# Analyzing Seed Weight, Fatty Acid Composition, Oil, and Protein Contents in *Vernonia galamensis* Germplasm by Near-Infrared Reflectance Spectroscopy

**Tesfaye Baye* and Heiko C. Becker**

Institute of Agronomy and Plant Breeding, University of Goettingen, D-37075 Goettingen, Germany

**ABSTRACT:** *Vernonia galamensis* is a potential new industrial oilseed crop from the Asteraceae family. The interest in this species is due to the presence of a high vernolic acid content of its seed oil, which is useful in the oleochemical industry for paints and coatings. The development of a rapid, precise, robust, nondestructive, and economical method to evaluate quality components is of major interest to growers, processors, and breeders. NIR reflectance spectroscopy (NIRS) is routinely used for the prediction of quality traits in many crops. This study was conducted to establish a rapid analytical method for determining the quality of intact seeds of *V. galamensis.* A total of 114 *Vernonia* accessions were scanned to determine seed weight, FA composition, oil, and protein contents using NIRS. Conventional chemical analysis for FA composition, total oil, and protein contents were performed by GC, Soxhlet extraction, and the Dumas combustion method, respectively. Calibration equations were developed and tested through cross-validation. The coefficient of determination in cross-validation for FA ranged from 0.47 (linoleic acid) to 0.55 (vernolic acid), and for oil, protein, and seed weight from 0.71 (oil) to 0.86 (seed protein). It was concluded that NIRS calibration equations developed for seed weight and seed quality traits can be satisfactorily used as early screening methods in *V. galamensis* breeding programs.

Paper no. J10682 in *JAOCS 81,* 641–645 (July 2004).

**KEY WORDS:** Fatty acid composition, intact seeds, meal protein, near-infrared reflectance spectroscopy, new crop, NIRS, oil content, seed protein, seed weight, *Vernonia galamensis.*

*Vernonia galamensis* (Cass.) Less., an underexploited oil-producing annual plant of the Asteraceae family, is distributed widely in East African countries, in both semiarid zones and areas with abundant rainfall (1). The seed of *V. galamensis* contains 35–40% oil with 72–80% vernolic acid (18:1>O or 12,13-epoxy-*cis*-9-octadecenoic acid), more than 50% of it as trivernolin, a natural epoxy FA. This long-chain unsaturated FA is potentially useful in the oleochemical industry in the formulation of additives, coatings, epoxy resins, lubricants and lubricant additives, sealants, toughened elastomers, and other industrial raw materials (2). The advantages of using industrial oil crops over fossil fuel-derived products include biorenewability, biodegradability, and nontoxicity (3). Research in new crops often starts with unevaluated small seed samples of wild germplasm. Access to vast and diverse germplasm either from a local collection or introduction is essential in the search for desirable traits of interest and in overcoming the barriers of domestication. In addition to the challenge of domesticating new crops, breeders also face the difficulty of meeting minimum oil and meal quality requirements for a largely untried market (4). To reach these goals, effective techniques to screen large numbers of germplasm for various quality parameters are necessary.

The conventional analytical techniques for evaluating germplasm accessions for quality traits are frequently limited by the time-consuming, tedious, destructive nature of the technique and the small amount of seeds available. For example, most official Soxhlet oil extraction methods require 5–10 g of seeds, about 100 mL of solvent (*n*-hexane), and take 8–10 h to complete. The solvents/reagents are considered to be toxic and environmentally hazardous waste products (5). Determination of 1000-seed weight by counting and weighing large sample numbers is also time consuming and costly.

NIR spectroscopy (NIRS) appears to be an attractive alternative for early screening of quality traits (6). Currently, NIRS is used for a wide range of analytical applications in various agricultural, petrochemical, pharmaceutical, clinical, biomedical, and environmental sectors and in process control analysis (7–10). Studies on several species including *Brassica* (11,12), sunflower (13*)*, rice (14), and *Camelina* (15) have demonstrated that NIRS may provide a simple, reliable, and economical way to estimate seed weight, seed color, moisture content, FA and amino acid compositions, oil, and protein contents nondestructively and concurrently from only one spectrum and from single seeds or small samples of intact seed. The present study was aimed at evaluating the potential of NIRS to estimate seed weight, FA composition, oil, and protein contents in *V. galamensis* germplasm accessions. A detailed account of *Vernonia* taxonomy, biology, ecological distribution, and potential uses can be found elsewhere (1,16).

