University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Agronomy & Horticulture -- Faculty Publications | Agronomy and Horticulture Department

2017

Genomic-enabled Prediction Accuracies Increased by Modeling Genotype × Environment Interaction in Durum Wheat

Sivakumar Sukumaran
International maize and Wheat Improvement Center

Diego Jarquin
University of Nebraska-Lincoln, jhernandezjarquin2@unl.edu

José Crossa
International Maize and Wheat improvement Center (CIMMYT), j.crossa@cgiar.org

Matthew Reynolds
International Maize and Wheat Improvement Center

Follow this and additional works at: http://digitalcommons.unl.edu/agronomyfacpub

 Part of the Agricultural Science Commons, Agriculture Commons, Agronomy and Crop Sciences Commons, Botany Commons, Horticulture Commons, Other Plant Sciences Commons, and the Plant Biology Commons

Sukumaran, Sivakumar; Jarquin, Diego; Crossa, José; and Reynolds, Matthew, "Genomic-enabled Prediction Accuracies Increased by Modeling Genotype × Environment Interaction in Durum Wheat" (2017). *Agronomy & Horticulture -- Faculty Publications*. 1197.
http://digitalcommons.unl.edu/agronomyfacpub/1197

This Article is brought to you for free and open access by the Agronomy and Horticulture Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Agronomy & Horticulture -- Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published April 26, 2018

ORIGINAL RESEARCH

Genomic-enabled Prediction Accuracies Increased by Modeling Genotype × Environment Interaction in Durum Wheat

Sivakumar Sukumaran, Diego Jarquin,* Jose Crossa,* Matthew Reynolds

ABSTRACT

Genomic prediction studies incorporating genotype × environment (G×E) interaction effects are limited in durum wheat. We tested the genomic-enabled prediction accuracy (PA) of Genomic Best Linear Unbiased Predictor (GBLUP) models—six non-G × E and three G × E models—on three basic cross-validation (CV) schemes— in predicting incomplete field trials (CV2), new lines (CV1), and lines in untested environments (CV0)— in a durum wheat panel grown under yield potential, drought stress, and heat stress conditions. For CV0, three scenarios were considered: (i) leave-one environment out (CV0-Env); (ii) leave one site out (CV0-Site); and (iii) leave 1 yr out (CV0-Year). The reaction norm models with G × E effects showed higher PA than the non-G × E models. Among the CV schemes, CV2 and CV0-Env had higher PA (0.58 each) than the CV1 scheme (0.35). When the average of all the models and CV schemes were considered, among the eight traits— grain yield, thousand grain weight, grain number, days to anthesis, days to maturity, plant height, and normalized difference vegetation index at vegetative (NDVIvg) and grain filling (NDVIIIg)—, plant height had the highest PA (0.68) and moderate values were observed for grain yield (0.34). The results indicated that genomic selection models incorporating G × E interaction show great promise for forward prediction and application in durum wheat breeding to increase genetic gains.

Core Ideas

- Genomic-enabled prediction accuracy of G×E models was superior to the non-G×E models.
- Forward prediction and sparse testing in durum wheat shows great promise.
- Genomic-enabled prediction accuracy of yield and components traits were highly associated with heritability.

In past years, the breeding technology referred to as 'genomic selection' (GS) (Meuwissen et al., 2001) has been implemented in plant breeding where several species of economic importance including wheat have been shown increased genomic-enabled prediction accuracy (PA) for several traits (Crossa et al., 2017). Genomic selection uses dense molecular markers to predict the breeding value of

S. Sukumaran, M. Reynolds, Global Wheat Program, International Maize and Wheat Improvement Center, Apdo. Postal 6-641, 06600 Mexico City; J. Crossa, Biometrics and Statistics Unit, Genetic Resources Program, Apdo. Postal 6-641, 06600 Mexico City; D. Jarquin, Dep. of Agronomy and Horticulture, Univ. of Nebraska, Lincoln, NE 68583. Received 10 Dec. 2017. Accepted 9 Feb. 2018. *Corresponding authors (diego.jarquin@gmail.com, j.crossa@cgiar.org).

Abbreviations: CV0, cross-validation predicting the performance of previously tested lines in untested locations; CV1, cross-validation evaluating the performance of lines that have not been evaluated in any of the observed environments; CV2, cross-validation evaluating the performance of lines that have been evaluated in some environments but not in others; DT, drought stress; DTH, days to heading; DTM, days to maturity; GBLUP, Genomic best linear unbiased predictor; G×E, genotype × environment; GNO, grain number; GS, genomic selection; HT, heat stress; M1 to M9, Model 1 to Model 9; NDVI, normalized difference vegetative index; NDVIvg, normalized difference vegetative index – vegetative; NDVI11g, normalized difference vegetative index – grain filling; PA, prediction accuracy; PH, plant height; TGW, thousand-grain weight; TRN, training population; TST, testing population; YLD, grain yield; YP, yield potential.

Plant Genome 11:170112
doi: 10.3835/plantgenome2017.12.0112

© Crop Science Society of America
5585 Guilford Rd., Madison, WI 53711 USA
This is an open access article distributed under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

individuals that have been genotyped but not phenotyped (testing population, TST) by means of a population that has both genotypic and phenotypic data (training population, TRN), and by fitting a statistical model which is used to predict breeding values of the non-phenotyped selection candidates. Performances for various complex traits of the TST population are predicted using allelic identity with loci that were found to be associated with the phenotype in the TRN. It is necessary to intensively phenotype and genotype diverse lines from a breeding program to provide potential TRN and TST sets to robustly and precisely calibrate genomic prediction models (Crossa et al., 2017).

Wheat breeding researchers have significantly improved PA over pedigree breeding for several economically important traits such as grain yield, maturity, and grain quality (de los Campos et al., 2009; Crossa et al., 2010; Pérez-Rodríguez et al., 2012). These studies have used random cross-validation of data sets comprising individuals being phenotyped and genotyped to mimic what breeders will face when performing GS-assisted breeding. These empirical results obtained by random cross-validation suggest that GS can increase genetic gains by shortening the breeding cycle and/or enhancing testing efficiency in field evaluations. In general, results of using random cross-validation on genomic wheat breeding data based on the standard Genomic Best Linear Unbiased Predictor (GBLUP) indicate that GS can significantly increase prediction accuracy related to pedigree and marker-assisted selection for low heritability traits. Nevertheless, these initial empirical results were obtained using single environments and thus do not exploit information across environments.

Standard GBLUP models were extended to multi-environment settings. Burgueño et al. (2012) used a multi-environment version of the GBLUP where G × E for grain yield in bread wheat was modeled using genetic correlations; the authors found that the multi-environment GBLUP had a higher PA than the single-environment GBLUP. It should be pointed out that Burgueño et al. (2012) did not attempt to incorporate environmental variables as surrogates for environments. Jarquín et al. (2014) proposed an extension of the GBLUP G × E random effects models where the main effects of genomic (markers) and environmental covariables, as well as their first order interactions (marker ×environmental covariates), are introduced using covariance structures that are functions of marker genotypes and environmental covariables. The studies by Burgueño et al. (2012) and Jarquín et al. (2014) employed grain yield of bread wheat data and applied random cross-validation to assess two prediction problems: (1) the performance of lines that have been evaluated in some environments, but not in others (cross-validation 2, CV2) and (2) the performance of lines that have not been evaluated in any of the observed environments (cross-validation 1, CV1). However, other prediction problems that do not involve random cross-validation are considered in what we call CV0: (1) predicting an environment (i.e., site-year combination) that was not included in the usual set of testing environments in the evaluation system (leave-one-environment-out); and (2) predicting a year using information from previous years (forward prediction). These prediction problems were recently studied by Jarquín et al. (2017) in bread wheat lines evaluated in the Kansas State University Hard Red Winter Wheat Breeding Program for different sites and years. Results of Jarquín et al. (2017) showed that the GBLUP G × E models had relatively high prediction accuracy values (0.4) when predicting the yield performance in untested environments and also high prediction ability (0.54) when predicting yield in incomplete field trials for sites with a moderate number of lines (sparse testing).

