Published online February 22, 2018

RESEARCH

Single-Step Reaction Norm Models for Genomic Prediction in Multienvironment Recurrent Selection Trials

Odilon P. Morais Júnior,* João Batista Duarte, Flávio Breseghello, Alexandre S. G. Coelho, Orlando P. Morais, and Ariano M. Magalhães Júnior

ABSTRACT

In recurrent selection programs, progeny testing is done in multienvironment trials, which generates genotype × environment interaction (G × E). Therefore, modeling G × E is essential for genomic prediction in the context of recurrent genomic selection (RGS). Developing single-step, best linear unbiased prediction-based reaction norm models (termed RN-HBLUP) using data from nongenotyped and genotyped progenies, can enhance predictive accuracy. Our objectives were to evaluate: (i) a class of RN-HBLUP models accommodating combined relationship of pedigree and genomic data, environmental covariates, and their interactions for prediction of phenotypic responses; (ii) the predictive accuracy of these models and the relative importance of main effects and interaction components; and (iii) the influence of different grouping strategies of genetic–environmental data (within selection cycles or across cycles) on prediction accuracy of the merit for untested progenies. The genetic material comprised 667 $S_{1:3}$ progenies of irrigated rice (*Oryza sativa* L.) and six check cultivars. These materials were evaluated in yield trials conducted in 10 environments during three selection cycles. Genomic information was derived from single-nucleotide polymorphism markers genotyped on 174 progenies in the third cycle. We evaluated six predictive models. Environmental covariates and G × E interaction explained a significant portion of the phenotypic variance, increasing accuracy and decreasing the bias of phenotypic prediction. Within-cycle data were sufficient for accurate prediction of untested progenies, even in untested environments. We concluded that the RN-HBLUP model, with the comprehensive structure, could be useful in improving the prediction accuracy of quantitative traits in RGS programs.

O.P. Morais Júnior, J.B. Duarte, and A.S.G. Coelho, Dep. of Genetics and Plant Breeding, Escola de Agronomia, Federal Univ. of Goiás, Campus Samambaia, Ave. Esperança, s/n, CEP 74690-900, Goiânia, Goiás, Brazil; F. Breseghello and O.P. Morais, Embrapa Arroz e Feijão, Rodovia GO-462, km 12, Rural Zone, CP 179, CEP 75375-000, Santo Antônio de Goiás, Goiás, Brazil; A.M. Magalhães Júnior, Embrapa Clima Temperado, Rodovia BR-392, km 78, CP 403, CEP 96010-971, Pelotas, Rio Grande do Sul, Brazil. Article dedicated to the memory of O.P. Morais. Received 15 June 2017. Accepted 5 Nov. 2017. *Corresponding author (odilonpmorais@gmail.com). Assigned to Associate Editor Manjit Kang.

Abbreviations: **A**, pedigree-based relationship matrix; ABLUP, pedigree-based best linear unbiased prediction model; BLUP, best linear unbiased prediction; CV, cross-validation; DTF, days to flowering; EBV, estimated breeding value; EC, environmental covariate; Embrapa, Brazilian Agricultural Research Corporation; **G**, marker-based relationship matrix; G × E, genotype × environment interaction; GBLUP, marker-based best linear unbiased prediction model; GS, genomic selection; GY, grain yield; **H**, combined relationship matrix; HBLUP, single-step best linear unbiased prediction model; H × E, combined relationship × environment interaction; H × W, combined relationship × environmental conditions interaction; MCMC, Markov chain Monte Carlo; MET, multienvironment trial; PH, plant height; RGS, recurrent genomic selection; RN-HBLUP, single-step best linear unbiased prediction-based reaction norm model; SNP, single-nucleotide polymorphism.

Recurrent selection programs are aimed at improving populations by increasing the frequency of favorable alleles that determine quantitative traits through consecutive selection–recombination cycles while ensuring that genetic variability is not exhausted through the selection cycles (Bernardo, 2010). In recurrent selection schemes based on the genomic selection (GS; Meuwissen et al., 2001) approach, known as recurrent genomic selection (RGS; Bernardo, 2010), the most challenging phase is

Published in Crop Sci. 58:592–607 (2018).
doi: 10.2135/cropsci2017.06.0366

the training of genomic-enabled prediction models. In this phase, in addition to accurate genotyping, high-quality phenotypic data must be obtained on traits of interest in multienvironment trials (METs). Progeny testing is usually done in multiple, replicated field trials involving many entries, which are aimed at selecting progenies with broad adaptability and stability for the target environments. This strategy can reduce the risk of discarding certain progenies that potentially perform well in some but not all target environments.

Genotype × environment interaction (G × E) has great influence on the phenotypic prediction of progenies in METs, with consequent reduction of genetic progress from selection (Allard and Bradshaw, 1964; Kang and Magari, 1996; Xu, 2016). The relevance of G × E might increase in the face of global climate change that can generate pronounced differences between environments (Heslot et al., 2014; Trenberth et al., 2014). Breeding programs produce a large volume of phenotypic data across time, usually from unbalanced randomized trials conducted in different testing locations, which may result from planned unbalanced designs or missing observations from the unequal numbers of progenies and replications (Bernardo, 2010; Dawson et al., 2013; Lado et al., 2016). Nevertheless, use of appropriate statistical models and historical unbalanced datasets can be useful in increasing the efficiency of the breeding program, especially for modeling G × E in genomic prediction. Considering that progeny testing is one of the most expensive and time-consuming phases in a RGS program, designing efficient strategies for modeling G × E and improving prediction accuracy are desirable.

Statistical treatment of G × E has evolved across time, in accordance with the development of biometrical methods and the increasing availability of high-quality genetic, genomic, and environmental data (Eberhart and Russell, 1966; van Eeuwijk et al., 1996; Burgueño et al., 2012; Jarquín et al., 2014; Pérez-Rodríguez et al., 2017). Originally, genetic prediction models were developed for predicting phenotypes of individuals on the basis of phenotypic information from single environments or multiple environments without considering G × E by using pedigree-based (**A**-matrix; Henderson, 1976) or marker-based (**G**-matrix; VanRaden, 2008) relationship matrices. Those models assumed constant progeny behavior across the target environments for phenotypic prediction, although it is well known that crossover type of G × E commonly causes genotype rank changes in MET data (Kang, 1997; Yang, 2007; Xu, 2016). Thus, there is a need for the development of multienvironment models that consider modeling G × E, with the aim of improving ability to predict phenotypes accurately, even in environments where progeny or line was not cultivated or tested.

Several genetic or metabolic processes associated with chemical reactions mediated by enzymes are affected by the fluctuations of environmental factors (Hasanuzzaman et al., 2013). Currently, high-dimensional data about those factors (e.g., climate and soil factors that influence plant growth and development) that are obtained by so-called "envirotyping" technology, are available (Xu, 2016). With this approach, it is possible to dissect the environmental component into individual factors, thus generating several environmental covariates (ECs). Therefore, by incorporating precise measurements of environmental effects across developmental stages of the crop, coupled with precise phenotyping, it is possible to decompose G × E into interrelated factors, optimizing phenotype prediction (Kang and Gauch, 1996; Kang, 1997; Xu, 2016). The use of marker and/or pedigree data and EC data from historical trials conducted with good experimental quality allows for more accurate predictions by modeling G × E in genomic-enabled prediction models. Based on this approach, the phenotypic prediction can be done in three ways: untested genotypes in tested environments, tested genotypes in untested environments, and untested genotypes in untested environments (Bustos-Korts et al., 2016).

