DE GRUYTER OPEN

Nova Biotechnol Chim (2017) 16(2): 132-137
DOI: 10.1515/nbec-2017-0018

Nova Biotechnologica *et* Chimica

Distinguishing between juniper-flavoured spirit drinks from different producers

Veronika Uríčková✉, Kinga Dobóová and Jana Sádecká

Institute of Analytical Chemistry, Faculty of Chemical and Food Technology, Slovak University of Technology, Radlinského 9, Bratislava, SK-812 37, Slovak Republic

Article info

Article history:
Received: 27th March 2017
Accepted: 6th December 2017

Keywords:
Beverage
Chemometrics
Fluorescence spectroscopy

Abstract

A simple method for classifying juniper-flavoured spirit drinks is proposed based on the ratio of fluorescence intensity values in synchronous fluorescence spectra. Receiver operating curves (ROC) and linear discriminant analysis (LDA) were used to compute the performance of the classification. Significant differences in the fluorescence intensity ratios (I_{316}/I_{287} and I_{324}/I_{287}) observed in the spectra recorded using wavelength difference 10 nm were evaluated by ROC analysis to identify cut-off values that gave ideal AUCs equal to one, thus allowing for 100% correct classification of the samples according to producer criteria. LDA showed that drinks of different producers could be distinguished (100% correct classification) on the basis of their differences in the fluorescence intensity ratios (I_{323}/I_{287}, I_{324}/I_{287}, I_{316}/I_{287} and I_{325}/I_{287}). These results show that complete synchronous spectra are not required to discriminate between producers. Instead of them, fluorescence intensity could be measured at selected wavelengths.

Introduction

Juniper-flavoured spirit drinks are spirit drinks produced by flavouring ethyl alcohol of agricultural origin with juniper (*Juniperus communis* L. and/or *Juniperus oxicedrus* L.) berries. The minimum alcoholic strength by volume of juniper-flavoured spirit drinks shall be 30%. For a process without fermentation juniper berries are first slightly squeezed and then prepared with 30% drinkable alcohol. After a resting period the distillation is initiated. In enterprises where juniper mashes are still fermented the production of the brandy is done in a two-step distillation. A gas chromatograph coupled with a mass spectrometer is a configuration often used in the analysis of alcoholic beverages (Vichi *et al.* 2008). Automated sequential multidimensional GC/MS is a technique capable of producing matrix-specific libraries of complex products. Spectral deconvolution of GC/MS data based on these libraries provides a reliable, unambiguous means of tracking the genealogy of juniper berry content from raw materials to final products (gins) and provides a more rationale means for detecting adulterants (Robbat *et al.* 2011). However multidimensional GC/MS method is relatively expensive, time-consuming and requires highly skilled operators. Up to date, several simple and rapid methods such as ultraviolet (UV), visible (VIS), infrared (IR)and fluorescence spectroscopies have been tried for determining the origin of beverages (Shen *et al.* 2012; Azcarate *et al.* 2013; Martelo-Vidal *et al.* 2013). The last technique is particularly attractive because of its high sensitivity and excellent specificity. By combining fluorescence

✉ *Corresponding author:* veronika.urickova@stuba.sk

spectroscopy and chemometric method discrimination of red wines according to grape variety (Airado-Rodríguez *et al.* 2011; Saad *et al.* 2016; Silvestri *et al.* 2014; Yin *et al.* 2009), typicality (Dufour *et al.* 2006; Yin *et al.* 2009), manufactures (Yin *et al.* 2009) and geographical origin (Dufour *et al.* 2006) or reliable classification of white wines according to grape variety (Azcarate *et al.* 2015) can be successfully achieved. Furthermore, adulterations of brandy can be identified and determined by using chemometric methods even if slight fluorescent spectral variations are observed for the samples (Markechová *et al.* 2014). The combination of fluorescence spectroscopic data with UV/VIS and near IR data has improved the grouping of single-malt whiskies according to their geographic origin (Mignani *et al.* 2012). In our previous study, we applied general discriminant analysis and support vector machine to synchronous fluorescence spectra of juniper-flavoured spirit drinks from different producers, thereby obtaining 100% correct classification of juniper-flavoured spirit drinksof three producers (Uríčková *et al.* 2015).

In this paper, a simplified method for classifying juniper-flavoured spirit drinks is proposed based on the ratio of fluorescence intensity values in synchronous fluorescence spectra. Receiver operating curves (ROC) and linear discriminant analysis (LDA) were used to compute the performance of the classification.

