World Scientific
www.worldscientific.com

Nondestructive classification of mung bean seeds by single kernel near-infrared spectroscopy

Kaewkarn Phuangsombut[*], Nattaporn Suttiwijitpukdee[†]
and Anupun Terdwongworakul[*,‡]

[*]Department of Agricultural Engineering
Faculty of Engineering at Kamphaengsaen
Kasetsart University, Kamphaengsaen
Nakhon Pathom, 73140 Thailand

[†]Kasetsart Agricultural and Agro-Industrial
Product Improvement Institute
Kasetsart University, Bangkok, 10900 Thailand
[‡]fengant@ku.ac.th

Near-infrared spectroscopy (NIRS) in the range 900–1700 nm was performed to develop a classifying model for dead seeds of mung bean using single kernel measurements. The use of the combination of transmission-absorption spectra and reflection-absorption spectra was determined to yield a better classification performance (87.88%) than the use of only transmission-absorption spectra (81.31%). The effect of the orientation of the mung bean with respect to the light source on its absorbance was investigated. The results showed that hilum-down orientation exhibited the highest absorbance compared to the hilum-up and hilum-parallel-to-ground orientations. We subsequently examined the spectral information related to the seed orientation by developing a classifying model for seed orientation. The wavelengths associated with classification based on seed orientation were obtained. Finally, we determined that the re-developed classifying model excluding the wavelengths related to the seed orientation afforded better accuracy (89.39%) than that using the entire wavelength range (87.88%).

Keywords: Mung bean; germination; near-infrared spectroscopy; classification; single kernel.

1. Introduction

The mung bean (*Vigna radiata*) is one of the important crops used as an ingredient (in the form of sprouts) in various foods for consumption in Thailand due to its high nutritional value. The sprouts of mung beans resulting from germination improve the nutritional and medicinal qualities of

‡Corresponding author.

mung beans[1] through changes in the chemical constituents, nutritional value, and metabolites occurring during the sprouting process.[2] Consequently, the presence of hard seeds during seed collection poses a major problem to mung bean consumption, particularly during the sprouting process or when boiling. In this context, the seed size also affects the germinating ability. Larger seeds (with relatively heavier and larger heads and stems) were found to correspond to a higher germination percentage.[3] The "hardness" of seeds is a genetically and physiologically controlled trait.[4] The seed coats of hard seeds contain seven times more lignin, 23% higher silica and 12% higher fiber content than those of normal seeds. Moreover, the hard seed coat consists of tough and highly dense palisade cells. The thickness of the seed coats of the hard seeds is greater and more solid at the ultrastructure level. Further, the micropyle of the hard seed is closed and covered with palisade cells, in which covering makes hard seeds impermeable to water. In the seed production process, hard seeds need to be separated from normal seeds as the hard seeds will not grow as they are impervious to water, and they do not germinate in the sprouting process. The moist hard seeds that do not sprout can develop mold and infect neighboring sprouting seeds in the process. Thus, the separation of hard seeds from the normal ones forms an important aspect of seed collection and separation. At present, the only method used for the separation of hard seeds from normal ones involves the use of a reciprocating force; the inclined and reciprocating movement of a sieve plate separates the hard seeds from the sound ones. The working principle of this approach is based on the difference in the densities of the hard seeds and normal seeds. However, this technique is not very effective.

One of the techniques for nondestructive assessment of internal quality of agricultural products that is becoming increasingly popular is near-infrared spectroscopy (NIRS). NIRS has been widely applied to evaluate the internal quality of beans with the primary advantage that the technique requires minimum sample preparation.[5–9] In one mung bean breeding program, NIRS reflectance spectra were obtained from intact grains and the palmitic, linoleic, and total fatty acid contents were successfully predicted with reliable accuracy.[10] Further, NIRS has been utilized to evaluate the contents of protein, starch, and amylose in

intact mung bean seeds for seed conservation purposes.[11] The investigators reported that the NIRS method is robust, affords nondestructive determination of the seed quality, and the nondestructive aspect preserves the seed's vitality. NIRS has also been utilized to assess the internal quality of certain other beans and seeds. For examples, models based on partial least squares regression (PLSR) have been developed to analyze the nutritional composition of the Chinese Faba bean seed. In one such study, the authors reported that a model based on the ground-seed powder generally performed better than the model based on the intact seed in the estimation of protein, starch, oil, and total polyphenol.[12] NIRS has also been reported to show significant potential as a rapid and nondestructive method for the analysis of crude protein and fat content in whole-kernel soybean.[13] From a qualitative viewpoint, NIRS based on transmittance measurements has been useful in discriminating among the insect-infested, filled, and empty seeds of three species of *Larix*.[14] Classification accuracy with the method has been reported to be equal to or nearly 100%.