Based on MET data, multienvironmental genomic prediction models and methods have been proposed for wheat (*Triticum aestivum* L.) breeding. For example, to deal with G × E in prediction models, Burgueño et al. (2012) proposed a multivariate version of the genomic best linear unbiased predictor (GBLUP) on the basis of a mixed model with factor analytic structure. Heslot et al. (2014) proposed an extension of the factorial regression with inclusion of ECs associated with different phenological phases in a crop model. Jarquín et al. (2014) proposed a class of Bayesian models that accommodates high-dimensional genomic and EC data, plus their interactions, using reaction norm-based GBLUP. López-Cruz et al. (2015) considered a GS model that incorporates marker × environment interaction, which can be used not only for genomic prediction, but also for identifying chromosomal regions that produce stable effects across environments and other regions that are responsible for G × E, via genome-wide association analysis (Crossa et al., 2016). Cuevas et al. (2016) modeled G × E with multi-environment Bayesian genomic prediction kernel models. These previous studies highlighted the potential of multi-environment models to investigate the genetic basis of G × E to determine genotype sensitivity and to increase the prediction accuracy, compared with models that consider G × E as part of the experimental error.

Among the multienvironmental genomic-prediction approaches proposed so far, the Jarquín et al. (2014) models are among the most comprehensive ones, especially where high-dimensional EC data have been integrated. In this approach, genotypic and environmental gradients are

described as linear functions of genetic (**G**-matrix or **A**-matrix) and EC effects, allowing the use of phenotypic, genomic, full-pedigree, and environmental data accumulated across time to increase prediction accuracy. That approach has proven efficient in different cereal crops. In maize (*Zea mays* L.), this class of reaction norm models has been shown to be promising in structuring G × E, as it allows borrowing genetic information across environments (Zhang et al., 2014). In cotton (*Gossypium hirsutum* L.), Pérez-Rodríguez et al. (2015) extended this class of Bayesian multiplicative model to pedigree-based analysis (reaction norm-based pedigree-based best linear unbiased prediction [ABLUP]), accommodating the **A**-matrix. Sukumaran et al. (2017) used the Jarquín's models in a wheat population phenotyped under a diverse set of environments around the world for studying models that include (or do not include) genomic and pedigree interactions with environments but without EC data. All these studies detected a marked increase in predictive accuracy with the most comprehensive prediction model. Despite these advances, it is still necessary to assess this class of reaction norm models for other quantitative traits in other self-pollinating species, especially in RGS programs.

Use of all phenotypic information available for a breeding population increases accuracy for multienvironmental genomic prediction. Considering that not all individuals are genotyped or phenotyped, the procedure known as "single-step" (Legarra et al., 2009; Misztal et al., 2009) provides a unified computational framework to retrieve and integrate phenotypic, full-pedigree, and genomic information for more accurate genomic prediction in single or multienvironmental context. According to this approach, the **G**-matrix from genotyped individuals is blended with the **A**-matrix from all individuals (genotyped and nongenotyped), generating a single combined relationship matrix (**H**-matrix), which can be used to compute best linear unbiased prediction (BLUP) based on Henderson's (1976) standard mixed-model equations, giving rise to a single-step best linear unbiased prediction (HBLUP) method.

Legarra et al. (2014) demonstrated that use of the single-step method, either by frequentist or Bayesian framework, for large-scale multiple regression allows higher predictive accuracy than expected with GBLUP or ABLUP models due to the larger population size, because it uses both genotyped and nongenotyped individuals. This approach has been extensively applied with good results in animal breeding (Aguilar et al., 2010; Legarra et al., 2014; Mota et al., 2016), but in plant breeding, examples are still scant. Ashraf et al. (2016) used a multivariate Gaussian mixed model to deal with G × E in wheat. Pérez-Rodríguez et al. (2017) used single-step-based reaction norm models (but not incorporating ECs) for exploiting G × E in a large reference set of wheat lines that had available DNA markers only for about half of the lines. Therefore, the Bayesian reaction norm models of Jarquín et al. (2014) have the potential to be fit with any genetic relationship information (DNA marker and/or pedigree) and EC data, especially by using the HBLUP method.

Genome-enabled prediction of quantitative traits based on MET data is a great challenge in plant breeding. In this context, the potential of single-step-based reaction norm models (termed RN-HBLUP) by using high-dimensional EC data has not yet been investigated in self-pollinating species in RGS programs. Therefore, we fitted here some of these models for multienvironment prediction in a RGS rice (*Oryza sativa* L.) program, with the following objectives: (i) to evaluate a class of RN-HBLUP models, including combined pedigree and genomic information, EC data, and their interactions; (ii) to evaluate the predictive accuracy of these models and the relative importance of additive (main effects of combined relationship and EC data) and multiplicative components (interactions between the main effects); and (iii) to evaluate the influence of different methods of grouping genetic–environmental information (within selection cycle or across cycles) on the accuracy of the models to predict untested progenies in tested and untested environments.

MATERIALS AND METHODS

Phenotypic Data

Historic data from three selection cycles of the CNA12S population of irrigated rice synthesized by Embrapa (Brazilian Agricultural Research Corporation) were used in this study. The dataset comprises 10 progeny yield trials conducted during 2005 (Cycle 1), 2009 (Cycle 2), and 2014 and 2015 (Cycle 3) at six different locations. Each location-year combination was considered one environment. Details about the genealogy of this population, environments, and number of $S_{1:3}$ progenies and check cultivars evaluated in each recurrent selection cycle are described in Morais Júnior et al. (2017).

The dataset was unbalanced because of an unequal number of progenies in each environment per selection cycle, with only two or four common check cultivars and missing data. Additionally, different experimental designs were also used: augmented block design (Federer, 1956) with unreplicated progenies (Cycles 1 and 2), square lattice design (15 × 15), with two replications (2014, Cycle 3), and augmented square lattice design with common check cultivars between blocks (14 × 17) and two replications for progenies (2015, Cycle 3).

Yield trials were conducted in lowland area with continuous flooding until grain maturity was achieved. Fertilizer was applied at farmer's level, including topdressing with nitrogen at final vegetative stage, V5 (Counce et al., 2000). Weed and insect pest controls were done via mechanized spraying. No fungicides were applied. Three phenotypic traits were measured: grain yield at 13% moisture (GY, kg ha^{-1}); plant height from ground level to panicle tip at prematurity (R7) from a sample of six plants per plot (PH, cm); and days to flowering, from sowing to 50% of plants at anthesis (DTF).

