Published January 29, 2016

RESEARCH

Modeling Genotype × Environment Interaction for Genomic Selection with Unbalanced Data from a Wheat Breeding Program

Bettina Lado, Pablo González Barrios, Martín Quincke, Paula Silva, and Lucía Gutiérrez*

ABSTRACT

Genomic selection (GS) has successfully been used in plant breeding to improve selection efficiency and reduce breeding time and cost. However, there is not a clear strategy on how to incorporate genotype × environment interaction (GEI) to GS models. Increased prediction accuracy could be achieved using mixed models to exploit GEI by borrowing information from other environments. The objective of this work was to compare strategies to exploit GEI in GS using mixed models. Specifically, we compared strategies to predict new genotypes by borrowing information from other environments modeling the correlation matrix across environments and to design sets of environments aiming for low GEI to predict genomic performance in new environments. We evaluated 1477 advanced wheat (*Triticum aestivum* L.) lines for yield in 35 location–year combinations genotyped with genotyping-by-sequencing (GBS). Mixed models were used to obtain either overall or by-environment predictions for different sets of environments. Overall accuracy was high (0.5). Borrowing information from relatives evaluated in multiple environments and modeling the correlation matrix across environments was the best strategy to predict new genotypes. On the other hand, the best strategy for predicting the performance of genotypes in new environments was either to predict across locations for single years or to predict within defined mega-environments (MEs) for any year or location. In summary, higher predictive ability was obtained by characterizing and by modeling GEI in the GS context.

B. Lado, P. González Barrios, and L. Gutiérrez, Dep. of Statistics, College of Agriculture, Universidad de la República, Garzón 780, Montevideo, Uruguay; L. Gutierrez, Dep. of Agronomy, Univ. of Wisconsin–Madison, 1575 Linden Dr., Madison, WI 53706; M. Quincke and P. Silva, Programa Nacional de Investigación Cultivos de Secano, Instituto Nacional de Investigación Agropecuaria, Est. Exp. La Estanzuela, Colonia 70000, Uruguay. Received 2 Apr. 2015. Accepted 14 Sept. 2015. *Corresponding author (gutierrezcha@wisc.edu).

Abbreviations: ALL, all location–year combinations; AYT, advanced yield trial; BLUE, best linear unbiased estimation; CV, cross-validation; DOL, location Dolores; DZ, location Durazno; $GBLUP_{g\times e}$, by-environment genomic predictions; $GBLUP_{M}$, overall genomic predictions for the mean of the environments; GBS, genotyping-by-sequencing; GEI, genotype × environment interaction; GGE, genotypic main effect and genotype × environment interaction matrix; GS, genomic selection; INIA, National Agricultural Research Institute; LE, location La Estanzuela; ME, mega-environment; MET, multienvironment trial; PYT, preliminary yield trial; R2, location Ruta2; WBP, Wheat Breeding Program; YET, elite yield trial; YOU, location Young.

Genomic selection has proven highly successful, and plant breeding programs are implementing GS strategies as part of their breeding routines (Heffner et al., 2010; Lorenz et al., 2012; Hayes et al., 2013; Rutkoski et al., 2015). Genomic selection was first developed in animal breeding (Meuwissen et al., 2001) and has since been transferred into plant breeding (Heffner et al., 2009). While most of the GS studies have focused on evaluating GS model performance (Bennewitz et al., 2009; Solberg et al., 2009; Heslot et al., 2012; de los Campos et al., 2013) and optimizing the training population in terms of population size (Lorenzana and Bernardo, 2009; Asoro et al., 2011; Lorenz et al., 2012), number of markers (Lorenzana and Bernardo, 2009; Asoro et al., 2011; Heffner et al., 2011), and its structure (Asoro

Published in Crop Sci. 56:1–15 (2016).
doi: 10.2135/cropsci2015.04.0207

et al., 2011; Crossa et al., 2013; Isidro et al., 2014); there are few reports demonstrating results from implementing GS in the full context of a plant breeding program. One of the key differences between animal and plant breeding is that in the latter, GEI is large (Cooper and Delacy, 1994; Mathews et al., 2008). Multienvironment trials (MET) are routinely conducted in plant breeding to capture and account for GEI. However, these trials are highly unbalanced because of poor lines being removed from evaluation while promising lines are further evaluated (Bernardo, 2010). To implement GS strategies in plant breeding, GEI needs to be accounted for, modeled, and predicted (Resende et al., 2011; Burgueño et al., 2012; Heslot et al., 2014; Lopez-Cruz et al., 2015).

Several strategies have been proposed to deal with GEI: ignore, reduce, or exploit it (Eisemann et al., 1990; DeLacy et al., 1996). Ignoring or reducing GEI is possible when the proportion of the total phenotypic variance as a result of GEI is low (DeLacy et al., 1996) or when selection for broad adaptation is conducted (Lal, 2013; Lal, 2014). However, the most powerful strategy is to exploit GEI to either develop locally adapted material (Mathews et al., 2008) or to use GEI to better characterize the genotypes (Dreccer et al., 2007; Mathews et al., 2008; Cooper et al., 2014). Genotype × environment interaction was first characterized with linear–bilinear models using the singular value decomposition of either the GEI matrix (i.e., additive main effect and multiplicative interaction, AMMI; Gollob, 1968; Mandel, 1969; Gauch and Zobel, 1988; van Eeuwijk, 1995; Gauch, 2006; Gauch et al., 2008; Crossa, 2012); the genotypic main effect and GEI matrix (GGE; Yan et al., 2000, 2007; Yan and Tinker, 2006); or site regression (Crossa and Cornelius, 1997). Linear–bilinear models could be graphically represented through biplots (Gabriel, 1971; Yan and Tinker, 2006) or enhanced biplots (Gauch, 2006), where genotypes and environments are plotted in a single graph. Later, mixed models were introduced to allow more flexibility and to model different correlation structures among environments (Piepho, 1998; Burgueño et al., 2008; Kelly et al., 2009; Cullis et al., 2010), increasing statistical power (Wang and Schmidt, 2002). A number of correlation structures can be modeled with mixed models, from a single correlation among environments (i.e., compound symmetry models; Piepho, 2005), heterogeneous variances (Piepho, 1998; Malosetti et al., 2004; Burgueño et al., 2008; Cullis et al., 2010), up to a unique correlation for each pair of environments (i.e., unstructured variance–covariance; Burgueño et al., 2008; Malosetti et al., 2013). Finally, MEs were defined as groups of environments with minimal GEI within (Braun et al., 1996). Both linear–bilinear and mixed models are powerful tools to characterize GEI and have successfully been used in plant breeding.

Breeding programs produce large MET datasets that are highly unbalanced with heterogeneous quality because of the degree of replication in different trials (Bernardo, 2010). Early generations usually consist of a large number of genotypes evaluated in a few environments with few replications. Advanced trials, on the other hand, consist of few genotypes evaluated in a large number of environments with more replications (Bernardo, 2010; Lorenz et al., 2012). This creates a high level of unbalance in the data sets that can be properly managed with mixed models using restricted maximum-likelihood estimates (Patterson and Thompson, 1971). Mixed models can also model individual-plot variability to account for variability in data quality from MET (Malosetti et al., 2004; Mathews et al., 2008). In summary, mixed models are flexible enough to model MET data with high levels of unbalance, with correlation between environments and genotypes, and with variability in data quality from multiple trials.

Genotype × environment interaction has been exploited in genomic studies to map quantitative trait loci in balanced populations (Piepho, 2000; Malosetti et al., 2004; van Eeuwijk et al., 2005; Boer et al., 2007; Mathews et al., 2008) and genome-wide association studies (Gutiérrez et al., 2015) using mixed models. Additionally, GS studies have included GEI information by performing overall predictions across environments (Heffner et al., 2011; Resende et al., 2011), within environments or groups of environments (Burgueño et al., 2012; Ly et al., 2013; Dawson et al., 2013; Heslot et al., 2014), or using marker-by-environment predictions (Jarquín et al., 2014; Lopez-Cruz et al., 2015). While mixed models have successfully been used to model GEI in GS, the best strategy to predict new genotypes or new environments is not clear. The aim of this work was to improve the predictive ability of the models comparing strategies to exploit GEI in GS to predict new genotypes and new environments by evaluating different groups of environments and modeling the variance–covariance matrix across environments in a large, highly unbalanced, historical dataset from a wheat breeding program. We specifically aimed to answer the following questions: Could we improve the predictive ability by modeling the variance–covariance matrix across environments and predicting the performance of new genotypes (i.e., unphenotyped genotypes)? Could we design groups of environments where overall predictions outperformed environment-specific predictions? Could we borrow information from some environments to predict the performance in new environments (i.e., untested environments)?

MATERIALS AND METHODS

Plant Material

A total of 1477 advanced inbred lines from the National Agricultural Research Institute (INIA) Wheat Breeding Program (WBP) were extensively phenotyped and also genotyped to be

Table 1. Description of the wheat breeding program trials indicating the number of genotypes in each trial or group of trials, mean yield (kg ha^{-1}) per trial, and standard deviation for 1477 genotypes evaluated in 35 environments in preliminary (PYT), advanced (AYT), and elite (EYT) yield trials grouped by cycle (S, short; I, intermediate; L, long).

