, and style diversification has also been more and more strict at the same time. How to improve the wine quality and enrich the diversification of the wine market has become an emergency issue [2]. The quality of the wines is determined by grape varieties, grape growing conditions, vintage, and the wine-making techniques. However, wine fraud such as adulteration and lack of authentication is constantly interfering the wine market. Hence, a robust analytical method in the verification of the wine authenticity and adulteration is necessary.

Recently the most popular and advanced method is metabolomics technology based on nuclear magnetic resonance spectroscopy. This method has great advantages of information analysis and structure determination, which can assess the entire process of fruit growing and wine making from a more direct perspective [3,4]. In previous reports, Son et al. confirmed the applicability of metabolomics based on NMR and multivariate statistical data sets in determining the grape varietals and production areas [5]. López et al. used nuclear magnetic resonance (NMR) technique to analyze the metabolites of wine samples from nine wineries in the same producing area [6]. It was found that NMR technology could accurately distinguish the wines from different wineries which were geographically very close, and isopentanol and isobutanol compounds were the key biomarkers for this discrimination. Godelmann et al. investigated a total of

about 600 samples of wines produced in Germany and the results indicated that the grape variety, the geographical origin, and the year of vintage of wines could be predicated 89% correctly on average by ^{1}H NMR spectroscopy combined with multivariate statistical analysis [7]. Licia et al. also identified the authenticity of the typical Italian *Denominazione di origine controllata* (DOC) wines by the same method [8]. At present, metabolomics technology based on nuclear magnetic resonance spectroscopy (^{1}H NMR) has entered a new stage and each NMR spectrum and could be regarded as the individual "fingerprint" of a wine sample, containing the information of variety, origin, vintage, physiological state, technological treatment, fermentation and other environmental factors, such as climate, soil composition and humidity, light and rainfall [9-12].

In emergent viticultural countries, the annual consumption of wine is very high, and most of their qualities are quite uneven, and wine fraud is also commonly seen in recent years. The widespread market of the corresponding wines deserves deeper studies on the traceability and authenticity of the wines [13]. The application of NMR-based metabolomics with multivariate statistical data sets in verification of the wine authenticity and adulteration should be expanded. In addition, it is necessary to establish "Fingerprints" library of wine in the main production areas in China. As two of the most popular and widely grown grapes in China, Cabernet Sauvignon and Shiraz from Linfen of Shanxi province was chosen in this research. Their characteristic metabolites were analyzed by NMR technique, and the biomarkers for the differentiation were identified by the pattern recognition analysis. This research provides a benchmark for further comparative study on the wine quality and the verification of the wine authenticity.

2 Methodology

2.1 Sampling

Cabernet Sauvignon and Shiraz grapes were grown in the Linfen area of Shanxi Province. The climate was warm and the accumulated temperature above 10°C was 4101 to 4600°C·d. Heat in this producing area was sufficient and the annual rainfall is 550-600mm, with the hydrothermal coefficient less than 1.5. The frost-free period was 180-220 days and the temperature difference between day and night was 12-20°C, which was suitable for grapes to grow and accumulate sugar. The mean content of soil available phosphorus, potassium and alkali-hydrolyzed nitrogen at 0-20 cm soil depths were 14.15, 141.80 and 63.29 mg/kg, respectively.

Single varieties of wines were all brewed in 2016 by a standard process using the following techniques: De-stemmed and crushed the fruits, and then added the yeast to ferment at 25°C for 8-10 days, followed by pressing the pomace gently. The wine was separated and tank-switched, and then it was sampled for pretreatment. Physical and chemical indicators of the wine samples were all in line with the standards of the International Organisation of Vine and Wine [14]. All samples were stored at -4°C.

2.2 Reagents and Apparatus

Oxalic acid, sodium oxalate (Shanghai Suyi Chemical Reagent Co., Ltd., China), heavy water (D_2O, deuterium content > 99.9%; Qingdao Tenglong Microwave Technology Co., Ltd., China), 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS).