## MATERIALS AND METHODS

*Seed materials*. In 2000, *V. galamensis* seed samples were collected across the natural growing regions of Ethiopia and evaluated in field experiments during the 2001/2002 main crop season at Alemaya University of Agriculture, Ethiopia.

*To whom correspondence should be addressed at University of Florida, P.O. Box 110690, Gainesville, FL 32611-0690. E-mail: tbaye@ifas.ufl.edu

Copyright © 2004 by AOCS Press

The harvested seeds from 114 accessions were analyzed for seed quality traits at the Institute of Agronomy and Plant Breeding, University of Goettingen, Germany.

*Seed weight (SW).* From each sample 1000 seeds were counted and weighed on a laboratory balance model M2P (Sartorius, Goettingen, Germany).

*FA.* FA were determined by standard GC of the FAME following the procedure of Thies (17) and analyzed on a PerkinElmer gas chromatograph model 8600 (PerkinElmer Corp., Norwalk, CT) equipped with a fused-silica capillary column FFAP, 25 m × 0.25 mm × 0.25 μm film thickness (Macherey Nagel GmbH + Co. KG, Düren, Germany). The oven, detector, and injector temperatures were 200, 250, and 250°C, respectively. The carrier gas was hydrogen, at a pressure of 70 kPa. Samples (2 μL) were injected at a split ratio of 1:70. Individual FA including palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2), and vernolic (18:1>O) acids were expressed as percentage of the total FA, including minor FA.

*Oil content.* The seed oil content of milled seed samples was determined for about 5–10 g by the Soxhlet oil extraction technique with *n*-hexane for about 10 h.

*Seed and meal protein contents.* Nitrogen concentration was determined for both seed and meal protein from about 200 mg flour in two replicates by the Dumas combustion method using an automated CN analyzer (Heraeus CN-Rapid, Hanau, Germany). Total nitrogen was converted to percent protein using a factor of 6.25.

*NIRS scanning and spectral analysis.* About 3–5 g cleaned and intact seed from each of the 114 accessions of *Vernonia* was placed in a small ring cup (i.d. 4.7 cm) and scanned with NIRSystems model 6500 monochromator instrument (NIRSystems, Inc., Silver Spring, MD). Reflectance spectra (log 1/*R*) were collected between 1100 and 2500 nm. The data obtained from measurement of three replications were averaged and used for calibration and validation. NIRS spectra were matched with reference data and analyzed using modified partial least squares (MPLS) regression, math treatment 2, 5, 5, 1, and WinISI IIv. 1.04 (Infrasoft International, Port Matilda, PA) software. Calibration performance was assessed by SEC (SE of calibration), $r^2$ (coefficient of determination), and SECV (SE of cross-validation) following Shenk and Westerhaus (18).