Year	Location†	EYT_{SI}	EYT_L	AYT_S	AYT_I	AYT_L	PYT_S	PYT_I	PYT_L	No. genotypes	Yield	SD
2010	DOL	25			64	36				111	5358	455
	LE1	25	16		64	36	300	197	326	617	4835	1284
	LE2	25	16							41	5566	1265
	LE3	25	16							41	4875	1151
	LE4	25	16							41	4972	1455
	R2	25								25	5336	526
	YOU	25	16		64	36				117	5366	681
2011	DOL	30	16	81	49	42				159	4028	857
	DZ	30	16							44	5758	688
	LE1	30	16	81	49	42	239	128	264	957	4652	1416
	LE2	30	16							44	5560	631
	LE3	30	16							44	5439	719
	LE4	30	16							44	3182	1092
	R2	30	16							44	6476	504
	YOU	30	16	81	49	42				110	4546	649
2012	DOL	30	20	49	49	42				138	2624	604
	LE1	30	20	49	49	42				138	2482	1138
	LE2	30	20							48	3342	1029
	LE3	30	20							48	3437	1050
	LE4	30	20							48	4042	1048
	R2	30	20							48	3323	1050
	YOU	30	20	49	49	42				138	1920	734
2013	DOL	30	20							50	5782	470
	LE1	30	20				233	199	149	631	6737	1191
	LE2	30	20							50	7819	672
	LE3	30	20							50	5667	867
	LE4	30	20							50	5055	862
	R2	30	20							50	7714	655
	YOU	30	20							50	4444	694
2014	DOL	30								30	3538	728
	LE1	30	16	64	49	36	132	69	132	528	4341	1125
	LE2	30	16							46	4581	1034
	LE3	30	16							45	4166	1063
	R2	30	16	64	49	36				195	4305	614
	YOU	30	16	64	49	36				195	3910	814

† DOL, Dolores; DZ, Durazno; LE, La Estanzuela; R2, Ruta2; YOU, Young.

used as a training set for genomic prediction models. These genotypes were chosen to represent the advanced inbred lines at a given time in the breeding program and consisted of all the lines from the preliminary yield trials (F_7; PYT) from 2010, 2011, and 2013 as well as the advanced (F_8; AYT) and elite (F_9; EYT) yield trials from 2010.

Phenotypic Data

Historic data from the INIA-WBP was used to evaluate yield of genotypes. Phenotypic evaluation from five locations in Uruguay evaluated in 5 yr, including one location with four sowing dates, were used to define a total of 35 environments (Table 1). Dolores (DOL; 33°50′ S, 58°14′ W, 15 m asl), Durazno (DZ; 33°33′ S, 56°31′ W, 91 m asl), La Estanzuela (LE; 34° 20′ S, 57° 42′ W, 81 m asl), Young (YOU; 32°76′ S, 57°57′ W, 85 m asl), and Ruta2 (R2; 33°45′ S, 57°90′ W, 95 m asl) were the five locations used to evaluate the genotypes. Four sowing dates were evaluated in LE (LE1, LE2, LE3, and LE4). The evaluation years were 2010 to 2014. However, the data is typical for a breeding program, as not all genotypes were evaluated in all the environments, and therefore, the data is highly unbalanced in years and locations as a result of promising genotypes being evaluated in multiple environments and poor genotypes being evaluated in a single environment (Table 1; Fig. 1).

Within each year, genotypes were evaluated in PYT, AYT, or EYT. Preliminary yield trials consisted of approximately 600 genotypes each year evaluated in a sequence of trials. Three to twelve PYTs were conducted every year with 50 to 90 genotypes in each trial. Genotypes were grouped in trials by heading date (i.e., short, PYT_S; intermediate, PYT_I; or long, PYT_L cycle). Each trial consisted of a resolvable incomplete block design with two replications where 6 to 10 incomplete blocks nested within each complete replications were used. Only location LE1 was used for PYT (Table 1).

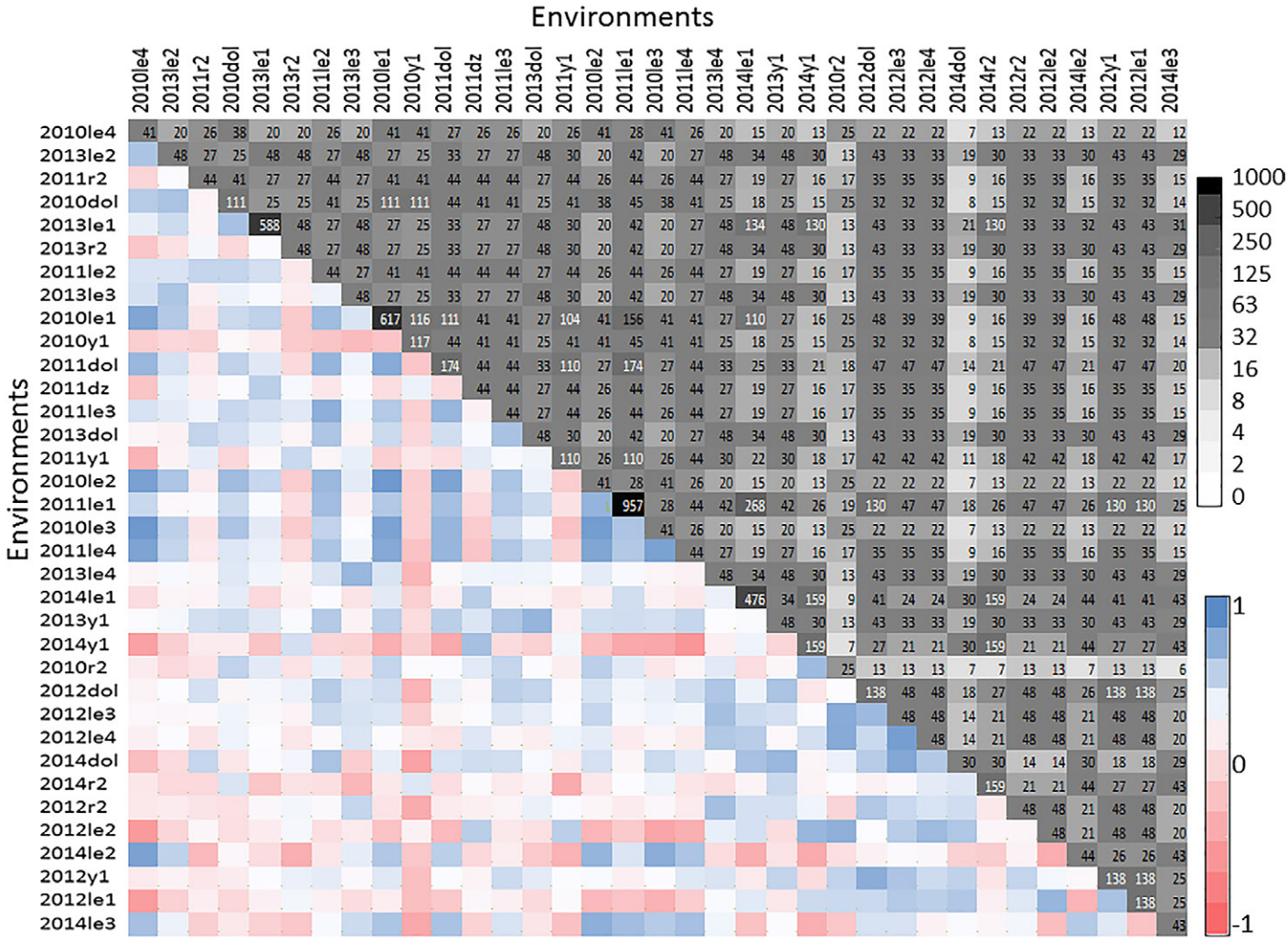


Figure 1. Characterization of genetic correlation across environments and degree of unbalance of 1477 wheat genotypes evaluated in 35 environments. Lower diagonal heat plot indicates genetic correlations among pairs of environments, while upper diagonal indicates the number of shared genotypes evaluated in each pair of environments. Environments are labeled by their year (i.e., 2010–2014) followed by their location (i.e., Dolores, Duranzo, La Estanzuela [LE1–LE4], Ruta2, and Young).

Advanced yield trials consisted of approximately 150 genotypes each year evaluated in a series of trials. Three AYTs were conducted every year with 35 to 80 genotypes in each trial with genotypes grouped in trials by heading date (i.e., short, AYT_S; intermediate, AYT_I; or long, AYT_L cycle). Each trial consisted of a resolvable incomplete block design with three replications where six to nine incomplete blocks nested within each complete replication were used.

Elite yield trials consisted of approximately 50 genotypes each year evaluated in a series of trials. Three EYTs were conducted every year with 16 to 30 genotypes in each trial with genotypes grouped by heading date (i.e., short and intermediate, EYT_{SI}, or long, EYT_L, cycle). Each trial consisted of a resolvable incomplete block design with four replications where four to five incomplete blocks nested within each complete replication were used. All five locations were used for EYTs (Table 1). A total of 52 historic as well as commercial varieties were used as checks in the trials having each trial between eight and 20 checks at a time.

Genotypic Data

The 1477 wheat lines were genotyped by GBS (Elshire et al. [2011] modified by Poland et al. [2012a] for wheat). The SNP identifications were conducted using the Tassel pipeline (Glaubitz et al. [2014] as in Lado et al. [2013]). Briefly, markers with a minor allele frequency smaller than 0.01 and more than 80% of missing data were discarded. We identified 81,999 SNPs. Marker-data imputation was conducted using the realized relationship matrix through the multivariate normal expectation maximization method using the rrBLUP package (Endelman, 2012) from R software (R Development Core Team, 2015).

Phenotypic Data Analysis

Phenotypic best linear unbiased estimations (BLUEs) were obtained for all genotypes present in each trial. Field analysis was conducted according to experimental design. Since PYTs consisted of a series of smaller α-design trials connected through common checks, the following model was used to estimate genotypic means for each heading date group in each environment:

$$y_{ijkl} = \mu + g_i + t_j + r_{k(j)} + b_{l(jk)} + \varepsilon_{ijkl} \quad [1],$$

where μ is the overall mean, g_i is the fixed effect of the ith genotype, t_j is the random effect of the jth trial, $r_{k(j)}$ is the fixed effect of the kth replicate nested within the jth trial, $b_{l(jk)}$ is the random effect of the lth incomplete block nested within the jth trial and the kth replicate, and ε_{ijkl} is the residual error for the ith genotype in the lth incomplete block within the kth replicate in the jth trial with t_j, $b_{l(jk)}$ and ε_{ijkl} as random variables being $t_j \sim N(0,\sigma_t^2)$, $b_{l(jk)} \sim N(0,\sigma_b^2)$ and $\varepsilon_{ijkl} \sim N(0, \sigma_e^2)$.

The advanced and elite yield trials (AYT and EYT) consisted of α-designs, and therefore the following model was used to estimate genotypic means for each heading date group in each environment:

$$y_{ikl} = \mu + g_i + r_k + b_{l(k)} + \varepsilon_{ikl} \quad [2],$$

where μ, g_i, r_k, $b_{l(k)}$, and ε_{ikl} were defined as in Eq. [1], with $b_{l(k)}$ and ε_{ikl} as random variables being $b_{l(k)} \sim N(0,\sigma_b^2)$ and $\varepsilon_{ikl} \sim N(0, \sigma_e^2)$. Best linear unbiased estimates were estimated for each genotype in each heading date group and each environment.

