Wageningen Academic Publishers

The multi-elemental isotope ratios analysis of oranges by ICP-MS and their geographic origin identification

X. Xu[1*], Q. Guo[1] and E. Duhoranimana[2]

[1]Dali Comprehensive Inspection Centre of Quality and Technical Supervision, Dali 671000, Yunnan, China P.R.; [2]State Key Laboratory of Food Science and Technology, School of Food Science and Technology, Jiangnan University, Wuxi 214122, Jiangsu, China P.R.; xuxing1983@163.com

Received: 26 April 2018 / Accepted: 13 June 2018

RESEARCH ARTICLE

Abstract

Inductively coupled plasma mass spectrometry was utilised to analyse the isotope ratios of 67 orange samples from different regions in China. To achieve geographic origin identification of the oranges, principal component analysis linear discriminant analysis (LDA) and partial least squares regression (PLSR) have been used. Twelve isotope ratios were considered as the significant variables by PLSR. A percentage of original classification was 97.0%, and the cross-validated rate of classification was 92.5% by the approach of LDA combine with PLSR. The results indicated that chemometrics techniques, combined with isotope ratios, were effective for the classification of oranges with geographical features. This study provided a theoretical reference for the discrimination and classification of oranges in China. It could be used as a theoretical reference for oranges discrimination worldwide.

Keywords: inductively coupled plasma mass spectrometry, chemometrics, geographic origin, orange

1. Introduction

The authenticity of agricultural products has been received increasing attention due to the geographic origin of agricultural products which is regarded as their quality assurance (Luykx and Van Ruth, 2008). Consumers are interested in agricultural products with clear labelling of geographic origin, and they often purchase agricultural products according to this information. Additionally, the usage of geographical labels can make producers obtain market credit and premium prices. Therefore, many countries legislate to protect the products of geographical indication (Martin *et al.*, 2012; Selih *et al.*, 2014). For instance, European Union has provided 'Protected Designation of Origin (PDO)' to protect food and agricultural products since 1996 (Mandrile *et al.*, 2016). At present, 'Trademark Law' and 'Provisions of the Product Protection of Geographical Indications' are both used to protect geographical indications in China (Yang, 2011). Therefore, it is imperative to confirm the authenticity of the label information for agricultural products.

In these past years, most researches about the geographical origin were focus on wine, coffee, milk, honey and rice (Ariyama *et al.*, 2012; Burns *et al.*, 2017; Geana *et al.*, 2013; Jandric *et al.*, 2017; Sun *et al.*, 2014; Zain *et al.*, 2016). But there are few literatures about origin traceability of oranges. In China, there are thirty-seven kinds of certified geographical indication oranges. Thus, it becomes significant to prevent deceptive labelling of these certified geographical indication oranges. Recently, various analysis techniques were applied to study the geographical origin, including spectroscopy, chromatography, mass spectrometry, nuclear magnetic resonance and sensor (Ben Mansour *et al.*, 2016; Cossignani *et al.*, 2018; Gamboa-Becerra *et al.*, 2017; Li *et al.*, 2017; Oliveira *et al.*, 2016; Vetrova *et al.*, 2017). Among these techniques, trace element composition and isotope ratios were analysed by mass spectrometry were considered as the most effective method for identifying the provenance of agricultural products. Additionally, some reports suggested that soil, source of water and mineral could affect the isotopes compositions of lead (Pb) and strontium (Sr) in crops (Ariyama *et al.*, 2012). Thus, the isotopes of Pb and Sr are extensively used to

identify the geographical origin (Geana *et al.*, 2017; Manton, 2010; Song *et al.*, 2014). Nowadays, some researchers have found that the isotopes of other elements could be also used in determining the provenance of agricultural products, such as ^{60}Ni, ^{66}Zn, ^{85}Rb, ^{95}Mo and ^{138}Ba (Martin *et al.*, 2012). However, the isotopes of other elements have not been widely attempted in the study of origin traceability.

