ELSEVIER

Journal of Cereal Science 43 (2006) 236–243

Journal of CEREAL SCIENCE

www.elsevier.com/locate/jnlabr/yjcrs

Development of a calibration to predict maize seed composition using single kernel near infrared spectroscopy

Tesfaye M. Baye [a,1], Tom C. Pearson [b], A. Mark Settles [a,*]

[a] *Horticultural Sciences Department, University of Florida, P.O. Box 110690, Gainesville, FL 32611-0690, USA*
[b] *USDA-ARS Grain Marketing and Production Research Center, 1515 College Avenue, Manhattan, KS 66502, USA*

Received 8 August 2005; received in revised form 10 November 2005; accepted 16 November 2005

Abstract

The relative composition of protein, oil, and starch in the maize kernel has a large genetic component. Predictions of kernel composition based on single-kernel near infrared spectroscopy would enable rapid selection of individual seed with desired traits. To determine if single-kernel near infrared spectroscopy can be used to accurately predict internal kernel composition, near infrared reflectance (NIR) and near infrared transmittance (NIT) spectra were collected from 2160 maize kernels of different genotypes grown in several environments. A validation set of an additional 480 kernels was analyzed in parallel. Constituents were determined analytically by pooling kernels of the same genotype grown in the same environment. The NIT spectra had high levels of noise and were not suitable for predicting kernel composition. Partial least squares regression was used to develop predictive models from the NIR spectra for the composition results. Calibrations developed from the absolute amount of each constituent on a per kernel basis gave good predictive power, while models based on the percent composition of constituents in the meal gave poor predictions. These data suggest that single kernel NIR spectra are reporting an absolute amount of each component in the kernel.

Keywords: Maize; Near infrared spectroscopy; Seed composition; Single kernel

1. Introduction

Cereal grains contribute to over 60% of the total world food production (Lásztity, 1999). Cereals are predominantly composed of carbohydrates, mostly in the form of starch, with considerable amounts of protein as well as some lipids, vitamins, and minerals. Both genetic and environmental effects create significant variation in the amount and quality of each of these constituents. Multiple methods have been developed to help breeders screen crops for various seed composition traits (e.g. Baenziger et al., 2001; Dunlap et al., 1995). Chemical analysis procedures are the most widely accepted reference methods for determining seed composition. However, these methods frequently are destructive and require large samples of grain.

Near infrared spectroscopy provides an alternative, non-destructive technology for measuring constituents of biological materials. Organic molecules have specific absorption patterns in the near infrared region that can report the chemical composition of the material being analyzed (Williams and Norris, 2001). Near infrared spectra can be collected either from the reflectance (NIR) of a sample or the transmittance (NIT) through a sample (Delwiche, 1995; Williams, 1979). Both NIR and NIT measurements allow the simultaneous determination of multiple constituents in a sample and are commonly used to predict the composition of bulk whole grain samples in maize (Orman and Schumann, 1991). Bulk whole-grain samples can be screened rapidly, require no sample preparation, and preserve the kernels after the measurement for further analysis or for propagation (Baye and Becker, 2004; Velasco et al., 1999). However, use of whole grain samples does not allow the identification of individual kernels that deviate significantly from the mean composition within

Abbreviations: bt1, *brittle1*; *bt2*, *brittle2*; *dek*, *defective kernel*; NIR, near infrared reflectance; NIT, near infrared transmittance; PLS, partial least squares regression; *rgh3*, *rough endosperm3*; *sh2*, *shrunken2*; *su1*, *sugary1*.

* Corresponding author. Tel.: +1 352 392 7571; fax: +1 352 392 5653.
E-mail address: settles@ufl.edu (A.M. Settles).
[1] Present Address: Department of Biostatistics, University of Alabama at Birmingham, 1665 University Blvd, RPHB 445B, Birmingham, AL 35294, USA.

0733-5210/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jcs.2005.11.003

a population. In addition, bulk samples give no indication of an abnormal distribution of kernels within the sample, such as a few kernels contaminated by a fungus or a segregating population of kernels with differing composition (Dowell et al., 2002).

Non-destructive analysis of single kernel composition is valuable for identifying outlying individuals both for breeding and for industrial seed sorting applications. In maize, single-kernel nuclear magnetic resonance (NMR) technology has been used to select oil traits (Alexander et al., 1967), and it has been demonstrated that the rate of improvement for maize oil traits can be enhanced over bulk sample analysis through single-kernel selection (Silvela et al., 1989). However, NMR is limited to detecting liquid constituents such as oil and moisture in maize seeds. NIR and NIT spectroscopy give the possibility for rapid screening of individual kernels for multiple chemical constituents. Single kernel NIT spectra between 920 and 950 nm were found to correlate very well with moisture content, with a 2% standard error of prediction of moisture content (Finney and Norris, 1978). More recently, single kernel NIR and NIT spectroscopy has been used to sort individual maize kernels for different types of fungal infections (Dowell et al., 2002; Pearson et al., 2001, 2004) or to identify genetically modified kernels (Munck et al., 2001).