While nondestructive NIRS has been successfully applied in the abovementioned contexts, there has been no research on developing nondestructive techniques for the discrimination of hard seeds from the normal seeds of the mung bean. Consequently, in this study, we investigated the possibility of applying the NIRS technique to separate hard seeds from normal seeds. Furthermore, we proposed that the classifying model could compensate for the effect of bean orientation on the NIR spectra and that the performance of the classifying model could be improved by combining the transmission-absorption spectra and reflection-absorption spectra.

2. Materials and Methods

2.1. *Sample preparation*

Two crops of mung bean (Kamphaengsaen II variety) were harvested and threshed to separate the seeds from the peels at a crop research center in the Chainat province in the central region of Thailand. Two hundred seeds were subsequently obtained. The seed samples were maintained at 25°C and a relative humidity of approximately 85% in an air-conditioned laboratory overnight to ensure that the

seeds attained temperature equilibrium prior to measurement.

2.2. Spectra acquisition of single kernel

As different orientations of the hilum with respect to the light source yield different reflection and transmission characteristics for the same bean, a uniquely designed holder (Fig. 1) was used to collect one transmission spectrum and two reflection spectra at two positions of each sample. This sample holder was designed so as to allow the attachment of three fiber optic cables to capture the transmitted light from underneath the sample and two reflected light from the top surface of the mung bean. In the measurement, the kernel was placed in the holder and illuminated from the top. The transmitted light was collected through a hole at the bottom and the
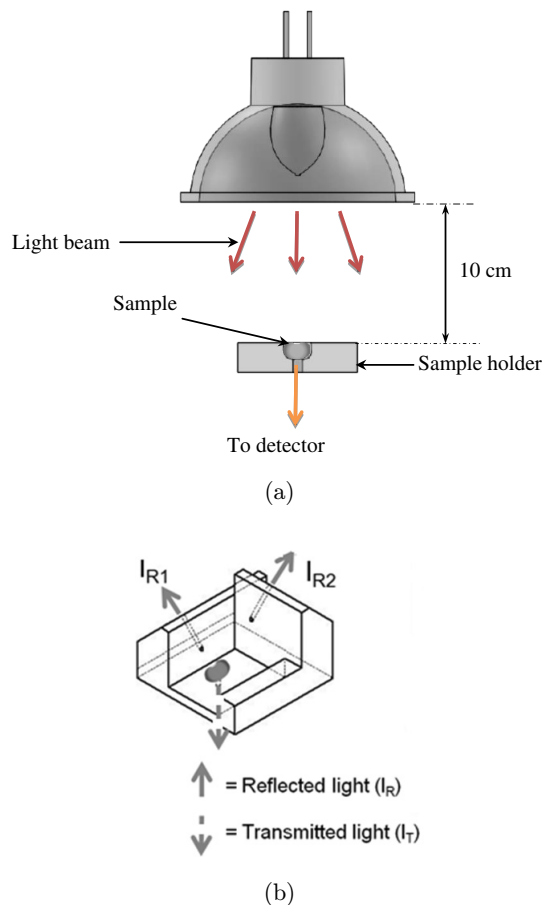

Fig. 1. (a) Mung bean holder showing light source and (b) isometric view of mung bean holder uniquely designed to collect both reflected and transmitted light.

reflected light was collected at one end and the middle part of the seed surface from two fiber optic units installed at 45° with respect to the illumination and spaced 90° apart above the samples.

A light source (100 W, tungsten halogen) was directed onto the kernel and set at a distance of 10 cm above the seed. The transmitted light or reflected light signals were conveyed via an optical fiber to a spectrophotometer (NIRQuest; Ocean Optics; FL, USA) operating in the range 900–1700 nm with a 512-element InGaAs detector and 3.1 nm spectral resolution. The integration time was optimized to 250 ms with a spectral average count of six. Teflon material was used as the reference for each set of 10 seeds.

