Near Infrared Spectroscopy (NIRS): A tool for Quality Evaluation of Intact Safflower Germplasm

Sangita Yadav, Poonam Suneja, Ranbir Singh, Vandana Joshi and SK Mishra
National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi-110012

A key element of successful development of new cultivars is availability of inexpensive and rapid methods for measurement of fatty acids in seeds. Oilseeds are important sources of vegetable oils. Genetic variation for fatty acid composition is essential for genetic improvement of the oil quality and developing new cultivars. Screening of the large number of germplasm collections requires use of non-destructive analytical technique. About 400 accessions of the safflower were scanned by NIRS as intact seed, and their reference values regressed against different spectral transformations by modiûed partial least squares regression. A calibration set of thirty two accessions (n=32) was analysed by both NIRS and gas-liquid chromatography (GLC) and calibration equations for the major fatty acids were developed. Equation was checked with external validation set of 30 samples. Calibration was focused on the possibility of screening seed samples of different composition of oleic acid (C18:1) and linoleic acid (C18:2) using NIRS analysis.Calibration equation demonstrated a close relationship between NIRS and GLC data for oil (r^2=0.891), palmitic (r^2=0.940), stearic (r^2=0.921), oleic (r^2=0.780) and linoleic acid (r^2=0.939). The results indicated that NIRS could be used to rapidly determine the fatty acid composition in rapeseed seeds in the breeding programmes for high quality rapeseed oil. The results demonstrated that NIRS is a non-destructive, reliable and rapid selection tool for oil quality evaluation. We concluded that a reliable estimation of oil content, palmitic acid, stearic acid, oleic acid and linoleic acid content in intact seeds of safflower is possible by using NIRS technique.

Key Words: Fatty acids, Near Infrared Spectroscopy, Oil content, Safflower

Introduction

Safflower (*Carthamus tinctorius* L.), a member of the family of the Asteraceae, is one of the oldest cultivated crops. Apart from the use as edible oil, safflower is a crop for multiple purposes. Dyes extracted from florets can be used as natural food colour or to tint natural cosmetics. Safflower has very valuable oil for human nutrition, because of high contents of polyunsaturated fatty acids, especially linoleic acid (Honermeier, 2006). Safflower seed oil content varied from 20% to 45% (Belgin *et al.*, 2007). Safflower oil contains about 71% to 75 % linoleic acid and 16% to 20% oleic acid followed by palmitic and stearic acids. Improvement of nutritional and/or functional properties of Safflower oil by modiûcation of safflower fatty acid (FA) composition is a current objective of plant breeders. Regardless of the target, identification and tracking of traits is a major element of plant breeding. Therefore, availability of inexpensive and rapid methods for determination of fatty acid composition of seed samples is a key element of success for development of new grain cultivars.

Research papers published over the last decade demonstrated applicability of NIR spectroscopy for FA profiling in oilseeds. Calibration models for single rapeseed developed by Velasco *et al.* demonstrated comparatively close relationships between GLC measurements and those of NIR spectroscopy for oleic (r^2= 0.85, calibration set size n = 530) and erucic (r^2= 0.88, n = 219) FA. However, no reliable correlation existed for linoleic (r^2= 0.56, n = 530) and linolenic (r^2= 0.53,n = 530) acids. An earlier experiment with bulk rapeseeds conducted by Velasco and Becker resulted in excellent cross validation results for oleic, linoleic, linolenic, and erucic acids (r^2 = 0.95–0.98, n = 220). In contrast, determination coefficients for palmitic, stearic, and eicosenoic acids in bulk rapeseeds were not as high: 0.76, 0.62, and 0.69, respectively (all n= 220).

In safflower, the predictive ability of NIR spectroscopy for FA analysis is not well documented. The objectives of this study were to investigate the applicability of NIR for analysis of FA composition in whole safflower

Materials and Methods

Plant Materials

Four hundred accessions of safflower received from Germplasm Evaluation Division, NBPGR, New Delhi were used to develop a NIRS prediction model for the determination of oil quality parameters.

