USE OF NEAR INFRARED REFLECTANCE SPECTROSCOPY TO EVALUATE QUALITY CHARACTERISTICS IN WHOLE-WHEAT GRAIN

Uso de la espectroscopía de reflectancia en el infrarrojo cercano para evaluar características de calidad en trigo

Daniel Cozzolino[1] *, Inés Delucchi[1], Moham Kholi[2] , Daniel Vázquez[1]

A B S T R A C T

The aim of this work was to explore the potential of visible (Vis) and near infrared reflectance (NIR) spectroscopy to measure quality characteristics in whole grain wheat (*Triticum aestivum* L.) as a tool in breeding programs. A total of 100 samples were analyzed by the reference methods for crude protein (CP), wet gluten (WG) and sodium dodecyl sulfate (SDS) sedimentation test. Whole grain samples were scanned in a NIR monochromator instrument (400-2500 nm) in reflectance. Partial least squares (PLS) were used to develop calibration equations for the quality characteristics in whole wheat. Calibration models were validated using an independent set of samples (n = 50) randomly selected from the population set. The uncertainty of the PLS models was evaluated by the standard error of prediction (SEP). The SEP obtained were 0.35% for CP, 2.04 for SDS and 4.14% for WG. It was concluded that NIR spectroscopy might be used as a screening tool to segregate early generations of wheat genotypes. At a later stage is needed to improve the accuracy of the NIR calibrations, broadening the calibration spectra with the incorporation of more genotypes and different crop years.

Key words: whole wheat, NIR, protein, wet gluten, grain quality, SDS, *Triticum aestivum*.

R E S U M E N

El objetivo de este trabajo fue explorar el potencial de la espectroscopía en el visible (Vis) e infrarrojo cercano (NIR) para medir características de calidad en el trigo (*Triticum aestivum* L.) para su uso en programas de mejoramiento. Cien muestras fueron analizadas por el método de referencia para proteína cruda (CP), gluten húmedo (WG) y sulfato de dodecil de sodio (SDS) o prueba de sedimentación. Las muestras de trigo se analizaron en un instrumento NIR (400-2500 nm) en reflectancia. El método de los cuadrados mínimos parciales (PLS) fue utilizado para desarrollar las ecuaciones de calibración para las características de calidad en trigo. Los modelos de calibración se validaron utilizando un conjunto independiente de muestras (n = 50) aleatoriamente escogido del conjunto de la población. La incertidumbre de los modelos PLS de calibración fue evaluada usando el error estándar de la predicción (SEP). El SEP obtenido fue de 0,35% para CP, 2,04 para SDS y 4,14% para WG. Se concluyó que la espectroscopía de NIR podría utilizarse como una herramienta de selección para segregar genotipos de trigo en generaciones tempranas. En una etapa posterior se necesita mejorar la precisión de los análisis NIR, ampliando el espectro de calibración con la incorporación de más genotipos y diferentes años de cultivo.

Palabras clave: grano de trigo, NIR, proteína, gluten húmedo, calidad de grano, SDS, *Triticum aestivum*.

[1] Instituto Nacional de Investigación Agropecuaria, Estación Experimental INIA La Estanzuela, Ruta 50 Km 12, Colonia, Uruguay.
Actual address: The Australian Wine Research Institute, Waite Campus, Urrbrae, PO Box 197. Glen Osmond, South Australia, 5064.
E-mail: Daniel.Cozzolino@awri.com.au *Corresponding author.
[2] International Center for Wheat and Maize Improvement (CIMMYT).
Recibido: 17 de octubre de 2005 Aprobado: 30 de marzo de 2006.

