Published January 14, 2016

Near-Infrared Calibration of Soluble Stem Carbohydrates for Predicting Drought Tolerance in Spring Wheat

Julia L. Piaskowski,* David Brown, and Kimberly Garland Campbell

ABSTRACT

Soluble stem carbohydrates are a component of drought response in wheat (*Triticum aestivum* L.) and other grasses. Near-infrared spectroscopy (NIR) can rapidly assay for soluble carbohydrates indirectly, but this requires a statistical model for calibration. The objectives of this study were: (i) to build a robust calibration between the NIR spectra and soluble carbohydrate concentration of ground wheat stems; and (ii) to determine whether soluble stem carbohydrates are correlated with yield rankings of drought-stricken wheat grown in the northwestern United States. Five spring wheat cultivars were grown in field trials conducted at six environments in the state of Washington varying in annual precipitation from 212 to 474 mm. Wheat stems were harvested from all plots at the onset of grain fill and assayed for NIR reflectance. Soluble stem carbohydrates were determined on a subset of the samples. The NIR data were calibrated to soluble stem carbohydrates using multiple linear regression, partial least squares regression, ridge regression with best linear unbiased prediction, random forest, least absolute shrinkage and selection operator (lasso), elastic net, and Bayesian lasso regression. Partial least squares regression provided the most accurate and reliable predictions for soluble carbohydrates. Correlations between soluble stem carbohydrates and grain yield were consistent across environments (r = 0.904 and ρ = 0.80). The effect of environment on the variation in response variables was lower for soluble carbohydrates than yield (5.93 and 71.7%, respectively) across environments. These data provide evidence that stem carbohydrates can aid in selecting cultivars with enhanced drought resilience.

Published in Agron. J. 108:285–293 (2016)
doi:10.2134/agronj2015.0173
Received 8 Apr. 2015
Accepted 26 Aug. 2015
Supplemental material available online

Much of the elite wheat germplasm developed for the Pacific Northwest states of Washington, Oregon, and Idaho possesses one or more physiological or anatomical traits that improve performance under water-limiting conditions (Blum, 2011). Due to low heritability of grain yield in drought environments, direct selection for drought tolerance is challenging (Lopez et al., 2003; Schramm et al., 2010). An important trait contributing to improved grain yield of wheat grown under drought is the accumulation of soluble carbohydrates in stems after anthesis but before grain fill (Zadoks growth stage 60) (Zadoks et al., 1974; Rebetzke et al., 2008). During stem elongation, wheat and other grasses store extra photoassimilates as nonstructural carbohydrates, mostly fructans, in the upper internodes of stems (Batten et al., 1993; Olien and Clark, 1993; Schnyder, 1993; Wardlaw and Willenbrink, 2000). These carbohydrates can be remobilized for grain filling during drought or heat stress, when photosynthetic activity is low. The concentration of these soluble stem carbohydrates has been positively correlated to grain yield under drought, contributing up to 50% of the variation for yield and comprising up to 40% of the stem dry weight (Gallagher et al., 1976; Brooks et al., 1982; Aggarwal and Sinha, 1984; Blum et al., 1991; Schnyder, 1993).

Soluble stem carbohydrates have been used as an indirect selector for grain yield in the early stages of wheat breeding when yield plots are not feasible. Previous research has demonstrated that the genotypic rankings for soluble stem carbohydrates in wheat in large part have been constant across drought-affected environments in Australia and have high broad-sense heritability (h^2 = 0.9) (Foulkes et al., 2002; Ruuska et al., 2006; Rebetzke et al., 2008). Predicting wheat yield under drought through soluble stem carbohydrate assays has become a standard practice in Australian wheat breeding programs (van Herwaarden and Richards, 2002). While post-anthesis soluble stem carbohydrates have been associated with grain yield under drought for some North American wheat

J.L. Piaskowski and D. Brown, Dep. of Crop and Soil Sciences, Washington State Univ., Pullman, WA 99164; and K.G. Campbell, USDA–ARS, Wheat Genetics, Quality, Physiology, and Disease Research Unit, Pullman, WA 99164. *Corresponding author (jpiaskowski@wsu.edu).

Abbreviations: BLUP, best linear unbiased prediction; BLR, Bayesian lasso regression; KM, Kubelka–Munk; lasso, least absolute shrinkage and selection operator; MLR, multiple linear regression; NIR, near infrared; PLSR, partial least square regression; PRESS, predicted sum of squares; RPD, residual prediction deviance; rrBLUP, ridge regression with best linear unbiased prediction; SEP, standard error of prediction; SVDPC, singular value decomposition and principal component algorithm..

genotypes (Ehdaie et al., 2006, 2008), the use of stem carbohydrates for selection is not widespread in the United States. This is because (i) more research is needed to understand the role of soluble stem carbohydrates in the drought response of US wheat germplasm, and (ii) the assay for soluble carbohydrates is time consuming and requires the use of a caustic chemicals (Parnell and White, 1983; Brink and Marten, 1986; Brown et al., 1987).

Near-infrared spectroscopy has been used to estimate organic compounds such as protein content, digestible fiber content, and starch in small grains, forages, and other agronomic crops. It requires little sample preparation and is relatively inexpensive (Delwiche, 2004; Roberts et al., 2004; Osborne, 2006), but NIR spectral data must be statistically calibrated to a set of laboratory assay results for a given trait. Two common modeling techniques used for NIR calibration are (i) best subset selection by multiple linear regression (MLR) (Westerhaus et al., 2004) and (ii) partial least squares regression (also called predictive latent space regression) (Wold et al., 2001).

In MLR, the number of spectral data points that can serve as predictors (p) is limited by the number of observations (n) due to a degrees of freedom deficit. Best subsets selection is a process for variable selection in MLR that optimizes a chosen statistical criterion (e.g., r^2, the predicted sum of squares [PRESS] statistic, or Bayes' information criterion). A drawback of MLR is its inability to obtain a robust estimate of an individual predictor's slope when there is high multicollinearity among the predictors, a common feature of NIR spectral data. As a result, a model from MLR cannot be easily extended to data independent of the training set.

Partial least squares regression (PLSR) can include all of the NIR spectral information by utilizing singular value decomposition of either the correlation matrix of the NIR spectral data or the covariance matrix between NIR and the soluble carbohydrate data to summarize the variation into latent variables. Partial least squares regression methods have been applied with success to many NIR calibrations of soluble carbohydrates in wheat and other grasses (Brink and Marten, 1986; Centner et al., 2000; Shetty et al., 2012), in which both high levels of accuracy (r^2 = 0.92–0.98, residual prediction deviance [RPD] = 3.3) (Jafari et al., 2003; Chen and Wang, 2004; Decruyenaere et al., 2012) and moderate levels of accuracy (r^2 = 0.71–0.78, RPD = 1.56–1.97) (Nie et al., 2009) have been reported. Partial least squares regression methods can summarize variation effectively, but they are computationally intensive.

