

# Single Kernel Near-Infrared Analysis of Tetraploid (Durum) Wheat for Classification of the Waxy Condition

Stephen R. Delwiche,<sup>1,2</sup> Robert A. Graybosch<sup>3</sup>, Lavern E. Hansen<sup>3</sup>, Edward Souza<sup>4</sup>, and Floyd E. Dowell<sup>5</sup>

## ABSTRACT

Cereal Chem. 83(3):287–292

Plant breeding programs are active worldwide in the development of waxy hexaploid (*Triticum aestivum* L.) and tetraploid (*T. turgidum* L. var. *durum*) wheats. Conventional breeding practices will produce waxy cultivars adapted to their intended geographical region that confer unique end use characteristics. Essential to waxy wheat development, a means to rapidly and, ideally, nondestructively identify the waxy condition is needed for point-of-sale use. The study described herein evaluated the effectiveness of near-infrared (NIR) reflectance single-kernel spectroscopy for classification of durum wheat into its four possible waxy alleles: wild type, waxy, and the two intermediate states in which a null allele occurs at either of the two homologous genes (*Wx-1A* and *Wx-1B*)

that encodes for the production of the enzyme granule bound starch synthase (GBSS) that controls amylose synthesis. Two years of breeders' samples (2003 and 2004), corresponding to 47 unique lines subdivided about equally into the four GBSS genotypes, were scanned in reflectance (1,000–1,700 nm) on an individual kernel basis. Linear discriminant analysis models were developed using the best set of four wavelengths, best four wavelength differences, and best four principal components. Each model consistently demonstrated the high ability (typically >95% of the time) to classify the fully waxy genotype. However, correct classification among the three other genotypes (wild type, *wx-A1* null, and *wx-B1* null) was generally not possible.

Starch in cereal endosperm consists of two main macromolecules: amylose, and amylopectin. Amylose, which consists of long linear chains of  $\alpha$ -(1-4)-linked D glucan units, is synthesized in the amyloplast by the action of the enzyme called the waxy protein, known as granule bound starch synthase (GBSS) (Shannon and Garwood 1984; Nakamura and Yamamori 1995). Without the presence of GBSS, amylose concentration becomes near zero, commonly referred to as the waxy condition. Under customary natural conditions, GBSS is encoded by *Wx* genes, such that each genome (A, B, or D for hexaploid wheat; A or B for tetraploid wheat) is an original contributor to GBSS. Evolution of wheat produced a reciprocal translocation of loci of chromosomes 4A and 7B, resulting in the translocation of *Wx-B1* from 7B to 4A. Therefore in hexaploid wheats (*Triticum aestivum* L.), the waxy genes are found on chromosomes 4A, 7A, and 7D, while in tetraploid wheats (*Triticum turgidum* L.), they are located on chromosomes 4A and 7A. The condition of a partial waxy line occurs when at least one but not all waxy genes are null alleles. Waxy wheat breeding programs are currently underway in several countries for the development of adapted bread and durum wheats. Potential uses of waxy wheat include that of being a waxy maize substitute (Reddy and Seib 2000), a blending flour to confer unique processing characteristics (Hoshino et al 2000; Epstein et al 2002), a possible shelf life extender (Nakamura and Yamamori 1995), and various industrial applications. Zhao et al (1998) discovered that the *wx-B1* null trait commonly occurs in Australian wheat cultivars and that in a survey of more than 100 of these, all of the good quality cultivars for white salted Japanese udon noodles possessed this null allele. The *wx-B1* null cultivars tended to have a higher starch pasting viscosity (as measured by the Rapid Visco Analyser) and higher flour swelling volume (defined in Crosbie

1991). Much of the early history of waxy wheat breeding research is reviewed in Graybosch (1998). Recent analysis by Graybosch et al (2003) determined the role of genetics and environment in bread wheats, with the findings that protein-quality traits of waxy lines were not grossly different from nonwaxy check cultivars, and that starch-related traits were relatively stable across environments. On durum lines, Vignaux et al (2004) recently found that while protein quality of fully waxy lines was slightly lower than wild type and partial waxy lines, the waxy lines were still within the strong gluten range. A stronger effect was observed for the milling properties, with waxy lines more susceptible to starch damage. Interestingly, partial waxy lines did not have statistically different amylose contents compared with wild type lines and, in fact, other measurements such as kernel ash percentage, SDS sedimentation volume, and wet gluten also demonstrated equivalence between partial waxy and wild type lines.

