Pak. J. Bot., 44: 219-222, Special Issue March 2012.

STUDIES ON GENETIC DIVERSITY FOR SEED QUALITY IN RAPESEED (*BRASSICA NAPUS* L.) GERMPLASM OF PAKISTAN THROUGH NEAR INFRARED SPECTROSCOPY AND PRINCIPAL COMPONENT ANALYSIS

IFTIKHAR ALI[*], SYED ANWAR SHAH, HAFIZ MUNIR AHMED, KIFAYAT-UR-REHMAN AND MUMTAZ AHMAD

Nuclear Institute for Food and Agriculture (NIFA), P.O. Box 446, Peshawar, Pakistan
[*]*Corresponding author: e-mail: iali63@yahoo.com; Tel: +92912964060; Fax: +92912964059*

Abstract

To optimize the exploitation of plant genetic resources for improvement of oilseed crops in Pakistan, seed quality characterization of germplasm of rapeseed is being executed at Nuclear Institute for Food and Agriculture (NIFA), Peshawar. The aim of this study was to determine the feasibility of using Near Infrared Spectroscopy (NIRS) to identify seed quality traits of oilseed rape germplasm. Seventeen hundred and fifty three (1753) accessions were collected from different sources across Pakistan. Whole seeds of each collection were scanned (1100-2500 nm) on NIR system 6500 in reflectance mode by spinning module. The NIRS spectrum can easily identify different quality characteristics of rapeseed germplasm by applying principle component analysis (PCA). Principle Component (PC) 1 showed better results in separating germplasm for protein, glucosinolate, oleic acid, and erucic acid. PC2 and PC3 identified oil content and linolenic acid respectively. The PCA score plots were clear to identify differences in seed quality of indigenous and exotic germpalsm. It can be concluded from this study that the NIRS and PCA techniques can be used to identify genetic diversity for quality traits of oilseed brassica germplasm.

Introduction

The Brasscicaceae family consists of approximately 375 genera and 3200 species of plants and brassica genus consists of approximately 100 species, including *Brassica napus* L., spp. *oleifera*, commonly known as oilseed rape or rapeseed or canola. *Brassica napus* is not native to Pakistan, and originated in either the Mediterranean area or Northern Europe. It is thought to have originated from a cross where the maternal donor was closely related to two diploid species, *Brassica oleracea* and *Brassica rapa* (Gupta & Partab, 2007).

The screening large germplasm collections for biochemical and nutritional traits are normally limited by the small number of seeds available. Although analytical techniques such as gas-liquid chromatography of fatty acid methyl esters allow determination of oil content and its fatty acid composition in very small samples, the use of a representative sample requires the destruction of a high number of seeds (Garces & Mancha, 1993; Agnihotri & Kumar, 2004). In Brassica species, recent studies have demonstrated that near-infrared reflectance spectroscopy (NIRS) may provide a reliable estimation of the fatty acid composition of the seed oil of intact seeds, with the main advantages that the analysis is nondestructive, simultaneous with the analysis of other seed constituents such as oil, protein, and glucosinolate content (GSL), and suitable for small samples (Velasco & Becker, 1998; Mika *et al*., 2003). The characteristics of NIRS technique make it optimal for the nondestructive evaluation of oil content and composition in germplasm accessions of cruciferous seeds. However, it would require the development of calibration equations integrating large taxonomic variation, since the development of specific equations for single species or even for single genera would be impracticable (Shenk & Westerhaus, 1993; Velasco *et al*., 1999).

The objective of this study was to evaluate the potential of NIRS to estimate the oil content and its fatty acid composition with GSL in a germplasm collection of brassicaceae through the development of calibration equations integrating large taxonomic variability. To optimize the exploitation of plant genetic resources for improvement of oilseed crops in Pakistan, seed quality characterization of germplasm of rapeseed crop is being executed at Nuclear Institute for Food and Agriculture (NIFA), Peshawar.

Materials and Methods

Sample collection: The study was carried out with 1753 collections belonging to *Brassica napus* of the family brasssicaceae. Samples, in replications, of different accessions / lines / cultivars / varieties of oilseed crops collected from different R&D Institutions, growers and open markets countrywide.