The GBLUP G × E model (Jarquín et al., 2014) can also be applied with pedigree data where the numerical relationship matrix ($\boldsymbol{A}$) is derived from pedigree relationship information (Pérez-Rodríguez et al., 2015). Recently, in bread wheat, the GBLUP G × E based on $\boldsymbol{A}$ information was applied to large-scale screening of international nurseries focusing on the Wheat Yield Consortium Yield Trial and the Stress Adapted Trait Yield Nursery of CIMMYT, which were both grown in major spring wheat production areas worldwide (Sukumaran et al., 2017b). The authors showed that higher predictive ability was achieved by modeling G × E using the pedigree information given in the numerical relationship matrix $\boldsymbol{A}$. Furthermore, the GBLUP G × E model can be used to include both genomic and pedigree information and thereby increase genomic-enabled and pedigree-enabled prediction accuracy. In bread wheat breeding, Sukumaran et al. (2017b) showed that the genomic prediction models with interaction terms due to genomic (G) × environment and pedigree ($\boldsymbol{A}$) × E were the best models for grain yield prediction.

A recent study by Pérez-Rodríguez et al. (2017) showed a method for combining genomic and pedigree information in a single-step model and assessing the PA of a large number of bread wheat lines (58,798), spanning years and evaluated in several environments, for predicting grain yield performance in several South Asian sites (in India, Pakistan, and Bangladesh), using the GBLUP G × E model of Jarquín et al. (2014). The results indicated that PA achieved by models using only pedigree information, only genomic information, or both pedigree and genomic information to predict environments in India, Pakistan, and Bangladesh is higher (0.25–0.38) than prediction accuracy using only phenotypic correlations (0.20). The results of this study indicated that the single-step approach combining pedigree and marker information is useful for reducing genotyping costs while maintaining the prediction accuracy of unobserved individuals at relatively intermediate levels.

Most of the previously mentioned genomic-enabled prediction results refer to the complex trait—grain yield—measured in bread wheat. However, genomic-enabled results applied to durum wheat have not been very abundant. A recent study on 1184 lines from the North Dakota State University durum wheat program was done with the main objectives of identifying QTL to be used in marker-assisted selection and also for studying genomic-enabled predictions on quality traits (e.g., test weight, semolina

color, and gluten, etc.) (Fiedler et al., 2017). The authors found that PA for quality traits ranged from 0.20 to 0.66. Although the durum wheat lines were evaluated in several site-year combinations, no genomic-enabled prediction incorporating genomic × environment interaction was assessed. A genomic-enabled prediction model using durum wheat lines was proposed by Crossa et al. (2016); the model is based on Bayes B model and can be used for genomic prediction under penalized regression or as selection variables (marker selection).

Based on the above considerations and the fact that PA studies in durum wheat are scarce, we conducted a genomic-enabled prediction study using the GBLUP G × E model of Jarquín et al. (2014), on durum wheat lines evaluated in yield potential, drought, and heat stress environments over 2 yr (cycles 2014–15 and 2015–16), with the main objective of examining the PA of models for complex traits (e.g., grain yield), as well as less complex traits (e.g., days to heading and days to maturity). Another objective of this research was to study several prediction problems: (1) random cross-validations CV1 and CV2, and (2) cross-validation CV0 for leaving-one-environment-out, leaving-one-site-out, and the prediction of future years (forward prediction).

MATERIALS AND METHODS

Germplasm

We used a durum panel (*Triticum turgidum* subsp. *durum*) that consisted of 208 entries, which was a subset of the 15,000-durum accessions characterized from CIMMYT's gene bank. These accessions were screened for visual biomass, grain yield, flowering time, and plant height and a subset of 208 lines were developed which were closer to the durum checks in agronomic performance. The panel also consisted of lines from CIMMYT's International Wheat Improvement Network (IWIN) nurseries (http://www.cimmyt.org/international-wheat-improvement-network-iwin/): 2IDYN, 3IDYN, 15IDYN, 33EDUYT, 34IDSN, and 24EDYT-SA. These lines originated from different countries—Chile, Ethiopia, Ecuador, Lebanon, Iran, Mexico, and Syria— as per International Wheat Information System records. The present panel was also used for a genome-wide association study using DArTseq markers and QTL hotspots were identified for agronomic traits under yield potential, drought stress, and heat stress conditions (Sukumaran et al., 2018).

Phenotypic Experimental Data

Phenotyping was conducted at the Campo Experimental Norman E Borlaug, CIMMYT's main research station at Cd. Obregon, Sonora, Mexico, under yield potential (well-watered and high radiation, YP), drought stress (DT), and heat stress (HT) environments. These conditions were achieved by changing the planting date and irrigation schedule (Table 1). The panel was grown in 2m plots with 0.75cm between the rows, in a raised bed system. The diseases and pests prevalent in the region were controlled by relevant fungicide and pesticide applications as needed. The following traits were measured; grain yield (YLD), thousand-grain weight (TGW), grain number m^{-2} (GNO), days to anthesis (DTA under YP but days to heading - DTH- under DT and HT), plant height (PH), days to maturity (DTM), and normalized difference vegetative index (NDVI) at vegetative (NDVIvg) and grain filling (NDVIllg) stages for 2 yr (2014–15 and 2015–16) following established protocols (Pask et al., 2012). More details about the phenotypic measurements are described in an earlier publication (Sukumaran et al., 2018).

Table 1. Information about the durum panel grown under yield potential (YP), drought stress (DT), and heat stress (HT) conditions and weather parameters (Sukumaran et al., 2018).†

Year	Env.	Planting date	Harvest date	T_{mean}	T_{Range}	Prec.	Irrig.	T_{max} > 35
2014–15	YP	28-Nov-14	22-May-15	19.4	11.8–28.2	84.2	3	0
	DT	09-Dec-14	07-May-15	19.4	11.8–28.1	84.0	0	0
	HT	20-Mar-15	14-Jul-15	25.9	18.0–34.2	61.2	6	47
2015–16	YP	16-Dec-15	17-May-16	16.1	08.1–26.2	20.4	3	0
	DT	02-Dec-15	13-Apr-16	16.2	08.2–26.4	20.2	0	0
	HT	26-Feb-16	14-Jun-16	22.4	13.4–32.1	17.4	6	26

† Env. = environments; T_{mean} = Mean temperature during the crop cycle; T_{Range} = mean minimum and maximum temperatures; Prec. = precipitation; Irrig. = No. of irrigations; T_{max} > 35, Number of days when temperatures were above 35°C.

Genotypic Experimental Data

Sukumaran et al. (2018) detailed about the genotyping of the panel. In short, we collected fresh leaves from each line and a modified cetyltrimethylammonium bromide method was used for DNA extraction (Saghai-Maroof et al., 1984). Electrophoresis in 1% agarose gel was used to determine the DNA quality and concentration. High-throughput genotyping was conducted using DArTseqTM technology (Sansaloni et al., 2011) at the facility in CIMMYT, Mexico; Genetic Analysis Service for Agriculture. The genomic DNA was digested with a combination of two restriction enzymes, *PstI* (CTGCAG) and *HpaII* (CCGG) and a genomic representation of the samples was generated by ligating barcoded adapters to identify each sample to run within a single lane of an Illumina HiSeq2500 instrument (Illumina Inc., San Diego, CA). Approximately 500,000 unique reads per sample were generated by sequencing up to 77 bases of the amplified fragments. We used a proprietary analytical pipeline—developed by DArT P/L—to generate SNPs.

