한작지(Korean J. Crop Sci.), 51(S): 304~309(2006)

Nondestructive Prediction of Fatty Acid Composition in Sesame Seeds by Near Infrared Reflectance Spectroscopy

Kwan-Su Kim*†, Si-Hyung Park*, Myoung-Gun Choung**, and Sun-Lim Kim***

*Department of Medicinal Plant Resources, Mokpo National University, Muan 534-729, Korea
**Department of Pharmacognosy Material Development, Kangwon National University, Samcheok 245-711, Korea
***National Institute of Crop Science, RDA, Suwon 441-857, Korea

ABSTRACT Near infrared reflectance spectroscopy (NIRS) was used to develop a rapid and nondestructive method for the determination of fatty acid composition in sesame (*Sesamum indicum* L.) seed oil. A total of ninety-three samples of intact seeds were scanned in the reflectance mode of a scanning monochromator, and reference values for fatty acid composition were measured by gas-liquid chromatography. Calibration equations were developed using modified partial least square regression with internal cross validation (n=63). The equations obtained had low standard errors of cross-validation and moderate R^2 (coefficient of determination in calibration). Prediction of an external validation set (n=30) showed significant correlation between reference values and NIRS estimated values based on the SEP (standard error of prediction), r^2 (coefficient of determination in prediction) and the ratio of standard deviation (SD) of reference data to SEP. The models developed in this study had relatively higher values (more than 2.0) of SD/SEP(C) for oleic and linoleic acid, having good correlation between reference and NIRS estimate. The results indicated that NIRS, a nondestructive screening method could be used to rapidly determine fatty acid composition in sesame seeds in the breeding programs for high quality sesame oil.

Keywords : NIRS, fatty acid composition, sesame (*Sesamum indicum* L.)

Sesame (*Sesamum indicum* L.) is one of the most important oilseed crops and has been cultivated in Korea since ancient times. Sesame plays an important role as a source of edible vegetable oil and provides a healthy food for humans. The oil and whole seed of sesame are mainly used for the commercial products, and dehulled seeds are often used to obtain higher purity of oil and the protein rich flour, as a healthy food removing the seed coat with fiber and pigments. Sesame oil, consisting of approximately 50% seed weights, is mainly composed of fatty acids such as palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2), and linolenic (18:3) acids (Kim *et al.*, 2004). The major components in sesame oil are unsaturated fatty acids (mainly oleic and linoleic acid), which contribute significantly to human nutrition (Were *et al.*, 2006). Quality of sesame oil is related to the fatty acid composition and oxidative stability of oil. The higher oil content and different fatty acid composition, especially of oleic acid and linoleic acid, was of great importance in the sesame breeding for improving qualitative traits (Baydar *et al.*, 1999).

Fatty acid composition was usually determined by gas-liquid chromatography and flame ionization detector (GLC-FID) system with a capillary column. However, these methods require stepwise sample preparation. For example, fatty acids must be converted to fatty acid methyl esters (FAMEs) for GLC analysis. These analytic methods are time-consuming, costly, and labor-intensive, and it is disadvantageous for selecting superior lines of sesame. Thus, the rapid and non-destructive method is in high demand to evaluate quality traits of sesame seed for sesame breeding program.

Near-infrared spectroscopy (NIRS) is known to play a fundamental role in simplifying the analysis of chemical and physical properties without sample preparation. For example, NIRS was applied for the analysis of quality characteristics in food and agricultural commodities (Batten, 1998; Williams & Noriss, 2001). NIRS has been applied for the determination of diverse compounds and the classifications in numerous foods and industrial crops, such as sesame (Sato *et al.*, 2003; Kim *et al.*, 2006), soybean

†Corresponding author: (Phone) +82-61-450-2661
(E-mail) kskim@mokpo.ac.kr

(Choung *et al.*, 2005), perilla and peanut (Oh *et al.*, 2000), sunflower (Fassio & Cozzonlino, 2004), and rapeseed (Font *et al.*, 2003). Though the application of NIRS method to predict fatty acid composition of intact sesame seeds has been reported by Sato *et al.* (2003), it is needed to develop new prediction model to apply to our breeding program for domestic sesame germplasm.