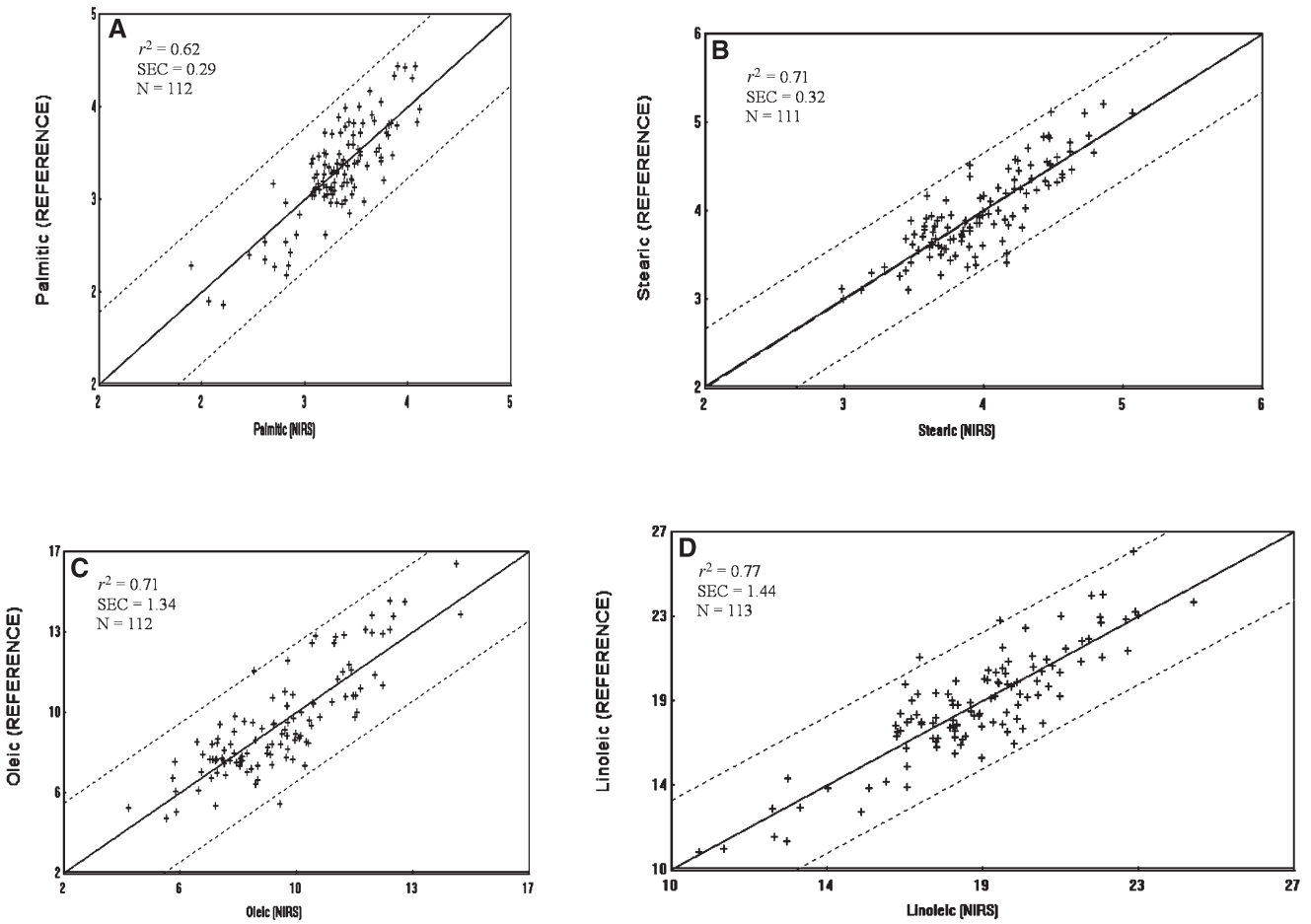


**FIG. 1.** Reference vs. NIR reflectance spectroscopy (NIRS)-predicted values of calibration plot for FA in intact seed samples of *Vernonia galamensis*. *N*, number of samples; SEC, SE of calibration; $r^2$, coefficient of determination.

