

Identification of Wheat Lines Possessing the 1AL.1RS or 1BL.1RS Wheat-Rye Translocation by Near-Infrared Reflectance Spectroscopy

Stephen R. Delwiche,^{1,2} Robert A. Graybosch,³ and C. James Peterson^{3,4}

ABSTRACT

Cereal Chem. 76(2):255–260

Wheat-rye chromosomal translocations, particularly those involving the short arm of rye chromosome 1R, have been used during the past 25 years to instill resistance to plant pathogens and insects and improve the hardness, adaptation, and yield of wheat. Unfortunately, the presence of the 1AL.1RS or 1BL.1RS rye translocations in wheat has been shown to impart inferior dough handling and baking characteristics. Although numerous analytical techniques (e.g., HPLC, monoclonal antibody tests, high-performance capillary electrophoresis) have been developed for detecting these translocations, the complexity of the analytical procedures restricts their use to research and analytical laboratories. The purpose of this study was to examine the potential of diffuse reflectance near-infrared spectroscopy, a well-accepted technique in the grain industry, for detecting IRS-containing genotypes. This research used three independent groups of wheat

samples, ranging in genetic diversity from sister lines derived from IRS breeding populations to commercial cultivars. Based on the diffuse reflectance spectra (1,100–2,500 nm) of flour, partial least squares (PLS) models, through cross-validation, exhibited misclassification rates as low as 0%, particularly for commercial cultivars. Misclassification rates for corresponding, but separate, test sets were as low as 1%. When the same modeling procedure was applied to samples of more closely related genetic backgrounds, cross-validation misclassification rates rose to 15–20%. Most problematic were samples that were heterogeneous for IRS such as the cultivar Rawhide. Incorporating heterogeneous samples into a calibration equation improved the classification accuracy of these samples but diminished the prediction accuracy of nonheterogeneous samples.

Wheat-rye (*Triticum aestivum* L and *Secale cereale* L) chromosomal translocations, particularly those involving the short arm of rye chromosome 1R, are used in the United States and elsewhere to instill resistance to plant pathogens and insects and improve the hardness, adaptation, and yield of wheat (Zeller 1973, Lukaszewski 1990). In the United States, translocations of this chromosome commonly fall into two categories: 1AL.1RS, derived from the Oklahoma germplasm line Amigo, and 1BL.1RS, derived from Russian cultivars Kavkaz and Aurora. Resistance to powdery mildew, leaf rust, stem rust, and greenbug (*Schizaphis graminum* Rondani) are conferred by 1AL.1RS, while the same types of resistance, with exception of greenbug, are conferred by 1BL.1RS. Rye translocations in wheat have been shown to impart inferior dough handling and baking characteristics in hard wheats (Zeller and Hsam 1983, Dhaliwal et al 1987, Graybosch et al 1993a, Fenn et al 1994) and, to a lesser extent, soft wheats (McKendry et al 1996). Dough stickiness, lower mixing tolerance, and lowered loaf volume are associated with 1BL.1RS, while less severe conditions arise from 1AL.1RS (Graybosch et al 1993a).

Numerous techniques have been investigated for detecting IRS in wheat lines. These include traditional electrophoretic methods (Dhaliwal et al 1988, Johansson et al 1994), HPLC (Lookhart et al 1991), γ -gliadin-specific monoclonal antibodies (Howes et al 1989), secalin-specific monoclonal antibodies (Graybosch et al 1993b, Andrews et al 1996), and high-performance capillary electrophoresis (HPCE) (Lookhart et al 1996). While these methods have been successful at detecting IRS, and in some cases, at discriminating between 1AL.1RS and 1BL.1RS (Andrews et al 1996, Lookhart et al 1996), the complexity of the analytical procedures restricts their use to research

and analytical laboratories. The purpose of this study was to examine the potential of diffuse reflectance near-infrared (NIR) spectroscopy, a technology commonly used by the trade, regulatory, and processing industries, for detection of IRS-containing genotypes. This feasibility study has involved the use of three independent groups of wheat samples, ranging from sister lines derived from IRS breeding populations to commercial cultivars.