Successful release and marketing of commercial waxy or partial waxy cultivars of durum wheat will ultimately depend on the market's perceived value of these wheats and a coincident reliable and rapid test to authenticate their waxiness condition. Presently, breeders typically determine the GBSS genotype by iodine staining of the nongerm region of the endosperm for recognition of waxy seeds, however this procedure is not definitive for typing partial waxy lines. Detection of the different GBSS isoforms is typically performed by SDS-PAGE (Zhao and Sharp 1996) or alternatively by ELISA (Rahman et al 1995). Recently identified genetic markers for each of the waxy null alleles in hexaploid wheat have resulted in the feasibility of using multiplex PCR techniques for genotyping all GBSS states (Nakamura et al 2002). However, in common with other methods for GBSS genotyping, the PCR methodology is expensive and dependent on highly trained analysts. For commercial purposes as well as the desire for faster screening procedures in breeding, genotyping should be simple, inexpensive, and possible to perform at the point of sale. Near-infrared (NIR) spectroscopy is a well-entrenched methodology of the cereals industry of the past 30 years, and is used extensively for prediction of protein and moisture at grain elevators and a host of other properties in the analytical laboratories of processing facilities. In recent years, we evaluated the ability of NIR reflectance spectroscopy on the ground meal of hard red winter (hexaploid) wheat to classify by GBSS genotype (five of the eight genotypes available), these being wild type (functional alleles at all three *wx* loci), single null (null alleles at *wx-A1* or *wx-B1* loci), double null (null alleles at *wx-A1* and *wx-B1* loci), or waxy (null alleles at all three *wx* loci) (Delwiche and Graybosch

<sup>1</sup> USDA-ARS, Beltsville Agricultural Research Center, Instrumentation and Sensing Laboratory, Building 303, BARC-East, Beltsville, MD 20705-2350. Names are necessary to report factually on available data; however, the USDA neither guarantees nor warrants the standard of the product, and the use of the name by the USDA implies no approval of the product to the exclusion of others that may also be suitable.

<sup>2</sup> Corresponding author. Phone: 301-504-8450. E-mail: delwiche@ba.ars.usda.gov

<sup>3</sup> USDA-ARS, Department of Agronomy, University of Nebraska, Lincoln, NE.

<sup>4</sup> University of Idaho, Plant Breeding and Genetics Department, Aberdeen Research and Extension Center, Aberdeen, ID.

<sup>5</sup> USDA-ARS, Grain Marketing and Production Research Center, Manhattan, KS.

2002). Difficulties arose in correctly classifying partial waxy lines, with misclassifications usually being assignments into neighboring gene classes such as a double null assigned to the single null class. Relative to the conventional methods for the identification of waxy segregants in breeding programs, the NIR-based methods are simpler, more rapid, and adaptable to existing automated sorting technologies (Dowell et al, *unpublished*). As a means to circumvent the issue of single and double nulls, the study described herein examined waxy allele classification of tetraploid species, based on the NIR reflectance of intact whole durum kernels. The objective of this research was to assess the potential of NIR practices in classifying durum wheat according to the presence or absence of the functional alleles at the two waxy loci. Success of such practices could lead to immediate adoption in waxy seed breeding programs, especially in the nondestructive selection of seed from segregating samples.

## MATERIALS AND METHODS

**Plant materials.** Sets of wild type, partial waxy (*wx-A1* null or *wx-B1* null), and waxy (*wx-A1* and *wx-B1* double-null) durum wheat lines were developed in several genetic backgrounds. The hexaploid partial waxy (*wx-A1* and *wx-B1* null) wheat cultivar Kanto 107 was used as a female parent in matings with the wild type tetraploid durum wheat cultivars Botno and Monroe. F<sub>1</sub> generation plants were self-pollinated in a greenhouse at Lincoln, NE, and F<sub>2</sub> waxy seed were identified by staining with I<sub>2</sub>KI, and again greenhouse grown. F<sub>2</sub> waxy plants were used as females in a second round of matings with one of three durum wild type cultivars, Renville, Cando, or Vic. The BC<sub>1</sub>F<sub>1</sub> generation was sown in the greenhouse in Lincoln, NE, in the fall of 1999, and allowed to self-pollinate. BC<sub>1</sub>F<sub>2</sub> seed were stained with I<sub>2</sub>KI solution (1:10 dilution of stock solution formed by 200 mg of I<sub>2</sub> + 2 g of KI/100 mL of H<sub>2</sub>O), and waxy segregants were identified and sown in the field at Aberdeen, ID, in the spring of 2000. Idaho-grown waxy lines were mated in the field again to various wild type durum cultivars, including Kyle, Utopia, and Cando, respectively, forming three populations, a single population, and another single population. BC<sub>2</sub>F<sub>1</sub> plants were greenhouse-sown in the fall of 2000, self-pollinated, and BC<sub>2</sub>F<sub>2</sub> seed were sown as bulk populations in greenhouse flats. Materials were then advanced to the BC<sub>2</sub>F<sub>3</sub> generation through single-seed descent. Subsequent single heads (BC<sub>2</sub>F<sub>4</sub> seed) were harvested and used to identify granule-bound starch synthase (GBSS) genotypes. Six seeds were removed from each BC<sub>2</sub>F<sub>4</sub>-derived head, ground in a coffee grinder, and pooled for analysis. Starch was purified and GBSS isoforms extracted and separated by SDS-PAGE and identified by silver staining (Graybosch et al 1998). Wild type durum wheat produces two GBSS isoforms with apparent molecular masses of ≈60 kDa; *wx-A1* and *wx-B1* null lines each produce one detectable GBSS band, and waxy genotypes produce no GBSS. Three lines having each of the four possible tetraploid wheat GBSS genotypes were identified per population. Remnant seed of each line was saved and used to sow subsequent field experiments. In the fall of 2002, each of the 60 lines was sown in single, unreplicated 2.5-m rows in Yuma, AZ. Fifty-three lines were successfully harvested in the spring of 2003 and planted the following fall in a randomized complete block design with three replicates. Two of the three field replicates were used for subsequent analyses. Samples are hereafter designated by harvest Year 1 (2003) or Year 2 (2004). GBSS genotypes of Year 1 and Year 2 samples were confirmed by testing bulk samples of 8–12 seeds per line. In addition, wild type, *wx-A1* null, and *wx-B1* null lines were further analyzed by evaluation of 6–8 single seeds (as independent samples). Each of the first of two field replicates of Year 2 was examined. Two samples were still segregating and consequently were removed (along with their field replicate counterparts) from the analysis.