The objectives of this study were to investigate NIRS application for predicting fatty acid composition and to develop massive screening technique in intact seed samples for sesame breeding program.

MATERIALS AND METHODS

Sesame seed samples

A total of 93 samples of sesame were obtained from breeding lines of Jeonnam Agricultural Research and Extension Services (Naju, Korea) including 20 Korean recommended sesame varieties and were used to develop NIRS prediction model for the determination of individual fatty acid composition. The sesame plants were grown in the greenhouse and harvested in 2004. The harvested seeds were cleaned, dried in the laboratory, and then stored in desiccators until NIRS and GLC analysis.

Chemical analysis for fatty acid composition

About 2 g of each sample was homogenized using an Ultra-Turax T8 homogenizer (IKA-Werke GmbH & Co. KG, Germany), extracted in 30 mL of *n*-hexane for 1 day by shaking at 100 rpm using a VS-8480 SRN horizontal shaker (Vison Co., Korea), and filtered with filter paper. The residues were extracted two more times, and the final volume of each extract solution was exactly adjusted to 100 mL. For fatty acid analysis, a 1 mL aliquot of *n*-hexane extract was transferred into the reaction vial and concentrated under nitrogen flow at 70℃. For the saponification of the oil, 0.5 mL of 0.5N NaOH solution was added and allowed to react at 100℃ for 10 min and then cooled. After 0.5 mL of 14% BF_3 (boron trifluoride) solution was added, it was allowed to react at 100℃ for 10 min for the esterification of fatty acids. After cooling, 1.5 mL of *n*-hexane and 1.0 mL of distilled water were mixed, partitioned in *n*-hexane, and the upper layer (containing fatty acid methyl esters in *n*-hexane extract) was transferred into a 2 mL autosampler vial before GC injection for fatty acid analysis. The GLC system 6890N GC (Agilent Technologies Co., USA) equipped with a flame ionization detector (FID) and a HP-Innowax capillary column (30 m length × 0.25 mm i.d., film 0.25 μm, J&W Scientific, Agilent Technologies Co., USA) was used. The oven temperature was raised from 160℃ (holding for 1 min) to 230℃ at a constant rate of 5℃/min, and then held for 10 min. The injector and detector port temperatures were kept at 230 and 250℃, respectively. The carrier gas was nitrogen at a flow rate of 1.0 mL/min and the split ratio at the injector port was 20:1. The running time was about 20 min for each sample. Individual fatty acids were expressed as percentage of the total fatty acids.

Spectra collection and pretreatment

The NIR spectroscopic analysis was performed using a NIRSystem Model 6500 near infrared scanning monochromator (Foss NIRSystems Inc., USA) in the reflectance mode. Intact seed samples (about 4 g) were placed in a standard ring cup and then scanned. Reflectance energy readings were references to corresponding readings from an internal ceramic disc. Each spectrum was recorded once from each sample, and was obtained as average of 32 successive scans over the sample, plus 16 scans over the standard ceramic before and after scanning the samples. All spectral data were recorded as the logarithm of reciprocal of reflectance (log 1/R) in the wavelength range between 400 and 2500 nm at 2 nm intervals to give a total of 1050 data points per sample. The scanning procedure could be finished within about 1.5 min per sample when NIRS instrument was warmed-up and were confirmed the stability of NIRS through photometric repeatability (noise test) and wavelength accuracy test.

The NIRS manipulation for scanning, mathematical processing, and statistical analysis was performed with the WinISI II software (Windows version 1.60, Foss and Infrasoft International LLC, USA). The samples (n=93) were randomly split into two sets using WinISI program. The calibration set (63 samples) was used to calibrate and

cross-validate the equation derived, and the other 30 samples as an external validation set were used to test the goodness of fit of the developed equations.