Other statistical modeling approaches have been developed that handle both high multicollinearity and instances where $p >> n$. Shrinkage methods accomplish this by applying a biasing penalty that shrinks the magnitude of the regression coefficients. Ridge regression (and its mathematical equivalent best linear unbiased prediction, rrBLUP) applies a penalty that shrinks all the coefficients by a constant factor. Least absolute shrinkage and selection operator (lasso) selects variables and shrinks coefficients with a small absolute value more than larger coefficients. Elastic net is a compromise between the two using both the L1-norm (lasso) and L2-norm (ridge) for penalization of the regression coefficients (Tibshirani, 1996; Zou and Hastie, 2005). Bayesian lasso regression is a variation on traditional lasso where the variance of the predictors is not assumed constant and there is no variable selection of the predictors (Park and Casella, 2008).

Decision tree learning is another effective approach for generating robust predictions less sensitive to overfitting and for modeling nonlinear effects. Random forest is a type of nonparametric decision tree that uses bootstrapping for creating the training set and samples the predictor variables randomly at each node to determine the best split (Breiman, 2001). Although there are several published reports of these methods outperforming PLSR in chemometric applications, they have not been widely used for NIR calibration (Chen and Martin, 2009; Dyar et al., 2012; Lee et al., 2013).

To address these issues, we (i) evaluated different statistical approaches for predicting soluble stem carbohydrates using NIR spectroscopy and (ii) evaluated the ability of NIR-predicted stem carbohydrates to predict wheat yield rankings under drought. We conducted two field experiments to evaluate the performance of five spring wheat genotypes in multiple cropping systems and locations for a total of six environments over 2 yr in the semiarid Pacific Northwest of the United States. We used a diverse array of statistical models for the NIR calibration based on their ability to produce robust predictions when there is multicollinearity in the predictors and to use the entire data set for predictive models when $p >> n$. A reliable, fast, and simple method for assessing drought performance in early generation wheat breeding is essential for efficiently developing drought-tolerant wheat.

MATERIALS AND METHODS

Genotypes

Five cultivars were used in the study: Alpowa, Blanca Grande, Louise, Otis, and Walworth. Alpowa (PI 566596) was originally bred for the low-rainfall regions of the state of Washington (Li et al., 2011). Louise (PI 634865) was developed for the high-rainfall region of Washington (>400 mm annual rainfall) but has exhibited widespread adaptation (Kidwell et al., 2006b). Otis (PI 634866) was bred for the intermediate rainfall zones (<400 mm) but has also shown wide adaptation in the Pacific Northwest (Kidwell et al., 2006a). Blanca Grande (PVP 200200240) is a hard white wheat marketed for the western United States by Resource Seeds, Inc. Walworth (PI 630938) was bred in the Northern Great Plains, and its response to drought is unknown (Hall, 2003). Blanca Grande, Alpowa, and Louise have all been found to maintain similar yields across different rainfall regimes (Li et al., 2011).

Field Experiments

Field experiments were conducted in 2009 and 2010 at five locations in eastern Washington. These locations represent the rainfall gradient of Washington's dryland wheat production region (Table 1). The experiments were arranged in a randomized complete block design with four blocks and four replicates within each block. The experiments were conducted under diverse management systems. The conventionally managed locations were the Washington State University Dryland Experiment Station at Lind, WA, in 2009 and 2010 (47°0′3.96″ N, 118°33′33.48″ W). The certified organic management sites were G&L Farms near Benge, WA, in 2009 (46°52′9.84″ N, 118°1′0.84″ W) and the Washington State

Plantstress note: Precipitation amount decreases towards grain filling

University managed Boyd Farm (located near Pullman, WA, 46°45′0.72″ N, 117°4′54.48″ W) in 2010, which was in transition to certified organic production. No-till management sites were the Juris farm in Bickleton, WA (46°0′16.2″ N, 120°14′26.16″ W), in 2009 and the Aeschliman Farm at Colfax, WA, in 2010 (46°48′6.48″ N, 117°27′7.56″ W) (Table 1). The plots in Lind, 2009 and Benge, 2009 were planted in rows of five with 25-cm spacing using a Wintersteiger planter. The plots in Lind, 2010 and Pullman, 2010 were planted in rows of 10 with 18.5-cm spacing with a Fabro planter (Fabro Enterprises). The plots at the no-till sites of Bickleton and Colfax were planted in rows of seven with 16-cm spacing by a custom-built planter fitted with cross-slot openers (Baker No-Tillage Ltd.). All plots were 460 cm long and 135 cm wide, except at Pullman, where they were 365 cm long to adjust for the greater yield potential for that environment, and Lind and Pullman in 2010, where they were 210 cm wide due to the planter used (Fabro). Plot widths include spacing between rows of plots. Previous evidence indicates that including the extra space results in yield estimates that more accurately reflect large-scale yield trials (data not shown). Plots were managed to control weeds and diseases. The plots were weeded by hand throughout the season. In 2010, the plots at Colfax where sprayed with propiconazole (1-[[2-(2,4-dichlorophenyl)-4-propyl-1,3-dioxolan-2-yl]methyl]-1*H*-1,2,4-triazole) at labeled rates in response to stripe rust infection (*Puccinia striiformis* f. sp. *tritici*). Before planting, the upper 33 cm of soil was sampled and tested for N, P, K, and S. Additional samples for N analysis (NO_3 and NH_4) were collected from the 33- to 66- and 66- to 100-cm soil depths. Each site was fertilized at planting and planted at densities and fertilizer levels appropriate to the soil conditions and the expected grain yield for the rainfall zone (Koenig, 2005; Washington State University Extension, 2011).

Plant height at maturity was measured by averaging the heights of three plants within each plot that were representative of the plot. The plots were harvested at maturity with a Wintersteiger research plot combine. Whole-plot grain yield was measured for each plot and extrapolated to kilograms per hectare.

Table 1. Field sites used and precipitation at each site during the experiment.

Location	Year	Avg. annual precipitation†	Actual precipitation
		mm	
Lind	2009	247	212
	2010	247	293
Benge	2009	301	276
Bickleton	2009	338	272
Pullman	2010	538	474
Colfax	2010	500	340

† Average rainfall estimates provided by the Western Regional Climate Center using data gathered from regional weather stations for a minimum of 50 yr (www.wrcc.dri.edu, accessed August 2014).

Laboratory Analyses

The upper nodes of 5 to 10 stems were sampled from each plot 180 degree days after the onset of anthesis (approximately 10 calendar days), when the wheat was just beginning to undergo grain fill (Zadoks growth stage 72) (Zadoks et al., 1974). Immediately after sampling, the stem tissue was dried for 72 h at 60°C, ground to a fine powder (0.2 mm), and stored at room temperature.