**Equipment.** A diode array spectrometer (Zeiss MCS511, Jena, Germany) with a 128-element indium gallium arsenide array was used to collect reflectance readings (943–1,705 nm). Kernel reflectance signals were referenced to corresponding reflected energy readings from a white polytetrafluoroethylene tile (Spectralon, Lab-sphere, Sutton, NH). Illumination was supplied by two externally controlled 5V, 150 mA tungsten filament lamps with gold-coated parabolic reflectors. These lamps, each opposing the other and oriented 45° with respect to horizontal were situated <10 mm from the kernel. An integration time of 6.8 msec per scan was used. A single fiber (1 m × 600 μm diameter), with one end positioned directly above the kernel (≈5 mm distance), collected radiation reflected by the kernel. Forty-eight indented slots in a bakelite plate were used to house the kernels from each sample. Because of the parallel alignment of the slots, the axes of the kernels were always in a common alignment; however, the placement of the kernel within the slot was random with respect to the rotational angle about its long axis. Therefore, with respect to the overlying fiber probe, crease up, crease down, and crease sideways positions were equally likely. This semirandom alignment approach was used because it was thought that such positioning is realistically achievable in an automated system. Furthermore, previous work of the author indicated that rotational alignment about the kernel axis did not have a significant effect on the accuracy of other classification models, such as those for mold damage (Delwiche 2003).

Year 2 samples were also scanned in diffuse reflectance (1,100–2,498 nm) by an analytical spectrometer (model 6500, Foss-NIR-Systems, Laurel, MD) equipped with a rotating sample drawer assembly. Before scanning, the samples were ground in a cyclone mill. Scans from duplicate packs of a standard forage cell (≈5 g meal/pack) were averaged.

**Procedure.** Kernels were scanned in batches of 48, with this number being one less than the number of slots (7 × 7 array) in the horizontal plate. At uniform time intervals, a computer-controlled two-axis movable stage positioned each kernel in line with the light source assembly. Thirty-two successive scans were averaged and transformed to log(1/R) and eventually stored to computer file. Afterward, linear interpolation was applied to each spectrum so that a uniform wavelength spacing of 6 nm between neighboring array elements was established, instead of the nonuniform wavelength spacing (ΔWL = 5.85–6.23 nm) of the diode array. Low signal in the low wavelength region prevented the use of the corresponding data, which left an effective wavelength region at 1,002–1,704 nm, for a total of 118 spectral values.

**Genotyping.** Ten kernels were drawn at random from each sample of one field replicate from Year 2. SDS-PAGE was performed on the endosperm from half of each kernel to confirm the sample's genotype. Any sample with fewer than eight kernels possessing the anticipated genotype was discarded from classification analysis. The sample's corresponding field replicate was also discarded, based on the conservative assumption that both sample and replicate were still segregating.

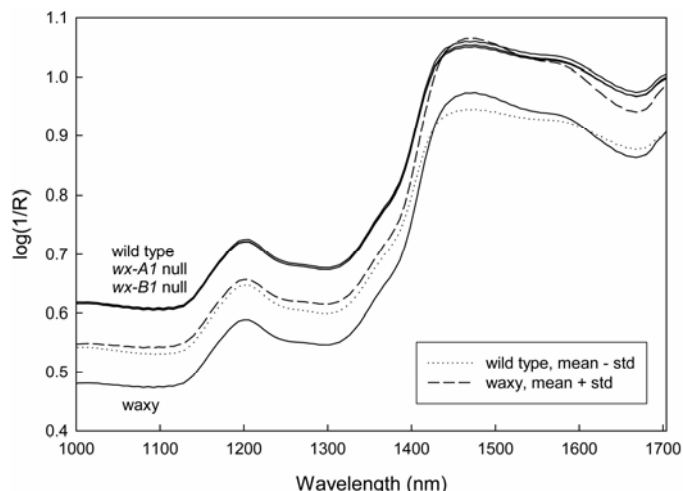
**Amylose and protein analysis.** Amylose content (reported as the percentage of apparent amylose in total starch) of Year 1 samples and two field replicates of Year 2 samples were measured by iodine-binding blue complex colorimetry, using slight modifications to the method of Knutson and Grove (1994). To enhance precision of diluting the sample for complex formation, the DMSO-iodine reagent volume was adjusted by using a positive displacement pipettor. A standard calibration curve was developed using serial dilutions of a crude wheat amylose extract obtained by the procedure of Klucinec and Thompson (1998). All samples were measured in duplicate, with the second set of assays completed approximately one month after the first set. The repeatability of the chemical procedure, as gauged by the standard deviation of a control sample inserted approximately every eight sample assays, was 1.1%. Protein content (N × 5.7, reported in

percentage units) was measured by combustion on 150-mg portions of ground meal (AACC International 2000). Duplicate assays were performed one day apart. Defined in a similar manner to the amylose colorimetry procedure, the repeatability standard deviation for the combustion procedure was 0.1%.