INTRODUCTION

Near infrared reflectance (NIR) spectroscopy has been used in plant selection in forage breeding programmes, for commodity trading, soil analysis and segregation of materials in cereal breeding programs (Deaville and Flinn, 2000). In all these applications, NIR spectroscopy has a role in increasing the number of samples analyzed, reducing cost and time taken, and the amount of sample required for testing (Murray, 1993). The NIR method relies on calibrations, which utilise absorbances at several, even many wavelengths, to predict composition of a sample (Murray, 1993). Calibration is a mathematical process, which uses multivariate regression techniques (e.g., partial least squares: PLS) relating NIR optical measurements (absorbance values) at selected wavelengths (1100 to 2500 nm) to reference values measured by conventional chemical or physical methods (Deaville and Flinn, 2000).

Food and agricultural products analysed by NIR spectroscopy capitalize on the less intense absorbance of energy in the region of the electromagnetic spectrum in the wavelength range between 750 and 2500 nm (Murray, 1993). Molecular bonds which dominate the NIR region containing H attached to atoms such as N, O and C, provide information about the relative proportions of C-H, N-H and O-H bonds, which are constituents of the organic matrix. Today, the grain from cereal crops is segregated on the basis of crude protein (CP) and moisture content using NIR instruments based on either reflectance or transmittance (Dardenne *et al*., 1994; Pawlinsky and Williams, 1998; Delwiche, 1998 and 2004).

Wheat (*Triticum aestivum* L.) is the most consumed of all cereal grains, in terms of range of products for human and animal consumption that can be derived from it (Pawlinsky and Williams, 1998). Wheat functionality its defined by the ability of wheat to produce the final product (baked, extruded) for which it is purchased. Therefore, the prediction of wheat functionality is an essential step of breeding programs to produce improved wheat varieties (Pawlinsky and Williams, 1998).

Grain protein is of primary importance in determining the bread making quality of wheat flour. Although total protein is the primary factor in determining the end use of wheat, there is often a need to measure properties that are indicative of the protein quality (Delwiche, 2004). Early applications of NIR to predict functionality parameters in wheat were based on the use of ground grain (Williams *et al*., 1988; Pawlinsky and Williams, 1998; Delwiche, 1998). It is well documented that NIR spectroscopy is used routinely in the grain industry to measure the CP content in whole wheat grain, however, few studies report the application to measure other quality properties of wheat using the whole grain.

The aim of this work was to explore the potential of visible (Vis) and near infrared reflectance (NIR) spectroscopy to measure quality characteristics in whole grain wheat as a tool in breeding programs.

MATERIALS AND METHODS

Wheat samples and reference methods

A total of 100 whole-wheat samples were selected from a wide range of genotype materials grown at the Instituto Nacional de Investigación Agropecuaria (INIA), La Estanzuela Experimental Research Station, Uruguay (2000-2001). The wheat genotypes were part of a wide range of materials used by the breeding program of the International Center for Wheat and Maize Improvement (CIMMYT) based at INIA La Estanzuela, Colonia, Uruguay. Uruguay (33 °00' S and 56 °00' W) is located entirely within the temperate zone with average high and low temperatures in summer (January) of 28 °C and 17 °C, and in winter (July) 14 °C and 6 °C, respectively. Rainfall is fairly evenly distributed throughout the year with mean annual precipitation of 950 mm.

Crude protein (CP) and wet gluten (WG) were determined according to the approved American Association of Cereal Chemist reference methods (AACC, 1993). Gluten strength was measured by the sodium dodecyl sulfate (SDS) sedimentation test (Peña *et al*., 1990). Crude protein was corrected to 12% moisture using the dry mass weight (AACC, 1993).

Visible and near infrared analysis

Whole-wheat spectra were collected using a NIR scanning spectrophotometer (NIRSystems 6500, NIRSystems, Silver Spring, Maryland, USA) in reflectance mode (400–2500 nm). Whole kernel samples were placed in the sample transport module in a rectangular (1/4) quartz cup (NIRSystems, part

number 0IH–0379). Samples were scanned in duplicate and the average spectrum was used to develop the multivariate models. Reflectance data were stored as log (1/R) (R = reflectance) at 2 nm intervals (1050 data points). Two pairs of lead sulphide detectors collected the reflectance spectra in the NIR region (1100-2500 nm). Reflectance energy readings were referenced to corresponding readings from an internal ceramic disk provided by the instrument manufacturer. The spectrum of each sample was the average of 32 successive scans. Computer operation and spectral data collection were manipulated using ISI DOS version 3.1 software (InfraSoft International, Port Matilda, Pennsylvania, USA).