Broad-sense heritability was estimated for each trial. We estimated variance component using restricted maximum likelihood method estimates from Eq. [1] or [2] but with genotypes as random effect. All of the analyses were conducted in R software (R Development Core Team, 2015) with package nlme (Pinheiro et al., 2013).

Characterization of Genotype × Environment Interaction

Several strategies were used to characterize GEI including variance component estimation, correlation between environments, AMMI and GGE models, and graphical representation through augmented biplots.

The magnitude of GEI for yield was first characterized with variance components for all the years and locations. A random effect model using the PROC MIXED procedure of SAS (SAS Institute, 2005) was used to estimate variance components for each yield trial (i.e., PYT, AYT, and EYT). Second, Pearson correlations between all pairs of environments using all the shared lines between each pair of environments were estimated.

Multiplicative models such as AMMI (Gauch and Zobel, 1988; Zobel et al., 1988) and GGE (Yan et al., 2000) were used to further characterize GEI. Since most of the GEI structure is given by the EYT that are evaluated in multiple locations, a subset consisting of 60 elite genotypes were chosen to construct both AMMI and GGE models. These genotypes were chosen to have at least 50% coverage in the genotype × environment table of means and because the GEI structure they create is similar to that of the complete set of genotypes that have a 7% coverage (Fig. 1; Supplemental Fig. S1). A mixed model with the realized additive relationship matrix to model covariance among genotypes and the estimated correlation matrix across environments was fitted to predict genotypic performance of all genotypes in all the environments before fitting AMMI and GGE models (Henderson, 1984; Bernardo, 2010). The genetic correlation matrix between environments was estimated using the spectral decomposition method described by Rebonato and Jäckel (1999) and subsequently used as the variance–covariance matrix. The mixed model for genomic prediction was fit with the rrBLUP package (Endelman, 2012) using the Kronecker product between the additive relationship matrix among genotypes and the estimated variance–covariance matrix among environments. This model assumes equal genotypic relationship matrix in all environments (for further description of this model, see the following section). The AMMI1 to AMMI10 and GGE1 to GGE10 were fitted in SAS statistical software (SAS Institute, 2005) and the Gollob's (1968) statistical test was used to choose the number of significant axis from both AMMI and GGE model families. Environments were grouped in MEs with the best AMMI and GGE models. Additionally, we compared the MEs formed with the best models to those of the simplest models: GGE2 and AMMI2. Finally, since the MEs created by the different strategies were similar, a GGE2 augmented biplot representation was used to visually represent the GEI structure using the R statistical software (R Development Core Team, 2015) with the GGEBiplotGUI package (Frutos et al., 2014).

Genomic Predictions and Prediction Accuracy

To compare strategies to deal with GEI, we partitioned the data in 13 sets of environments: by year (i.e., all locations for 2010, 2011, 2012, 2013, or 2014), by location (i.e., all years for DOL, LE1, R2, or YOU), by ME (i.e., all location–year combinations within each ME1, ME2, and ME3), and all environments (all location–year combinations [ALL]). We used two general approaches to make genomic predictions: overall predictions for the mean performance within each set of environments ($GBLUP_M$, i.e., predictions of genomic estimated breeding values [GEBVs] for the mean performance of genotypes in each set of environments: years, locations, ME, or ALL), and predictions by environment ($GBLUP_{g\times e}$, i.e., predictions of GEBVs for each environment within sets created by years, locations, ME, or ALL). We used the following general model structure to make predictions:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{g} + \varepsilon \quad [3]$$

For $GBLUP_M$: $\mathbf{y}_{N\times1}$ is the vector of mean yield for each genotype in the set of environments (i.e., the BLUEs from a model accounting for field design and environment) of length N (N = population size or number of genotypes in the set), $\mathbf{1}_{N\times1}$ is a vector of ones of length N; μ is the overall mean; $\mathbf{g}_{N\times1}$ is a random vector of genotypic predictors with $\mathbf{g} \sim N(0, \mathbf{A}_{N\times N}\sigma_g^2)$, where $\mathbf{A}$ is the realized additive relationship matrix; $\mathbf{Z}_{N\times N}$ is the corresponding incidence matrix; and ε is the residual errors vector with $\varepsilon \sim N(0,\mathbf{R}_{N\times N})$ where $\mathbf{R}$ is a diagonal matrix with elements as the reciprocals of the variances of the adjusted means for each genotype from the phenotypic analysis stage to account for heterogeneity in mean estimate precision following Rodrigues et al. (2014).

In $GBLUP_{g\times e}$, $\mathbf{y}_{n\times1}$ is the mean vector of yield for each genotype in each environment (i.e BLUEs from field experiments) with length n (n = number of genotypes [N] by number of environment [k]: $N \times k$); $\mathbf{1}_{n\times1}$ is a vector of ones of length n; μ is the overall mean; $\mathbf{g}_{n\times1}$ is a random vector of genotype × environment predictors with $\mathbf{g} \sim N[0, (\mathbf{A}_{N\times N} \otimes \boldsymbol{\rho}_{k\times k})\sigma_g'^2]$, where $\boldsymbol{\rho}_{k\times k}$ is the estimated correlation matrix among environments and k is the

number of environments; $\mathbf{Z}_{n\times n}$ is the corresponding incidence matrix; and ε is the residual errors vector with $\varepsilon \sim N(0,\mathbf{R}_{n\times n})$, where **R** is a diagonal matrix with elements as the reciprocals of the variances of the adjusted means for each genotype from the phenotypic analysis stage to account for heterogeneity in mean estimate precision following Rodrigues et al. (2014).

Accuracy of GEBVs was estimated as the correlation between predicted and observed genotypic values using cross-validation (CV, Legarra et al., 2008). A *k*-fold CV (CV_1) was used following Burgueño et al. (2012). Briefly, the observations were randomly divided into *k* non-overlapping subsets. Then, $k - 1$ groups were used as training set, and the remaining group was used as the validation set (i.e., GEBV are obtained for each individual in the validation set). This procedure is followed until breeding values from individuals in all *k* subsets are estimated. A total of 100 replications of the CV with $k = 7$ were performed and the correlation between GEBV and observed genotypic values was used to estimate the accuracy of GEBV.

Predictions of New Genotypes (Unphenotyped Genotypes)

To predict new genotypes we compared the $GBLUP_M$ with the $GBLUP_{g\times e}$ models. To make unbiased comparisons, the predictive ability of the $GBLUP_M$ model was estimated with the sevenfold CV (CV_1). In the case of the $GBLUP_{g\times e}$ model, the sevenfold CV was modified (CV_2) to predict breeding values for a genotype without any phenotypic information per se from the same genotype in other environments. Briefly, we split the genotype × environment data into seven subsets. However, to ensure fairness when predicting new genotypes, no information from a given genotype was used to predict its own performance. Therefore, yield data from multiple environments for a single genotype were grouped in the same *k*th subset. Additionally, sets were constructed to have the same amount of information (i.e., data points) and not necessarily the same number of genotypes. Therefore, the number of genotypes in each subset is slightly different. We also used 100 replications of the CV to estimate the correlation between GEBV and observed genotypic values.

Finally, because the sets of environments include different number of genotypes (i.e., population size), we compared the predictive ability of $GBLUP_M$ and $GBLUP_{g\times e}$ models for all sets of environments using a series of random samples with different sizes for all sets (i.e., $N = 99$, $N = 200$, $N = 500$, and $N = 900$). We compared the predictive ability using the sevenfold CV described earlier (CV_1 and CV_2).

Predictions of New Environments (Unphenotyped Environments)

Different sets of environments were used to predict genotypes in new environments (unphenotyped environments) using both GBLUP models. First, we considered environments grouped by year and we predicted the genomic breeding values for the tested genotypes in each location within a year using all the remaining locations for that year. This way, the training set consisted of all genotypes in all the locations less one, and the testing set consisted of the same genotypes in the remaining location. We followed this procedure to predict genomic breeding values for all genotypes in all locations one at a time. The accuracy of the genomic prediction was estimated as the correlation between genomic estimated breeding values for a specific location and the observed yield performance of genotypes for that location. Second, we used the same principle but for environments grouped by location. We used all years less one for a specific location as the training set while we used the remaining year as the testing set. Finally, we used the same principle but for environments grouped in MEs. We trained the model with all environments (i.e., combination of locations and years) in a ME less one, and we tested it with the remaining environment for that ME.

RESULTS

Phenotypic Data Analysis

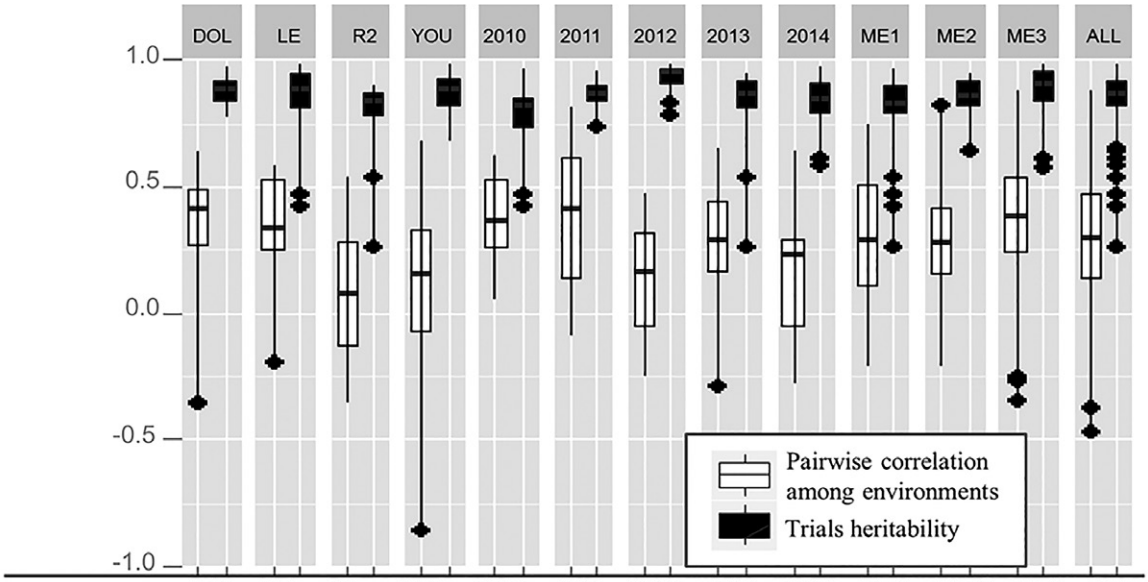
The dataset is highly unbalanced having pairs of environments sharing from six to 268 genotypes (Fig. 1). The LE1 location included PYT, AYT, and EYT, it is therefore the environment with the largest number of genotypes (Table 1; Fig. 1). The other locations with the largest number of genotypes were YOU in all years, DOL in 2010 to 2013, and R2 in 2014.