Analyses of variance ANOVA (mixed procedure, SAS, v. 8.02) were performed on amylose content and protein content, as well as kernel mass, to evaluate the effect of GBSS genotype, with year as a blocking (random) factor. When the effect was significant, pairwise comparisons of means ( $\alpha = 0.01$ ) were subsequently performed.

**Classification modeling.** Linear discriminant analysis (LDA) (Fisher 1936) was primarily used in classification model development. All linear discriminant analysis was performed with one-sample-out cross-validation using SAS procedures (Stepdisc, Princomp, and Discrim). Preliminary trials with nonlinear forms of discriminant analysis, such as *k*-nearest neighbor, did not demonstrate superior performance and were therefore dropped from subsequent analyses. Classification models were developed for each crop year individually and for both years combined. Models were based on four independent groups, wild type, *wx-A1* null, *wx-B1* null, and waxy. The numbers of kernels per allele were wild type = 528 (Year 1) + 960 (Year 2), *wx-A1* null = 672 + 912, *wx-B1* null = 528 + 768, waxy = 528 + 960. Stepwise discriminant analysis was performed on three separate parameter structures: 1) the best set of individual wavelengths, 2) the best set of wavelength differences [ $\log(1/R)_{\lambda 1} - \log(1/R)_{\lambda 2}$ ], and 3) the best set of principal component scores. Before formation of principal components, the spectra were treated with a Savitzky-Golay second difference convolution function (7-point window) for the purpose of reducing spectral variation caused by conditions not related to the GBSS state. Preliminary analyses indicated that

cross-validation accuracy did not increase as the number of wavelengths, wavelength differences, or principal components increased beyond four. Therefore, these results are limited to four-term models. Additionally, models for all combinations of two GBSS groups at a time were developed in anticipation of such classification needs.



**Fig. 1.** Mean single kernel Year 2 spectra (solid lines) of the waxy alleles, based on averages of 960, 912, 768, and 960 individual spectra from wild type, *wx-A1* null, *wx-B1* null, and waxy groups, respectively. The + one standard deviation (dashed line) and - one standard deviation (dotted line) envelopes are included for the waxy and wild type groups, respectively. Standard deviation envelopes for the partial waxy groups are omitted because of their similarity to the wild type.

**TABLE I**  
Cross-Validation Accuracies (%) of Four-Class Linear Discriminant Analyses<sup>a</sup>

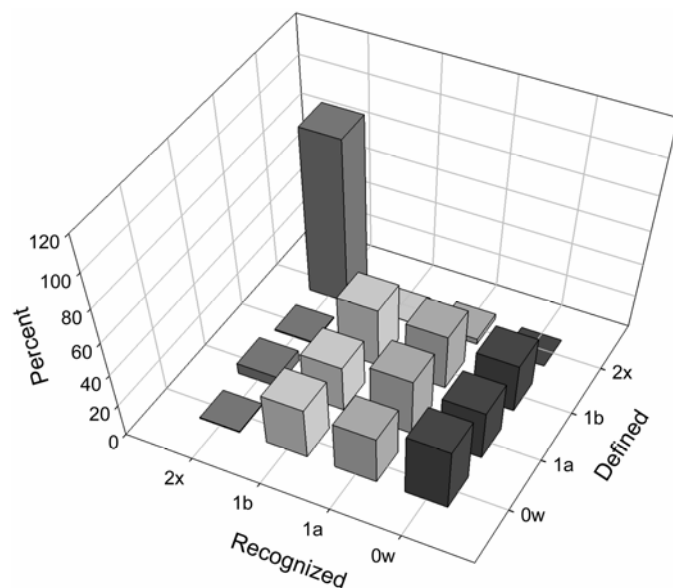
Model	Type of Spectra	Wavelengths (nm) or PC	Year	Accuracy by Allele (%)				
				Wild Type	<i>wx-A1</i> Null	<i>wx-B1</i> Null	Waxy	Average
LDA on best 4 WL	SK	1002, 1140, 1470, 1440	1	28.2	33.0	54.0	94.7	52.5
	SK	1002, 1140, 1470, 1428	2	40.1	36.4	26.8	99.1	50.6
	SK	1002, 1140, 1476, 1446	Both	41.3	26.3	36.0	97.2	50.2
LDA on best 4 WL differences	SK	(1494-1440), (1326-1242), (1476-1470), (1320-1278)	1	49.6	36.0	51.7	95.8	58.3
	SK	(1674-1410), (1338-1224), (1470-1452), (1344-1338)	2	28.8	37.7	34.5	98.2	49.8
	SK	(1494-1440), (1338-1182), (1356-1350), (1470-1170)	Both	46.8	32.1	45.5	96.8	55.3
LDA on best 4 PC, on 7pt SG 2nd der.	SK	PC2, PC8, PC7, PC4	1	51.0	38.0	36.6	70.1	48.9
	SK	PC2, PC4, PC1, PC8	2	38.1	34.9	35.9	96.7	51.4
	SK	PC2, PC6, PC8, PC1	Both	43.8	11.4	48.5	77.2	45.2
LDA on best 4 PC, on 7 pt SG 2nd der.	Avg of 48 SK	PC1, PC2, PC6, PC4	2	30.0	26.3	43.7	100	54.4
LDA on best 4 PC, 15 pt SG 2nd der.	Ground meal	PC2, PC3, PC5, PC6	2	80.0	47.4	56.2	100	70.9

<sup>a</sup> LDA, linear discriminant analysis; WL, wavelength; PC, principal component; 7 pt. SG 2nd der., Savitzky-Golay second derivative convolution with a 7-point window width; SK, single kernel.