Total Oil Content of Seed

The safflower seeds were dried to 4-5% moisture level in oven at 108°C for 16 to 18 h. The oil content of the seed samples were determined by a non-destructive

method using a Newport NMR analyser (Model-4000) from Oxford Analytical Instruments Ltd. U.K. after calibrating with pure safflower oil.

Fatty Acid Analysis by Gas Liquid Chromatograph (GLC)

Samples of safflower seeds were freshly grounded (Remi homogenizer) and weighed so that 40 mg oil are obtained when extracted with 10 ml solvent mixture consisting of chloroform: hexane:methanol (8:5:2 v/v/v). The extracts obtained were dried at 60°C in nitrogen gas for 30 min. Methyl esters of oil samples were prepared according to the method of Neff *et al*. (1994) with slight modifications. 1µl of the hexane extract was injected into a highly polar HP Innowax capillary column of 30m length (inner diameter: 0.32m, film thickness: 0.5 µm, split: 1:80). A Hewlett Packard gas chromatograph, model 6890 equipped with flame ionization detector (FID) was used. The injector and detector temperatures were 260°C and 275°C respectively. Oven temperature was programmed from 150°C holding at 1 min. to 210°C at the rate of 15°C/min., followed by 210°C to 250°C at the rate of 5°C/min. for 12 min. Peaks of fatty acid methyl esters were identified by comparing their retention time with that of the known standards, run under similar separation conditions, peak integration was performed applying HP3398A software.

NIR Reflectance Analysis

Collection of Spectra

About 2 g intact seed of safflower were scanned on a Monochromator NIR Systems Model TR-3712-6 (Foss Tecator Near Infrared Analyzer Systems: Silver Springs MD 20904, USA) in reflectance mode. The reflectance spectra (Log 1/R) from 400 to 2500 nm were recorded at 2 nm interval.

Calibration and Validation

Calibration models were developed on the entire dataset (n=400) using cross validation. Further, from the overall set, samples were selected using the algorithms available in the WinISI (version 3.10) (Infrasoft International, Port Matilda PA, USA) software package (Shenkand Westerhaus,1993) where 32 samples were used as the training or calibration set, while the remaining 30 as the external validation set for testing the calibration models.

Calibration was performed by using modified partial least square (MPLS) on the range from 400 to 2500 nm. The log 1/R spectra were transformed into their first and second derivative and Detrend scatter correction (Barnes *et al*, 1989) was computed at the gaps of 4 nm and smoothing over segments of length of 4 nm.

Calibration statistics calculated included standard error of calibration (SEC) and coefûcient of determination (R^2). The equations were then validated using the external set of 30 accessions to assess standard error of performance (corrected for bias) (SEP(C)) and coefûcient of determinations (r^2). The prediction accuracy of the models was tested on the validation set using the root mean square of the standard error of prediction (RMSEP), bias and the coefûcient of correlation (R) (Fearn, 2002, Williams, 2001).The residual predictive deviation (RPD=SD/SECV), deûned as the ratio between the standard deviation of the population (SD) and the SECV for the NIR predictions, is a useful statistic that is often applied to evaluate how well a calibration model can predict chemical data (Fearn, 2002; Williams, 2001).

Results and Discussion

Spectral Analysis

The raw NIR reflectance and second derivative spectra of intact seed samples are shown in Fig. 1. The main absorption bands are observed at 1214 nm related to C-H stretching 2nd overtone (-CH_2), 1496 nm related to C-H stretching 1st overtone, 1724 nm related to CO (oil) and C-H stretching 1st overtone (-CH_2), 1936 nm related to OH 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. Lipid has absorption band around 1360 nm, 1725 nm, 1760 nm, 2140 nm, 2190 nm, 2310 nm, 2323 nm and 2347 nm. Some of these absorption bands can be seen 1356 nm, 1388 nm, 1508 nm, 1720 nm, 2225 nm and 2290 nm due to lipid reflected in chemical structure of the constituent. 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, and sesame, especially, in near-infrared region (Choung *et al*., 2005; Kim *et al*., 2006). The second derivative spectra had a trough corresponding to each peak in the original spectra, removing the overlapping peaks and baseline effects (Osborne *et al.,* 1993). The average spectrum of the second derivative (Fig. 1) showed absorption bands at 1212, 1726, 1760, and 2306 nm related to hydrocarbon (-CH) in the NIR region.

