

## Evaluation of Millet (*Panicum miliaceum* subsp. *miliaceum*) Germplasm For Seed Fatty Acids Using Near-Infrared Reflectance Spectroscopy

Young-Yi Lee\*<sup>†</sup>, Jung-Bong Kim\*\*, Ho-Sun Lee\*, Young-A Jeon\*, Sok-Young Lee\*, and Chung-Kon Kim\*

\*Agrobiodiversity Center, National Academy of Agricultural Science, RDA, Suwon 441-717, Korea

\*\*Dept. of Agro-Food Resources, National Academy of Agricultural Science, RDA, Suwon 441-857, Korea

**ABSTRACT** The objective of this study was to rapidly evaluate fatty acids in a collection of millet (*Panicum miliaceum* subsp. *miliaceum*) of different origins so that this information could be disseminated to breeders to advance germplasm use and breeding. To develop the calibration equations for rapid and nondestructive evaluation of fatty acid content, near-infrared reflectance spectroscopy (NIRs) spectra (1104-2494 nm) of samples ground into flour ( $n=100$ ) were obtained using a dispersive spectrometer. A modified partial least-squares model was developed to predict each component. For foxtail millet germplasm, our models returned coefficients of determination ( $R^2$ ) of 0.89, 0.89, 0.89, and 0.92 for palmitic acid, oleic acid, linoleic acid, and total fatty acids, respectively. The prediction of the external validation set ( $n=10$ ) showed significant correlation between references values and NIRs values ( $r^2=0.64, 0.90, 0.79,$  and  $0.89$  for palmitic acid, oleic acid, linoleic acid, and total fatty acids, respectively). Standard deviation/standard errors of cross-validation (SD/SECV) values were close to 3 (2.62, 2.40, 1.85, and 2.23 for palmitic acid, oleic acid, linoleic acid, and total fatty acids, respectively). These results indicate that these NIRs equations are functional for the mass screening and rapid quantification of the oleic and total fatty acids characterizing millet germplasm. Among the samples, IT153514 showed an especially high content of fatty acids ( $48.14 \text{ mg g}^{-1}$ ), whereas IT123909 had a very low content ( $34.44 \text{ mg g}^{-1}$ ).

**Keywords** : fatty acids, millet, germplasm, near-infrared reflectance spectroscopy

**Minor** coarse grains serve as supplementary foods for humans and have functional traits that differ from major grain foods. Antioxidant, antitumor, and antimutagenic effects have been reported for some minor cereal crops including buckwheat, sorghum, millet, and Job's tears (Kwak *et al.*,

2004), likely resulting from functional phytochemicals such as vitamins, phenolics, tocopherols, and phytosterols. These chemical traits could be targeted as principal breeding characters. Thus, characterization of coarse grain germplasms for these traits should be performed so that useful germplasms can be identified and recommended to breeders or users.

Rapid characterizations and evaluations of breeding-targeted characters are very important for conservation, management, and utilization of plant germplasm (Vines *et al.*, 2005; Lee *et al.*, 2009). The Rural Development Administration (RDA) genebank at the National Agrobiodiversity Center (NAC) of Korea conserves 165,000 seed germplasms, more than 75% (165,753 accessions) of which are grain food crops (RDA, 2011). For many years, rice and soybean germplasms have been the focus of food crop breeding in Korea, and the germplasms of these crops have been actively characterized and evaluated. However, other grain crop germplasms have been less investigated. The National Agrobiodiversity Center conserves germplasms of 23,070 accessions of coarse cereal crops, including adzuki and mung beans.

Near-infrared reflectance spectroscopy (NIRs) is a rapid technique that can be used to measure many traits routinely tested in cereal breeding programs. This technique has several advantages over conventional laboratory methods. For some major components of rice grains such as starch, protein, amino acids and fats, NIRs calibration models are well developed. However, no NIRs models have been developed for other components such as phenolic and tocopherol compounds in more minor crop grains, including adzuki bean, mung bean, and foxtail millet, even though NIRs methodology has proved feasible in measuring the phenolic and flavonoid contents of other systems and crops.