MATERIALS AND METHODS

Samples

Group 1. Thirty genotypes of hard red winter wheats representing popular releases or promising lines for the Great Plains, were grown at eight locations (Alliance, Clay Center, Grant, Lincoln, McCook, North Platte, Scottsbluff, and Sidney) throughout Nebraska during the 1989–90 and 1990–91 crop years. The second year Alliance crop was absent because of winter-kill, while Sidney produced separate sets of irrigated and nonirrigated wheat for both years. Twenty-two genotypes (Abilene, Arapahoe, Arkan, Centura, Centurk 78, Cheyenne, Chisholm, Cimarron, Eagle, Karl, Lamar, Lancota, Newton, N86L022, N86L177, NE83407, Plainsman V, Red Chief, Redland, Scout 66, TAM-101, and Vona) did not contain IRS, while five (TAM-107, TAM-200, TAM-202, Century, and TXGH12588) contained 1AL.1RS derived from Amigo, two (Siouxland, KS801072) contained 1BL.1RS derived from Kavkaz, and one (Rawhide, also derived from Kavkaz) was heterogeneous for 1BL.1RS with theoretically one-half of the kernels possessing the translocation. Using a randomized complete block design, each genotype-location-year (considered a sample during NIR modeling) was grown in five blocks; however, only two of these blocks were used in the present study. The identity of each sample was confirmed by HPLC (Graybosch et al 1996). Equal numbers ($n = 113$ each) of IRS (Rawhide included) and non-IRS samples were selected from the same block so that all spectrally acceptable (i.e., sufficient material for NIR scanning, similar fineness of flour as determined by cursory human inspection) IRS samples were used during model development. Non-IRS samples were selected from a larger number of candidates (because of the inherent imbalance in the pool of IRS and non-IRS samples) based on the criterion that the flour protein content distributions of IRS and non-IRS subsets be as similar as possible. Upon selection, the average absolute difference in protein content between a IRS sample and its non-IRS counterpart was 0.034%. The purpose for matching these distributions was to minimize the chance that classification would be based on protein content. All

¹ U.S. Department of Agriculture, Agricultural Research Service, 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.

² Corresponding author. E-mail: sdelwiche@asrr.arsusda.gov

³ USDA-ARS, Department of Agronomy, University of Nebraska East Campus, Lincoln, NE.

⁴ Current address: Department of Crop and Soil Science, Oregon State University, Corvallis, OR.

spectrally acceptable (same rules as above) samples from the second block ($n = 109$ for IRS, 311 for non-IRS) were reserved for model testing. Numbers of each genotype within model development (i.e., cross-validation set) and test sets are summarized in Table I. Additional information on the origin, growth, and protein composition of these samples is provided in Graybosch et al (1996).

Group II. Nine populations of hard red winter wheat, consisting of six sister lines per population, were grown in unreplicated plots at the University of Nebraska Agronomy Farm, Lincoln, NE, and harvested in 1990. The pedigree of each population is listed in Table II. For each population, at least one parent possessed either 1BL.1RS or 1AL.1RS. Half of the lines within a population possessed IRS (protein content range of 12.4–15.7%); the other half did not (protein content range of 13.1–15.8%). During NIR modeling, each line was considered as a unique sample. All lines ($n = 54$) were used during cross-validation.

Group III. Experimental hard red winter wheat lines, derived from seven distinct cultivars or breeders lines (Arapahoe, GRS1201, N86L177, Redland, Siouland, TAM-202, and Raeder-s), were grown during 1996–97 at McCook, NE (Table III). Each of five unique pedigrees contained between three and 12 sibling lines, yielding a total of 40 lines of progeny, of which, 25 lines possessed IRS (either 1AL or 1BL), 10 did not possess IRS, and five were heterogeneous (50:50) for IRS. Additionally, six of the seven parental cultivars or breeders lines (Raeder-s excluded), with half possessing IRS, were grown and harvested. Plots were randomized in three complete blocks, with the blocks contributing equally to the cross-validation set. During NIR modeling, each plot represented a sample. To maintain an equal number of non-IRS samples (13 lines altogether) and IRS samples within each block, only 10 of the possible 25 IRS sibling lines, plus the three IRS parental lines, were used

during cross-validation. Protein content ranges for non-IRS and IRS cross-validation samples were 9.7–13.9 and 10.5–14.6%, respectively. The remaining 15 IRS sibling lines and the five heterogeneous sibling lines, with protein content ranges of 10.2–14.3 and 10.4–15.1%, respectively, were reserved for testing the NIR model. The exact assignment of all sibling lines is detailed in Table III.

Procedure

Seeds (35–50 g) from each sample were tempered to 15% moisture (wb), then milled on an experimental mill (Quadraplex, Brabender, South Hackensack, NJ). Flour was sifted through a No. 70 U.S. standard sieve on a shaker (Strand Shaker Co., Minneapolis, MN) at 225 rpm for 90 sec. Flour samples were stored at 0–5°C in 0.1-mm thick resealable plastic bags.

For Group I, protein content ($N \times 5.7$) was measured in duplicate on 150-mg portions from each flour sample by the combustion technique using a nitrogen analyzer (FP-428, Leco Mfg. Co., St. Joseph, MI). Protein content reported on an as-is moisture basis (≈ 13 –14%, wb) was subsequently used when forming the Group I cross-validation set, as described earlier. Protein contents of Group II and Group III samples, while similarly measured, were not used during data set structuring because of the smaller number of samples in these groups.

NIR log(1/R) spectra (1,100–2,498 nm, 32 repetitive scans) were collected at 2-nm intervals on two successive fillings of a standard ring cell (≈ 4 g/filling) with each sample. Spectra were collected on a spectrophotometer (model 6500, NIRSystems, Silver Spring, MD) equipped with a spinning sample module. Diffuse reflectance readings were referenced to those from a ceramic disk. The spectrum of each sample was the mathematical average of the log(1/R) values of the respective fillings.