Data Availability

The phenotypic and genetic data are available at http://hdl.handle.net/11529/11053. A consensus map from Diversity arrays was used for the present study and physical positions of the markers are available at the link above.

Phenotypic Data Analysis

We used META-R software to estimate the variance components—analysis of variance— and to predict the

Best Linear Unbiased Predictions (BLUPs) (Alvarado et al., 2015). Lines (L), environments (E) and $L \times E$ interaction, were considered as random factors, while location, block, and replication were considered as fixed factors to estimate the BLUPs. In addition, BLUPs for YLD, TGW, and GNO were calculated using DTH as a co-variate. The following formula was used to estimate broad-sense repeatability (H^2):

$$H^2 = \frac{\sigma_L^2}{\sigma_L^2 + \sigma_{LE}^2 / I + \sigma_e^2 / rI}$$

where σ_L^2 is the variance for the line effect, σ_{LE}^2 is the line by environment interaction variance, σ_e^2 is error variance, r is the number of replications and I is the number of environments.

Genomic-Enabled Prediction Models

Baseline Model

The following linear predictor represents the response of the j^{th} ($j = 1,\ldots,J$) line tested in the i^{th} ($i = 1,\ldots,I$) environment $\{y_{ij}\}$ as the sum of an overall mean μ plus random deviations around zero due to environmental $\left[E_i \overset{iid}{\sim} N\left(0,\sigma_E^2\right)\right]$ and line effects $\left[L_j \overset{iid}{\sim} N\left(0,\sigma_L^2\right)\right]$, the interaction between the i^{th} environonment and the j^{th} line $\left[LE_{ij} \overset{iid}{\sim} N(0,\sigma_{LE}^2)\right]$ and a random error term $\left[e_{ij} \overset{iid}{\sim} N\left(0,\sigma_e^2\right)\right]$.

$$y_{ij} = \mu + E_i + L_j + LE_{ij} + e_{ij}$$

where $N(.,.)$ denotes the normal density, *iid* stands for independent and identically distributed responses and σ_E^2, σ_L^2, σ_{LE}^2, σ_e^2 are the corresponding variances for environment, line, line × environment and residual terms. This model does not allow borrowing of information among lines because they were treated as independent outcomes. The models used in this study were derived from the baseline model by either subtracting terms or modifying the underlying assumptions.

Main Effects Models

Model 1 (M1). Environment + Line Main Effects (L + E)

This model is obtained by retaining the first three components from the baseline model while their underlying assumptions remain unchanged.

$$y_{ij} = \mu + E_i + L_j + e_{ij} \qquad [1]$$

Here environments were considered as site-by-year combinations.

Model 2 (M2).Site + Line Main Effects (L + S)

The previous model considers environments as independent outcomes; thus, borrowing information among environments is not possible. In an attempt to recover information from the same site observed in different years, a model that includes this site effect was considered.

$$y_{kj} = \mu + S_k + L_j + e_{kj} \qquad [2]$$

Here y_{kj} represents the response of the j^{th} line observed in the k^{th} site ($k = 1,\ldots,K$) with $S_k \overset{iid}{\sim} N\left(0,\sigma_S^2\right)$ and σ_S^2 acting as its correspondent variance component. This model intends to recover information from the same site but observed in different years by ignoring the year effect.

Model 3 (M3).Year + Line Main Effects (L + Y)

This model ignores the site effect by assuming no soil/environmental changes between sites in the same year but across years. Thus, all sites in the same year are treated as the same and the only variation affecting the responses (besides the lines) is the one that occurs from one yr to another.

$$y_{lj} = \mu + Y_l + L_j + e_{kj} \qquad [3]$$

where y_{lj} denotes the response of the the j^{th} line observed in the l^{th} year ($l = 1,\ldots,Y$) with $Y_l \overset{iid}{\sim} N\left(0,\sigma_l^2\right)$ and σ_l^2 its variance component. This model intends to borrow information between sites observed in the same year.

Model 4 (M4). Environment + Marker Main Effects (L + G + E)

Considering an alternative representation of the line effect L_j in Model [1] as a linear combination between markers and their correspondent marker effects, $g_j = \sum_{m=1}^{p} x_{jm} b_m$, genomic information can be introduced using the following linear predictor

$$y_{ij} = \mu + E_i + L_j + g_j + e_{ij} \qquad [4]$$

where $b_m \overset{iid}{\sim} N\left(0,\sigma_b^2\right)$ represents the random effect of the m^{th} ($m = 1,\ldots,p$) marker and σ_b^2 its correspondent variance component. Using the results from the multivariate normal distribution, $\mathbf{g} = (g_1,\ldots,g_J)'$, the vector of genetic effects, follows a normal density with zero mean vector and co-variance matrix $Cov(\mathbf{g}) = \mathbf{G}\sigma_g^2$ with $\mathbf{G} = \frac{\mathbf{XX}'}{p}$ as the genomic relationship matrix. It describes genetic similarities among pairs of individuals. Here $\mathbf{X}$ represents the centered and standardized (by columns) genomic matrix and $\sigma_g^2 = p \times \sigma_b^2$ acts as the correspondent variance component such that $\mathbf{g} = \{g_j\} \sim N\left(0,\mathbf{G}\sigma_g^2\right)$. In this model, the line effect L_j is retained in the model to account for imperfect information and model mis-specification due to imperfect linkage disequilibrium.

Model 5 (M5). Site + Marker Main Effects (L + S + G)

This model was built using M2 as the initial starting point but adding a genetic component (marker information), as shown in the previous model. Thus, M5 becomes

$$y_{kj} = \mu + S_k + L_j + g_j + e_{kj} \quad [5]$$

Model 6 (M6). Year + Marker Main Effects (L + Y + G)

Similar to the previous two models, M6 adds molecular marker information to model M3, so that it becomes

$$y_{lj} = \mu + Y_l + L_j + g_j + e_{kj} \quad [6]$$

Models with Interaction

A disadvantage of the previous models is that they only consider the main effect of the lines across environments, thereby avoiding specific responses of each line in each environment and these relies predictive ability exclusively on genetics. In this case, each line would show the same genetic values across environments. To overcome this issue, the informed (at the genetic level) G × E interaction is introduced via co-variance structures, as shown by Jarquín et al. (2014). Here the naïve interaction component EL_{ij} is replaced by gE_{ij}, where $\boldsymbol{gE} = \{gE_{ij}\} \sim N\left(0, (\boldsymbol{Z_g G Z'_g}) \circ (\boldsymbol{Z_E Z'_E})\boldsymbol{\sigma^2_{gE}}\right)$ and $\boldsymbol{Z_g}$ and $\boldsymbol{Z_E}$ are the correspondent incidence matrices for molecular markers and environments, σ^2_{gE} is the associated variance component for this interaction and '∘' represents the Hadamard or Schur product (element-to-element product) between two matrices. Conceptually, this component allows the inclusion of all first order interactions between each marker and each observed environment.

Model 7 (M7). Genomic × Environment Interaction [L + E + G + (G × E)]

This model extends model M4 by adding the previously introduced interaction term as follows:

$$y_{ij} = \mu + E_i + L_j + g_j + gE_{ij} + e_{ij} \quad [7]$$

Model 8 (M8). Genomic × Site Interaction; L + S + G + (G × S)

Similar to model M7, this model extends M5 by adding the interaction term between genotypes and sites. Here the interaction between each molecular marker and each site is included via $\boldsymbol{gS} = \{gS_{kj}\} \sim N\left(0, (\boldsymbol{Z_g G Z'_g}) \circ (\boldsymbol{Z_S Z'_S})\boldsymbol{\sigma^2_{gS}}\right)$, where $\boldsymbol{Z_S}$ and σ^2_{gS} are the incidence matrix of the sites and the associated variance component, respectively.

$$y_{kj} = \mu + S_k + L_j + g_j + gS_{kj} + e_{kj} \quad [8]$$

Model 9 (M9). Genomic × Year Interaction [L + Y + G + (G × Y)]

This model is an extension of model M6 that includes the interaction between each molecular marker and each year. For this, the vector $\boldsymbol{gY} = \{gY_{lj}\} \sim N\left(0, (\boldsymbol{Z_g G Z'_g}) \circ (\boldsymbol{Z_Y Z'_Y})\boldsymbol{\sigma^2_{gY}}\right)$ was added.

$$y_{lj} = \mu + Y_l + L_j + g_j + gY_{lj} + e_{kj} \quad [9]$$

where $\boldsymbol{Z_Y}$ represents the incidence matrix for years and $\sigma^2_{\boldsymbol{Yg}}$ is the corresponding variance component associated with this random effect.