Environmental Covariates

Ten environmental factors, including solar radiation, air temperature, humidity, rainfall, and wind speed (Table 1), were obtained from NASA orbital sensors (Stackhouse Jr., 2014). From those, 120 ECs were produced, resulting from the division of each environmental factor into four classes (with limits corresponding to the first, second, and third quartiles of the distribution of each environmental factor, per environment) and also from combination of the environmental factors in pairs across the four classes, corresponding to the incidence (0 = absence and 1 = presence) of the measured value in each combination. Aiming to exploit the temporal variation within each environment, the crop cycle was divided into five phenological phases: (i) from sowing to V4, beginning of tillering; (ii) from V4 to R1, differentiation of the panicle; (iii) from R1 to R4, anthesis; (iv) from R4 to R7, milky grain; and (v) from R7 to R9, the harvest point (Counce et al., 2000). Also, progenies and check cultivars were classified into maturation groups (early, middle, and late) according to the phenology of reference cultivars in irrigated rice (Freitas et al., 2006). A total of 650 ECs was generated for each genotype (10 environmental factors plus 120 ECs, multiplied by five phenological phases), where each final EC measure corresponded to the sum of daily measurements of the respective variable within the interval for each phenological phase. The same procedure was performed independently for genotypes of each maturity group.

As quality control, ECs with >30% of repeated values or >5% outside the range of mean ± 4 SDs were removed, resulting in a final set of 401 ECs, which were standardized. The matrix $\mathbf{W}$, which contains information on genotype-EC combination for each environment, was generated according to Jarquín et al. (2014) method. From matrix $\mathbf{W}$, the matrix $\mathbf{\Omega}$ was derived:

$$\mathbf{\Omega} = \frac{\mathbf{W}\mathbf{W}^{\mathrm{T}}}{q}$$

where q is the number of ECs and the superscript T represents transposed. Because of the prior centralization of the matrix $\mathbf{W}$ and standardization of $\mathbf{\Omega}$ by the q-value, matrix $\mathbf{\Omega}$ has average diagonal coefficients equal to one. This matrix describes the similarity among environmental conditions for combinations of genotypes and environments.

Pedigree and Genotypic Data

Pedigree information for all of the 667 $S_{1:3}$ progenies evaluated was recovered up to the 10th generation, representing all known ancestors of the 16 parents used to synthesize the CNA12S population, comprising 951 genotypes. With this information, we calculated the **A**-matrix, corresponding to an average numerator relationship matrix (Wright, 1922) between all the genotypes involved in the pedigree. For genetic analysis, we considered only the portion of the **A**-matrix corresponding to the 667 $S_{1:3}$ progenies.

Of the 196 $S_{1:3}$ progenies from the third selection cycle (2015), 174 progenies were genotyped with 7735 single-nucleotide polymorphisms (SNPs) at Diversity Arrays Technology (DArT), Canberra, ACT, by the DArTseq method using an Illumina HiSeq 2500 sequencer. By filtering polymorphic markers with thresholds of call rate >75% and minor allele frequency ≥5%, 6174 SNPs were retained for genetic analyses. Missing values (~4.5% of the data points) were replaced by the marker expected value computed from estimates of allele frequencies in the dataset (Pérez-Rodríguez and de los Campos, 2014).

The unification of the **A**-matrix with the **G**-matrix in a single matrix was performed via construction of the **H**-matrix using the single-step procedure (Legarra et al., 2009; Misztal et al., 2009). To calculate the **H**-matrix, we used the following general equation:

$$\mathbf{H} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{G}_{\mathrm{w}} \end{bmatrix} = \mathbf{A} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_{\mathrm{w}} - \mathbf{A}_{22} \end{bmatrix}$$

where 1 corresponds to nongenotyped individuals and 2 to genotyped individuals; **A** is the additive relationship matrix derived from the pedigree; and $\mathbf{G}_{\mathrm{w}}$ is the weighted **G**-matrix, $\mathbf{G}_{\mathrm{w}} = (1 - w)\mathbf{G}_{\mathrm{n}} + w\mathbf{A}_{22}$, where w is the weight attributed to the portion of the genetic variance not explained by the markers (Christensen and Lund, 2010) and $\mathbf{G}_{\mathrm{n}}$ is the normalized **G**-matrix. Because marker-based relationships are more efficient in explaining relationships among progenies, capturing the effects of Mendelian segregation, we used $w = 0.05$. Following the VanRaden (2008) method, the matrix $\mathbf{G}_{\mathrm{n}}$ was obtained, accommodating inbreeding (Forni et al., 2011), as follows:

$$\mathbf{G}_{\mathrm{n}} = \frac{(\mathbf{M} - \mathbf{P})(\mathbf{M} - \mathbf{P})'}{\mathrm{tr}\left[(\mathbf{M} - \mathbf{P})(\mathbf{M} - \mathbf{P})'\right] \Big/ \sum_{i=1}^{n}(1 + F)}$$

where **M** is the matrix $n \times m$ (n = number of progenies and m = number of markers), such that $m_{ij} \in \{-1,0,1\}$ specifies the SNP genotype at each locus; **P** is the observed frequency matrix of the allele i (p_i), expressed as $2(p_i - 0.5)$; tr[] is the trace of a matrix; and F is the individual inbreeding coefficient

Table 1. List of environmental factors considered in the study, estimated from NASA orbital sensors (Stackhouse Jr., 2014).

Environmental factor	Unit of measurement
Average top-of-atmosphere insolation	MJ m^{-2} d^{-1}
Average insolation incident on a horizontal surface	MJ m^{-2} d^{-1}
Average downward longwave radiative flux	MJ m^{-2} d^{-1}
Average air temperature at 2 m above the surface of the earth	°C
Minimum air temperature at 2 m above the surface of the earth	°C
Maximum air temperature at 2 m above the surface of the earth	°C
Relative air humidity at 2 m above the surface of the earth	%
Dew–frost point temperature at 2 m above the surface of the earth	°C
Average precipitation	mm d^{-1}
Wind speed at 10 m above the surface of the earth	m s^{-1}

derived from the pedigree for n genotyped progenies. With such a strategy, matrices $\mathbf{G}_w$ and $\mathbf{A}_{22}$ have very similar diagonal element averages. The **H**-matrix was normalized by dividing it by the mean of its main diagonal, which allows the same expectation of variance for breeding values of genotyped and nongenotyped progenies.

Phenotypic Data Analyses

For fitting phenotypic values per environment, single-environment analyses were performed using the following mixed linear model:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}_1\mathbf{r} + \mathbf{Z}_2\mathbf{b} + \mathbf{Z}_3\mathbf{p} + \mathbf{e} \quad [1]$$

where $\mathbf{y}$ is the vector of phenotypic data for the trait; β is the vector of fixed effects (i.e., intercept, genotype group [one group of check cultivars and another group of progenies], and effect of check cultivars); $\mathbf{r}$ is the vector of random effects of replication, with $\mathbf{r} \sim \mathrm{MN}\left(0, \sigma_r^2\mathbf{I}\right)$, where MN stands for multivariate normal, and $\mathbf{I}$ is the identity matrix; $\mathbf{b}$ is the vector of random effects of blocks within replication, with $\mathbf{b} \sim \mathrm{MN}\left(0, \sigma_b^2\mathbf{I}\right)$; $\mathbf{p}$ is the vector of random effects of progenies, $\mathbf{p} \sim \mathrm{MN}\left(0, \sigma_g^2\mathbf{I}\right)$; $\mathbf{X}$ and $\mathbf{Z}_1$ through $\mathbf{Z}_3$ are the respective incidence matrices for these effects; and $\mathbf{e}$ is the vector of random residual error, with $\mathbf{e} \sim \mathrm{MN}\left(0, \sigma_e^2\mathbf{I}\right)$.

