Published March 11, 2016

RESEARCH

Genomic Prediction using Phenotypes from Pedigreed Lines with No Marker Data

Bilal Ashraf, Vahid Edriss, Deniz Akdemir, Enrique Autrique, David Bonnett, Jose Crossa, Luc Janss, Ravi Singh, and Jean-Luc Jannink*

ABSTRACT

Until now, genomic prediction (GP) in plant breeding has only used information from individuals that have been genotyped. Information from nongenotyped relatives of genotyped individuals can also be used. Single-step GP combines marker and pedigree information into a single relationship matrix to perform GP. The objective of this study was to evaluate single-step GP in a wheat breeding program. We compared the performance of pedigree-based, marker-based, and single-step models (ABLUP, GBLUP, and HBLUP, respectively). Data consisted of 1176 genotyped (via genotyping-by-sequencing) and 11,131 nongenotyped wheat lines replicated in five management environments at the CIMMYT experiment station in Obregon, Mexico. Analyses involved three scenarios: (i) all lines had pedigree information but only some were genotyped, with phenotypes from one or two environments in the 2011–2012 season, (ii) all lines had genotype and pedigree information and phenotypes from four or five environments in the 2012–2013 season, and (iii) the combination of Scenarios 1 and 2. Prediction accuracies were calculated by five-fold cross validation on plant height, maturity, heading date, and grain yield. The single-step HBLUP outperformed GBLUP and pedigree-based ABLUP in all cases. We conclude that the single-step procedure combining pedigree and genomic marker data should be favored where appropriate data is available for GP in wheat breeding programs.

B. Ashraf, V. Edriss, and L. Janss, Aarhus Univ., Dep. of Molecular Biology and Genetics–Center for Quantitative Genetics and Genomics, 8830 Tjele, Denmark; B. Ashraf, V. Edriss, D. Akdemir, and J.-L. Jannink, Plant Breeding and Genetics Section, School of Integrative Plant Science, 240 Emerson Hall, Cornell Univ., Ithaca, NY 14853; E. Autrique, D. Bonnett, J. Crossa, and R. Singh, International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, 06600 El Batan, Mexico; J.-L. Jannink, USDA-ARS, R.W. Center for Agriculture and Health, Ithaca, NY 14853. B. Ashraf and V. Edriss contributed equally to this work. Received 19 Feb. 2015. Accepted 10 Nov. 2015. *Corresponding author (jeanluc.jannink@ars.usda.gov).

Abbreviations: ABLUP, best linear unbiased prediction using pedigree information only; GBS, genotyping-by-sequencing; GBLUP, best linear unbiased prediction using genomic information only; GE, genotype × environment; GEBV, genomic estimated breeding value; GP, genomic prediction; HBLUP, best linear unbiased prediction using both genomic and pedigree information; MAF, minor allele frequency.

Genomic prediction has been suggested as an effective approach to increase selection gain in livestock breeding (Meuwissen et al., 2001), and is also emerging as a powerful tool in plant breeding (Heffner et al., 2009; Jannink et al., 2010). Genomic prediction uses genome-wide molecular markers to estimate breeding values and select individuals before phenotyping, or used to increase selection accuracy given limited phenotyping. Accuracy of GP mainly depends on the amount of information used to derive the prediction equation. Generally, GP methods assume that the whole reference population is genotyped and phenotyped. In practice, however, there are situations when not all individuals are genotyped. These situations are common in animal breeding programs (Su et al., 2012) and are also likely in plant breeding. To use all phenotypic information, it is appealing

Published in Crop Sci. 56:1–8 (2016).
doi: 10.2135/cropsci2015.02.0111

Table 1. Number of genotyped and nongenotyped wheat lines for four traits, recorded in 2 yr, and replicated under different Environments (E).†

		2011–2012		2012–2013				
		E1	E2	E3	E4	E5	E6	E7
Trait	Lines	FlatDrip	Bed5IR	Bed2IR	Bed5IR	BedHeat	Flat5IR	FlatDrip
Height								
Genotyped	1160	0	952	1160	1160	1160	1160	1160
Nongenotyped	9315	0	9312	3	0	0	0	3
Maturity								
Genotyped	1160	0	952	496	1160	1159	1160	0
Nongenotyped	9309	0	9306	3	0	0	0	0
Heading date								
Genotyped	1176	224	952	1160	1160	1160	1160	1160
Nongenotyped	11131	1819	9312	3	0	0	0	3
Grain yield								
Genotyped	1176	224	951	1160	1160	1160	1160	1160
Nongenotyped	10965	1819	9146	0	0	0	0	0

† Environment names consist of two parts: First, "Flat" vs. "Bed" refers to plants sowed on flat land or raised on 80-cm-wide beds with three rows. Second, the IR number refers to the number of times (two or five) the field was irrigated. One irrigation was equivalent to 100 mm of water. "Drip" means the total water (including the water present in the soil) was adjusted to 180 mm. Sowing wheat in Mexico normally starts at the end of autumn. In contrast, the heat management is planted at the end of winter such that the growing period temperature was higher than for other managements.