Modeling

Classification modeling was performed by partial least squares (PLS) regression (PLSplus IQ software, Galactic Industries, Salem, NH). Non-IRS samples were assigned the value -0.5 . Similarly, 1AL.1RS and 1BL.1RS samples were assigned the value 0.5 . The modeling approach varied slightly between data groups, with details for each group described separately. For Group I, one-sample-out cross-validation was performed on the member set ($n = 226$). Specifically, during cross-validation each sample's spectrum was temporarily removed and PLS regressions (1–20 factors) were formed with the remaining samples ($n - 1$). Each regression equation at each factor level was applied to the removed sample. Equation predictions at each factor level were accumulated over n cross-validation

TABLE I

Number of Group I Cultivar and Breeding Line Samples Used in Cross-Validation and Testing of Near-Infrared Reflectance Models for Detecting 1AL.1RS or 1BL.1RS Wheat-Rye Translocation

| Cultivar or Line ^a | Cross-Validation Set (n) | Test Set (n) |
|-------------------------------|------------------------------|------------------|
| Non-IRS | | |
| Abilene | 4 | 16 |
| Arapahoe | 9 | 16 |
| Arkan | 6 | 14 |
| Centura | 6 | 16 |
| Centurk78 | 5 | 13 |
| Cheyenne | 8 | 16 |
| Chisholm | 6 | 12 |
| Cimarron | 4 | 12 |
| Eagle | 4 | 16 |
| Karl | 2 | 16 |
| Lamar | 5 | 16 |
| Lancota | 4 | 16 |
| N86L022 | 2 | 16 |
| N86L177 | 5 | 16 |
| NE83407 | 11 | 15 |
| Newton | 8 | 16 |
| Plainsman V | 3 | 16 |
| RedChief | 1 | 0 |
| Redland | 5 | 11 |
| Scout66 | 6 | 13 |
| TAM-101 | 5 | 16 |
| Vona | 4 | 13 |
| IRS | | |
| Century ^{AL} | 16 | 15 |
| KS801072 ^{BL} | 13 | 11 |
| Rawhide ^{BL} | 16 | 16 |
| Siouland ^{BL} | 13 | 13 |
| TAM-107 ^{AL} | 13 | 13 |
| TAM-200 ^{AL} | 13 | 12 |
| TAM-202 ^{AL} | 13 | 13 |
| TXGH12588 ^{AL} | 16 | 16 |

^a Type of translocation: ^{AL} = 1AL.1RS, ^{BL} = 1BL.1RS (with Rawhide heterogeneous for the IRS translocation, being $\approx 50:50$ mixture of 1BL.1RS and non-IRS kernels).

TABLE II
Pedigrees of Group II Sample Populations Used in Cross-Validation and Testing of Near-Infrared Reflectance Models for Detecting 1AL.1RS or 1BL.1RS Wheat-Rye Translocation

| Population | Pedigree ^a | Type of IRS Translocation in Siblings ^b |
|------------|--|--|
| 1 | Siouland*/Plainsman V | 1BL.1RS |
| 2 | Siouland*/Sumner | 1BL.1RS |
| 3 | New 66.1.8/TAM-107* | 1AL.1RS |
| 4 | New 66.1.8/Siouland* | 1BL.1RS |
| 5 | OK81322*/4/Atlas 66/Cmn/2/Nebraska Restorer/3/Aurora* | 1BL.1RS |
| 6 | NE7060/Siouland* | 1BL.1RS |
| 7 | Aurora*/NE701154/3/NE7060/2/Rannaya 12/Bezostaya 4/4/Agrotriticum \times Triticum aestivum | 1BL.1RS |
| 8 | Lindon/NE7060/2/TAM-107* | 1AL.1RS |
| 9 | Lindon/NE7060/2/NE80413* | 1BL.1RS |

^a Pedigree notation according to Purdy et al (1968). Asterisks indicate source of IRS translocation.

^b Six sibling line samples per population, of which, three do not possess the IRS translocation and three possess the translocation indicated.

cycles. The simple rule for classification was that a prediction lying beyond the midpoint (>0 for non-1RS samples and <0 for 1RS samples) was deemed a misclassification. Cross-validation performance was defined as the percentage of misclassified samples. Percentages were determined separately for non-1RS and 1RS samples. By inspection, the optimal number of factors was chosen as that which produced either the lowest average misclassification percentage or nearly the lowest percentage, but with fewer factors than the former and with a better balance between non-1RS and 1RS misclassification percentages. The optimal model was subsequently applied to the test set and the misclassification percentages were determined by the same midpoint rule.

For Group II, modeling also was performed by PLS one-sample-out, 1–20 factors cross-validation. Because of the smaller overall number of Group II samples ($n = 54$), all were used during cross-validation with none allocated for a separate test set. For modeling Group III samples ($n = 78$), the cross-validation procedure was slightly different from that of Groups I and II in that all three field replicates of a line were removed during a cross-validation cycle. The Group III 1RS samples not used during cross-validation were placed in a test set ($n = 60$). As in Group I modeling, the optimal PLS model determined by cross-validation was applied to these test samples.