Experimental

Samples

A total of thirty-two commercially available samples from three Slovak producers (code S1–S3) were collected in 2015 – 2016. Different products from the same producer and four bottles of the same product were sampled: 16 samples [Klasik Slovenská borovička, Klasik Inovecká borovička, Klasik borovička jemná, Borovička Borec (S1)], 12 samples [Trenčianska Juniperus borovička, Juniperus borovička, Koniferum borovička (S2)] and 4 samples [Borovička Zlatá (S3)]. The alcoholic degree ranged within 35 – 42% ethanol. The samples were stored at room temperature and analyzed without any prior treatment.

Fluorescence spectroscopy

The acquisition of fluorescence spectra was performed using a Perkin-Elmer LS 50 Luminescence spectrometer equipped with a Xenon lamp, a 10 mm x 10 mm x 45 mm quartz cell and FL Data Manager Software for spectral acquisition and data processing. The slits of monochromators, scan speed, acquisition interval and integration time were set at 5 nm, 200 nm min^{-1}, 1 nm and 0.1 s, respectively. Synchronous fluorescence (SF) spectra were collected by simultaneously scanning the excitation and emission monochromators in the excitation wavelength range from 200 to 450 nm, with constant wavelength differences $\Delta\lambda$ between them. The values of $\Delta\lambda$ were varied from 10 to 100 nm, in steps of 10 nm. SF spectrum was a plot of the variations in fluorescence intensity as a function of the excitation wavelength for a fixed $\Delta\lambda$. Three spectra were recorded for each sample and the average of the three replicates was used for further analysis. Fluorescence intensities were plotted as a function of the excitation wavelength.

Software

Data were exported to ASCII and processed with the Microsoft Office Excel 2010 software. STATISTICA version 7.0 (StatSoft, USA, 2004) was used for LDA. Univariate ROC curve analysis available on MetaboAnalyst home page http://www.metaboanalyst.ca was used (Xia and Wishart 2016).

Results and Discussion

Synchronous fluorescence spectra

In SF, a signal is observed only when $\Delta\lambda$ is in accord with the interval between one excitation band and one emission band. Thus, the shape and intensity of the SF spectra depend on the $\Delta\lambda$ value used. Fig. 1 shows the averaged SF spectra of samples from the three producers recorded at $\Delta\lambda$ = 10 nm, 20 nm, 30 nm and 40 nm. Regarding $\Delta\lambda$ = 10 nm (Fig. 1A), SF spectrum of the S1 brands showed two overlapping bands in the wavelength range from 308 nm to 330 nm with