Near-infrared spectral readings of ground wheat stems were taken in duplicate using an optical-grade petri dish illuminated and interrogated from below using an Analytical Spectral Devices (ASD Inc.) high-intensity contact probe foreoptic. The dish was rotated 90° between readings and 10 internal scans integrated per recorded spectra. The spectra were recorded with an ASD AgriSpec full-range spectrometer (350–2500 nm, 1-nm sampling, with 3-nm spectral resolution to 1000 nm and 10-nm spectral resolution at longer wavelengths) after a warm-up time of at least 2 h and calibration with a white and dark reference. The two scans per sample were averaged, and a smoothing spline—weighted inversely to wavelength-specific instrument noise—was fit to the spectra as described by Brown et al. (2006), with reflectance (R) values extracted at 2- and 10-nm intervals. Commonly used transformations were applied to the reflectance data to reduce multicollinearity, including first derivative, second derivative, absorbance [log(1/R)] and the Kubelka–Munk (KM) function, derived for ideal diffuse reflectance spectra. All preprocessing of data was performed manually in R versions 3.1.0 to 3.1.3 (R Development Core Team, 2011) following the procedures described by Brown et al. (2006).

Approximately 25% of the ground stem samples were randomly selected from each environment and assayed directly for soluble carbohydrates. Between 19 and 23 samples were chosen from each environment, for a total of 124 samples in the training population. The remaining 353 samples constituted the unknown population, totaling 477 samples (three experimental plots were lost due to planting error).

Soluble carbohydrates were extracted using a modification of the methods of Xue et al. (2008). Samples of 0.5 g each were shaken in a hot solution of 80% ethanol, centrifuged for 15 min at 10,000 rpm, and the supernatants saved. This procedure was repeated three times. The supernatants were combined for each sample and brought to 10 mL with 80% ethanol. Soluble carbohydrate concentration was determined in the combined supernatants using the anthrone method (Yemm and Willis, 1954), with glucose as a standard. The assay was independently performed twice on each extracted sample. For each assay, two aliquots were removed, their absorbances were measured at 620 nm, and the results were averaged together. A linear calibration curve of the glucose standard was applied to the absorbance of each assay ($r^2 = 0.98$), and the resultant concentrations were averaged across all technical replicates to give the final measured soluble stem carbohydrate concentration for each sample.

Near-Infrared Spectral Calibration and Model Evaluation

Seven modeling techniques to predict stem carbohydrates from NIR spectra were tested: MLR, PLSR with the SIMPLS algorithm (de Jong, 1993; Martens, 2001), PLSR with the singular value decomposition and principal component (SVDPC) algorithm (Martens, 2001), best linear unbiased predictor using the ridge regression kernel (rrBLUP) (Whittaker et al., 2000), lasso (Tibshirani, 1996), elastic net (Zou and Hastie, 2005), Bayesian lasso regression (BLR) (Park and Casella, 2008), and random forest (Breiman, 2001).

All model building, testing, and other analyses of the spectral data were performed in R 3.1.0 to 3.1.3. Spectral data points in 2- and 10-nm intervals for absorbance, first derivative and second derivative of reflectance, and the KM transformation were used in the calibration models. Previous studies reported using spectral data in 2-nm intervals (Ruuska et al., 2006). However, because the spectral resolution of most spectrophotometers in the NIR range is 10 nm, both wavelength intervals were tested. The spectral data points for each wavelength were standardized to a mean of zero and standard deviation of one across all samples before analysis. To assess the extent of multicollinearity in the data, principal component analysis was performed on the 10-nm spectral data set. The variance inflation factor could not be calculated for any of the data sets, regardless of the transformation, due to perfect collinearity between some of the predictor variables.

Model selection for MLR was performed using the Bioconductor package maSigPro (Conesa et al., 2006) and confirmed in Proc Reg of SAS Version 9.3 (SAS Institute). Partial least squares regression was performed using the PLSR package (Mevik and Wehrens, 2007). The number of latent variables as predictors in the model was chosen by minimizing the PRESS statistic, calculated as

$$\sum_{i=1}^{n}\left(y_i-\hat{y}_{i,-i}\right)^2$$

Ridge regression in the mixed model with BLUP was performed in the rrBLUP package (Endelman, 2011). The Glmnet package was used for finding lasso and elastic net estimates (Friedman et al., 2010). The optimal lamba values for lasso and elastic net were chosen by minimizing the average mean square error in leave-one-out cross-validation. The tuning parameter α was set at zero for lasso and 0.5 for elastic net. Bayesian lasso regression was performed with the BLR package (de los Campos et al., 2009) using the lambda from the model fitting as a starting point for the Gibbs sampler. The average of the estimates after 20,000 iterations was used, discarding the first 10,000 as burn-in. The R package *randomForest* (Liaw and Weiner, 2002) was used to predict the data. The default settings were used (maximum number of trees = 500, sample with replacement = true, maximum number of terminal nodes = n) with the exception that when the spectral variables were sampled and tested for their efficacy in predicting y, the maximum number of predictors was attempted.

All calibration models were evaluated for changes in their predictive accuracy due to the effects of (i) wavelength interval, (ii) data transformation, and (iii) the modeling technique using n-fold cross-validation (equivalent to leave-one-out). In addition, the effects of modeling techniques using only the 10-nm first derivative transformed data were tested using k-fold cross-validation of 10-fold, fourfold and twofold and by removing the data for a single experiment factor: environment, year, or cultivar. To calculate these, the data for a specific year, cultivar, or environment were omitted during the model building process, and then those data were predicted from the resultant calibration model. This process was repeated until all data points were predicted. The evaluation criteria for all cross-validation procedures were standard error of prediction (SEP), the coefficients of determination for n-fold cross-validation ($CV_{-i}\,r^2$) and k-fold cross-validation ($CV_{-k}\,r^2$), RPD (the ratio between the standard deviation of the measured values and the standard error of prediction), and bias. All statistical terminology and equations are summarized in Supplementary Table S1.

Statistical Analysis of the Field Trials

To evaluate the effects of wheat cultivar and environment on the variation for soluble stem carbohydrates, the predicted soluble carbohydrate response data were subject to an ANOVA in the context of the multi-environmental trial described above. The data were analyzed both within and across sites in a generalized randomized complete block design using Proc Mixed in SAS, where the block was treated as a random effect and all other effects were treated as fixed effects:

$$Y_{hij}=\mu+\alpha_i+\epsilon_{i(h)}+\beta_j+\left(\alpha\beta\right)_{ij}+\varepsilon_{j(hi)}$$

(h = 1,2,3,4 blocks, i = 1,2...,6 sites, j = 1,2,...,5 cultivars)

where Y_{ij} is the response variable, subject to the main effects of environment, α_i, and cultivar, β_j, and their interaction, $(\alpha\beta)_{ij}$. There were four blocks nested within each environment, with effects estimated as $\epsilon_{i(h)}$, and four replicates within each block, with error estimated as $\varepsilon_{j(hi)}$, for a total of 80 plots per site and 480 plots across the entire experiment. The relative portion of variation explained by the model terms were estimated by taking the ratio between the marginal sums of squares in Type III sums of squares analysis and the sums of squares total for each dependent variable. Supplementary Fig. S1 provides a diagram of the experimental workflow.

