A Nondestructive Approach for Discrimination of the Origin of Sesame Seeds Using ED-XRF and NIR Spectrometry with Chemometrics

Young Hee Choi¹, Chae Kyu Hong¹, Geon Yong Park¹, Chang Kyu Kim¹, Jung Hun Kim¹, Kwon Jung¹, and Joong-Ho Kwon*

School of Food Science and Biotechnology, Kyungpook National University, Daegu 41566, Korea

¹Special Inspection Team, Seoul Metropolitan Government Research Institute of Public Health and Environment, Seoul 13818, Korea

Received August 28, 2015
Revised October 26, 2015
Accepted December 9, 2015
Published online April 30, 2016

*Corresponding Author
Tel: +82-53-950-5775
Fax: +82-53-950-6772
Email: jhkwon@knu.ac.kr

pISSN 1226-7708
eISSN 2092-6456

© KoSFoST and Springer 2016

Abstract An energy dispersive X-ray fluorescence (ED-XRF) spectrometer and a near infrared (NIR) spectrometer combined with chemometrics were applied for origin discrimination of 48 Korean, 44 Chinese, and 21 Indian sesame seed samples used for development of a discriminant calibration model. Multi-elemental ED-XRF analysis based on Mg, Al, Si, P, S, Cl, K, Ca, Mn, Fe, and Cu was used for comparisons among origins. All elements, except for Fe, showed differences and 96.5% of seed samples were assigned to the correct origin using discriminant analysis based on chemical analytical results. NIR measurements were performed for spectral scanning. Classification of seeds using NIR discriminant analysis achieved 89.4% of seed samples assigned to the correct origin. Both ED-XRF and NIR are useful as nondestructive tools for discrimination of sesame seed origins.

Keywords: energy dispersive X-ray fluorescence spectrometer, near infrared spectrometer, chemometric, sesame seed, geographical origin

Introduction

Sesame (*Sesamum indicum* L.) is an important edible oil crop in Asia and Africa. Sesame seeds have been an important seasoning in Korea for a long time and daily intake per person can reach 2.49 g (1), usually consumed as sesame seed oil. According to the Ministry of Agriculture, Food and Rural Affairs, Republic of Korea, the cultivation area and production amount of sesame seeds have both decreased each year in Korea from 52,000 ha and 32,000 tons in 1995, to 25,000 ha and 10,000 tons in 2012 (2). However, 30,000 tons of seeds were imported from China and 40,000 tons were imported from India in 2012 (3). Korea primarily imports sesame seeds from China and India as the price of sesame seeds cultivated in Korea is 3x to 4x higher than for seeds cultivated in China and India. Because of the price difference, imported sesame seeds can be mislabeled as domestic, which is illegal in Korea. To prevent illegal marketing of sesame seeds, analytical methods for distinguishing domestic from imported sesame seeds should be established. Contents of oil, protein, sesamin, and sesamol are different, depending on the origin of sesame seeds (4). However, analysis is time-consuming, labor-intensive, costly, and requires large amounts of organic solvents for discrimination of seed origin.

Numerous approaches for provenance analysis in foods have already been reported. Methods are often based on multi-elemental signatures (5–10), isotope ratios of elements, or some combination of multi-elemental signatures and isotope ratios (11–13), or on organic compound characteristics coupled with chemometrics (14–16). Commonly, trace elements in foods are determined using inductively coupled plasma-atomic emission spectrometry/mass spectrometry (ICP-AES/MS), which allows simultaneous analysis of different multi-elements. However, analytical sample preparation for ICP-AES/MS includes acid digestion, which requires a chemical laboratory with a liquid-waste treatment facility. Moreover, analysis requires a skilled analyst experienced with sample processing, required chemicals, and equipment (17).

In this study, an energy dispersive-X ray fluorescence (ED-XRF) spectrometer was used for analysis of multi-elements and discrimination of the geographic origin of sesame seeds. ED-XRF spectrometers can be used for direct analysis of solid samples with a multi-elemental capability, a wide dynamic range, high throughput, and a low cost per determination without previous chemical treatment (18).

