Near infrared reflectance spectroscopy as a new screening tool to increase durum wheat yield

Ferrio J.P., Araus J.L., Bertran E., Nachit M., Royo C.

in

Royo C. (ed.), Nachit M. (ed.), Di Fonzo N. (ed.), Araus J.L. (ed.).
Durum wheat improvement in the Mediterranean region: New challenges

Zaragoza : CIHEAM
Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 40

2000
pages 141-144

Near infrared reflectance spectroscopy as a new screening tool to increase durum wheat yield

J.P. Ferrio*, E. Bertran, M.M. Nachit***, C. Royo**** and J.L. Araus***
*Unitat de Fisiologia Vegetal, Universitat de Barcelona, Avgda. Diagonal 645, 08028 Barcelona, Spain
**Unitat de Química Analítica, Universitat Autònoma de Barcelona, Spain
***Germplasm Program, CIMMYT/ICARDA, P.O. Box 5466, Aleppo, Syria
****Area de Conreus Extensius, Centre UdL-IRTA, Alcalde Rovira Roure 177, 25198 Lleida, Spain

SUMMARY – Near Infrared Reflectance Spectroscopy (NIRS) is widely used in routine screening as it provides a quick, non-destructive assay requiring little or no sample pretreatment. Here, we present a new application of NIRS as an indirect selection tool during the early generations of a breeding program addressed to increase durum wheat yield under Mediterranean conditions. To assess the ability of NIRS in yield evaluation, spectral reflectance was measured in kernels from a large set of genotypes, cultivated in three different environments in NW-Syria, providing a wide range of yield (from about 600 to 9500 kg/ha). Partial Least Squares Regression (PLSR) was used to model the association between grain yield (GY) and these spectra, and strong significant correlations ($P < 0.001$) were found between NIRS-predicted and measured yield, with r^2 values of 0.88 and 0.76 for calibration and prediction sets, respectively.

Key words: NIRS, yield, durum wheat, PLSR, regression.

RESUME – *"Le NIRS comme nouvel outil de criblage pour augmenter le rendement du blé dur". La Near Infrared Reflectance Spectroscopy (NIRS) est une technique très utilisée pour faire des analyses routinières puisqu'elle permet un essai rapide et non destructif sans pré-traitement des échantillons. Ici nous présentons une nouvelle application de la NIRS comme outil de sélection pendant les premières générations d'un programme d'amélioration dirigé à augmenter la production de blé dur en milieux méditerranéens. Afin d'évaluer la capacité d'estimer la production moyennant NIRS nous avons mesuré la réflectance spectrale en graines d'un grand nombre de génotypes, cultivés dans trois milieux différents du Nord-Ouest de la Syrie, tout en obtenant une grande marge de production (entre 600 et 9500 kg/ha aprox). Nous avons utilisé la Partial Least Squares Regression (PLSR) pour modéliser l'association entre la production (GY) et ces spectres, trouvant des corrélations très significatives ($P < 0{,}001$) entre les valeurs prédites par NIRS et celles qui ont été mésurées, avec une r^2 de 0,88 et 0,76 pour les sets de calibration et prédiction respectivement.*