Nuclear Magnetic Resonance Spectrometer (AVANCE600, Bruker Co., Ltd., German); Ultra-low Temperature Freezer (ULT178-6-V49, Revco Co., Ltd., USA); Vacuum Freeze Dryer (SNL315SV-230, Thermo Co., Ltd., USA); Micro Vortex Mixer (WH-3, Shanghai Huxi Analytical Instrument Co., Ltd., China); Desktop High-speed Centrifuge (TG16A-WS, Luxiangyi Centrifuge Instrument Co., Ltd., China);

2.3 ^{1}H NMR Spectroscopic Analysis

2.3.1 Sample Pretreatment

10 mL of wine samples were taken and then centrifuged at the speed of 3000 rpm for 20 min at -4°C. After the centrifugation, 3 mL of the supernatant was transferred and pre-frozen at -80°C overnight. After 48 h of freeze-drying, 400 μL of oxalate buffer (pH 4), 140 μL of D_2O and 60 μL of 0.5% DSS were added, and the mixture was centrifuged at 13,000 rpm for 20 min. At last, 500 μL of the supernatant was loaded into a 5 mm nuclear tube for NMR analysis. Each sample was tested 8 times.

2.3.2 NMR Data Collection

The ^{1}H NMR spectra of the wine samples were collected from the AVANCE 600 nuclear magnetic resonance

spectrometer. NMR analysis was carried out at a constant temperature of 298 K, and an H, C-sensitive cryogenic probe was used. The proton frequency and spectral width was 600.23 MHz and 7183.9 Hz, respectively. The number of sampling points was 32 K. The relaxation delay was set to 2 s and the sampling time was set to 2.3 s. For the mixing time, 100 ms was chosen. The linewidth enhancement factor was 0.3 Hz. The Noesygpprld sequence was used to suppress the water peak signal, and the number of scans was set at 256 times.

2.4 Multivariate Statistical Analysis

1. The chemical shift interval between 0-10.0 ppm in the nuclear magnetic resonance spectrum was integrated at the section of 0.005 ppm using Software AMIX. The residual ethanol peaks of 1.18-1.22 ppm and 3.57-3.72 ppm, residual DSS peaks of -0.5-0.5 ppm, 1.74-1.84 ppm and 2.90-2.95 ppm, and residual water peaks of 4.8-4.96 ppm were all removed.
2. The data were normalized and then introduced into Software SIMCA-P12.0 for pattern recognition analysis. Principal components analysis (PCA) was performed to visualize the acquired data, reduce the dimension of the high-dimensional data, remove the signal noise, and observe the discrete trend between samples.
3. To create a more reasonable regression model, partial least squares discriminant analysis (PLS-DA) was performed to sharpen the separation between observations groups. A maximum separation among classes could be obtained by rotating PCA components. In addition, PLS-DA is also helpful to understand which component carries the class separating information.
4. An external model validation experiment was carried to verify the degree of fitting of PLS-DA.

Metabolite content was obtained by calculating the ratio of the peak area of the protons on a given group of the substance to be measured, to that of the internal standard DSS in the one-dimensional ^{1}H-NMR spectrum [15]. The mass concentration u (g /L) of metabolites was calculated as follows:

$$u = \frac{m_S}{V} = \frac{(A_S/n_S)\times M_s}{(A_R/n_R)\times M_R} \times \frac{m_R}{V}$$

Note: A_S is the integral area of the selected signal for the tested sample; A_R is the integral area of the selected signal for the internal standard DSS; n_s is the number of protons contained in the tested sample by the integral signal; n_R is the number of protons contained in the internal standard DSS by the integral signal; M_S is the relative molecular mass of the tested sample; M_R is the relative molecular mass of the internal standard DSS; m_R is the mass of the internal standard DSS.

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

3 Result and Discussion

3.1 Pattern Recognition of Cabernet Sauvignon and Shiraz Dry Red Wine from the Same Producing Area

As illustrated from the ^{1}H NMR spectrum of Cabernet Sauvignon and Shiraz (Figure 1), the sensitivity of the proposed method was good and the signal of the major metabolites of Cabernet Sauvignon and Shiraz red wine could be well separated and recognized. Combining relevant literature [16-19] and the NMR spectra obtained in this experiment, the associated chemical shift of the main metabolites was present in Table 1.