RESULTS AND DISCUSSION

Cross-validation results, expressed as the percentage of misclassified samples at PLS factors of 1–20 are depicted in Fig. 1 with separate graphs for Groups I–III. Each graph contains horizontal dashed lines at 0 and 50% misclassification rates that respectively represent the conditions of a perfect classification model and a model that is no better on average than that predicted by random event. For Group I (Fig. 1A), the percentage of misclassified samples decreased as the number of factors increased to 14. Beyond that, the misclassification rate from cross-validation was relatively steady at 2–4%. From eight factors onward, the misclassification rates for 1RS and non-1RS samples were equivalent, with 14 factors considered optimal. At 14 factors, only three 1RS and two non-1RS samples from a combined set of 226 were misclassified. All three 1RS samples were from the heterogeneous cultivar Rawhide. A subsequent model, based on the same cross-validation samples, but without the 16 Rawhide samples and their corresponding protein-matched non-1RS samples, demonstrated that perfect classification during cross-validation was obtained from 15 factors onward.

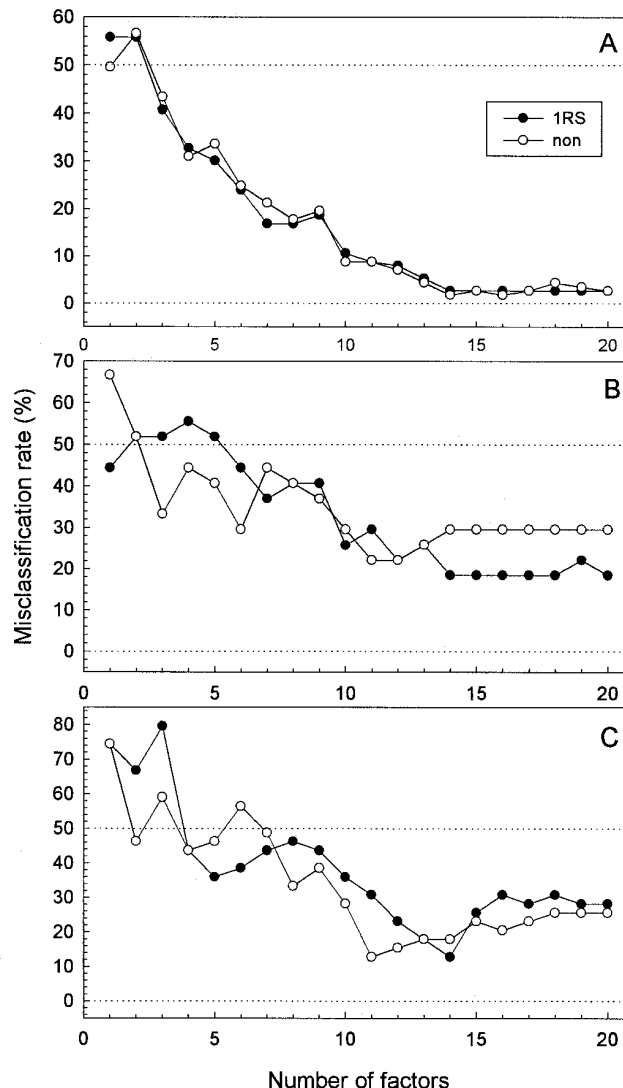


Fig. 1. One-sample-out cross-validation results of partial least squares models expressed as percentage of misclassified samples with (●) and without (○) 1RS translocation. A–C: Groups I–III, respectively. (See Tables I and III.)

TABLE III
Pedigrees and Assignment of Group III Samples Used in Cross-Validation and Testing of Near-Infrared Reflectance Models for Detecting 1AL.1RS or 1BL.1RS Wheat-Rye Translocation

| Parent or Pedigree ^a | Number of Parent or Sibling Lines | Type of 1RS and Number of Sibling Lines Possessing Translocation for Pedigrees | Assignment ^b |
|---------------------------------|-----------------------------------|--|--|
| Arapahoe | 1 | None | <i>xv</i> |
| GRS1201 | 1 | 1AL.1RS | <i>xv</i> |
| N86L177 | 1 | None | <i>xv</i> |
| Redland | 1 | None | <i>xv</i> |
| Siouxland | 1 | 1BL.1RS | <i>xv</i> |
| TAM-202 | 1 | 1AL.1RS | <i>xv</i> |
| Raeder-s/Siouxland/2/N861022 | 6 | 1 none, 3 with 1BL.1RS, 2 heterogeneous for 1BL.1RS | <i>xv</i> for the none and 1 1BL.1RS siblings; <i>t</i> for 2 1BL.1RS and the heterogeneous siblings |
| Siouxland/N86L177/2/N861177 | 7 | 2 none, 3 with 1BL.1RS, 2 heterogeneous for 1BL.1RS | <i>xv</i> for the none and 2 1BL.1RS siblings; <i>t</i> for 1 1BL.1RS and the heterogeneous siblings |
| Redland/GRS1201 | 12 | 6 none, 5 with 1AL.1RS, 1 heterogeneous for 1AL.1RS | <i>xv</i> for the none and 1AL.1RS siblings; <i>t</i> for the heterogeneous sibling |
| Redland/TAM-202 | 3 | 3 with 1AL.1RS | <i>xv</i> for 1 1AL.1RS sibling; <i>t</i> for the remaining siblings |
| GRS1201/TAM-202 | 12 | 1 none, 11 with 1AL.1RS | <i>xv</i> for the none and 1 1AL.1RS siblings; <i>t</i> for the remaining siblings |

^a Pedigree notation according to Purdy et al (1968).