To evaluate the magnitude of G × E, we obtained the Type B genetic correlations (Yamada, 1962) between pairs of environments by adjusting the following mixed linear model:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{r} + \mathbf{Z}_3\mathbf{b} + \mathbf{Z}_4\mathbf{p} + \mathbf{Z}_5\mathbf{d}_1 + \mathbf{Z}_6\mathbf{d}_2 + \mathbf{Z}_7\mathbf{d}_3 + \mathbf{e} \quad [2]$$

where, in addition to the terms already defined from Eq. [1], $\mathbf{a}$ is the vector of random effects of environments, with $\mathbf{a} \sim \mathrm{MN}\left(0, \sigma_a^2\mathbf{I}\right)$; $\mathbf{d}_1$ is the vector of random effects of interaction between genotype groups and environments, with $\mathbf{d}_1 \sim \mathrm{MN}\left(0, \sigma_{va}^2\mathbf{I}\right)$; $\mathbf{d}_2$ is the vector of random effects of interaction of progenies with environments, with $\mathbf{d}_2 \sim \mathrm{MN}\left(0, \sigma_{ga}^2\mathbf{I}\right)$; and $\mathbf{d}_3$ is the vector of random effects of interaction of check cultivars with environments, with $\mathbf{d}_3 \sim \mathrm{MN}\left(0, \sigma_{ca}^2\mathbf{I}\right)$.

To estimate genetic and phenotypic correlations between traits, a multivariate version of the mixed linear model described in Eq. [2] was used to obtain genetic (**H**) and residual (**R**) matrices of (co)variance associated with the progenies $\mathbf{p}_i \sim \mathrm{MN}\left(0, \mathbf{G} \otimes \mathbf{I}\right)$ and residual effects $\mathbf{e}_i \sim \mathrm{MN}\left(0, \mathbf{R} \otimes \mathbf{I}\right)$, where $\otimes$ denotes the Kronecker product. This model was adjusted for the dataset per selection cycle and for the dataset involving all cycles.

All the models used to fit phenotypic values were adjusted via Bayesian approach, using that Gibbs sampler algorithm (Liu and Daniels, 2006), a Markov chain Monte Carlo (MCMC) method, to obtain genotypic values and variance and covariance components. We assumed normal distributions for the location parameters of fixed effects, multivariate normal distribution for the data and location parameters of random effects, and inverse-Wishart distributions for variance and covariance components. Based on these assumptions, the fixed effect priors were normally distributed, with mean zero and variance equal to 10^8, thus fixing it in a prior large value. For variance and covariance parameters, we assigned a k-dimensional improper uniform (uninformative) inverse-Wishart distribution, by setting the hyperparameters $\nu = -(k + 1)$ and $\mathbf{S} = \mathrm{diag}(k) \times 0$, where ν is the degree of belief, $\mathbf{S}$ is the scale matrix, and k is the number of traits considered.

The use of such improper prior information leads to proper posterior distributions, as observed data contain enough information to surmount the prior information (Sorensen and Gianola, 2002). Thus, we obtained marginal posterior distributions equivalent to the restricted maximum likelihood (REML) procedure, to obtain variance and covariance components and to compute the BLUP (Yang, 2007) of genotypic values. In the Gibbs sampling setting, based on previous analyses of some MCMCs, we considered 210,000 samples collected from the posterior distribution after discarding 10,000 for burn-in, considering 20 as thinning. The conversion of the chains for each parameter was assessed by checking for potential autocorrelations of consecutive values in the chain, and its convergence was assessed by Geweke's convergence diagnostic (Geweke, 1992).

Based on marginal posterior distribution, we obtained posterior mean and 95% highest posterior density interval for the following parameters: coefficient of genetic variation, $\mathrm{CV}_g = \left[\sigma_g^2\right]^{0.5} / \overline{P}$; coefficient of relative variation, $\mathrm{CV}_r = \left[\sigma_g^2 / \sigma_e^2\right]^{0.5}$; broad-sense heritability based on progeny mean per environment, $h_{\overline{p}}^2 = \sigma_g^2 / \left(\sigma_g^2 + \sigma_e^2/k\right)$; selective accuracy, $r_{gg} = \left[h_{\overline{p}}^2\right]^{0.5}$; and Type B genetic correlation between environments, $r_b = \sigma_g^2 / \left(\sigma_g^2 + \sigma_{ga}^2\right)$ (Yamada, 1962). In these equations, σ_g^2, σ_{ga}^2, and σ_e^2 are marginal posterior distributions for components of genetic variance among progenies, variance to progeny × environment interaction, and residual variance between plots, respectively; $\overline{P}$ is the grand mean of progenies; and k is the number of weighted replications in the unbalanced dataset.

Predictive Models

We extended the class of reaction norm-based models proposed by Jarquín et al. (2014) to obtain variance components associated with different sources of variation. Unlike Jarquín's original models, we used mixed models that assumed the effects of check cultivars common to different selection cycles as fixed and all the other effects as random. Additionally, our approach makes use of the combined relationships (**H**-matrix), instead of marker-based (**G**-matrix) or pedigree-based (**A**-matrix) relationships. For all models, homogeneous residual variance was assumed, as genomic prediction under homogeneous or moderately heterogeneous variances between environments is highly correlated. The adjusted models, all characterized as RN-HBLUP models, are described as follows:

1. M1 (ETP): model including effects of environments (E_i, location-year combination), progenies (P_j) and check cultivars (T_k):

$$\hat{y}_{ijk} = \mu + E_i + P_j + T_k + \varepsilon_{ijk} \quad [3]$$

where $\hat{y}_{ijk}$ is the genotypic value (BLUP) previously obtained; μ is the intercept; $E_i \sim \mathrm{NIID}\left(0, \sigma_E^2\right)$ is the random effect of the environment i; $P_j \sim \mathrm{NIID}\left(0, \sigma_g^2\right)$ is the random effect of the progeny j; T_k is the fixed effect of the check cultivar k; and $\varepsilon_{ijk} \sim \mathrm{NIID}\left(0, \sigma_\varepsilon^2\right)$ is the random residual effect; NIID() stands for normal independent and identically distributed. As the effects of both progenies and environments are

independent, there is no information sharing between progenies or between environments.

2. M2 (ETH): derived from model M1 (Eq. [3]), with the effect of progeny (P_j) replaced by the effect of the combined relationship (**H**-matrix):

$$\hat{y}_{ijk} = \mu + E_i + a_j + T_k + \varepsilon_{ijk} \qquad [4]$$

where $a_j \sim MN(\mathbf{0}, \mathbf{H}\sigma^2_a)$ is the additive random effect of progeny j, with vector $\mathbf{a} = (a_1, \ldots, a_n)'$, which accounts for the additive relationship; thus, there is information sharing across progenies.

3. M3 (ETPH): combination of the models M1 (Eq. [3]) and M2 (Eq. [4]), including effects of progenies and combined relationship (**H**-matrix):

$$\hat{y}_{ijk} = \mu + E_i + P_j + a_j + T_k + \varepsilon_{ijk} \qquad [5]$$

Although the **H**-matrix captures a large portion of genetic variation among progenies, the remainder of the variation can be captured by progeny effects. Imperfect linkage disequilibrium between markers and causal loci and pedigree or genotyping errors can inflate the variation not captured by the **H**-matrix.