to combine information of genotyped and nongenotyped relatives in a single relationship matrix to perform GP. Genomic best linear unbiased prediction (BLUP) models have become common in practical GP across animal and plant breeding research due to their relatively low computational demands (Crossa et al., 2010; Gao et al., 2012; Su et al., 2012). Several approaches have been suggested to estimate GEBV (genomic estimated breeding value; Ducrocq and Liu, 2009; Harris and Johnson, 2010). The most sophisticated approach to estimate GEBV that integrates genomic, pedigree, and phenotypic information is known as the "single-step procedure" (Aguilar et al., 2010; Christensen et al., 2012; Legarra et al., 2009). In a dairy cattle study, Su et al. (2012) reported that this single-step approach gave a more accurate GEBV than a multistep procedure. The single-step procedure has been investigated in many livestock studies for GP (Gao et al., 2012; Liu et al., 2014). However, to the best of our knowledge, this approach has not been introduced in plant research for GP. Evaluation procedures can include genomic information only (GBLUP), pedigree information only (ABLUP), or both (HBLUP). Here, we compare these three models for the first time for wheat (*Triticum aestivum* L.) performance prediction using GBS (Poland et al., 2012) data.

In plant breeding, the issue of genotype × environment (GE) interaction is a known challenge. Trials in a breeding program are usually conducted in multiple environments to reduce the risk of discarding lines that potentially perform well in some but not in all environments (Ceccarelli et al., 1994). One way to account for GE is to consider each instance of a trait measured in different environments as different traits (Burgueño et al., 2012; Falconer, 1952). With that kind of model, GE manifests as lack of correlation in the trait between environments (Cooper and DeLacy, 1994). This modeling is useful because it shows the extent to which predictions for one environment can "borrow information" from measurements in another environment (Burgueño et al., 2012). In the end, such information borrowing depends on the genetic architecture of the trait and the extent to which it is similarly conditioned in different environments (El-Soda et al., 2014).

The main aim of this study was to investigate the impact on prediction accuracy when some wheat lines are not genotyped. In addition, we compared the performance of different prediction models (GBLUP, ABLUP, and HBLUP), and determined the impact of different weights for the pedigree relationship on prediction accuracy. Finally, we compared these univariate models to a multivariate GBLUP model where measurements in each environment were considered separate variables.

MATERIALS AND METHODS

Phenotypic Data

Data for this study came from CIMMYT international multienvironmental wheat trials. Four traits were analyzed: plant height, maturity, heading date, and grain yield. Phenotypic data for these traits were recorded in the years 2011 to 2013 in five management environments (Table 1). All trials were established at CIMMYT's main wheat breeding station at Ciudad Obregon, Sonora, Mexico (27°22′17″ N, 109°55′44″ W, and 35 m above sea level). Wheat lines were evaluated under three irrigation regimes: 2IR, two irrigations giving moderate drought stress; 5IR, five irrigations representing an optimally irrigated crop; and Drip, drip irrigation resulting in high drought stress. Two planting systems were used: Bed, bed planting; Flat, planting on the flat, and two planting dates: No symbol, normal; Heat, late planting resulting in high temperatures at the grain-filling stage.

Plot sizes, seeding, and irrigation were as follows: Bed5IR, two adjacent beds of 0.80 m each by 2.8 m (4.5 m^2 area). Three

rows of seed were planted in each bed. Beds were irrigated before planting. After planting, irrigation was applied at 35, 65, 90, and 110 d. The total amount of water available for the crop was targeted to 530 to 550 mm. Bed2IR, plot size, and planting as for Bed5IR. A single postplanting irrigation was applied 35 to 37 d after planting. The total amount of water available for the crop was targeted to 280 to 300 mm. Flat5IR, Plot size was 1.3 by 4.5 m (5.9 m^2 area) with six rows of seed spaced 16 cm apart. Planting was in dry soil and irrigation was applied by flooding small basins after planting. Irrigations were programmed at 10, 38 to 40, 70, 95, and 115 d after the first irrigation. The amount of water available for the crop was about the same as for Bed5IR. FlatDrip, plot size was 1.3 by 4 m (5.2 m^2) with six rows of seed spaced 16 cm apart. Planting was done in dry soil and three drip tapes were used for irrigation. Available water in the soil was measured before planting. First irrigation was about 90 mm, and the rest of the water was applied in one or two additional irrigations before heading to target 180 mm. The previous four environments were regularly planted during the last week of November. BedHeat environment, plot size was the same as Bed5IR. Preplant irrigation was applied 2 wk before planting and 5 to 6 additional irrigations were planned after planting. The total amount of water available for the crop was about the same as for Bed5IR. This management was planted the last week of February, 3 mo after the Bed5IR.

Genotypic and Pedigree Data

Genotyping-by-sequencing (Elshire et al., 2011) was implemented according to the protocol described in Poland et al. (2012), identifying a total of 45,818 SNP markers. The GBS method generates much missing data. Missing marker scores were replaced with the mean of the nonmissing values for that marker.