There are too many data from instrumental analysis, so researchers need to do some statistical analysis for the data. Recently, chemometrics techniques provide great advantages in identifying the provenance of agricultural products (Barbosa *et al.*, 2014; Guo *et al.*, 2016). For instance, principal component analysis (PCA), linear discriminant analysis (LDA), clustering analysis and artificial neural network have been carried out in the study of origin traceability (Giannetti *et al.*, 2017; Kruzlicova *et al.*, 2013; Selih *et al.*, 2014; Zain *et al.*, 2016). However, many researchers have used a single technique, or compared the discriminant rates of different techniques. There are few researchers who have combined these chemometrics techniques to analyse and determine the geographic origin of agricultural products.

Therefore, the aim of this study was to investigate the multi-elemental isotope ratios of oranges by inductively coupled plasma mass spectrometry (ICP-MS), and develop an effective method to identify the provenance of oranges in China. The combination of PCA, LDA and partial least squares regression (PLSR) was applied to establish the discriminant model. Moreover, the correlation between the isotope ratios of analysed orange samples and their provenance was systematically investigated.

2. Materials and methods

Apparatus and reagents

Analyses of samples were performed using a Nexlon300x inductively coupled plasma mass spectrometer of PerkinElmer (Waltham, MA, USA) equipped with a micro-concentric nebuliser, nickel cones, and a peristaltic sample delivery pump. A Mars 6 Microwave System was from CEM Corporation (Matthews, NC, USA) equipped with the capacity of 100 mL closed Teflon digestion vessel. A BHW-09C constant temperature heater was from Shanghai Botong Chemical Technology Co., Ltd (Shanghai, China). A vortex shaker was from Thermo Fisher Scientific (Waltham, Ma, USA).

Ultrapure grade HNO_3 (71%) was from Suzhou Crystal Clear Chemical Co., Ltd (Suzhou, China). Ultrapure water (18.2 MΩ·cm) was obtained from a Biogen Ultrapure Water Type1 (Weston, FL, USA). Aqueous phase needle type filters (0.22 μm polyethersulfone) were from ANPEL Scientific Instrument Co., Ltd (Shanghai, China).

Sampling and pretreatment

A total of 67 orange samples were collected from 5 regions (BCYN: Binchuan in Yunnan province, NNGX: Nanning in Guangxi province, QZZJ: Quzhou in Zhejiang province, ZGHB: Zigui in Hubei province, PJSC: Pujiang in Sichuan province) in China during November 2016 to May 2017. A map of the distribution of orange samples is shown in Figure 1. All fresh samples were analysed in time and temporarily stored in the dark at 4 °C. The storage time was less than two weeks.

Peeled orange samples were homogenised at 200 rpm for 30 s using an electric blender (Dragon Laboratory Instruments Limited, Beijing, China), and then the homogeneous mixture were used for subsequent analysis. An aliquot of 1.0 g homogenised sample was weighed in a 100 ml Teflon digestion vessel. Then, 7 ml of ultrapure grade HNO_3 was added to the vessel, and the lid was tightly screwed. Afterwards, the mixture was shaken for 1 min by a vortex shaker. Meanwhile, the same volume of HNO_3 without sample was used as the reagent blank. The blank and samples were pre-treated by a Mars 6 Microwave System equipped with a microwave acid digestion, they were heated in an oven at 200 °C for 1 hour. When the digestion program was finished, the digestion solution was placed in a constant temperature heater at 160 °C for 2 hours, in order to evaporate the HNO_3. After cooling to ambient temperature, the digestion solution was transferred into a volumetric flask and the solution volume was adjusted to 25 ml with ultrapure water. Finally, the digestion solution was filtered by aqueous phase needle type filters for ICP-MS analysis.