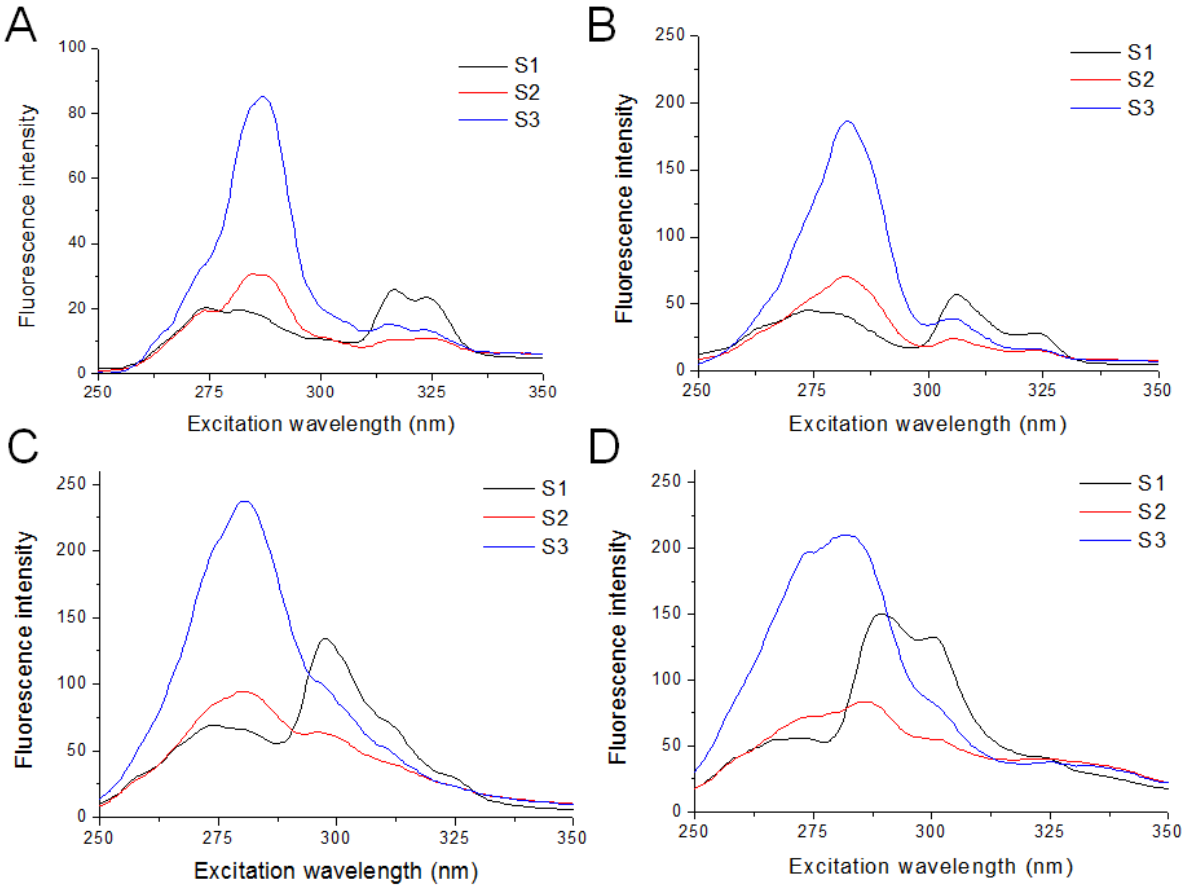

Fig. 1. Averaged synchronous fluorescence spectra recorded at (a) Δλ = 10 nm, (b) Δλ = 20 nm, (c) Δλ = 30 nm and (d) Δλ = 40 nm.

Table 1. The ratio of fluorescence intensity (I) values for spirits of different producers.

Parameter	Producer	Mean	Standard deviation	Minimum	Median	Maximum
I_{316}/I_{287} (Δλ = 10 nm)	S1	1.290	0.407	0.890	1.131	2.028
	S2	0.415	0.124	0.261	0.424	0.634
	S3	0.176	0.006	0.170	0.175	0.183
I_{324}/I_{287} (Δλ = 10 nm)	S1	1.197	0.347	0.879	1.024	1.920
	S2	0.438	0.135	0.260	0.470	0.640
	S3	0.156	0.006	0.148	0.156	0.161
I_{306}/I_{282} (Δλ = 20 nm)	S1	1.200	0.437	0.818	1.033	1.978
	S2	0.354	0.031	0.316	0.364	0.406
	S3	0.209	0.005	0.203	0.210	0.214
I_{298}/I_{280} (Δλ = 30 nm)	S1	1.839	0.615	1.378	1.545	2.953
	S2	0.697	0.076	0.601	0.705	0.816
	S3	0.408	0.008	0.401	0.408	0.418

maxima at 316 and 324 nm as well as two overlapping less intense bands in the wavelength range from 260 nm to 290 nm with maxima at about 274 and 281 nm. SF spectrum of the S2 brands showed a maximum at about 286 nm, a shoulder at 273 nm and less intense band in the wavelengths from 310 nm to 330 nm with slightly observable maxima at about 315 and 324 nm. SF spectrum of the S3 brand showed the highest fluorescence of all the samples with a maximum at about 287 nm and two less intense overlapping bands with maxima at 314 and 324 nm. Regarding Δλ = 20 nm (Fig. 1B), broadening in spectral bands, increasing fluorescence intensity of all

Table 2. Area under the ROC curve and cut-off value for discriminating different producers using fluorescence intensity ratios.

Δλ (nm)	Ratio of fluorescence intensity	S1 and S2 versus S3		S1 versus S2 and S3	
		AUC	cut-off	AUC	cut-off
10	I_{316}/I_{287}	1	0.222	1	0.762
10	I_{324}/I_{287}	1	0.211	1	0.759
20	I_{306}/I_{282}	0.964	0.265	0.938	0.612

bands and changes in their relative intensities were observed. In addition, fluorescence maxima were blue-shifted to 262, 274, 306 and 323 nm for S1 brands; to 263, 282 and 304 nm for S2 brands or to 282, 305 and 323 nm for S3 brand. At higher Δλ values, a further increase in fluorescence intensity together with broadening or disappearance of the bands was observed. Fig. 1 suggests the potential of SF spectroscopy to differentiate between juniper-flavoured spirit drinks from different producers, in particular on the basis of SF spectra recorded at small Δλ value.