All lines evaluated have pedigree data going back as far as 20 generations, and pedigree-based relationships were calculated using the BROWSE software as described in (McLaren et al., 2005). The entries of this matrix equal twice the kinship coefficient (or coefficient of parentage) between pairs of lines. The numbers of lines genotyped and phenotyped were different for each trait. In height and maturity, 1160 were genotyped among 10,475 and 10,469 lines, respectively. Similarly, for heading date and grain yield, 1176 lines were genotyped out of 12,307 and 12,141 lines (Table 1).

Scenarios for Analyses

To avoid confounding method comparisons with GE interaction, we first evaluated each method on data taken from a single year. Thus, we considered three scenarios. Scenario 1: phenotypes were from the first year (2011–2012). Scenario 2: phenotypes were from the second year (2012–2013). Since almost all lines evaluated in the second year had genotypic data, we excluded from Scenario 2 the few lines that did not. Scenario 3: combined information from Scenarios 1 and 2. The genotyped lines had 2 yr of phenotypic information while the rest had only 1 yr.

Line Effect Estimation

Multiple observations were made of each line. To assess the breeding value prediction methods, however, we wanted to deal with just a single value per line. In Scenario 1, for height and maturity, each line was evaluated only once, so raw phenotypes were used as the response variable for height and maturity. Heading date and grain yield were phenotyped in two environments, so environment was used as a fixed effect in Model [1] below. In Scenario 2, each line was observed in five environments. Line effects were estimated using Model [1]:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}(\mathbf{y}^{\star}) + \mathbf{e} \qquad [1]$$

where $\mathbf{y}$ is a vector of observations, $\mathbf{X}$ is an incidence matrix relating observations to environments, $\boldsymbol{\beta}$ is a vector of environment effects, $\mathbf{Z}$ is an incidence matrix relating observations to lines, $\mathbf{y}^{\star}$ is a vector of line effects, and $\mathbf{e}$ is a vector of residuals. Both $\boldsymbol{\beta}$ and $\mathbf{y}^{\star}$ were considered fixed effects. The same model was used in Scenario 3 as Scenario 2, except that $\boldsymbol{\beta}$ then contained effects for both years and environments.

Statistical Models

Data of the four traits were analyzed by using GBLUP, ABLUP, and HBLUP methods, and prediction accuracies were calculated using a fivefold cross-validation scheme. All analyses were performed in R package v. 3.1.2 (R Development Core Team, 2014) using the EMMREML package.

GBLUP

The basic GBLUP method (Van Raden, 2008) used to predict genomic breeding value was:

$$\mathbf{y}^{\star} = \mu + \mathbf{Z}\mathbf{u} + \mathbf{e} \qquad [2]$$

where $\mathbf{y}^{\star}$ is the vector of line effects for the focal trait, μ is the overall mean, $\mathbf{Z}$ is a design matrix relating line effects to lines, and $\mathbf{u}$ is a vector of genomic breeding values. We assumed $\mathbf{u} \sim N(0, \mathbf{G}\,\sigma^2_{\mathbf{u}})$, where $\mathbf{G}$ is the genomic relationship matrix and $\mathbf{e}$ is a vector of residuals, with the assumption $\mathbf{e} \sim N(0, \mathbf{I})$. We used Van Raden's (2008) first method to calculate $\mathbf{G}$: denoting the centered marker score matrix $\mathbf{M}$ (column means have been subtracted from all columns), and p_i is the allele frequency for marker i,

$$\mathbf{G} = \frac{\mathbf{M}\mathbf{M}^{\mathrm{T}}}{2\sum p_i(1-p_i)}.$$

Narrow-sense line-mean heritabilities (h^2) were calculated using these variance components:

$$h^2 = \sigma^2_{\mathbf{u}} / (\sigma^2_{\mathbf{u}} + \sigma^2_{\mathbf{e}}).$$

ABLUP

The model was the same as Eq. [2] for GBLUP, with the only difference that the covariance of $\mathbf{u}$ was assumed proportional to a relationship matrix calculated from pedigree, $\mathbf{A}$ with $\mathbf{u} \sim N(0, \mathbf{A}\,\sigma^2_{\mathbf{u}})$, instead of the $\mathbf{G}$ matrix calculated from markers.