Multivariate analysis

Spectral data were exported from the ISI software for chemometric analysis to the Unscrambler software (version 7.5, CAMO ASA, Oslo, Norway). Principal component analysis (PCA) was performed before partial least squares (PLS) models were developed. PCA is a mathematical procedure for resolving sets of data into orthogonal components, whose linear combinations approximate the original data to any desired degree of accuracy (Naes *et al*., 2002). In PCA, the information in the data is projected down to a small number of latent variables called principal components (PCs). PCA was used to derive the first 20 principal components from the spectral data. These were used in further analysis to examine the relevant and interpretable structure in the data as well as outlier detection. PCA was used to select samples for allocation in either the calibration or validation set, as well as to detect outlier samples. The PLS models were developed between the reference method and the NIR data for the different parameters evaluated.

The optimum number of terms in the PLS calibration models were determined by cross validation and defined by the PRESS (predicted residual error sum of squares) function in order to avoid over-fitting of the models. Cross validation estimates the prediction error by splitting the calibration samples into groups (four in this case). One group was reserved for validation and the remaining groups were used for calibration. The process was repeated until all groups had been used for validation once. During calibration development, raw and second derivative were used. Spectra were pre-processed using the second derivative to reduce baseline variation and enhance the spectral features. Analysis of the second derivative was performed using Savitzky-Golay derivation and smoothing (10 point and second order filtering operation) (Savitzky and Golay, 1964). Calibration statistics calculated included the coefficient of determination in calibration (R^2_{CAL}), and the standard error of cross validation (SECV). The ratio of standard deviation and standard error in cross validation (RPD) were used to test the accuracy of the calibration models developed (Williams, 2001). An RPD value higher than three is considered adequate for analytical purposes in most NIR applications for agricultural products. A validation set (n = 50) was used to test the NIR calibration models developed. The correlation coefficient (R), the standard error of prediction, the slope and bias were calculated and used to evaluate the accuracy of the calibrations.

RESULTS AND DISCUSSION

The mean, range and standard deviation (SD) of quality characteristics measured in the wheat samples used into calibration and validation sets are summarised in Table 1. A wide variation in quality characteristics was observed related with the different genotypes analyzed in the present study. High and positive correlations (Pearson, $p < 0.05$) were found between CP and WG ($r = 0.71$), intermediate

Table 1. Mean, range and standard deviation (SD) of crude protein (CP), sedimentation test (SDS) and wet gluten (WG) for whole wheat samples used in calibration and validation.

Cuadro 1. Media, rango y desviación estándar (SD) para proteína cruda (CP), test de sedimentación (SDS) y gluten húmedo (WG) de las muestras de grano de trigo usadas en calibración y validación.

	Calibration (n = 50)			Validation (n = 50)		
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**
CP, %	12.8	1.2	9.6-15.2	12.9	1.3	9.0-15.2
SDS, %	12.4	2.9	7.5-21.5	12.6	2.85	8-22.0
WG, %	31.7	4.9	19.9-43.9	31.9	4.85	19-41.7

CP: crude protein; SDS: sodium dodecyl sulfate sedimentation test; WG: wet gluten

between CP and SDS (r = 0.60) and low between WG and SDS (r = 0.12). It was observed that both wet and dry gluten increased with increasing protein content. These results agreed with those reported by other authors (Bechere *et al*., 2002; Lerner *et al*., 2004). However, contradictory reports were found in the literature (e.g., low correlation CP and SDS).