## RESULTS AND DISCUSSION

On average, the spectra of the nonwaxy categories were very similar, as shown in Fig. 1 for Year 2. A one-sided standard deviation envelope is also included for the waxy and wild type categories. The partial waxy categories had standard deviation envelopes of the same size as the wild type envelope and are therefore omitted. Implied from the broad width and overlapping nature of the nonwaxy envelopes is the difficulty associated with spectrally based classification. Even the waxy category's envelope is sufficiently large to cause overlap with spectra of the three other categories, thus prohibiting the use of single wavelength classification models. Although GBSS genotype had a significant overall effect ( $P < 0.01$ ) on kernel mass, pairwise comparisons of means indicated that the difference between either partial waxy condition and the wild type condition was not significant. Based on kernel mass alone, four-category LDA models correctly classified only 57 and 53% of the waxy kernels for Years 1 and 2, respectively. In contrast with Vignaux et al (2004), who found the appearance of waxy durum kernels to be unique, the waxy kernels of the present study did not have a large difference in appearance with respect to wild type or partial waxy kernels.

Classification results for the four-group spectral LDA models are summarized in Table I. Based on the average of the percentage of correctly classified kernels of each group, the classification accuracy of single kernel models was between 45.2% (4-factor PCA, both years) and 58.3% (4-wavelength difference, Year 1), with an overall average of 51.4%. On average, Year 1 models were slightly more accurate than Year 2 or both year models (53.2% vs. 50.6% or 50.2%, respectively). Likewise, the models based on four wavelength differences were slightly better on average than the models based on four wavelengths or four principal components (54.5% vs. 51.1% or 48.5%, respectively). Most evident was the high accuracy associated with identification of waxy kernels, with accuracies often in excess of 95%. However, these high accuracies were offset by the very low classification accuracies for the three other groups, ranging from 11.4% (4-factor PCA, both years, *wx-A1* null) to 54.0% (4-wavelength, Year 1, *wx-B1* null). Using the 4-factor PCA Year 2 model as an example (average classification accuracy = 51.4%), the classification

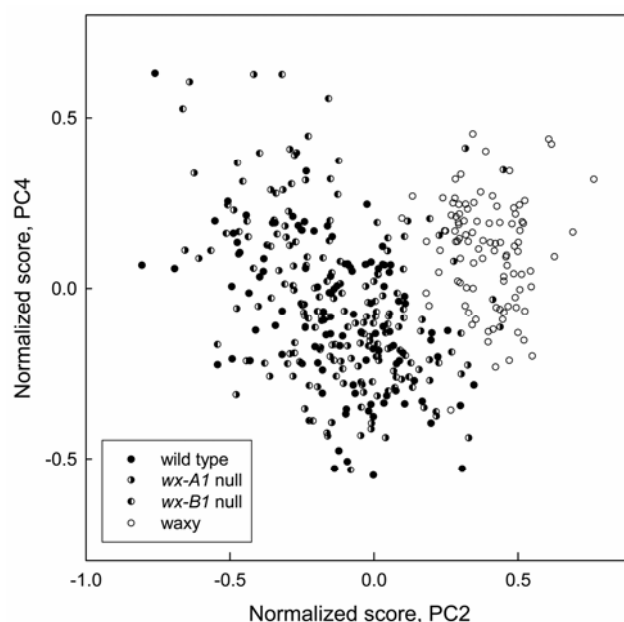


**Fig. 2.** Cross-validation classification assignment percentages for a linear discriminant analysis using four principal component scores (PC2, PC4, PC1, and PC8), Year 2 (spectral pretreatment = 7-point Savitzky-Golay second derivative, 0w = wild type, 1a = *wx-A1* null, 1b = *wx-B1* null, 2x = waxy).

assignments are demonstrated in Fig. 2. As evident from this figure, misclassification of waxy kernels seldom occurred, and misclassification of the wild type and partial waxy genotypes occurred frequently, but rarely were these assignments into the waxy group. In fact, a chi-square test performed on the *wx-B1* null kernels that were classified into the three nonwaxy groups (229, 257, and 276 kernels into wild type, *wx-A1* null, and *wx-B1* null, respectively) failed to find that the classified fractions differed from an expected fraction (1/3 : 1/3 : 1/3) caused by chance alone.