The number of peaks a signal has was classified: s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet) and dd (doublet of doubles).

At present, NMR fingerprinting technique combined with pattern recognition analysis has been widely used in many fields [20-22]. NMR data of the dry red wine samples were imported into SIMCA P-12.0 software for principal component analysis. PCA scores plot derived from the ^{1}H NMR spectra of Cabernet Sauvignon and Shiraz dry red wines was illustrated in Figure 2. As seen from the score plot, there was an obvious distinction between

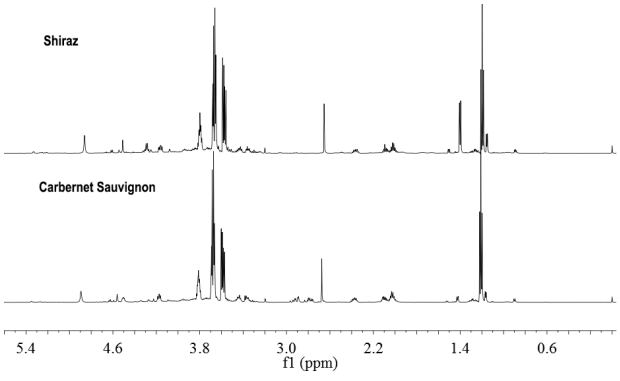

Figure 1: ^{1}H NMR spectrum of Cabernet Sauvignon and Shiraz dry red wines.

Table 1 ^{1}H NMR assignment of metabolites in Cabernet Sauvignon and Shiraz dry red wines.

No.	Metabolites	NMR Chemical Shifts
1	Valine	0.90(d, C4H_3), 1.02(d, C5H_3)
2	Ethanol	1.19(t, C2H_3), 3.67(q, C1H_2)
3	2,3-Butanediol	1.16(d,C1H_3+C4H3)
4	Succinic acid	2.68(s, C2H_2+C3H_2)
5	Proline	2.00(m, u, γ-CH_2), 2.07(m, u, β-CH), 2.35(m, u, β'-CH), 3.35(m, u, δ-CH), 3.40(m, u, δ-CH), 4.13(m, u, α-CH)
6	Ethyl acetate	1.26(t, C4H_3), 4.18(q, C3H_2)
7	Tartaric acid	4.57(s, C2H+C3H)
8	α-Glucose	5.25(d, αC1H)
9	β-Glucose	4.64 (d, βC1H)
10	Lactic acid	1.42(d,C3H_3), 4.29(m,C2H)
11	Gallic acid	7.15(s,C2H+C6H)
12	Glycerin	3.58(q,C2H_2), 3.67(q,C3H_2), 3.81(m,C1H)
13	Citric acid	2.82(d, C2H_a+C4H_a), 2.94(d, C2H_b+C4H_b)
14	Choline	3.20(s,N-CH_3), 3.52(t,αCH_2), 4.09(t, βCH_2)
15	α-D-Glucuronic acid	5.35(d,C1H)
16	Malic acid	2.78(dd, βCH_2), 2.90(dd, β'CH_2), 4.50(q, CH)
17	Alanine	1.52(d, βCH_3)

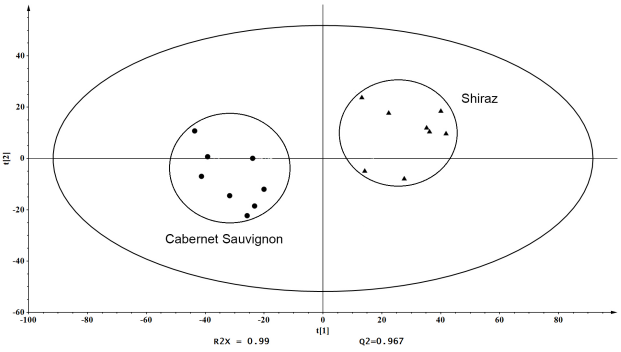

Figure 2: PCA scores plot based on the ^{1}H NMR spectra of Cabernet Sauvignon and Shiraz dry red wines. Black solid circles and black solid triangles represent Cabernet Sauvignon and Shiraz dry red wines samples, respectively.