^b Refers to set assigned during partial least squares modeling: *xv* = cross-validation set, *t* = test set.

Misclassification rates for Group II cross-validation samples decreased, albeit erratically, from 1 to 12 factors (Fig. 1B). Noticeable differences in the misclassification rate between IRS and non-IRS samples were apparent at 1–11 factors with the relative order dependent on the factor number. At 12 factors, the misclassification rate was equal (22%) between IRS and non-IRS and was lower than the average misclassification rates at all other factors. Beyond 12 factors, the misclassification rate increased on average and also became different between IRS and non-IRS samples. With the 12-factor model deemed as optimal, an examination of the misclassified samples revealed that populations 6 (NE7060/Siouxland) and 8 (Lindon/NE7060/2/TAM-107) (Table II) contained the highest number of misclassifications (4 of 6 and 3 of 6, respectively), while populations 1, 3, 4, 7, and 9 each had one misclassified sample, and populations 2 and 5 had no misclassified samples. No apparent difference was observed for the misclassification rates of 1AL. IRS vs. 1BL. IRS samples. Subsequent analysis of the misclassified samples by SDS-PAGE revealed that two of the IRS samples (from populations 6 and 8) were actually heterogeneous for IRS in the same manner as Rawhide in Group I. In this case, however, removal of these two samples from cross-validation did not result in the reduction of misclassified samples as it had in Group I.

Cross-validation misclassification rates for Group III samples (Fig. 1C) were similar to those of Group II. Again, the relative order

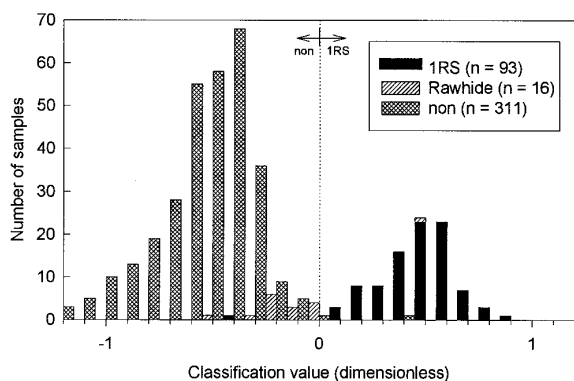


Fig. 2. Histogram of Group I test set classification predictions with and without IRS translocation from application of the 15-factor partial least squares model developed without Rawhide samples. One non-IRS value (−1.47) is not plotted.

of the misclassification rate for IRS and non-IRS samples changed over the range of 1–12 factors. At 13 factors, the percentage of misclassified samples was equal among the subgroups and only slightly higher (17.9%) than the average at 14 factors (15.4%). Therefore 13 factors was selected as optimal for this group.

Model testing was performed on Group I and III samples. Test set misclassification percentages along with corresponding cross-validation values are summarized in Table IV. For Group I, application of the 14-factor PLS model, inclusive of Rawhide, to that group's test set resulted in 2 of 109 (1.8%) IRS samples and 5 of 311 (1.6%) non-IRS samples misclassified, rates that were very similar to those estimated by cross-validation (2.7 and 1.8% for IRS and non-IRS samples, respectively). Misclassifications were not associated with protein content or geographical origin (data not shown). Although less apparent in trend than during cross-validation, one of the two misclassified IRS samples was from Rawhide. With application of the 15-factor Rawhide-excluded model to the test samples, the number of misclassified non-IRS samples dropped to 2 of 311 (0.6%), while the number of non-Rawhide IRS misclassified samples remained the same. However, because of the absence of Rawhide in the calibration of this second model, 15 of 16 (93.8%) Rawhide test samples were misclassified. A histogram of this model's raw output test-set values (Fig. 2) shows that the Rawhide samples were primarily clustered between −0.5 and 0.0. For unknown reasons, unless incorporated into a calibration, Rawhide had the tendency to be classified as a non-IRS wheat. However, if the calibration set included Rawhide, then the classification error rate for this cultivar decreased, though to the detriment of the classification accuracy for the other cultivars.

For the non-IRS samples, output values ranged from −1.47 to 0.40, with 5th and 95th percentiles being −0.92 and −0.21, respectively (Fig. 2). Likewise, IRS values ranged from −0.42 to 0.83, with 0.12 and 0.69, respectively, representing the 5th and 95th percentiles. Without the outlying (and misclassified) IRS sample, the IRS values were normally distributed as determined by a Shapiro-Wilk test (Proc Univariate, SAS Institute, Cary, NC).