For the purpose of demonstrating the influence of waxiness on NIR spectra, principal component scores from the two factors that were identified as being most useful for classification, namely PC2 and PC4 (with 7-point Savitzky-Golay second derivative spectral pretreatment) are depicted in Fig. 3. To avoid crowding of plotted points, this graph contains a random selection of just 100 kernels from each category of Year 2 data. Plots of Year 1 data (not shown) were similar. Apparent from Fig. 3 is a distinct clustering of the waxy kernels. In contrast, wild type and partial waxy kernels showed significant overlap. Scores plots using any of the other principal components did not reveal any better separation within the nonwaxy categories.

Given the difficulty of classifying nonwaxy genotypes, the question arose on whether classification could be improved by spectral averaging. To address this question, cross-validation accuracies from a model that utilized the average spectrum of each sample's set of 48 single kernel spectra are included in Table I. Whereas the correct identification of waxy samples improved to the level of perfect classification, the accuracies of the nonwaxy samples were on par with the single kernel models, such that the overall accuracy was 54.4%. In contrast, the same spectral pretreatment and classification approach on ground meal spectra yielded improved accuracy in all four groups (80.0, 47.4, 56.2, and 100% for wild type, *wx-A1* null, *wx-B1* null, and waxy groups, respectively). Possible reasons for the superior performance of ground meal models include a better sensitivity of the analytical spectrometer to internal composition; ground particle size differences among groups, which produces a physical effect that is sensed in diffuse reflectance of the meal; and the elimination of surface reflectance and background differences caused by variation in kernel size and shape.



**Fig. 3.** Principal components scores of 100 randomly selected Year 2 kernels from each waxy allele (spectral pretreatment = 7-point Savitzky-Golay second derivative).

The results of two-class models, each based on 4 PC with 7-point Savitzky-Golay second derivative pretreatment structure, are summarized in Fig. 4. All six combinations of group pairs are shown. As seen by the bars in this graph, the best models were those that compared waxy with wild type, or waxy with either partial waxy group. Classification of *wx-A1* null and *wx-B1* null kernels resulted in the largest fraction of errors, with each group with <60% correct assignment.

Based on the amylose fraction of ground meal samples, the partial waxy lines for Year 1 were not significantly different, nor were their protein contents (Table II). This is in contrast with Sharma et al (2002), who found that durum lines with the *wx-B1* null allele had significantly lower amylose content, higher starch pasting peak viscosity, and higher semolina swelling power than comparable wild type lines. Working with hexaploid wheat, Yamamori and Quynh (2000) also found that the *wx-B1* allele has a stronger effect on amylose production than either the *wx-A1* or *wx-D1* alleles. Zhao et al (1998) postulated that the increases in flour swelling volume and peak pasting viscosity that are associated with the *wx-B1* null allele are not necessarily explained by a change in amylose content, but rather by a subtle structural change to the starch brought on by changes in the distribution of

glucose chain lengths. Although amylose fraction differences between partial lines were significant for Year 2 and aligned to be consistent with Sharma et al (2002), it is apparent that such differences are not sufficient to be spectrally sensed.

Generally, the difficulty of distinguishing between the partial waxy genotypes is consistent with published research. Vignaux et al (2004) found no reduction in the amylose fraction of partial waxy durum lines, nor did they find changes in pasta properties (e.g., cooking time, water absorption, firmness, cooking loss) with respect to wild type pasta (Vignaux et al 2005). They suggested three possible reasons for why durum lines that possess one null allele do not reduce the level of GBSS activity to the point where amylose content is reduced: dosage compensation, a plateau in amylose synthesis that is reached with just one allele active, and such differences are too small to be detectable. On potato, Filpse et al (1996) observed that whereas GBSS activity increased linearly with the number of wild type GBSS genes, the amylose content response was not linear, but instead reached a plateau. Furthermore, protein content differences among wild type and partial waxy lines were not statistically significant. With regard to amylose and protein contents of the nonwaxy lines, such agreement between the current study and published research suggests