RESULTS

Soluble Carbohydrate Assay

The results from the anthrone assay showed that the soluble carbohydrate data were normally distributed, with a mean of 283.4 mg glucose g^{-1} dried stem tissue and a standard deviation of 75.87. Measurement error due to the extraction procedure and anthrone assay were quantitatively similar, with standard errors of 9.32 and 9.48, respectively. Two outlying observations were removed from the experiment due to having a Studentized residual with an absolute value >3. Both samples had an inadequate amount of ground stem material for NIR analysis, resulting in visually deviant NIR spectra.

Near-Infrared Spectral Transformation and Resolution

Principal component analysis revealed substantial multicollinearity among the spectral data points, particularly for the untransformed reflectance and the KM-transformed data. Principal component analysis of the first derivative of the spectral data indicated that the first three components summarized just under 60% of the variation across the entire data set. In contrast, 98% of the variation in the untransformed reflectance and the KM-transformed data could be summarized in the first three principal components. The second derivative of the spectral data was the least correlated across the x variables.

The influence of the data transformations on model performance including the untransformed reflectance were

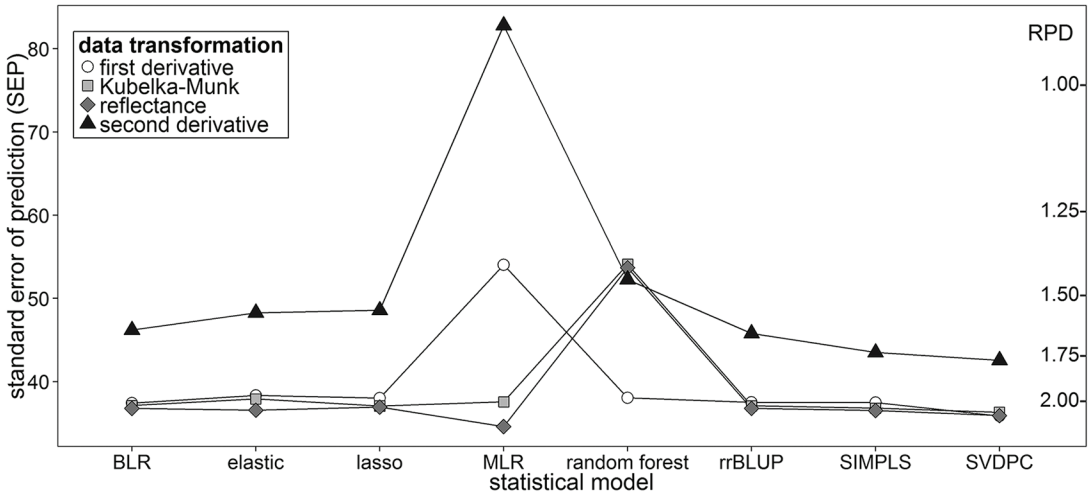

Fig. 1. Comparison of prediction accuracy due to modeling technique using n-fold cross-validation. On the left axis is the standard error of prediction (SEP), and the right axis contains the residual prediction deviation (RPD), the ratio between the standard deviation and SEP. The modeling techniques included Bayesian lasso regression (BLR), elastic net, least absolute shrinkage and selection operator (lasso), multiple linear regression (MLR), random forest, ridge regression with best linear unbiased prediction (rrBLUP), the SIMPLS algorithm in partial least squares regression, and the singular value decomposition with principal components (SVDPC) algorithm in partial least squares regression.

evaluated by the SEP, CV_{-i} r^2, and the RPD values from *n*-fold cross-validation. Both the first-derivative-transformed and the untransformed reflectance data had the highest CV_{-i} r^2 and RPD across the different modeling approaches (Supplementary Table S2; Fig. 1). Most of the data transformations resulted in similar CV_{-i} r^2 values, ranging from 0.44 to 0.71. However, the second-derivative transformation resulted in a lower CV_{-i} r^2 and a higher SEP. The first-derivative-transformed data were used for further analysis because the predictions from those data had the highest estimated accuracy in cross-validation.

Wavelength interval (2 or 10 nm) had very little impact on model accuracy. The results of the SEP statistic, CV_{-i} r^2, and RPD all indicate that there was little gain in predictive accuracy from sampling spectra more intensively (2 nm). When the second-derivative-transformed data were used in 2-nm intervals for calibration, the predictive accuracy was lower (data not shown). Because there was no improvement in prediction from using the 2-nm spectra, the smaller and less computationally demanding 10-nm spectra were used.

Calibration Model

We found that the PLSR methods were the most effective modeling techniques for predicting the known data with high precision (r^2 = 0.95, CV_{-i} r^2 = 0.85) and capturing the phenotypic variation present in the training population (Fig. 2; Supplementary Table S2). The singular value decomposition with principal components (SVDPC) kernel, a spectral decomposition method that uses the correlation matrix between only the *x* variables, had a higher prediction accuracy than the SIMPLS algorithm, which is based on the cross-product matrix between *x* and *y*. The remaining statistical modeling techniques had very similar predictive accuracies. The shrinkage-based models (rrBLUP, lasso, elastic net, and BLR) and random forest models showed similar performance in terms of predicting the data accurately and minimizing SEP cross-validation (Table 2; Supplementary Table S3). Multiple linear regression and random forest had low predictive accuracy in *n*-fold cross-validation, probably due to overfitting the small data set. The relative ranking of the models for predictive accuracy considering four statistics, SEP, CV_{-i} r^2, CV_{-k} r^2, and RPD, was PLSR-SVDPC > PLSR-SIMPLS > BLR > rrBLUP > lasso > elastic net > random forest > MLR.

Despite different approaches to capturing the variation present in NIR spectra, the predicted values for the first-derivative-transformed data were quite similar for all models except MLR (Supplementary Fig. S2). The PLSR-SVDPC approach had the highest predictive accuracy across all cross-validation approaches. All model techniques produced stable models without a substantive drop in the coefficient of determination under different cross-validation approaches. The largest decrease in CV_{-k} r^2 was observed when a year, site, or combination of the two was omitted during model building.