A detailed analysis of the spectral features of 32 samples obtained at Δλ = 10 nm showed that the ratio of fluorescence intensity values at 316 nm and 287 nm (I_{316}/I_{287}) was 1.209 ± 0.407 (mean ± SD), 0.415 ± 0.124 and 0.176 ± 0.006 for S1, S2 and S3 samples, respectively. Other relevant differences in the ratio of fluorescence intensity values at 324 nm and 287 nm (I_{324}/I_{287}), at 306 nm and 282 nm (I_{306}/I_{282}) and at 298 nm and 280 nm (I_{298}/I_{280}) were observed in the spectra recorded using Δλ = 10 nm, 20 nm and 30 nm, respectively (Table 1). It is noticed that both fluorescence intensity ratios calculated from S1 group are always higher than those of S2 and S3 groups.

Receiver operator characteristic curve

The main criteria (cut-off) for classifying the samples were based on receiver operator characteristic (ROC) curves, which are often used to determine the cut-off point based on which subjects will be classified as either a positive or negative outcome (Xia *et al.* 2013). Univariate ROC curve analysis available on MetaboAnalyst home page (http://www.metaboanalyst.ca) was used (Xia and Wishart 2016). MetaboAnalyst accepts a table with the sample values in rows and the feature labels in columns. The first column was a set of sample labels (S1, S2 and S3), the second column was a set of class labels (0 or 1) and next four columns was the ratio of fluorescence intensity values matrix. To find cut-off point for S1 samples, the class label was assigned a value of 0 for S1 samples and 1 for all other samples. To find cut-off point for S3 samples, the class label was assigned a value of 0 for S1 and S2 samples and 1 for S3 samples. The results for distinguishing S1 samples from S2 and S3 samples, and for separating S1 and S2 samples from S3 samples are shown in Table 2. Univariate ROC curve analysis resulted in the areas under the curve (AUC) and cut-off values (Table 2) with a probability of 0.95. In general, the AUC close to 0.5 means poor discrimination, whereas the AUC higher than 0.9 indicates excellent separation between the two classes. Regarding ROC analysis of data based on spectra recorded at Δλ = 10 nm, ideal AUCs equal to one were obtained for I_{316}/I_{287} ratio.

The first cut-off value was 0.222, below which the sample was deemed as belonging to S3. Above the second cut-off value, 0.762, the sample was deemed as belonging to S1. Similar cut-off values were obtained for I_{324}/I_{287} ratio (Table 2) again based on ideal AUCs. Regarding ROC analysis of data based on spectra recorded at Δλ = 20 nm, slightly smaller AUCs were obtained for I_{306}/I_{282} ratio. Based on cut-off values 0.265 and 0.612, one sample was incorrectly classified in both cases. ROC analysis showed that I_{298}/I_{280} ratio (Δλ = 30 nm) was not significant either for distinguishing S1 samples from S2 and S3 samples or for separating S1 and S2 samples from S3 samples.

Linear discriminant analysis (LDA)

In the next part, LDA was applied to the ratio of fluorescence intensity values $I_{315.5}/I_{287}$, I_{316}/I_{287}, $I_{316.5}/I_{287}$, I_{317}/I_{287}, I_{323}/I_{287}, I_{324}/I_{287} and I_{325}/I_{287} based on spectra recorded at Δλ = 10 nm. LDA was

Table 3. Stepwise discriminant analysis and raw (standardized) coefficients of discriminant functions.

Step	Variable	Wilks' lambda	Partial lambda	p-level	Raw (standardized) coefficients	
					Root 1	Root 2
1	I_{323}/I_{287}	0.177	0.634	0.002	37.37 (9.22)	-17.09 (-4.22)
2	I_{324}/I_{287}	0.142	0.791	0.047	-33.71 (-8.88)	23.88 (6.29)
3	I_{316}/I_{287}	0.160	0.700	0.009	2.65 (0.82)	-22.90 (-7.05)
4	I_{325}/I_{287}	0.123	0.913	0.307	-1.47 (-0.37)	19.62 (4.99)
				Constant	-3.88	-1.46

used to determine which variables most significant discriminate between three naturally occurring groups and in addition to find classification functions to predict group membership. To select the least number of variables, stepwise LDA was performed and Wilks' lambda was calculated at each step. The variable with the smallest Wilks' lambda that improves classification was entered into the analysis. The results of the applied LDA, according to producer criteria for each step are summarized in Table 3. Variables I_{315}/I_{287}, $I_{316.5}/I_{287}$ and I_{317}/I_{287} were excluded by a stepwise method, taking into account their lower variability observed through the different producer samples.

