Genomic Selection for Processing and End-Use Quality Traits in the CIMMYT Spring Bread Wheat Breeding Program

Published June 15, 2016

ORIGINAL RESEARCH

Genomic Selection for Processing and End-Use Quality Traits in the CIMMYT Spring Bread Wheat Breeding Program

Sarah D. Battenfield, Carlos Guzmán, R. Chris Gaynor, Ravi P. Singh, Roberto J. Peña, Susanne Dreisigacker, Allan K. Fritz, and Jesse A. Poland*

Abstract

Wheat (*Triticum aestivum* L.) cultivars must possess suitable end-use quality for release and consumer acceptability. However, breeding for quality traits is often considered a secondary target relative to yield largely because of amount of seed needed and expense. Without testing and selection, many undesirable materials are advanced, expending additional resources. Here, we develop and validate whole-genome prediction models for end-use quality phenotypes in the CIMMYT bread wheat breeding program. Model accuracy was tested using forward prediction on breeding lines (n = 5520) tested in unbalanced yield trials from 2009 to 2015 at Ciudad Obregon, Sonora, Mexico. Quality parameters included test weight, 1000-kernel weight, hardness, grain and flour protein, flour yield, sodium dodecyl sulfate sedimentation, Mixograph and Alveograph performance, and loaf volume. In general, prediction accuracy substantially increased over time as more data was available to train the model. Reflecting practical implementation of genomic selection (GS) in the breeding program, forward prediction accuracies (r) for quality parameters were assessed in 2015 and ranged from 0.32 (grain hardness) to 0.62 (mixing time). Increased selection intensity was possible with GS since more entries can be genotyped than phenotyped and expected genetic gain was 1.4 to 2.7 times higher across all traits than phenotypic selection. Given the limitations in measuring many lines for quality, we conclude that GS is a powerful tool to facilitate early generation selection for end-use quality in wheat, leaving larger populations for selection on yield during advanced testing and leading to better gain for both quality and yield in bread wheat breeding programs.

Core Ideas

- Genomic selection applied for wheat quality in CIMMYT spring bread wheat breeding program.
- All wheat quality traits predicted and validated using forward genomic selection.
- Dough and loaf traits have moderately high predictive ability in CIMMYT breeding program.
- Genomic selection genetic gain 1.4 to 2.7 times higher than phenotypic selection.

THE HUMAN POPULATION is growing exponentially, with current projections predicting a population of >9 billion by 2050 (Gerland et al., 2014). An intersection of improved agronomic practices and improved crop varieties will be imperative to ensure food security in the coming decades. While overall production must increase, there

S.D. Battenfield and J.A. Poland, Wheat Genetics Resource Center, Dep. of Plant Pathology, Kansas State Univ., 4024 Throckmorton Plant Science Center, Manhattan, KS, 66506; C. Guzmán, R.P. Singh, R.J. Peña, and S. Dreisigacker, Global Wheat Program, International Maize and Wheat Improvement Center, Mexico, D.F., Mexico; R.C. Gaynor, The Roslin Institute and Royal (Dick) School of Veterinary Studies, Univ. of Edinburgh, Easter Bush, Midlothian, UK; J.A. Poland and A.K. Fritz, Dep. of Agronomy, Kansas State Univ., 2004 Throckmorton Plant Science Center, Manhattan, KS, 66506. Received 15 Jan. 2016. Accepted 25 Apr. 2016. *Corresponding author (jpoland@ksu.edu).

Abbreviations: AACC, American Association of Cereal Chemists; ALVPL, Alveograph ratio of height (P) to length (L) of the curve (tenacity/extensibility) (P/L); ALVW, dough strength, Alveograph work value under the curve ; FLRPRO, flour protein; FLRSDS, flour sodium dodecyl sulfate sedimentation; GAUSS, Gaussian kernel; GRNHRD, grain hardness; GRNPRO, grain protein; GS, genomic selection; LOFVOL, bread loaf volume; MIXTIM, Mixograph mix time; MP, Mixograph percentage torque; NIRS, near-infrared spectroscopy; PLSR, partial least squares regression; RRBLUP, ridge regression best linear unbiased predictor; SDS, sodium dodecyl sulfate; SNP, single-nucleotide polymorphism; TESTWT, test weight; TKW, 1000-kernel weight.

Published in Plant Genome
Volume 9. doi: 10.3835/plantgenome2016.01.0005

is also growing demand to produce higher-quality, more nutritious food. Bread wheat is a staple of many diets, with milled flour used for a variety of products including leavened and unleavened breads, noodles, cookies, cakes, and pastries. Each of these products demands wheat with a specific best-fit quality profile specifically considering the protein concentration, grain hardness, and gluten strength (Peña, 2002; Peña et al., 2002).

Processing and end-use quality for wheat is a combination of many defined parameters. Multiple phenotypic traits of the grain, flour, dough, and final products must be assessed to determine an overall quality and best end-use product. Typically, hard grain with high protein and strong and extensible gluten is acceptable for making leavened breads and industrial pan bread, hard or medium hard grain with intermediate levels of protein and good dough extensibility make good flat breads, and soft grain with low protein and weak and extensible gluten is used for cookies, cakes, and pastries (Peña, 2002; Peña et al., 2002). Many laboratory tests must be considered to ensure that candidate wheat varieties meet the quality profile for a given end-use product. Some of these tests are direct targets of selection with defined thresholds, whereas other should be interpreted collectively or are used to inform for further stages of testing. CIMMYT uses grain hardness, grain protein, Alveograph W (dough strength) and P/L (tenacity vs. extensibility) values, along with final loaf volume collectively to approximate the product for which a wheat line would be best suited (R.J. Peña, unpublished data, 2011).

Wheat grain is assessed for premilling characteristics, which impact marketing. These tests include kernel weight, weight per volume, color, hardness, vitreousness of the kernel, and total protein content. Many of these characteristics are strongly correlated with grain yield with varying levels of heritability. Grain weight was found to have increased in CIMMYT bread wheats over time and was significantly correlated with yield (Aisawi et al., 2015). In contrast, grain protein content is often negatively correlated with yield and is highly impacted by environment and agronomic management (Terman et al., 1969).

Wheat endosperm texture also plays an important role in milling and end-use targets. Hard and soft wheat differ in the strength of which starch granules are attached to the protein matrix. Hard wheat has much stronger attachment, thus requiring more energy expenditure during milling and damaging more starch than in soft endosperm wheat (Giroux and Morris, 1997). Higher damaged starch increases the amount of water that is absorbed by the dough, which is favored in baking leavened breads compared with making cookies and pastries. The majority of genetic variation for hardness controlled by *Ha* hardness genes located on the short arm of chromosome 5D, but is also impacted by other small-effect loci (Pasha et al., 2010). In industrial markets, kernel size, volume, and protein tests are often used for bulk purchasing and allow the wheat to be sorted into marketing classes (e.g., hard white, hard red winter, soft white, etc.). This segregation to market classes has a large impact for milling and flour mixing by millers to ensure consistent end-use products over time. In local markets in the developing world, the visual characteristics of a cultivar are extremely important, as much of this wheat will be purchased as grain then milled and used at home.

The next stage of testing focuses on milling extraction, protein concentration, and moisture of the flour. Increase in flour yield is profitable for millers, but it is important to note that optimal flour yield is attained when mill rollers and sieves are set appropriately for the common shape and size of a specific wheat line. As such, experimental test mills cannot be reset for each genotype and commercial mills mill a mixture of all different varieties. Protein and moisture tests of the flour are often used for estimating water absorption in the dough.

Wheat dough is special among cereals for its viscoelastic ability to rise and extend while still retaining shape and connectivity. The viscoelastic properties of wheat are primarily conditioned by the storage proteins, glutenins, and gliadins (Delcour and Hoseney, 2010; Garg et al., 2006; Payne et al., 1987; Zheng et al., 2009). Glutenins are responsible for the elasticity and resistance to extension properties of wheat dough, whereas gliadins are responsible for the cohesive properties of wheat dough, which allow it to rise and retain gas (Delcour and Hoseney, 2010). Additionally, other constituents of the wheat kernel, such as nonstarch polysaccharides, enzymes, oligosaccharides, and damaged starch may also have impacts on dough rheology and end-use quality (Delcour and Hoseney, 2010). Dough rheology and end-use tests involve mixing flour to dough to determine the viscoelastic properties of strength, elasticity, tolerance, and final outcomes of wheat when optimally mixed. These tests are time consuming, costly, and require large quantities of flour to conduct. However, each of these tests are collectively necessary to determine the quality profile and a suitable end-use product for a specific wheat line (Peña, 2002; Peña et al., 2002).