The broad-sense heritability estimated on a line-mean basis was high for most trials (i.e., >0.65 with an average of 0.82; Fig. 2), and only four trials had low heritability (i.e., 0.26 and 0.54 in EYT in R2 in 2013 and 0.42 and 0.47 in LE1 in 2010; Fig. 2).

Environmental Characterization of Genotype × Environment Interaction

Genotype × environment interaction contributed a large proportion of the total variance, being 7.6, 16.9, and 15.5% of the total variance in the PYT, AYT, and EYT, respectively (Table 2). Specifically, GEI explained much more than the genotypic variance in AYT and EYT, where several years were evaluated. The year component was large in all cases and GEI among years is larger than GEI among locations (Table 2).

Eight components were significantly different from zero in both AMMI and GGE models (Table 3). The best GGE model family (i.e., GGE8) identified three MEs that were the same as the simplest model (i.e., GGE2, data not shown); ME1 (with 11 location–year combinations) and ME2 (with 10 location–year combinations) included all locations for 2010, 2011, and 2013 while ME3 (with 14 location–year combinations) included all locations for 2012 and 2014 (Fig. 3). In GGE models, the first two components explained 74% and the first eight explained 97% of total variance (genotype plus GEI variance). The best one of the AMMI model family (i.e., AMMI8) grouped environments into two MEs collapsing ME1 and ME2 into a single group and discriminating years 2012 and 2014 from the remaining years (data not shown). In the AMMI models, the first two components explained 74%, and the first eight components explained 96% of the total variance (GEI variance). The environments with the largest discriminative ability were LE4 in 2011, LE2 and YOU in 2013, LE1 and YOU



	DOL	LE	R2	YOU	2010	2011	2012	2013	2014	ME1	ME2	ME3	ALL
N Trials	16	32	12	19	19	28	23	17	22	36	43	42	121
N Locations	1	1	1	1	7	8	7	7	6	7	7	7	8
N Years	4	5	4	5	1	1	1	1	1	3	4	4	5
N Genotypes	248	1477	159	287	387	747	99	412	422	906	914	214	1477
N GenoxEnv	389	2164	262	428	640	1180	450	739	740	1395	1523	901	3820
Proportion GxE	0.31	0.28	0.33	0.25	0.24	0.20	0.65	0.26	0.29	0.14	0.17	0.30	0.07

Figure 2. Pairwise correlations among environments and heritability for all trials (PYT_S, PYT_I, PYT_L, AYT_S, AYT_I, AYT_L, EYT_{SI}, and EYT_L) within sets of environments. The number of trials, years, locations, and the proportion of the genotype × environment interaction for each set of environments is indicated. The 35 environments are grouped by either years (i.e., 2010–2014), locations (i.e., Dolores [DOL], La Estanzuela [LE1], Ruta2 [R2], Young [YOU]), mega-environment (ME1–ME3), or all together (ALL).

Table 2. Yield variance component estimation and proportion of the total variance for preliminary (PYT), advanced (AYT), and elite (EYT) yield trial for 1477 wheat genotypes evaluated in 35 environments that are a combination of locations and years.

	PYT		AYT		EYT	
Variance Component	Yield	Proportion	Yield	Proportion	Yield	Proportion
	kg ha^{-1}	%	kg ha^{-1}	%	kg ha^{-1}	%
Location	–	–	100576	3.2	307877	10.2
Year	1640866	46.6	2069766	66.5	1221367	40.4
Genotype	921628	26.2	101345	3.3	172070	5.7
Genotype × location	–	–	85633	2.8	61219	2.0
Genotype × year	266829	7.6	163231	5.2	156339	5.2
Genotype × location × year	–	–	278384	8.9	252138	8.3
Replication	72844	2.1	36548	1.2	458388	15.2
Incomplete block	56121	1.6	51841	1.7	40882	1.4
Residual	561969	16.0	225249	7.2	351171	11.6
Total	3520257	100.0	3112573	100.0	3021451	100.0

in 2014, while the environments with the least discriminative ability were YOU in 2010 and LE2 in 2014.

Genomic Predictions of Unphenotyped Genotypes

Predicting genotypic performance of new genotypes for each environment with the $GBLUP_{g\times e}$ was better than predicting for the mean of the environment with the $GBLUP_M$ for most sets of environments (Fig. 4). The advantage of the $GBLUP_{g\times e}$ model was more evident for larger population sizes (Fig. 4, 5). Furthermore, LE1 (the location with the largest number of genotypes) had the largest difference in mean prediction comparing both GBLUP models (Fig. 5), $GBLUP_{g\times e}$ being better than the $GBLUP_M$. On the other hand, most of the sets grouping environments by years have higher accuracies with the $GBLUP_M$ model (i.e., 2012, 2013, and 2014; Fig. 4, 5). In addition, 2012 shows the highest accuracy for the $GBLUP_M$ model, with an accuracy of 0.5 (Fig. 5). Finally, no differences in prediction ability between $GBLUP_M$ and $GBLUP_{g\times e}$ were found for large data sets ($N = 500$ or $N = 900$; Fig. 5).

Table 3. Additive main effect and multiplicative interaction (AMMI) and GGE family models for the genotype × environment interaction analysis of wheat lines evaluated in 35 environments including locations in years. The sum of squares (SS), proportion of the total GEI variance, and test of significance for the first 10 principal components is shown.

	AMMI			GGE		
Model	SS	Proportion	*P*-value	SS	Proportion	*P*-value
1	1396.2	51.46	<0.0001	1889.8	44.02	<0.0001
2	616.6	22.72	<0.0001	1306.2	30.42	<0.0001
3	198.0	7.27	<0.0001	544.2	12.67	<0.0001
4	148.9	5.48	<0.0001	186.8	4.35	<0.0001
5	88.4	3.26	<0.0001	88.5	2.06	<0.0001
6	66.0	2.43	<0.0001	68.2	1.58	<0.0001
7	44.6	1.71	0.0003	47.3	1.10	0.0002
8	36.2	1.33	0.0332	36.3	0.84	0.0312
9	24.5	0.90	0.6731	27.1	0.63	0.4363
10	22.1	0.81	0.8108	23.3	0.54	0.7328

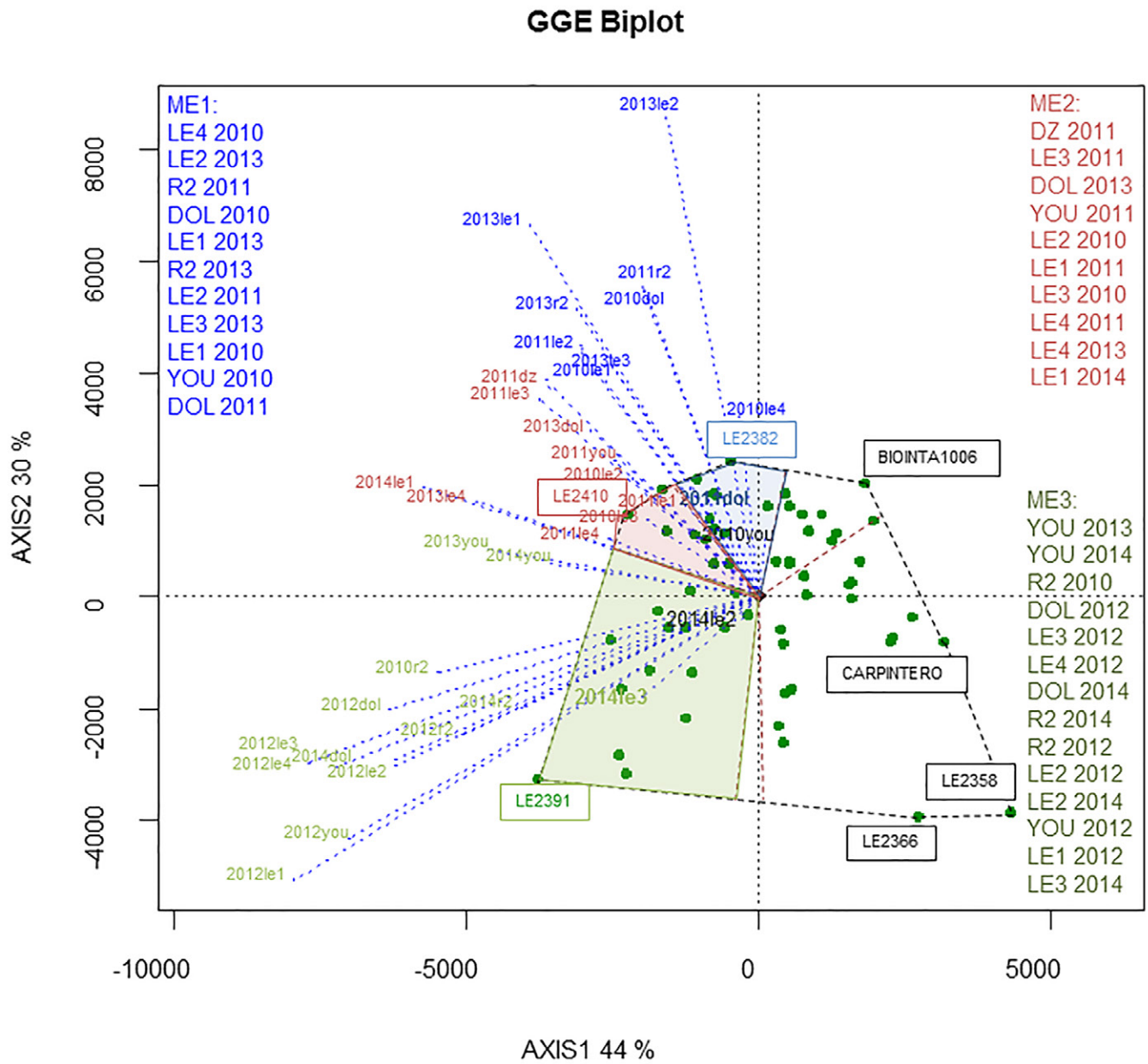


Figure 3. GGE biplot for yield in a wheat dataset considering 60 genotypes evaluated in 35 environments. Grouping in three mega-environments (MEs) based on winning genotypes (i.e., indicated within square) are shown (i.e., represented as polygon): ME1, ME2, and ME3. The environments for each ME are labeled outside the polygon.