Mots-clés : *NIRS, production, blé dur, PLSR, régression.*

Introduction

Increasing yield performance under Mediterranean conditions is a target in durum wheat breeding programs. However, these programs require extensive field testing, and direct grain yield (GY) determinations are usually only attained in late generations of the breeding program, when the number of genotypes is low and there are enough kernels for yield plots. Therefore, as the production of high-yielding genotypes by direct selection (i.e. based on the yield itself) requires progressively larger and more costly procedures, the use of indirect approaches has been postulated. Indeed, certain indirect selection criteria could be used in early stages to improve the genetic gains of any breeding program (Austin, 1993; Araus, 1996; Araus *et al.,* 1998). These criteria should be easy to measure, preferably non-destructive and applicable to a large number of plants in a relatively short time. In this context, the use of Near Infrared Reflectance Spectroscopy (NIRS) in mature kernels as an indirect selection tool for higher yield in durum wheat grown in Mediterranean conditions may be a good alternative. In laboratory settings, NIRS is currently the basis for quick, accurate, non-destructive and highly repeatable assays of many biological traits and compounds related with the organic (also inorganic, i.e. ash content) composition of plant matter (see references in Finney *et al.,* 1987; Shenk *et al.,* 1992). The ability of NIRS to determine these various compounds in plant samples is due to vibrational and rotational energies associated with H bonds (Osborne and Fearn, 1986). The quantitative and qualitative composition of proteins, carbohydrates, lipids, and minerals in grain cereals is affected by growing conditions, which also determine the final yield (see references in Finney *et al.,* 1987; Araus *et al.,* 1998; Stone and Savin, 1999). This is particularly evident for cereals in Mediterranean conditions, where drought during grain filling is the main environmental constraint affecting GY and grain composition. Therefore, provided that the chemical composition of kernels covaries with GY due to their common influence from growing conditions during the crop cycle, it may be possible to correlate the spectral reflectance signal of kernels in the near infrared region with GY. NIRS analysis of complex characteristics (such as GY) with an undefined chemical basis, requires the use of chemometric tools which are independent of any previous knowledge of the chemical nature of the samples analyzed. In this work, Partial Least Squares Regression (PLSR; Marterns and Naes, 1989) was used to find a model able to associate high GY values with the chemical information inherent to the near infrared spectrum. The intrinsic features of PLSR, which subjects spectral data and predictive variables to a simultaneous Principal Component Analysis (PCA), allow the dimensional reduction needed to address a given problem with no appreciable loss of relevant information.

Materials and methods

A large set of genotypes of durum wheat (*Triticum turgidum* L. var. *durum*) were cultivated in rain-fed conditions in two sites (Breda and Tel Hadya) in NW Syria, with consistent differences in rainfall and evaporative demand. A third trial with the same genotypes was planted at Tel Hadya under support irrigation. These genotypes corresponded to the durum core collection assembled at the International Center for Agricultural Research in the Dry Areas (ICARDA) by the CIMMYT/ICARDA durum breeding program of the Centro Internacional de Mejoramiento de Maiz y Trigo (CIMMYT). Growth conditions, harvesting procedures and yield attained are detailed in a previous paper (Araus *et al.,* 1997). The different kinds of environment, as well as the large set of genotypes assayed allowed us to obtain a wide range of environmental and genotypical variation for yield. Mature kernels were ground and oven-dried at 60ºC for 48 h before being analysed on a NIRSystems 6500 spectrophotometer equipped with a fiber-optic probe for quantitative analysis. The instrument was controlled by the spectrophotometer's bundled software (NSAS version 3.20). Reflectance spectra were obtained directly from 271 samples with 32 triplicate scans performed at 2 nm intervals over the wavelength range 1100-2500 nm. An average spectrum was subsequently computed from collected data. These spectra were regressed by direct application of PLSR. Spectra were imported from the NSAS file into the programme UNSCRAMBLER v 6.0, developed by CAMO, which included the PLSR algorithm used in the building of models. Samples were divided in two sets: one set of 93 samples was used for calibration purposes, while an independent set (178 samples) was used to evaluate the predictive ability of the model. Samples were selected in order to obtain a calibration set with more or less regular increments for GY (i.e. flat calibration), including samples from different sites in the overlapping ranges to minimise the weight of sample origin in the model. In order to correct baseline shifts, often related with particle size, different spectral treatments were tested: 1st and 2nd derivative, SNV (Barnes *et al.,* 1989; McClure, 1994). The number of factors used in each model was determined with cross-validation (Wold, 1978). The Standard Error for Calibration (SEC) was calculated to assist in model selection:

$$SEP = \sqrt{\frac{\sum (Y_{ref} - Y_{nirs})^2}{(N-1)}}$$

where Y_{ref}, Y_{nirs} stand for the GY value of each sample determined either by harvest in the field (reference) or predicted from NIRS, respectively, and N is the number of samples. The coefficient of determination (r^2), the intercept and the slope of the linear regression between predicted and measured values also were used in selecting the best-fit model. Finally, in order to evaluate performance of the selected model, Standard Error for Prediction (SEP) and coefficient of determination (r^2) were also calculated for the prediction set.