NIR spectra acquisition: The samples were scanned on a monochromator (NIR Systems, model 6500) equipped with sample autochanger. The standard ring cup, which requires a seed volume of about 5g, was used. For each samples the reflectance spectrum (log 1/R) from 400 to 2500 nm was recorded at 2-nm intervals. Calibration and validation procedures were carried out with ISI software, version 1a.1 (Infrasoft International) as described by Anon., (1998).

Statistical analysis: The collected data were analyzed by numerical taxonomic techniques using the procedures of principal component analysis (Sneath & Sokal, 1973). The standardized and transformed data matrix was prepared using software STATISTICA (2004). The large dimensions of generated data were reduced to only three leading principal components (PC1, PC2, and PC3) which have extracted major share of variation in the data of three selected traits for rapeseed genetic variability. The PCA was used to compute variance in data and to determine the most variability-influencing trait. The standard deviation and mean of the eigenvector of principal components were employed to study the different constituents of the available rapeseed germplasm in Pakistan.

Results and Discussion

The NIR spectrum of rapeseed germplasm was captured in visible (400-700 nm) and near infrared (701-2500 nm) range (Fig. 1). In the visible range seed texture and seed coat color was observed. Absorption bands in the NIR region were observed at around 1200, (O-H stretch second overtone, associated with water, C-H first overtone, associated with lipids and oil), 1580 (C-H combination tones, associated with amino acids) 1700 nm (C-H first overtone, associated with fatty acids and oil) 1950 nm (C-H, N-H sulfur compounds and amino acids). The 2300 nm band acts as a reference point whose wavelength does not change across the spectra of a set of similar samples as described by Murray (1986) and it is true in our case as all analyzed germplasm were belonged to same species (*Brassica nuapus* L.) The standard deviation of the NIR spectra showed that wavelengths associated with water (variable moisture content) and oil are the most important in explaining the variation among the samples.

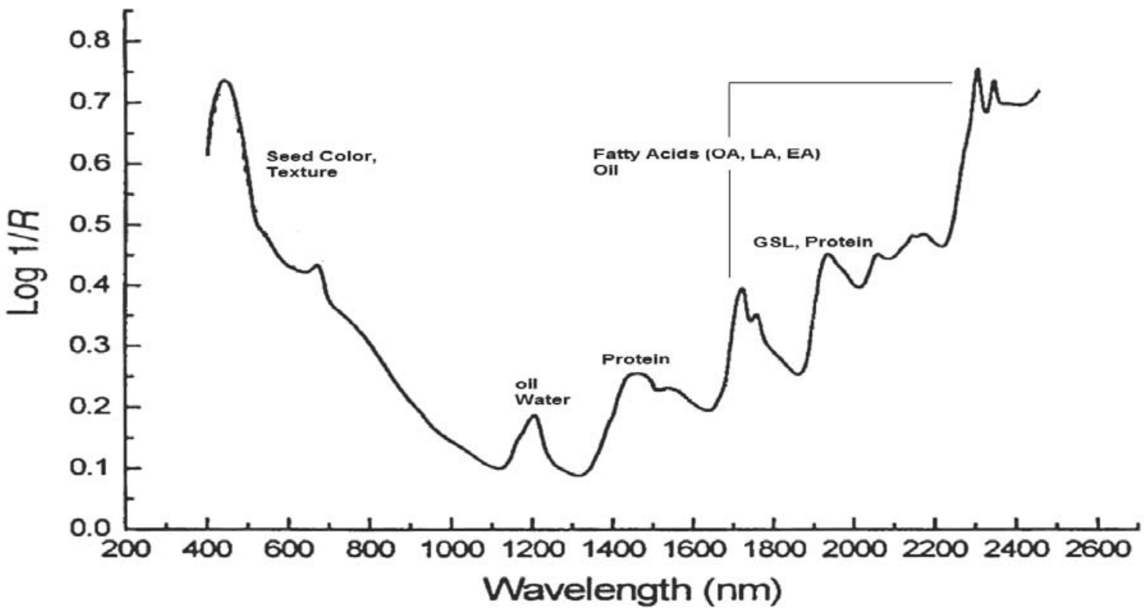

Fig. 1. Mean near-infrared spectrum of rapeseed (*Brassica napus* L.) germplasm.