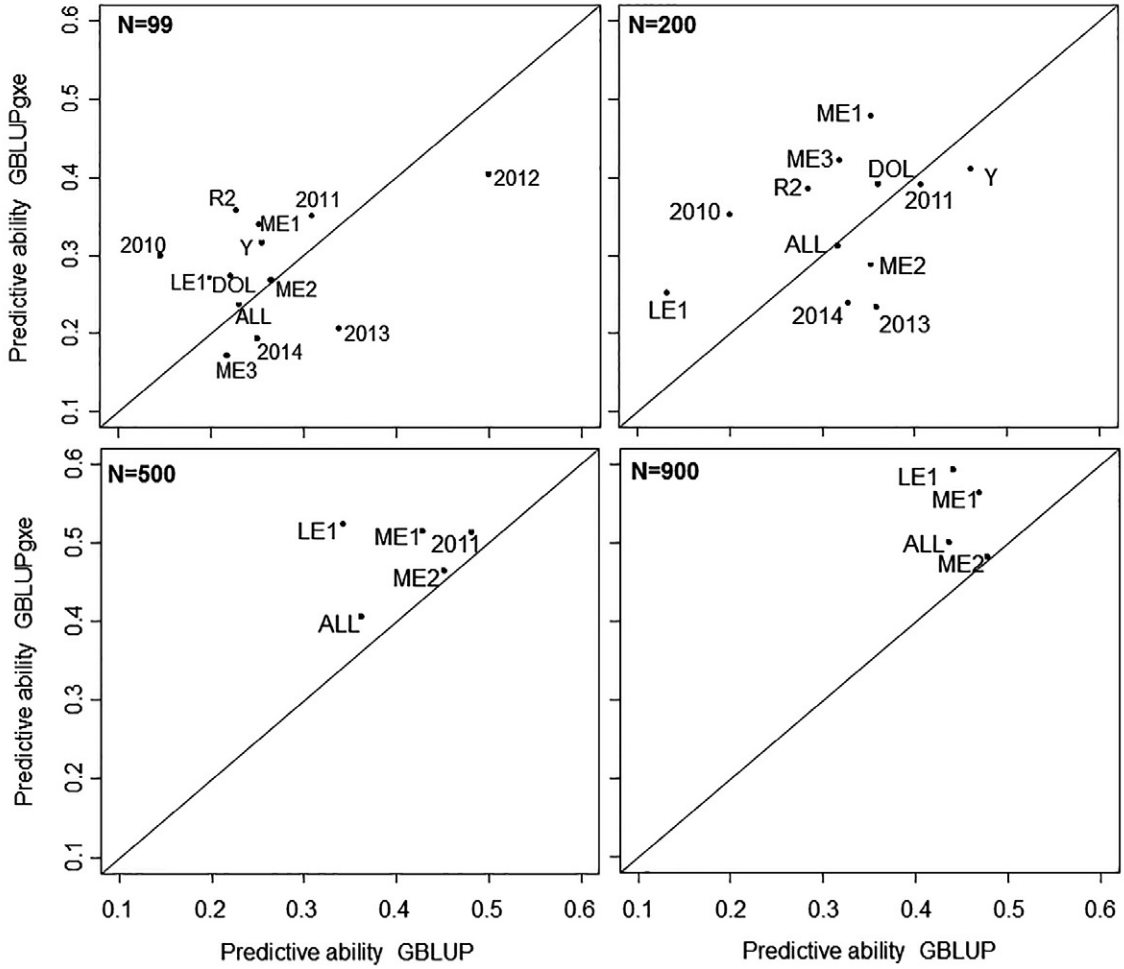


Figure 4. Accuracy of genomic breeding values predictions for yield in different sets of environments predicted either for the mean ($GBLUP_M$) or each environment within a set ($GBLUP_{g\times e}$) in 1477 wheat genotypes evaluated in 35 environments. The accuracy was estimated for different sample sizes (N = 99, N = 200, N = 500 and N = 900). Environments are grouped in sets by either years (i.e., 2010–2014), locations (i.e., DOL, LE1, R2, YOU), mega-environments (ME1–ME3), or all together (ALL).

Genomic Predictions of Unphenotyped Environments

When comparing independent training and testing sets and predicting for new environments, predictions for specific locations were high with either the $GBLUP_M$ or the $GBLUP_{g\times e}$ models using either information from other locations in a single year or using other environments within a ME (Table 4). However, predictions across years are poorly achieved using single locations with either GBLUP models and are only reasonably achieved predicting within ME with either the $GBLUP_M$ or the $GBLUP_{g\times e}$ models (Table 4). The $GBLUP_M$ model should be used for predicting performance of genotypes in new environments (i.e., environments without previous phenotypic information) because the $GBLUP_{g\times e}$ model requires an estimation of genotypic correlations between environments that would be unavailable if no phenotypic information is available.

DISCUSSION

One of the most challenging issues in plant breeding is the use of highly unbalanced existing phenotyping in the presence of GEI (Burgueño et al., 2012; Dawson et al., 2013; Heslot et al., 2014). Breeding programs produce large MET datasets that are highly unbalanced and with different levels of quality because of the degree of replication in different trials (Bernardo, 2010). We used a large MET dataset with 1477 genotypes extensively phenotyped in 35 environments to compare different strategies for handling GEI. Our dataset had very good experimental design yield data quality with an average heritability on a line-mean basis of 0.82, and our predictive ability was moderate to high for a complex trait like grain yield with an overall prediction accuracy of 0.5. We specifically aimed to answer the following questions: Could we improve the predictive ability by modeling the variance–covariance matrix across environments and predicting performance of genotypes by environment? Could we design groups of environments were overall predictions

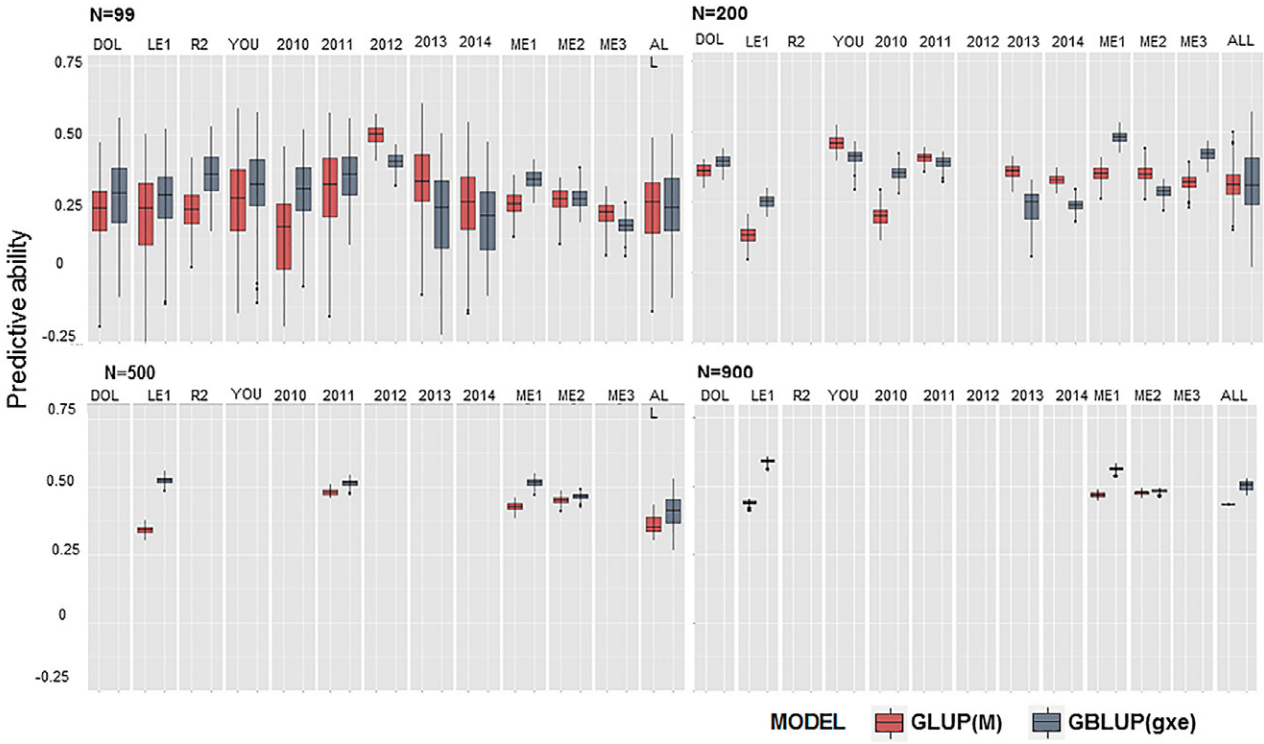


Figure 5. Accuracy of genomic breeding values predictions for yield in different sets of environments predicted either for the mean ($GBLUP_M$) or each environment within a set ($GBLUP_{g\times e}$) in 1477 wheat genotypes evaluated in 35 environments indicating the standard deviation of the accuracy. The accuracy was estimated for different samples sizes ($N = 99$, $N = 200$, $N = 500$, and $N = 900$). Environments are grouped in sets by either years (i.e., 2010–2014), locations (i.e., Dolores [DOL], La Estanzuela [LE1], Young [YOU], Ruta2 [R2]), mega-environments (ME1–ME3) or all together (ALL). The range of prediction accuracy is for 100 replications of the sevenfold cross-validation.