Calibration Models for Fatty Acid Composition

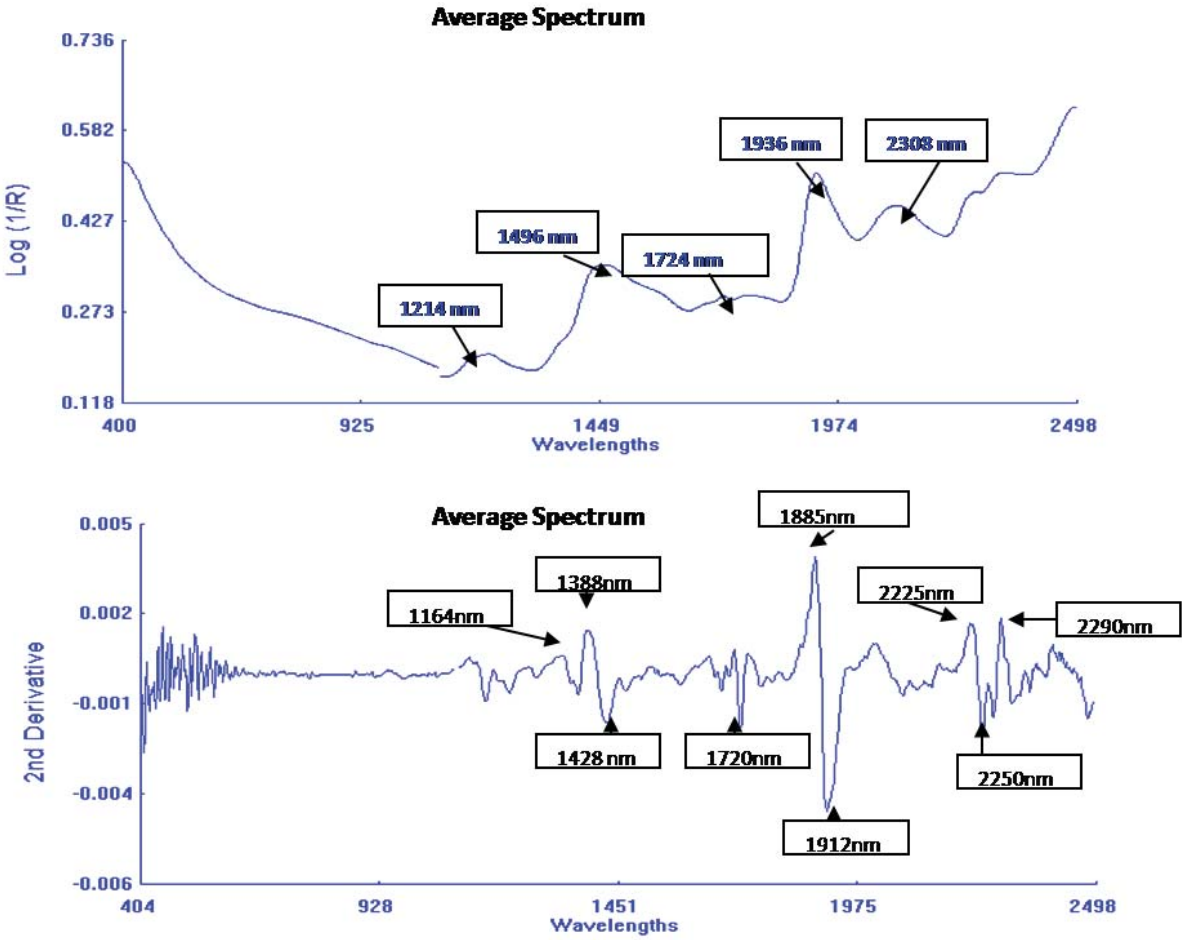

Fig. 1: Raw spectrum (log 1/R; A) and second derivative (B) of NIRS average spectrum of intact seeds of safflower