<sup>†</sup>Corresponding author: (Phone) +82-31-299-1814 (E-mail) youngyi@korea.kr

<Received 9 November 2011; Revised 29 November 2011; Accepted 5 December 2011>

For example, phenolic compounds in red wine fermentation have been accurately predicted by NIRs (Cozzolino *et al.*, 2004). Prediction models for total phenolics in eucalyptus leaves, a forage legume, green rooibos, and green tea leaves have also been reported.

In this study, foxtail millet (*Panicum miliaceum* subsp. *miliaceum*) germplasm collections were characterized and evaluated *via* NIRs. Our objective was to investigate whether fatty acids (FAs) of a coarse cereal grain could be predicted by NIRs. Successful prediction will contribute to more effective application of NIRs in coarse cereal germplasm use and crop breeding programs.

## MATERIALS & METHODS

### Germplasm seed samples

The 516 accessions of millet germplasm conserved at the NAC, RDA genebank were NIR-scanned as grains. Samples were ground, NIR-scanned again as flour, and then analyzed for FAs using gas chromatography (GC).

### Spectroscopic analysis

To obtain NIR reflectance spectra, samples were scanned immediately after grinding in triplicate spectral analysis cells using a dispersive NIR Systems 6500 spectrometer (FOSS North America, Eden Prairie, MN, USA). The instrument was fitted with a spinning cup sampling device. Each cell was scanned 16 times, and the spectra were averaged and transformed to  $\log(1/R)$ . After visual inspection, the spectra of the triplicate subsamples were averaged.

### Fatty acid analysis

Fatty acid methyl esters (FAMES) were prepared using potassium hydroxide in dry methanol (2 N) and extracted with n-hexane, as described by American Oil Chemists' Society (AOCS) Official Method Ce 2-66 (AOAC, 1990). FAMES (1  $\mu$ l) were analyzed using an HP 5890 series II gas chromatograph (GC; Hewlett-Packard, Wilmington, DE, USA) equipped with a flame ionization detector (FID) and an HP 7673A automatic injector (Agilent Technologies, Palo Alto, CA, USA). A fused silica DB23 capillary column (60 m  $\times$  0.25 mm i.d., 0.25  $\mu$ m film thickness; J&W Scientific, Folsom, CA, USA) was used. Oven temperatures were

programmed as follows: 170°C for 3 min, increased to 220°C at 3°C/min, and kept at 220°C for 15 min. The injection and detector temperatures were 250°C and 260°C, respectively. The carrier gas was nitrogen, and the flow rate was 1 mL min<sup>-1</sup>. The split ratio was 1/65. FAME identification was based on retention times compared with those of the standard FAME mixture. Results were expressed without correction as percentage of peak area. Fatty acid analysis of each sample of each variety was performed in triplicate, and average values are reported.

### Calibration Development

The NIRs reflectance model was developed using a commercial spectral analysis program (ISI40 NIRS 2 version 4.01 and WINISI software, FOSS North America). Preprocessing of the spectral data (1104-2494 nm) of 100 samples consisted of a normal multiplicative scatter. The data were centered on the mean spectrum and mean reference value using modified partial least-squares (PLS) regression. The PLS regression was modified so that reference values and reflectance data were scaled at each wavelength to a standard deviation of 1.0 before each PLS regression term. The preprocessing methods chosen for the model were the optimum to obtain minimum error following cross-validation (four cross-validation groups per germplasm collection). The optimum number of PLS regression terms for the calibration, also determined by cross-validation, was the number of factors yielding the minimum error between predicted and reference values (standard error of cross-validation, SECV). The modified PLS regression model was tested using independent validation samples (n = 10). Statistics used to assess the model were the standard error of performance (SEP, not bias-corrected), coefficient of determination ( $R^2$ ), and slope and intercept of the linear regression of NIRs predicted versus analyzed values.