A near infrared (NIR) spectroscope as well as ED-XRF was used for non-destructive discrimination of sesame seed origins. Application of NIR spectroscopy allows analysis without previous chemical or

physical treatments and has attracted considerable attention for discrimination between sets of similar biological materials. In the framework of discrimination of geographical origins, this method has already been successfully used for discrimination of the geographical origins of pistachio (14), Chinese medicine (16,19), virgin olive oil (15,20) and *Torreyia grandis* seeds (21).

The aim of this study was investigation of using ED-XRF and NIR spectrometry coupled with chemometric classification methods to perform rapid, relatively cheap, and nondestructive analysis for discrimination of the geographical origins of sesame seeds.

Materials and Methods

Analytical sample preparation Korean ($n=52$), Chinese ($n=48$), and Indian ($n=23$) sesame seeds were purchased from local grocery stores and from sesame seed oil manufacturers in Seoul, Korea in 2014. A total of 123 seed samples were randomly divided into a calibration group and a validation group. A total of 113 seed samples were used for creation of discriminant calibration models and 10 seed samples consisting of 4 Korean, 4 Chinese, and 2 Indian were used for model validation. All sesame seeds were unprocessed and used in original shapes, which were long and round of 2 mmx5 mm size. Seed samples were densely packed into 30 mm diameter plastic XRF sample cells (Chemplex, Palm City, FL, USA) with a 4 μ m thick prolene film (Chemplex) and 12 mm diameter NIR sample cells (Thermo Fisher Scientific, Waltham, MA, USA). Cell lids were closed and sample cells were placed on spectrometer sample holders.

Multi-elemental XRF determination The NEX-CG spectrometer used in this study was from Applied Rigaku Technologies, Inc. (Austin, TX, USA) with an energy-dispersive XRF equipped with a 50 W power palladium X-ray tube, 4 secondary targets of RX9, Cu, Mo, and Al, an air-cooled end window, and a PC based multi-channel analyzer (MCA). The X-ray tube, the secondary target, and all analytical samples had triaxial geometry for analysis under a vacuum of <50 Pa.

In general, quantitative analysis is carried out using the calibration-curve method. However, it is difficult to obtain sufficient certified standards with matrices similar to study materials. In this study, standard-less quantitative procedures based on use of mathematical algorithms were used. The fundamental parameter (FP) method is a mathematical method that relates a spectrometer response to pure elements as a means for prediction of the fluorescence intensity for a given matrix composition. This methodology, therefore, can be used to obtain elemental compositions, even if standards are not available. In a general way, the relative precision and accuracy of results using this quantitative method are usually better than 5-10% (21,22). Marques *et al.* (18) used the FP method for quantification of multi-elements in vegetations. FP analysis is available using the discrimination method for geographical origins because multi-elemental contents were compared among groups. The 11 abundant elements Mg, Al, Si,

P, S, Cl, K, Ca, Mn, Fe, and Cu that were identified in amounts greater than the limit of quantitation were selected for use after 56 elements in sesame seeds were analyzed using ED-XRF.

Statistical multi-elemental analysis Discriminant analysis (DA), a form of multivariate statistical analysis has been frequently applied for determination of geographical origins of foodstuffs based on multi-elemental data (10,12,17,23). Multi-elemental concentrations of sesame seeds were reported as mean values of triplicate measurements. A one-way analysis of variance (ANOVA) was performed for determination of differences between sesame seeds from Korea, China, and India at $p<0.05$. Discriminant analysis was performed following the stepwise method and 3 linear Fisher's functions generated in 3 geographical origins respectively and were used for discrimination of unknown samples (24,25). All statistical analyses were conducted using IBM SPSS Statistics 20 (SPSS Inc., Chicago, IL, USA).

Near Infrared Spectrometry FT-NIR spectra were scanned using Result Integration Software on an Antaris II NIR Spectrometer (Thermo Fisher Scientific) with an integrating sphere. NIR measurements were performed within the wave number region of 4,000-10,000 $1/\text{cm}$. Each spectrum represented an average of 32 absorbance scans and raw data were measured in 3.857 $1/\text{cm}$ intervals, which resulted in 1,556 variables. Calibration using 113 analytical samples and validation using 10 analytical samples was performed simultaneously for discriminant analysis of sample spectra using TQ Analysis software (Thermo Fisher Scientific).