Data processing

The equations for NIRS prediction were developed using the Global program in WinISI software with the modified partial least squares (MPLS) regression using wavelengths of entire visible (400-1100 nm) and near infrared (1100-2500 nm) region at every 8 nm. Besides MPLS, the regression methods like PLS (partial least squares), principle component regression, and multiple linear regression were tested to develop calibration models for fatty acid composition in sesame seeds. No scatter correction (log 1/R) and scatter correction using standard normal variate and detrending (SNVD) were evaluated. The SNVD was designed to remove additive baseline and multiplicative signal effects resulting in a spectrum with zero mean and a variance equal to one. Application of SNVD transformation to raw spectral data reduces the differences in spectra related to physical characteristics such as particle size and path length of samples (Barnes *et al.* 1989; Shenk & Westerhaus, 1991a). Various mathematical treatments were applied for equation development. For example, 2,4,4,1, the first number indicated the order of derivative (two is second derivative of log 1/R), the second number was the gap in data points over which the derivative was calculated, and the third and fourth numbers represented the number of data points used in first smoothing and second smoothing (Shenk & Westerhaus, 1991b).

RESULTS AND DISCUSSION

Spectral analysis

The raw NIR reflectance and second derivative spectra of the intact seed samples of sesame are shown in Fig. 1. Main absorption bands are observed at 1208 nm related to C-H stretching 2nd overtone (-CH_2), 1496 nm related to C-H stretching 1st overtone, 1724 nm related to C-O (Oil) and C-H stretching 1st overtone (-CH_2), 1942 nm related to O-H bending 2nd overtone (Water), and 2308 nm related to C-H bending 2nd overtone (Oil). The information of functional group in spectrum was searched from WinISI software. The overall spectrum shows strong absorption bands related with oil and water, and is similar to those of other oil crops such as perilla, peanut, soybean, especially, in near-infrared region (Oh *et al.*, 2000; Choung *et al.*, 2005).

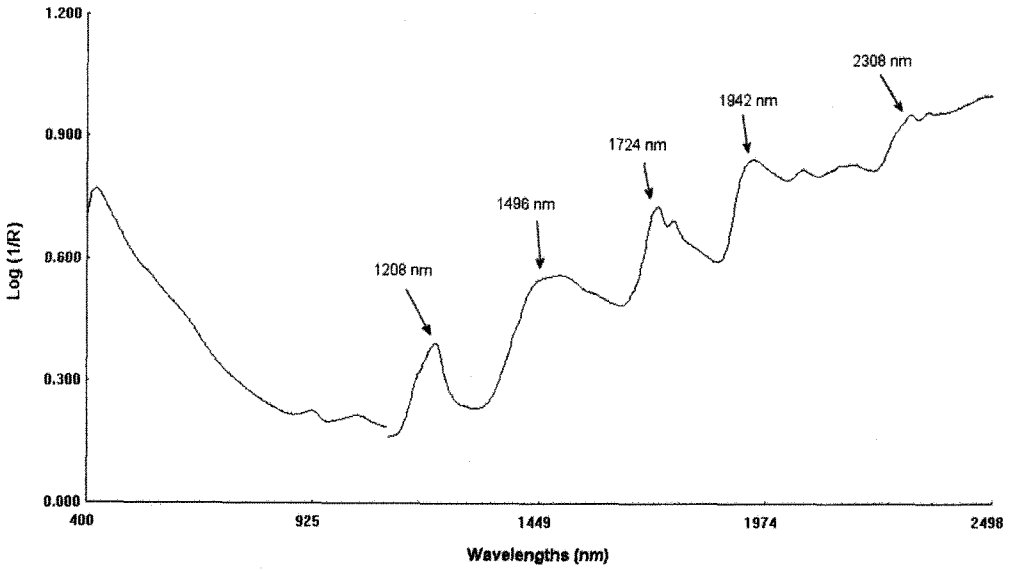

Fig. 1. Raw NIR average spectrum (log 1/R) of intact sesame seeds.