4. M4 (ETPHW): extension of model M3 (Eq. [5]) to incorporate EC information through the $\mathbf{W} = \{w_{ij}\}$ matrix, where $w_{ij} = \sum_{q=1}^{Q} W_{ijq}\lambda_q$ corresponding to the regression on EC, where Q is the total number of ECs, W_{ijq} is the measure of the qth EC in the environment–genotype ij combination, and λ_q is the effect of the qth EC, with $\lambda_q \sim NIID(0, \sigma^2_\lambda)$. Therefore, model M4 becomes:

$$\hat{y}_{ijk} = \mu + E_i + P_j + a_j + w_{ij} + T_k + \varepsilon_{ijk} \qquad [6]$$

where $\mathbf{w} \sim MN(0, \mathbf{\Omega}\sigma^2_w)$. In this way, the $\mathbf{w} = \mathbf{W}\lambda$ vector follows normal multivariate distribution, with zero mean and covariance matrix $\mathbf{\Omega}\sigma^2_w$; where $\mathbf{\Omega}$ is the matrix of similarity among environmental conditions. Through this model, there is information sharing across environments and across progenies.

5. M5 (ETPHW-H×E): extension of model M4 (Eq. [6]) to incorporate the interaction between additive effect of progeny (a_j) and effect of environment (E_i), here represented by aE_{ij}. According to Jarquín et al. (2014), based on covariance functions (Kempthorne, 1954), the structure of the interaction aE_{ij} may be generated by the Hadamard product $\left[\mathbf{Z}_p\mathbf{H}\mathbf{Z}_p'\right] \circ \left[\mathbf{Z}_E\mathbf{Z}_E'\right]$, where $\mathbf{Z}_p$ is the incidence matrix for additive genetic effects, which connects progeny observations with the **H**-matrix; and $\mathbf{Z}_E$ is the incidence matrix for environmental effects, which connects progeny observations to environments. Model M5 becomes:

$$\hat{y}_{ijk} = \mu + E_i + P_j + a_j + w_{ij} + aE_{ij} + T_k + \varepsilon_{ijk} \qquad [7]$$

where $aE_{ij} \sim MN\left[\mathbf{0}, \left(\mathbf{Z}_p\mathbf{H}\mathbf{Z}_p'\right) \circ \left(\mathbf{Z}_E\mathbf{Z}_E'\right)\sigma^2_{aE}\right]$ is the random effect of the interaction between additive effect of progeny j and effect of environment i.

6. M6 (ETPHW-H×E-H×W): extension of model M5 (Eq. [7]) to incorporate the interaction between additive effect of progeny (a_j) and effect of EC (w_{ij}), represented by aw_{ij}. Like the strategy used for the interaction aE_{ij}, the covariance structure of the interaction aw_{ij} may be generated by the Hadamard product $\left[\mathbf{Z}_p\mathbf{H}\mathbf{Z}_p'\right] \circ \mathbf{\Omega}$. Thus, model M6 becomes:

$$\hat{y}_{ijk} = \mu + E_i + P_j + a_j + w_{ij} + aE_{ij} + aw_{ij} + T_k + \varepsilon_{ijk} \qquad [8]$$

where $aw_{ij} \sim MN\left[0, \left(\mathbf{Z}_p\mathbf{H}\mathbf{Z}_p'\right) \circ \mathbf{\Omega}\sigma^2_{aw}\right]$ is the random effect of the interaction between the additive effect of progeny j and the effect of EC i.

To obtain estimates of variance components for each source of variation of random effect, the six prediction models were adjusted for each trait, considering the entire dataset (i.e., without cross-validation [CV]). Prior to adjustment, genotypic values (BLUP) were centralized and standardized across the environments to obtain estimates of variance components, with sum equal to one. We calculated the magnitudes of the variance components within environments (i.e., the percentage of the variation within environments, not considering the variance component for environment). We used this strategy to evaluate the ability of models to predict progeny performance in each environment.

Predictive Accuracy of Models

We assessed predictive accuracy of each model using training-testing random partitions of the dataset as a CV procedure (Pérez-Rodríguez and de los Campos, 2014). For this scheme, after previous evaluations of different random partitions, 50 random partitions were sampled, assigning 90% of the progenies to the training set and 10% to the testing set. Check cultivars were not allocated to the random partitions of the testing set to avoid unbalance in check cultivar effects while fitting the prediction models. The same training-testing partitions were used to assess the prediction accuracy of each model and, for each partition, models were fitted to the training set.

As proposed by Burgueño et al. (2012), we considered two CV schemes for prediction of progenies: CV1 to predict the performance of untested progenies (i.e., those that were not evaluated in any environment), and CV2 to predict the performance of progenies that were tested in some environments, but not in others. CV1 represents the case of progenies not included in field trials or when total phenotypic information is missing, and CV2 represents the frequent problem of unbalanced data between environments. In CV1, progeny information was randomly assigned to the partitions, so that all data for a progeny were assigned to the same partition. In CV2, plot observations were randomly assigned to partitions, so that information from a progeny could be assigned to different partitions.

Like the approach used by Jarquín et al. (2014), we wanted to evaluate the ability of models to predict progeny performance in each environment. Therefore, the predictive accuracy of each model was estimated after correcting the genotypic values ($\hat{y}_{ijk}$, BLUP) and the breeding values ($\hat{g}_{ijk}$, EBV) derived from the training set for main effects of environment and EC. The

correction vector corresponded to the sum of the estimates of the main effects $\left(\hat{n}_{ij} = \hat{E}_j + \hat{w}_{ij}\right)$ derived from model M6 (Eq. [8]), and the corresponding partition in the CV scheme (CV1 or CV2). The predictive accuracy of each model and CV scheme was assessed via the Pearson's product-moment correlation between corrected genotypic values $\left(\tilde{\hat{y}}_{ijk} = \hat{y}_{ijk} + \hat{n}_{ij}\right)$ and corrected breeding values $\left(\tilde{\hat{g}}_{ijk} = \hat{g}_{ijk} + \hat{n}_{ij}\right)$ in the testing set, within environment. Based on 50 correlation estimates for each fitted model, we computed the average accuracy and standard error. All models were evaluated for phenotype prediction in relation to both CV schemes, the predictions having been based on 55,000 samples collected from the posterior distribution after discarding 5000 for burn-in and considering four as thinning.

We also evaluated the bias associated with the prediction for each model-trait combination. We defined the linear regression coefficient (b, slope) of $\tilde{\hat{y}}_{ijk}$ on $\tilde{\hat{g}}_{ijk}$ as a measurement of the degree of bias of a model. In this case, $b = 1$ indicates no bias, $b < 1$ indicates overestimation of EBVs (i.e., more variance among the predicted values), and $b > 1$ indicates underestimation of EBVs.

Prediction of Untested Progenies in Tested Environments

To evaluate the efficiency of multienvironment and single-environment models for phenotype prediction of untested progenies for each environment, we estimated the predictive accuracy via two approaches: (i) single-environment HBLUP (M_single) using only combined relationships (**H**-matrix) for phenotype prediction of progenies in target environments; and (ii) model M6, for prediction of phenotypes in target environments. For the second approach, two strategies of combinations of environments were tested, with dataset within the cycle of the target environment (model M6_within_cycle), and with a general dataset across cycles (model M6_across_cycles). Accuracy of phenotype prediction of untested progenies was computed with the CV1 scheme. Partitioning of the progenies into the testing set was common to all models, corresponding to the progeny sample within each target environment.