## RESULTS AND DISCUSSION

### Spectra collection via NIRs scanning

In the present study, the NIR spectra of 516 accessions of millet germplasm were obtained as grain status which is not ground. And then about 100 accessions of each of those were selected by spectral data, origin and color of

seed coat. Fig. 1 shows the spectra of selected samples are almost similar to those of whole spectra in diversity that we can see as the width of spectra collection. It means the selected samples could be representative to whole collections of germplasm. About at 1974 nm, a peak was observed in all spectra. So the absorbance at the wavelength of 1974 nm is thought to be decisive value that concern chemical composition including FAs composition. Because little distinct peaks were observed in raw spectra without math treatment, these spectra were modified to find out more correlative wavelength with determination of protein by mathematic treatment in variable peak and shape. After mathematic treatment, some absorption bands displayed clearer peak (data is not shown). Subsequently selected 100 samples were ground and NIR scanned again before GC analysis of FAs was started. In the spectra of flour ground samples, their peaks are sharper and clearer than in those of grain samples not ground. So these two patterns of spectrum could be compared and investigated which one was more appropriated for NIR model regression. Zhang *et al.* (2008) reported that the spectra of the dehulled rice grain were very similar to those of milled rice grain, but

the grain sample clearly had stronger energy absorption than flour samples.

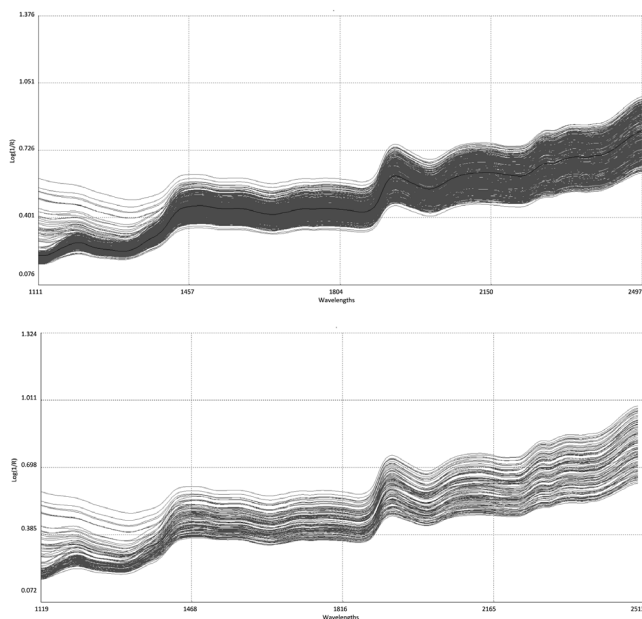
#### Determination of reference values

Fig. 2 shows the FA contents of the millet samples. Total FAs averaged  $39.13 \text{ mg g}^{-1}$ , with a range of  $34.44\text{--}48.14 \text{ mg g}^{-1}$ . Among the millet collections, the IT123909 and IT153514 accessions showed especially low and high FA contents, respectively. Constituent acid content ranges were as follows: palmitic acid,  $3.77\text{--}6.00$ ; stearic acid,  $0.94\text{--}1.94$ ; oleic acid,  $5.50\text{--}12.48$ ; linoleic acid,  $22.18\text{--}31.61$ ; and linolenic acid,  $0.38\text{--}0.88 \text{ mg g}^{-1}$ . Oleic and linoleic acid values varied widely among the accessions, suggesting that the NIRs equation modeling values of the two acids would be significant, as collection diversity correlates positively with regression models.