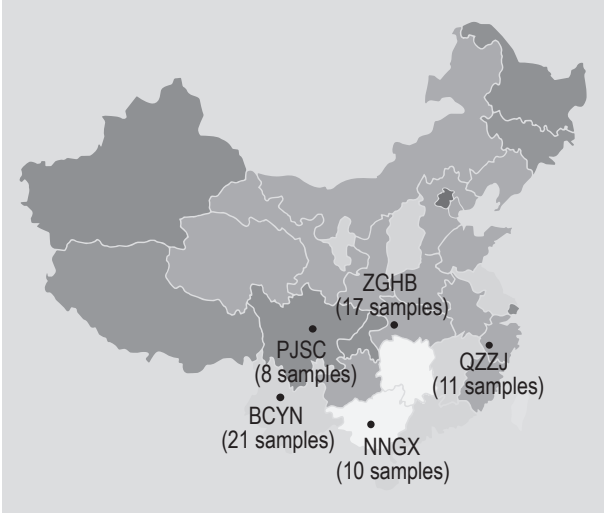

Figure 1. Map showing the distribution of orange samples which were collected in China. BCYN: Binchuan in Yunnan province; NNGX: Nanning in Guangxi province; QZZJ: Quzhou in Zhejiang province; ZGHB: Zigui in Hubei province; PJSC: Pujiang in Sichuan province.

Inductively coupled plasma mass spectrometry analysis

All orange samples were analysed using PerkinElmer Nexlon300x ICP-MS. The gaseous argon was used to form the plasma in the ICP-MS with purity over 99.9999%, and the collision gas was nitrogen with purity over 99.999%. Isotopes determined by ICP-MS were ^{60}Ni, ^{61}Ni, ^{62}Ni, ^{63}Cu, ^{65}Cu, ^{66}Zn, ^{67}Zn, ^{68}Zn, ^{84}Sr, ^{86}Sr, ^{87}Sr, ^{88}Sr, ^{95}Mo, ^{97}Mo, ^{107}Ag, ^{109}Ag, ^{135}Ba, ^{137}Ba, ^{204}Pb, ^{206}Pb, ^{207}Pb, ^{208}Pb. Signal response abundance of the isotopes was analysed in duplicate for each sample. Then, the isotope ratios were calculated from the response abundance of the isotopes.

Statistical analysis

Statistical analysis of the data was carried out using SPSS Statistics 17.0 (SPSS Inc., Chicago, IL, USA) and Unscrambler 9.7 (CAMO ASA, Oslo, Norway). PCA is a technique for considering all variability, then reducing and compressing the dimension. The main factors (PCs) identified will account for a major variable in the data matrix. PCA is often used for classification of samples (Chung *et al.*, 2016).

LDA is a superior technique for group classifications. LDA could generate the discriminant model according to the group assignment, then reclassify the data and provide the accuracy of classification for samples (Zhao *et al.*, 2012). The stepwise selection process was used to improve the accuracy of the discriminant models, and eliminate the variables which could not obviously contribute to the classification. Finally, the robustness of the discriminant model was assessed by a cross-validation test.

PLSR is a technique that generalises and combines features of PCA and multiple regressions. PLSR is especially useful when a cluster of dependent variables need to be predicted from a large number of independent variables. All the variables were standardised (1/Sdev) before PLSR analyses. The detailed procedure of PLSR analysis was referred to Song *et al.* (2010). This technique is often used to analyse the canonical correlation for multiple factors.

3. Results and discussion

Principal component analysis

Principal component analysis is generally used for the first step in statistical analysis. This method can be applied to reduce the dimensionality of the data, and find a correlation between the geographical regions and the variables. There were 22 isotope ratios of 67 samples as original variables. As a result of SPSS Statistics, the 7 principal components can explain 88.6% of the variance in the dataset. The first principal component (PC1) carries 31.0% of the data variance, the PC2 carries 22.7%, the PC3 11.1%, the PC4 carries 7.5%, the PC5 carries 6.6%, the PC6 carries 5.2% and the PC7 carries 4.5% of the original data variance. These results indicated that the PC1 and the PC2 could explain more than 50% of the variance in the dataset, and the contribution rate of other five principal components was poor. Therefore, the PC1 and the PC2 were applied to draw a two-dimensional principal component diagram. After plotting the samples as presented in Figure 2, all orange samples from different geographic regions can be visually clustered. However, the figure does not assuredly represent the total variation for classifying the orange samples from different geographic regions.