Prediction of Progenies in Untested Environments

We calculated the accuracy of phenotype prediction of progenies in untested environments (Lado et al., 2016; i.e., environments without phenotyping) using model M6 based on two strategies of combinations of environments: (i) a dataset from all environments of the respective cycle, except the untested environment (model M6_within_cycle); and (ii) a dataset from all environments of all cycles, except the untested environment (model M6_across_cycles). For each untested environment, predictive accuracy was obtained by computing Pearson's product-moment correlation between all genotypic values ($\hat{y}_{ijk}$, BLUP) and estimated breeding values ($\hat{g}_{ijk}$, EBV) of the respective untested environment. This procedure was performed only once for each environment (i.e., without CV).

Software and Statistical Packages

All statistical analyses were implemented in the R platform (R Core Team, 2017). The following packages were used: 'MCMCglmm' for fitting phenotypic data in mixed models through Bayesian inference; 'coda' for the diagnosis of Markov chains; a modified version of the package 'pedigreemm' (Bates and Vazquez, 2009), accounting for self-pollination, for estimation of the **A**-matrix; 'pedigreeR', for estimation of inbreeding coefficients (F) from pedigree; and 'BGLR' (Pérez-Rodríguez and de los Campos, 2014) to fit genome-enabled prediction models.

RESULTS

Descriptive Statistics

The dataset was highly unbalanced for all phenotypic traits (Table 2). The lowest number of entries was evaluated in Cycle 1, and the largest number in Cycle 2. Although the progenies evaluated were not common to cycles, common check cultivars allowed connecting the dataset and adjusting means for environmental and cycle effects. Within each cycle, the Type B genetic correlation for genotypes between pairs of environments varied widely, indicating large variability in G × E effects for these traits.

The distribution of the standardized data per environment (not adjusted for environmental effect) varied throughout the cycles, especially for GY and DTF (Fig. 1). For GY, higher variation was observed in Cycle 1 than in Cycle 2, whereas for DTF, lower variation was observed in Cycle 3, which indicates that selection reduced phenotypic variance for these traits across cycles. The median of distributions indicated a tendency to increase GY and reduce DTF, suggesting genetic gain from selection. Despite the high variation in the distribution of these traits across cycles, the overall empirical distribution of each trait was approximately normal (data not shown), allowing the use of Gaussian distribution models.

No population structure was verified in the singular value decomposition of the combined relationship (**H**-matrix) by the eigenvalues and vector charges of the first two eigenvectors (see Fig. 2, Panels i and ii). The fact that 44% of the total variation was explained by the first two eigenvectors and 80% of the variation was explained by 10 eigenvectors suggested that the population was not highly diverse. This result was not surprising, considering that CNA12S was a closed population, synthesized by backcrossing and recombined via circulant diallel among progenies. For the variation attributed to the matrix of environmental similarity (**Ω**), the first two eigenvectors explained ~68% of the total variation (Fig. 2, Panels iii and iv). This can be explained by the limited number of environments sampled or by the great importance of some environmental factors.

Broad-sense heritability based on progeny-mean ranged from 0.28 to 0.63 for GY, 0.44 to 0.86 for PH, and 0.83 to 0.99 for DTF (Appendix A in the supplemental material). The values of relative coefficient of

Table 2. Number of $S_{1:3}$ progenies and check cultivars per environment (diagonal, bold) and common to pairs of environments (above diagonal), and Type B genetic correlation between environments (below diagonal), for grain yield (GY), plant height (PH), and days to flowering (DTF) in the CNA12S population.

		Cycle 1‡				Cycle 2‡		Cycle 3			
Trait	Environment†	ALE-2005	CAC-2005	CLE-2005	ITA-2005	CAC-2009	STM-2009	ALE-2014	ALE-2015	CLE-2015	GOI-2015
	ALE-2005	**111**	110	107	108	4	4	4	2	2	2
	CAC-2005	0.71	**133**	127	130	4	4	4	2	2	2
	CLE-2005	0.55	0.86	**128**	125	4	4	4	2	2	2
	ITA-2005	0.37	0.84	0.76	**131**	4	4	4	2	2	2
GY	CAC-2009	–	–	–	–	**277**	260	4	2	2	2
	STM-2009	–	–	–	–	0.74	**301**	4	2	2	2
	ALE-2014	–	–	–	–	–	–	**215**	198	198	198
	ALE-2015	–	–	–	–	–	–	0.78	**200**	200	200
	CLE-2015	–	–	–	–	–	–	0.36	0.39	**200**	200
	GOI-2015	–	–	–	–	–	–	0.30	0.29	0.20	**200**
	ALE-2005	**113**	–	202	113	4	4	4	2	2	2
	CAC-2005	–	–	–	–	4	4	4	2	2	2
	CLE-2005	0.55	–	**132**	132	4	4	4	2	2	2
	ITA-2005	0.51	–	0.76	**134**	4	4	4	2	2	2
PH‡	CAC-2009	–	–	–	–	–	–	4	2	2	2
	STM-2009	–	–	–	–	–	**299**	4	2	2	2
	ALE-2014	–	–	–	–	–	–	**225**	198	198	198
	ALE-2015	–	–	–	–	–	–	0.89	**200**	200	200
	CLE-2015	–	–	–	–	–	–	0.71	0.82	**200**	200
	GOI-2015	–	–	–	–	–	–	0.76	0.24	0.68	**200**
	ALE-2005	**125**	86	120	125	4	4	4	2	2	2
	CAC-2005	0.58	**88**	83	88	4	4	4	2	2	2
	CLE-2005	0.67	0.45	**127**	127	4	4	4	2	2	2
	ITA-2005	0.72	0.46	0.55	**134**	4	4	4	2	2	2
DTF	CAC-2009	–	–	–	–	**320**	315	4	2	2	2
	STM-2009	–	–	–	–	0.68	**315**	4	2	2	2
	ALE-2014	–	–	–	–	–	–	**225**	198	198	198
	ALE-2015	–	–	–	–	–	–	0.92	**200**	200	200
	CLE-2015	–	–	–	–	–	–	0.89	0.87	**200**	200
	GOI-2015	–	–	–	–	–	–	0.83	0.81	0.80	**200**

† ALE, Alegrete, Rio Grande do Sul; CAC, Cachoeirinha, Rio Grande do Sul; CLE, Capão do Leão, Rio Grande do Sul; ITA, Itajaí, Rio Grande do Sul; STM, Santa Maria, Rio Grande do Sul; GOI, Goianira, Goiás.

‡ In the CAC-2005 and CAC-2009 environments, plant height (PH) was not evaluated.

variation ranged from 0.38 to 1.41 for GY, 0.67 to 2.95 for PH, and 1.50 to 10.4 for DTF. The lowest values of relative coefficient of variation, heritability, and selective accuracy for GY were observed in Cycle 2. Its magnitudes, however, were approximately maintained between the selection Cycles 1 and 3, as it did not differ significantly, given the overlapping 95% highest posterior density intervals.