Historically, the primary focus of wheat breeding is grain yield combined with visual selection for lines with improved agronomic performance and disease resistance. In many breeding programs, quality traits are evaluated as a final performance test because the tests are expensive and usually cannot occur until later in the breeding program because of the large amount of grain necessary. This often results in promising wheat cultivars that cannot be released because of poor quality. In addition, there is limitation for developing any wheat cultivars with good and specialized end-use traits. Accurate processing and end-use quality prediction models would allow breeding programs to cull unacceptable lines or target specific lines earlier in the pipeline before time and resources are invested in lines that will not pass the final test.

In the scope of breeding for quantitative traits, marker-assisted selection with previously identified significant markers has limited prediction power for complex traits (Heffner et al., 2011b). Genomic selection

models, however, use high-density genotype data sets and simultaneously model all additive genetic variance. These models use entries with known phenotype and genotype to train an algorithm, cross-validate the prediction, and then predict traits in materials with only genotype information available. This approach was first introduced into animal breeding by Meuwissen et al. (2001) demonstrating that ridge regression and Bayesian approaches could be used to model the total additive variance and predict breeding values. Their claim that attaining genome-wide marker profiles would become cheaper than phenotyping each individual is becoming a reality (Poland and Rife, 2012). Taking all this into consideration, GS could serve as a way to predict processing and end-use quality phenotypes earlier in the pipeline before breeders have enough seed for testing and allow predictions of more individuals than would be possible to phenotype.

Genomic selection has been evaluated many times for wheat yield and disease resistance (Arruda et al., 2015; Crossa et al., 2010, 2014; Dawson et al., 2013; Heffner et al., 2009; Poland et al., 2012b; Rutkoski et al., 2010, 2012, 2014) but not thoroughly for wheat processing and end-use quality. Genomic selection was tested in soft wheat for end-use quality in a biparental population and a small breeding population (Heffner et al., 2009, 2011a). These studies relied on cross-validation, rather than forward prediction approaches, to assess the prediction accuracy of the GS models. They did find processing and end-use quality traits to be more highly predictive than grain yield.

Here we conducted forward prediction in the breeding program with GS models on all important processing and end-use quality traits that are regularly assessed by the CIMMYT bread wheat breeding program. The objective of this study was to determine prediction accuracy of several GS models for these complex processing and end-use quality traits, assess the accuracy of forward prediction into the next year, and introduce GS for end-use quality to the CIMMYT bread wheat breeding program.

Materials and Methods

Germplasm

Wheat lines used in training and testing the GS models were from F_5–derived F_7 lines in first-year yield trials grown in Ciudad Obregon, Sonora, Mexico, and advanced to quality testing in the CIMMYT spring bread wheat breeding program. In concordance with the selection and advancement of material in the program, a given line was only evaluated for quality in 1 yr. Materials were planted in lattice designs with 28 entries to every two checks in two replications. Only those selected for superior yield or other agronomic performance were advanced to processing and end-use quality testing. A single sample from one replication was used to measure grain, flour, dough, and end-use quality phenotypes for each selected wheat line.

Phenotypes

Processing and end-use quality phenotypes were assessed from first-year yield trials in the CIMMYT bread wheat breeding program. Thus, as noted below, some methods have been altered from American Association of Cereal Chemists (AACC) standards to increase throughput, decrease sample size, or increase variance present among samples for breeding selection. Near-infrared spectroscopy (NIRS) data were also used to train the GS models, since this was the data made available on large sample sizes in the breeding program.

Grain morphological characteristics were evaluated with the digital image system SeedCount SC5000 (Next Instruments) and weighed to obtain 1000-kernel weight (TKW [g]), which has high correlation between hand-counted and image-analyzed TKW ($R^2 > 0.95$; C. Guzmán, unpublished data, 2016). A 37.81-mL sample was weighed to obtain test weight (TESTWT, kg hL^{-1}), which has high correlation between small-scale and full-scale TESTWT ($R^2 > 0.95$; C. Guzmán, unpublished data, 2016). Grain protein (GRNPRO), hardness (GRNHRD) based on particle size index (R^2 0.8–0.9; C. Guzmán, unpublished data, 2016), and moisture content were determined by NIRS using NIR System 6500 (Foss) by the official methods AACC 39-10, 39-70A, and 39-00, respectively (AACC, 2000). The GRNPRO was reported at 12.5% moisture basis. Grain samples were optimally tempered to 13 to 16.5%, depending on grain hardness, and milled using Brabender Quadrumat Jr. (C. W. Brabender OHG). Flour protein (FLRPRO) and moisture content were estimated by NIRS using the Antaris II FT-NIR analyzer (Thermo). Both NIRS instruments were calibrated for particle size index (AACC Method 55-30), moisture (AACC Method 44-15A), and protein (AACC Method 46-11A). Sodium dodecyl sulfate (SDS) sedimentation (FLRSDS) was conducted as in Peña et al. (1990). Dough rheology was assessed using the Swanson and Working Mixograph (National Mfg. Co.) according to AACC method 54-40A (AACC, 2000), and the Chopin Alveograph (Tripette & Renaud), AACC method 54-30A (AACC, 2000). These methods were adjusted to allow for unified optimal water absorption based on solvent retention capacity instead of the AACC standard method based on protein concentration of the sample (Guzmán et al., 2015). Optimal mix time (MIXTIM) and torque (MP) were measured by Mixograph. Dough strength, work value under the curve (ALVW), and tenacity vs. extensibility, the ratio of height to length of the curve (P/L), were measured using Alveograph. Alveograph P/L values (ALVPL) were log transformed prior to analysis for normalization then untransformed for data presentation. To assess end-use quality for yeast-leavened bread, pup loaves were baked as pan bread with AACC method 10-09 (AACC, 2000) using the Guzmán et al. (2015) adjustment for optimal water absorption. Bread loaf volume (LOFVOL) was measured by rapeseed displacement in accordance with AACC method 10-05.01 (AACC, 2000).

Genotypes

Leaf tissue was collected and bulked from five plants per line and DNA was extracted using a modified CTAB protocol (Saghai-Maroof et al., 1984). For genotyping-by-sequencing, DNA was quantified, normalized to 10 µL at 10 ng μL^{-1}, digested with two-enzymes *PstI* and *MspI*, ligated with barcoded adapters, amplified, and then sequenced following the protocol of Poland et al. (2012a). Sequences were trimmed to 64 bp, unique sequence tags were aligned, and single-nucleotide polymorphisms (SNPs) were recoded numerically as (–1, 0, 1) using TASSEL 5 v2 (Bradbury et al., 2007). The SNPs were aligned with pseudo-positions of the wheat genome using POPSEQ (Chapman et al., 2015; International Wheat Genome Sequencing Consortium, 2014). The SNPs were investigated for percentage missing and heterozygosity. Markers with >20% missing data or >20% heterozygous calls were discarded. Individuals with >80% missing data were also removed from further analysis. Remaining missing SNPs were imputed using mean imputation based on marker frequency using R (R Development Core Team, 2014) package rrBLUP (Endelman, 2011).

Analyses

Genomic selection models were developed using packages in R (R Development Core Team, 2014). Ridge regression best linear unbiased predictor (RRBLUP) and reproducing kernel Hilbert space, here referred to as Gaussian kernel (GAUSS), models were conducted using the package rrBLUP, as described in Endelman (2011). Partial least squares regression (PLSR), elastic net, and random forest were tested using R packages pls (Mevik and Wehrens, 2007), glmnet (Friedman et al., 2009), and randomForest (Liaw and Wiener, 2002), respectively. The PLSR, elastic net, and random forest required model training or specification before implementation. We used a 10-fold cross-validation to train the PLSR model for optimal number of components to be used in the prediction algorithm. Elastic net was also trained using a 10-fold cross-validation in training data to tune the alpha, mixing, and lambda regularization parameters. The random forest predictions were made using 1000 trees. These models were combined by Gaynor (2015) into R package GSwGBS. Correlations for all phenotypes were estimated across years in JMP Genomics 7.1 (SAS Institute Inc., 2013). Models were tested using cross-validation and temporal forward predictions. Cross-validation predictions were conducted on all data across all years with 20% random masked, which was replicated 10 times. Prediction correlations were assessed between predicted phenotypes and empirical phenotypes, and accuracies were determined by dividing the correlation coefficient by the square root of heritability. Forward predictions were conducted using data as it would have historically become available to predict the following year (i.e., 2009 predicts 2010 and 2009 and 2010 predict 2011, etc.).