HBLUP

The single-step method (Christensen et al., 2012; Legarra et al., 2009) includes information from genotyped and nongenotyped individuals by blending the genomic relationship matrix **G** and the pedigree-based matrix **A**. The HBLUP method uses the same linear model as the GBLUP methods with the exception that $u \sim N(0,\mathbf{H}\sigma_g^2)$, where **H** is a combined relationship matrix (Christensen et al., 2012; Legarra et al., 2009).

$$\mathbf{H} = \begin{bmatrix} \mathbf{G}_w & \mathbf{G}_w\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{A}_{12}^T\mathbf{A}_{11}^{-1}\mathbf{G}_w & \mathbf{A}_{22} + \mathbf{A}_{12}^T\mathbf{A}_{11}^{-1}(\mathbf{G}_w - \mathbf{A}_{11})\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix} \quad [3]$$

Partitioning of the above combined relationship matrix **H** is based on whether the lines are genotyped or not. The matrices $\mathbf{A}_{11}$, $\mathbf{A}_{22}$, and $\mathbf{A}_{12}$ are submatrices of **A** (the pedigree-based relationship matrix; the superscripted T represents transposed, superscripted –1 indicates invserse) containing relationships among genotyped, nongenotyped, and between genotyped and nongenotyped lines, respectively. For the subset of genotyped lines,

$$\mathbf{G}_w = (1 - w)\mathbf{G} + w\mathbf{A}_{11} \quad [4]$$

where w is a weight that corresponds to the fraction of the genetic variance not captured by markers. Further details about these adjustments were studied in Gao et al. (2012). We used $w = 0.05$ to place the majority of weight on the genomic relationship matrix as this matrix should be able to track Mendelian segregation and we therefore assumed that it would be a more accurate estimate of the true relationship. A similar w was also tested in Christensen et al. (2012). **G** is the genomic relationship matrix described above. The matrix $\mathbf{G}_w$ is a combination of **G** and $\mathbf{A}_{11}$ and these two matrices need to be scaled appropriately: the **G** matrix was scaled such that the average of its diagonal elements was equal the average of diagonal elements of $\mathbf{A}_{11}$ (Forni et al., 2011). The system of equations for these adjustments was discussed in detail by Christensen et al. (2012).

HBLUP-Weight

In the HBLUP model, the weight for the polygenic effect was fixed at 0.05. For our Scenario 2, however, all lines evaluated had both **G** and **A** matrices, such that $\mathbf{H} = \mathbf{G}_w$ because the partitioning of Eq. [3] above means that $\mathbf{A}_{22}$ has dimension 0. We could therefore optimize the relative weighting of these two known covariances (G and A matrices) by fitting them simultaneously as random effects as shown in Eq. [5]. The two random effect models were fitted using the EMMREML package of R (R development Core Team, 2014) with the function emmremlMultiKernel. This function can optimize w in a way to maximize the likelihood. This function was used only in Scenario 2 to see whether optimization of the relative weights would increase accuracy. Data were analyzed by using the linear model:

$$\mathbf{y}^\star = \mu + \mathbf{Z}(\mathbf{u}_\mathbf{A} + \mathbf{u}_\mathbf{G}) + \mathbf{e} \quad [5]$$

where $\mathbf{u}_\mathbf{A} \sim N(0, \mathbf{A}\sigma_\mathbf{A}^2)$, $\mathbf{u}_\mathbf{G} \sim N(0, \mathbf{G}\sigma_\mathbf{G}^2)$, and $\mathbf{e} \sim N(0,\mathbf{I}\sigma_\mathbf{e}^2)$ are independently distributed. We can reparameterize this model writing $\mathbf{u}_\mathbf{A} \sim N(0, \mathbf{A}w\sigma^2)$, $\mathbf{u}_\mathbf{G} \sim N[0, \mathbf{G}(1 - w)\sigma^2]$, and $\mathbf{e} \sim N(0,\mathbf{I}\sigma_\mathbf{e}^2)$, where $\sigma^2 = \sigma_\mathbf{A}^2 + \sigma_\mathbf{G}^2$ and $w = \sigma_\mathbf{A}^2/(\sigma_\mathbf{A}^2 + \sigma_\mathbf{G}^2)$. Estimates of the parameters are obtained by iteratively maximizing the restricted maximum likelihood (REML) on w, σ^2, and $\sigma_\mathbf{e}^2$.

Multivariate Gaussian Mixed Model

In Scenario 1, lines were phenotyped in only one environment each, but in Scenarios 2 and 3, the same lines were tested in multiple environments. To account for these environments in Scenario 2, we used a multivariate mixed model:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{ZU} + \mathbf{E}, \quad [6]$$

with **Y** being the matrix of responses with dimension of $n \times q$, where n is number of lines and q is number of environments; **X** is an $n \times q$ design matrix showing the incidence of observations in environments; **B** is a $q \times q$ matrix of corresponding fixed environment effects; **Z** is an $n \times n$ design matrix for random effects; **U** is an $n \times q$ matrix of additive genetic effects with $\mathbf{U} \sim MN(0, \mathbf{G}, \mathbf{V}_u)$; **G** is described above; and $\mathbf{V}_u$ is a $q \times q$ additive genetic covariance matrix among observations in the q environments, and **E** is $n \times q$ residual effect matrix with $\mathbf{E} \sim MN(0, \mathbf{I}, \mathbf{V}_e)$, **I** is an identity matrix and $\mathbf{V}_e$ is a $q \times q$ error covariance matrix among observations. Further detail about these models can be found in the EMMREML package of R (R development Core Team, 2014). The final prediction from this model was the row mean of **U**.