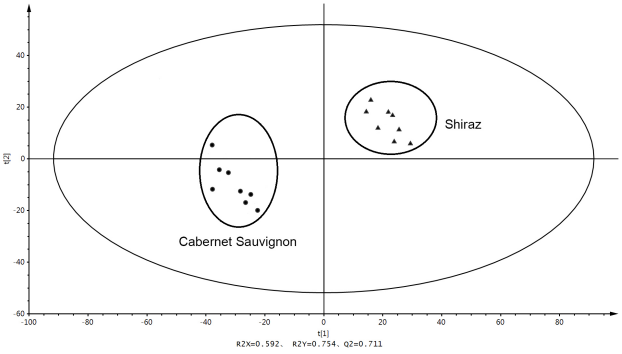

Figure 3: PLS-DA scores plot based on the ^{1}H NMR spectra of Cabernet Sauvignon and Shiraz dry red wines. Black solid circles and black solid triangles represent Cabernet Sauvignon and Shiraz dry red wines samples, respectively.

Cabernet Sauvignon and Shiraz dry red wine, revealing the significant difference of the metabolites from these two wines. The cumulative contribution rate, R^2X = 0.99, and Q^2 = 0.967, indicating that the established PCA model was of good quality.

After the orthogonality correction, the PLS-DA model was established and the PLS-DA scores plot was shown in Figure 3. The cumulative contribution rate R^2X = 0.592, R^2Y = 0.754 and Q^2 = 0.711. All three values were greater than 0.5, indicating that the model is valid. As seen from the PLS-DA scores plot, the distinction between Cabernet Sauvignon and Shiraz dry red wines was more obvious than that in PCA scores plot.

A permutation test was used to verify the fitting degree of PLS-DA model, mainly evaluated by the slope of the regression line and the intercept of the regression line with the vertical axis. When the slope of the regression line was greater and the intercept was smaller, it indicated that there were more data to interpret the model and the predictive ability of the model was better [23]. Besides, it could also be evaluated by comparing the difference between R^2 and Q^2. The smaller the difference between the two values was, the smaller the difference between the data explained by the model and the predicted data was, indicating that the predictive ability of the model was excellent [24]. As seen from the validation plots of the permutation test in PLS-DA model (Figure 4), the values of R^2 and Q^2 did not exceed the quality parameter of the actual model in any arrangement, once again demonstrating that the model was reliable and predictive.

3.2 Metabolites analysis in Cabernet Sauvignon and Shiraz red wines

Major metabolites that made difference could be obtained from the PLD-DA loading plot. In the modeling process, values unrelated to the classification would be filtered out by orthogonal signal correction (OSC) [25].

PLS-DA loading plot of Cabernet Sauvignon and Shiraz dry red wines vinified at Shanxi in 2016 was shown in Figure 5. The higher peak in the loading plot indicated

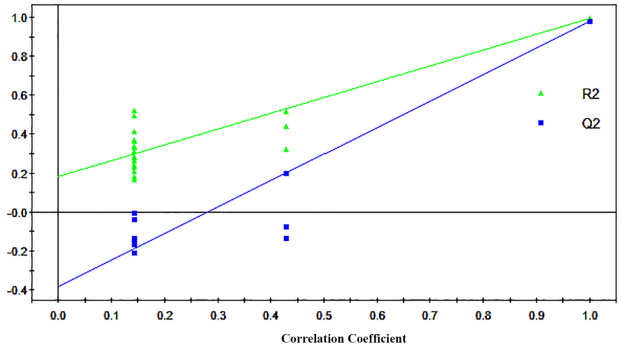

Figure 4: Validation plots based on the ^{1}H NMR spectra of Cabernet Sauvignon and Shiraz dry red wines. Green triangles and blue squares represent R^2 and Q^2, respectively.