A PCA analysis was performed on the quality characteristics measured by the reference method in order to observe variation among the quality characteristics in the wheat samples. Two principal components (PCs) explained 99% of the variation in the set of samples. WG was the variable that explains 73% of the variation among samples in the PC1. Variation among samples along the PC2 (26%) is related with SDS and CP content.

The highest SD in the mean spectrum was observed at wavelength regions related with water content at 980, 1430 and 1930 nm related with O-H overtones (Murray, 1986). High SD were also observed around 1700 nm, associated with C-H first overtones (oil content), and around 2200-2300 nm associated with C-H combinations tones (Murray, 1986).

Table 2 shows the Vis and NIR PLS calibration models for quality parameters in whole wheat using raw and after applying the second derivative as a mathematical treatment. It was observed that a high correlation between the NIR spectral data and the reference values were obtained for CP. The predicted CP values were highly correlated to the laboratory reference values with low SECV (0.54-0.65%). Both WG (SECV = 4.2-4.5%) and SDS (SECV = 1.8-2.2) gave intermediate calibration models. The SECV obtained for the prediction of SDS and WG content in the different PLS models developed correspond to an average error of 14% for WG and of 17% for SDS relative to the average (Table 1).

The RPDs obtained in this study showed that intermediate accuracy was obtained for the measurement of CP content in the whole wheat samples by the NIR method ($2 < RPD < 3$), and low accuracy for SDS and WG ($RPD \leq 2$). An RPD lower than 2 will be considered acceptable for rough screening. It is well known that WG is related positively with CP content in wheat, therefore it is likely that wavelength regions associated with N-H and O-H absorptions might contribute in order to develop a calibration for both WG and SDS. The regression coefficients associated with the N-H region did not have a big effect in the calibration models (Figure 1).

Table 2. Partial least squares (PLS) calibration statistics for quality characteristics measured in whole wheat.

Cuadro 2. Estadísticas de calibración obtenidas mediante cuadrados medios parciales (PLS) en grano de trigo.

	R^2_{CAL}	SECV	T	RPD
Vis+NIR §				
CP, %	0.92	0.63	10	2.0
SDS, %	0.70	1.8	7	1.6
WG, %	0.53	4.3	5	1.1
NIR §				
CP, %	0.92	0.65	9	1.8
SDS, %	0.81	2.0	10	1.4
WG, %	0.53	4.5	6	1.1
Vis+NIR &				
CP,%	0.94	0.59	9	2.0
SDS, %	0.67	2.2	2	1.3
WG,%	0.79	4.4	4	1.1
NIR &				
CP, %	0.94	0.54	4	2.2
SDS, %	0.79	2.1	5	1.3
WG, %	0.70	4.2	3	1.2

CP: crude protein; SDS: sodium dodecyl sulfate sedimentation test; WG: wet gluten; Vis+NIR: visible and near infrared; NIR: near infrared; CP: crude protein; R^2_{CAL}: coefficient of determination in calibration; SECV: standard error in cross validation; RPD; SD/SECV; T: number of PLS terms used to develop the calibration models, § raw spectra, & second derivative.

Table 3 shows the NIR validation statistics for the quality characteristics measured in whole wheat. The SECV and SEP values obtained are unacceptably large to be used for analytical purposes. For the quality characteristics measured, the error obtained by the NIR method is considered high for laboratory use. However, such models might be used to evaluate the SDS and WG in early generations in a breeding program. Therefore, a high, medium and low scale might be developed in order to select the most promising wheat genotypes. The use and application of such scales need to be validated with field experiments. The SDS sedimentation is measured in some laboratories by electrophoresis technique, that is and indication of the strength of the gluten. During breeding, it is advantageous to select for SDS in early generations because lines chosen for high SDS values maintain good characteristics when grown in different locations and years.

Table 3. Validation statistics for quality characteristics measured in whole-wheat samples (n = 50).
Cuadro 3. Estadísticas de validación obtenidas para los parámeteros de calidad medidos en grano de trigo (n = 50).