Cross-Validation Scheme

Prediction accuracies were calculated by using fivefold cross-validation in the three scenarios. Because the genotyped lines were common to all scenarios, validation folds stayed constant across scenarios. The genotyped lines were split randomly into five folds. Each fold in turn was excluded from the training population, predictions were made for that fold from the remaining individuals, and predictions were correlated with the observed $\mathbf{y}^\star$.

RESULTS

Narrow-sense heritability of plant height, maturity, heading date, and grain yield were 0.46, 0.70 0.87, and 0.10, respectively. Prediction accuracies of different traits in three scenarios are reported in Table 2.

Height

In Scenario 1, the prediction accuracy by using GBLUP was 0.44 and increased by 29% when we included the information from nongenotyped lines in HBLUP. In Scenario 2, with five environments, prediction accuracies with GBLUP, ABLUP, HBLUP, and HBLUP-weight, were 0.39, 0.38, 0.40, and 0.42, respectively. In Scenario 3, by including all the information from Scenario 1 and 2, prediction accuracy with ABLUP was 0.49 and increased to 0.57 by fitting single-step HBLUP model.

Table 2. Prediction accuracies obtained via fivefold cross validation in three scenarios† for four traits with four models. Heritabilities are given in parentheses for each trait.

Genomic model	Height (0.46)			Maturity (0.70)			Heading date (0.87)			Grain yield (0.10)		
	SN 1	SN 2	SN 3	SN 1	SN 2	SN 3	SN 1	SN 2	SN 3	SN 1	SN 2	SN 3
GBLUP	0.44	0.39	–	0.39	0.40	–	0.43	0.38	–	0.46	0.23	–
ABLUP	0.43	0.38	0.49	0.38	0.30	0.56	0.37	0.33	0.40	0.54	0.20	0.56
HBLUP	0.57	0.40	0.57	0.59	0.40	0.66	0.54	0.39	0.52	0.68	0.23	0.59
HBLUP-weight												
Accuracy	–	0.42	–	–	0.40	–	–	0.40	–	–	0.25	–
Weight‡	–	0.64	–	–	0.43	–	–	0.55	–	–	0.46	–

† Scenario 1 (SN 1): phenotypes were from the first year (2011–2012). Scenario 2 (SN 2): phenotypes were from the second year (2012–2013). Since almost all lines evaluated in the second year had genotypic data, we excluded from Scenario 2 the few lines that did not. Scenario 3 (SN 3): combined information from Scenarios 1 and 2. The genotyped lines had 2 yr phenotypic information and the rest had only 1-yr phenotypic information.

‡ Value of the optimized weight obtained from the HBLUP-weight analysis.

Maturity

In this trait, HBLUP had the highest accuracy in Scenario 1 compared with ABLUP and GBLUP. In Scenario 2, prediction accuracies were 0.40, 0.30, 0.40, and 0.40 with GBLUP, ABLUP, HBLUP, and HBLUP-weight, respectively. Again in Scenario 3, HBLUP outperformed ABLUP.

Heading Date

Prediction accuracy by using GBLUP was 0.43 and decreased to 0.37 with ABLUP model, and increased to a level of 0.54 with HBLUP approach in Scenario 1. Like maturity, prediction accuracy was lowest with ABLUP and highest by using single-step HBLUP approach. HBLUP again provided better accuracy than ABLUP in Scenario 3.

Grain Yield

Prediction accuracy in grain yield was higher than other traits in Scenario 1. In Scenario 1, GBLUP, HBLUP and ABLUP had the accuracy of 0.46, 0.68, and 0.54, respectively. In Scenario 2, prediction accuracies were lower than other traits. The results in Scenario 2 were similar to maturity and heading date. Prediction accuracy using HBLUP was highest and followed by GBLUP and ABLUP. In Scenario 3, prediction accuracy with ABLUP was 0.56 and increased to 0.59 by fitting single-step HBLUP model.

Prediction accuracies by using single-step HBLUP outperformed GBLUP in all cases. HBLUP also outperformed ABLUP in all cases except in grain yield, where ABLUP did better in Scenario 3. Scenario 1 had higher accuracies than Scenario 2 in all cases. The polygenic effect weight, w, optimized by emmremlMultiKernal in the HBLUP-weight model, outperformed the fixed weight used in HBLUP for all traits. Prediction accuracies with HBLUP-weight were higher than GBLUP, ABLUP, and HBLUP for all four traits in Scenario 2. The optimized weights from HBLUP-weight were higher than our a priori set weight of 0.05 (0.64, 0.43, 0.55, and 0.46, for height, maturity, heading date, and grain yield, respectively). We found, however, that the loglikelihood declined only modestly when suboptimal weights were imposed: 95% confidence intervals based on a loglikelihood decrease of two spanned a range width of about 0.4 for grain yield and 0.3 for the other traits (Supplemental Fig. S1). This result indicates that optimal weights are not accurately estimated even when >1000 individuals are analyzed. Scenario 3 combined the information of Scenarios 1 and 2. There, the HBLUP model outperformed the ABLUP model for Height, Maturity, and Heading date. For Grain yield, HBLUP also produced higher prediction accuracy than ABLUP; however, this improvement in accuracy, compared with ABLUP, is not as high as with the rest of the traits. The HBLUP algorithm, however, used a weight $w = 0.05$ that, given observations from Scenario 2, was probably suboptimal.