Table 4. Accuracy for unphenotyped environments with $GBLUP_M$ (1) and $GBLUP_{g\times e}$ (2), using independent training and testing sets for 1477 genotypes evaluated in 35 environments. Environments were grouped in sets by either year (2010–2014), location (Dolores [DOL], La Estanzuela [LE1], Young [YOU], Ruta2 [R2]) or mega-environment (ME1–ME4). Different training sets were used to predict genotypic performance in new environments (i.e. the testing set).

			Testing location							
			LE		YOU		DOL		R2	
Set	Training set	Testing Year	1	2	1	2	1	2	1	2
By year†	All locations in testing year but one	2010	0.31	0.31	0.12	0.17	0.71	0.71	0.57	0.44
		2011	0.42	0.39	0.39	0.31	0.52	0.37	0.44	0.57
		2012	0.74	0.79	0.44	0.54	0.78	0.81	0.57	0.65
		2013	0.03	0.05	0.40	0.74	0.36	0.59	0.28	0.36
		2014	0.23	−0.14	0.30	0.19	0.58	0.71	0.30	0.19
By location‡	All years for a location but one	2010	0.16	0.17	0.01	−0.02	−0.06	0.04	0.01	0.11
		2011	0.23	0.18	0.27	0.34	0.08	0.14	0.19	0.29
		2012	0.50	0.31	0.10	0.04	0.33	0.29	0.05	0.15
		2013	0.23	0.10	−0.05	0.08	0.31	0.56	0	0.19
		2014	0.21	0.20	0.06	−0.05	0.08	0.36	−0.05	0.33
By mega-environment§	All environments for a specific mega-environment but one	2010	0.10 (1)¶	0.20 (1)	0.11 (1)	0.18 (1)	0.72 (1)	0.72 (1)	0.32 (3)	0.35 (3)
		2011	0.36 (2)	0.35 (2)	0.36 (2)	0.44 (2)	0.39 (1)	0.28 (1)	0.38 (1)	0.39 (1)
		2012	0.63 (3)	0.45 (3)	0.46 (2)	0.54 (2)	0.75 (3)	0.80 (3)	0.63 (3)	0.30 (3)
		2013	0.28 (1)	0.18 (1)	0.46 (3)	0.43 (3)	0.43 (2)	0.50 (2)	0.44 (1)	0.32 (1)
		2014	0.24 (2)	0.25 (2)	−0.01 (2)	0.11 (2)	0.63 (3)	0.78 (3)	−0.03 (3)	0.20 (3)

† When sets of environments are grouped by year, the training set includes all genotypes in all locations but one for a given year, while the testing set includes the remaining location; one location at a time is left out of the training set and is being predicted.

‡ When sets of environments are grouped by location, the training set includes all genotypes in all years but one for a given location, while the testing set includes the remaining year; 1 yr at a time is left out of the training set and is being predicted.

§ When sets of environments are grouped by ME, the training set includes all genotypes in all environments within a ME but one, while the testing set includes the remaining environment; one environment at a time is left out of the training set and is being predicted.

¶ Numbers 1 through 3 indicate the ME to which the specific environment belongs.

outperformed environment-specific predictions? Could we borrow information from other environments to predict performance in unphenotyped environments?

Correlation between Environments and Mega-Environments

Mixed Models are one of the most effective tools to handle unbalanced data while modeling correlation structures (Piepho, 1997, 1998; Yu et al., 2006; Piepho et al., 2007; Piepho and Möhring, 2007; Malosetti et al., 2007; Meyer, 2009). However, reliable estimates of correlation should be used to model the variance–covariance matrix among environments. In breeding programs, it is common to evaluate lines using less replication and environments during early testing (i.e., in PYT) saving resources for larger replication and environments of fewer lines in advanced trials (i.e., AYT or EYT; Bernardo, 2010). This creates a structure in GEI data where only promising lines are evaluated in multiple environments. Therefore, the estimation of the correlation among environments could be biased by the use of selected lines. We found that the genetic correlations matrices between environments estimated with 60 elite lines was similar to the correlation estimated with all available lines (i.e., $r = 0.7$, $p < 0.001$; Fig. 1; Supplemental Fig. S1). Furthermore, the genetic correlation between two environments for unselected lines (i.e., AYT) was similar to the correlation between the same two environments for selected lines (i.e., EYT; Supplemental Fig. S2). Piepho and Möhring (2006) argue that it is reasonable to assume that the change in allelic frequencies as a result of selection from training to testing populations is minimal, and therefore, prediction accuracy in the testing population is appropriate. Therefore, it is reasonable to assume that genotypic correlations among environments from selected individuals properly estimate correlations for the entire population. Other studies have shown that creating unbalanced data to sample more environments and genotypes produce higher prediction accuracy than having smaller population sizes and fewer environments with balanced data (Endelman et al., 2014).

Genotype × Environment Interaction Characterization

Genotype × environment interaction is a widespread phenomenon in plants, and therefore, a proper characterization of the GEI is key to decide on strategies to handle it (Crossa and Cornelius, 1997; Piepho, 1997, 1998; Burgueño et al., 2008; Malosetti et al., 2013). We found three MEs in our dataset: ME1 and ME2 have a strong correlation between environments and are composed of 2010, 2011, and 2013 environments; ME3 also has a strong correlation between environments and are composed of 2012 and 2014 environments. This pattern was detected by either the use of the best GGE or AMMI model, but also by the GGE2 model using the winning genotypes in each ME (i.e., genotypes LE2382, LE2410, and LE 2391; Fig. 3). Several studies have shown the advantages of either AMMI (Gauch, 2006, 2013; Gauch et al., 2008) or GGE models (Yan et al., 2000, 2007; Yang et al., 2009) to model GEI, but we identified the same general pattern with either an AMMI or a GGE model. We found that genotype × year interactions contributed a large proportion of the GEI in our study. Year-to-year variations are more difficult to predict (Dawson et al., 2013) and usually contribute to a large proportion of the GEI (Ceretta and van Eeuwijk, 2008; Heslot et al., 2014). In our study, years 2010, 2011, and 2013 were high-yielding environments (i.e., yield average of 5370 kg ha^{-1}), while years 2012 and 2014 were low- or medium-yielding environments (i.e., yield average of 3946 kg ha^{-1}). Other studies found a similar pattern of environments grouping by yield performance (Byth et al., 1976; Windhausen et al., 2012). Furthermore, we found that while the average monthly rainfall in these locations is 80 mm mo^{-1}, heavy rainfall occurred during anthesis and grain filling in years 2012 and 2014 (Supplemental Fig. S3). The heavy rainfall increased disease presence, causing a severe infection of Fusarium head blight, which, in turn, decreased yield performance of lines. Further environmental characterization will provide a better insight into the GEI structure, but, by and large, we were able to successfully predict for rainy–dry anthesis period environments and might be able to predict performance for Fusarium and non-Fusarium years in this dataset.

Predictive Ability

Plant breeders seek to develop varieties with the best performance for their target environments (Comstock, 1977). Most research on genomic selection has focused on accurately predicting the genotypic mean performance over all environments, assuming that all trial locations belong to the same target population of environments (Asoro et al., 2011; Charmet and Storlie, 2012; Storlie and Charmet, 2013; Crossa et al., 2013; Dawson et al., 2013). However, since GEI is widespread, the accuracy of genomic prediction could be lowered especially by the presence of crossover interaction (Burgueño et al., 2012; Dawson et al., 2013; Jarquín et al., 2014). Mixed models are powerful tools to make predictions by modeling the variance–covariance structure across environments to account for GEI (Piepho, 1997; Burgueño et al., 2008; Meyer, 2009; Malosetti et al., 2013). We compared two strategies to deal with GEI: reduce it (i.e., predict overall performance for a set of environments, $GBLUP_M$) and exploit it (i.e., predict performance by environment using the correlations among environments, $GBLUP_{g\times e}$). We obtained high overall predictions accuracies for grain yield ($GBLUP_M$) compared with other studies (Heffner et al., 2011; Burgueño et al., 2012; Poland et al., 2012b; Charmet and

Storlie, 2012; Dawson et al., 2013; Lado et al., 2013; Crossa et al., 2014; Lopez-Cruz et al., 2015). This probably is due to the large population size with extensive phenotyping and high quality experimental design with two to four replications and high heritability. High prediction ability has been found to be associated with large population sizes, a large number of sampled environments (Riedelsheimer and Melchinger, 2013) even at the cost of unbalanced designs (Heffner et al., 2011; Endelman et al., 2014), and good experimental designs (Cooper et al., 2014) with large numbers of replications (Lorenz, 2013).

To reduce GEI, environments were grouped into sets based on different criteria (i.e., year, location, or ME). The genotype × environment model (i.e., $GBLUP_{g\times e}$) outperformed the overall model ($GBLUP_M$) when predicting within MEs or locations. On the other hand, most of the overall predictions within years (i.e., $GBLUP_M$) were higher than by-environment predictions (i.e., $GBLUP_{g\times e}$), indicating low GEI within years. The only exception to this pattern is the location YOU. The improvement in the predictive ability of the $GBLUP_{g\times e}$ can be explained by the shared information among environments. The $GBLUP_M$ model is equivalent to the $GBLUP_{g\times e}$ model with a correlation between environments equal to one. Therefore, when correlations among environments are high (like correlations among locations within years), similar predictive ability is found between $GBLUP_M$ and $GBLUP_{g\times e}$ models.

Furthermore, we found that predicting unphenotyped environments is better achieved within MEs and years. This result is not unexpected because lower GEI is found within years or MEs, while higher GEI is found across years. The use of multienvironment models to make overall genomic predictions has been documented elsewhere (Jarquín et al., 2014). However, the use of predictions within MEs has improved predictive ability in some studies (Burgueño et al., 2012; Dawson et al., 2013) but not in others (Windhausen et al., 2012). We found that grouping environments in MEs that are mainly based on year-to-year repeatable variations and predicting overall performance yielded high predictive ability in most of the situations. Therefore, using locations to sample GEI is not a good strategy to capture all GEI for this program. Furthermore, predictions based on single-year phenotypic evaluation were unreliable, and historic data from breeding programs in the form of MET data is the key to characterize GEI and to increase predictive ability of GS models. However, we found that even random environmental factors (i.e., years) could be better modeled when GEI is included in the prediction model with the $GBLUP_{g\times e}$ model. Furthermore, when the size of population increases, all sets of environments and models tended to have the same performance with close predictive ability.