Heritability could not be calculated by traditional ANOVA methods because of lack of replication. The RRBLUP model's additive genetic variance (Endelman, 2011) and the error variance of this model (V_e) were calculated. We assume that error variance included dominance and epistatic genetic variance along with all environmental and measurement error. Additive genetic variance was estimated by subtracting RRBLUP error variance from phenotypic variance (V_p). Following this, a standard calculation of heritability, additive genetic variance divided by phenotypic variance, was performed (Falconer and Mackay, 1996):

$$h^2 = \frac{V_p - V_e}{V_p} \quad [1]$$

We calculated the relative gain from indirect selection using the genotype data, as possible, with GS compared with direct selection on the quality phenotypes. In the CIMMYT program, it is currently possible to genotype the entire set of up to 10,000 lines in first-year yield testing. Less than 2000 of these lines are advanced for quality testing. A set of ~1000 lines are advanced for second-year replicated testing in multiple environments. Based on these population sizes for selection and quality testing, we have a conservative estimate of 10% selection intensity using GS ($i = 1.755$) and a maximum selection intensity of 50% possible for phenotypic selection ($i = 0.798$).

From Falconer and Mackay (1996) relative gain from indirect selection was then calculated as the ratio of CR_X to R_X based on the following:

$$\frac{CR}{R} = \frac{i_Y h_Y r_A}{i_X h_X} \quad [2]$$

where CR is the correlated response to indirect selection, R is the response to direct selection, i designates the selection intensity for direct selection on X and indirect selection on Y, h is the square root of heritability, and r_A is the additive genetic correlation.

The phenotypic correlation between two traits can be noted as a function of the additive genetic correlation and the environmental correlation (Falconer and Mackay, 1996):

$$r_P = h_X h_Y r_A + e_X e_Y r_E \quad [3]$$

As we are predicting to new environments, we assume no environmental correlation between the GEBVs and phenotypic observations leaving the following:

$$r_P = h_X h_Y r_A \quad [4]$$

Substituting to Eq. [2] gives an estimate of relative gain based on phenotypic correlation:

$$\frac{CR}{R} = \frac{i_Y r_P}{i_X h_X^2} \quad [5]$$

Using this equation, we calculated the relative gain of GS for each quality trait using increased selection intensity.

Table 1. Materials available for genomic selection modeling.

Trial harvest year	Total no. lines in yield trial	No. lines screened for quality	Phenotype and genotype available
2010	4956	1258	250
2011	6685	1000	995
2012	10,196	1580	850
2013	9436	1215	886
2014	7672	1345	1114
2015	8872	1460	1425
Total	47,817	7858	5520

Results and Discussion

Materials and Genotypes

Phenotypic assessment was conducted on 7858 lines in first-year yield trials between 2009 and 2015 for processing and end-use quality. In the first year of the project, 2009, only individuals promoted to advanced testing were genotyped, whereas for other years, all individuals in the first-year yield test were genotyped. This resulted in many fewer individuals present in the first year. Filtering for large amounts of missing genotypic data per individual resulted in 5520 individuals with high-quality genotype and phenotype for GS (Table 1). Originally, 20,833 SNPs were found using the TASSEL 5 v2 pipeline. Since no significant differences were found in model accuracy as marker number increased, SNPs were then restricted to no more than 20% missing to ensure higher accuracy through reduced reliance on imputation. This resulted in 3075 SNPs that were used in the GS models. Markers were used as nonpositioned de novo markers, though A and B chromosomes were more densely covered than D chromosomes.

Phenotype Means

Phenotype distributions of all traits within all years followed an approximately normal distribution (Supplemental Fig. S1, S2), except Alveograph P/L (Supplemental Fig. S2), which was log transformed for subsequent analysis (Box and Cox, 1964). Phenotype mean and standard errors are presented by year (Table 2). Heritability estimates made from RRBLUP error and phenotypic variances were moderate, ranging between 0.41 and 0.68 (Table 2). Previous reports have also estimated the heritability of most processing and end-use quality traits as intermediate (Breseghello and Sorrells, 2006; Kuchel et al., 2006).

Protein assessments were highly correlated (Table 3) as could be expected, since the majority of the protein in the wheat kernel is stored in the endosperm (Delcour and Hoseney, 2010). Most dough rheology traits evaluated here were highly correlated with the exception of Alveograph P/L (Table 3). Phenotypic correlations in this study again demonstrate that no single quality test is a substitute for end-use testing, as the correlations from all other parameters are present but not strongly correlated to final pup loaf volume (Table 3). This further supports classification systems for end use as a function of several quality phenotypes (Peña, 2002).

Genomic Selection

Reflecting the power of large training populations, GS prediction accuracy in forward prediction increased over time as more lines and environments were added to the models. Cross-validation accuracy was, on average, 31.8% greater than all models in all years and 19% more accurate, on average, than 2015 forward predictions. This higher accuracy varied across traits, where ALVPL demonstrated <1% change in accuracy between cross-validation and forward prediction in contrast with a 44% difference for the TESTWT predictions. This demonstrates that forward prediction is much more difficult, and cross-validations likely overinflate results breeders may expect with implementation of GS. With unbalanced data from breeding programs and confounded cohorts of full-sib and closely related lines being evaluated under common environments (years), the cross-validation approach of genotype-only models are likely overestimating the predictive ability for GS.

Genomic selection models used in this study produced very similar results in forward prediction, with the exception of random forest having lower prediction accuracy (Fig. 1, 2). It is known that models have differing performance with varying genetic architecture of traits. The GAUSS was the best prediction model for all traits in cross-validation (Fig. 3). However, GAUSS did not have the same significant advantage in forward predictions (Fig. 1, 2). This indicates that GAUSS could be benefitting from the full siblings present within year but not across year as is common in breeding programs. The full-sib testing within a year would have confounded environment with cross-validation that would benefit from weighting of kinship in the GAUSS models, but this model would not predict as well into new years (environments) with more distantly related breeding lines.

Response to selection using GS as a correlated trait indicated that for all traits, GS has increased response to selection than phenotypic selection. Increases in section response ranged from 35 (TESTWT) to 147% (ALVPL) over phenotypic selection (Table 4). For CIMMYT wheat quality screening, 10% selection intensity (i = 1.755) was used for GS, and 50% selection intensity (i = 0.798) was used for a phenotypic selection value. These increases in response to selection from GS largely are due to the increased selection intensity, where the full yield trial may be genotyped, rather than phenotyping only those individuals passing the yield test.

For the kernel traits, data for TKW was only available starting in 2012 (Table 2). Prediction correlations for TKW increased from 0.40 to 0.49 over time (Table 5). Random forest performed worst for this trait (Fig. 1). The TESTWT predictions also increased with time from 0.1 in the first year to 0.34 in 2015 (Table 5; Fig. 1). Initially, there was no predictive ability for GRNHRD in 2011. As with TESTWT, a larger training set over time increased

Table 2. Phenotype means and standard errors by year and narrow-sense heritability (h^2) across years.

	2010		2011		2012		2013		2014		2015		
	No. of entries												
	250		995		850		886		1114		1460		ALL
Phenotype†	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	h^2
TKW					48.30	0.13	46.57	0.11	47.79	0.11	44.60	0.09	0.60
TESTWT	82.43	0.06	80.15	0.05	82.37	0.03	81.83	0.03	81.74	0.03	80.81	0.03	0.56
GRNHRD	40.75	0.36	45.77	0.15	40.31	0.16	42.95	0.11	43.56	0.09	45.86	0.06	0.41
GRNPRO	12.07	0.05	11.73	0.02	11.31	0.02	11.70	0.02	12.23	0.02	12.45	0.05	0.43
FLRYLD			67.55	0.11	68.83	0.08	69.35	0.06	70.57	0.06	70.92	0.02	0.55
FLRPRO	10.22	0.05	10.20	0.02	9.57	0.02	9.99	0.02	10.71	0.02	11.03	0.02	0.57
FLRSDS	14.86	0.15	14.35	0.07	13.83	0.08	14.05	0.26	13.68	0.06	14.79	0.06	0.62
MIXTIM	2.75	0.04	3.15	0.02	3.11	0.02	3.35	0.03	2.97	0.02	3.27	0.02	0.68
MP			106.12	1.00	116.41	0.92	123.02	1.10	113.16	0.83	126.90	0.83	0.63
ALVW	285.88	5.70	256.68	2.17	271.58	2.33	291.74	3.10	253.06	2.49	263.70	2.19	0.65
ALVPL	1.04	0.02	0.93	0.01	1.03	0.01	0.99	0.01	0.96	0.01	0.74	0.01	0.46
LOFVOL	746.12	4.22	785.25	1.49	752.46	2.48	807.83	1.85	822.59	1.72	821.64	1.30	0.63

† TKW, 1000-kernel weight (g); TESTWT, test weight (kg hL^{-1}); GRNHRD, grain hardness (PSI); GRNPRO, grain protein (at 12.5% moisture basis); FLRYLD, flour yield from milling (percentage recovered); FLRPRO, flour protein (at 14% moisture basis); FLRSDS, flour sodium dodecyl sulfate sedimentation volume (mL); MIXTIM, optimum mix time (min); MP, torque at the integral of the midline peak; ALVW, work value from Alveograph curve (J); ALVPL, Alveograph P (tenacity) divided by L (extensibility) (mm mm^{-1}); LOFVOL, pup loaf volume (cm^3).