Reference analysis of fatty acid composition

The descriptive statistics including mean, standard deviation (SD), and range, for each fatty acid composition of sesame samples used in the calibration and validation sets were shown in Table 1. Each reference value of fatty acid composition in a validation sample set was similar to those in the calibration sample set. Mean values of individual fatty acid composition were 7.86% of palmitic acid (PA), 5.22% of stearic acid (ST), 42.8% of oleic acid (OL), 43.7% of linoleic acid (LN), and 0.45% of linolenic acid (LNL) in the calibration set, and 7.80% of PA, 5.27% of ST, 42.7% of OL, 43.8% of LN, and 0.44% of LNL in the validation set.

Calibration models for fatty acid composition

In developing NIRS models for each fatty acid composition, the statistics of calibrations and cross-validations were shown in Table 2. The MPLS regression model in the whole NIR spectra range (400-2500 nm) using the second derivative transformation with scatter correction (SNVD) of raw reflectance spectra yielded the equations of each fatty acid composition showing higher R^2 (0.875-0.976) and 1-VR, lower SEC and SECV values than the different mathematical treatments tested. The equations for each fatty acid composition using mathematical treatment 2,4, 4,1 were selected considering SD/SECV values with the highest values, 1.45 of PA, 1.58 of ST, 2.26 of OL, 2.96 of LN, and 1.82 of LNL as the selection criteria of models. The results shown in Table 2, were to get reliable equations for OL (42.8% in oil) and LN (43.7% in oil), major fatty acids in sesame seed oil, having high values of R^2 (0.927 and 0.976, respectively) and SD/SECV (2.26 and 2.96, respectively), and showed close relationship between reference values and NIRS estimate values. The best cali-

Table 1. Descriptive statistics for fatty acid composition of intact sesame seed samples used in both calibration and validation.

Constituents	Calibration (n=63)			Validation (n=30)		
	Mean[a]	Range	SD	Mean	Range	SD
PA	7.86	6.67-9.37	0.516	7.80	7.12-8.96	0.401
ST	5.22	4.22-5.77	0.381	5.27	4.74-5.74	0.297
OL	42.8	39.3-46.6	1.770	42.7	36.9-46.7	2.246
LN	43.7	40.2-46.8	1.806	43.8	39.2-49.7	2.301
LNL	0.45	0.33-0.73	0.073	0.44	0.35-0.66	0.065

[a]Individual fatty acid composition are expressed as percentage of the total fatty acids in the seed oil, respectively; SD, standard deviation of mean; PA, palmitic acid; ST, stearic acid; OL, oleic acid; LN, linoleic acid; LNL, linolenic acid.

Table 2. Equation statistics using regression model (MPLS) and scatter correction for NIRS prediction of fatty acid composition in the calibration set (n=63) of intact sesame seed samples.

Constituents	T	N	Calibration		Cross-validation		RSC
			SEC	R^2	1-VR	SECV	
PA	8	61	0.192	0.865	0.539	0.361	1.45
ST	6	60	0.142	0.838	0.605	0.223	1.58
OL	7	59	0.438	0.927	0.802	0.716	2.26
LN	8	59	0.265	0.976	0.885	0.579	2.96
LNL	5	59	0.027	0.815	0.701	0.034	1.82

T, Number of terms, number of PLS loading factors in the regression model MPLS; N, number of samples used to develop the model; SEC, standard error of calibration; R^2, coefficient of determination of calibration; 1-VR, one minus the ratio of unexplained variance divided by variance; SECV, standard error of cross-validation; RSC, SD/SECV : the ratio of SD (standard deviation of reference data) to SECV in the calibration set; PA, palmitic acid; ST, stearic acid; OL, oleic acid; LN, linoleic acid; LNL, linolenic acid.