We were able to successfully predict new genotypes borrowing information from relatives evaluated in multiple environments using the $GBLUP_{g\times e}$ model. This model proved to be superior for predicting unphenotyped genotypes. However, where genomic predictions for new environments is desirable, the use of genetic correlations between environments is not possible, and therefore the $GBLUP_M$ should be used. We showed that for predicting performance of genotypes in new environments, the best strategy was to either predict across locations for single years or to predict within ME for any year or location.

CONCLUSIONS

Modeling GEI increased predictive ability. The use of MEs to identify groups of environments with small GEI, and overall predictions within MEs produces higher prediction accuracy compared with overall predictions across years. Borrowing information from environments by modeling the variance–covariance matrix is beneficial for predicting new genotypes before phenotyping. Finally, when large data sets were used, the sets of environments and models tend to perform similar.

Acknowledgments

We would like to thank three anonymous reviewers for their comments that contributed to this manuscript. We would also like to thank Jesse Poland for his review of this manuscript. Funding for this project was provided by INIA (Project L2-2011) and the Research and Innovation Agency of Uruguay (ANII), through Grant POS_NAC_2013_1_11261. We express our appreciation for the effort of the technical personnel of INIA wheat breeding program Leonardo Hernández, María Ferreira, Dumas Laun, Abel López, and José Flores.

References

Asoro, F.G., M.A. Newell, W.D. Beavis, M.P. Scott, and J.L. Jannink. 2011. Accuracy and training population design for genomic selection on quantitative traits in elite North American oats. Plant Gen. 4:132–144. doi:10.3835/plantgenome2011.02.0007

Bennewitz, J., T. Solberg, and T. Meuwissen. 2009. Genomic breeding value estimation using nonparametric additive regression models. Genet. Sel. Evol. 41:20–32. doi:10.1186/1297-9686-41-20

Bernardo, R. 2010. Breeding for quantitative traits in plants. 2nd ed. Stemma Press, Woodbury, MN. ISBN 978-0-9720724-1-0.

Boer, M.P., D. Wright, L. Feng, D.W. Podlich, L. Luo, M. Cooper, and F.A. van Eeuwijk. 2007. A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. Genetics 177:1801–1813. doi:10.1534/genetics.107.071068

Braun, H.J., S. Rajaram, and M. van Ginkel. 1996. CIMMYT's approach to breeding for wide adaptation. Euphytica 92:175–183. doi:10.1007/BF00022843

Burgueño, J., J. Crossa, P.L. Cornelius, and R.C. Yang. 2008. Using factor analytic models for joining environments and genotypes without crossover genotype × environment interaction. Crop Sci. 48:1–15. doi:10.2135/cropsci2007.11.0632

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

Byth, D., R. Eisemann, and I. DeLacy. 1976. Two-way pattern analysis of a large data set to evaluate genotypic adaptation. Heredity (Edinb) 37:215–230. doi:10.1038/hdy.1976.84

Ceretta, S., and F.A. van Eeuwijk. 2008. Grain yield variation in malting barley cultivars in Uruguay and its consequences for the design of a trials network. Crop Sci. 48:167–180. doi:10.2135/cropsci2006.06.0428

Charmet, G., and E. Storlie. 2012. Implementation of genome-wide selection in wheat. Russ. J. Genet. Appl. Res. 2:298–303. doi:10.1134/S207905971204003X

Comstock, R.E. 1977. Quantitative genetics and the design of breeding programs. In: E. Pollack, O. Kempthorne, T.B. Bailey Jr., editors, Proceedings of the International Conference on Quantitative Genetics. 16–21 Aug. 1976. Iowa State Univ. Press, Ames, IA. p. 705–718.

Cooper, M., and I.H. 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

Cooper, M., C. Messina, and D. Podlich. 2014. Predicting the future of plant breeding: Complementing empirical evaluation with genetic prediction. Crop Pasture Sci. 65:311–336. doi:10.1071/CP14007

Crossa, J. 2012. From genotype × environment interaction to gene × environment interaction. Curr. Genomics 13:225–244. doi:10.2174/138920212800543066

Crossa, J., Y. Beyene, S. Kassa, P. Pérez, J.M. Hickey, C. Chen, G. de los Campos, J. Burgueño, V.S. Windhausen, E. Buckler, J.L. Jannink, M. a Lopez Cruz, and R. Babu. 2013. Genomic prediction in maize breeding populations with genotyping-by-sequencing. G3: Genes, Genomes, Genet. 3:1903–1926.

Crossa, J., and P.L. Cornelius. 1997. Sites regression and shifted multiplicative model clustering of cultivar trial sites under heterogeneity of error variances. Crop Sci. 37:406–415. doi:10.2135/cropsci1997.0011183X003700020017x

Crossa, J., P. Pérez, J. Hickey, and J. Burgueño. 2014. Genomic prediction in CIMMYT maize and wheat breeding programs. Heredity (Edinb) 112:48–60. doi:10.1038/hdy.2013.16

Cullis, B.R., A.B. Smith, C.P. Beeck, and W.A. Cowling. 2010. Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaction using factor analysis. Genome 53:1002–1016. doi:10.1139/G10-080

Dawson, J.C., J.B. Endelman, N. Heslot, J. Crossa, J. Poland, S. Dreisigacker, Y. Manès, M.E. Sorrells, and J.L. Jannink. 2013. The use of unbalanced historical data for genomic selection in an international wheat breeding program. F. Crop. Res. 154:12–22. doi:10.1016/j.fcr.2013.07.020

DeLacy, I.H., K.E. Basford, M. Cooper, J.K. Bull, and C.G. McLaren. 1996. Analysis of multi environment trials: An historical perspective. Plant adaptation and crop improvement. 39–124.

de los Campos, G., J.M. Hickey, R. Pong-Wong, H.D. Daetwyler, and M.P.L. Calus. 2013. Whole-genome regression and prediction methods applied to plant and animal breeding. Genetics 193:327–345. doi:10.1534/genetics.112.143313

Dreccer, M.F., M.G. Borgognone, F.C. Ogbonnaya, R.M. Trethowan, and B. Winter. 2007. CIMMYT-selected derived synthetic bread wheats for rainfed environments: Yield evaluation in Mexico and Australia. F. Crop. Res. 100:218–228. doi:10.1016/j.fcr.2006.07.005

Eisemann, R.L., M. Cooper, and D.R. Woodruff. 1990. Beyond the analytical methodology: Better interpretation of genotype by environment interaction. In: M.S. Kang, editor, Genotype by environment interaction and plant breeding. Louisiana State University, Baton Rouge, LA. p. 108–l 17.

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

Endelman, J. 2012. Using rrBLUP 4.0. http://www.afhalifax.ca/bete/revive-eponge/Rblkfct/vignette.pdf (accessed 28 Oct. 2015).

Endelman, J.B., G.N. Atlin, Y. Beyene, K. Semagn, X. Zhang, M.E. Sorrells, and J.L. Jannink. 2014. Optimal design of preliminary yield trials with genome-wide markers. Crop Sci. 54:48–59. doi:10.2135/cropsci2013.03.0154

Frutos, E., M.P. Galindo, and V. Leiva. 2014. An interactive biplot implementation in R for modeling genotype-by-environment interaction. Stoch. Environ. Res. Risk Assess. 28:1629–1641.

Gabriel, K.R. 1971. The biplot graphic display of matrices with application to principal component analysis. Biometrika 58:453–467. doi:10.1093/biomet/58.3.453

Gauch, H.G. 2006. Statistical analysis of yield trials by AMMI and GGE. Crop Sci. 46:1488–1500. doi:10.2135/cropsci2005.07-0193

Gauch, H.G. 2013. A simple protocol for AMMI analysis of yield trials. Crop Sci. 53:1860–1869. doi:10.2135/cropsci2013.04.0241

Gauch, H.G., H.P. Piepho, and P. Annicchiarico. 2008. Statistical analysis of yield trials by AMMI and GGE: Further considerations. Crop Sci. 48:866–889. doi:10.2135/cropsci2007.09.0513

Gauch, H.G., and R.W. Zobel. 1988. Predictive and postdictive success of statistical analyses of yield trials. Theor. Appl. Genet. 76:1–10. doi:10.1007/BF00288824

Glaubitz, J.C., T.M. Casstevens, F. Lu, J. Harriman, R.J. Elshire, Q. Sun, and E.S. Buckler. 2014. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. PLoS ONE 9:E90346. doi:10.1371/journal.pone.0090346

Gollob, H. 1968. A statistical model which combines features of factor analytic and analysis of variance techniques. Psychometrika 33:73–115. doi:10.1007/BF02289676

Gutiérrez, L., S. Germán, S. Pereyra, P.M. Hayes, C.a. Pérez, F. Capettini, A. Locatelli, N.M. Berberian, E.E. Falconi, R. Estrada, D. Fros, V. Gonza, H. Altamirano, J. Huerta-Espino, E. Neyra, G. Orjeda, S. Sandoval-Islas, R. Singh, K. Turkington, and A.J. Castro. 2015. Multi-environment multi-QTL association mapping identifies disease resistance QTL in barley germplasm from Latin America. Theor. Appl. Genet. 128:501–516. doi:10.1007/s00122-014-2448-y

Hayes, B.J., N.O.I. Cogan, L.W. Pembleton, M.E. Goddard, J. Wang, G.C. Spangenberg, and J.W. Forster. 2013. Prospects for genomic selection in forage plant species. Plant Breed. 132:133–143. doi:10.1111/pbr.12037

Heffner, E.L., J. Jannink, and M.E. Sorrells. 2011. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. Plant Gen. 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. 49:1–12. doi:10.2135/cropsci2008.08.0512

Henderson, C.R. 1984. Applications of linear models in animal breeding models. 3rd ed. University of Guelph, Guelph, ON.