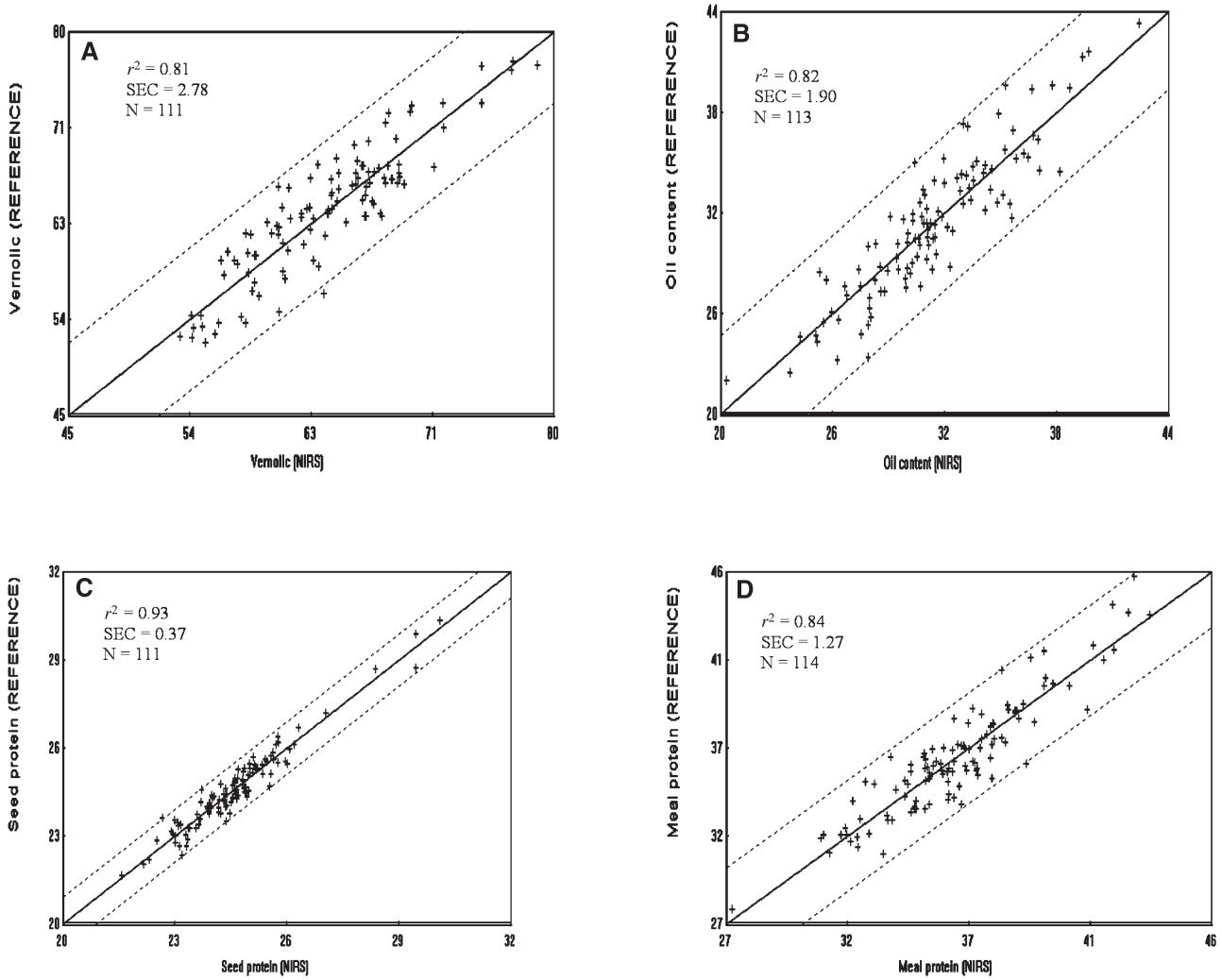


**FIG. 2.** Reference vs. NIRS-predicted values of calibration plot for vernolic acid, oil, and protein content in intact seed samples of *V. galamensis*. For abbreviations see Figure 1.

## RESULTS AND DISCUSSION

The 114 samples varied for oil, seed protein, and meal protein content from 22 to 45%, 21 to 30%, and 28 to 46%, respectively (data not shown). Variations for the various FA and seed weight were also observed. There were significant differences among the 114 samples for all the traits, which suggest the presence of sufficient variation among the samples to develop NIRS equations.

Table 1 shows the calibration statistics for the various traits considered. The calibration equations for seed weight, FA, oil, and protein contents showed relatively low SEC and SECV. A higher SEC value was recorded for vernolic acid, which may be due to the broader range of variation for this trait than others. The scatter plots of predicted vs. measured values for each trait are presented in Figures 1 to 3. For instance, the reference values for protein content were widely spread and showed a strong relationship ($r^2$ = 0.93) with the NIRS values (Fig. 2C). The NIRS value for oil content gives a reasonable ability to differentiate among high, intermediate, and low oil content (Fig 2B).

The FA 16:0, which is a minor constituent owing to its low concentration, tended to have lower $r^2$ values as a result of

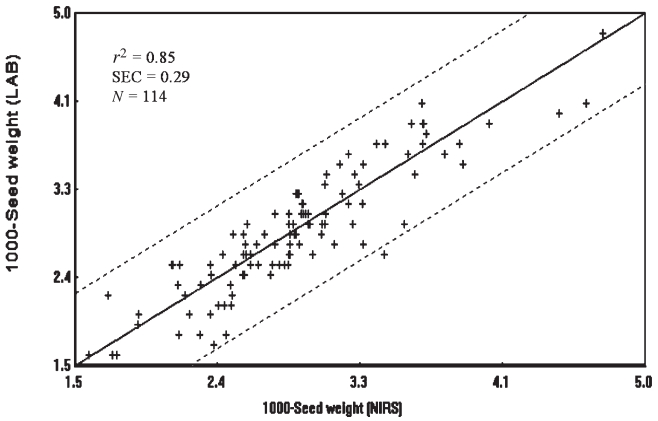


**FIG. 3.** Reference vs. NIRS-predicted values of calibration plot for 1000-seed weight in seed samples of *V. galamensis*. For abbreviations see Figure 1.