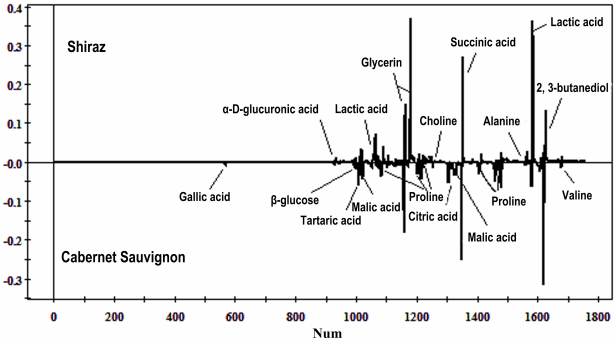

Figure 5: PLS-DA loading plot based on the ^{1}H NMR spectra of Cabernet Sauvignon and Shiraz dry red wines.

that content of the corresponding metabolite was higher in dry red wine, and the lower one indicated that content of the corresponding metabolite was lower. As seen from the loading plot, compared with Shiraz red wine vinified in Shanxi, Cabernet Sauvignon contained higher levels of proline, valine, tartaric acid, citric acid, malic acid, gallic acid, β-glucose and ethyl acetate, while content of 2,3-butanediol, lactic acid, choline, glycerin, α-D-glucuronic acid, succinic acid and alanine was lower. In Shiraz dry red wine, content of 2, 3-butanediol, choline, glycerol, succinic acid, lactic acid, alanine content was relatively higher whereas proline, tartaric acid and gallic acid was in lower levels. The main contributors to the differences were proline, tartaric acid, glycerin, lactic acid, choline, succinic acid and gallic acid.

The main metabolites of Cabernet Sauvignon and Shiraz were quantitatively analyzed and the result was as shown in Figure 6. Significant difference between the main metabolites content was observed, consistent with the result obtained from PLS-DA loading plot.

Proline was the main amino acid in wines [26], and it was also the amino acid with the largest difference in content between the two samples. Proline has both, sweetness and bitterness, and is a complex amino acid. The content of proline had a direct impact on the depth and complexity of wine taste and flavor. Since proline in wines was not consumed by fermentation and maturation,