Results and Discussion

Comparison of multi-elemental compositions among Korean, Chinese, and Indian sesame seeds Mg, Al, Si, P, S, Cl, K, Ca, Mn, Fe, and Cu multi-elemental contents of Korean, Chinese, and Indian sesame seeds are shown in Table 1. All elements, except for Fe, showed significant ($p<0.05$) differences among the 3 origins. Korean seeds contained 494, 361, and 19.2 mg/kg of Al, P, and Mn, respectively and were significantly ($p<0.05$) higher than for seeds from the other locations. Seeds from India contained the highest levels of 2,886 mg/kg Mg, 691 mg/kg Cl, 21,004 mg/kg Ca, and 19.0 mg/kg Cu. Mg, P, and K contents of sesame seeds from China were significantly ($p<0.05$) lower than for seeds from the other countries. The Fe content was not significantly ($p>0.05$) different among the 3 origins.

Discrimination of sesame seed origin based on multi-elemental discriminant analysis Geographic origins of sesame seeds were determined using multi-elemental data in combination with discriminant analysis (DA). DA was performed on chemical analytical results for the 10 variables Mg, Al, Si, P, S, Cl, K, Ca, Mn, and Cu. The

Fe content was excluded from analysis due to no significant ($p>0.05$) difference among the 3 origin countries based on ANOVA. Only the 7 variables Mg, Al, P, Cl, Ca, Mn, and Cu were applied to DA as a stepwise method was used. Two canonical discriminant functions were generated because 3 groups were used in the discriminant analysis method.

Coefficients of a standardized canonical discriminant function reflect the relative contribution of a discriminant variable in the context of other variables in the function. The higher the absolute value of the coefficient, the greater the degree of contribution of the variable to the discriminant function (24,25). All variables contributed similarly to Function 1 because the standardized coefficients of Mg, Al, P, Cl, Ca, Mn, and Cu were 0.378, -0.320, -0.415, 0.433, 0.340, -0.446, and 0.513, respectively. On the other hand, the standardized coefficients of Mg, Al, P, Cl, Ca, Mn, and Cu in Function 2 were 0.351, 0.059, 0.685, 0.400, -0.071, 0.064, and 0.291. P (0.685), Cl (0.400), and Mg (0.351) contributed more than other elements to the discrimination power of Function 2. Function 1 explained 73.8% of the variability while Function 2 accounted for 26.2% of the variability. When a total of 113 sesame seeds were classified using discriminant scores of Functions 1 and 2, 96.5% of the 113 samples were assigned to the actual origin (Table 2). All Korean sesame seeds were classified correctly and only one of 21 Indian sesame seeds was classified incorrectly as Korean. Of 44 Chinese sesame seeds, 2 were misclassified as Indian and a single seed was misclassified as Korean. The discriminant score plot shown in Fig. 1 clearly distinguishes the 3 groups, although 4 sesame seeds were misclassified (Fig. 1).

Fig. 1. Plot of discriminant analytical values (functions 1 and 2) for sesame seeds cultivated in Korea, China, and India (◆: Korea, ●: China, ▲: India, ☆: misclassified seed samples).

For validation of the discrimination model, concentrations of 7 elements obtained from analysis of 4 Korean, 4 Chinese, and 2 Indian seed samples were used as variables for Fisher's linear discriminant functions. The geographical origin of an unknown seed sample was determined based on the largest value calculated in the Fisher's linear discriminant functions using 3 origins. Origins of all 10 seed samples were correctly identified (Table 3).

Many previous studies have effectively used combined chemical

Table 1. Ranges, mean values, and standard deviation (SD) of elements measured in sesame seeds from Korea, China, and India