Assessing Different Prediction Problems using Various Cross-Validation Strategies

Nine GBLUP models were used to compare three basic cross-validation schemes that mimic real prediction problems that breeders might face in the field. These problems are presented here: starting from the easiest to the most difficult one: (1) incomplete field trials (CV2) where some lines are observed in some environments but not in others; the goal here is to predict the crop performance of these lines in environments where these have not yet been observed, (2) prediction of newly developed lines (CV1) in an attempt to measure the predictive ability of new lines that have not yet been observed in any field, predictive ability between observed and unobserved lines is based primarily on genetic similarities as main source of information, and (3) predicting already observed lines in unobserved environments (CV0). Here, the main interest is to predict the crop performance of lines in potentially new environments. The latter cross-validation scheme gives an idea of the stability of the lines across a diverse set of environmental conditions. Three scenarios were considered for CV0 depending on whether site, year, or site-year combination (environment) are considered for prediction: (i) leaving one environment out (CV0-Env); (ii) leaving one site out (CV0-Sites); and (iii) leaving 1 year out (CV0-Year). For the three different CV0 no random cross-validations is performed and the observed values in one site or year or environments are directly correlated with the predicted values on those trials (e.g., environment, site, year).

For random cross-validation CV1 and CV2, the prediction accuracies of the nine models were calculated by performing random fivefold cross-validation where 20% of the durum wheat (testing set) were predicted and 80% were observed and used as training set. For CV1 none of the 20% of the lines in the testing set were observed in any of the environments (site and year combination), whereas for CV2 the 20% of the lines in the testing set were observed in some environments but not in the others. The prediction accuracy is computed as the correlations between the observed and predicted values within same environments.

RESULTS

Agronomic Performance of the Durum Panel

The highest grain yield was observed under YP (5.79 t/ha; $H^2 = 0.80$), followed by DT (2.33 t/ha; $H^2 = 0.47$), and HT (1.64 t/ha; $H^2 = 0.30$). TGW observed also varied among the different environments, under YP (44.4 g; $H^2 = 0.87$), DT (40.8 g; $H^2 = 0.69$), and HT (31.8 g; $H^2 = 0.63$). The same trend—yield followed by DT and HT—was observed for all traits. The DTA and DTM also followed a similar pattern, where the shortest duration

Table 2. Descriptive statistics and repeatability (H^2) estimated through the best linear unbiased predictions (BLUPs) of a durum panel grown under yield potential, drought, and heat stress conditions in 2014–15 and 2015–16 (Sukumaran et al., 2018).

	Yield potential			Drought stress			Heat stress		
Traits†	Mean	Range	H^2	Mean	Range	H^2	Mean	Range	H^2
YLD	5.79	2.67–7.42	0.80	2.33	1.16–3.54	0.47	1.64	0.2–2.56	0.30
TGW	44.43	31.98–57.24	0.87	40.8	31.1–51.1	0.69	31.8	24.08–40.65	0.63
GNO	12999	2090–18,379	0.79	8914	4318–12,174	0.22	4621	811–7574	0.41
DTA	76	67–102	0.94	71	61–79	0.71	55.5	47–81.2	0.86
DTM	113	104–144	0.90	100	95–108	0.81	84.1	75.5–101.2	0.78
PH	96.8	81.8–134.2	0.95	68.47	56.5–97.1	0.83	54.2	40.1–77.3	0.66
NDVIvg	0.41	0.31–0.48	0.30	0.38	0.30–0.44	0.72	0.32	0.22–0.44	0.58
NDVIllg	0.56	0.44–0.65	0.37	0.39	0.31–0.57	0.82	–	–	–

† YLD = grain yield (t/ha); TGW = thousand-grain weight (g); GNO = grain number/m^2; DTA = days to anthesis; DTM = days to maturity; PH = plant height (cm); normalized difference vegetative index at vegetative (NDVIvg) and grain filling (NDVIllg) stages.

of the crop was in HT (DTA = 55.5 d; DTM = 84.1 d) followed by to DT (DTA = 71 d; DTM = 100 d) and YP (DTA = 76 d; DTM = 113 d). Under YP conditions, High H^2 values were observed for all traits (0.79 to 0.95) except for NDVIvg and NDVIllg (H^2 = < 0.40). However, under the DT and HT stress conditions, NDVI values had moderate to high (> 0.58) H^2 values (Table 2).

The traits which showed very high correlation under drought (r = 0.72) and heat (r = 0.94) were GNO and YLD, and DTA and DTM were the traits with the highest correlation (r = 0.91) under YP condition. The TGW and GNO were negatively associated (r = -0.50) in YP and DT, but showed significantly less association (r = -0.06) under HT. The association of YLD and TGW was the highest under HT (r = 0.24), followed by YP (r = 0.14) and DT (r = 0.12). The YLD was negatively associated with DTA (r = -0.35), DTM (r = -0.26), and PH (r = -0.34) under YP conditions, but the effects were not significant under DT and HT. Under HT, PH was positively associated with YLD (r = 0.44). The NDVIvg was significantly associated with PH (r = 0.65) under HT (data not shown).

Genomic Prediction

We used the GBLUP G × E model to predict the lines using five different prediction problems methods: (1) predicting years (CV0-Year); (2) predicting sites (CV0-Sites); (3) predicting environments (CV0-Env); (4) predicting lines untested in the environment (CV1); and (5) predicting lines in incomplete trials (CV2) (sparse testing). The results of the correlations between the predicted values and observed values for each environment are shown in the Supplementary Tables: CV0-Year (Supplementary Table 1), CV0-Sites (Supplementary Table 2), CV0-Env (Supplementary Table 3), CV1 (Supplementary Table 4), and CV2 (Supplementary Table 5).

When years were predicted using the information from other year (CV0-Year; predict 2015 from 2016, or vice versa), high prediction accuracies were observed between the predicted values and observed values in each environment for each trait (Supplementary Table 1). The average correlations between the predicted and observed values for CV0-Year from six environments indicated high correlations for models that included the G × E term (Table 3). The trait with the highest average PA in six environments was PH (0.750) and the lowest was YLD (0.362). The TGW, DTA, and DTM had PA > 0.60, whereas the PA for GNO, NDVIllg, and NDVIvg were 0.37 to 0.43. Among the different models, M8, which had the G × S term, was the best model for four out of eight traits. For each trait, the models with the interaction term were the best models: M7 (PH and TGW), M8 (DTA, DTM, GNO, NDVIvg, and YLD), and M9 (NDVIllg).

Predicting the sites (CV0-Sites) based on all other sites—choosing one site among YP, DT, and HT and predicting a site using the other two sites—also showed medium-to-high prediction accuracies (Supplementary Table 2). The highest PA across six environments was for PH (0.755) and lowest was for YLD (0.300) (Table 4). The best models for DTM (0.651) and TGW (0.568) showed high PA, while NDVIvg (0.476), NDVIllg (0.444), and GNO (0.361) were moderate. For three traits, the models with the interaction term were the best: M8 (DTA) and M9 (GNO and YLD).