Table 3. Phenotypic correlations over all years.

	TKW†	TESTWT	GRNHRD	FLRYDL	GRNPRO	FLRPRO	FLRSDS	MIXTIM	MP	ALVW	ALVPL
TESTWT	0.08										
GRNHRD	–0.23	–0.20									
FLRYLD	–0.21	0.03	0.18								
GRNPRO	0.04	–0.04	–0.17	–0.13							
FLRPRO	0.04	–0.05	–0.20	–0.09	0.94						
FLRSDS	0.02	0.00	–0.20	–0.18	0.45	0.46					
MIXTIM	–0.13	0.01	–0.12	–0.12	–0.09	–0.09	0.40				
MP	–0.10	0.04	–0.16	–0.16	–0.01	0.00	0.48	0.97			
ALVW	–0.06	0.04	–0.22	–0.20	0.20	0.22	0.61	0.84	0.90		
ALVPL	0.11	0.16	–0.28	–0.26	–0.19	–0.19	0.03	0.36	0.41	0.43	
LOFVOL	–0.13	–0.17	–0.05	–0.03	0.51	0.53	0.58	0.30	0.34	0.44	–0.27

† TKW, 1000-kernel weight (g); TESTWT, test weight (kg hL^{-1}); GRNHRD, grain hardness (PSI); GRNPRO, grain protein (at 12.5% moisture basis); FLRYLD, flour yield from milling (percentage recovered); FLRPRO, flour protein (at 14% moisture basis); FLRSDS, flour sodium dodecyl sulfate sedimentation volume (mL); MIXTIM, optimum mix time (min); MP, torque at the integral of the midline peak; ALVW, work value from Alveograph curve (J); ALVPL, Alveograph P (tenacity) divided by L (extensibility) (mm mm^{-1}); LOFVOL, pup loaf volume (cm^3).

this prediction to 0.32 correlation between the observed and predicted in 2015 (Table 5; Fig. 1). The GRNHRD also had one of the lowest predictive accuracies in cross-validation (Fig. 4). Though hardness is under the control of major-effect loci, within a given breeding program, lines are generally fixed for the major alleles, leaving minimal genetic variance and low heritability. The results observed here corroborate with Heffner et al. (2011a) who found that softness had lower prediction accuracy than other quality traits. Although there was a normally distributed phenotypic range for GRNHRD (Supplemental Fig. S1), most materials in this data set are still classified as hard or semihard with few soft lines present. A high proportion of the CIMMYT historical and breeding lines previously tested had the haplotype *Pina-D1b* and *Pinb-D1a* alleles for the hardness (*Ha*) genes on the short arm of chromosome 5D (Lillemo et al., 2006). Protein concentration, where more protein leads to harder grain, may be one of the factors responsible for some of the smaller differences found within hardness class (Pasha et al., 2010).

With the milling traits, FLRYLD data was first available in 2011 for prediction in 2012 (Table 2, 3). This was the only trait that did not show a marked increase in prediction accuracy over time. The predictions for FLRYLD were highest in the first year of testing and dropped slightly in the following years (Table 5; Fig. 1). Grain and flour protein are very highly correlated phenotypes (Table 3) and follow very closely to one another in predictive ability for GS (Table 5; Fig. 1). In these traits, we have seen a general increase over time. The FLRSDS, which is correlated to both protein and dough rheology traits (Table 3), has intermediate prediction accuracy (Fig. 2, 3) but may have come to a forward accuracy plateau of between 0.5 and 0.6 (Fig. 2; Table 5).

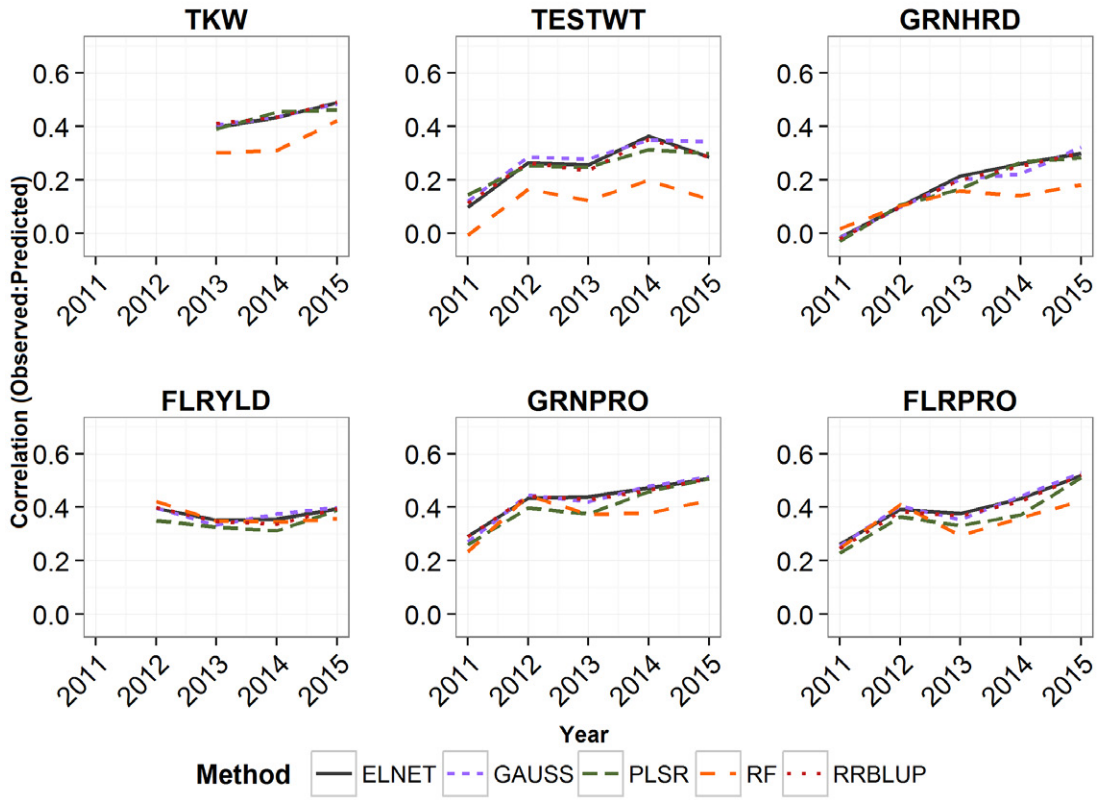

Fig. 1. Genomic selection forward prediction correlations for 1000-kernel weight (TKW), test weight (TESTWT), grain hardness (GRNHRD), flour yield (FLRYLD), grain protein (GRNPRO), and flour protein (FLRPRO) over time.