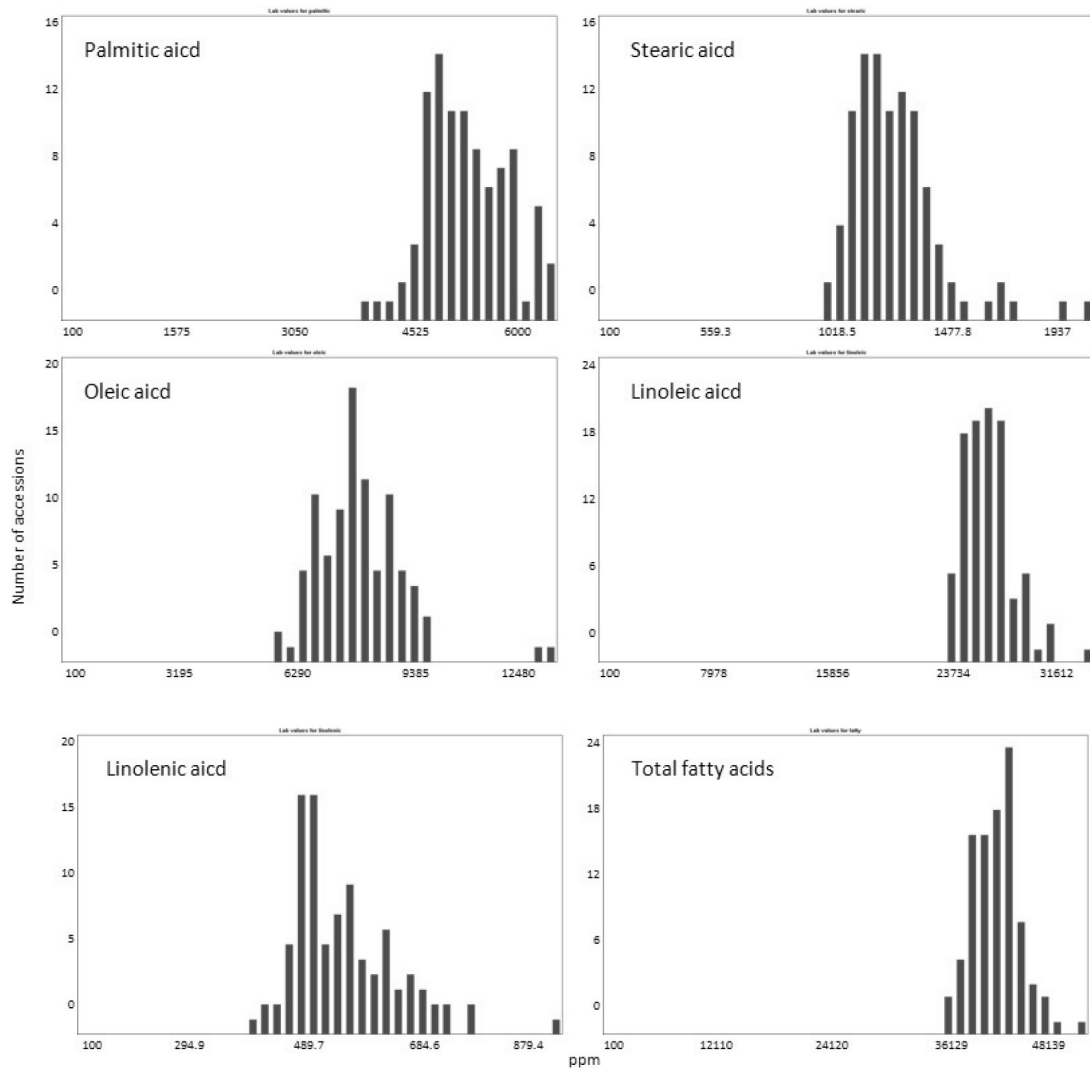
#### NIRs calibration for modeling

The reference values of all of the samples were used in mathematical treatments designed to create prediction equations using modified PLS. After using the equations to calculate predicted values, we conducted an external validation using an independent sample set ( $n=10$ ) to ensure that the equations could be applied to the prediction of FAs. Fig. 3 shows the validation graphics, and Fig. 4 gives the predicted values.