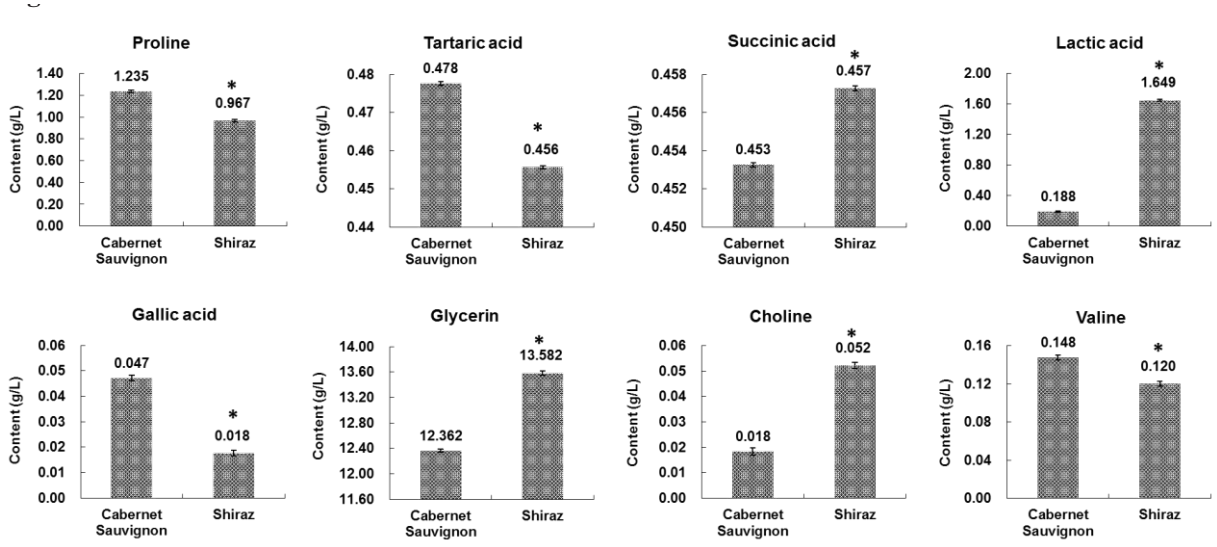

Figure 6: Content of the main metabolites in Cabernet Sauvignon and Shiraz dry red wines. Asterisks indicate the significant difference ($p < 0.05$) between the metabolite content of Cabernet Sauvignon and Shiraz dry red wine.

the difference in the content of proline here is mainly attributed to the different wine varietals. For tartaric acid, it is a non-volatile acid, which could impart a dark red color to the wine. Its degree of enrichment in wines often depends on the grape varieties and the soil of the vineyard. For malic acid and lactic acid, the content of both does not simply depend on the type of grape. Unlike tartaric acid, in the process of grape maturation, malic acid is consumed by respiration, resulting in the decrease in acidity. In addition, malic acid will be further converted to softer lactic acid during the lactic fermentation of wine brewing [27]. According to the PLS-DA loading plot, we know that the content of citric acid in Cabernet Sauvignon is significantly higher than that in Shiraz, but the content of lactic acid is lower. It may be attributed to the difference in grape varieties, or the degree of fermentation.

Compared to Cabernet Sauvignon, Shiraz dry red wine contained a higher level of succinic acid and glycerol. Succinic acid is a by-product of yeast nitrogen metabolism during wine fermentation which has a mild fruity flavor, while glycerin is non-volatile and unscented, it does not influence the aroma of wine, but its sweetness and stickiness will also lead to the difference of sweetness and wine body thickness between Cabernet Sauvignon and Shiraz dry red wines.

4 Conclusion

In this study samples of Cabernet Sauvignon and Shiraz red wines vinified at Shanxi in 2016 were analyzed by ^{1}H NMR, and the results showed that there was a notable difference between the metabolites of Cabernet Sauvignon and Shiraz. NMR technique combined with pattern recognition could excellently distinguish the two wines from each other, and the model established in this experiment was reliable, accurate and predictive.

The results provide a benchmark for further comparative study on the wine quality and the verification of the wine authenticity. The key contributors to the difference identified in this study may also be used to establish a quality evaluation system specifically for Cabernet Sauvignon and Shiraz dry red wines in China.

Acknowledgements: The present work is supported by the National Natural Science Foundation of China (Project No.31271857).

Conflict of interest: Authors state no conflict of interest.