**TABLE 1**
**Calibration and Cross Validation Statistics[a] in NIRS Equations for the Analysis of FA, Oil, and Protein Contents (%), and Seed Weight (g) in Intact Seed Samples of *Vernonia galamensis***

| | Calibration | | | | | | Cross-validation | |
|---|---|---|---|---|---|---|---|---|
| Trait | *N* | Mean | Range | SD | $r^2$ | SEC | 1 – VR | SECV |
| 16:0 | 112 | 3.6 | 2.2–5.1 | 0.5 | 0.62 | 0.29 | 0.51 | 0.32 |
| 18:0 | 111 | 3.8 | 2.1–5.7 | 0.6 | 0.71 | 0.32 | 0.52 | 0.40 |
| 18:1 | 112 | 8.9 | 1.6–16.2 | 2.4 | 0.71 | 1.35 | 0.48 | 1.75 |
| 18:2 | 113 | 18.2 | 9.6–26.8 | 2.9 | 0.77 | 1.43 | 0.47 | 2.12 |
| 18:1 >O | 111 | 63.4 | 45.2–81.6 | 6.1 | 0.81 | 2.74 | 0.55 | 4.04 |
| Oil | 113 | 31.5 | 18.5–44.4 | 4.3 | 0.82 | 1.90 | 0.71 | 2.37 |
| SP | 111 | 24.6 | 20.4–31.5 | 1.4 | 0.93 | 0.37 | 0.86 | 0.52 |
| MP | 114 | 35.9 | 26.7–45.2 | 3.1 | 0.84 | 1.27 | 0.77 | 1.50 |
| SW | 114 | 2.8 | 0.6–5.1 | 0.73 | 0.85 | 0.29 | 0.78 | 0.35 |

[a]NIRS, NIR reflectance spectroscopy; *N*, number of samples; $r^2$, coefficient of determination of calibration; SEC, SE of calibration; SECV, SE of cross-validation; 1 – VR, coefficient of determination of cross-validation; 16:0, palmitic acid; 18:0, stearic acid; 18:1, oleic acid; 18:2, linoleic acid; 18:1>O, epoxy FA; SP, seed protein; MP, meal protein; SW, 1000-seed weight.

restricted variance or relatively lower range and SD. It also exhibited a low SEC value. The addition of new samples with lower and higher 16:0 concentration may improve the calibration statistics for this trait. A lower $r^2$ for palmitic acid compared with stearic, oleic, linoleic, and linolenic acids was also reported previously (19). Thus, the best performance in calibration equations for individual FA corresponded to those FA for which the variability in the calibration set was wider (Table 1), indicating that successful calibration equations using NIRS depend on the variability of constituents under investigation.

Prediction of FA seems less reliable than prediction of oil and protein contents. The large similarity of the spectra for linoleic and vernolic acids—because vernolic acid (18:1 epoxy) is synthesized from linoleic acid (18:2) and/or the unusual epoxy group of vernolic acid is attached to 18:1 molecules—might have influenced the accuracy of the prediction for other FA. Similar problems of detecting the epoxy group, because linoleate gives a signal at approximately the same position as epoxy, were reported by Hopkins and Bernstein (20). A second explanation could be that the chemical information is obscured by changes in the spectra caused by differences in seed size and shape. Seeds of *Vernonia*, also referred to as cypsela, are elliptic in shape, 5–7 mm long and 1–2 mm broad (at the middle). Pazdernik *et al*. (19) reported that calibration equations based on ground soybean samples were more accurate than those based on intact seed samples, which might be explained by diversity of the surface area and size of the seeds used. Thus, NIRS is useful mainly for preselecting accessions for FA composition; confirmation by GC analysis is still required. NIRS as an early screening technique is extremely useful for predicting oil content, as this is the most difficult trait in terms of time and sample size requirements in conventional chemical analysis. However, NIRS is known to be sensitive to changes in genotypes, growing locations, and seasons, which necessitate frequent recalibration. Future research should aim to improve the accuracy of the prediction by adding a broader range of samples to the calibration set for each of the traits from different genotypes, regions, and years.

## ACKNOWLEDGMENTS

We thank Thomas zum Felde for his valuable help in NIRS and the German Academic Exchange Service (DAAD) for a scholarship to the first author.