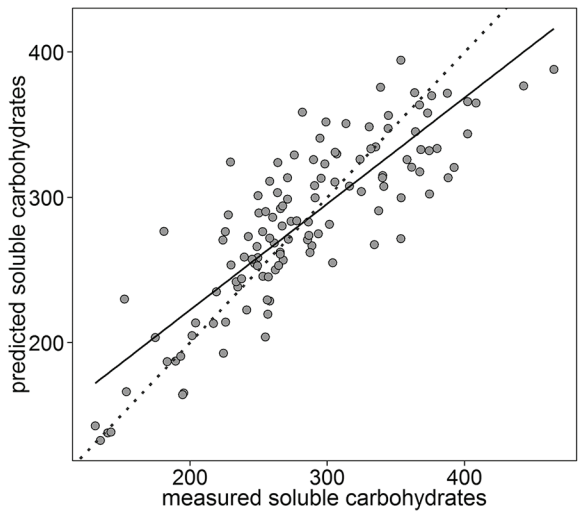

Fig. 2. Predicted stem carbohydrates vs. their laboratory-measured values using the singular value decomposition and principal component (SVDPC) calibration model. Both axes are in milligrams of carbohydrate per gram of stem dry weight. The dotted line is the location of a perfect fit.

Table 2. Standard error of prediction (SEP) of near-infrared-predicted soluble stem carbohydrates using Bayesian lasso regression (BLR), elastic net, least absolute shrinkage and selection operator (lasso), multiple linear regression (MLR), random forest, ridge regression with best linear unbiased prediction (rrBLUP), the SIMPLS algorithm, and the singular value decomposition with principal components (SVDPC) algorithm and different cross-validation approaches. The first derivative-transformed data in 10-nm intervals were used.

Validation approach	BLR	elastic	lasso	MLR	Random forest	rrBLUP	SIMPLS	SVDPC
n-fold	37.44	38.35	38.03	54.03	38.05	37.52	37.47	35.87
k-fold								
10-fold (repeated 50 times)	37.59	37.82	38.12	50.28	38.16	37.62	37.77	36.43
Fourfold (repeated 50 times)	37.67	38.31	38.52	49.59	38.80	37.84	38.14	37.15
Twofold (repeated 50 times)	38.92	39.58	39.92	46.17	41.10	39.60	40.45	39.79
Leave out a cultivar	37.88	38.46	38.57	48.78	40.64	39.37	38.14	37.21
Leave out a field site	43.36	44.01	43.68	60.83	43.33	43.58	45.00	43.21
Leave out an environment (year–site combination)	43.43	44.13	43.46	59.23	44.40	43.73	44.04	44.15
Leave out a year	45.93	46.45	46.91	62.36	47.99	46.48	48.56	48.39

Field Trial

Both 2009 and 2010 were dry years, with lower than average annual precipitation for most of the field sites ranging from 8.3% below normal (Benge) to 32.0% below normal (Pullman) (Table 1). A notable exception was Lind in 2010, which received 18.6% more total precipitation than usual. The experimental wheat plots experienced high emergence (90–100%) and normal growth and development from the vegetative phase to maturity. All sites were subject to a late-season drought during which very little or no precipitation was recorded between anthesis and harvest. The predicted stem carbohydrates from the PLSR-SVDPC model had a mean of 278.3 and standard deviation of 57.54. Their distributions were similar across the different environments. As a precautionary measure, we tested and found that plant height had no influence on the concentration of soluble stem carbohydrates and did not need to be included as a covariate for an analysis of covariance.

The effects of environment and cultivar on soluble stem carbohydrates and grain yield were statistically significant across all sites, and cultivar had a statistically significant effect on those response variables within all environments (Table 3). The effect of environment was responsible for 71.7% of the total variation in yield and 5.93% of the variation in stem carbohydrates (Table 3). The values for stem carbohydrates ranged between 150 and 375 mg g^{-1} dried stem tissue across environments. The effect of cultivar on stem carbohydrates and yield varied within each environment and for the different dependent variables (Table 3). Occasionally, the Pearson correlation coefficient, ρ, was higher than the Spearman correlation coefficient, *r*. These patterns were inconsistent and were not correlated with year, management, or annual precipitation.

Soluble carbohydrate concentration was effective in predicting yield rankings across sites and within sites when using least square means (ρ = 0.6–1, *r* = 0.683–0.922) (Table 4; Supplementary Table S3). The highest correlations were in 2009 at Benge and Bickleton, which were both water stressed, and at Colfax in 2010, which was the most water sufficient of all the environments (Table 1). Regression analysis between the least square means for soluble stem carbohydrates and yield showed that the average yield increased 1 kg ha^{-1} for every increase of 1.72 to 7.88 mg soluble carbohydrates g^{-1} plant tissue, depending on the environment (Fig. 3). The two highest yielding cultivars in this study, Alpowa and Louise, had the highest mean levels of soluble stem carbohydrates across environments (yield of 1718.62 and 1837.65 kg ha^{-1}, soluble carbohydrates of 328.18 and 321.92 mg g^{-1}, respectively). Blanca Grande, the lowest yielding

Table 3. Results of ANOVA for soluble stem carbohydrates and yield, and the relative amount of variation of the variables explained in the linear model using Type III sums of squares.

Effect		Stem carbohydrates		Yield	
		F-test	Variation	*F*-test	Variation
			%		%
Treatments across all environments					
Cultivar (Cv)		84.90***	33.05	37.50***	8.39
Environment (E)		8.14**	5.93	85.91***	71.71
Cv × E		6.84***	13.86	2.55**	2.58
Cultivar within each environment					
2009	Lind	76.39***	79.77	26.95***	58.29
	Benge	3.22*	15.72	23.17***	48.76
	Bickleton	16.35***	47.19	31.18***	59.79
2010	Lind	11.90***	38.98	15.33***	36.99
	Pullman	25.30***	54.83	25.06***	49.64
	Colfax	13.87***	10.38	3.45*	15.06

* $P \leq 0.05$.
** $P \leq 0.01$.
*** $P \leq 0.001$.

Table 4. Correlation coefficients for Pearson and Spearman correlations between soluble stem carbohydrates and grain yield using the raw data and least square means for each cultivar.

Source of variation		Raw data		Least square means	
		Pearson	Spearman	Pearson	Spearman
Within each environment					
2009	Lind	0.466***	0.469***	0.683	0.7
	Benge	0.236*	0.233*	0.786	0.9*
	Bickleton	0.555***	0.589***	0.889*	1***
2010	Lind	0.189†	0.186	0.922*	0.7
	Pullman	0.538***	0.504***	0.893	0.7
	Colfax	0.408**	0.361*	0.873†	0.9*
Across all environments		0.292***	0.298***	0.904*	0.8

† $P \leq 0.10$.
* $P \leq 0.05$.
** $P \leq 0.01$.
*** $P \leq 0.001$.

cultivar (1347.72 kg ha^{-1}), had the lowest mean soluble stem carbohydrates (232.32 mg g^{-1}) (Fig. 4).