5 References

[1] Alañón M.E., Pérez-Coello M.S., Marina M.L., Wine science in the metabolomics era, TrAC-Trend Anal. Chem., 2015, 74, 1-20.
[2] Famularo B., Bruwer J., Li E., Region of origin as choice factor: wine knowledge and wine tourism involvement influence, Int. J. Wine Bus. Res., 2010, 22, 362-385.
[3] López-Rituerto E., Savorani F., Avenoza A., et al., Investigations of La Rioja terroir for wine production using 1H NMR metabolomics, J. Agric. Food Chem., 2012, 60, 3452-3461.
[4] Young-Shick H., NMR-based metabolomics in wine science, Magn. Reson. Chem., 2011, 49, 13-21.
[5] Son H.S., Kim K.M., Van Den Berg F., et al., ^{1}H nuclear magnetic resonance-based metabolomic characterization of wines by grape varieties and production areas, J. Agric. Food Chem., 2008, 56, 8007-8016.
[6] López-Rituerto E., Savorani F., Avenoza A., et al., Investigations of La Rioja terroir for wine production using ^{1}H NMR metabolomics, J. Agric. Food Chem., 2012, 60, 3452-3461.
[7] Godelmann R., Fang F., Humpfer E., et al., 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.
[8] Caruso M., Galgano F., Castiglione Morelli M.A., et al., Chemical profile of white wines produced from "Greco bianco" grape variety in different Italian areas by nuclear magnetic resonance (NMR) and conventional physicochemical analyses. J. Agric. Food Chem., 2011, 60, 7-15.
[9] Son H.S., Hwang G.S., Kim K.M., et al., ^{1}H NMR-based metabolomic approach for understanding the fermentation behaviors of wine yeast strains, Anal. Chem., 2008, 81, 1137-1145.
[10] Son H.S., Hwang G.S., Kim K.M., et al., Metabolomic studies on geographical grapes and their wines using ^{1}H NMR analysis coupled with multivariate statistics, J. Agric. Food Chem., 2009, 57, 1481-1490.
[11] Cassino C., Tsolakis C., Bonello F., et al. Effects of area, year and climatic factors on Barbera wine characteristics studied by the combination of ^{1}H-NMR metabolomics and chemometrics, J. Wine Res., 2017, 28, 259-277.
[12] Du Y.Y., Bai G.Y., Zhang X., et al. Classification of wines based on combination of ^{1}H NMR spectroscopy and principal component analysis, Chin. J. Chem., 2007, 25, 930-936.
[13] Viggiani L., Morelli M.A.C., Characterization of wines by nuclear magnetic resonance: a work study on wines from the Basilicata region in Italy, J. Agric. Food Chem., 2008, 56, 8273-8279.
[14] Oiv O., Compendium of International methods of wine and must analysis, International Organisation of Vine and Wine: Paris, France, 2009, 154-196.
[15] Hu B., Yue Y., Zhu Y., et al., Proton nuclear magnetic resonance-spectroscopic discrimination of wines reflects Genetic homology of several different grape (*V. vinifera* L.) cultivars, PloS one, 2015, 10, e0142840.
[16] Brescia M A., Caldarola V., De Giglio A., et al., Characterization of the geographical origin of Italian red wines based on traditional and nuclear magnetic resonance spectrometric determinations, Anal. Chim. Acta, 2002, 458, 177-186.

[17] Anastasiadi M., Zira A., Magiatis P., et al., ^{1}H 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.
[18] Košir I.J., Kidric J., Identification of amino acids in wines by one-and two-dimensional nuclear magnetic resonance spectroscopy, J. Agric. Food Chem.,2001, 49, 50-56.
[19] Košir I.J., Kidrič J., Use of modern nuclear magnetic resonance spectroscopy in wine analysis: determination of minor compounds, Anal. Chim. Acta, 2002, 458, 77-84.
[20] Fotakis C., Kokkotou K., Zoumpoulakis P., et al., NMR metabolite fingerprinting in grape derived products: An overview, Food Res. Int., 2013, 54, 1184-1194.
[21] Ward J.L., Harris C., Lewis J, et al., Assessment of ^{1}H NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of Arabidopsis thaliana, Phytochemistry, 2003, 62, 949-957.
[22] Krishnan P., Kruger N.J., Ratcliffe R G., Metabolite fingerprinting and profiling in plants using NMR, J. Exp. Bot., 2004, 56, 255-265.
[23] Szymańska E., Saccenti E., Smilde A.K., et al., Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies, Metabolomics, 2012, 8, 3-16.
[24] Bylesjö M., Rantalainen M., Cloarec O., et al., OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification, J. Chemom., 2006, 20, 341-351.
[25] Esteban-Dıez I., González-Sáiz J.M., Pizarro C., An evaluation of orthogonal signal correction methods for the characterisation of arabica and robusta coffee varieties by NIRS, Anal. Chim. Acta, 2004, 514, 57-67.
[26] Lehtonen P., Determination of amines and amino acids in wine—a review, Am. J. Enol. Vitic., 1996, 47, 127-133.
[27] Bauer R., Dicks L.M.T., Control of malolactic fermentation in wine. A review, S. Afr. J. Enol. Vitic, *25*, 74-88.