Linear discriminant analysis

In order to build a good discriminant model, and decrease the complexity of the classification, LDA was carried out using isotope ratios in this study. Then, a cross-validated procedure was used to evaluate the accuracy of the discriminant model. The original classification and cross-validated rates from LDA are listed in Table 1. Results revealed that, the LDA with isotope ratios can obtain a

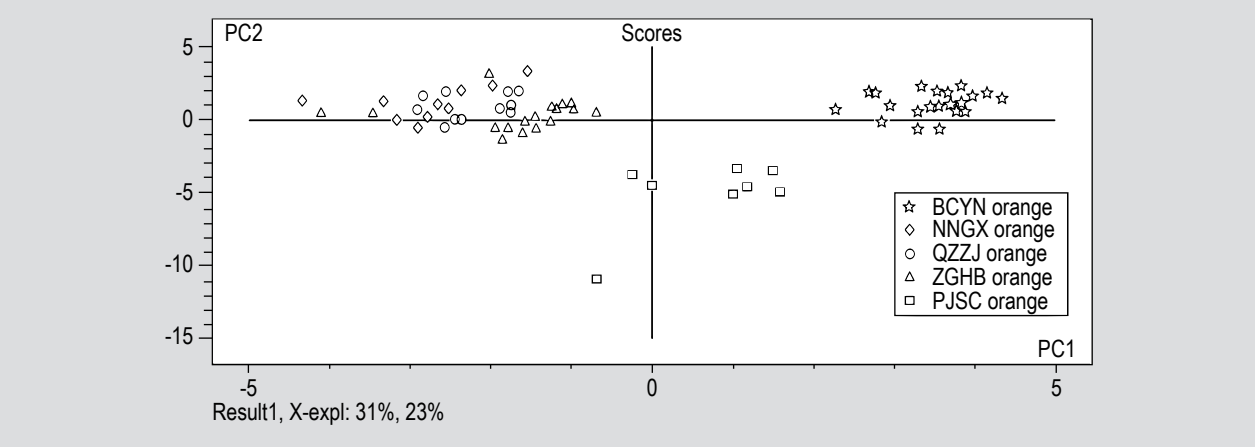

Figure 2. Score plot of the principal component analysis in the orange samples (n=67).

reasonable original classification rate (total 97.0%), and cross-validated rate (total 88.1%). From these results, BCYN orange and PJSC orange could be perfectly discriminated. But NNGX orange, QZZJ orange and ZGHB orange cannot be discriminated very well. A classified illustration of the orange samples for five categories (BCYN, NNGX, QZZJ, ZGHB or PJSC) in the plane is depicted in Figure 3, and the classification could be visualised. The discriminant model can discriminate orange samples from different geographic regions.

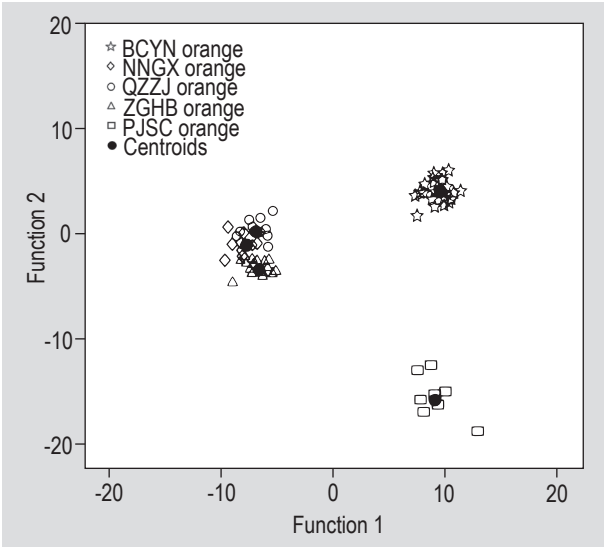

Figure 3. Linear discriminant analysis of orange samples from different regions based on isotope ratios (22 variables).