As a mung bean seed is not symmetrical along one axis, each individual seed was scanned in three orientations relative to the hilum of the seed (Fig. 2). The orientations were designated as hilum-down (HD), hilum-up (HU), and hilum-parallel-to-ground (HP). The absorbance spectra for the three orientations were compared for the effective classification of the hard (dead) mung bean from the normal (sound) bean.

2.3. Germination test

After spectral acquisition, the mung been seeds were germinated for seven days. Each seed was placed on moist paper in a Petri dish at 25°C. The numbers of germinated seeds and hard seeds were recorded on the third day after the start of the germination test.[15]

2.4. Development of classification model

We allocated the samples of single seeds to a calibration set and a prediction set based on the first principal component score derived from principal component analysis. The germinated seeds were sorted according to the first principal component score, which explained 90% of the variance, and subsequently, the seeds were alternately allocated to the calibration and prediction sets. Similarly, the nonviable seeds were also allocated to the calibration and prediction sets with the use of the first principal component score. In total, 134 seeds were allocated to the calibration set (67 germinated seeds and 67 nonviable seeds), and the remaining 66 seeds were allocated to the prediction set (33 germinated
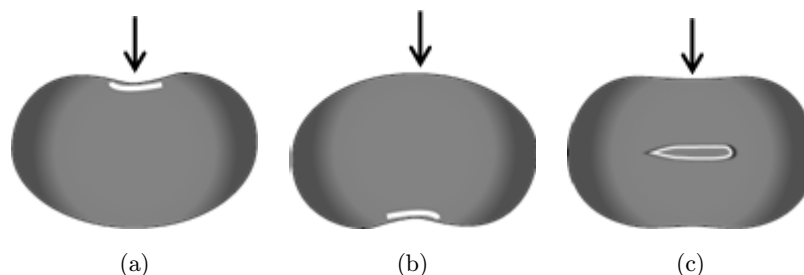
Fig. 2.   Three orientations of mung bean seed for measurement: (a) hilum face-up, (b) hilum face-down and (c) hilum-parallel-to-the-ground.

seeds and 33 nonviable seeds). Models for classification were constructed by implementing partial least squares discriminant analysis (PLS-DA) with the use of the calibration set. The constructed classification models were cross-validated to determine the optimum number of PLS factors using (PLSR) (The Unscrambler v.9.8; Camo; Oslo, Norway), with values of 1 and 0 being set for the germinated and nonviable seeds, respectively. The classification performance of the models was calculated as the percentage correctness via prediction of the germination of the seed samples in the prediction set.

Different pre-treatment techniques were also examined to reduce the effect of scattering (which results in spectral variations) and to improve the classification performance. The techniques, which were applied either individually or in combination, comprised smoothing, calculation of the first and second derivatives of the spectra, multiplicative scatter correction (MSC), and standard normal variate (SNV). Therefore, apart from the raw spectra used for the development of the classification model, pre-treated spectra were also utilized for comparison with regard to the model performance.

As the hilum is present on one side of the seed surface, the scattering effect on the hilum side is different from that of the smooth-sided surface of the mung bean seed. In addition, each placement of the seed on the table does not assume the same orientation. Therefore, in order to investigate the effect of the hilum orientation on the classification performance, four models were developed using the spectra obtained from seeds oriented in the hilum-down (HD), hilum-up (HU), hilum-parallel-to-the-ground (HP) positions, and a combination of the HD, HU, and HP orientations.

Initially, only the transmission-absorption spectra (AT) were used in the PLS-DA analysis to develop a classification model. However, different orientations of the hilum with respect to the light source yielded different reflection and transmission characteristics for the same bean. Therefore, in order to obtain more information on the absorbance, two more fiber optic units were employed to collect reflected light at one end (AR1) and the middle part of the seed surface (AR2) in addition to the transmitted light (Fig. 1(b)). Thus, combined absorbance, which was the combination of the transmission-absorption spectra and two reflection-absorption spectra (AT + AR1 + AR2), was also used as independent variables for further development of the classification model for the improvement of the classifying performance.

3.   Results and Discussion

3.1.   *Absorbance of mung bean seed*

In general, the absorbance characteristics of the mung bean seed (Fig. 3) are similar to those of wheat kernels,[16] with a large and relatively constant absorbance over the long wavelength region from 1430 nm for the sound beans. In our case, noise was present in the absorbance at low and high wavelengths due to the low sensitivity of the sensor. Consequently, the wavelength range from 923 nm to 1400 nm was selected for further analysis to avoid the upper and lower noisy absorbance regions.