Principal component analysis (PCA) is one method for identifying patterns in the data of NIR spectrum (Munck & Moller, 2005). Given these trends in the data, PCA was performed on raw absorbance spectra of the rapeseed germplasm and the eigenvectors for first three PCs plotted in Fig. 2. A total variation of 91.28% among the rapeseed germpalsm was captured by the first three PCs. The highest variance of 49.51% was observed for PC-1 and high influence on fatty acids including oleic acid (OA) and erucic acid (EA) along with GSL was identified. The EA and GSL both showed negative values and expressed effective low concentrations in the local germpalsm. The OA had a positive response, which confirmed desirable magnitude of this highly valuable monounsaturated fatty acid concentration in the germplasm. The PC-2 explained 26.42% variance and had influence on the oil and protein contents of the samples (Fig. 2). The highly unsaturated fatty acid LA related to PC-3 and this PC had 15.32% variance in the total data variability. The ability of the NIR-based model to classify the rapeseed germplasm is based on the vibrational responses of chemical bonds (O-H, N-H and C-H) in the near infrared sector. It is probable that greater variability among the germplasm for those chemical entities (e.g., oil content or protein or fatty acids) that respond in these sectors of the spectrum leads to better accuracy of the model (Murray, 1986; Miller, 2001).

In NIR spectral matrix, the length of the traits vectors, which are represented in graph (Fig. 3) by lines connecting the origin and individual trait coordinates is an approximation of the traits standard deviation, as a result can be interpreted as a measure of a trait's ability to discriminate among the tested germplasm (Sneath & Sokal, 1973; Martens & Naes, 1991). The cosine of the angle by two trait vectors is an indicator of how correlated overall germplasm response for these two traits. The traits with vectors forming right angle are not correlated, those forming acute angles are positively correlated, and those forming obtuse are negatively correlated.

The most important trait of the oilseed crop is the seed oil percentage and genotype with high seed oil content is extremely desirable. The available rapeseed germplasm in Pakistan has practically substantial seed oil concentration and diversity (Figs. 3&4). The most important two quality traits, in case of rapeseed are EA and GSL (canola or double zero quality). These two quality traits were closely correlated (Fig. 3) and had significant variability in indigenous germplasm (Fig. 4). The GSL content had no correlation but EA had small negative correlation with oil content (Fig. 3). The highly desirable monounsaturated omega fatty acid 'OA' is clearly negatively correlated with oil presence (Fig. 3). There is significant genetic diversity was identified in the analyzed rapeseed germplasm (Fig. 3). Our results indicated limited variability and diversity for LA trait in indigenous rapeseed germplasm as were reported by other researchers (Rakow, 1973; Rucker & Robbelen, 1997; Velasco *et al*., 1997; McVetty & Scarth, 2002; Scarth & Tang, 2006; Werteker *et al*., 2010; Akhtar *et al*., 2011).

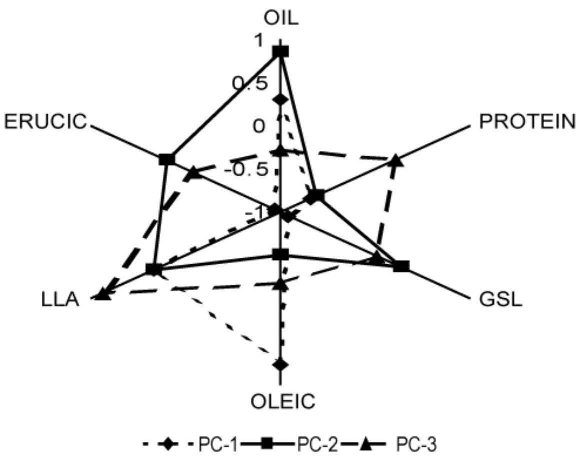

Fig. 2. Eigenvectors of the first three principal components of the NIR spectrum of rapeseed germplasm.

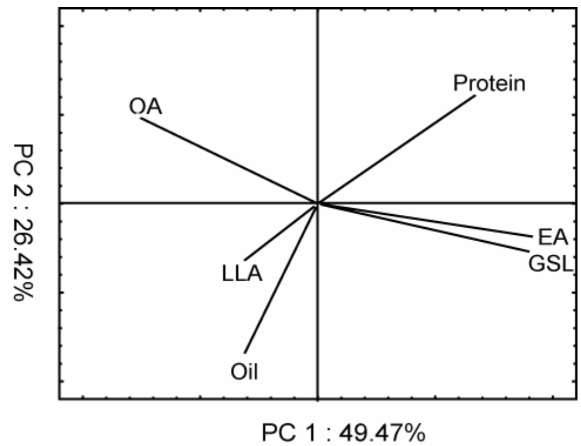

Fig. 3. Score plot of first two principal components of the NIR spectrum of rapeseed germplasm.