Table 3. Prediction accuracies in Scenario 2 using the multivariate (multienvironment) mixed model.

Model	Height	Maturity	Heading date	Grain yield
GBLUP	0.48	0.46	0.41	0.46
ABLUP	0.51	0.41	0.37	0.48
HBLUP	0.49	0.46	0.41	0.47

Multivariate Mixed Model Results

To account for multiple environments in Scenario 2, predictions were also obtained using a multivariate mixed model, considering different environments as traits (Table 3). Prediction accuracies improved notably relative to univariate predictions (Table 2). For instance, in height, prediction accuracy was 0.39 with univariate GBLUP and increased to 0.48 with the multivariate model. Similar increases were seen for ABLUP and HBLUP by using the multivariate-environment approach.

Genotypic Correlations among Environments

Correlations were calculated from the genetic covariances obtained from the emmremlMultivariate function of EMMREML package in R (R development Core Team, 2014). Pairwise genotypic correlations between different environments are given for all traits (Tables 4 and 5).

Table 4. Genotypic correlation for height (above diagonal) and maturity (below diagonal) between different Environments (E).

Environment	E2	E3	E4	E5	E6	E7
E2	–	0.22	0.40	0.08	0.44	0.17
E3	0.62	–	0.57	0.36	0.69	0.49
E4	0.25	0.15	–	0.55	0.93	0.53
E5	0.14	0.13	0.80	–	0.54	0.14
E6	0.25	0.03	0.87	0.71	–	0.44
E7	†	†	†	†	†	–

† Maturity was not recorded in E7.

Table 5. Genotypic correlation for heading date (above diagonal) and grain yield (below diagonal) between different Environments (E).

Environment	E2	E3	E4	E5	E6	E7
E2	–	0.63	0.48	0.56	0.47	0.57
E3	0.17	–	0.96	0.84	0.92	0.77
E4	0.33	0.36	–	0.80	0.93	0.71
E5	0.09	0.09	0.12	–	0.70	0.75
E6	0.15	−0.30	0.03	−0.04	–	0.64
E7	0.03	0.12	−0.19	0.23	−0.08	–

Height

Results indicate that all of the environments were positively correlated to each other. Some of the correlations were above 0.50. The highest correlation 0.93 was observed between the environments E4 and E6. The lowest correlation 0.08 was found between E2 and E5.

Maturity

For this trait, most of the environments were strongly correlated to each other, except the environment E2 from the 2011–2012 yr. The year 2011 was unusual in Obregon because the temperature was below freezing for a couple of days. This may explain why E2 has a lower correlation with the environment E5. The highest correlation of 0.85 was found between E4 and E6.

Heading Date

Heading date correlations followed a similar pattern to those for maturity, with E2 having lower correlations than other environments. No pair of environments had a negative correlation. Some of the correlations between environments reached 0.96 (between E3 and E4).

Grain Yield

In this trait, most of the correlations were lower than 0.50. The environments E3 and E4 were more highly correlated than the rest of environment correlations. Some of the environment pairs had a negative correlation.

DISCUSSION

The present study assessed the prediction accuracy of wheat lines in three Scenarios by applying GBLUP, ABLUP, and HBLUP (Christensen et al., 2012; Gao et al., 2012; Legarra and Misztal, 2008) methods. Prediction accuracies were reported as the correlation between GEBV and observed line effects. The procedure was demonstrated on four example traits (i.e., plant height, maturity, heading date, and grain yield). First, HBLUP accuracy was greater than GBLUP accuracy in all cases. Similar results were reported by Gao et al. (2012) in Nordic Holstein population. The likely reason for this superiority is that HBLUP was able to use much more phenotypic information than GBLUP. From Table 1, we can see that on average over all traits, there were close to 10,000 observations on nongenotyped lines. These observations were not available to GBLUP, for which there were generally <7000 observations across six environments. In principle, GP methods should do better than pedigree-based prediction methods because they can account for and predict the effects of Mendelian segregation (Daetwyler et al., 2007; Van Raden et al., 2008), and this is what we have seen for all traits.

Breeders often observe greater GE variance for yield than for height and maturity traits. This GE would reduce the heritability of yield in Scenario 2, where lines were evaluated in five management environments. In the analyses used here, high GE translates to low genetic correlation between environments (Cooper and DeLacy, 1994). This measure showed greater GE for yield than the other traits: average between-environment correlations were 0.44, 0.40, 0.72, and 0.07 for height, maturity, heading date, and grain yield, respectively (Tables 4 and 5). Another nonmutually exclusive possibility is that lines were selected more strongly for yield between the 2011–2012 season and the 2012–2013 season than for other traits. Such selection would deplete variance in yield more than in other traits and therefore decrease its heritability and make it more difficult to predict (Falconer and Mackay, 1996). Indeed, the selection criteria used to advance lines from the first year to multienvironment trials in the second year weight yield more heavily than other traits.