Regarding the absorption bands, water bands were observed clearly at wavelengths around 990 nm and 1440 nm (Fig. 3). In between the water bands was a significant starch band at 1200 nm associated with the second overtone of C–H stretching.[17] The presence of the strong starch band
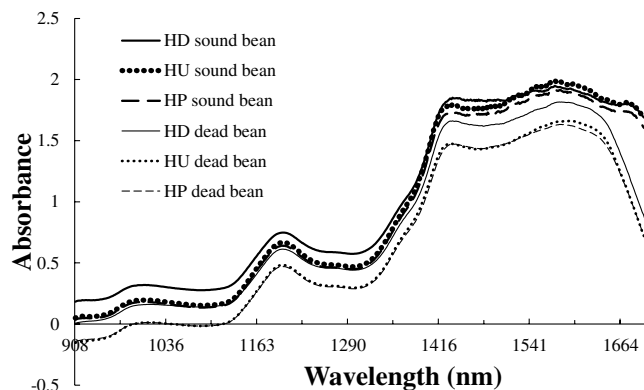
Fig. 3. Average absorbance of mung beans at different hilum orientations (HD: hilum face-down, HU: hilum face-up, and HP: hilum-parallel-to-the-ground).

confirmed the fact that the main composition of the mung bean was starch. In addition, sound seeds appeared to absorb more light than dead seeds, as shown in Fig. 3.

The second-derivative spectra of sound and dead seeds (Fig. 4), which aided in resolving the absorption peaks, exhibited prominent band peaks around 1194 nm and 1363 nm with the sound kernel peaks exhibiting higher absorbance. This result was partly in agreement with a previous report on the NIRS absorbance characteristics of wheat kernels.[16] These band peaks may be assigned to the second overtone and the combination of vibrations of the C–H groups of the structural and food reserve components of grains, such as starch and other carbohydrates and proteins.[18]
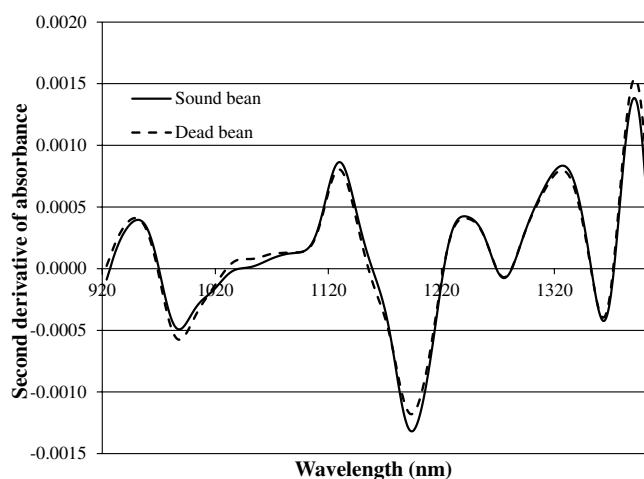

Fig. 4. Second derivative of absorbance for sound and dead mung bean seeds.

The seed orientation investigation revealed that the seed in the HD orientation exhibited the highest absorbance, and the seed in the HP orientation exhibited the lowest absorbance (Fig. 3). One contributing reason was that the HD orientation led to the highest scattering, allowing less light transmission through the seed.

3.2. *Classification*

The PLS-DA model constructed from the transmittance spectra of all orientations yielded a classification accuracy of 81.31% (Table 1). From Table 1, it can also be observed that all the PLS-DA models developed using the combined absorbance, which was calculated from a summation of both transmission-absorption spectra and reflection-absorption spectra, resulted in a better classification than the transmittance model. The PLS-DA model developed using the combined absorbance of all orientations yielded an accuracy of 87.88%, which was higher than that yielded by the model using only the transmittance (81.31%). This improvement in results justified the use of the combination of both transmission-absorption spectra and reflection-absorption spectra in preference to using only transmission-absorption spectra. This combined absorbance partially compensated for variation in the scattering due to the hilum orientation.