Partial least squares regression

Relationship between isotope ratios and geographic regions was studied by PLSR analysis. The X-matrix was designed as isotope ratios; the Y-matrix was designated as orange samples. The ellipses (inner and outer ellipse) indicate 50 and 100% of explained variance respectively. As indicated in Figure 4, there were 12 variables ($^{61}Ni/^{60}Ni$, $^{62}Ni/^{60}Ni$, $^{62}Ni/^{61}Ni$, $^{65}Cu/^{63}Cu$, $^{86}Sr/^{84}Sr$, $^{88}Sr/^{84}Sr$, $^{88}Sr/^{86}Sr$, $^{97}Mo/^{95}Mo$, $^{137}Ba/^{135}Ba$, $^{206}Pb/^{204}Pb$, $^{207}Pb/^{204}Pb$ and $^{208}Pb/^{204}Pb$) placed between the inner and outer ellipses, and it could be considered as the significant variables. There were 10 variables in the inner ellipse, which revealed a poor performance for the classification, and could be considered as the non-significant variables.

Combined linear discriminant analysis and partial least squares regression analysis

With the purpose of building the optimal discriminant model for orange samples from different geographic regions, LDA and PLSR analysis were combined in this study. According to the PLSR analysis results, the non-significant variables were not considered, and the significant variables were used for LDA. The data are listed in Table 2. The LDA performed with the significant variables can provide a well original classification rate (total 97.0%), and cross-validated rate (total 92.5%). This original classification rate was similar to the data of traditional LDA, but the obtained cross-validated rate (total 92.5%) was better with the data (total 88.1%) of traditional LDA. A classified illustration of the orange samples for five categories (BCYN, NNGX, QZZJ, ZGHB or PJSC) in the plane is depicted in Figure 5 (12 variables). Thus, it can be considered that the approach of combining LDA and PLSR is more promising than the approach of LDA only.

Table 1. Classification results of orange samples by linear discriminant analysis.

Original/cross-validated	Group[1]	Predicted group membership (%)					
		BCYN orange	NNGX orange	QZZJ orange	ZGHB orange	PJSC orange	Total
Original	BCYN orange	100.0	0.0	0.0	0.0	0.0	97.0
	NNGX orange	0.0	100.0	0.0	0.0	0.0	
	QZZJ orange	0.0	18.2	81.8	0.0	0.0	
	ZGHB orange	0.0	0.0	0.0	100.0	0.0	
	PJSC orange	0.0	0.0	0.0	0.0	100.0	
Cross-validated	BCYN orange	100.0	0.0	0.0	0.0	0.0	88.1
	NNGX orange	0.0	60.0	30.0	10.0	0.0	
	QZZJ orange	0.0	9.1	72.7	18.2	0.0	
	ZGHB orange	0.0	5.9	0.0	94.1	0.0	
	PJSC orange	0.0	0.0	0.0	0.0	100.0	

[1] BCYN: Binchuan in Yunnan province; NNGX: Nanning in Guangxi province; QZZJ: Quzhou in Zhejiang province; ZGHB: Zigui in Hubei province; PJSC: Pujiang in Sichuan province.