Table 1 shows that the predicted values of grain constituents and reference values of palmitic acid was better correlated ( $R^2=0.85$ ) than were other components. Table 2 shows that the predicted values of flour constituents and reference values of palmitic acid, oleic acid, linoleic acid, and total FAs were better correlated ( $R^2=0.89, 0.89, 0.89, \text{ and } 0.92$ , respectively) than were other components. Because the  $R^2$  values of stearic acid and linolenic acid (0.36 and 0.52, respectively) were lower than those of other components, we do not regard these acids as predictable by NIRs. Additionally, we used the SD/SECV ratio to evaluate the reliability of the calibration model (Moschner & Biskuper-Korell, 2006). The SD/SECV values for palmitic, oleic, and total fatty acids were 2.62, 2.40, and 2.23, respectively. These high values (close to 3.0) indicate a significant correlation between the reference values and NIRs estimated values, similar to those of the calibration. For other components, the SD/SECV values were below to 2, which could mean



**Fig. 1.** NIR spectra for grain condition of millet germplasm. Left, spectra for whole samples of 516 accessions of millet germplasm; Right, spectra for 100 accessions which were selected for NIR calibration.



**Fig. 2.** Histograms for germplasm distribution according to fatty acids contents in foxtail millet collection(n=100) used for the calibration equation.

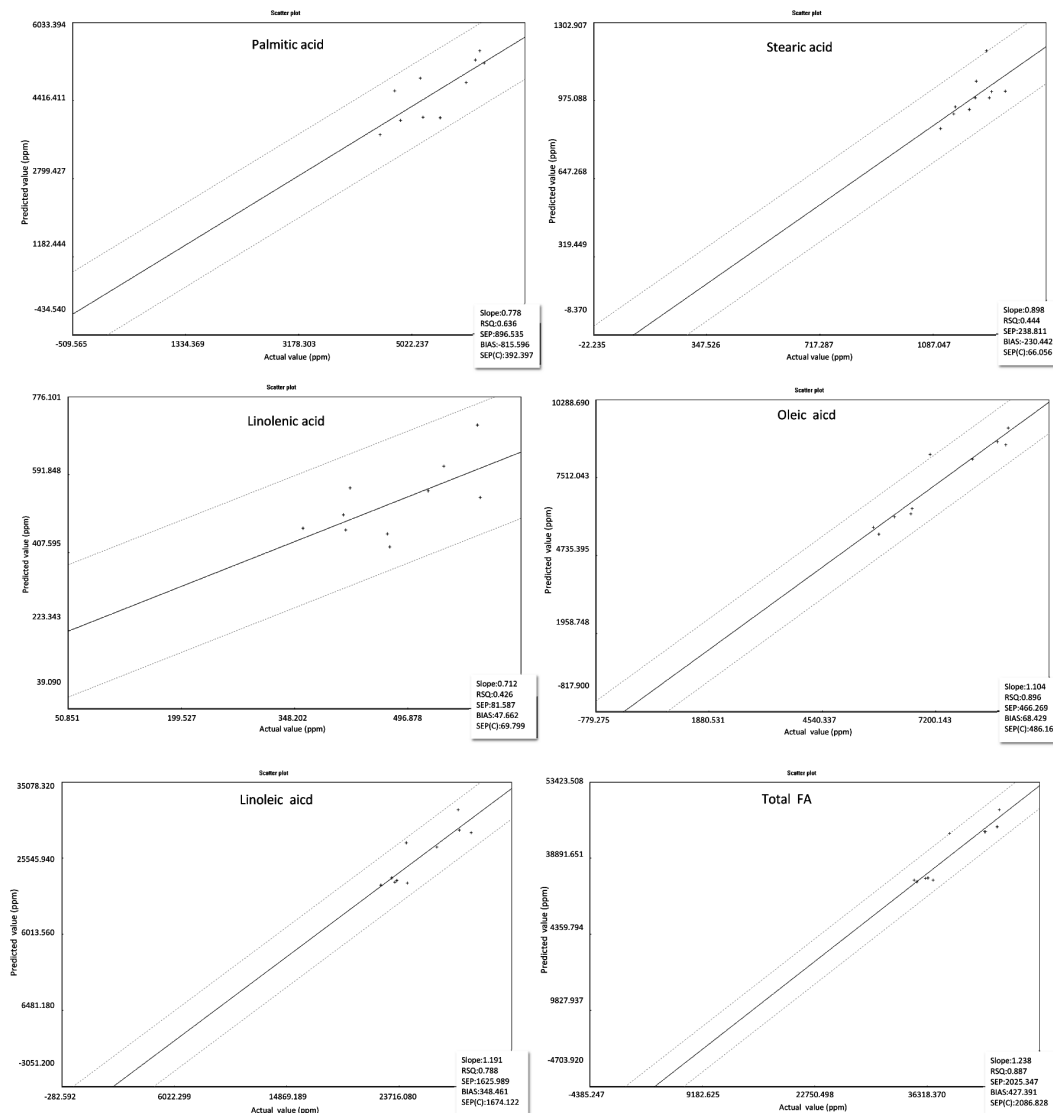
rare useful for prediction.