Optimizing the weight (w) for combining the pedigree and the genomic relationship matrices in Scenario 2 (termed HBLUP-weight here) gave higher accuracies than fixed weight HBLUP in all traits. The optimal weighting factors differed between the traits. The weighting factors optimized by emmremlMultiKernel were 0.64, 0.43, 0.55, and 0.46 for height, maturity, heading date, and grain yield, respectively. In a dairy cattle study, Liu et al. (2011) found that optimal residual polygenic variance in a GBLUP model with a polygenic effect appeared to be different among traits. In our study, we also found differences in optimal weights, but the likelihood curves for the weight parameters were fairly flat (Supplemental Fig. S1),

such that the confidence intervals for the weights overlapped substantially between traits. Nevertheless, these results suggest a possible benefit to trait-specific weighting in single-step HBLUP. One interpretation of the differing weights between traits is that this parameter is affected by the fraction of the variance caused by low minor allele frequency (MAF) loci whose effects are poorly captured by markers. For traits with a high fraction of the variance generated by such low MAF causal loci, the weight would be high on the pedigree component. Other research in animal systems, however, tempers the idea that trait-specific weights are needed for prediction (Gao et al., 2012) because differential weighting had a relatively small effect on accuracy. We also sought to identify optimal weights in Scenario 3 using likelihood curves (Supplemental Fig. S2). In that scenario, however, for three traits (Maturity, Height, and Grain Yield) the likelihood was maximized when all the weight went to the pedigree-based relationship matrix, **A** (Supplemental Fig. S2). This suggests that when data have a substantial majority of individuals without genotypes, optimization will not identify an intermediate weight.

Dealing with GE by considering measurements in each environment as separate traits and estimating the covariance among such traits in a multivariate approach gave us improved prediction accuracy relative to a standard univariate approach (Table 3). Genotypic correlations among different environments were determined by using a multivariate mixed model (Eqn. [6]). Results revealed that most of the experimental environments were strongly correlated to each other (Tables 4 and 5). We found higher correlations among environments in the same year than between environments across years. This is consistent with the idea that genotype by year interactions are often greater than GE interactions. In addition, in the year 2011–2012, very low temperatures were recorded over winter at the CIMMYT Obregon experimental station, making it an extreme environment.

The single-step HBLUP method has been applied in various livestock research studies (Aguilar et al., 2010; Chen et al., 2011; Liu et al., 2014). In a recent paper, Legarra et al. (2014) showed that the single-step method produced higher predictive ability than multistep and pedigree BLUP methods in different traits of animal species. Broadly speaking, our results were in agreement with these previous studies, particularly in situations where only a small fraction of the total population was genotyped.

CONCLUSIONS

The results from this study indicate that single-step HBLUP can provide higher prediction accuracies than GBLUP and pedigree-based ABLUP. The accuracy can be further increased when we have multienvironment data by applying a multivariate approach. The optimized weighting factors used in HBLUP-weight increased prediction accuracy in all traits. These results provide evidence that the single-step approach should be favored to combine information from genotyped and nongenotyped wheat lines using pedigree and marker data. Therefore, in practical genetic evaluation, HBLUP could be an alternative tool to improve prediction accuracy over GBLUP and ABLUP.

Supplemental Information Available

Supplemental information is included with this article.

Supplemental Fig. S1. Analysis of Scenario 2. Difference from the maximum loglikelihood as a function of the weight given to the pedigree-based relationship matrix. The horizontal line is at a loglikelihood of –2: all weight values above the line are in the 95% confidence interval for the weight parameter.

Supplemental Fig. S2. Analysis of Scenario 3. Difference from the maximum loglikelihood as a function of the weight given to the pedigree-based relationship matrix. The horizontal line is at a loglikelihood of –2: all weight values above the line are in the 95% confidence interval for the weight parameter.

Acknowledgments

This research received funds from a grant from the Bill and Melinda Gates Foundation. Deniz Akdemir was supported by USDA-NIFA-AFRI Award 2011-68002-30029, The Triticeae Coordinated Agricultural Project. We thank reviewers for constructive comments that required analyses that improved this study.