Since oil quality is a significant concern, particularly for the content of oleic and linoleic acids, which are considered to be good sources for human body. Several workers have reported the oil content and FA profile of safflower oil. The oil content in safflower seed ranged from 20% to 45% (Raie *et al.*, 2006, Belgin *et al.*, 2007). FA composition of oil can be an indicator of its stability, physical properties, and nutritional value. The calibration equations for all the constituents using mathematical treatment 2,4,4,1 were developed. 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 FA composition showing higher values of R^2 (0.960-0.971) indicating and the equations could be used for screening the FA composition in intact seeds of rapeseed. R^2 means the contribution ratio of the calibration equation to the whole vibrations, and if R is higher than 0.7, the calibration equation explains more than 50% of the component vibrations. Table 1 shows the calibration and validation statistic of the best model.

The equation exhibited the r^2 values of 0.891 for oil, 0.940 for palmitic, 0.921 for stearic, 0.780 for oleic, and 0.939 for linoleic acid. The established mean value for oil content is 33.23% with RPD value 3.0. Similarly the established mean value for palmitic, stearic, oleic and linoleic are shown in Table 1. The higher the value of the RPD the greater the probability of the model to accurately predict the chemical composition of samples outside the calibration set. It was reported that an RPD value greater than three (range 3.1–4.9) is considered fair and recommended for screening purposes, an RPD value greater than ûve (range 5–6.4) is considered good for quality control (Fearn, 2002; Williams, 2001). The RPD values were 4.06, 3.58 and 4.02 respectively for stearic acid, palmitic acid and linoleic acid, indicating the good reliability of the calibration equation. Figure

Table 1. Calibration and external validation statistics for intact safflower germplasm for oil and different fatty acid

	Calibration				External-validation			
	Min	Max.	Mean	SEC	R^2	SEP(C)	r^2	RPD
Oil	30.25	36.78	33.52	0.29	0.92	0.45	0.891	3.00
Palmitic acid	3.44	9.04	6.24	1.18	0.96	0.27	0.940	4.06
Stearic acid	0	7.03	3.44	0.23	0.96	0.313	0.921	3.58
Oleic acid	5.37	31.05	18.21	0.72	0.97	1.50	0.780	2.06
Linoleic acid	59.39	80.03	72.21	0.73	0.97	0.75	0.939	4.02

2 represented laboratory reference values against NIRS predicted values in the validation set for individual FA composition (oil, palmitic, stearic, oleic and linoleic), showing also the relationship between NIRS and reference.

The correlation of co-efficient showed highly significant negative correlation between oleic and linoleic acids (Table 2). Correlation coefficient greater than 0.71 or smaller than -0.71 have been suggested to be biologically meaningful (Skinner *et al.*, 1999). In the literature it is reported that oleic acid concentration is controlled by its conversion to linoleic acid, probably as a result of the enzymatic activity of oleic desaturase. A significant negative correlation was observed between oleic with linoleic and oil content, linoleic with stearic acid. Palmitic acid had a significant correlation with oil content and stearic acid.