Results and discussion

In order to address a broad explanation of the chemical basis that justifies the ability of NIRS in yield evaluation, correlation coefficients (r) between GY and each of the wavelengths assayed in the near infrared region of the spectrum were plotted (Fig. 1). We saw that the best correlations (r over 0.4) corresponded, according to the literature (Goddu and Delker, 1960; Osborne and Fearn, 1986; Shenk *et* *al.,* 1992), with the reflectance signature of starch (1900 and 2000 nm), cellulose (1780 and 2336 nm), or both (1580 and 2100 nm), water (1940 nm), peptide bonds (1980 nm) and amine groups (2132, 2245 nm). From these results, we can conclude that differences in the amount and/or composition of carbohydrates and proteins in kernels should be the main sources of correlation between NIRS spectra and GY values.

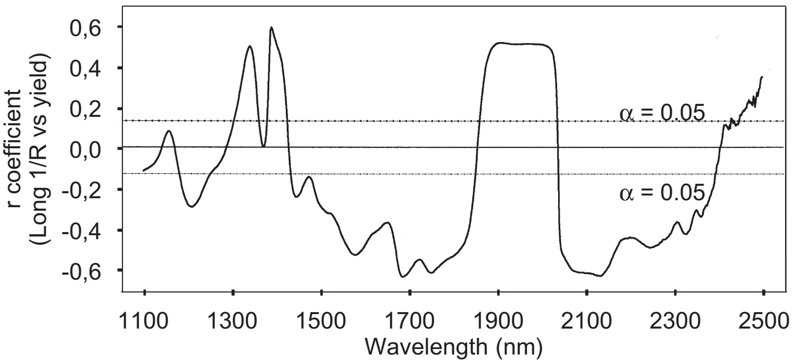

Fig. 1. Correlation spectrum (r) between each Near Infrared Wavelength assayed and Grain Yield.

The best results were found aplying SNV to the whole spectrum (1100-2500 nm) and using 9 factors for PLSR. NIRS-predicted vs measured plot for both calibration and prediction sets, as well as the regression line and statistics for calibration samples are shown in Fig. 2. Correlations between NIRS-predicted and field-measured GY were highly significant ($P < 0.001$), with r^2 values of 0.88 and 0.76 for calibration and validation sets, respectively. However, both SEC (790 kg/ha) and SEP (1190 kg/ha) were higher than it could be expected from these correlations. This was probably due to the large number of environmental and genetic factors affecting GY, some of them undetectable by NIRS analysis in kernels. Despite this, the model provides a broad picture of GY values, revealing a degree of association between NIRS information and measured GY. The major limitation of this approach lies in its empirical nature, which determines the necessity of a previous calibration procedure. However, these quick, non-destructive and standarized measures using NIRS could be used in early generations of a breeding program to discard most of the poor-yielding genotypes, thus increasing genetic gain. Moreover these NIRS evaluations could be combined with the grain quality evaluations which are performed routinely with NIRS in breeding programmes (see Hollamby and Bayraktar, 1996). Selection during later generations of the breeding program could be performed using the conventional procedures which involve multilocal assays.