Table 3. Monitoring statistics for fatty acid composition in the external validation set (n=30) of intact sesame seeds.

Constituents	N	Mean	SD	Bias	r^2	SEP(C)	Slope	RSP
PA	29	7.81	0.406	-0.151	0.615	0.254	0.903	1.60
ST	28	5.30	0.285	0.129	0.634	0.192	0.730	1.48
OL	25	43.3	1.663	0.111	0.800	0.762	0.901	2.18
LN	27	43.5	1.945	0.041	0.914	0.579	1.057	3.36
LNL	24	47.6	3.700	0.321	0.529	2.551	0.921	1.45

N, number of samples used to monitor the model; SD, standard deviation of mean; Bias, average difference between reference and NIRS values; r^2, coefficient of determination of cross-validation; SEP(C), the corrected standard error of prediction; Slope, the steepness of a straight line curve; RSP, SD/SEP(C): the ratio of SD of reference data to SEP(C) in the external validation set; PA, palmitic acid; ST, stearic acid; OL, oleic acid; LN, linoleic acid; LNL, linolenic acid.

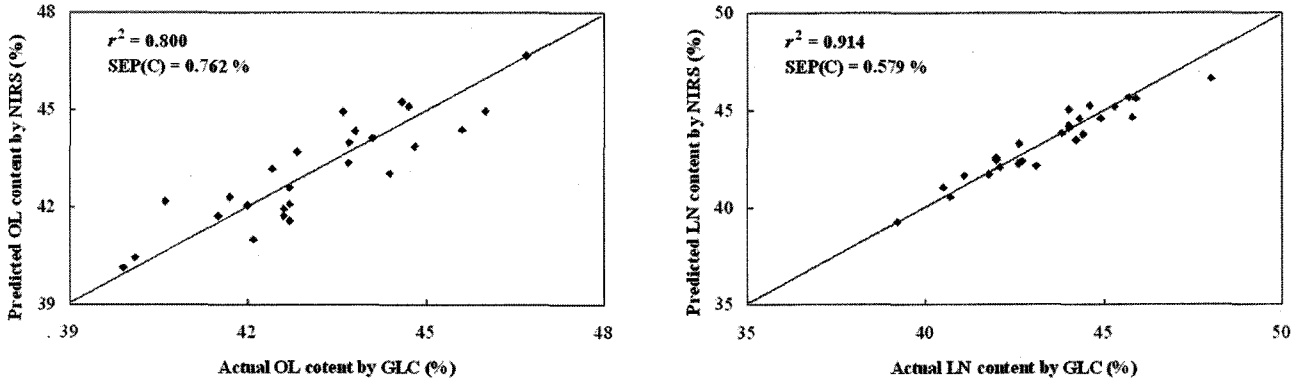

Fig. 2. Scatter plots of NIRS vs. reference values for oleic (OL) and linoleic acid (LN) in the external validation set (n=30) of intact sesame seed samples.

bration models for OL and LN were developed with the mathematical approach over the visible and near infrared segment (400-2500 nm), and the equations could be used for screening the fatty acid composition in intact sesame seeds.

External validation for fatty acid composition

The predicted statistics for fatty acid composition were shown in Table 3. The r^2 and SD/SEP(C) values for PA, ST, and LNL were lower, indicating not good correlation between reference values and NIRS estimated values as similarly appeared in those of calibration. The predictions for OL and LN were confirmed by higher r^2 (0.800 and 0.914) and SD/SEP(C) values (2.18 and 3.36). The results were similar to calibration parameters studied in sunflower (Velasco *et al.*, 1999). Figure 2 represented laboratory reference values against NIRS predicted values in the validation set for OL and LN in fatty acid composition, showing also the relationship between NIRS and reference.