Most oil and fatty acid equations developed for safflower seed showed sufûcient accuracy for using this technique as a valuable tool for the analysis of these components. The R^2 shown by the equations for oil and fatty acid determination in safflower seeds, indicated from good to excellent quantitative information (Shenk and Westerhaus, 1996). On the other hand, it was observed that the prediction statistics for all the individual

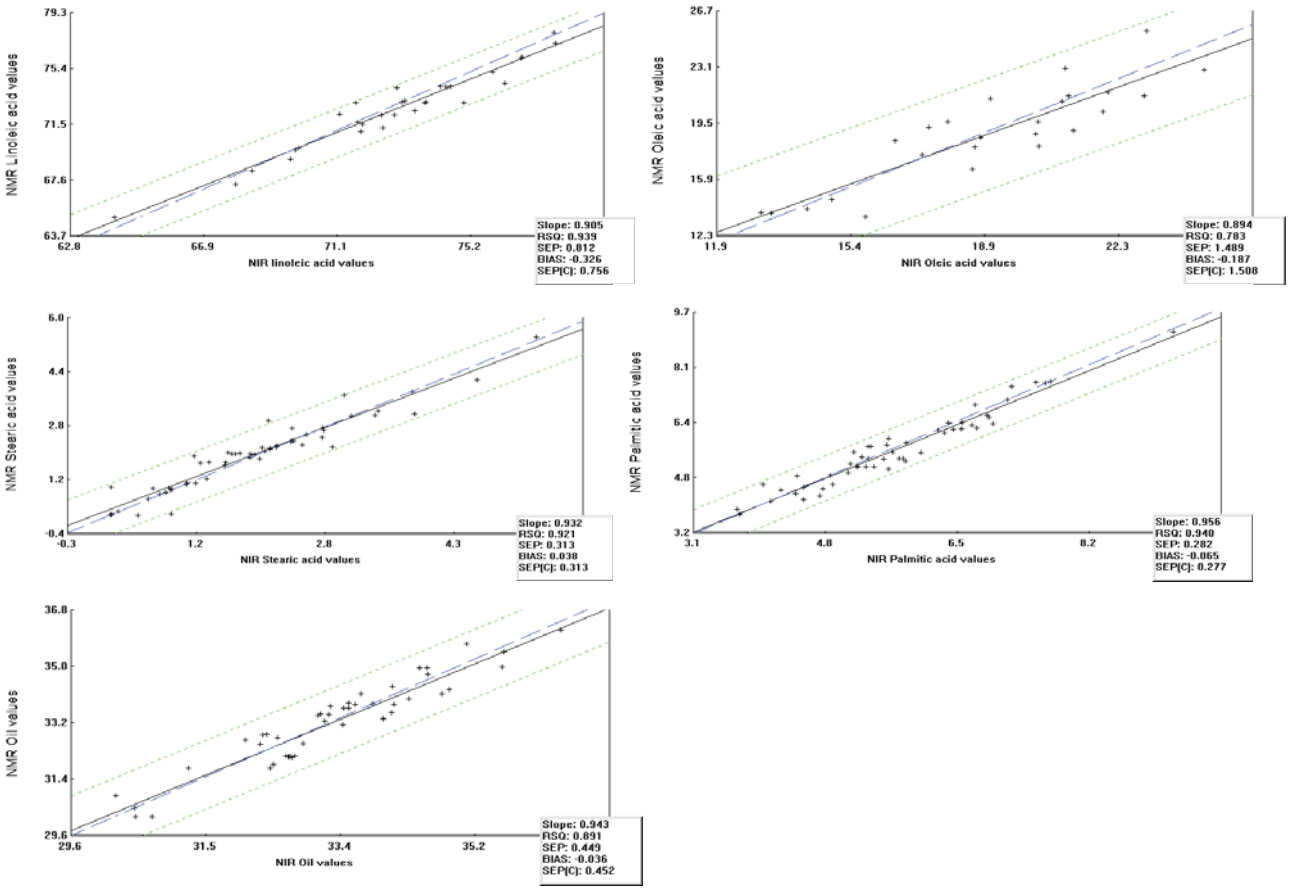

Fig. 2: Scatter plots of NIRS vs. reference values for oil and fatty acids in the cross validation set (n=32) of intact seed samples of safflower

Table 2. Correlation co-efficient between oil content and the fatty acid concentrations in safflower[a]

	oil	palmitic	stearic	oleic	linoleic
oil	1	0.27*	0.082	–0.12**	–0.015
palmitic		1	0.22*	–0.36*	0.025
stearic			1	–0.034	–0.26*
oleic				1	–0.91*
linoleic					1

[a]Correlation shown in bold face type are significant at 1%(*) and at 5%(**) level of significance

constituents are having high values of coefficient of determination (r^2) and the model developed in this study had higher RPD value (>3.0) for most of the constituents, which actually used to evaluate the reliability of calibration models and thus being useful for screening (Williams and Sobering, 1996).