Components of Variance

The main effect of environments made the largest contribution to the phenotypic variance for all traits (Appendices B, C, and D in the supplemental material). For GY, this contribution varied between 82% (model M1) and 66% (model M6). For DTF, the EC effects accounted for >70% of the phenotypic variance. This indicated that DTF was strongly influenced or controlled by environmental effects, possibly because we considered the variability within environment by maturity group. When incorporating EC effects into the models, there was a reduction in variance attributable to environment, for all traits, indicating that ECs were able to explain a considerable portion of the environmental variation.

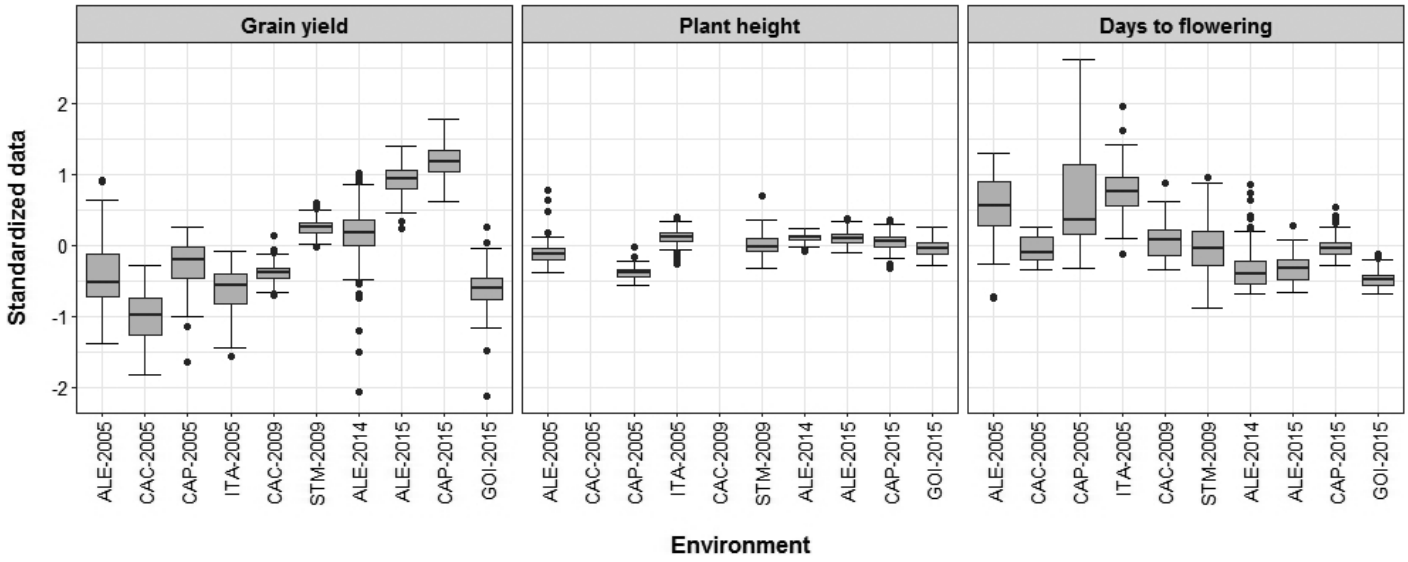

Fig. 1. Boxplot of standardized data for grain yield, plant height, and days to flowering of $S_{1:3}$ progenies of the CNA12S population, evaluated in 10 environments. In the CAC-2005 and CAC-2009 environments, plant height was not evaluated. See Table 2 for definition of environment abbreviations.

Accuracy of the Predictive Models

Predictive accuracy of each RN-HBLUP model, for each CV scheme (CV1 and CV2) is presented in Fig. 3. As expected, the accuracy was lower in CV1, because in this case, no information was available on progeny phenotypes across environments. In model M1, which did not use the **H**-matrix, negative values for accuracy were detected with CV1. On the other hand, when **H**-matrix was considered in model M2, there was a considerable increase in accuracy, even in CV1.

The most significant change in accuracy in both CV schemes resulted from the inclusion of EC effects in model M4 (Fig. 3), especially for DTF. A marked effect of ECs on DTF was expected, because the exploitation of the phenotypic variation within environment, by the definition of maturity groups, contributed to the increase in accuracy. The increase in predictive accuracy attributable to inclusion of EC effects was higher in CV1 than in CV2, because the prediction of progenies in the absence of any genetic information, as in CV1, is strongly benefited by the modeling of intra- and interenvironmental variations.

The models M5 and M6 were approximately equivalent for all traits with CV1, which indicated that consideration of combined relationship × environment interaction (H × E) effects (model M5) had a similar result as with model M6 that included H × E and combined relationship × environmental conditions interaction (H × W) effects (Fig. 3). On the other hand, in CV2, there was no increase in accuracy attributable to H × E effects, although a significant increase in accuracy occurred because of H × E and H × W effects for GY and PH in model M6. In this case, the rate of increase in

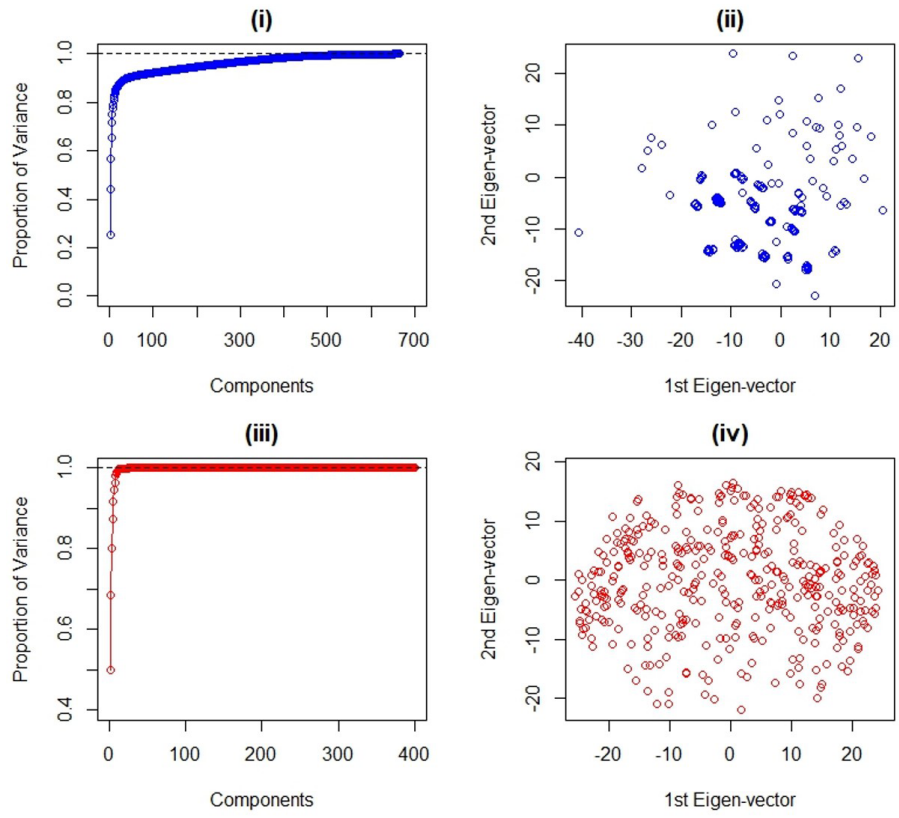

Fig. 2. Scree plot of eigenvalues and loadings of the first two eigenvectors for (Panels i and ii) the combined relationship **H**-matrix and (Panels iii and iv) the environmental similarity **Ω**-matrix, associated with the $S_{1:3}$ progenies of the CNA12S population, evaluated in 10 environments.