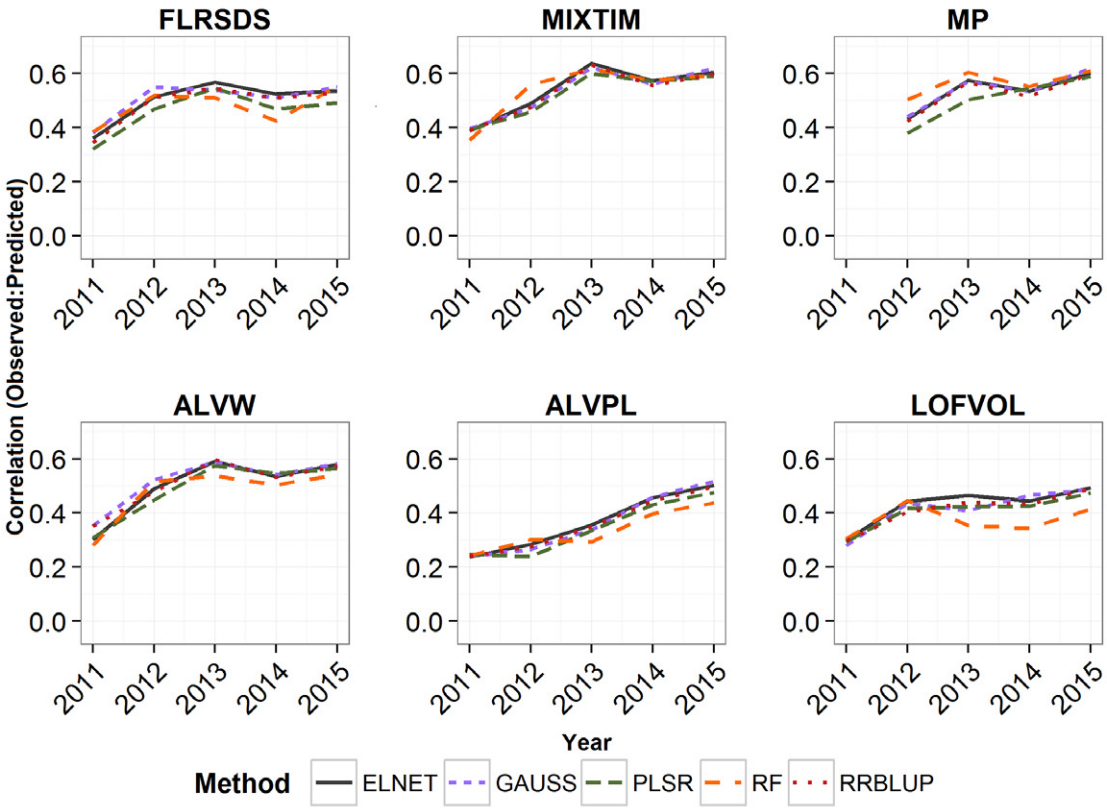

Fig. 2. Genomic selection forward prediction correlations for flour sodium dodecyl sulfate sedimentation (FLRSDS), Mixograph mix time (MIXTIM), Mixograph torque (MP), Alveograph dough strength (W) and tenacity vs. extensibility (P/L) (ALVW and ALVPL), and loaf volume (LOFVOL) over time.

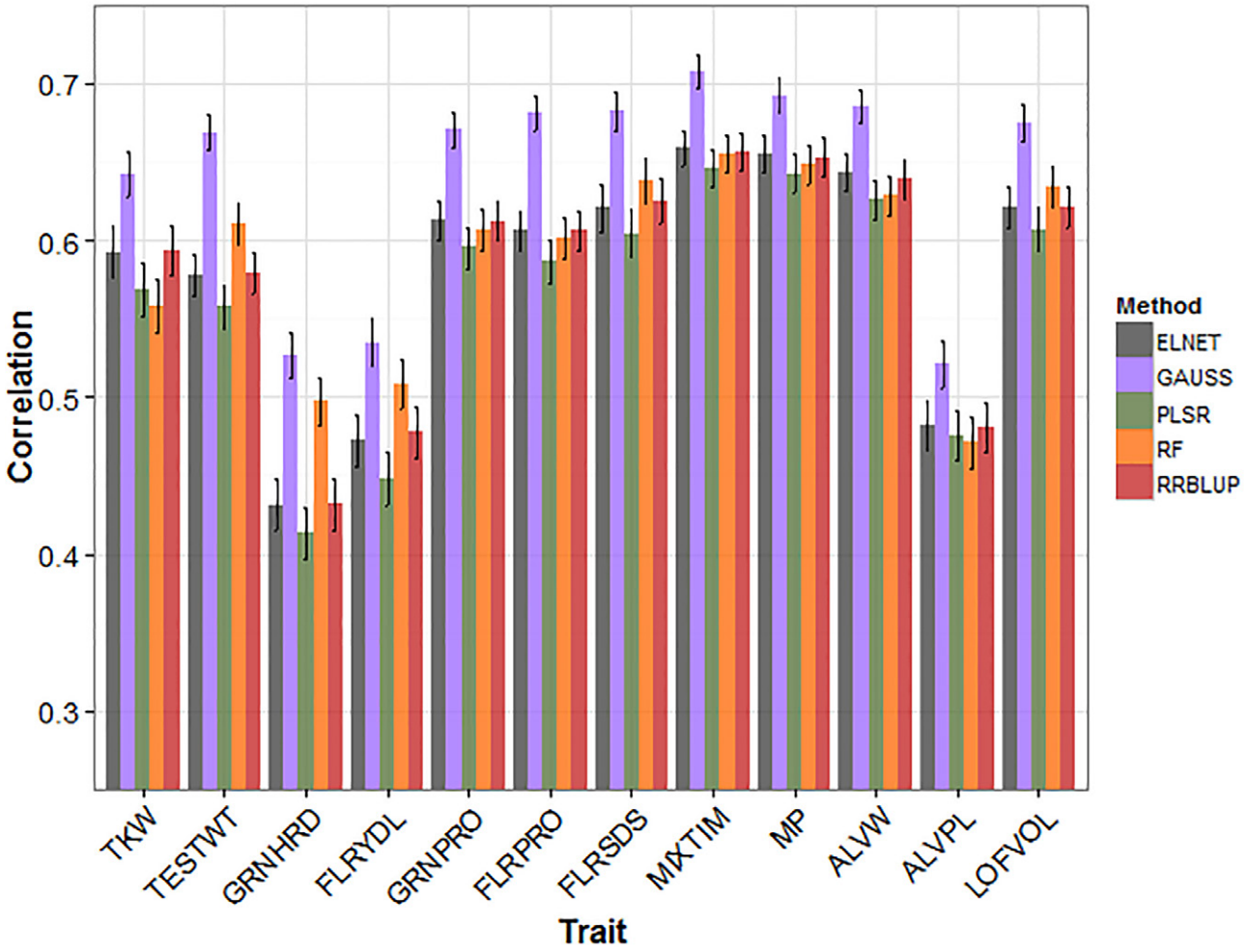

Fig. 3. Genomic selection cross-validation correlations for all methods. TKW, 1000-kernel weight (g); TESTWT, test weight (kg hL^{-1}); GRNHRD, grain hardness (PSI); GRNPRO, grain protein (at 12.5% moisture basis); FLRYLD, flour yield from milling (percentage recovered); FLRPRO, flour protein (at 14% moisture basis); FLRSDS, flour sodium dodecyl sulfate sedimentation volume (mL); MIXTIM, optimum mix time (min); MP, torque at the integral of the midline peak; ALVW, work value from Alveograph curve (J); ALVPL, Alveograph P (tenacity) divided by L (extensibility) (mm mm^{-1}); LOFVOL, pup loaf volume (cm^3). Method: ELNET, elastic net; GAUSS, Gaussian kernel; PLSR, partial least squares regression; RF, random forest; RRBLUP, ridge regression best linear unbiased predictor.

Dough rheology traits are the foundation of determining gluten type classification at CIMMYT (R.J. Peña, unpublished data, 2011), which is part of their assessment for best-suited end-use quality type. For example, strong gluten is typically favored in pan breads; medium strength gluten is better for flat breads and noodles; weak gluten is best for cakes, cookies, and pastries; and tenacious gluten is only acceptable as wheat for animal feed (Peña, 2002). For these traits, MIXTIM, MP, and ALVW were all highly correlated (Table 3). The Mixograph percentage torque and ALVW are measures of gluten strength, while ALVPL is a better indication of the balance of dough tenacity and extensibility. The Mixograph mix time, MP, and ALVW had good prediction accuracy with forward (Fig. 2; Table 5) and cross-validation (Fig. 3; Table 5). As ALVPL was not normally distributed, we employed log transformation, which lead to ~5% increase in prediction accuracy for this trait (data not shown). Still, ALVPL prediction correlations were lower than other dough rheology traits (Fig. 2, 3; Table 5), likely a result of this trait being a ratio of two independent measures. Unfortunately, data were not recorded separately as P and L values over the years, so this comparison could not be made. The ALVPL forward prediction correlations increased from 0.24 to 0.52 in 2015 with the addition of data over time (Fig. 2).