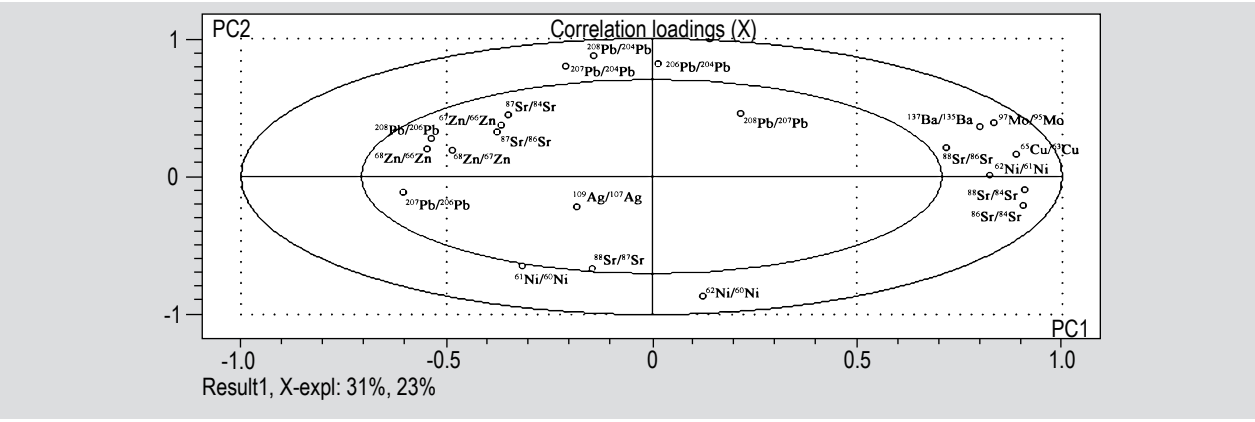

Figure 4. Partial least squares regression correlation loadings plot for orange samples.

Table 2. Classification results of orange samples by linear discriminant analysis combined with partial least squares regression.

Original/cross-validated	Group[1]	Predicted group membership (%)					
		BCYN orange	NNGX orange	QZZJ orange	ZGHB orange	PJSC orange	Total
Original	BCYN orange	100.0	0.0	0.0	0.0	0.0	97.0
	NNGX orange	0.0	100.0	0.0	0.0	0.0	
	QZZJ orange	0.0	9.1	81.8	9.1	0.0	
	ZGHB orange	0.0	0.0	0.0	100.0	0.0	
	PJSC orange	0.0	0.0	0.0	0.0	100.0	
Cross-validated	BCYN orange	100.0	0.0	0.0	0.0	0.0	92.5
	NNGX orange	0.0	80.0	20.0	0.0	0.0	
	QZZJ orange	0.0	9.1	81.8	9.1	0.0	
	ZGHB orange	0.0	0.0	5.9	94.1	0.0	
	PJSC orange	0.0	0.0	0.0	0.0	100.0	

[1] BCYN: Binchuan in Yunnan province; NNGX: Nanning in Guangxi province; QZZJ: Quzhou in Zhejiang province; ZGHB: Zigui in Hubei province; PJSC: Pujiang in Sichuan province.

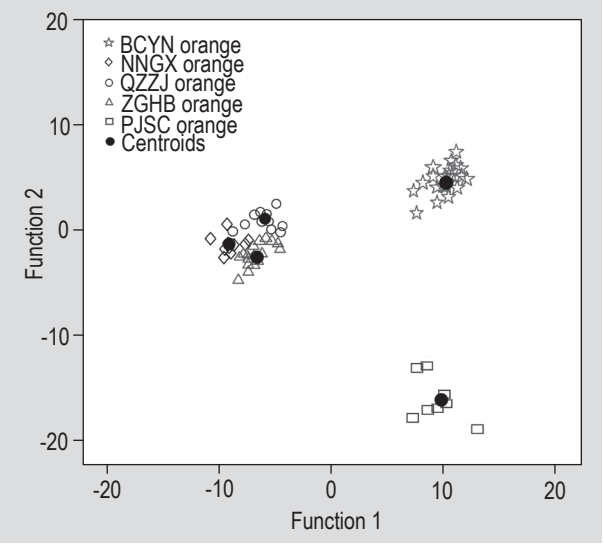

Figure 5. Linear discriminant analysis of orange samples from different regions based on isotope ratios (12 variables).