In the prediction scheme CV0-Env, each environment was predicted based on all other environments and the correlations were high between the observed and predicted values in each environment (Supplementary Table 3). When the average of six environments was taken, the PA was moderate to high, with PH being the highest (0.775) and YLD being the lowest (0.400) (Table 5). All other traits had a PA > 0.41 for the best model for each trait: DTA (0.724), DTM (0.696), GNO (0.419), NDVIllg (0.482), NDVIvg (0.515), and TGW (0.625). Six out of eight traits had high PA when models with the interaction term were used, M7 (NDVIvg and PH) and M8 (DTA, DTM, TGW, and YLD).

When the CV1 scheme was tested for each environment, the mean and standard deviations of the correlation between predicted and observed values were estimated (Supplementary Table 4). The average correlations between predicted and observed values in six environments for each trait are shown in Table 6. The highest average PA was observed for DTA (0.41) and lowest was

Table 3. Average correlations between predicted and observed values for the traits in cross-validation scenario (CV0-Year) where 1 year was left out (2015 or 2016) and the other year was used to predict the year that was left out (2015 or 2016). The values correspond to the average correlations between predicted and observed values of six environments (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015) for each trait. The best model for each trait is underlined. For details of each environment, see Supplementary Table 1.

Traits/Models	M1†	M2	M3	M4	M5	M6	M7	M8	M9
DTA	0.631	0.631	0.628	0.684	0.683	0.635	0.677	0.687	0.671
DTM	0.610	0.610	0.591	0.635	0.634	0.595	0.628	0.642	0.617
GNO	0.317	0.316	0.285	0.354	0.353	0.323	0.337	0.376	0.331
NDVIllg	0.441	0.439	0.418	0.479	0.479	0.457	0.451	0.429	0.480
NDVIvg	0.363	0.364	0.367	0.366	0.367	0.360	0.369	0.400	0.367
PH	0.740	0.739	0.732	0.745	0.743	0.662	0.750	0.748	0.706
TGW	0.596	0.596	0.591	0.599	0.600	0.580	0.609	0.608	0.594
YLD	0.278	0.277	0.237	0.325	0.325	0.305	0.305	0.362	0.306

† M1, E + L; M2, L + S; M3, L + Y; M4, L + G + E; M5, L + S + G; M6, L + Y + G; M7, (E + L + G + (G × E); M8, L + S + G + (G × S); M9, L + Y + G + (G × Y). L = line effect; E = environment (site–year combination) effect; G = main effect of genomic markers; G × E = genotype × environment interaction; S = site effect; G × S = genotype × site interaction; Y = year effect; G × Y = genotype × year interaction.

Table 4. Average correlations between predicted and observed values for the traits in cross-validation scenario (CV0-Sites) leaving one site out (HT, SQ, and YP) and other sites were used to predict the site that was left out (HT, SQ, and YP). The values below correspond to the average correlations between predicted and observed values of six environments (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015) for each trait. The best model for each trait is underlined. For details of each environment, see Supplementary Table 2.

Traits/Models	M1†	M2	M3	M4	M5	M6	M7	M8	M9
DTA	0.675	0.675	0.665	0.680	0.680	0.627	0.681	0.684	0.618
DTM	0.646	0.649	0.632	0.648	0.651	0.605	0.649	0.650	0.603
GNO	0.325	0.298	0.329	0.347	0.342	0.353	0.328	0.348	0.361
NDVIllg	0.423	0.398	0.424	0.444	0.436	0.442	0.427	0.429	0.433
NDVIvg	0.476	0.363	0.476	0.472	0.314	0.453	0.479	0.330	0.464
PH	0.750	0.742	0.743	0.755	0.731	0.684	0.752	0.740	0.679
TGW	0.568	0.568	0.567	0.566	0.549	0.536	0.566	0.555	0.535
YLD	0.261	0.231	0.265	0.284	0.293	0.293	0.267	0.293	0.300

† M1, E + L; M2, L + S; M3, L + Y; M4, L + G + E; M5, L + S + G; M6, L + Y + G; M7, (E + L + G + (G × E); M8, L + S + G + (G × S); M9, L + Y + G + (G × Y). L = line effect; E = environment (site–year combination) effect; G = main effect of genomic markers; G × E = genotype × environment interaction; S = site effect; G × S = genotype × site interaction; Y = year effect; G × Y = genotype × year interaction.

for NDVIvg (0.18). Among the models, the best model was M7 with the G × E interaction term for all traits.

When the CV2 scheme was tested for the lines, the correlation between the predicted and observed values for each environment was high (Supplementary Table 5). The average correlation between observed and predicted values for six environments for each trait indicated the highest PA was for PH (0.76) and the lowest was for YLD (0.40) (Table 7). The best models for the traits were M7 for GNO, NDVIllg, NDVIvg, PH, TGW, YLD, and M8 for DTA and DTM. The models with interaction terms were the best for all traits.

We compared the PA of different cross-validation schemes, traits, and all models. When comparing different traits in different CV schemes, CV1 had the lowest PA. The YLD was the trait with the lowest PA for the four CV schemes, but just like GNO, it did not show lower values in the CV1 scheme, whereas all other traits showed a decreasing trend (Fig. 1A). The trait with the lowest PA in the CV1 scheme was NDVIvg. When the cross-validation schemes were compared for each model, the lowest PA were for CV1 (Fig. 1B). Three models even had negative prediction accuracies in the CV1 scheme: M1 (L + E), M2 (L +S), and M3 (L + Y). In most cases, models with interaction terms performed better than the main effect models.

Comparing models among the CV schemes, the models with interaction terms (M7, M8, and M9) had higher prediction accuracies than their main effect models when all traits and all cross-validation schemes were considered (Fig. 2A). Among them, M7 had the highest PA > 0.51, followed by M8 (0.49) and M9 (0.46) combining all traits and cross-validation schemes. Models M1 (L + E), M2 (L +S), and M3 (L + Y) had the lowest PA < 0.40. The PH had the highest average PA among all traits (0.68) when the average of all models and all cross-validation schemes were used, and the lowest (0.34) was for YLD (Fig. 2B). Four traits—PH (0.68), DTA (0.61), DTM (0.57), and TGW (0.56)—had PA > 0.50 when the averages of all models and all cross-validation schemes were used. Accounting for all

Table 5. Average correlations between predicted and observed values for the traits in cross-validation scenario (CV0-Env) leaving one environment out (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015) and all other environments were used to predict the environment that was left out (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015). The values below correspond to the average correlations between predicted and observed values in six environments (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015) for each trait. The best model for each trait is underlined. For details of each environment, see Supplementary Table 3.

Traits/Models	M1†	M2	M3	M4	M5	M6	M7	M8	M9
DTA	0.711	0.710	0.697	0.715	0.715	0.670	0.717	0.724	0.652
DTM	0.686	0.689	0.665	0.688	0.691	0.643	0.689	0.696	0.635
GNO	0.404	0.386	0.392	0.419	0.394	0.389	0.412	0.404	0.393
NDVIllg	0.460	0.438	0.460	0.482	0.471	0.475	0.470	0.457	0.456
NDVIvg	0.511	0.410	0.503	0.509	0.366	0.494	0.515	0.437	0.484
PH	0.772	0.766	0.765	0.773	0.757	0.699	0.775	0.750	0.694
TGW	0.618	0.618	0.618	0.617	0.608	0.594	0.623	0.625	0.574
YLD	0.361	0.340	0.350	0.377	0.371	0.348	0.376	0.400	0.342

† M1, E + L; M2, L + S; M3, L + Y; M4, L + G + E; M5, L + S + G; M6, L + Y + G; M7, (E + L + G + (G × E); M8, L + S + G + (G × S); M9, L + Y + G + (G × Y). L = line effect; E = environment (site–year combination) effect; G = main effect of genomic markers; G × E = genotype × environment interaction; S = site effect; G × S = genotype × site interaction; Y = year effect; G × Y = genotype × year interaction.