	R	SEP	Bias	Slope
NIR [&]				
CP, %	0.95	0.35	0.12	0.93
SDS, %	0.60	2.04	0.15	0.43
WG, %	0.50	4.14	-1.09	0.41

CP: crude protein; SDS: sodium dodecyl sulfate sedimentation test; WG: wet gluten; R: correlation coefficient; SEP: standard error of prediction, [&] second derivative.

Protein content, on the other hand, is highly influenced by the environment. However, the method demonstrated that it will be quick, easy and practical enough to provide reasonable prediction of SDS and WG in wheat minimizing sample presentation.

It is well known that the NIR method requires little sample preparation and is extremely rapid when compared to the reference method. Predicting whole grain wheat quality in a small sample size by NIR spectroscopy might be used to discard early generation breeding cultivars that do not possess the desired characteristics. This will assist breeders in their efforts to combine high grain protein content and other quality characteristics in wheat breeding programs for evaluation of new varieties with good quality characteristics.

The results from this study suggested those quality characteristics, such as CP, SDS and WG content, that might be predicted in whole wheat samples using NIR spectroscopy. In view of these results, it might be expected that increased accuracy would result, if calibrations were developed for different years and genotypes. Such application offers the advantage of either reducing or eliminating the rather large effort that is now required for each NIR laboratory to assemble a calibration set and perform conventional chemical analysis in order to develop calibration equations for individual samples.

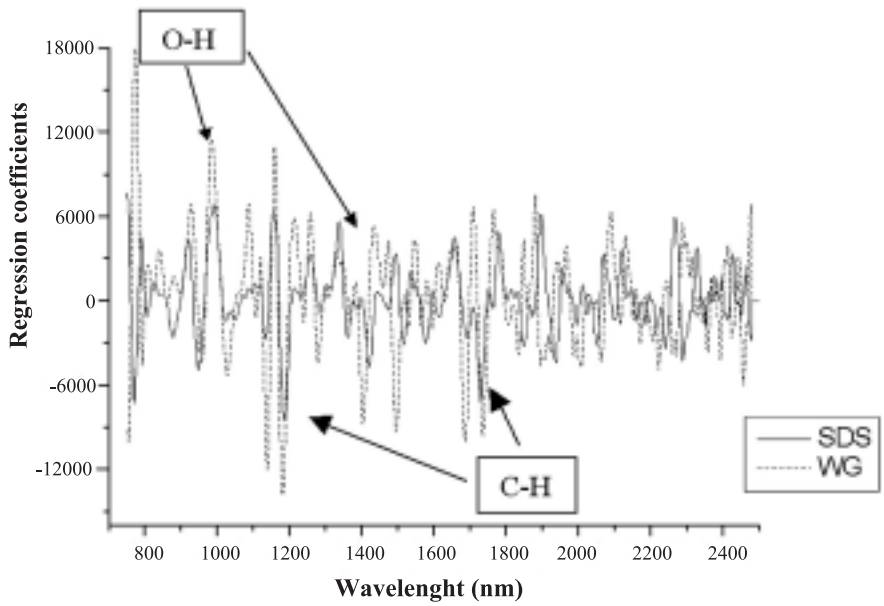

Figure 1. Partial least square regression coefficients for sedimentation test (SDS) and wet gluten (WG) in whole grain.
O-H: oxygen-hydrogen bond (moisture): C-H: carbon-hydrogen bond (carbohydrates/lipids); SDS: sodium dodecyl sulfate sedimentation test; WG: wet gluten
Figura 1. Coeficientes de regresión de los cuadrados medios parciales para test de sedimentación (SDS) y gluten húmedo (WG) en grano de trigo.
O-H: enlace oxígeno-hidrógeno (humedad); C-H: enlace carbono-hidrógeno (carbohidratos y/o lípidos); SDS: test de sedimentación; WG: gluten húmedo

CONCLUSIONS

The calibration and validation statistics obtained in this work showed the potential of NIR spectroscopy to predict quality characteristics in whole-wheat grain. The accuracy for the PLS models developed for SDS and WG were considered low for routine analysis, however NIR spectroscopy might be used as a screening tool to segregate early generations of wheat genotypes. NIR spectroscopy combined with multivariate techniques (PLS) might be very helpful for breeding programs where there is a need to analyze many samples in a short time and no sample destruction is required. If precise quality results are required, the NIR method will not be the best method to use. Further work is needed in order to improve the accuracy of the NIR analysis including more genotypes and years.