DISCUSSION

Soluble stem carbohydrates are advantageous for crop improvement because they can be assessed early in the breeding cycle on a small number of plants, when accurate yield estimates from large plots are not feasible. The success of using stem carbohydrates in spring wheat as an early-generation yield indicator relies on (i) obtaining a reliable NIR calibration for soluble carbohydrates, and (ii) a strong correlation between yield and soluble stem carbohydrates. These two requirements were met, confirming the importance of soluble storage carbohydrates for drought response in the spring wheat germplasm and environmental conditions of this study.

The importance of data quality rather than the statistical approach in determining model predictive accuracy was corroborated in this study. The relationship between NIR spectral data and soluble carbohydrates was linear and could be described efficiently using latent variables or a small number of variables with shrinkage approaches. Previous research with several of these regression and classification tools (BLR, lasso, elastic net, and random forest) indicated a tendency of these modeling techniques to capture the same variation present despite using different approaches for summarizing the variation (Heslot et al., 2012). A broad set of drought-prone environments was used in this study that reflect the target conditions of dryland wheat production. The wheat cultivars used are important to the regional wheat industry, but this panel could be too small to span the full extent of variation of soluble stem carbohydrates. More observations across a broader panel of wheat genotypes are needed to build a robust NIR calibration model for soluble stem carbohydrates and to clarify the relationship between soluble stem carbohydrates and grain yield.

Direct selection for yield under drought is challenging because it is a complex quantitative trait governed by innumerable genetic and non-genetic factors. Selecting indirectly for a correlated quantitative trait is more effective than direct selection when the genetic heritability is greater in the correlated trait than the target trait and the genetic correlation between the two traits is high (Falconer and Mackay, 1996). While grain yield and soluble stem carbohydrate concentration were highly correlated, the amount of phenotypic variation explained by cultivar for both traits varied across these experimental conditions. The amount of variation attributed to genotype in soluble stem carbohydrate concentration was fourfold higher than in yield (33.05 and 8.39%, respectively) when this analysis was performed across environments. However, when the statistical analysis was stratified by environment, the amount of variation explained by cultivar was inconsistent for both yield and soluble stem carbohydrates. The Spearman rank correlations between the traits were moderate (0.4–0.6) to high (0.7–1) across and within environments, and the rankings were largely consistent across the different environments. Together, our results suggest that indirect selection for stem carbohydrate rankings can be more effective than direct selection for yield if the selection is for environments similar to those evaluated in the training set. Given that this is a small data set, these results should be viewed as proof of concept that needs further research under US Pacific Northwest conditions before widespread usage is justified.

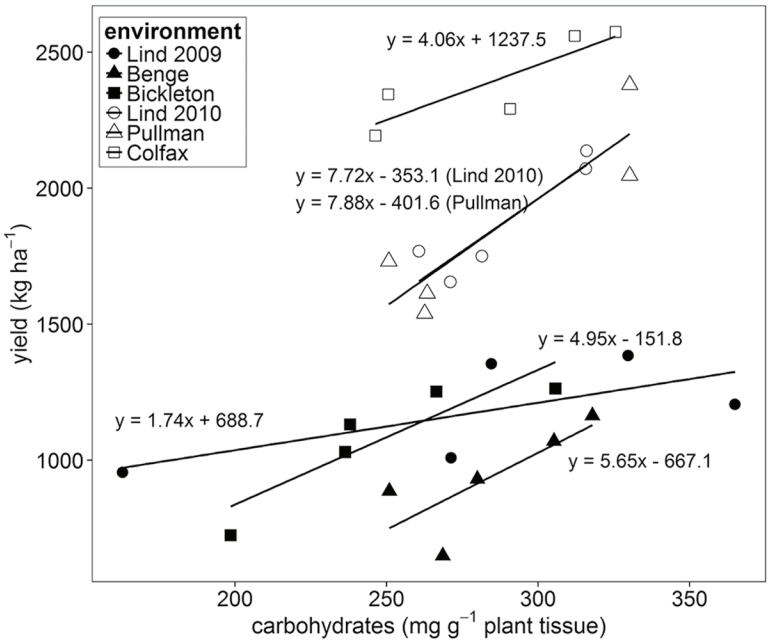

Fig. 3. Grain yield vs. soluble stem carbohydrates for each environment and their regression equations using least square means.

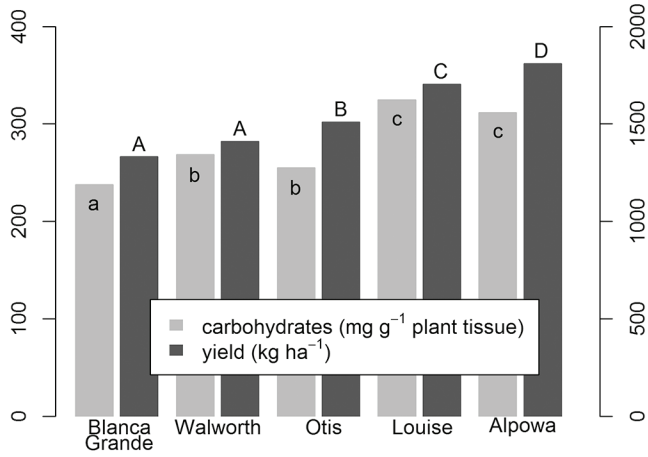

Fig. 4. Least square means for yield and stem carbohydrates by cultivar across all environments. The left and right axes indicate grain yield (kg ha^{-1}) and stem carbohydrate concentration (mg g^{-1} dry weight), respectively. Letters indicate pairwise differences (α = 0.05, Tukey's adjusted) within each response variable.

SUPPLEMENTARY MATERIALS

A supplementary file containing the following figures and tables is included: Fig. S1, a workflow diagram of the experiment; Table S1, abbreviations and statistics used; Table S3, an extensive set of results from the calibration tests; Table S4, a table of Spearman correlations between soluble stem carbohydrates and yield among the different environments; and Fig. S2, a panel of scatterplots for the measured vs. predicted stem carbohydrate concentration for each statistical modeling technique.

ACKNOWLEDGMENTS

We gratefully acknowledge the help of Erin Getzin in performing the laboratory measurements, Lesley Murphy for her work refining the anthrone assay protocol, and Ross Bricklemeyer for equipment and software training. This work was conducted with support from the following: USDA–ARS in-house Project 5348-21220-003-00D; Washington Grains Commission Project 6451; and WSU Center for Sustaining Agriculture and Natural Resources BioAg Project 110817-001.

REFERENCES

Aggarwal, P., and S.K. Sinha. 1984. Effect of water stress on grain growth and assimilate partitioning in two cultivars of wheat contrasting in their yield stability in a drought-environment. Ann. Bot. 53:329–340.