Heslot, N., D. Akdemir, M.E. Sorrells, and J.L. Jannink. 2014. Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor. Appl. Genet. 127:463–480. doi:10.1007/s00122-013-2231-5

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

Isidro, J., J.L. Jannink, D. Akdemir, J. Poland, N. Heslot, and M.E. Sorrells. 2014. Training set optimization under population structure in genomic selection. Theor. Appl. Genet. 128:145–158.

Jarquín, D., J. Crossa, X. Lacaze, P. Du Cheyron, J. Daucourt, J. Lorgeou, F. Piraux, L. Guerreiro, P. Pérez, M. Calus, J. Burgueño, and G. de los Campos. 2014. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor. Appl. Genet. 127:595–607. doi:10.1007/s00122-013-2243-1

Kelly, A.M., B.R. Cullis, A.R. Gilmour, J.A. Eccleston, and R. Thompson. 2009. Estimation in a multiplicative mixed model involving a genetic relationship matrix. Genet. Sel. Evol. 41:33–42. doi:10.1186/1297-9686-41-33

Lado, B., I. Matus, A. Rodríguez, L. Inostroza, J. Poland, F. Belzile, A. del Pozo, M. Quincke, M. Castro, and J. von Zitzewitz. 2013. Increased genomic prediction accuracy in wheat breeding through spatial adjustment of field trial data. G3: Genes, Genomes, Genet. 3:2105–2114.

Lal, R.K. 2013. Adaptability patterns and stable cultivar selection in menthol mint (*Mentha arvensis* L.). Ind. Crops Prod. 50:176–181. doi:10.1016/j.indcrop.2013.07.008

Lal, R.K. 2014. Breeding for new chemotypes with stable high essential oil yield in *Ocimum*. Ind. Crops Prod. 59:41–49. doi:10.1016/j.indcrop.2014.04.047

Legarra, A., C. Robert-Granié, E. Manfredi, and J.M. Elsen. 2008. Performance of genomic selection in mice. Genetics 180:611–618. doi:10.1534/genetics.108.088575

Lopez-Cruz, M., J. Crossa, D. Bonnett, S. Dreisigacker, J. Poland, J.L. Jannink, R.P. Singh, E. Autrique, and G. de los Campos. 2015. Increased prediction accuracy in wheat breeding trials using a marker × environment interaction genomic selection model. G3: Genes, Genomes, Genet. 5:569–582.

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.

Lorenz, A., K. Smith, and J. Jannink. 2012. Potential and optimization of genomic selection for Fusarium head blight resistance in six-row barley. Crop Sci. 52:1609–1621. doi:10.2135/cropsci2011.09.0503

Lorenzana, R.E., and R. Bernardo. 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor. Appl. Genet. 120:151–161. doi:10.1007/s00122-009-1166-3

Ly, D., M. Hamblin, I. Rabbi, G. Melaku, M. Bakare, H.G. Gauch, R. Okechukwu, A.G.O. Dixon, P. Kulakow, and J.L. Jannink. 2013. Relatedness and genotype × environment interaction affect prediction accuracies in genomic selection: A study in cassava. Crop Sci. 53:1312–1325. doi:10.2135/cropsci2012.11.0653

Malosetti, M., J.M. Ribaut, and F.A. van Eeuwijk. 2013. The statistical analysis of multi-environment data: Modeling genotype-by-environment interaction and its genetic basis. Front. Physiol. 4:44.

Malosetti, M., J.M. Ribaut, M. Vargas, J. Crossa, and F.A. van Eeuwijk. 2007. A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). Euphytica 161:241–257. doi:10.1007/s10681-007-9594-0

Malosetti, M., J. Voltas, I. Romagosa, S.E. Ullrich, and F.A. van Eeuwijk. 2004. Mixed models including environmental covariables for studying QTL by environment interaction. Euphytica 137:139–145. doi:10.1023/B:EUPH.0000040511.46388.ef

Mandel, J. 1969. The partitioning of interaction in analysis of variance. J. Res. Natl. Bur. Stand., Sect. B 73B:309–328.

Mathews, K.L., M. Malosetti, S. Chapman, L. McIntyre, M. Reynolds, R. Shorter, and F.A. van Eeuwijk. 2008. Multi-environment QTL mixed models for drought stress adaptation in wheat. Theor. Appl. Genet. 117:1077–1091. doi:10.1007/s00122-008-0846-8

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

Meyer, K. 2009. Factor-analytic models for genotype × environment type problems and structured covariance matrices. Genet. Sel. Evol. 41:21–32. doi:10.1186/1297-9686-41-21

Patterson, H., and R. Thompson. 1971. Recovery of inter-block information when block sizes are unequal. Biometrika 58:545–554. doi:10.1093/biomet/58.3.545

Piepho, H.P. 1997. Analyzing genotype-environment data by mixed models with multiplicative terms. Biometrics 53:761–766. doi:10.2307/2533976

Piepho, H.P. 1998. Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. Theor. Appl. Genet. 97:195–201. doi:10.1007/s001220050885

Piepho, H.P. 2000. A mixed-model approach to mapping quantitative trait loci in Barley on the basis of multiple environment data. Genetics 156:2043–2050.

Piepho, H.P. 2005. Statistical tests for QTL and QTL-by-environment effects in segregating populations derived from line crosses. Theor. Appl. Genet. 110:561–566. doi:10.1007/s00122-004-1872-9

Piepho, H.P., and J. Möhring. 2006. Selection in cultivar trials: Is it ignorable? Crop Sci. 46:192–201. doi:10.2135/cropsci2005.04-0038

Piepho, H.P., and J. Möhring. 2007. Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888. doi:10.1534/genetics.107.074229

Piepho, H.P., J. Möhring, A.E. Melchinger, and A. Büchse. 2007. BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161:209–228. doi:10.1007/s10681-007-9449-8

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and R Core Team. 2013. nlme: Linear and nonlinear mixed effects models. R package version 3.1-122. http://CRAN.R-project.org/package=nlme (accessed 1 Feb. 2015).

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, S. Dreisigacker, J. Crossa, H. Sánchez-Villeda, M. Sorrells, and J.L. Jannink. 2012b. Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Gen. 5:103–113. doi:10.3835/plantgenome2012.06.0006

R Development Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rebonato, R., and P. Jäckel. 1999. The most general methodology to create a valid correlation matrix for risk management and option pricing purposes. Appl. Econ. Lett. 17:1767–1768.

Resende, M.F.R., P. Muñoz, J.J. Acosta, G.F. Peter, J.M. Davis, D. Grattapaglia, M.D.V. Resende, and M. Kirst. 2011. Accelerating the domestication of trees using genomic selection: Accuracy of prediction models across ages and environments. New Phytol. 193:617–624. doi:10.1111/j.1469-8137.2011.03895.x

Riedelsheimer, C., and A.E. Melchinger. 2013. Optimizing the allocation of resources for genomic selection in one breeding cycle. Theor. Appl. Genet. 126:2835–2848. doi:10.1007/s00122-013-2175-9

Rodrigues, P.C., M. Malosetti, H.G. Gauch, and F.A. van Eeuwijk. 2014. A weighted AMMI algorithm to study genotype-by-environment interaction and QTL-by-environment interaction. Crop Sci. 54:1555–1570. doi:10.2135/cropsci2013.07.0462

Rutkoski, J., R.P. Singh, J. Huerta-Espino, S. Bhavani, J. Poland, J.L. Jannink, and M.E. Sorrells. 2015. Efficient use of historical data for genomic selection: A case study of stem rust resistance in wheat. Plant Gen. 8:1–10. doi:10.3835/plantgenome2014.09.0046

SAS Institute. 2005. JMP user guide. Release 6. SAS Institute Inc., Cary, NC.

Solberg, T.R., A.K. Sonesson, J.A, Woolliams, and T.H.E. Meuwissen. 2009. Reducing dimensionality for prediction of genome-wide breeding values. Genet. Sel. Evol. 41:29–37. doi:10.1186/1297-9686-41-29

Storlie, E., and G. Charmet. 2013. Genomic selection accuracy using historical data generated in a wheat breeding program. Plant Gen. 6:1–9. doi:10.3835/plantgenome2013.01.0001

van Eeuwijk, F.A. 1995. Linear and bilinear models for the analysis of multi-environment trials: I. An inventory of models. Euphytica 84:1–7. doi:10.1007/BF01677551

van Eeuwijk, F.A., M. Malosetti, X. Yin, P.C. Struik, and P. Stam. 2005. Statistical models for genotype by environment data: From conventional ANOVA models to eco-physiological QTL models. Aust. J. Agric. Res. 56:883–894. doi:10.1071/AR05153

Wang, H., and P. Schmidt. 2002. One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. J. Prod. Anal. 18:129–144. doi:10.1023/A:1016565719882

Windhausen, V.S., S. Wagener, C. Magorokosho, D. Makumbi, B. Vivek, H.P. Piepho, A.E. Melchinger, and G.N. Atlin. 2012. Strategies to subdivide a target population of environments: Results from the CIMMYT-Led maize hybrid testing programs in Africa. Crop Sci. 52:2143–2152. doi:10.2135/cropsci2012.02.0125

Yan, W., L. Hunt, Q. Sheng, and Z. Szlavnics. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. Crop Sci. 40:597–605. doi:10.2135/cropsci2000.403597x

Yan, W., M.S. Kang, B. Ma, S. Woods, and P.L. Cornelius. 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. Crop Sci. 47:643–655. doi:10.2135/cropsci2006.06.0374

Yan, W., and N. Tinker. 2006. Biplot analysis of multi-environment trial data: Principles and applications. Can. J. Plant Sci. 6:623–645. doi:10.4141/P05-169

Yang, R.C., J. Crossa, P.L. Cornelius, and J. Burgueño. 2009. Biplot analysis of genotype × environment interaction: Proceed with caution. Crop Sci. 49:1564–1576. doi:10.2135/cropsci2008.11.0665

Yu, J., G. Pressoir, W.H. Briggs, I. Vroh Bi, M. Yamasaki, J.F. Doebley, M.D. McMullen, B.S. Gaut, D.M. Nielsen, J.B. Holland, S. Kresovich, and E.S. Buckler. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38:203–208. doi:10.1038/ng1702

Zobel, R.W., M.J. Wright, and H.G. Gauch, Jr. 1988. Statistical analysis of a yield trial. Agron. J. 80:388–393. doi:10.2134/agronj1988.00021962008000030002x