These results demonstrated the accurate prediction capacities of the calibration models for fatty acid composition such as OL and LN using a non-destructive NIRS method in intact sesame seeds, though larger population of sesame seed samples was required to obtain the accurate prediction of PA, ST, and LNL. The calibration models developed in this study need to be updated for expanding and increasing the robustness of models by applying to new samples with different cultivation location, growth characteristics, harvest time, cross population, and qualitative enhancing condition as well as increasing sample size.

ACKNOWLEDGEMENTS

This study was supported by grants (Code No. 2005030 1034380) from BioGreen 21 Program of the Rural Development Administration, Republic of Korea.

REFERENCES

Barnes, R. J., M. S. Dhanoa, and S. J. Lister. 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Appl. Spectrosc. 43 : 772-777.

Batten, G. D. 1998. Plant analysis using near infrared reflectance spectroscopy : the potential and the limitations. Aust. J. Exp. Agr. 38 : 697-706.

Baydar, H., R. Marquard, and I. Turgut. 1999. Pure line selection for improved yield, oil content and different fatty acid composition of sesame, *Sesamum indicum*. Plant Breeding 118 : 462-464.

Choung, M. G., S. T. Kang, W. Y. Han, I. Y. Baek, H. K. Kim, K. S. Kim, and S. H. Park. 2005. Determination of fatty acid composition in soybean seed using near infrared reflectance spectroscopy. Korean J. Breeding 37 : 197-202.

Fassio, A. and D. Cozzolino. 2004. Non-destructive prediction of chemical composition in sunflower seeds by near infrared spectroscopy. Ind. Crop Prod. 20 : 321-329.

Font, R. M., M. del Rio-Celestino, J. M. Fernandez, and A. del Haro-Bailon. 2003. Acid detergent fiber analysis in oilseed *Brassicas* by near-infrared reflectance spectroscopy. J. Agric. Food Chem. 51 : 2917-2922.

Kim, G. S., D. H. Kim, M. R. Jeong, I. B. Jang, K. B. Shim, C. H. Kang, S. E. Lee, N. S. Seong, and K. S. Song. 2004. Quantitative analysis of sesamin and sesamolin in various cultivars of sesame. Korean J. Crop Sci. 49 : 496-502.

Kim K. S., S. H. Park, and M. G. Choung. 2006. Nondestructive determination of lignans and lignan glycosides in sesame seeds by near infrared reflectance spectroscopy. J. Agric. Food Chem. 54 : 4544 -4550.

Oh, K. W., M. G. Choung, S. B. Pae, C. S. Jung, B. J. Kim, Y. C. Kwon, J. T. Kim, and Y. H. Kwack. 2000. Determination of seed lipid and protein contents in perilla and peanut by near-infrared reflectance spectroscopy. Korean J. Crop Sci. 45 : 339-342.

Sato, T., A. A. Maw, and M. Katsuta. 2003. NIR reflectance spectroscopic analysis of the FA composition in sesame (*Sesamum indicum* L.) seeds. J. Am. Oil Chem. Soc. 80 : 1157-1162.

Shenk, J. S. and M. O. Westerhaus. 1991a. Population structuring of near infrared spectra and modified partial least squares regression. Crop Sci. 31 : 1548-1555.

Shenk, J. S. and M. O. Westerhaus. 1991b. New standardization and calibration procedures for NIRS analytical systems. Crop Sci. 31 : 1694-1696.

Velasco, L., B. Perez-Vich, and J. M. Fernandez-Martinez. 1999. Nondestructive screening for oleic and linoleic acid in single sunflower achenes by near-infrared reflectance spectroscopy. Crop Sci. 39 : 219-222.

Were, A. B., O. A. Onkware, S. Gudu, M. Welander, and S. A. Carlsson. 2006. Seed oil content and fatty acid composition in East African sesame (*Sesamum indicum* L.) accessions evaluated over 3 years. Field Crop Res. 97 : 254-260.

Williams, P. and K. Norris. 2001. Near infrared technology in the agricultural and food industries, 2nd ed., American Association of Cereal Chemists Inc., St. Paul, MN, USA.