Batten, G.D., A.B. Blakeney, V.B. McGrath, and S. Ciavarella. 1993. Non-structural carbohydrate: Analysis by near infrared reflectance and its importance as an indicator of plant growth. Plant Soil 155–156:243–246. doi:10.1007/BF00025029

Blum, A. 2011. Plant breeding for water-limited environments. Springer, New York.

Blum, A., L. Shpiler, G. Golan, J. Mayer, and B. Sinmena. 1991. Mass selection of wheat for grain filling without transient photosynthesis. Euphytica 54:111–116. doi:10.1007/BF00145637

Breiman, L. 2001. Random forests. Mach. Learn. 45:5–32. doi:10.1023/A:1010933404324

Brink, G.E., and G.C. Marten. 1986. Analysis of alfalfa root carbohydrate concentration by near infrared spectroscopy. Crop Sci. 26:159–161. doi:10.2135/cropsci1986.0011183X002600010037x

Brooks, A., C.F. Jenner, and D. Aspinall. 1982. Effect of water deficit on endosperm starch granules and grain physiology of wheat and barley. Aust. J. Plant Physiol. 9:423–436. doi:10.1071/PP9820423

Brown, D., K. Shepherd, M. Walsh, M. Mays, and T. Reinsch. 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132:273–290. doi:10.1016/j.geoderma.2005.04.025

Brown, W.F., J.E. Moore, and P. Mislevy. 1987. Near infrared reflectance analysis of total nonstructural carbohydrates concentration in tropical grasses. Crop Sci. 27:786–788. doi:10.2135/cropsci1987.0011183X002700040036x

Centner, V., J. Verdú-Andrés, B. Walczak, D. Jouan-Rimbaud, F. Despagne, L. Pasti, et al. 2000. Comparison of multivariate calibration techniques applied to experimental NIR data sets. Appl. Spectrosc. 54:608–623. doi:10.1366/0003702001949816

Chen, C., and S. Wang. 2004. Nutritional component determination in Napier grass using near-infrared spectroscopy analysis. J. Taiwan Livest. Res. 37:195–204.

Chen, T., and E. Martin. 2009. Bayesian linear regression and variable selection for spectroscopic calibration. Anal. Chim. Acta 631:13–21. doi:10.1016/j.aca.2008.10.014

Conesa, A., M.J. Nueda, A.A. Ferrer, and T. Talon. 2006. maSigPro: A method to identify significant differential expression profiles in time-course microarray experiments. Bioinformatics 22:1096–1102. doi:10.1093/bioinformatics/btl056

Decruyenaere, V., C. Clément, R. Agneessens, C. Losseau, and D. Stilmant. 2012. Development of near-infrared spectroscopy calibrations to quantify starch and soluble sugar content in the roots of *Rumex obtusifolius*. Weed Res. 52:1–5. doi:10.1111/j.1365-3180.2011.00895.x

de Jong, S. 1993. SIMPLS: An alternative approach to partial least squares regression. Chemom. Intell. Lab. Syst. 18:251–263. doi:10.1016/0169-7439(93)85002-X

de los Campos, G., H. Naya, D. Gianola, J. Crossa, A. Legarra, E. Manfredi, et al. 2009. Predicting quantitative traits with regression models for dense molecular markers and pedigree. Genetics 182:375–385. doi:10.1534/genetics.109.101501

Delwiche, S.R. 2004. Analysis of small grain crops. In: C.A. Roberts et al., editors, Near-infrared spectroscopy in agriculture. Agron. Monogr. 44. ASA, CSSA, and SSSA, Madison, WI. p. 269–320. doi:10.2134/agronmonogr44.c11

Dyar, M.D., M.L. Carmosino, E.A. Breves, M.V. Ozanne, S.M. Clegg, and R.C. Wiens. 2012. Comparison of partial least squares and lasso regression techniques as applied to laser-induced breakdown spectroscopy of geological samples. Spectrochim. Acta B 70:51–67.

Ehdaie, B., G.A. Alloush, M.A. Madore, and J.G. Waines. 2006. Genotypic variation for stem reserves and mobilization in wheat: II. Postanthesis changes in internode water-soluble carbohydrates. Crop Sci. 46:2093–2103. doi:10.2135/cropsci2006.01.0013

Ehdaie, B., G.A. Alloush, and J.G. Waines. 2008. Genotypic variation in linear rate of grain growth and contribution of stem reserves to grain yield in wheat. Field Crops Res. 106:34–43. doi:10.1016/j.fcr.2007.10.012

Endelman, J.B. 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome 4:250–255. doi:10.3835/plantgenome2011.08.0024

Falconer, D., and T. Mackay. 1996. Quantitative genetics. 4th ed. Prentice Hall, New York.

Foulkes, M.J., R.K. Scott, and R. Sylvester-Bradley. 2002. The ability of wheat cultivars to withstand drought in UK conditions: Formation of grain yield. J. Agric. Sci. 138:153–169. doi:10.1017/S0021859601001836

Friedman, J., T. Hastie, and R. Tibshirani. 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33:1–22.

Gallagher, J.N., P.V. Biscoe, and B. Hunter. 1976. Effects of drought on grain growth. Nature 264:541–542. doi:10.1038/264541a0

Hall, R. 2003. Walworth hard red spring wheat. South Dakota State Univ., Brookings.

Heslot, N., H.-P. Yang, M.E. Sorrells, and J.-L. Jannink. 2012. Genomic selection in plant breeding: A comparison of models. Crop Sci. 52:146–160. doi:10.2135/cropsci2011.06.0297

Jafari, A., V. Connolly, A. Frolich, and E.J. Walsh. 2003. A note on estimation of quality parameters in perennial ryegrass by near infrared reflectance spectroscopy. Ir. J. Agric. Food Res. 42:293–299.

Kidwell, K.K., V.L. DeMacon, G.S. Shelton, J.W. Burns, B.P. Carter, X.M. Chen, et al. 2006a. Registration of 'Otis' wheat. Crop Sci. 46:1386–1387. doi:10.2135/cropsci2005.06-0177

Kidwell, K.K., G.S. Shelton, V.L. Demacon, J.W. Burns, B.P. Carter, X.M. Chen, et al. 2006b. Registration of 'Louise' wheat. Crop Sci. 46:1384–1386. doi:10.2135/cropsci2005.06-0176

Koenig, R.T. 2005. Eastern Washington nutrient management guide: Dryland winter wheat. Washington State Univ. Ext., Pullman.