The Partial Wilks' Lambda (the smaller the Partial Wilks' Lambda, the greater is the contribution to the overall discrimination) indicates that variable I_{323}/I_{287} contributes most, variable I_{316}/I_{287} second most, variable I_{324}/I_{287} third most, and variable I_{325}/I_{287} contributes least to the overall discrimination. New two discriminant functions (Roots), linear combinations of variables selected, were obtained to discriminate the different producers. The following two eigenvalues and canonic correlations (given in parentheses) were calculated: 4.964 (0.912) and 0.492 (0.574) explaining about 90.9 % and 100% of the total variance, respectively. Raw canonical discriminant function coefficients obtained are also reported in Table 3.

A two-dimensional plot of discriminant functions derived from the four selected variables is shown in Fig. 2. The first discriminant function (Root1) mostly discriminates between S1 and the two others (S2 and S3). The second function (Root2) provides some discrimination between S3 (all show negative values) and S2 (which have mostly positive values). However, the discrimination is not as clear as that provided by the Root1. Standardized canonical discriminant function coefficients (given in parentheses in Table 3) indicate that the first discriminant function is marked mostly by variables I_{323}/I_{287} and I_{324}/I_{287}, while the second function is weighted mostly by variables I_{324}/I_{287} and I_{316}/I_{287} and to a lesser extent by the other two variables.

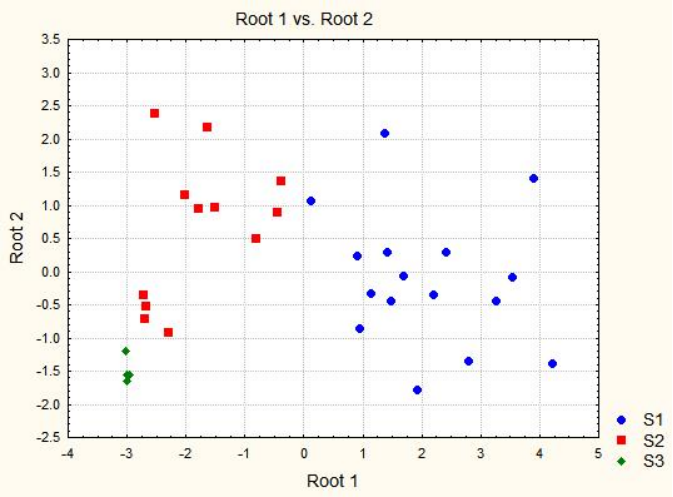

Fig.2. Projections of samples according to producer criteria in the space formed by the two discriminant functions.

Finally, the two classification functions were used for the classification of samples, with the result that 100% of samples were classified correctly according to producer criteria. The performance of the LDA model was evaluated using the leave-one-out-cross-validation approach, which is the best alternative for small number of samples – less than 50 (Molinaro *et al.* 2005). In this kind of validation, the sample set, itself, was used to validate the model. The model was repeatedly 32 times calculated leaving out a single sample and then used to predict the left-out sample. In cross-validation step, all S1 and S3 samples were again classified correctly, however, three of S2 samples

were specified as belonging to S3 group, leading to 90.6 % correct classification

Conclusions

Significant differences in the fluorescence intensity ratios (I_{316}/I_{287} and I_{324}/I_{287}) observed in the spectra recorded using $\Delta\lambda = 10$ nm were evaluated by ROC analysis to identify cut-off values that gave ideal AUCs equal to one, thus allowing for 100% correct classification of the samples according to producer criteria. LDA showed that drinks of different producers could be distinguished on the basis of their differences in the fluorescence intensity ratios (I_{323}/I_{287}, I_{324}/I_{287}, I_{316}/I_{287} and I_{325}/I_{287}). All of the 32 samples that were used as input data for the analysis were also classified correctly. In addition, the results obtained by both ROC and LDA were similar. These results show that complete synchronous spectra are not required to discriminate between producers. Instead of them, fluorescence intensity could be measured at selected wavelengths.

