Use of near infrared spectroscopy for the estimation of total isoflavone content on a single seed soybean analysis

Tetsuo Sato,[a] Kentaro Eguchi[a] and Yoichi Nishiba[b]

[a]*Analysis and Monitoring Office, National Agricultural Research Center for Kyushu Okinawa Region (KONARC), National Agriculture and Food Research Organization (NARO), Koshi-city, Kumamoto, Japan. E-mail: sato@affrc.go.jp*

[b]*Crop Functionality and Utilization Research Team, KONARC, NARO, Koshi-city, Kumamoto, Japan*

Introduction

Soybean (*Glycine max* L.) is a major oilseed crop, and a good source of nutritional constituents, such as oil and protein. In Japan, soybean is used for producing excellent traditional foods (such as miso and tofu), and also for development of new industrial foods (such as snack foods and supplements). In order to increase consumption of soybeans, it is desirable to increase the contents of physiologically functional elements, such as isoflavones, which give added value to soybeans. Isoflavones have a function in prevention of osteoporosis.[1] Recently, farmers and consumers are increasingly focusing their attentions not only on the major nutritional constituents and palatability, but also on the physiologically functional constituents. Conventional methods for the determination of these physiologically functional elements are very labor intensive. A simple and rapid method for the estimation of them is necessary for screening soybean varieties for plant breeding. Furthermore, analysis of small amounts of sample is demanded. In this report, the feasibility of near infrared (NIR) spectroscopy for the estimation of the total isoflavone content in single soybean seeds was examined.

Materials and methods

Samples and chemical analyses

The soybean samples were the same as those used in an earlier study.[2] Forty-eight samples were collected from various areas of Japan, and sent to our research center to be analyzed. The cultivated areas of the samples were from northern (Hokkaido) to southern (Kumamoto) Japan. The samples were milled with an ultra-centrifugal mill ZM1000 (Retsch Co., Germany) through a screen ($\phi = 1.0$ mm).

Figure 1. A single grain cup for a single seed (left) and a whole grain cup for plural seeds (right).

The contents of the individual isoflavones were determined by a high-performqance liquid chromatography (HPLC) method.[2] The total isoflavone content was the sum of the contents of the individual isoflavones. The unit is mg (100 g DW)$^{-1}$. The samples were divided into two sets: a calibration set ($n = 36$) and a prediction set ($n = 12$). The fundamental statistics of the total isoflavone contents of the samples were respectively as follows (minimum–maximum, mean ± standard deviation): 133.44–633.42, 313.26 ± 116.83 for the calibration set, and 156.96 – 482.24, 286.07 ± 107.79 for the validation set.

Near infrared spectroscopic measurements

An InfraAlyzer 500 (IA500) (Bran + Luebbe (B+L) GmbH, Norderstedt, Germany) was used to measure the NIR reflectance spectra over the wavelength range from 1100 to 2500 nm at 2-nm intervals. A single soybean seed was placed in the hole of a single seed cup (Figure 1, left) on the standard drawer for the NIR measurements. Five individual seeds per sample were measured, and average spectra were calculated for the statistical analysis.

Statistical analysis

Multiple linear regression (MLR) analysis of the NIR data with the chemical data was carried out using IDAS software (B+L) on the calibration set. When the first- and second-derivative NIR spectra were calculated, the default parameters were used. Validation of the calibration equations was carried out using the prediction sample set. The Unscrambler (version 9.6; Camo Co., Norway) was also used on the IA500 data for partial least squares regression (PLSR) or principal component regression (PCR) analysis. The data were analysed not only on the original spectra, but also on the derivatised, i.e., pretreated spectral data. The derivative dimensions were as follows: gap 11, segment 10 for the first derivative (abbreviated as d1); and gap 10, segment 11 for the second derivative (d2).

Further, the calibrations developed for multiple soybean seeds, which were previously reported,[3] were also applied to the single seed spectral data in order to examine their performance on single seeds.

Table 1. Results of statistical analysis for total isoflavone content by NIR on a single seed analysis.

	Calibration			Prediction			
		r	*SEC*	*r*	*SEP*	Bias	MC-*SEP*
Single seed analysis	MLR Analysis by IDAS	0.77	79.62	0.80	69.92	23.25	68.87
	PLS analysis by Unscrambler	0.89	53.70	0.79	69.01	19.33	69.20
Multiple seeds analysis applying to single seeds spectra	MLR analysis by IDAS	—	—	0.83	62.99	18.21	62.98
	PLS analysis by Unscrambler	—	—	0.96	158.84	155.33	34.70

The unit is mg $(100\,gDW)^{-1}$

Results and discussion

Statistical analysis for total isoflavone content on a single seed analysis

For single seed analysis, based on the standard error of prediction (*SEP*), NIR spectroscopy may be available for making a rough non-destructive estimate of the total isoflavone content by either MLR- or PLSR-analysis (Table 1, upper columns). The scatter plots of these results are shown in Figure 2. The same level of *SEP* was obtained in single seed analysis as in the analysis of multiple soybean seeds.[3] Kudou *et al*.[4] reported that isoflavones are distributed mainly in hypocotyls of soybean seeds, i.e., on the surface of a seed, which may account for the reasonable results obtained for single soybean seed analysis.