References

Aguilar, I., I. Misztal, D. Johnson, A. Legarra, S. Tsuruta, and T. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci. 93:743–752. doi:10.3168/jds.2009-2730

Burgueño, J., G. de los Campos, K. Weigel, and J. Crossa. 2012. Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. Crop Sci. 52:707–719. doi:10.2135/cropsci2011.06.0299

Ceccarelli, S., W. Erskine, J. Hamblin, and S. Grando. 1994. Genotype by environment interaction and international breeding programmes. Exp. Agric. 30:177–187. doi:10.1017/S0014479700024121

Chen, C.-Y., I. Misztal, I. Aguilar, A. Legarra, and W. Muir. 2011. Effect of different genomic relationship matrices on accuracy and scale. J. Anim. Sci. 89:2673–2679. doi:10.2527/jas.2010-3555

Christensen, O.F., P. Madsen, B. Nielsen, T. Ostersen, and G. Su. 2012. Single-step methods for genomic evaluation in pigs. Animal 6:1565–1571. doi:10.1017/S1751731112000742

Cooper, M., and I. DeLacy. 1994. Relationships among analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. Theor. Appl. Genet. 88:561–572. doi:10.1007/BF01240919

Crossa, J., G. de los Campos, P. Pérez, D. Gianola, J. Burgueño, J.L. Araus, D. Makumbi, R.P. Singh, S. Dreisigacker, J. Yan, V. Arief, M. Banziger, and H.-J. Braun. 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

Daetwyler, H.D., B. Villanueva, P. Bijma, and J.A. Woolliams. 2007. Inbreeding in genome-wide selection. J. Anim. Breed. Genet. 124(6):369–376.

Ducrocq, V., and Z. Liu. 2009. Combining genomic and classical information in national BLUP evaluations. Proc. Interbull Meeting, Barcelona, Spain, 21–24 Aug. 2009. Interbull Bull. 40:172–177.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6:e19379 doi:10.1371/journal.pone.0019379

El-Soda, M., M. Malosetti, B.J. Zwaan, M. Koornneef, and M.G. Aarts. 2014. Genotype × environment interaction QTL mapping in plants: Lessons from *Arabidopsis*. Trends Plant Sci. doi:10.1016/j.tplants.2014.01.001

Falconer, D.S. 1952. The problem of environment and selection. Am. Nat. 86:293–298. doi:10.1086/281736

Falconer, D., and T. Mackay. 1996. Introduction to quantitative genetics. 4th. Pearson Education, Harlow, UK.

Forni, S., I. Aguilar, and I. Misztal. 2011. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. Genet. Sel. Evol. 43:1. doi:10.1186/1297-9686-43-1

Gao, H., O.F. Christensen, P. Madsen, U.S. Nielsen, Y. Zhang, M.S. Lund, and G. Su. 2012. Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. Genet. Sel. Evol. 44:8. doi:10.1186/1297-9686-44-8

Harris, B., and D. Johnson. 2010. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. J. Dairy Sci. 93:1243–1252. doi:10.3168/jds.2009-2619

Heffner, E.L., M.E. Sorrells, and J.-L. Jannink. 2009. Genomic selection for crop improvement. Crop Sci. 49:1–12. doi:10.2135/cropsci2008.08.0512

Jannink, J.-L., A.J. Lorenz, and H. Iwata. 2010. Genomic selection in plant breeding: From theory to practice. Briefings Funct. Genomics 9:166–177. doi:10.1093/bfgp/elq001

Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. J. Dairy Sci. 92:4656–4663. doi:10.3168/jds.2009-2061

Legarra, A., O.F. Christensen, I. Aguilar, and I. Misztal. 2014. Single Step, a general approach for genomic selection. Livest. Sci. 166:54–65. doi:10.1016/j.livsci.2014.04.029

Legarra, A., and I. Misztal. 2008. Technical note: Computing strategies in genome-wide selection. J. Dairy Sci. 91:360–366. doi:10.3168/jds.2007-0403

Liu, Z., M. Goddard, F. Reinhardt, and R. Reents. 2014. A single-step genomic model with direct estimation of marker effects. J. Dairy Sci. 97:5833–5850. doi:10.3168/jds.2014-7924

Liu, Z., F.R. Seefried, F. Reinhardt, S. Rensing, G. Thaller, and R. Reents. 2011. Impacts of both reference population size and inclusion of a residual polygenic effect on the accuracy of genomic prediction. Genet. Sel. Evol. 43:9. doi:10.1186/1297-9686-43-19

McLaren, C.G., R. Bruskiewich, A.M. Portugal, and A.B. Cosico. 2005. The International Rice Information System. A platform for meta-analysis of rice crop data. Plant Physiol. 139:637–642. doi:10.1104/pp.105.063438

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

Poland, J.A., P.J. Brown, M.E. Sorrells, and J.-L. Jannink. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7(2):e32253. doi:10.1371/journal.pone.0032253

R Development Core Team. 2014. EMMREML: Fitting mixed models with known covariance structures. v. 3.1.2. Available at: https://cran.r-project.org/web/packages/EMMREML/index.html (verified 30 Nov. 2015). R Development Core Team, Vienna, Austria.

Su, G., P. Madsen, U.S. Nielsen, E.A. Mäntysaari, G.P. Aamand, O.F. Christensen, and M.S. Lund. 2012. Genomic prediction for Nordic Red Cattle using one-step and selection index blending. J. Dairy Sci. 95:909–917. doi:10.3168/jds.2011-4804

Van Raden, P.M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423. doi:10.3168/jds.2007-0980