4. Conclusions

Multivariate analysis combined with instrumental technique provided an excellent approach for handling complicated variables at once. In this study, isotope ratios were used to classify 67 orange samples from different geographic regions by several chemometrics techniques (PCA, LDA and PLSR). Twelve isotope ratios were considered as the significant variables by PLSR. The percentage of original classification was 97.0%, and the cross-validated rate of classification was 92.5% by the approach of combining LDA and PLSR. The results demonstrated the usefulness of isotope ratios as indicators for authenticating the geographical origin of oranges. Therefore, the present study provided a promising and useful theoretical reference for the discrimination and authentication of oranges.

Acknowledgements

This work was supported by the Science and Technology Project of Yunnan Quality Technical Supervision Bureau (grant number 2017ynzjkj06).

References

Ariyama, K., Shinozaki, M. and Kawasaki, A., 2012. Determination of the geographic origin of rice by chemometrics with strontium and lead isotope ratios and multielement concentrations. Journal of Agricultural and Food Chemistry 60: 1628-1634.

Barbosa, R.M., Batista, B.L., Varrique, R.M., Coelho, V.A., Campiglia, A.D. and Barbosa, F., 2014. The use of advanced chemometric techniques and trace element levels for controlling the authenticity of organic coffee. Food Research International 61: 246-251.

Ben Mansour, A., Gargouri. B., Melliou. E., Magiatis, P. and Bouaziz, M., 2016. Oil quality parameters and quantitative measurement of major secoiridoid derivatives in Neb Jmel olive oil from various Tunisian origins using qNMR. Journal of the Science of Food and Agriculture 96: 4432-4439.

Burns, D.T., Tweed, L. and Walker, M.J., 2017. Ground roast coffee: review of analytical strategies to estimate geographic origin, species authenticity and adulteration by dilution. Food Analytical Methods 10: 2302-2310.

Chung, I.M., Kim, J.K., Prabakaran, M., Yang, J.H. and Kim, S.H., 2016. Authenticity of rice (*Oryza sativa* L.) geographical origin based on analysis of C, N, O and S stable isotope ratios: a preliminary case report in Korea, China and Philippine. Journal of the Science of Food and Agriculture 96: 2433-2439.

Cossignani, L., Blasi, F., Simonetti, M.S. and Montesano, D., 2018. Fatty acids and phytosterols to discriminate geographic origin of lycium barbarum berry. Food Analytical Methods 11: 1180-1188.

Gamboa-Becerra, R., Montero-Vargas, J.M., Martinez-Jarquin, S., Galvez-Ponce, E., Moreno-Pedraza, A. and Winkler, R., 2017. Rapid classification of coffee products by data mining models from direct electrospray and plasma-based mass spectrometry analyses. Food Analytical Methods 10: 1359-1368.

Geana, E.I., Sandru, C., Stanciu, V. and Ionete, R.E., 2017. Elemental profile and Sr-87/Sr-86 isotope ratio as fingerprints for geographical traceability of wines: an approach on Romanian Wines. Food Analytical Methods 10: 63-73.

Geana, I., Iordache, A., Ionete, R., Marinescu, A., Ranca, A. and Culea, M., 2013. Geographical origin identification of Romanian wines by ICP-MS elemental analysis. Food Chemistry 138: 1125-1134.

Giannetti, V., Mariani, M.B., Mannino, P. and Marini, F., 2017. Volatile fraction analysis by HS-SPME/GC-MS and chemometric modeling for traceability of apples cultivated in the Northeast Italy. Food Control 78: 215-221.

Guo, X., Huang, F., Zhang, H., Zhang, C.J., Hu, H.H. and Chen, W.B., 2016. Classification of traditional Chinese pork bacon based on physicochemical properties and chemometric techniques. Meat Science 117: 182-186.