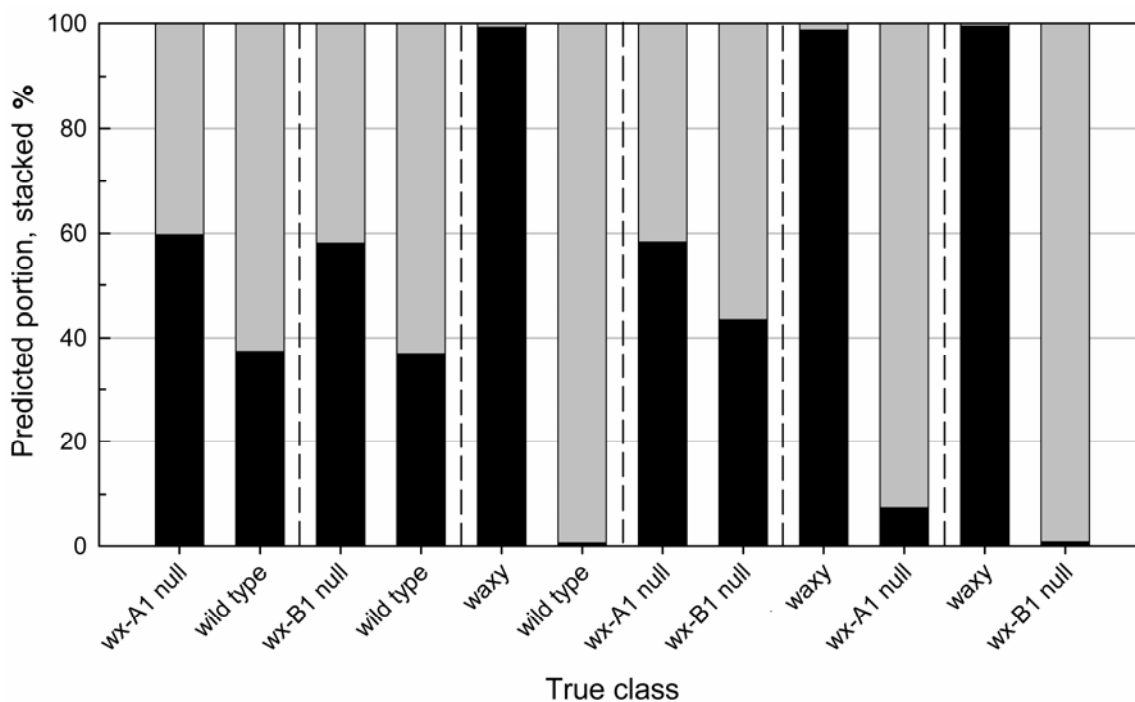


Fig. 4. Family of two-class linear discriminant analysis classification cross validations. The two-class models are separated by dashed lines, such that for each model, the black region of the left bar and the gray region of the right bar correspond to the correct assignment for the respective bar.

TABLE II  
Summary Statistics for Durum Wheat Samples<sup>a</sup>

Genotype	Year	N <sup>b</sup>	Mean Kernel Mass (mg)	Mean Protein Content (%)	Mean Amylose Fraction (%)
wild type	1	11	50.9	11.50	28.0
<i>wx-A1</i> null	1	14	51.6	11.51	23.2
<i>wx-B1</i> null	1	11	51.7	11.99	22.5
waxy	1	11	48.1	12.78	2.6
wild type	2	20	44.1	16.60	29.0
<i>wx-A1</i> null	2	19	43.0	16.94	26.8
<i>wx-B1</i> null	2	16	44.3	16.80	24.1
waxy	2	20	42.0	16.96	3.0
wild type	Both	31	46.5ab	14.79a	28.7a
<i>wx-A1</i> null	Both	33	46.6b	14.63a	25.3b
<i>wx-B1</i> null	Both	27	47.3a	14.84a	23.4b
waxy	Both	31	44.2c	15.48a	2.9c

<sup>a</sup> Significant pairwise (LSD) differences in means indicated by different letters (at  $P = 0.01$ ).

<sup>b</sup>  $N$  = number of samples, which is the basis for the means of protein content and amylose content. For kernel mass, the basis is  $N \times 48$ .

that the NIR-based approach has utility as a tool in wheat breeding when the genetic trait of interest is chemically expressed.

Methods such as iodine staining, SDS-PAGE separations of waxy proteins, or the DNA polymerase chain reaction at best can process a few hundred seeds per day. NIR and automated sorting technologies can process thousands of seeds per day, and are far less labor intensive. Even though NIR separations might not be 100% accurate, they can provide early generation populations that are highly enriched in lines with the waxy phenotype. Breeders will thereby have access to large populations of waxy lines, from which they can make additional selections based on disease resistance, grain yield, etc. All breeding programs ultimately rely on the presence of sufficient numbers of progeny from which one can identify the few lines bearing all desired traits. It also should be noted that conventional methods of identifying waxy lines are not foolproof. In some plant species such as sorghum, waxy mutants exist that produce the granule-bound starch synthase, albeit in nonfunctional form (Pedersen et al 2005). SDS-PAGE evaluation would miss such mutants, while NIR procedures, based on grain amylose fraction, will not.

## CONCLUSIONS

This study examined the potential of near-infrared spectroscopy of single tetraploid (durum) wheat kernels for classification by waxy allele: waxy, partial waxy (*wx-A1* null or *wx-B1* null), or wild type. Based on 48 single kernel spectra of 47 durum wheat lines grown in two successive seasons, NIR-based identification of the waxy genotype was often >95% correct, whereas the classification accuracy of the three non-waxy genotypes was very low, typically <50%. Fortunately, in waxy breeding programs, the identification of the waxy seed is the most critical. Therefore, this technique stands as a demonstrated tool for immediate use in breeding programs and eventual use in commerce for maintenance of identity preservation of waxy cultivars.

## ACKNOWLEDGMENTS

We give thanks to B. Stetzler (ARS, Beltsville) for spectral data collection and to B. Stetzler and S. Mookherji (ARS, Beltsville) for combustion protein analyses.

## LITERATURE CITED

AACC International. 2000. Approved Methods of the American Association of Cereal Chemists, 10th Ed. Method 46-30. The Association: St. Paul, MN.

Crosbie, G. B. 1991. The relationship between starch swelling properties, past viscosity and boiled noodle quality in wheat flours. *J. Cereal Sci.* 13:145-150.

Delwiche, S. R. 2003. Classification of scab- and other mold-damaged wheat kernels by near-infrared reflectance spectroscopy. *Trans. ASAE* 46:731-738.