For Group III samples, the 13-factor optimal model applied to the 60 IRS test samples resulted in the misclassification of six samples, with half of these being heterogeneous for IRS (i.e., three of the 15 heterogeneous siblings misclassified) (Table IV). Compared to Group I samples, misclassification of heterogeneous lines in Group III was not as problematic, though classification accuracy for the non-heterogeneous IRS samples (three of 45 misclassified) was poorer than that for the corresponding Group I samples (one of 93 misclassified).

TABLE IV
Optimal Near-Infrared Reflectance Partial Least Squares Classification Performance for Identifying IRS Translocation Among Models for Group I and III Wheat Samples^a

| Group, Model ^b | Subgroup | Cross-Validation Set | | Test Set | |
|---------------------------|--|----------------------|-------------------|----------|-------------------|
| | | n | Misclassified (%) | n | Misclassified (%) |
| I, 14-factor | Non-IRS | 113 | 1.8 | 311 | 1.6 |
| | IRS, excluding Rawhide | 97 | 0.0 | 93 | 1.1 |
| | Rawhide | 16 | 18.8 | 16 | 6.2 |
| I, 15-factor | Non-IRS | 97 | 0.0 | 311 | 0.6 |
| | IRS, excluding Rawhide | 97 | 0.0 | 93 | 1.1 |
| | Rawhide | ... | ... | 16 | 93.8 |
| III, 13-factor | Non-IRS | 39 | 17.9 | ... | ... |
| | IRS, excluding heterogeneous siblings | 39 | 17.9 | 45 | 6.7 |
| | Heterogeneous siblings | ... | ... | 15 | 20.0 |
| I and III, 16-factor | Non-IRS of Group I | 39 | 0.0 | 311 | 1.0 |
| | IRS of Group I, excluding Rawhide | 39 | 0.0 | 93 | 7.5 |
| | Rawhide of Group I | ... | ... | 16 | 81.2 |
| | Non-IRS of Group III | 39 | 7.7 | ... | ... |
| | IRS of Group III, excluding heterogeneous siblings | 39 | 12.8 | 45 | 2.2 |
| | Heterogeneous siblings of Group III | ... | ... | 15 | 40.0 |

^a Groups defined in Tables I and III.

^b Optimal number of factors in partial least squares model developed from one-sample-out cross-validation of samples within the groups indicated.

To examine the effect of including a new crop year on model accuracy, an intergroup cross-validation set consisting of equal numbers of Group I and Group III samples was used to develop a similar classification model. Thirty-nine IRS samples (excluding Rawhide) and an equal number of non-IRS samples were drawn from the pool of Group I cross-validation samples using the criterion that their protein contents matched those of the 78 Group III cross-validation samples as best as possible (average absolute difference [Group I – Group III] = 0.266% protein). Cross-validation and test results are also shown in Table IV. Optimal model performance occurred at 16 factors, which was one and three factors more than the individual Group I and III models, respectively. Although cross-validation performance either remained the same (Group I) or improved slightly (Group III) with use of the multiyear calibration, test set performance was less predictable. Except for the samples of Rawhide, the test set classification accuracy for Group I samples declined with use of the multiyear model; the opposite trend occurred with the Group III samples.

To identify wavelengths within the PLS equations with the greatest influence on IRS classification accuracy, the scaled regression vectors corresponding to the optimal PLS model for Groups I (Rawhide excluded), II, and III samples are shown in Fig. 3. Scaling consisted of, on a wavelength by wavelength basis, multiplying each regression coefficient by the mean spectral value of the cross-validation set, then dividing all terms in this product spectrum by the product term of greatest absolute value. Note that with exception of the final step, this procedure is similar to that used for prediction of the classification value of an unknown sample. Specifically, an unknown's spectrum is substituted for the mean spectrum, whereupon the sum of all product terms and an additional constant yields the classification value. All three graphs in Fig. 3 indicate the lack of importance of wavelengths between 1,000 and 1,400 nm. Corresponding to the region of second overtones of CH and NH vibrational frequencies (Osborne and Fearn 1986), these bands are inherently much weaker than their first overtone counterparts and therefore are not useful for classification. Wavelengths possessing local absolute maxima that are common among all three groups are identified by circles (Fig. 3A–C). The three most apparent common bands occurred at 1,686, 2,264, and 2,300 nm, which correspond to vibrational frequencies for the CH stretch first overtone, OH stretch + OH deformation combination or OH stretch + CC stretch combination, and CH stretch + CH deformation combination, respectively (Osborne and Fearn 1986). Of these three wavelengths, the shortest (1,700–1,720 nm peak range) and longest (2,290–2,310 nm peak range) were observed in gluten, starch, and lipid fractions of wheat (Law and Tkachuk 1977).