CITED LITERATURE

AACC. 1993. Approved methods of the American Association of Cereal Chemists. 1100 p. 7th ed. American Association of Cereal Chemist (AACC), St. Paul, Minnesota, USA.

Bechere, E., R.J. Peña, and D. Mitiku. 2002. Glutenin composition, quality characteristics and agronomic attributes of durum wheat cultivars released in Ethiopia. Afr. Crop Sci. J. 10:173-182.

Dardenne, P., G. Sinnaeve, L. Bollen, and R. Biston. 1994. Reduction of wet chemistry for NIR calibrations. p. 154 - 160. *In* G.D. Batten, P.C. Flinn, L.A. Welsh, and A.B. Blakeney (eds.). Leaping ahead with near infrared spectroscopy. NIR Spectroscopy Group, Royal Australian Chemical Institute, Melbourne, Victoria, Australia.

Deaville, E.R., and P.C. Flinn. 2000. Near infrared (NIR) spectroscopy: an alternative approach for the estimation of forage quality and voluntary intake. p. 301-320. *In* D.I. Givens, E. Owen; R.F.E. Axford, and H.M. Omed (eds.). Forage evaluation in ruminant nutrition. CABI Publishing, Wallingford, Oxon, UK.

Delwiche, S.R. 1998. Protein content of single kernels of wheat by near infrared reflectance spectroscopy. J. Cereal Sci. 27:241-254.

Delwiche, S.R. 2004. Analysis of small grains. p. 269-345. *In* C.A. Roberts, J. Workman, J.B. Reeves (eds.). Near infrared spectroscopy in agriculture. ASA, CSSA, SSSA, Madison, Wisconsin, USA.

Lerner, S.E., M. Cogliatti, N.R. Ponzio, M.L. Seghezzo, E.R. Molfese, and W.J. Rogers. 2004. Genetic variation for grain protein components and industrial quality of durum wheat cultivars sown in Argentina. J. Cereal Sci. 40:161-166.

Murray, I. 1993. Forage analysis by near infrared spectroscopy. p. 285-312. *In* Davies, A., Baker, R.D., Grant, S.A., and Laidlaw, A.S. (eds.) Sward measurement handbook. 2nd ed. The British Grassland Society, Reading, UK.

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

Naes, T., T. Isaksson, T. Fearn, and T. Davies. 2002. A user-friendly guide to multivariate calibration and classification. 344 p. NIR Publications, Chichester, UK.

Pawlinsky, T. and P.C. Williams. 1998. Prediction of wheat bread baking functionality in whole kernels using near infrared reflectance spectros. J. Near Infrared Spectroscopy 6:121-127.

Peña, R.J., A. Amaya, S. Rajaram, and A. Mujeeb-Kazi. 1990. Variation in quality characteristics associated with some spring 1B/1R translocation wheats. J. Cereal Sci. 12:105-112.

Savitzky, A., and M.J.E. Golay. 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36:1627-1632.

Willliams, P.C., J. El-Haramein, G. Ortiz-Ferreira, and J.P. Srivatsva. 1988. Preliminary observations on the determination of wheat strength by near infrared spectroscopy. Cereal Chem. 65:109-114.

Williams, P.C. 2001. Implementation of near-infrared technology. p. 145-169. *In* Williams, P.C., and K.H. Norris (eds.). American Association of Cereal Chemist, St. Paul, Minnesota, USA.