Elements (mg/kg)	Korean (n=48)		Chinese (n=44)		Indian (n=21)	
	Range	Mean (±SD)	Range	Mean (±SD)	Range	Mean (±SD)
Mg ¹⁾	1178-3707	2383 ^b ±470	1171-3230	1976 ^{a2)} ±475	2279-4577	2886 ^c ±640
Al	214-1332	494 ^b ±217	184-624	372 ^a ±110	249-924	356 ^a ±144
Si	285-2185	760 ^b ±323	232-980	604 ^a ±200	454-1598	677 ^{ab} ±248
P	240-533	361 ^c ±68.0	165-366	217 ^a ±41.3	235-356	310 ^b ±34.5
S	869-1431	1077 ^b ±138	689-1040	885 ^{ab} ±67.7	707-1005	872 ^a ±74.7
Cl	91.2-397	203 ^a ±64.8	69.8-1265	295 ^b ±198	77.7-1075	691 ^c ±262
K	2856-6454	3986 ^b ±712	2454-4594	3242 ^a ±501	3499-4465	3996 ^b ±299
Ca	6131-21677	15586 ^a ±2844	16758-24496	19934 ^b ±1513	18055-22495	21004 ^c ±1117
Mn	12.7-37.6	19.2 ^b ±4.01	12.3-21.7	16.2 ^a ±1.78	14.5-26.4	17.4 ^a ±2.28
Fe	48.6-128	66.7±12.0	49.8-177	62.6±19.3	56.1-77.7	63.7±5.17
Cu	8.1-18.4	14.4 ^a ±2.20	15.1-20.0	16.9 ^b ±1.03	17.0-20.1	19.0 ^c ±0.75

¹⁾Mg, Al, Si, P, S, Cl, K, Ca, Mn, Fe, and Cu elemental concentrations were determined at the mg/kg level when C was used as a balance for the FP method in ED-XRF.

^{2)a-c}Mean values show differences at $p<0.05$ among origins based on ANOVA.

Table 2. Classification results for the origin of sesame seeds using multi-elemental canonical discriminant functions

	Predicted origin			Total	Correctly classified (%)
	Korea	China	India		
Korean sesame seeds	48	0	0	48	100
Chinese sesame seeds	1	41	2	44	93.2
Indian sesame seeds	1	0	20	21	95.2
Total				113	96.5

Table 3. Validation results for discrimination models based on Fisher's linear discriminant functions

	Concentration of elements (mg/kg)							Actual group	Predicted group
	Mg	Al	P	Cl	Ca	Mn	Cu		
Sample-1	2653	431	310	128	16900	18.6	13.6	Korea	Korea
Sample-2	2813	455	372	184	14867	23.0	14.3	Korea	Korea
Sample-3	2056	362	294	283	17567	20.8	16.6	Korea	Korea
Sample-4	2120	417	283	268	17367	18.1	14.9	Korea	Korea
Sample-5	3013	623	225	159	19933	18.6	18.6	China	China
Sample-6	1833	325	240	260	18900	16.4	15.7	China	China
Sample-7	1986	298	170	206	18967	15.6	15.8	China	China
Sample-8	1960	290	184	318	19767	16.1	15.7	China	China
Sample-9	2744	321	297	737	2147	16.1	19.4	India	India
Sample-10	2633	305	281	666	20167	17.3	18.9	India	India

and statistical analyses for determination of food geographic origins. Jeon *et al.* (23) reported 89.3% discrimination for sesame seed oil based on combining carbon, hydrogen, and oxygen stable isotope ratios and DA. Bong *et al.* (13) demonstrated 87.15% discrimination for cabbages, while Yan *et al.* (10) reported 91.30% discrimination for pomelos using AAS and ICP-AES. Otaka *et al.* (17) reported 91.3% discrimination for soybeans using XRF in combination with discriminant and multi-elemental analysis. Herein, a higher discrimination rate than previous studies was achieved demonstrating that the discriminant model used could distinguish 3 origins using multi-elemental concentrations determined with XRF.

NIR spectral investigation and discrimination analysis

The discriminant analysis classification technique in NIR spectroscopy can be used for classification of analytical samples based on computation of the distance from the centroid of each group to sample in Mahalanobis distance units (24). Multiple standards and multiple spectral ranges for analysis are typically used to describe each group to be separated. It is difficult to find specific bands in raw NIR spectra based on geographical origin because there are hundreds of active ingredients in sesame seeds. Because spectra are affected by physical properties of analyzed products and other noise, it is necessary to perform mathematical processing to reduce systematic noise and to enhance the contribution of the chemical composition (16).

In this study, the secondary derivative was applied to each absorbance spectrum as a mathematical transformation for elimination

Fig. 2. Calibration results plot based on computation of the distance from each class center in Mahalanobis distance units (2A: Indian and Korean, 2B: Chinese and Korean, 2C: Chinese and Indian, ◆: Korean sesame seed, ●: Chinese sesame seed, ▲: Indian sesame seed).