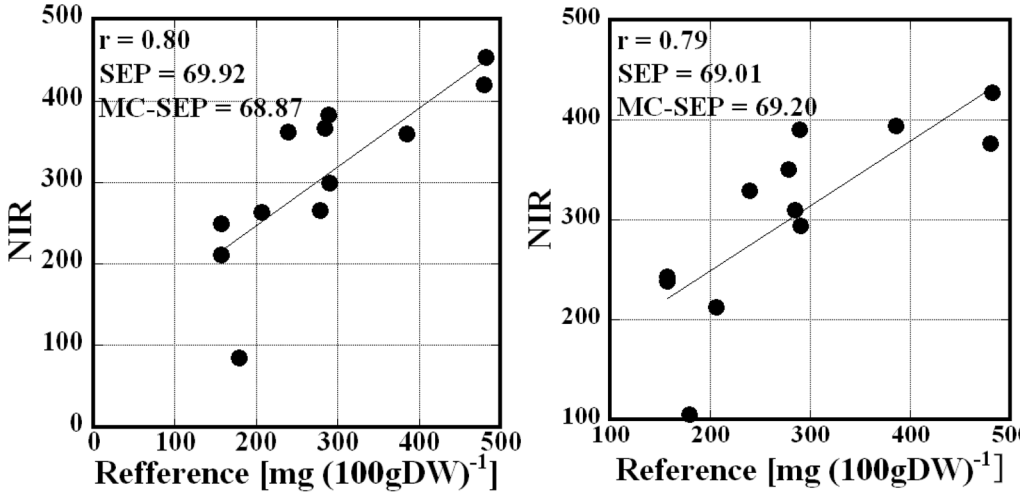

Figure 2. Scatter plots of the results of statistical analysis for total isoflavone content by NIR on a single seed analysis. (left) MLR-analysis and (right) PLSR-analysis.

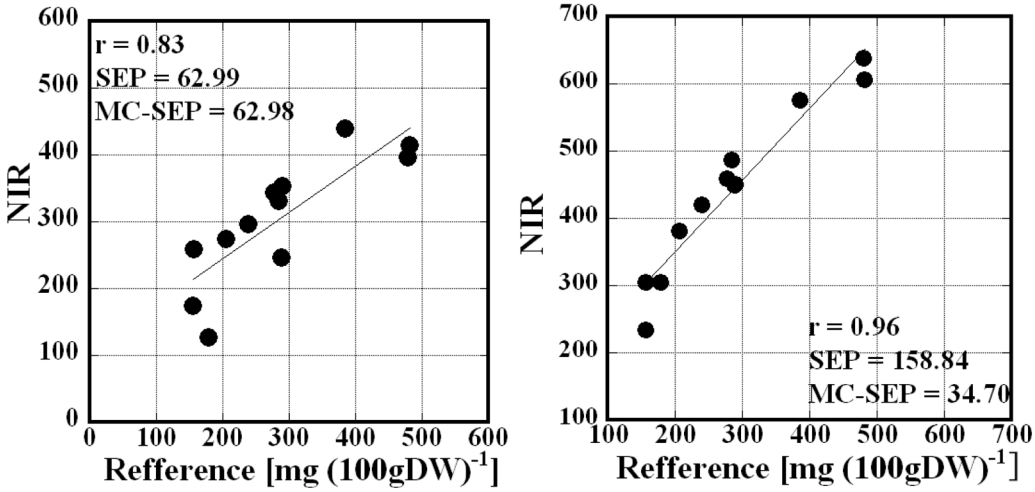

Figure 3. Scatter plots of adaptation of calibrations obtained by multiple seeds analysis to single seeds spectra, (left) MLR-analysis and (right) PLSR-analysis.

Adaptation of calibrations obtained by multiple seeds analysis to single seeds spectra

The results of calibrations developed for multiple seeds when applied to single seed analysis by both MLR- and PLSR-analysis are described in the lower columns of Table 1. The scatter plots of these results are shown in Figure 3. The spectra of multiple soybean seeds have stronger variations than do those of single seeds. In the case of PLSR, all spectral data were used, and the difference was accumulated in bias. This then made the bias very large. However, the relative level of total isoflavone maintained magnitude relation, and the MC-*SEP* was drastically improved by correcting bias. Based on the results of MC-*SEP*, the level of the total isoflavone content could be estimated by applying the multiple seed analysis NIR calibration to the analysis of single seeds. The same level of MC-*SEP* was obtained in the single seed analysis case as in multiple soybean seed analysis.

In conclusions, NIR spectroscopy may be available for enabling a rough nondestructive estimate of the total isoflavone content by either MLR- or PLSR analysis on a single seed spectrum.

References

1. Y. Yamori, in *Soybean Isoflavones,* Ed by Y. Yamori, S. Ohta and S. Watanabe. Saiwai. Shobou Co., Tokyo, Japan, p.79 (2001), in Japanese.
2. Y. Nishiba, I. Suda, T. Oki and T. Sugawara, *Nippon Shokuhin Kagaku Kogaku Kaishi.* **54,** 295 (2007), in Japanese.
3. T. Sato, K. Eguchi, T. Hatano and Y. Nishiba, *Plant Prod. Sci.,* **11,** 481 (2008).
4. S. Kudou, Y. Fleury, D. Welti, D. Magnolato, T. Uchida, K. Kitamura and K. Okuba, *Agric. Biol. Chem.* **55,** 2227 (1991).