Upon investigation of the effect of hilum orientation, we observed that better classification was achieved for all three models based on the HD, HU, and HP orientations. The HD model yielded the best classification accuracy of 95.45%. Among the three orientation models, the model based on the HU orientation yielded the lowest accuracy of 89.39%, probably due to having the highest scattering with respect to other orientations.

The model based on the spectral information of all orientations was more practical for further application as there was no need to place the seed at any specific orientation for measurement. Therefore, we next attempted to compensate for variation in the hilum orientation. Initially, spectral information related to the orientation ("orientation spectra") was required. This was obtained by performing PLS-DA to construct a model capable of separating the mung bean seeds into each orientation. Next, the regression coefficients of the model were examined to select the wavelengths corresponding to high values of the coefficients (Fig. 5).

K. Phuangsombut, N. Suttiwijitpukdee & A. Terdwongworakul

Table 1.   Results of classification performance of calibration models based on absorbance of single kernel of mung bean.

Model*	Number of factors	Calibration		Prediction		Number of misclassified hard seeds***	Number of misclassified normal seeds***	Accuracy (%)
		r**	RMSECV**	r	RMSEP**			
Transmittance with all orientation spectra	4	0.616	0.395	0.648	0.382	24 (99)	13(99)	81.31
T + R*** with hilum–down orientation	5	0.797	0.303	0.845	0.271	1 (33)	2 (33)	95.45
T + R with hilum-up orientation	5	0.672	0.372	0.761	0.327	4 (33)	3 (33)	89.39
T + R with hilum-parallel-to-ground orientation	5	0.734	0.342	0.778	0.314	3 (33)	2 (33)	92.42
T + R based on all wavelengths and all orientation spectra	6	0.771	0.319	0.745	0.336	12 (99)	12 (99)	87.88
T + R with all orientation spectra omitting wavelengths related to seed orientation	6	0.782	0.313	0.749	0.334	11 (99)	10 (99)	89.39

Notes: *All models were developed based on spectra pretreated with MSC and the Savitzky–Golay second derivative with the number of smoothing points equal to 15.
**r = Correlation coefficient, RMSECV = Root mean square error of cross validation and RMSEP = Root mean square error of prediction.
***Number in parenthesis indicates the total number of seeds for classification.
****T = Transmittance, R = Reflectance.

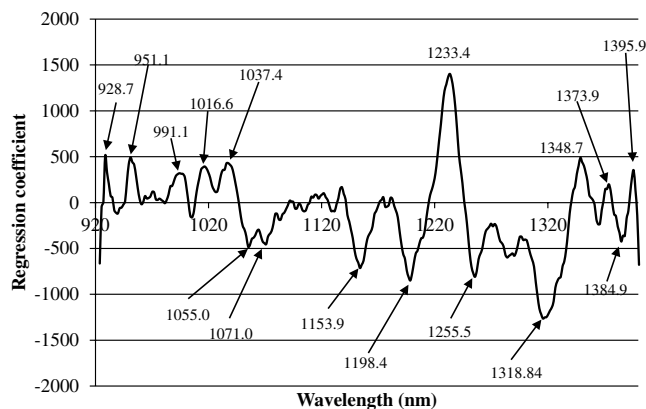Fig. 5. Regression coefficient plot with respect to wavelength of model predicting mung bean seed orientation.

Fig. 7. Regression coefficient PLS-DA model based on both transmittance and reflectance of all orientation spectra omitting wavelengths related to hilum orientation.

These wavelengths were associated with the hilum orientation. Sixteen wavelengths ranging from 928.7 nm to 1395.9 nm were chosen from the regression coefficients. Thus, the PLS-DA model based on all orientation spectra was re-developed with the use of the spectra with the absorbance pertinent to the above selected wavelengths being omitted. The classification accuracy obtained with the model without orientation-related wavelengths was 89.39%, which showed an improvement over that from the previous model (87.88%).