Contrary to the wavelengths possessing absolute maxima common to all groups, one wavelength in Group I (1,760 nm) and another in Group III (2,054 nm, identified by squares in Fig. 3A,C) possessed a local absolute maximum in regression coefficient that was unique to its group. Cereal lipids have been reported as having an absorption band at 1,760 nm (Williams and Norris 1988), which suggests that the assumed larger variation of lipid concentration (because of the larger genetic variation) of Group I samples might be used in the regression equation. Less clear is an explanation for the uniqueness of the Group III 2,054-nm regression term, a wavelength that is typically attributed to protein and arises from the NH symmetrical stretch + amide II combination or NH asymmetrical stretch + amide III combination bands (Osborne and Fearn 1986).

Success in the detection of IRS by NIR spectroscopy is corroborated in light of known biochemical effects of the translocation in wheat. Working with seven of the nine populations within Group II (omitting the 1AL.IRS populations), but with a greater number of sister lines per population (ranging from 32 to 67 lines), Lee et al (1995) determined that protein content was unaffected by the presence of 1BL.IRS. However, the concentration of glutenin was generally lower in IRS lines than in non-IRS lines. Similarly, SDS sedimentation volume values were lower in IRS lines. Recently,

the feasibility of measuring glutenin content or SDS sedimentation volume by NIR spectroscopy has been demonstrated (Delwiche et al 1998). Hence, it is likely that some of the same absorption bands used in the PLS models for these two biochemical parameters are also relied upon in the PLS models for IRS classification. If, in fact, the models identify IRS indirectly due to the lower glutenin content, the misclassifications would become less of a concern in the commercial application of this procedure. Misclassifications might represent IRS wheats with higher than expected glutenin (perhaps due to a favorable environment) or non-IRS wheats with lower glutenin. In either case, the effects on quality are minimized when glutenin content is the critical factor that determines quality. Low glutenin non-IRS would be binned with IRS, whereas high glutenin IRS would be binned with non-IRS, thus maintaining separation of wheats by dough-handling and baking performance.

CONCLUSIONS

Diffuse reflectance NIR spectroscopy (1,100–2,500 nm) on wheat flour can be used to identify genotypes possessing the 1AL.IRS or 1BL.IRS wheat-rye translocation. Based on PLS regression with a midpoint value cutoff criterion, sample classification accuracy was demonstrated in the range of 78–99%, with the actual accuracy value dependent on the genetic similarities of the wheat lines tested. Wavelengths <1,400 nm were not useful in the classification models. Commercial cultivars and advanced breeders lines were most accurately classified, while separation of near-isogenic lines was the most difficult. Classification of cultivars that are heterogeneous for the IRS translocation may be particularly problematic.

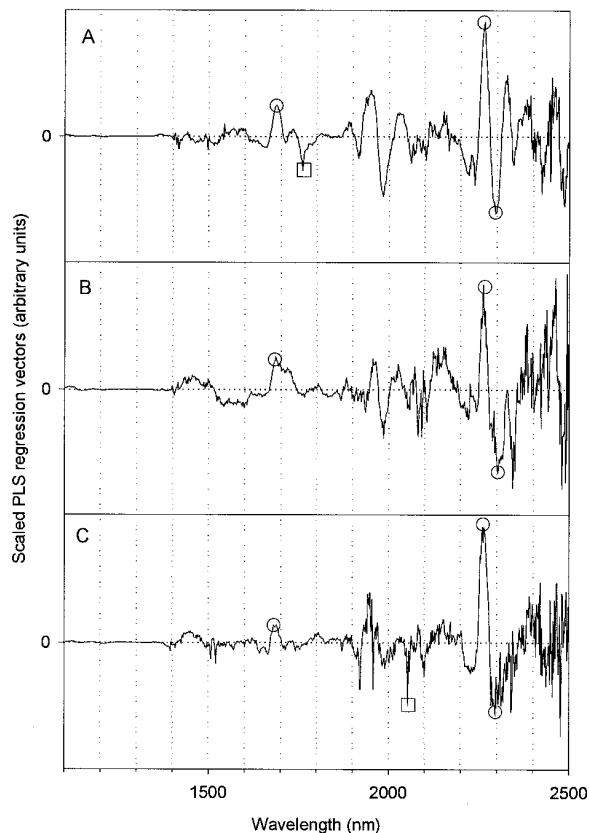


Fig. 3. Scaled partial least squares regression vectors corresponding to optimal models. **A**, Group I, 15-factor, without Rawhide; **B**, Group II, 12-factor; and **C**, Group III, 13-factor. Scaling was performed by multiplying each regression coefficient by that wavelength's corresponding mean (of cross-validation set) spectral value then dividing all products by the product with the greatest absolute value. ○ = peaks common to all groups; □ = peak unique to a particular group.

ACKNOWLEDGMENTS

Partial support of this work was by a grant from Anheuser-Busch Companies, St. Louis, MO. We are thankful for the technical assistance of Joyce Shaffer (Beltsville) and Vern Hansen (Lincoln).