Thus, we can conclude that the NIRS equations developed over safflower germplasm have allowed accurate predictions of multiple quality components in a simultaneous way.

Thousands of seed samples are generated by plant selection and breeding programmes which commonly look to develop progress in several traits simultaneously. Low cost, rapid methods are required to support such programs. Reflectance spectroscopy makes NIRS an ideal tool for screening for quality traits in safflower plant breeding programmes.

References

Barnes RJ, MS Dhanoa and SJ Lister (1989) Standard normal variate transformation and de-trending of Near infrared diffuse reflectance spectra. *Applied Spectroscopy* **43:** 772-777

Belgin C, G Bilal and K Mustafa (2007) Oil content and fatty acid composition if some safflower (*Carthamus tinctorius* L.) varieties sown in spring and winters. *Int. J. Natural Eng. Sci.* **1:** 11-15.

Choung MG, ST Kang, WY Han, IY Baek, HK Kim, KS Kim and SH Park (2005) Determination of fatty acid composition in soybean seed using near infrared reflectance spectroscopy. *Korean J. Breed.* **37:** 197-202

Fearn T (2002) Assessing calibrations: SEP, RPD, RER and R2. *NIR News* **13**: 12-14.

Honermeier B (2006) Saflor In: Heyland KU, H Hanus, ER Keller (eds.). Handbuch des Pflanzenbaues 4. Ölfrüchte, Faserpflanzen, Arzneipflanzen und Sonderkulturen. Verlag Eugen Ulmer, Stuttgart, 246 p.

Kim KS, SH Park and MG Choung (2006) Nondestructive determination of lignans and lignan glycosides in sesame seeds by near infrared reflectance spectroscopy. *J. Agric. Food Chem.* **54:** 4544-4550.

Neff WE, RO Adolf, GR List and M EL-Agaimy (1994) Analysis of vegetable oil triglycerols by silver ion high performance liquid chromatography with flame ionization of the detector. *J. Liquid Chromatography* **17:** 3951-68.

Osborne BG, T Fearn and PH Hindle (1993) Practical NIR spectroscopy with applications in food and beverage analysis, Longman Scientific and Technical, Harlow, UK.

Raie MY and SA Muhammad Khan (2006) Study of safflower seed oils. *Fette, Seifen, Anstrichmittel* **87:** 282-283.

Shenk J and MO Westerhaus (1993). Analysis of agriculture and food products by Near infrared reflectance spectroscopy. Infra soft International. Port Matilda, 116 p.

Velasco L and HC Becker (1998) Estimating the fatty acid composition of the oil in intact-seed rapeseed (*Brassica napus* L.) by Near-Infrared reflectance spectroscopy, *Euphytica* **101:** 221-230.

William PC and DC Sobering (1996) How do we do it: a brief summary of the methods we use in developing near infrared calibration: In Davis AMC and PC Williams (eds.) Near Infrared Spectroscopy; The Future Waves, NIR publications, Chichester, UK. pp 185-188,

Williams PC (2001) Implementation of near-infrared technology. In: Near Infrared Technology in the Agricultural and Food Industries. American Association of Cereal Chemist, St. Paul, Minnesota, USA, pp 145-169.

Windham WR, DR Mertens and FE Barton (1989). Protocol for NIRS calibration: sample selection and equation development and validation. In Marten GC (Ed.) Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality, USDA Agriculture Handbook, Washington DC, USA. 643 p.