The PLS-DA model obtained eight PLS components, which accounted for 91% of the variance in the spectra for the classification. The scatter plot of the first two PLS scores with 52% of explained variance is displayed in Fig. 6, which shows that most of the samples were separated diagonally into a sound seed cluster in the upper right and a dead
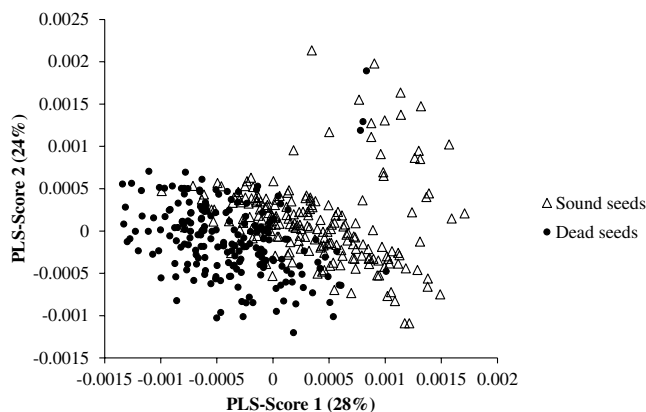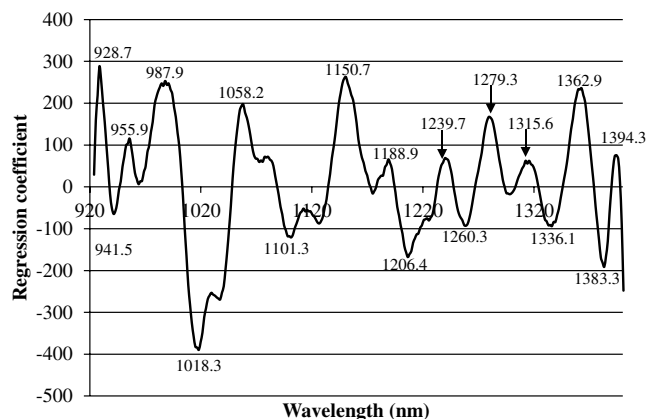
seed cluster in the lower left. Some of the samples were overlapping with respect to the first two scores. Clearer separation would be realized in eight dimensions which could not be visually displayed.

Regarding band assignments, Fig. 7 shows the plot of the regression coefficients, which was obtained from the PLS-DA model based on both the transmittance and reflectance of all orientation spectra (accuracy of 87.88%). The high peaks in the regression coefficients indicate the strong effect of bond vibrations on the classification between the sound and dead seeds. The coefficient peaks and their corresponding bond vibrations and structures are listed in Table 2.[19] Bands associated with the vibration of starch, oil, and protein were observed. Some bands were observed to relatively differ from the bands of chemical assignment. This difference

Fig. 6. PLS score plot of score 1 versus score 2 showing two mostly separated clusters of sound mung bean seeds and dead mung bean seeds.

Table 2. Chemical assignments of near-infrared absorption bands associated with high regression coefficients.

Wavelength at high regression coefficient (nm)	Wavelength of chemical assignment (nm)[19]	Bond vibration[19]	Structure[19]
928.7	928	C–H str. third overtone	oil
987.9	990	O–H str. second overtone	starch
1018.3	1020	N–H str. second overtone	protein
1058.2	1060	N–H str. second overtone	RNH_2
1150.7	1152	C–H str. second overtone	CH_3
1206.4	1215	C–H str. second overtone	CH_2
1362.9	1360	$2 \times$ C–H str. + C–H def.	CH_3
1383.3	1395	$2 \times$ C–H str. + C–H def.	CH_2

arose because the grain is a complicated system that includes several major constituents.[17]

4. Conclusions

We examined the use of NIRS to classify mung bean seeds. The NIRS results indicated the method's potential in the classification of mung bean seeds into sound and dead seeds. The classification model based on a combination of both transmission-absorption spectra and reflection-absorption spectra yielded a better performance than the model based on only transmission-absorption spectra. Furthermore, the orientation of the bean in relation to the light source affected the absorbance. The model using the spectral information of all orientations — hilum face-up, hilum face-down, and hilum-parallel-to-the-ground — yielded a lower classification accuracy than those using the spectral information of only one orientation. Improvement of the performance of the model combining the spectral information of all three orientations was achieved by excluding wavelengths associated with the orientation. We believe that our findings can significantly contribute to better mung bean classification.