Table 4. Classification results for the origin of sesame seeds using the discriminant analysis classification of NIR spectra

	Predicted origin			Total	Correctly classified (%)
	Korea	China	India		
Korean sesame seeds	42	4	2	48	85.7
Chinese sesame seeds	2	39	3	44	88.6
Indian sesame seeds	0	1	20	21	95.2
Total				113	89.4

of baseline drifts and enhancement of small spectral differences (16,26,27). Calibration results obtained based on computation of the distance from the centroid of each group to sample in Mahalanobis distance units are shown in Fig. 2. Origins were predicted based on the nearest distance from the centroid of each origin. When a total of 113 sesame seeds were classified using discriminant analysis, 85.7% of 48 Korean samples, 88.6% of 44 Chinese samples, and 95.2% of 21 Indian seeds were classified correctly with total discrimination rate of 89.4% (Table 4). Four Korean seed samples were incorrectly classified as Chinese and 2 Korean seed sample as Indian because 4 incorrectly identified Korean sesame seed samples were closer to the Chinese group centroid than the Korean group centroid, and 2 Korean sesame seed samples were closer to the Indian group centroid than the Korean group centroid. Of 44 Chinese sesame seeds, 2 seed samples were misclassified as Korean and 3 seed samples were misclassified as Indian. Only 1 Indian seed sample was misclassified as Korean.

Another 10 seed samples consisting of 4 Korean, 4 Chinese, and 2 Indian were analyzed for validation of the discrimination model using TQ analysis software. Only 1 Korean seed sample out of the 10 was incorrectly identified as Chinese.

NIR spectroscopy combined with chemometrics can be used for discrimination of geographical origins. For example, Lin *et al.* (16) reported that NIR was used to identify genuine crude *Radix Pseudostellariae* with a 92.5% identification rate, while Vitale *et al.* (14) discriminated origins of pistachio nuts using NIR spectroscopy combined with chemometric with classification accuracies of more than 90%. Woo *et al.* (19) reported that Korean and Chinese *Angelicae gigantis Radixes* samples were clearly identified with 100% accuracy using NIR combined with chemometrics. Although these studies had higher discrimination rates than reported herein, coupling of NIR spectroscopy with chemometric classification techniques in this study was shown to be a useful tool for identification of the origin of sesame seeds in a rapid, relatively cheap, and non-invasive manner. Further research is required to improve the discrimination accuracy based on analysis of more seed samples.

Disclosure The authors declare no conflict of interest.