Acknowledgement

This publication was supported by the Competence Center for SMART Technologies for Electronics and Informatics Systems and Services, ITMS 26240220072, funded by the Research & Development Operational Programme from the ERDF. This work was supported by the STU Grant Scheme for Support of Excellent Teams of Young Researchers.

References

Airado-Rodríguez D, Durán-Merás I, Galeano-Díaz T, Wold J.P (2011) Front-face fluorescence spectroscopy: A new tool for control in the wine industry. J. Food Compos. Anal. 24: 257-264.

Azcarate SM, De Araújo Gomes A, Alcaraz M, Ugulino de Araújo MC, Camiña JM, Goicoechea HC (2015) Modeling excitation–emission fluorescence matrices with pattern recognition algorithms for classification of Argentine white wines according grape variety. Food Chem. 184: 214-219.

Azcarate SM, Cantarelli MA, Pellerano RG, Marchevsky EJ, Camiña JM (2013) Classification of Argentinean Sauvignon blanc wines by UV spectroscopy and chemometric methods. J. Food Sci. 78: 432-436.

Dufour E, Letort A, Laguet A, Lebecque A, Serra JN (2006) Investigation of variety, typicality and vintage of French and German wines using front-face fluorescence spectroscopy. Anal. Chim. Acta 563: 292-299.

Markechová D, Májek P, Kleinová A, Sádecká J (2014) Determination of the adulterants in adulterant-brandy blends using fluorescence spectroscopy and multivariate methods. Anal. Methods 6: 379-386.

Martelo-Vidal MJ, Domínguez-Agis F, Vázquez M (2013) Ultraviolet/visible/near-infrared spectral analysis and chemometric tools for the discrimination of wines between subzones inside a controlled designation of origin: a case study of Rías Baixas. Aust. J. Grape Wine Res. 19: 62-67.

Mignani AG, Ciaccheri L, Gordillo B, Mencaglia AA, González-Miret ML, Heredia FJ, Culshaw B (2012) Identifying the production region of single-malt Scotch whiskies using optical spectroscopy and pattern recognition techniques. Sens. Act. B. 171-172: 458-462.

Molinaro AM, Simon R, Pfeiffer RM (2005) Prediction error estimation: a comparison of resampling methods. Bioinformatics 21: 3301-3307.

Robbat JrA, Kowalsick A, Howell J (2011) Tracking juniper berry content in oils and distillates by spectral deconvolution of gas chromatography/mass spectrometry data. J. Chromatogr. A 1218: 5531-5541.

Saad R, Bouveresse DJ-R, Locquet N, Rutledge DN (2016) Using pH variations to improve the discrimination of wines by 3D front face fluorescence spectroscopy associated to Independent Components Analysis. Talanta 153: 278-284.

Shen F, Yang DT, Ying YB, Li BB, Zheng YF, Jiang T (2012) Discrimination between shaoxing wines and other chinese rice wines by near-infrared spectroscopy and chemometrics. Food Bioprocess. Tech. 5: 786-795.

Silvestri M, Elia A, Bertelli D, Salvatore E, Durante C, Li Vigni M, Marchetti A, Cocchi M (2014) A mid level data fusion strategy for the Varietal Classification of Lambrusco PDO wines. Chemom. Intell. Lab. Syst. 137: 181-189.

Uríčková V, Sádecká J, Májek P (2015) Classification of Slovak juniper-flavoured spirit drinks. J. Food Nutr. Res. 54: 298-307.

Vichi S, Aumatell MR, Buxaderas S, López-Tamames E (2008) Assessment of some diterpenoids in commercial distilled gin. Anal. Chim. Acta 628: 222-229.

Xia Laboratory. MataboAnalyst. [Online]. Available at: http://www.metaboanalyst.ca. [Accessed: 2017, January 18].

Xia J, Broadhurst DI, Wilson M, Wishart DS (2013) Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics 9: 280-299.

Xia J, Wishart DS (2016) Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. Curr. Protoc. Bioinformatics 55:14.10.1-14.10.91.

Yin Ch, Li H, Ding Ch, Wang H (2009) Preliminary investigation of variety, brewery and vintage of wines using three-dimensional fluorescence spectroscopy. Food Sci. Technol. Res. 15: 27-38.