## REFERENCES

1. Gilbert, M.G., Notes on Eastern Africa Vernonieae (Compositae). IV. A Revision of the *Vernonia galamensis* Complex, *Kew Bull. 41*:19–35 (1986).
2. Ayorinde, F.O., B.D. Butler, and M.T. Clayton, *Vernonia galamensis*: A Rich Source of Epoxy Acid, *J. Am. Oil Chem. Soc. 67*:844–845 (1990).
3. Janick, J., New Crops for the 21st Century, in *Crop Science: Progress and Prospects*, edited by J. Nosberger, H.H. Geiger, and P.C. Struik, CABI Publishing, Wallingford, Oxfordshire, United Kingdom, 2001, pp. 307–327.
4. Thompson, A.E., Breeding New Industrial Crops, in *Advances in New Crops*, edited by J. Janick and J.E. Simon, Timber Press, Portland, OR, 1990, pp. 100–103.
5. Panford, J.A., and J.M. deMan, Determination of Oil Content of Seeds by NIR: Influence of Fatty Acid Composition on Wavelength Selection, *J. Am. Oil Chem. Soc. 67*:473–482 (1990).
6. Starr, C., A.G. Morgan, and B.D. Smith, An Evaluation of Near Infrared Reflectance Analysis in Some Plant Breeding Programs, *J. Agric. Sci. 97*:107–118 (1981).
7. Shenk, J.S., J.J. Workman, and M.O. Westerhaus, Application of NIR Spectroscopy to Agricultural Products, in *Handbook of Near-Infrared Analysis*, edited by D.A. Burns and E.W. Ciurczak, Marcel Dekker, New York, 1992, pp. 383–431.
8. Allen, B.I., The History and Background of NIR, *Cereal Chem. 4*:238–240 (1983).
9. Krishnan, P.G., W.J. Park, K.D. Kephart, D.L. Reeves, and G.L. Yarrow, Measurement of Protein and Oil Content of Oat Cultivars Using Near-Infrared Reflectance Spectroscopy, *Cereal Foods World 39*:105–108 (1994).
10. Blanco, M., and I. Villarroya, NIR Spectroscopy: A Rapid-Response Analytical Tool, *Trends Anal. Chem. 21*:240–250 (2002).
11. Velasco, L., F.D. Goffman, and H.C. Becker, Development of Calibration Equations to Predict Oil Content and Fatty Acid Composition in Brassicaceae Germplasm by Near-Infrared Reflectance Spectroscopy, *J. Am. Oil Chem. Soc. 76*:25–30 (1999).
12. Van Deynze, A.E., and K.P. Pauls, Seed Colour Assessments in *Brassica napus* Using a Near-Infrared Reflectance Spectrometer

Adapted for Visible Light Measurements, *Euphytica 76*:45–51 (1994).
13. Sato, T., Y. Takahata, T. Noda, T. Tanagisawa, T. Morishita, and S. Sakai, Nondestructive Determination of Fatty Acid Composition of Husked Sunflower *(Helianthus annua* L.) Seeds by Near-Infrared Spectroscopy, *J. Am. Oil Chem. Soc*. *72*:1177–1183 (1995).
14. Wu, J.G., C. Shi, and X. Zhang, Estimating the Amino Acid Composition in Milled Rice by Near-Infrared Reflectance Spectroscopy, *Field Crops Res. 75*:1–7 (2002).
15. Vollmann, J., A. Damboeck, S.J.H. Kuyt, and P. Ruckenbauer, Determination of *Camelina* Seed Weight Using Near-Infrared Reflectance Spectroscopy, *Plant Varieties Seeds 10*:95–101 (1997).
16. Baye, T., H. Kebede, and K. Belete, Agronomic Evaluation of *Vernonia galamensis* Germplasm Collected from Eastern Ethiopia, *Ind. Crops Prod*. *14*:179–190 (2001).
17. Thies, W., Schnelle und einfache Analysen der Fettsaeurezusammensetzung in einzelnen Raps-Kotyledonen. I. Gaschromatographische und papierchromatographische Methoden, *Z. Pflanzenzüchtg. 65*:181–202 (1971).
18. Shenk, J.S., and M.O. Westerhaus, *Analysis of Agriculture and Food Products by Near Infrared Spectroscopy*, Infrasoft International (ISI), Port Matilda, PA, 1993, pp. 63–66.
19. Pazdernik, D.L., A.S. Killam, and J.H. Orf, Analysis of Amino Acid and Fatty Acid Composition in Soybean Seed, Using Near-Infrared Reflectance Spectroscopy, *Agron. J*. *89*:679–685 (1997).
20. Hopkins, C.Y., and H.J. Bernstein, Applications of Proton Magnetic Resonance Spectra in Fatty Acid Chemistry, *Can. J. Chem*. *37*:775–782 (1959).

[Received June 23, 2003; accepted June 21, 2004]