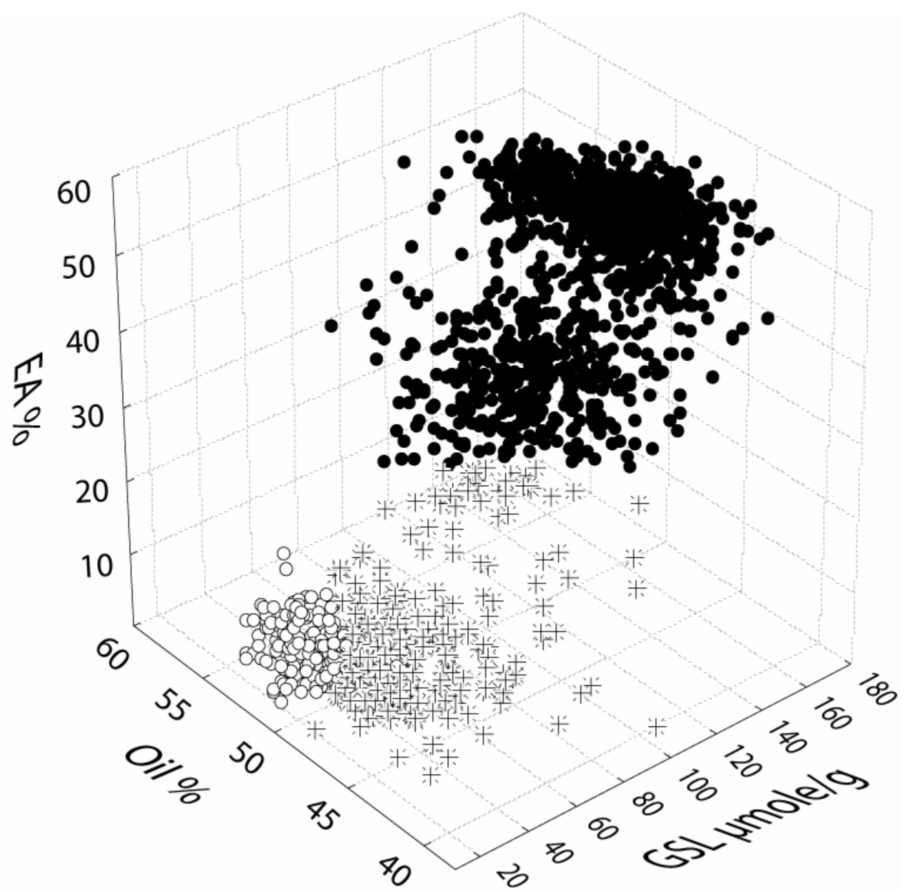

Fig. 4. 3D plot of oil, GSL and erucic acid content of rapeseed germplasm.

Analysis of genetic diversity through NIR spectroscopy and PCA in germplasm collections can facilitate reliable classification of accessions and identification of subsets of core accessions with possible utility for specific breeding purposes. Significant emphasis has been given under this work to comprehensive analysis of genetic diversity in oilseed brassica crop. Our studies validated the meaningful explanation of discrimination among samples based on selected characteristics (e.g., oil, fatty acids, protein, GSL and fatty acids). As suggested by other authors, NIR spectroscopy is a powerful technique to identify the seed contamination (e.g., mixing of seeds of different cultivars/species) and fraudulent practices or for screening of the breeding material. NIR spectroscopic screening can be easily implemented for the oilseed crop quality identification and quantification. This technique may be most useful for initial screening at early stages in the oilseed variety development breeding chain, enabling more costly chemistry methods to be used selectively for further examination of the valuable products. To be able to develop local adaptive oilseed crop varieties with better oil quality, it is important to measure and analyze the chemical composition of the breeding material.

Acknowledgments

The authors acknowledge PAEC and ALP for providing the technical and financial support. Thanks are extended to all donor Institutions who supplied seed samples for study.