Acknowledgments

References

1. T. El-Adawy, E. Rahma, A. El-Bedawey, A. El-Beltagy, "Nutritional potential and functional properties of germinated mung bean, pea and lentil seeds," *Plant Food Hum. Nutr.* **58**, 1–13 (2003).
2. D. Tang, Y. Dong, H. Ren, L. Li, C. He, "A review of phytochemistry, metabolite changes, and medicinal uses of the common food mung bean and its sprouts (Vigna radiata)," *Chem. Cent. J.* **8**(4), 1–9 (2014).
3. P. N. Chiangmai, P. Laosuwan, A. W. Anyuwat, "The effect of mung bean seed size on germination ability, bean sprout production and agronomic characters," *Silpakorn Univ. Int. J.* **6**(1–2), 170–189 (2006).
4. F. M. Rodriguez, E. M. T. Mendoza, "Physico-chemical basis for hardseededness in mung bean (Vigna radiata (L.) Wilczek)," *J. Agric. Food Chem.* **38**(1), 29–32 (1990).
5. G. Downey, J. Boussion, "Authentication of coffee bean variety by near-infrared reflectance spectroscopy of dried extract," *J. Sci. Food Agric.* **71**, 41–49 (1996).
6. S. R. Delwiche, L. O. Pordesimo, A. M. Scaboo, V. R. Pantalone, "Measurement of inorganic phosphorus in soybeans with near-infrared spectroscopy," *J. Agric. Food Chem.* **54**, 6951–6956 (2006).
7. Hacisalihoglu, B. Larbi, A. M. Settles, "Near-infrared reflectance spectroscopy predicts protein, starch, and seed weight in intact seeds of common bean (Phaseolus vulgaris L.)," *J. Agric. Food Chem.* **58**, 702–706 (2010).
8. S. A. Haughey, S. F. Graham, E. Cancouët, C. T. Elliott, "The application of Near-Infrared Reflectance Spectroscopy (NIRS) to detect melamine adulteration of soya bean meal," *J. Food Chem.* **136**, 1557–1561 (2013).
9. S. Kaliramesh, V. Chelladurai, D. S. Jayas, K. Alagusundaram, N. D. G. White, P. G. Fields, "Detection of infestation by Callosobruchus maculatus in mung bean using near-infrared hyperspectral imaging," *J. Stored Prod. Res.* **52**, 107–111 (2013).
10. Y. Y. Lee, J. B. Kim, S. Y. Lee, M. H. Kim, J. W. Lee, H. S. Lee, H.C. Ko, D.Y. Hyun, J. G. Gwag, C. K. Kim, Y. B. Lee, "Determination of seed fatty acids using near-infrared reflectance spectroscopy (NIR) in mung bean (*Vigna radiata*) germplasm," *Korean J. Food Nutr.* **23**(4), 582–587 (2010).
11. L. J. Wang, S. C. Liu, Y. Yao, G. X. Ren, X. Z. Cheng, "The development of near-infrared reflectance spectroscopy (NIRS) prediction model for the quality components of flour and intact seed in mungbean," *J. Plant Genet. Resour.* **14**(5), 833–838 (2013).
12. J. Wang, H. Liu, G. Ren, "Near-infrared spectroscopy (NIRS) evaluation and regional analysis of Chinese faba bean (Vicia faba L.)," *Crop J.* **2**, 28–37 (2014).
13. Y. Dong, S. Y. Qu, "Nondestructive method for analysis of the soybean quality," *Int. J. Food Eng.* **8**(4), 1–6 (2012).
14. M. Tigabu, P. C. Oden, "Simultaneous detection of filled, empty and insect infested seeds of three Larix

species with single seed near-infrared transmittance spectroscopy," *New Forest* **27**, 39–53 (2004).

15. ISTA, International rule for seed testing, *Handbook on Seeding Evaluation*, 3rd Edition, Zurich, Switzerland (2006).

16. K. H. S. Peiris, M. O. Pumphrey, F. E. Dowell, "NIR absorbance characteristics of deoxynivalenol and of sound and *Fusarium*-damaged wheat kernels," *J. Near Infrared Spec.* **17**, 213–221 (2009).

17. P. Williams, Grains and seeds, *Near-Infrared Spectroscopy in Food Science and Technology*,

Y. Ozaki, W. F. McClure, A. A. Christy, Eds., pp. 165–217, John Wiley & Sons, US (2007).

18. J. Workman, L. Weyer, *Practical Guide to Interpretive Near-Infrared Spectroscopy*, CRC Press, Boca Raton, FL, USA (2008).

19. B. Osborne, T. Fearn, P. H. Hindle, *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*, Longman Scientific and Technical, Harlow, UK (1993).