LITERATURE CITED

- Andrews, J. L., Blundell, M. J., and Skerritt, J. H. 1996. Differentiation of wheat-rye translocation lines using antibody probes for *Gli-B1* and *Sec-1*. *J. Cereal Sci.* 23:61-72.
- Delwiche, S. R., Graybosch, R. A., and Peterson, C. J. 1998. Predicting protein composition, biochemical properties, and dough handling properties of hard red winter wheat flour by near-infrared reflectance. *Cereal Chem.* 75:412-416.
- Dhaliwal, A. S., Mares, D. J., and Marshall, D. R. 1987. Effect of 1B/1R chromosome translocation on milling and quality characteristics of bread wheats. *Cereal Chem.* 64:72-76.
- Dhaliwal, A. S., Mares, D. J., Marshall, D. R., and Skerritt, J. H. 1988. Protein composition and pentosan content in relation to dough stickiness of 1B/1R translocation wheats. *Cereal Chem.* 65:143-149.
- Fenn, D., Lukow, O. M., Bushuk, W., and DePauw, R. M. 1994. Milling and baking quality of 1BL/1RS translocation wheats. I. Effects of genotype and environment. *Cereal Chem.* 71:189-195.
- Graybosch, R. A., Peterson, C. J., Hansen, L. E., Worrall, D., Shelton, D. R., and Lukaszewski A. J. 1993a. Comparative flour quality and protein characteristics of 1BL/1RS and 1AL/1RS wheat-rye translocation lines. *J. Cereal Sci.* 17:95-106.
- Graybosch, R. A., Seo, Y. W., and Peterson, C. J. 1993b. Detection of wheat-rye chromosomal translocations using an antiseccalin monoclonal antibody. *Cereal Chem.* 70:458-463.
- Graybosch, R. A., Peterson, C. J., Shelton, D. R., and Baenziger, P. S. 1996. Genotypic and environmental modification of wheat flour protein composition in relation to end-use quality. *Crop Sci.* 36:296-300.
- Howes, N. K., Lukow, O. M., Dawood, M. R., and Bushuk, W. 1989. Rapid detection of the 1BL/1RS chromosome translocation in hexaploid wheats by monoclonal antibodies. *J. Cereal Sci.* 10:1-4.
- Johansson, E., Jonsson, J. O., Henriksson, P., Brismar, K., and Heneen, W. K. 1994. Comparison of electrophoretic methods for identification of the 1BL.1RS rye translocation in wheat, simultaneous screening for high-molecular-weight glutenins, and detection of chromosome deletions. *J. Genet. Breed.* 48:291-300.
- Law, D. P., and Tkachuk, R. 1977. Near infrared diffuse reflectance spectra of wheat and wheat components. *Cereal Chem.* 54:256-265.
- Lee, J. H., Graybosch, R. A., and Peterson, C. J. 1995. Quality and biochemical effects of a 1BL/1RS wheat-rye translocation in wheat. *Theor. Appl. Genet.* 90:105-112.
- Lookhart, G. L., Graybosch, R. A., Peterson, C. J., and Lukaszewski, A. J. 1991. Identification of wheat lines containing the 1BL/1RS translocation by high-performance liquid chromatography. *Cereal Chem.* 68:312-316.
- Lookhart, G. L., Bean, S. R., Graybosch, R. A., Chung, O. K., Morena-Sevilla, B., and Baenziger, S. 1996. Identification by high-performance capillary electrophoresis of wheat lines containing the 1AL.1RS and the 1BL.1RS translocation. *Cereal Chem.* 73:547-550.
- Lukaszewski, A. J. 1990. Frequency of 1RS.1AL and 1RS.1BL translocations in United States wheats. *Crop Sci.* 30:1151-1153.
- McKendry, A. L., Tague, D. N., Finney, P. L., and Miskin, K. E. 1996. Effect of 1BL.1RS on milling and baking quality of soft red winter wheat. *Crop Sci.* 36:848-851.
- Osborne, B. G., and Fearn, T. 1986. Pages 20-42 in: *Near Infrared Spectroscopy in Food Analysis*. Longman Scientific and Technical, Essex, UK.
- Purdy, L. H., Loegering, W. Q., Konzak, C. F., Peterson, C. J., and Allan, R. E. 1968. A proposed standard method for illustrating pedigrees of small grain varieties. *Crop Sci.* 8:405-406.
- Williams, P. C., and Norris, K. H. 1988. Qualitative applications of near-infrared reflectance spectroscopy. Pages 241-246 in: *Near-Infrared Technology in the Agricultural and Food Industries*. P. C. Williams and K. H. Norris, eds. Am. Assoc. Cereal Chem.: St. Paul, MN.
- Zeller, F. J. 1973. 1B/1R Wheat-rye chromosome substitutions and translocations. Pages 209-221 in: *Proceedings of the 4th International Wheat Genetics Symp.* E. R. Sears and L. M. Sears, eds. University of Missouri: Columbia, MO.
- Zeller, F. J., and Hsam, S. L. K. 1983. Broadening the genetic variability of cultivated wheat by utilizing rye chromatin. Pages 161-173 in: *Proceedings of the 6th International Wheat Genetics Symposium*, S. Sakamoto, ed. Kyoto University: Kyoto, Japan.

[Received July 1, 1998. Accepted December 4, 1998.]