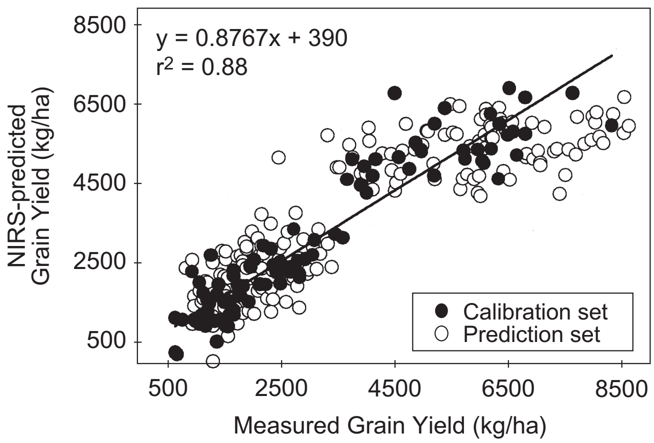

Fig. 2. Relationship between Grain Yield predicted by NIRS and measured in the field.

Acknowledgements

This study was supported by the CICYT research projects AGF96-1137-C02-01 and PB97-0865 (Spain). We appreciate the generous contribution of A. Asbati from ICARDA.

References

Araus, J.L. (1996). Integrative physiological criteria associated with yield potential. In: *Increasing Yield Potential in Wheat: Breaking the Barriers*, Reynolds, M.P., Rajaram, S. and McNab, A. (eds). CIMMYT, Mexico D.F., pp. 155-166.

Araus, J.L., Amaro, T., Casadesús, J., Asbati, A. and Nachit, M.M. (1998). Relationships between ash content, carbon isotope discrimination and yield in durum wheat. *Australian Journal of Plant Physiology,* 25: 835-842.

Araus, J.L., Amaro, T., Zuhair, Y. and Nachit, M.M. (1997). Effect of leaf structure and water status on carbon isotope discrimination in field-grown durum wheat. *Plant Cell and Environment,* 20: 1484-1494.

Austin, R.B. (1993). Augmenting yield-based selection. In: *Plant Breeding, Principles and Prospects*. Hayward, M.D., Bosemark N.O. and Romagosa, I. (eds). Chapman-Hall, London, pp. 391-405.

Barnes, R.J., Dhanoa, M.S. and Lister, S.J. (1989). Standard normal variate transformation and de-trending of Near-Infrared diffuse reflectance spectra. *Applied Spectroscopy*, 43(5): 772-777.

Finney, K.T., Yamazaki, W.T., Youngs, V.L. and Rubenthaler, G.L. (1987). Quality of hard, soft, and durum wheats. In: *Wheat and Wheat Improvement,* 2nd edn, Heyne, E.G. (ed.), Number 13 in the Series Agronomy. ASA, Inc., CSSA, Inc., SSSA, Inc. Publishers, Madison, pp. 677-748.

Goddu, R.F and Delker, D.A. (1960). Spectra-structure correlations for the Near-Infrared region. *Analytical Chemistry*, 32(1): 140-141.

Hollamby, G. and Bayraktar, A. (1996). Breeding objectives, philosophies and methods in South Australia. In: *Increasing Yield Potential in Wheat: Breaking the Barriers*, Reynolds, M.P., Rajaram, S. and McNab, A. (eds). CIMMYT, Mexico D.F., pp. 52-65.

Marterns, H. and Naes, T. (1991). *Multivariate Calibration*. Wiley, New York.

McClure, W.F. (1994). Near-Infrared spectroscopy: The giant is running strong. *Analytical Chemistry,* 66(1): 43A-53A.

Osborne, B.G. and Fearn, T. (1986). *Near Infrared Spectroscopy in Food Analysis*. Wiley, New York.

Shenk, J.S., Workman, J.J. and Westerhans, M.O. (1992). Application of NIRS to agricultural products. In: *Handbook of Near Infrared Spectroscopy Analysis*, Burns, D.A. and Ciurczak, E.W. (eds). Vol. 13 in Practical Spectroscopy Series. Dekker, Inc., New York, pp. 383-431.

Stone, P.J. and Savin, R. (1999). Grain quality and its physiological determinants. In: *Wheat: Ecology and Physiology of Yield Determination*, Satorre, E.H. and Slafer, G.A. (eds). FPP, New York, pp. 84-120.

Wold, S. (1978). Cross validatory estimation of the number of components in factor and principal components models. *Technometrics*, 20: 397-406.