Baking a pup loaf is the final end-use quality test to determine appropriateness of a wheat line for industrial pan bread. This test gives quantitative and qualitative results not only of the loaf size but also how the loaf and crumb structure appear and can then also be used in staling studies. Here we demonstrate that forward prediction accuracy of LOFVOL was 0.49 in 2015 (Fig. 2; Table 5).

Prediction accuracy of whole-genome models was lower in forward prediction than cross-validation for all traits (Fig. 4; Table 5). This likely is due to cross-validation models using training and testing data containing all years thus better accounting for environmental variation in the test set. Another reason could be due to the possibility of full siblings being randomly assigned to

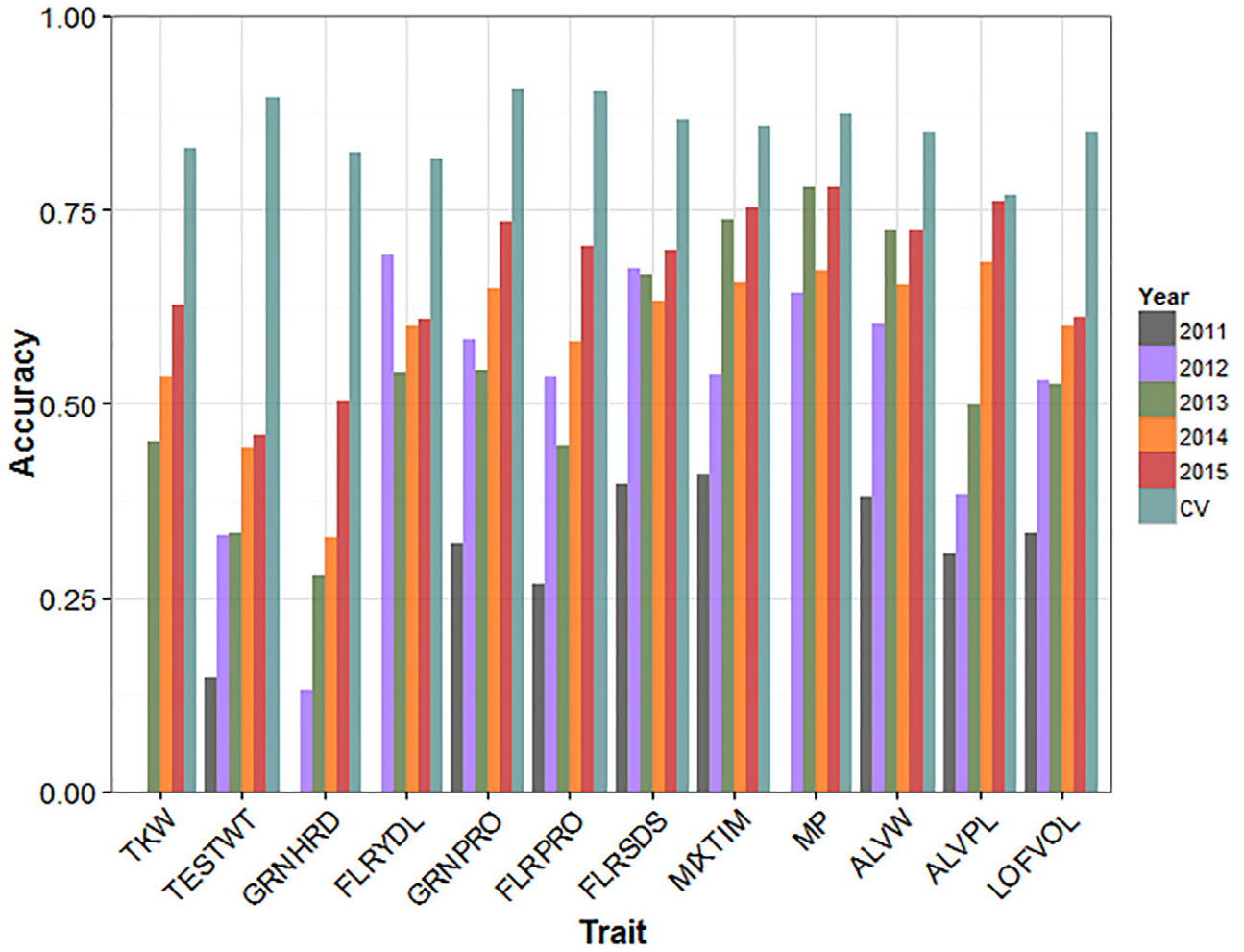

Fig. 4. Genomic selection accuracy over time using Gaussian kernel genomic selection method. TKW, 1000-kernel weight (g); TESTWT, test weight (kg hL^{-1}); GRNHRD, grain hardness (PSI); GRNPRO, grain protein (at 12.5% moisture basis); FLRYLD, flour yield from milling (percentage recovered); FLRPRO, flour protein (at 14% moisture basis); FLRSDS, flour sodium dodecyl sulfate sedimentation volume (mL); MIXTIM, optimum mix time (min); MP, torque at the integral of the midline peak; ALVW, work value from Alveograph curve (J); ALVPL, Alveograph P (strength) divided by L (extensibility) (mm mm^{-1}); LOFVOL, pup loaf volume (cm^3).

training and testing sets in cross-validation. The selection procedure in the breeding program keeps all good material, regardless of their relationship, and sometimes favors advancement of large groups of full siblings. In the full yield trial of 2014 (n = 7672), there was an average of 5.3 entries per cross with a maximum of 51 full siblings for one specific cross (data not shown). Thus, we assume cross-validation represents an overinflation compared with forward predictions not accounting for genotype-by-environment interactions and leveraging more information from relatives than would be available for prediction and selection into new crosses and generations of the breeding program.

Conclusions

Wheat quality is typically not the primary breeding objective; it is often secondary to yield, agronomic performance, and disease resistance. Though quality traits are important selection criteria, the population sizes that can be assessed for quality are generally much smaller, often unreplicated, and testing occurs at a later stage in the breeding pipeline because of cost and quantity of seed needed. However, with the implementation of GS for wheat quality, predictions that are available in earlier generations can enable better selection for quality and even targeting of wheat lines to potential areas of specific end use. The models here demonstrate that GS for processing and end-use quality has sufficient accuracy for implementation in the breeding program. In addition, prediction correlations and accuracies increased over time, likely a result of increasing training population size and the incorporation of data from additional environments into the training model. Finally, GS is heavily favored as a selection tool having high response to selection because of the increase in selection intensity that can occur when screening all available materials.

There are known genes of large effect in wheat for milling and baking quality leading to potential for marker-assisted selection. However, there are also impacts on these traits by more quantitative sources as well. Furthermore, many of the large-effect loci for quality will be fixed within a given breeding program. Indeed, the

Table 4. Comparison of response to selection using direct (phenotypic) and indirect (genomic) selection. Phenotypic selection threshold of 50% and genomic selection threshold of 10% giving selection intensity (*i*) of 0.798 and 1.755, respectively.

Phenotype†	Narrow-sense heritability (h^2)	r_p‡	Response	
			CR/*R*§	Increase in response
				%
TKW†	0.60	0.485	1.78	78
TESTWT	0.56	0.343	1.35	35
GRNHRD	0.41	0.322	1.73	73
FLRYLD	0.43	0.399	2.04	104
GRNPRO	0.55	0.545	2.18	118
FLRPRO	0.57	0.530	2.04	104
FLRSDS	0.62	0.550	1.95	95
MIXTIM	0.68	0.620	2.01	101
MP	0.63	0.619	2.16	116
ALVW	0.65	0.583	1.97	97
ALVPL	0.46	0.516	2.47	147
LOFVOL	0.63	0.486	1.70	70

† TKW, 1000-kernel weight (g); TESTWT, test weight (kg hL^{-1}); GRNHRD, grain hardness (PSI); GRNPRO, grain protein (at 12.5% moisture basis); FLRYLD, flour yield from milling (percentage recovered); FLRPRO, flour protein (at 14% moisture basis); FLRSDS, flour sodium dodecyl sulfate sedimentation volume (mL); MIXTIM, optimum mix time (min); MP, torque at the integral of the midline peak; ALVW, work value from Alveograph curve (J); ALVPL, Alveograph P (tenacity) divided by L (extensibility) (mm mm^{-1}); LOFVOL, pup loaf volume (cm^3).

‡ r_p, correlation between forward-predicted and empirical value in 2015.