Lee, S., H. Choi, K. Cha, and H. Chung. 2013. Random forest as a potential multivariate method for near-infrared (NIR) spectroscopic analysis of complex mixture samples: Gasoline and naphtha. Microchem. J. 110:739–748. doi:10.1016/j.microc.2013.08.007

Li, P., J. Chen, and P. Wu. 2011. Agronomic characteristics and grain yield of 30 spring wheat genotypes under drought stress and nonstress conditions. Agron. J. 103:1619–1628. doi:10.2134/agronj2011.0013

Liaw, A., and M. Weiner. 2002. Classification and regression by randomForest. R News 2(3):18–22.

Lopez, G.C., G.M. Banowetz, C.J. Peterson, and W.E. Kronstad. 2003. Dehydrin expression and drought tolerance in seven wheat cultivars. Crop Sci. 43:577–582. doi:10.2135/cropsci2003.0577

Martens, H. 2001. Reliable and relevant modelling of real world data: A personal account of the development of PLS regression. Chemom. Intell. Lab. Syst. 58:85–95. doi:10.1016/S0169-7439(01)00153-8

Mevik, B.-H., and R. Wehrens. 2007. The pls package: Principal component and partial least squares regression in R. J. Stat. Softw. 18(2). doi:10.18637/jss.v018.i02

Nie, Z., G.F. Tremblay, G. Bélanger, R. Berthiaume, Y. Castonguay, A. Bertrand, et al. 2009. Carbohydrates in alfalfa–timothy mixtures predicted with near infrared reflectance spectroscopy equations developed for single species. Can. J. Anim. Sci. 89:279–283. doi:10.4141/CJAS08128

Olien, C.R., and J.L. Clark. 1993. Changes in soluble carbohydrate composition of barley, wheat, and rye during winter. Agron. J. 85:21–29. doi:10.2134/agronj1993.00021962008500010005x

Osborne, B.G. 2006. Application of near infrared spectroscopy in quality screening of early-generation material in cereal breeding programmes. J. Near Infrared Spectrosc. 14:93–101. doi:10.1255/jnirs.595

Park, T., and G. Casella. 2008. The Bayesian lasso. J. Am. Stat. Assoc. 103:681–686. doi:10.1198/016214508000000337

Parnell, R., and J. White. 1983. The use of near infrared reflectance analysis in predicting the digestibility and the protein and water soluble carbohydrate contents of perennial ryegrass. J. Natl. Inst. Agric. Bot. 16:221–227.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Found. Statist. Comput., Vienna. http://www.R-project.org/

Rebetzke, G.J., A.F. van Herwaarden, C. Jenkins, M. Weiss, D. Lewis, S. Ruuska, et al. 2008. Quantitative trait loci for water-soluble carbohydrates and associations with agronomic traits in wheat. Aust. J. Agric. Res. 59:891–905. doi:10.1071/AR08067

Roberts, C.A., J. Stuth, and P. Flinn. 2004. Analysis of forages and feedstuffs. In: C.A. Roberts et al., editors, Near-infrared spectroscopy in agriculture. Agron. Monogr. 44. ASA, CSSA, and SSSA, Madison, WI. p. 231–267. doi:10.2134/agronmonogr44.c10

Ruuska, S.A., G.J. Rebetzke, A.F. Herwaarden, R.A. van. Richards, N.A. Fettell, L. Tabe, and C.L.D. Jenkins. 2006. Genotypic variation in water-soluble carbohydrate accumulation in wheat. Funct. Plant Biol. 33:799–809. doi:10.1071/FP06062

Schnyder, H. 1993. The role of carbohydrate storage and redistribution in the source–sink relations of wheat and barley during grain filling. New Phytol. 123:233–245. doi:10.1111/j.1469-8137.1993.tb03731.x

Schramm, E.C., J.C. Abellera, L.C. Strader, K.G. Campbell, and C.M. Steber. 2010. Isolation of ABA-responsive mutants in allohexaploid bread wheat (*Triticum aestivum* L.): Drawing connections to grain dormancy, preharvest sprouting, and drought tolerance. Plant Sci. 179:620–629. doi:10.1016/j.plantsci.2010.06.004

Shetty, N., R. Gislum, A.M.D. Jensen, and B. Boelt. 2012. Development of NIR calibration models to assess year-to-year variation in total non-structural carbohydrates in grasses using PLSR. Chemom. Intell. Lab. Syst. 111:34–38. doi:10.1016/j.chemolab.2011.11.004

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. J.R. Stat. Soc. 58:267–288.

van Herwaarden, A.F., and R.A. Richards. 2002. Water soluble carbohydrate accumulation in stems is related to breeding progress in Australia wheats. In: J.A. McComb, editor, Plant breeding for the 11th millennium: Proceedings of the 12th Australasian Plant Breeding Conference, Perth, WA, Australia. 15–20 Sept. 2002. WA Dep. of Agriculture, Bentley, WA. p. 878–882.

Wardlaw, I.F., and J. Willenbrink. 2000. Mobilization of fructan reserves and changes in enzyme activities in wheat stems correlate with water stress during kernel filling. New Phytol. 148:413–422. doi:10.1046/j.1469-8137.2000.00777.x

Washington State University Extension. 2011. Cereal and grain legume variety testing. Washington State Univ. Ext., Pullman. http://variety.wsu.edu/

Westerhaus, M., J. Workman, Jr., J.B. Reeves II, and H. Mark. 2004. Quantitative analysis. In: C.A. Roberts et al., editors, Near-infrared spectroscopy in agriculture. Agron. Monogr. 44. ASA, CSSA, and SSSA, Madison, WI. p. 133–174. doi:10.2134/agronmonogr44.c7

Whittaker, J.C., R. Thompson, and M.C. Denham. 2000. Marker-assisted selection using ridge regression. Genet. Res. 75:249–252. doi:10.1017/S0016672399004462

Wold, S., M. Sjostrom, and L. Eriksson. 2001. PLS-regression: A basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58:109–130. doi:10.1016/S0169-7439(01)00155-1

Xue, G.-P., C.L. McIntyre, C.L.D. Jenkins, D. Glassop, A.F. van Herwaarden, and R. Shorter. 2008. Molecular dissection of variation in carbohydrate metabolism related to water-soluble carbohydrate accumulation in stems of wheat. Plant Physiol. 146:441–454. doi:10.1104/pp.107.113076

Yemm, E.W., and A.J. Willis. 1954. The estimation of carbohydrates in plant extracts by anthrone. Biochem. J. 57:508–514. doi:10.1042/bj0570508

Zadoks, J.C., T. Chang, and C.F. Konzak. 1974. A decimal code for the growth stages of cereals. Weed Res. 14:415–421. doi:10.1111/j.1365-3180.1974.tb01084.x

Zou, H., and T. Hastie. 2005. Regression shrinkage and selection via the elastic net, with applications to microarrays. J.R. Stat. Soc. 67:301–320. doi:10.1111/j.1467-9868.2005.00503.x