We also compared the effectiveness of NIRs scanning of grain versus flour samples. In all cases, the predictive capability of flour samples exceeded that of whole grain samples, although the  $R^2$  pattern of flour and grain samples differed in some cases, most likely owing to the sample grinding process, as some phytochemicals are sensitive to high temperature or are located in the seed coat. The nature of the phytochemical components of the flour and grain samples has yet to be compared.

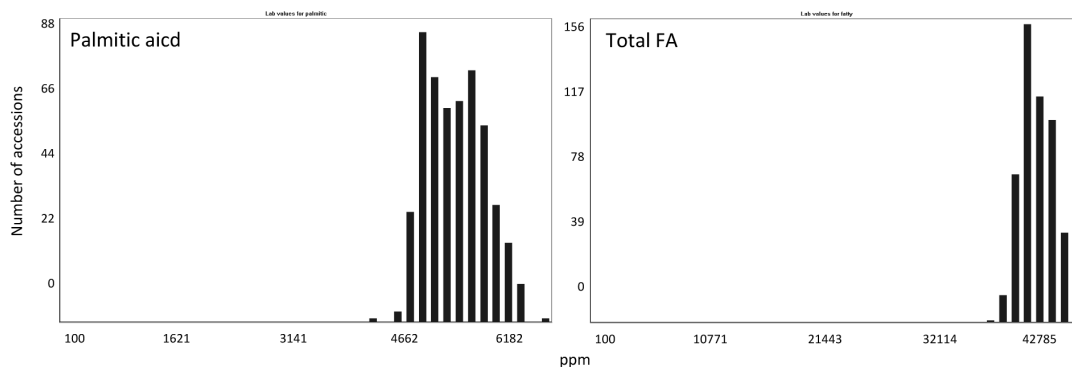
NIRs absorbance is greatly affected by OH groups. The moisture condition of all of the germplasm conserved at the NAC was assumed to be the same, as all germplasm

materials at the center are conserved in identical mid-term conservation conditions of 4°C. A future study of the seed moisture conditions of each accession would be warranted.

In conclusion, palmitic acid and total FA contents can be predicted with reliable accuracy using NIRs analysis of intact millet seeds. But a larger germplasm population covering a wide range of chemical values was required to accurately predict stearic, linoleic, and linolenic acid contents (Gotor *et al.*, 2007). With a larger sample size and broader variation of the reference dataset, calibration models for determining those FAs could be employed in a massive screening of breeding lines, despite the reduced accuracy of the models.



**Fig. 3.** Scatter plots of NIR predicted by developed equation versus actual-analyzed reference values for individual and total fatty acids ( $R^2=0.64, 0.44, 0.90, 0.79, 0.43,$  and  $0.89$  for palmitic, stearic, oleic, linoleic, linolenic, and total fatty acid, respectively) in the external validation set ( $n=10$ ) of flour samples of millet germplasm.



**Fig. 4.** Histogram for the distribution of whole 516 accessions according to NIR predicted value of fatty acids by the equation developed in this study.

**Table 1.** Equation statistics using regression model (MPLS) and scatter correction for NIRS prediction of chemical constituents in the calibration set (n = 100) of millet grain samples.