The present study was conducted to determine if single kernel NIR or NIT spectroscopy could also be used to predict kernel composition in maize. These predictions will help geneticists and breeders to screen large numbers of samples and then select and propagate single seeds with desirable composition traits. Single kernel NIR data have been used to develop a predictive model for wheat protein content (Delwiche and Hruschka, 2000). However, maize kernels have a much less uniform internal structure with the maize embryo comprising a larger proportion of the seed than in wheat. Indeed, attempts to develop calibrations for maize oil content using NIT data suggest that single kernel predictive models would be difficult to develop for maize (Cogdill et al., 2004; Orman and Schumann, 1992).

Here, we report the development calibration equations to predict accurately individual maize kernel constituents using NIR technology. Near infrared predictive models are best developed with samples that display a large range of compositional variation (reviewed in Willimas and Norris, 1987). In maize, *defective kernel* (*dek*) mutants have large effects on seed size and can effectively delete major constituents of the kernel by affecting starch biosynthesis (reviewed in Boyer and Hannah, 2001), storage protein accumulation (reviewed in Gibbon and Larkins, 2005), or by aborting embryo growth early in seed development (Magnard et al., 2004). In the mature kernel, the embryo contains the highest proportion of oil and removal of the embryo causes a reduction in total oil content. We reasoned that *dek* mutants would provide a large range of different maize kernel compositional variants for developing predictive models from near infrared spectra and focused on these mutants in this study.

2. Materials and methods

2.1. Seed stocks

All maize seeds used were obtained from plants grown at the University of Florida, Plant Science Research and Education Unit (Citra, Florida). Eight inbred lines commonly used for genetic studies were included as reference stocks. These lines included: W22, W23, Mo17, B73, A632, W64A as well as color-converted W22 and A632 stocks ($W22^{ACR}$ and $A632^{ACR}$). All inbred seeds were from self-pollinated ears. To develop the calibration and external validation sets, maize kernel mutants and their normal sibling seeds were used as a source of large variance in seed composition. The majority of the seed mutants were derived from the UniformMu transposon-tagging population (McCarty et al., 2005). The UniformMu mutants were selected based on visible defects in embryo or endosperm development, which are likely to cause altered kernel composition.

Seed for each of 24 UniformMu mutants were grown in two field seasons and self-pollinated to identify segregating ears. In each field season, there were variations in soil type, average temperature, and day length. A pair of segregating ears for each mutant was selected to include both environmental and genetic variation. In addition, six mutants with strong effects on starch accumulation were used as controls including: *shrunken2* (*sh2*) in a W22 background, *brittle1* (*bt1*) in a W23 background, and *sugary1* (*su1*), *bt1*, *brittle2* (*bt2*), and *sh2* in a W64A background. Two ears of *su1* in W64A and a single ear of each of the other mutants were included in the study for a total of 55 ears of corn segregating for 30 different seed mutant loci. A mutant and normal sample of 24 kernels each was selected from each ear for single-kernel near infrared spectroscopy and analytical determination of composition. A total of 110 samples were used to develop and test the calibrations. Fig. 1A shows examples of the mutant and corresponding normal kernels for eight of the ears used in this study, and a schematic of the overall study design is shown in Fig. 1B.

2.2. Near infrared data collection and pretreatments

Near infrared spectra were collected from 24 mutant and 24 normal sibling kernels from each of the 55 ears selected for the calibration development. Prior to collecting the spectra, the kernels were equilibrated to ambient humidity for 2 or more weeks in a controlled laboratory environment. NIR and NIT spectra were collected from the abgerminal side of individual kernels. The NIR spectra were measured on maize kernels placed manually onto the quartz window of a bifurcated interactance fiber-optic bundle (Fig. 2) attached to a diode-array spectrometer and light source (DA7000, Perten Instruments, Springfield, IL). The viewing area was 17 mm diameter, the illumination bundle was a 7 mm diameter ring, and the reflectance probe bundle was 2 mm diameter. The light source was chopped at a frequency of 30 Hz and the integration time for the spectrometer was set at 33 ms. A total of 15 spectra were acquired for each kernel without repositioning the kernel