References

1. Korea's Ministry of Food and Drug Safety. Available from: www.foodsafetykorea.go.kr. Accessed Jun. 15, 2015.
2. Ministry of Agriculture, Food and Rural Affairs. Statistical of imported agriculture, forestry and livestock. Available from: www.atfis.or.kr. Food information statistics system. Accessed Jun. 15, 2015.
3. Korea's Ministry of Food and Drug Safety. Imported food information. Available from: www.foodnara.go.kr/importfood/src/statistics/item_list.jsp. Accessed Jun. 15, 2015.
4. Shim KB, Hwang CK, Pae SB, Lee MH, Ha TJ, Park CH. Comparison of physiochemical characters of sesame seeds according to the different producing origin. *J. Korean Soc. Int. Agric.* 22: 371-375 (2010)
5. Kelly S, Heaton K. Tracing the geographical origin of food: The application of multi-element and multi-isotope analysis. *Trends Food Sci. Tech.* 16: 557-567 (2005)
6. Anderson KA, Smith BW. Effect of season and variety on the differentiation of geographic growing origin of pistachios by stable isotope profiling. *J. Agr. Food Chem.* 54: 1747-1752 (2006)
7. Camaro AB, Sara R, Eduardo JM, Juan ML. Use of the Argentinean garlic (*Allium Sativum L.*) germplasm mineral profile for determining geographic origin. *J. Food Compos. Anal.* 23: 586-591 (2010)
8. Rummel S, Hoelzl S, Horn P, Rossmann A, Schlicht C. The combination of stable isotope abundance ratios of H, C, N and S with $87\text{Sr}/86\text{Sr}$ for geographical origin assignment of orange juices. *Food Chem.* 118: 890-900 (2010)
9. Ariyama K, Shinozaki M, Kawasaki A. Determination of the geographic origin of rice by chemometrics with strontium and lead isotope ratios and multielement concentrations. *J. Agr. Food Chem.* 60: 1628-1634 (2012)
10. Yan J, Liu J, Xiong Y, Qin W, Tang C. Identification of the geographical origins of pomelos using multielement fingerprinting. *J. Food Sci.* 80: C228-C233 (2015)
11. Galgano F, Favat F, Carugo M, Scarpa T, Palma A. Analysis of trace elements in southern Italian wines and their classification according to provenance. *LWT-Food Sci. Technol.* 41: 1808-1815 (2008)
12. Liou HC, You CF, Chen CY, Liu YC, Chung MT. Geographic determination of coffee beans using multi-element analysis and isotope ratios of boron and strontium. *Food Chem.* 142: 439-445 (2014)
13. Bong YS, Song BY, Gautam MK, Jang CS, An HJ, Lee KS. Discrimination of the geographic origin of cabbage. *Food Control* 30: 626-630 (2013)
14. Vitale R, Bevilacqua M, Bucci R, Magri AD, Magri AL, Marini F. A rapid and non-invasive method for authenticating the origin of pistachio samples by NIR spectroscopy and chemometrics. *Chemometr. Intell. Lab.* 121: 90-99 (2013)
15. Galtier O, Dupuy N, Dreau YL, Ollivier D, Pinatel C, Kister J, Artaud J. Geographic origins and compositions of virgin olive oils determined by chemometric analysis of NIR spectra. *Anal. Chim. Acta* 595: 136-144 (2007)
16. Lin H, Zhao J, Chen Q, Zhou F, Sun L. Discrimination of *radix pseudostellariae* according to geographical origins using NIR spectroscopy and support vector data description. *Spectrochim. Acta A* 79: 1381-1385 (2011)
17. Otaka A, Hokura A, Nakai I. Determination of trace elements in soybean by X-ray fluorescence analysis and its application to identification of their production area. *Food Chem.* 147: 318-326 (2014)
18. Marques AF, Queralt I, Carvalho ML, Bordalo M. Total reflection X-ray fluorescence and energy-dispersive X-ray fluorescence analysis of runoff water and vegetation from abandoned mining of Pb-Zn ores. *Spectrochim. Acta B* 58: 2191-2198 (2003)
19. Woo YA, Kim HJ, Ze KR, Chung H. Near-Infrared(NIR) spectroscopy for the non-destructive and fast determination of geographical origin of *Angelicae*

- Gigantis Radix*. J. Pharmaceut. Biomed. 36: 955-959 (2005)
20. Mezghani LS, Vanloot P, Molinet J, Duppy N, Hammami M, Kamoun NG, Artaud J. Authentication of Tunisian virgin olive oils by chemometric analysis of fatty acid compositions and NIR spectra. Comparison with Maghrebian and French virgin olive oils. Food Chem. 173: 122-132 (2015)
 21. Xu L, Yan SM, Cai CB, Zhong W, Yu XP. Rapid analysis of geographical origins and age of *Torreya Grandis* seeds by NIR spectroscopy and pattern recognition methods. J. Spectrosc. Article ID 468509 (2013)
 22. Margui E, Queralto I, Hidalgo M. Application of X-ray fluorescence spectrometry to determination and quantitation of metals in vegetal material. Trends Analyt. Chem. 28: 362-372 (2009)
 23. Jeon HJ, Lee SC, Cho YJ, Oh JH, Kwon KS, Kim BH. A triple-isotope approach for discriminating the geographic origin of asian sesame oils. Food Chem. 167: 363-369 (2015)
 24. Yang BH. Understanding Multivariate Data Analysis. 4thed. Communication Books Inc., Seoul, Korea. pp. 107-124 (2011)
 25. Han KJ. Application of SPSS Statistical research. Baeksan Publisher, Seoul, Korea. pp. 381-432 (2012)
 26. Xu Z, Liu Y, Li X, Cai W, Show X. Discriminant analysis of Chinese patent medicines based on near-infrared spectroscopy and principal component discriminant transformation. Spectrochim. Acta A 149: 985-990 (2015)
 27. He Y, Li X, Deng X. Discrimination of varieties of tea using near infrared spectroscopy by principal component analysis and BP model. J. Food Eng. 79: 1238-1242 (2007)