§ CR/*R*, response to selection of correlated trait.

Pin alleles are essentially fixed in this material, removing variance from this large effect locus for grain hardness (data not shown). In the CIMMYT bread wheat program, marker-assisted selection is not conducted for quality except for SDS polyacrylamide gel electrophoresis for high- and low-molecular weight glutenins on parental lines. Currently, SDS polyacrylamide gel electrophoresis is more expensive than genotyping-by-sequencing and is conducted on far fewer selected entries. A comparison could be made between these marker systems, but it would severely reduce the population size and bias the training and testing populations, since the parental materials have already been selected for quality.

The CIMMYT bread wheat breeding program does not breed solely for one end-use quality product such as industrial leavened bread. Rather the focus is on wheat lines suitable to different wheat products found around the world. Therefore, the program strives to increase processing and end-use quality but does not necessarily cull materials for not fitting into industrial wheat bread-making standards. Following these ideas, we did not use a selection index for wheat quality specified to target a specific end-use product. Instead, all phenotypes were predicted with an indication of performance relative to standard checks and the mean of the predicted set for each trait, where one and two standard deviations from predicted mean were noted. The greatest advantage of predicted values was that they are available at harvest time on all materials in the yield trial, enabling selection for quality along with agronomic traits and yield in the progression of the breeding cycle. As quality testing can only be started after completing harvest, phenotypic data is typically not available in time for selection decisions for the next field nursery season. Additionally, traditionally, only those that were selected on yield performance were phenotyped, whereas with GS, all materials were screened.

Phenotyping for all wheat processing and end-use quality for the traits included in this study requires at least 700 g of seed and represents a considerable cost beyond that of yield testing. Wheat breeding programs may screen lines for traits that can be assessed in small samples, such as protein, SDS sedimentation, solvent retention capacity, or Mixograph, in early generations but typically do not have enough seed for most quality tests until after preliminary yield tests with a full yield plot. Genotyping a wheat line can be conducted immediately following line derivation in a nondestructive manner. This genomic profile gives potential for prediction of these quality traits years earlier than phenotyping as well as predicting many other traits. Still, we do not consider GS a replacement for phenotypic selection, but GS can be used to make more informed selection decisions for material advancement between harvest and planting of the next cycle and for prioritizing what is evaluated in the quality labs.

There is a distinct advantage for implementing GS for selection on quality, as considerably larger population sizes can be evaluated through genotyping than by phenotypic assessment of milling and baking. Reflecting this, the expected gain from selection is 1.4 to 2.7 times greater for GS when increasing the number of selection candidates from 2000 to 10,000. When comparing different breeding schemes, it is important to make a cost analysis to estimate an equal operating budget for each different breeding approach (Heffner et al., 2010; Lorenz, 2013). However, as selection for yield will be the primary driver of adoption of GS in breeding programs, the cost of genotyping will be largely offset by reduced or optimized yield testing, leaving a marginal indirect cost for predicting other traits. As the genotypes can be used to predict any trait, there is no additional cost for predicting quality phenotypes on top of the predictions for yield. Regardless, the actual cost associated with phenotyping 1000 breeding lines for quality at $100 to $200 per sample is roughly equivalent to genotyping 10,000 lines at $10 to $20 per sample.

Genomic selection for processing and end-use quality at CIMMYT has now been in development since 2012. In 2014, predictions for end-use and processing quality were available in the fall before phenotype assessments were completed. In 2015, quality phenotypes were predicted in the spring at the time of harvest for 9100 lines in first-year yield trials. These predicted phenotypes, with correlations and accuracies from the 2014 cycle, were used in the breeding program for selection decisions and line advancement. It is expected that predictive information regarding end-use quality earlier in the breeding program

Table 5. Genomic selection (GS) prediction correlations and accuracies of forward and cross-validation. Forward predictive models trained on all prior data, whereas cross-validation trained on a random 80% of the data to predict the remaining masked 20%. Cross-validation through 2014 was replicated 10 times. Both strategies used GAUSS GS model.

	Validation population											
	2011		2012		2013		2014		2015		Cross-validation	
	Training size											
	250		995		2095		2981		4095		3276	
	Testing size											
	995		850		886		1114		1.425		819	
Phenotype†	r‡	Accuracy	r	Accuracy	r	Accuracy	r	Accuracy	r	Accuracy	r	Accuracy
TKW	–	–	–	–	0.401	0.451	0.434	0.534	0.485	0.626	0.642	0.829
TESTWT	0.119	0.145	0.286	0.330	0.279	0.333	0.348	0.442	0.343	0.458	0.669	0.894
GRNHRD	–0.017	–0.018	0.098	0.131	0.202	0.277	0.223	0.329	0.322	0.503	0.527	0.823
FLRYLD	–	–	0.398	0.693	0.333	0.540	0.375	0.600	0.399	0.608	0.535	0.816
GRNPRO	0.270	0.320	0.447	0.582	0.421	0.544	0.480	0.647	0.545	0.735	0.671	0.905
FLRPRO	0.256	0.267	0.404	0.535	0.354	0.446	0.441	0.579	0.530	0.702	0.681	0.902
FLRSDS	0.381	0.395	0.551	0.673	0.538	0.667	0.510	0.633	0.550	0.699	0.682	0.866
MIXTIM	0.396	0.408	0.474	0.537	0.621	0.737	0.561	0.657	0.620	0.752	0.707	0.857
MP	–	–	0.440	0.642	0.572	0.778	0.533	0.672	0.619	0.780	0.693	0.873
ALVW	0.352	0.380	0.522	0.603	0.589	0.725	0.542	0.652	0.583	0.723	0.686	0.851
ALVPL	0.240	0.307	0.266	0.384	0.338	0.498	0.458	0.683	0.516	0.761	0.521	0.768
LOFVOL	0.279	0.333	0.436	0.529	0.407	0.525	0.466	0.602	0.486	0.612	0.675	0.850

† TKW, 1000-kernel weight (g); TESTWT, test weight (kg hL^{-1}); GRNHRD, grain hardness (PSI); GRNPRO, grain protein (at 12.5% moisture basis); FLRYLD, flour yield from milling (percentage recovered); FLRPRO, flour protein (at 14% moisture basis); FLRSDS, flour sodium dodecyl sulfate sedimentation volume (mL); MIXTIM, optimum mix time (min); MP, torque at the integral of the midline peak; ALVW, work value from Alveograph curve (J); ALVPL, Alveograph P (tenacity) divided by L (extensibility) (mm mm^{-1}); LOFVOL, pup loaf volume (cm^3).

‡ *r*, forward prediction accuracy.

will enable selections that will be made for specific end-use quality products in the near future of wheat breeding.

Acknowledgments

This research and graduate assistantship for SB was supported through the Monsanto Beachell-Borlaug International Scholars Program. Funding for this project was provided by US Agency for International Development Feed the Future Initiative (USAID Cooperative Agreement No. AID-OAA-A-13-0005) and the Bill & Melinda Gates Foundation through a grant to Cornell University for "Genomic Selection: The next frontier for rapid gains in maize and wheat improvement." Support for phenotyping of quality traits was provided by CGIAR CRP WHEAT, Durable Rust Resistance Project, and Fondo Sectorial SAGARPA-CONACYT (No. 146788– "Sistema de mejoramiento genético para generar variedades resistentes a royas, de alto rendimiento y alta calidad para una producción sustentable en México de trigo") of the Mexican government. The computing for this project was performed on the Beocat Research Cluster at Kansas State University, which is funded in part by NSF grants CNS-1006860, EPS-1006860, and EPS-0919443. This work represents contribution number 16-037-J from the Kansas Agricultural Experiment Station.

References

Aisawi, K.A.B., M.P. Reynolds, R.P. Singh, and M.J. Foulkes. 2015. The physiological basis of the genetic progress in yield potential of CIMMYT spring wheat cultivars from 1966 to 2009. Crop Sci. 55:1749 doi:10.2135/cropsci2014.09.0601

American Association of Cereal Chemists. 2000. Approved methods of analysis. 11th ed. AACC International, St. Paul, MN. http://methods.aaccnet.org/toc.aspx (accessed 19 June 2015).

Arruda, M.P., P.J. Brown, A.E. Lipka, A.M. Krill, C. Thurber, and F.L. Kolb. 2015. Genomic selection for predicting head blight resistance in a wheat breeding program. Plant Genome 8: doi:10.3835/plantgenome2015.01.0003

Box, G.E.P., and D.R. Cox. 1964. An analysis of transformation. J. R. Stat. Soc. [Ser A] 26:211–252.