Table 6. Average correlations and standard deviations (μ ± σ) between predicted and observed values for the traits in cross-validation scenario (CV1) using five-fold cross-validation for each environment (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015) using nine different models (M1 to M9). The values below correspond to the average correlations between predicted and observed values in six environments (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015) for each trait. The best model for each trait is underlined. For details of each environment, see Supplementary Table 4.

Traits/Models	M1†	M2	M3	M4	M5	M6	M7	M8	M9
DTA	-0.09 ± 0.07	-0.08 ± 0.06	-0.10 ± 0.06	0.39 ± 0.03	0.39 ± 0.04	0.32 ± 0.04	0.41 ± 0.04	0.40 ± 0.04	0.33 ± 0.04
DTM	-0.09 ± 0.06	-0.07 ± 0.06	-0.10 ± 0.06	0.31 ± 0.03	0.32 ± 0.03	0.22 ± 0.04	0.33 ± 0.04	0.32 ± 0.04	0.25 ± 0.04
GNO	-0.10 ± 0.06	-0.07 ± 0.07	-0.07 ± 0.07	0.27 ± 0.02	0.25 ± 0.02	0.25 ± 0.02	0.36 ± 0.03	0.29 ± 0.02	0.27 ± 0.02
NDVIllg	-0.12 ± 0.05	-0.10 ± 0.06	-0.10 ± 0.06	0.30 ± 0.03	0.28 ± 0.03	0.28 ± 0.03	0.30 ± 0.03	0.29 ± 0.03	0.28 ± 0.02
NDVIvg	-0.12 ± 0.06	-0.07 ± 0.07	-0.10 ± 0.05	0.13 ± 0.04	0.08 ± 0.03	0.13 ± 0.04	0.18 ± 0.03	0.13 ± 0.03	0.14 ± 0.04
PH	-0.08 ± 0.08	-0.09 ± 0.06	-0.10 ± 0.05	0.50 ± 0.02	0.50 ± 0.01	0.49 ± 0.01	0.51 ± 0.02	0.50 ± 0.02	0.49 ± 0.01
TGW	-0.09 ± 0.07	-0.08 ± 0.07	-0.09 ± 0.06	0.31 ± 0.03	0.29 ± 0.03	0.26 ± 0.03	0.39 ± 0.03	0.33 ± 0.03	0.28 ± 0.03
YLD	-0.10 ± 0.06	-0.07 ± 0.07	-0.06 ± 0.07	0.26 ± 0.01	0.26 ± 0.01	0.25 ± 0.01	0.33 ± 0.03	0.29 ± 0.02	0.26 ± 0.01

† M1, E + L; M2, L + S; M3, L + Y; M4, L + G + E; M5, L + S + G; M6, L + Y + G; M7, (E + L + G + (G × E); M8, L + S + G + (G × S); M9, L + Y + G + (G × Y). L = line effect; E = environment (site–year combination) effect; G = main effect of genomic markers; G × E = genotype × environment interaction; S = site effect; G × S = genotype × site interaction; Y = year effect; G × Y = genotype × year interaction.

Table 7. Average correlations and standard deviations (μ ± σ) between predicted and observed values for the traits in cross-validation scenario (CV2) using five-fold cross-validations for the environments (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015) using nine different models (M1 to M9). The values below correspond to the average correlations between the predicted and observed values in six environments (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015) for each trait. The best model for each trait is underlined. For details of each environment, see Supplementary Table 5.

Traits/Models	M1†	M2	M3	M4	M5	M6	M7	M8	M9
DTA	0.69 ± 0.02	0.69 ± 0.02	0.56 ± 0.04	0.70 ± 0.02	0.70 ± 0.02	0.60 ± 0.03	0.72 ± 0.02	0.72 ± 0.02	0.58 ± 0.03
DTM	0.66 ± 0.03	0.66 ± 0.03	0.44 ± 0.06	0.67 ± 0.03	0.67 ± 0.03	0.48 ± 0.05	0.67 ± 0.03	0.69 ± 0.03	0.48 ± 0.05
GNO	0.37 ± 0.03	0.29 ± 0.05	0.25 ± 0.05	0.40 ± 0.02	0.35 ± 0.03	0.33 ± 0.03	0.44 ± 0.03	0.37 ± 0.03	0.34 ± 0.03
NDVIllg	0.42 ± 0.03	0.36 ± 0.04	0.31 ± 0.05	0.45 ± 0.02	0.42 ± 0.04	0.39 ± 0.04	0.46 ± 0.02	0.42 ± 0.04	0.38 ± 0.04
NDVIvg	0.49 ± 0.02	0.27 ± 0.05	0.45 ± 0.03	0.49 ± 0.02	0.27 ± 0.04	0.46 ± 0.03	0.52 ± 0.02	0.28 ± 0.04	0.47 ± 0.03
PH	0.75 ± 0.01	0.71 ± 0.02	0.61 ± 0.04	0.76 ± 0.01	0.72 ± 0.02	0.64 ± 0.02	0.76 ± 0.01	0.71 ± 0.02	0.63 ± 0.02
TGW	0.60 ± 0.01	0.58 ± 0.02	0.54 ± 0.03	0.61 ± 0.01	0.59 ± 0.02	0.56 ± 0.02	0.66 ± 0.01	0.62 ± 0.02	0.56 ± 0.02
YLD	0.32 ± 0.04	0.25 ± 0.05	0.20 ± 0.06	0.36 ± 0.03	0.33 ± 0.04	0.30 ± 0.04	0.40 ± 0.03	0.36 ± 0.04	0.31 ± 0.04

‡ M1, E + L; M2, L + S; M3, L + Y; M4, L + G + E; M5, L + S + G; M6, L + Y + G; M7, (E + L + G + (G × E); M8, L + S + G + (G × S); M9, L + Y + G + (G × Y). L, line effect; E, environment (site–year combination) effect; G, main effect of genomic markers; G × E, genotype × environment interaction; S, site effect; G × S, genotype × site interaction; Y, year effect; G × Y, genotype × year interaction.

traits and all models, the best cross-validation schemes were CV0-Env and CV2, both with an average PA of 0.58 (Fig. 3). CV0-Sites and CV0-Year had similar PA (0.53), but higher PA than the CV1 scheme (0.35).

DISCUSSION

The application of genomic prediction models in breeding programs can increase genetic gains by shortening the breeding cycle. In our study, we validated genomic prediction models incorporating G × E interaction term to the GBLUP model of Jarquín et al. (2014, 2017) in a CIMMYT durum wheat panel, which was phenotyped under contrasting environments; well-watered, drought stress and heat stress conditions. Crossa et al. (2016) did perform genomic-enabled prediction studies in durum wheat using the Bayes B variable selection model together with the GBLUP G × E that considers the main effects of markers across all environments and the marker specific environment effect.

We tested nine genomic prediction models—six models non-G × E and three with G × E interaction terms in five cross-validation schemes for eight traits in three different environments. Results indicated high PA when models with G × E terms were used (Burgueño et al., 2011; Pérez-Rodríguez et al., 2015; Jarquín et al., 2017). The prediction accuracies reported here in durum wheat are relatively high and for some traits slightly higher than those reported in spring wheat. The environmental covariables on the environments used in this study varies and this might have influenced the response traits like NDVI, grain yield, etc. In addition, the amount of precipitation for the 2 yr was different, could have influence the response of some traits, especially grain yield, and contributed to the G × E interaction. The GBLUP G × E could have incorporated environmental covariables (Jarquín et al., 2014) and studied their influence, however in the study no environmental covariables were added into the model.