Constituent (mg g <sup>-1</sup> )	N	Mean	Calibration		Cross-validation		Math treatment
			SEC	R <sup>2</sup>	SECV	1-VR	
Palmitic acid	97	4.95	0.18	0.85	0.22	0.79	2,4,4,1
Stearic acid	94	1.17	0.09	0.32	0.10	0.18	2,8,8,1
Oleic acid	96	7.47	0.50	0.67	0.61	0.52	2,4,4,1
Linoleic acid	94	24.67	1.08	0.26	1.16	0.13	3,5,5,1
Linolenic acid	98	0.52	0.05	0.48	0.06	0.41	1,4,4,1
Total fatty acid	96	38.88	1.55	0.49	1.80	0.31	3,5,5,1

N, number of samples used to develop the model; SD, standard deviation; SEC, standard error of calibration; R<sup>2</sup>, coefficient of determination of calibration; SECV, standard error of cross-validation; 1-VR, coefficient of determination for cross-validation

**Table 2.** Equation statistics using regression model (MPLS) and scatter correction for NIRS prediction of chemical constituents in the calibration set (n = 100) of millet flour samples.

Constituent (mg g <sup>-1</sup> )	N	Mean	Calibration		Cross-validation		Math treatment
			SEC	R <sup>2</sup>	SECV	1-VR	
Palmitic acid	96	4.96	0.47	0.89	0.18	0.86	3,5,5,1
Stearic acid	95	1.18	0.12	0.36	0.12	-0.01	3,5,5,1
Oleic acid	94	7.45	0.87	0.89	0.36	0.83	1,4,4,1
Linoleic acid	94	24.86	1.42	0.89	0.77	0.71	3,5,5,1
Linolenic acid	96	0.52	0.08	0.52	0.05	0.48	1,4,4,1
Total fatty acid	96	38.99	2.21	0.92	0.99	0.80	3,5,5,1

N, number of samples used to develop the model; SEC, standard error of calibration; R<sup>2</sup>, coefficient of determination of calibration; SECV, standard error of cross-validation; 1-VR, coefficient of determination for cross-validation.

## REFERENCES

- Cozzolino D., M. J. Kwiatkowskia, M. Parkera, W. U. Cynkara, R. G. Dambergasa, M. Gishena, and M. J. Herderich. 2004. Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Anal. Chim. Acta* 513(1) : 73-80.
- Gotor A. A., M. Berger, E. Farkas, F. Labalette, S. Centis, and A. Calmon. 2007. Quantification of sunflower minor components by near-infrared spectrometry (NIR). *Helia*. 30(47) : 183-190.
- Kim Y. H., C. S. Kang, and Y. S. Lee. 2004. Quantification of tocopherol and tocotrienol content in rice bran by near infrared reflectance Spectroscopy. *Korean J. Crop Sci.* 49(3) : 155-275.
- Lee S. Y., Y. Y. Lee, S. K. Lee, Y. H. Cho, K. H. Ma, H. K. Kang, J. G. Gwag, Y. H. Park, and H. S. Lee. 2009. Selecting high amylose lines of rice varieties using NIR spectroscopy at the RDA gene bank conserved. *Korean journal of breeding science*, 41(1) : 108.
- Moschner, C. and B. Biskupek-Korel. 2006. Estimating the content of free fatty acids in high oleic sunflower by near-infrared spectroscopy. *Eur. J. Lipid Sci. Technol.* 108(7) : 606-613.
- RDA. 2011. The status of plant genetic resource conservation.
- Uauy C., A. Distelfeld, T. Fahima, A. Blechl, and J. Dubcovsky. 2006. A NAC Gene regulation senescence improve grain protein, zinc, and iron content in wheat. *Science* 314 : 1298-1301.
- Vines L., S. E. Kays, and P. E. Koehler. 2005. Near-infrared reflectance model for the rapid prediction of total fat in cereal foods, *J. Agric. Food Chem.* 53 : 1550-1555.
- Zhang C., Y. Shen, J. Chen, P. Xiao, and J. Bao. 2008. Nondestructive prediction of total phenolics, flavonoid contents, and antioxidant capacity of rice grain using near-infrared spectroscopy. *J. Agric. Food Chem.* 56 : 8268-8272.