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635 doi:10.1093/bioinformatics/btm308

Breseghello, F., and M.E. Sorrells. 2006. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* l.) cultivars. Genetics 172:1165–1177 doi:10.1534/genetics.105.044586

Chapman, J.A., M. Mascher, A. Buluç, K. Barry, E. Georganas, A. Session, et al. 2015. A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. Genome Biol. 16:26 doi:10.1186/s13059-015-0582-8

Crossa, J., L. Campos Gde, P. Perez, D. Gianola, J. Burgueno, J.L. Araus, et al. 2010. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics 186:713–724 doi:10.1534/genetics.110.118521

Crossa, J., P. Perez, J. Hickey, J. Burgueno, L. Ornella, J. Ceron-Rojas, et al. 2014. Genomic prediction in CIMMYT maize and wheat breeding programs. Heredity 112:48–60 doi:10.1038/hdy.2013.16

Dawson, J.C., J.B. Endelman, N. Heslot, J. Crossa, J. Poland, S. Dreisigacker, et al. 2013. The use of unbalanced historical data for genomic selection in an international wheat breeding program. Field Crops Res. 154:12–22 doi:10.1016/j.fcr.2013.07.020

Delcour, J., and R.C. Hoseney. 2010. Principles of cereal science and technology. 3rd Ed. AACC International, St. Paul, MN. doi:10.1094/9781891127632

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

Falconer, D.S., and T.F. Mackay. 1996. Introduction to quantitative genetics. Prentice Hill, London.

Friedman, J., T. Hastie and R. Tibshirani. 2009. Glmnet: Lasso and elastic-net regularized generalized linear models. R package version 1.

Garg, M., H. Singh, H. Kaur, and H.S. Dhaliwal. 2006. Genetic control of high protein content and its association with bread-making quality in wheat. J. Plant Nutr. 29:1357–1369 doi:10.1080/01904160600830134

Gaynor, R.C. 2015. GSwGBS: An R package genomic selection with genotyping-by-sequencing. Genomic selection for Kansas wheat. K-State Research Exchange, Manhattan, KS.

Gerland, P., A.E. Raftery, H. Ševčíková, N. Li, D. Gu, T. Spoorenberg, et al. 2014. World population stabilization unlikely this century. Science 346:234–237 doi:10.1126/science.1257469

Giroux, M., and C. Morris. 1997. A glycine to serine change in *puroindoline b* is associated with wheat grain hardness and low levels of starch-surface friabilin. Theor. Appl. Genet. 95:857–864. doi:10.1007/s001220050636

Guzmán, C., G. Posadas-Romano, N. Hernandez-Espinosa, A. Morales-Dorantes, and R.J. Pena. 2015. A new standard water absorption criteria based on solvent retention capacity (SRC) to determine dough mixing properties, viscoelasticity, and bread-making quality. J. Cereal Sci. 66:59–65 doi:10.1016/j.jcs.2015.10.009

Heffner, E.L., J.L. Jannink, H. Iwata, E. Souza, and M.E. Sorrells. 2011a. Genomic selection accuracy for grain quality traits in biparental wheat populations. Crop Sci. 51:2597 doi:10.2135/cropsci2011.05.0253

Heffner, E.L., J.L. Jannink, and M.E. Sorrells. 2011b. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. Plant Genome 4:65–75 doi:10.3835/plantgenome.2010.12.0029

Heffner, E.L., A.J. Lorenz, J.L. Jannink, and M.E. Sorrells. 2010. Plant breeding with genomic selection: Gain per unit time and cost. Crop Sci. 50:1681–1690 doi:10.2135/cropsci2009.11.0662

Heffner, E.L., M.E. Sorrells, and J.L. Jannink. 2009. Genomic selection for crop improvement. Crop Sci. 49:1 doi:10.2135/cropsci2008.08.0512

International Wheat Genome Sequencing Consortium. 2014. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science 345:1251788 doi:10.1126/science.1251788

Kuchel, H., P. Langridge, L. Mosionek, K. Williams, and S. Jefferies. 2006. The genetic control of milling yield, dough rheology and baking quality of wheat. Theor. Appl. Genet. 112:1487–1495 doi:10.1007/s00122-006-0252-z

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

Lillemo, M., F. Chen, X. Xia, M. William, R.J. Peña, R. Trethowan, et al. 2006. Puroindoline grain hardness alleles in CIMMYT bread wheat germplasm. J. Cereal Sci. 44:86–92 doi:10.1016/j.jcs.2006.03.004

Lorenz, A.J. 2013. Resource allocation for maximizing prediction accuracy and genetic gain of genomic selection in plant breeding: A simulation experiment. G3: Genes, Genomes, Genet. 3:481–491. doi:10.1534/g3.112.004911

Meuwissen, T.H., B. Hayes, and M. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Mevik, B.H., and R. Wehrens. 2007. The pls package: Principal component and partial least squares regression in R. J. Stat. Softw. 18:1–24. doi:10.18637/jss.v018.i02

Pasha, I., F. Anjum, and C. Morris. 2010. Grain hardness: A major determinant of wheat quality. Food Sci. Technol. Int. doi:10.1177/1082013210379691.

Payne, P.I., M.A. Nightingale, A.F. Krattiger, and L.M. Holt. 1987. The relationship between HMW glutenin subunit composition and the bread-making quality of British-grown wheat varieties. J. Sci. Food Agric. 40:51–65. doi:10.1002/jsfa.2740400108

Peña, R. 2002. Wheat for bread and other foods. Bread wheat improvement and production. Food and Agriculture Organization of the United Nations, Rome. p. 483–542.

Peña, R.J., A. Amaya, S. Rajaram, and A. Mujeeb-Kazi. 1990. Variation in quality characteristics associated with some spring 1B/1R translocation wheats. J. Cereal Sci. 12:105–112. doi:10.1016/S0733-5210(09)80092-1

Peña, R.J., R. Trethowan, W.H. Pfeiffer, and M.V. Ginkel. 2002. Quality improvement in wheat. Compositional, genetic, and environmental factors. In: A.S. Basra and L.S. Randhawa, editors, Quality improvement in field crops. Food Product Press, New York. p. 1–37.

Poland, J.A., P.J. Brown, M.E. Sorrells, and J.L. Jannink. 2012a. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7:E32253 doi:10.1371/journal.pone.0032253

Poland, J., J. Endelman, J. Dawson, J. Rutkoski, S. Wu, Y. Manes, et al. 2012b. Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genome 5:103 doi:10.3835/plantgenome2012.06.0006

Poland, J.A., and T.W. Rife. 2012. Genotyping-by-sequencing for plant breeding and genetics. Plant Genome 5:92 doi:10.3835/plantgenome2012.05.0005

R Development Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria

Rutkoski, J., J. Benson, Y. Jia, G. Brown-Guedira, J.L. Jannink, and M. Sorrells. 2012. Evaluation of genomic prediction methods for fusarium head blight resistance in wheat. Plant Genome 5:51 doi:10.3835/plantgenome2012.02.0001

Rutkoski, J.E., E.L. Heffner, and M.E. Sorrells. 2010. Genomic selection for durable stem rust resistance in wheat. Euphytica 179:161–173 doi:10.1007/s10681-010-0301-1

Rutkoski, J.E., J.A. Poland, R.P. Singh, J. Huerta-Espino, S. Bhavani, H. Barbier, et al. 2014. Genomic selection for quantitative adult plant stem rust resistance in wheat. Plant Genome 7:3. doi:10.3835/plantgenome2014.02.0006

Saghai-Maroof, M.A., K.M. Soliman, R.A. Jorgensen, and R. Allard. 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. USA 81:8014–8018. doi:10.1073/pnas.81.24.8014

SAS Institute. 2013. JMP Genomics 7.1. SAS Inst., Inc., Cary, NC.

Terman, G., R. Ramig, A. Dreier, and R. Olson. 1969. Yield-protein relationships in wheat grain, as affected by nitrogen and water. Agron. J. 61:755–759. doi:10.2134/agronj1969.00021962006100050031x

Zheng, S., P.F. Byrne, G. Bai, X. Shan, S.D. Reid, S.D. Haley, et al. 2009. Association analysis reveals effects of wheat glutenin alleles and rye translocations on dough-mixing properties. J. Cereal Sci. 50:283–290 doi:10.1016/j.jcs.2009.06.008