Among the cross-validation schemes, CV2 and CV0-Env had the highest PA and CV1 had the lowest PA. This

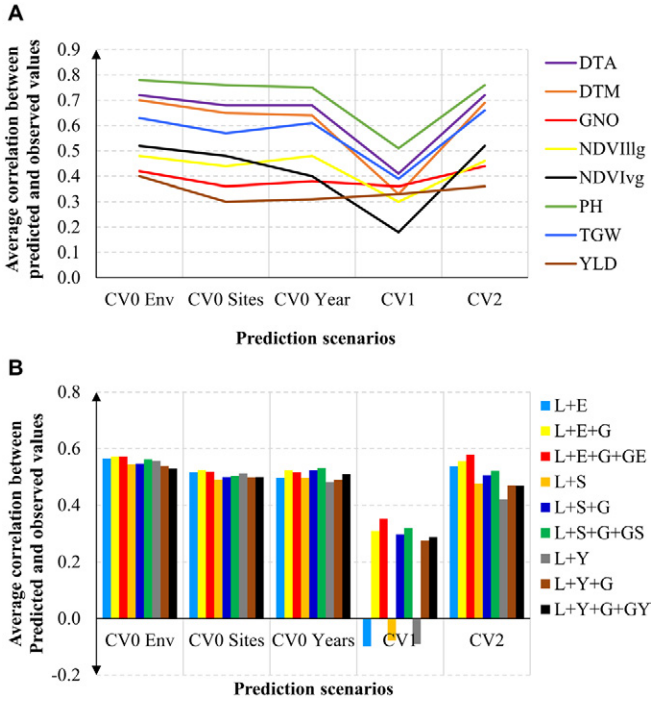

Fig. 1. Average correlations between predicted and observed values for five different cross-validation schemes: environments predicted using other environments (CV0-Env), sites predicted using other sites (CV0-Sites), years predicted using other years (CV0-Year), a set of lines predicted using other lines in the same environment (CV1), and sparse testing (CV2); (A) for different traits and (B) nine different models.

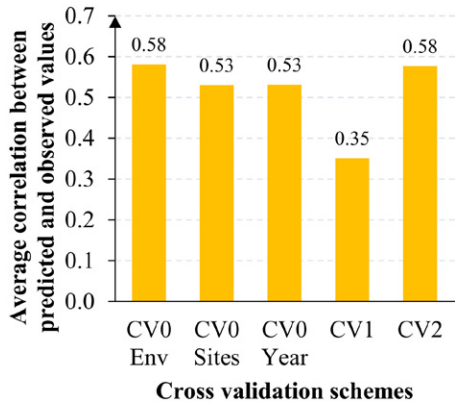

Fig. 3. Average correlations between predicted and observed values for five different cross-validation scenarios: environments predicted using other environments (CV0-Env), sites predicted using other sites (CV0-Sites), years predicted using other years (CV0-Year), a set of untested lines predicted using other lines in the same environment (CV1), and sparse testing (CV2) considering all models and all traits.

is similar to previous reports in spring wheat (Sukumaran et al., 2017a; b) and cotton (Pérez-Rodríguez et al., 2015). Earlier studies have reported low heritability as the reason for low PA for CV1, in the present experiment we have seen irrespective of the trait heritability the PA are lower for CV1. In durum wheat, earlier studies have shown PA for CV2 was higher than for CV1 (Crossa et al., 2016).

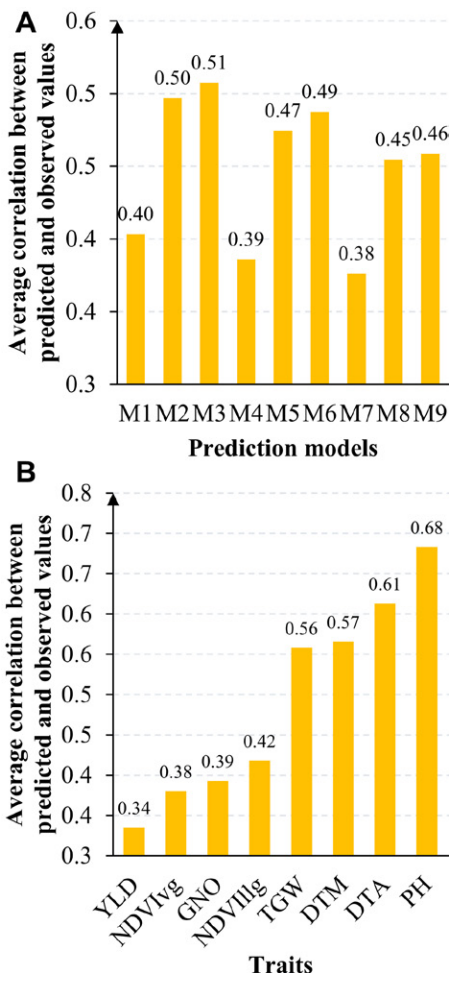

Fig. 2. Average correlations between predicted and observed values for the (A) nine different models considering all traits and all cross-validation scenarios and (B) for different traits. Model 1, E + L; M2, L + S; M3, L + Y; M4, L + G + E; M5, L + S + G; M6, L + Y + G; M7, (E + L + G + (G × E); M8, L + S + G + (G × S); M9, L + Y + G + (G × Y). L = line effect; E = environment (site–yr combination) effect; G = main effect of genomic markers; G × E = genotype × environment interaction; S = site effect; G × S = genotype × site interaction; Y = year effect; G × Y = genotype × year interaction.

Among the traits, PH was the highest predicted trait possibly because its heritability estimates were higher than the heritability estimates for YLD, which had the lowest PA. The TGW had very high PA similar to an earlier study in spring wheat and was highly predicted (Velu et al., 2016).

This study indicated that good PA could be obtained for YLD and its components even when predicting lines, environments, years, and sites are missing. According to the results, it is practical to implement genomic prediction and selection in a cost-effective manner by using more environments to test the germplasm by reducing replications or phenotyping only a fraction of the lines. The G × E models should be used to get good predicted values when using the CV1 scheme, where negative PA were observed when main effect models were used.

The present study in durum wheat demonstrated that environments, sites, and years can be predicted

reasonable well and thus facilitating an efficient use of time and funds for predicting unobserved lines (Yu et al., 2016; Crossa et al., 2017; Jarquín et al., 2017). The PA reported in this study is good enough to discard the lines in a real breeding program (Velu et al., 2016). As found in other crops like maize where GS has been successfully applied (Beyene et al., 2015) and the good genomic-enabled prediction found in spring and winter wheat, the results of this study indicated that GS in durum wheat can be successfully applied in breeding programs. (Habash et al., 2009).

CONCLUSIONS

In the present study, we used the reaction norm models for genomic prediction and applied them to a durum panel phenotyped under well-watered, drought stress, and heat stress conditions. Addition of G × E interaction terms to the model increased the PA in all cross validation schemes. The best cross validation scheme was predicting missing lines and predicting lines in untested environments (CV2). Forward prediction of years and sites (CV0) were also moderate to high than the CV1 scheme. High heritability traits showed high PA in all CV schemes. The results stress the importance of genomic prediction models incorporating G × E interactions to predict the performance of lines in forward breeding for application in durum-breeding programs.

Author Contributions

SS and JC conceived and designed the study; SS conducted the molecular marker and phenotypic data analysis; DJ conducted the genomic prediction analysis and proposed the different prediction scenarios/models; SS., DJ, and JC wrote the manuscript. All authors reviewed and commented on the manuscript.

Supplemental Information

Supplementary Table 1. Correlations between the predicted and observed values of the traits in six environments when 1 year was left out (2015 or 2016) and the other year was used to predict the year that was left out (2015 or 2016).

Supplementary Table 2. Correlations between predicted and observed values for the traits in cross-validation scenario (CV0) where one site was left out (HT, YP, or DT) and the other sites were used to predict the site that was left out (HT, YP, or DT).