Delwiche, S. R., and Graybosch, R. A. 2002. Identification of waxy wheat by near-infrared reflectance spectroscopy. *J. Cereal Sci.* 35:29-38.

Dowell, F. E., Maghirang, E. B., Graybosch, R. A., Baenziger, P. S., Baltensperger, D. D., and Hansen, L. E. 2006. An automated single-kernel near-infrared trait selection system. *unpublished*

Epstein, J., Morris, C. F., and Huber, K. C. 2002. Instrumental texture of white salted noodles prepared from recombinant inbred lines of wheat

differing in the three granule bound starch synthase (waxy) genes. *J. Cereal Sci.* 35: 51-63.

Fisher, R. A. 1936. The use of multiple measurements in taxonomic problems. *Ann. Eugenics* 7:179-188.

Flipse, E., Keetels, C. J. A. M., Jacobsen, E., and Visser, R. G. F. 1996. The dosage effect of the wildtype GBSS allele is linear for GBSS activity but not for amylose content: Absence of amylose has a distinct influence on the physico-chemical properties of starch. *Theor. Appl. Genet.* 92:121-127.

Graybosch, R. A. 1998. Waxy wheats: Origin, properties, and prospects. *Trends Food Sci Technol* 9:135-142.

Graybosch, R. A., Peterson, C. J., Hansen, L. E., Rahman, S., Hill, A., and Skerritt, J. H. 1998. Identification and characterization of U.S. wheats carrying null alleles at the *wx* loci. *Cereal Chem.* 75:162-165.

Graybosch, R. A., Souza, E., Berzonsky, W., Baenziger, P. S., and Chung, O. 2003. Functional properties of waxy wheat flours: Genotypic and environmental effects. *J. Cereal Sci.* 38:69-76.

Hoshino, T., Yoshikawa, R., Ito, S., Hatta, K., Nakamura, T., Yamamori, M., Hayakawa, K., Tanaka, K., Akashi, H., Endo, S., Tago, S., and Ishigami, S. 2000. Flour blends for breads, cakes, or noodles, and foods prepared from the flour blends. U.S. patent 6,042,867.

Klucinec, C. A., and Thompson, D. B. 1998. Fractionation of high-amylose maize starches by differential alcohol precipitation and chromatography of the fractions. *Cereal Chem.* 75:887-896.

Knutson, C. A., and Grove, M. J. 1994. Rapid method for estimation of amylose in maize starches. *Cereal Chem.* 71:469-471.

Nakamura, T., and Yamamori, M. 1995. Production of waxy (amylose-free) wheats. *Mol. Gen. Genet.* 248:253-259.

Nakamura, T., Vrinten, P., Saito, M., and Konda, M. 2002. Rapid classification of partial waxy wheats using PCR-based markers. *Genome* 45:1150-1156.

Pedersen, J. F., Bean, S. R., Graybosch, R. A., Park, S. H., and Tilley, M. 2005. Characterization of waxy grain sorghum lines in relation to granule-bound starch synthase. *Euphytica* 144:151-156.

Rahman, S., Kosarhashemi, B., Samuel, M. S., Hill, A., Abbott, D. C., Skerritt, J. H., Preiss, J., Appels, R., and Morell, M. K. 1995. The major proteins of wheat endosperm starch granules. *Aust. J. Plant Physiol.* 22:793-803.

Reddy, I., and Seib, P. A. 2000. Modified waxy wheat starch compared to modified waxy corn starch. *J. Cereal Sci.* 31:25-29.

Shannon, J. C., and Garwood, D. L. 1984. Genetics and physiology of starch development. Pages 26-86 in: *Starch: Chemistry and Technology*, 2nd Ed. R. L. Whistler, J. N. BeMiller, and E. F. Paschall, eds. Academic Press: New York.

Sharma, R., Sisson, M. J., Rathjen, A. J., and Jenner, C. F. 2002. The null-4A allele at the waxy locus in durum wheat affects pasta cooking quality. *J. Cereal Sci.* 35:287-297.

Vignaux, N., Doehlert, D. C., Hegstad, J., Elias, E. M., McMullen, M. S., Grant, L. A., and Kianian, S. F. 2004. Grain quality characteristics and milling performance of full and partial waxy durum lines. *Cereal Chem.* 81:377-383.

Vignaux, N., Doehlert, D. C., Elias, E. M., McMullen, M. S., Grant, L. A., and Kianian, S. F. 2005. Quality of spaghetti made from full and partial waxy durum wheat. *Cereal Chem.* 82:93-100.

Yamamori, M., and Quynh, N. T. 2000. Differential effects of *Wx-A1*, *-B1* and *-D1* protein deficiencies on apparent amylose content and starch pasting properties in common wheat. *Theor. Appl. Genet.* 100:32-38.

Zhao, X. C., and Sharp, P. J. 1996. An improved 1-D SDS-PAGE method for the identification of three bread wheat 'waxy' proteins. *J. Cereal Sci.* 23:191-193.

Zhao, X. C., Batey, I. L., Sharp, P. J., Crosbie, G., Barclay, I., Wilson, R., Morell, M. K., and Appels, R. 1998. A single genetic locus associated with starch granule properties and noodle quality in wheat. *J. Cereal Sci.* 27:7-13.

[Received October 31, 2005. Accepted February 17, 2006.]