References

Agnihotri, A. and Kumar, S. 2004. Value addition in rapeseed-mustard: Status and prospects. In: Rapeseed-Mustard research in India. (Eds.): A. Kumar and N.B. Singh. pp. 212-231. National Research Centre of Rapeseed and Mustard. ICAR, Sewar, Bharatpur.

Akhtar, L.H., M.A. Pervez and M. Nasim. 2011. Genetic divergence and inter-relationship studies in chickpea (*Cicer arietinum* L.). *Pak. J. Agri. Sci*., 48: 35-39.

Anonymous. 1998. NIR calibration in practice. Available: http://dx.doi.org/1-:1051/analusis:199826040038.

Anonymous. 2004. STATISTICA. StatSoft, Inc. (Data analysis software system), Version 7. www.statsoft.com.

Garces, R. and M. Mancha. 1993 One-step lipid extraction and fatty acid methyl esters preparation from fresh plant tissues. *Anal. Bioch*., 211: 139-143.

Gupta, S.K. and A. Partab. 2007. History, origin, and evolution. In: Rapeseed Breeding. (Ed.): S.K. Gupta. Advances in Botanical Research. Amsterdam: Elsevier Science, V. 45 1-20.

Martens, H. and T. Naes. 1991. Multivariate Calibration – John Wiley and Sons.

McVetty, P.B.E. and R. Scarth. 2002. Breeding for improved oil quality in Brassica oilseed species. *J. Crop Prod*., 5: 345-369.

Mika, V., P. Tillmann, R. Koprana, P. Nerusil and V. Kucera. 2003. Fast prediction of quality parameters in whole seeds of oilseed rape (*Brassica napus* L.). *Palnt and Soil Environment*, 49(4): 141-145.

Miller, Ch. E. 2001. Chemical principles of near infrared technology. Pages 19-39. In: *Near infrared technology in the agricultural and food industries*. (Eds.): P.C. Williams and K.H. Norris. Second Edition. American Association of Cereal Chemist, Minnesota, USA.

Munck, L. and B. Moller. 2005. Principal Component Analysis of Near Infrared Spectra as a Tool of Endosperm Mutant Characterisation and in Barley Breeding for Quality. *Czech J. Genet. Plant Breed*., 41(3): 89-95.

Murray, I. 1986. The NIR spectra of homologous series of organic compounds. pp 13-28. In: *Proceedings of the International NIR/NIT Conference*. (Eds.): J. Hollo, K.J. Kaffka and J.L. Gonczy. Akademiai Kiado, Budapest, Hungary.

Rakow, G. 1973. Selection for low linolenic rapeseed through mutation breeding. *Plant Breeding*, 69: 62-82.

Rucker, B. and G. Robbelen. 1997. Mutants of *Brassica napus* with altered seed lipid fatty acid composition. p. 316-318. In: *Proc. 12th Int. Symp. Plant Lipids, Toronto, Canada*, 8–12 July 1996, Kluwer Academic Publishers, Dordrecht, Netherlands.

Scarth, R. and J. Tang. 2006. Modification of Brassica oil using conventional and transgenic approaches. *Crop Sci*., 46: 1225-1236.

Shenk, J. and M.O. Westerhaus. 1993. Analysis of agriculture and Food products by near infrared reflectance spectroscopy. pp. 41-51. Infrasoft International, Port Matilda, USA.

Sneath, P.H.A. and R.R. Sokal. 1973. Numerical Taxonomy: The Principles and Practice of Numerical Classification. W.F. Freeman & Co., San Francisco, 573 p.

Velasco, L. and H. 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.

Velasco, L., Begon˜a Pe´ rez-Vich and Jose´ M. Ferna´ndez-Martı´nez. 1999. Nondestructive Screening for oleic and linoleic acid in single sunflower achenes by Near-Infrared Reflectance Spectroscopy. *Crop Sci*. 39: 219-222.

Velasco, L., J.M. Ferna´ndez-Martı´nez and A. de Haro. 1997a. Induce Variability for C18 unsaturated fatty acids in Ethiopian mustard. *Can. J. Plant Sci*., 77: 91-95.

Werteker, M., A. Lorenz, H. Johannes, E. Berghofer and C.S. Findlay 2010. Environmental and varietal influences on the fatty acid composition of rapeseed, soybeans and sunflowers. *J. Agronomy & Crop Science*, 196: 20-27.

.

(Received for publication 12 February 2011)