Supplementary Table 3. Correlations between predicted and observed values for the traits in cross-validation scenario (CV0) where one environment was left out (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015) and all other environments were used to predict the environment that was left out (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015).

Supplementary Table 4. Correlations and standard deviations (μ ± σ) between predicted and observed values for the traits in cross-validation scenario (CV1) using five-fold cross-validation for each environment (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015) using nine different models (M1 to M9).

Supplementary Table 5. Correlations and standard deviations (μ ± σ) between predicted and observed values for the traits in cross-validation scenario (CV2) using five-fold cross-validation for each environment (HT 2015, YP 2015, DT 2015, HT 2015, YP 2015, and DT 2015) using nine different models (M1 to M9)

Conflict of Interest Disclosure

The authors declare that there is no conflict of interest.

Acknowledgments

This study was implemented by CIMMYT as part of International Wheat Yield Partnership (IWYP), made possible by the generous support of Sustainable Modernization of Traditional Agriculture (MasAgro) initiative from the Secretariat of Agriculture, Livestock, Rural Development, Fisheries and Food and the WHEAT CRP. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of donors. We also acknowledge Hans Braun, Karim Ammar, Carolina Sansaloni, and Tom Payne for their support to this study. Thanks are due to Jacinta Gimeno Romeu, Jose Luis Gerardo Barrios Gonzalez, Araceli Torres Garcia, Nayeli Quiche Morales, Noel Pimienta Valenzuela, and the staff at Wheat Physiology lab at Campo Experimental Norman E Borlaug, Cd. Obregon, Mexico.

References

Alvarado, G., M. López, M. Vargas, Á. Pacheco, F. Rodríguez, J. Burgueño, and C. Crossa. 2015. META-R (Multi Environment Trail Analysis with R for Windows. Version 5.0. CIMMYT.

Beyene, Y., K. Semagn, S. Mugo, A. Tarekegne, R. Babu, B. Meisel, P. Sehabiague, D. Makumbi, C. Magorokosho, S. Oikeh, J. Gakunga, M. Vargas, M. Olsen, B.M. Prasanna, M. Banziger, and J. Crossa. 2015. Genetic gains in grain yield through genomic selection in eight bi-parental maize populations under drought stress. Crop Sci. 55:154–163. doi:10.2135/cropsci2014.07.0460

Burgueño, J., J. Crossa, J.M. Cotes, F.S. Vicente, and B. Das. 2011. Prediction assessment of linear mixed models for multienvironment trials. Crop Sci. 51:944–954. doi:10.2135/cropsci2010.07.0403

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

Crossa, J., G. De Los Campos, M. Maccaferri, R. Tuberosa, J. Burgueño, and P. Pérez-Rodríguez. 2016. Extending the marker × Environment interaction model for genomic-enabled prediction and genome-wide association analysis in durum wheat. Crop Sci. 56:2193–2209. doi:10.2135/cropsci2015.04.0260

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

Crossa, J., P. Pérez-Rodríguez, J. Cuevas, O. Montesinos-López, D. Jarquín, G. de los Campos, J. Burgueño, J.M. Camacho-González, S. Pérez-Elizalde, Y. Beyene, S. Dreisigacker, R. Singh, X. Zhang, M. Gowda, M. Roorkiwal, J. Rutkoski, and R.K. Varshney. 2017. Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. Trends Plant Sci. xx:1–15. 10.1016/j.tplants.2017.08.011

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

Fiedler, J.D., E. Salsman, Y. Liu, M. Michalak de Jiménez, J.B. Hegstad, B. Chen, F.A. Manthey, S. Chao, S. Xu, E.M. Elias, and X. Li. 2017. Genome-Wide Association and Prediction of Grain and Semolina

Quality Traits in Durum Wheat Breeding Populations. Plant Genome 0:0. doi:10.3835/plantgenome2017.05.0038

Habash, D.Z., Z. Kehel, and M. Nachit. 2009. Genomic approaches for designing durum wheat ready for climate change with a focus on drought. J. Exp. Bot. 60:2805–2815. doi:10.1093/jxb/erp211

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

Jarquín, D., C. Lemes da Silva, R.C. Gaynor, J. Poland, A. Fritz, R. Howard, S. Battenfield, and J. Crossa. 2017. Increasing Genomic-Enabled Prediction Accuracy by Modeling Genotype × Environment Interactions in Kansas Wheat. Plant Genome 10:0. doi:10.3835/plantgenome2016.12.0130

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

Pask, A., J. Pietragalla, and D. Mullan. 2012. Physiological Breeding II: A Field Guide to Wheat Phenotyping. Mexico, DF (Mexico).

Pérez-Rodríguez, P., J. Crossa, K. Bondalapati, G. De Meyer, F. Pita, and G. De Los Campos. 2015. A pedigree-based reaction norm model for prediction of cotton yield in multienvironment trials. Crop Sci. 55:1143–1151. doi:10.2135/cropsci2014.08.0577

Pérez-Rodríguez, P., J. Crossa, J. Rutkoski, J. Poland, R. Singh, A. Legarra, E. Autrique, G. de los Campos, J. Burgueño, and S. Dreisigacker. 2017. Single-Step Genomic and Pedigree Genotype × Environment Interaction Models for Predicting Wheat Lines in International Environments. Plant Genome 10:0. doi:10.3835/plantgenome2016.09.0089

Pérez-Rodríguez, P., D. Gianola, J.M. González-Camacho, J. Crossa, Y. Manès, and S. Dreisigacker. 2012. Comparison Between Linear and Non-parametric Regression Models for Genome-Enabled Prediction in Wheat. G3: Genes|Genomes|Genetics 2:1595–1605. doi:10.1534/g3.112.003665

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

Sansaloni, C., C. Petroli, D. Jaccoud, J. Carling, F. Detering, D. Grattapaglia, and A. Kilian. 2011. Diversity Arrays Technology (DArT) and next-generation sequencing combined: Genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. BMC Proc. 5:P54. doi:10.1186/1753-6561-5-S7-P54

Sukumaran, S., J. Crossa, D. Jarquin, M. Lopes, and M.P. Reynolds. 2017a. Genomic Prediction with Pedigree and Genotype × Environment Interaction in Spring Wheat Grown in South and West Asia, North Africa, and Mexico. G3 Genes|Genomes|Genetics 7:481 LP-495. doi:10.1534/G3.116.036251

Sukumaran, S., J. Crossa, D. Jarquín, and M. Reynolds. 2017b. Pedigree-based prediction models with genotype × environment interaction in multienvironment trials of CIMMYT wheat. Crop Sci. 57:1865–1880. doi:10.2135/cropsci2016.06.0558

Sukumaran, S., M.P. Reynolds, and C.P. Sansaloni. 2018. Genome-wide association analyses identify QTL hotspots for yield and component traits in durum wheat grown under yield potential, drought, and heat stress environments. Front. Plant Sci. 9:81. doi:10.3389/fpls.2018.00081

Velu, G., J. Crossa, R.P. Singh, Y. Hao, S. Dreisigacker, P. Perez-Rodriguez, A.K. Joshi, R. Chatrath, V. Gupta, A. Balasubramaniam, C. Tiwari, V.K. Mishra, V.S. Sohu, and G.S. Mavi. 2016. Genomic prediction for grain zinc and iron concentrations in spring wheat. Theor. Appl. Genet. 129:1595–1605. doi:10.1007/s00122-016-2726-y

Yu, X., X. Li, T. Guo, C. Zhu, Y. Wu, S.E. Mitchell, K.L. Roozeboom, D. Wang, M.L. Wang, G.A. Pederson, T.T. Tesso, P.S. Schnable, R. Bernardo, and J. Yu. 2016. Genomic prediction contributing to a promising global strategy to turbocharge gene banks. Nat. Plants 2:16150. doi:10.1038/nplants.2016.150