Jandric, Z., Frew, R.D., Fernandez-Cedi, L.N. and Cannavan, A., 2017. An investigative study on discrimination of honey of various floral and geographical origins using UPLC-QToF MS and multivariate data analysis. Food Control 72: 189-197.

Kruzlicova, D., Fiket, Z. and Kniewald, G., 2013. Classification of Croatian wine varieties using multivariate analysis of data obtained by high resolution ICP-MS analysis. Food Research International 54: 621-626.

Li, J.J., Fu, B.B., Huo, D.Q., Hou, C.J., Yang, M., Shen, C.H., Luo, H.B. and Yang, P., 2017. Discrimination of Chinese teas according to major amino acid composition by a colorimetric IDA sensor. Sensors and Actuators B: Chemical 240: 770-778.

Luykx, D. and Van Ruth, S.M., 2008. An overview of analytical methods for determining the geographical origin of food products. Food Chemistry 107: 897-911.

Mandrile, L., Zeppa, G., Giovannozzi, A.M. and Rossi, A.M., 2016. Controlling protected designation of origin of wine by Raman spectroscopy. Food Chemistry 211: 260-267.

Manton, W.I., 2010. Determination of the provenance of cocoa by soil protolith ages and assessment of anthropogenic lead contamination by Pb/Nd and lead isotope ratios. Journal of Agricultural and Food Chemistry 58: 713-721.

Martin, A.E., Watling, R.J. and Lee, G.S., 2012. The multi-element determination and regional discrimination of Australian wines. Food Chemistry 133: 1081-1089.

Oliveira, L.F., Braga, S., Augusto, F., Hashimoto, J.C., Efraim, P. and Poppi, R.J., 2016. Differentiation of cocoa nibs from distinct origins using comprehensive two-dimensional gas chromatography and multivariate analysis. Food Research International 90: 133-138.

Selih, V.S., Sala, M. and Drgan, V., 2014. Multi-element analysis of wines by ICP-MS and ICP-OES and their classification according to geographical origin in Slovenia. Food Chemistry 153: 414-423.

Song, B.Y., Ryu, J.S., Shin, H.S. and Lee, K.S., 2014. Determination of the source of bioavailable Sr using Sr-87/Sr-86 tracers: a case study of hot pepper and rice. Journal of Agricultural and Food Chemistry 62: 9232-9238.

Song, S.Q., Zhang, X.M., Hayat, K., Huang, M.G., Liu, P., Karangwa, E., Gu, F.L., Jia, C.S., Xia, S.Q., Xiao, Z.B. and Niu, Y.W., 2010. Contribution of beef base to aroma characteristics of beeflike process flavour assessed by descriptive sensory analysis and gas chromatography olfactometry and partial least squares regression. Journal of Chromatography A 1217: 7788-7799.

Sun, H.W., Xia, L.Y., Liang, S.Z. and Shen, S.G., 2014. The correlation of inorganic anion contents in rice and its soils based on four geographical origin. Food Analytical Methods 7: 1791-1797.

Vetrova, O.V., Kalashnikova, D.A., Melkov, V.N. and Simonova, G.V., 2017. Detection of honey adulterations with sugar syrups by stable isotope mass spectrometry. Journal of Analytical Chemistry 72: 756-760.

Yang, R., 2011. The economic value evaluation of geographical indication products. Chinese Agricultural Science Bulletin 27: 432-435.

Zain, S.M., Behkami, S., Bakirdere, S. and Koki, I.B., 2016. Milk authentication and discrimination via metal content clustering – a case of comparing milk from Malaysia and selected countries of the world. Food Control 66: 306-314.

Zhao, H.Y., Guo, B.L., Wei, Y.M. and Zhang, B., 2012. Effects of wheat origin, genotype, and their interaction on multielement fingerprints for geographical traceability. Journal of Agricultural and Food Chemistry 60: 10957-10962.