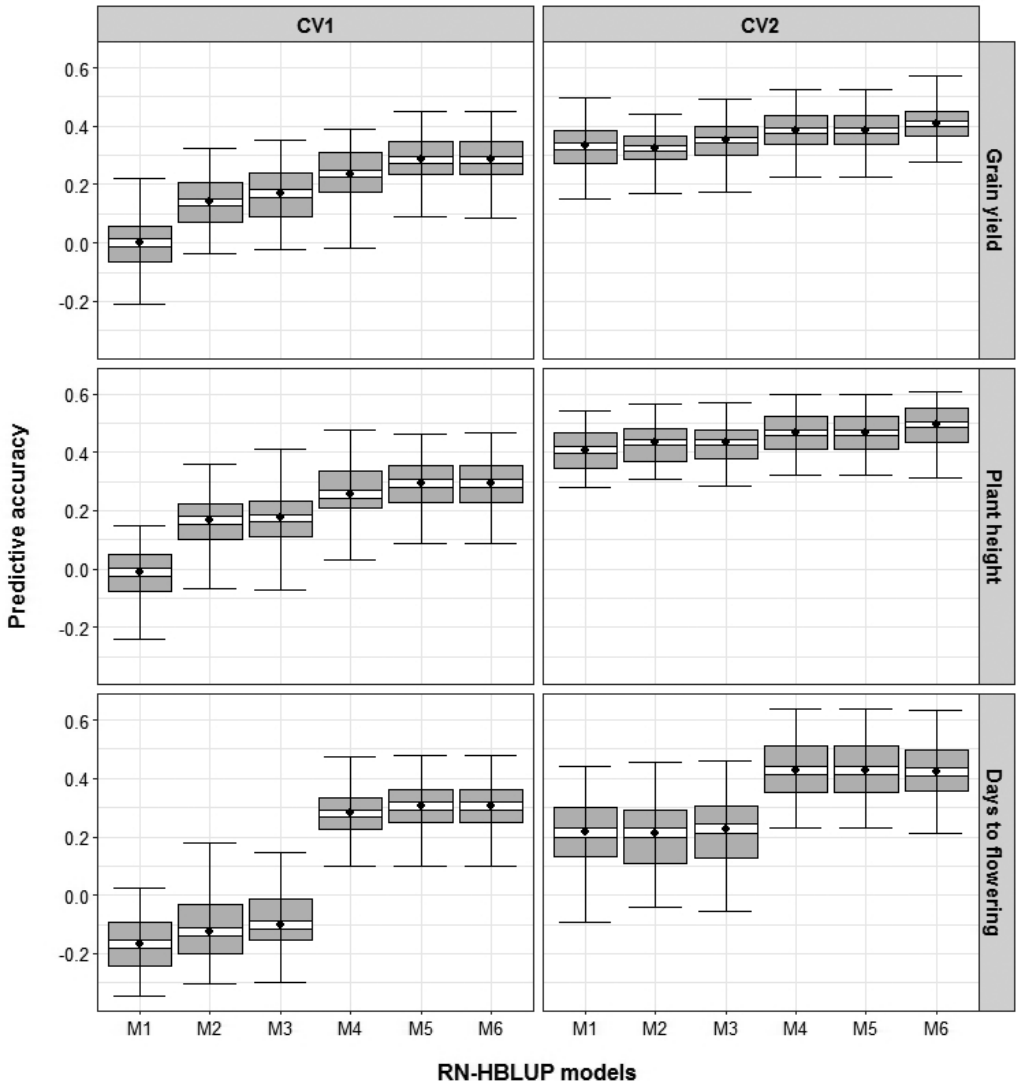

Fig. 3. Boxplot of predictive accuracy (average of correlations between genotypic values, best linear unbiased prediction [BLUP], and estimated breeding values [EBV] within environments) for six single-step BLUP-based reaction norm (RN-HBLUP) models (M1 to M6, see Materials and Methods) with two cross-validation schemes (CV1 and CV2) for grain yield, plant height, and days to flowering. Black dots indicate the mean, and white bands inside gray boxes represent the range (mean ± SE).

accuracy from model M6 in relation to model M4 was ~6% for GY and PH. In CV2, DTF was the only trait that did not benefit from H × E or H × W effects, although DTF benefited greatly from the EC effects.

Bias of the Predictive Models

As a measure of bias of the prediction, we calculated the regression coefficient (*b*) by regressing the genotypic values for the validation set on their EBVs for each model-trait combination and CV scheme (CV1 and CV2) (Table 3). In both CV schemes, the most significant biases in prediction were observed for models without EC effects and/or interaction terms (models M1, M2, and M3), especially for DTF, which presented a highly biased overestimation of EBVs. For the other models, the regression coefficients did not differ from one, indicating no significant bias in the prediction of EBVs for GY and PH with CV1 scheme and for PH with CV2 scheme. Therefore, with the inclusion of EC effects and/or interaction terms in RN-HBLUP models, there was a trend of underestimating EBVs for some traits.

Prediction of Untested Progenies

Regarding the accuracy in predicting phenotypes of untested progenies in each environment, complete RN-HBLUP models (M6_within_cycle and M6_across_cycles) presented higher accuracy than the single-environment model (M_single), especially for GY and DTF (Fig. 4 and 5). The superiority of RN-HBLUP was more evident for GY in the case of unreplicated trials in Cycles 1 and 2 (ALE-2005, CAC-2005, CLE-2005, ITA-2005, CAC-2009, and STM-2009). In all environments, accuracy of models M6_within_cycle and M6_across_cycles did not differ, except in environments of Cycle 3 for PH, where M6_within_cycle showed superiority (Fig. 4). For DTF, RN-HBLUP models showed marked superiority over single-environment models in all environments, with accuracy >0.75. The greater efficiency of this model in predicting phenotypes of progenies might be attributable to the phenotypic variability taken into account within environment by the use of EC for each environment, according to the maturity group of genotypes.

Table 3. Slope (regression coefficient) and standard error for the linear regression of best linear unbiased prediction (BLUP) values of the validation set on their estimated breeding values (EBVs) as a measurement of the bias of each single-step BLUP-based reaction norm (RN-HBLUP) model, according to cross-validation scheme (CV1 and CV2) for grain yield (GY), plant height (PH), and days to flowering (DTF).

	CV1			CV2		
Model†	GY	PH	DTF	GY	PH	DTF
M1 (ETP)	0.16** (0.149)	−0.38** (0.148)	−0.11** (0.018)	0.10** (0.133)	0.08** (0.124)	−0.11** (0.022)
M2 (ETH)	0.49** (0.069)	0.33** (0.044)	0.01** (0.015)	0.65** (0.065)	0.44** (0.050)	0.00** (0.023)
M3 (ETPH)	0.58** (0.092)	0.47** (0.064)	−0.03** (0.017)	0.69** (0.088)	0.62** (0.070)	−0.03** (0.022)
M4 (ETPHW)	1.11 (0.114)	0.88 (0.080)	1.39** (0.120)	1.38** (0.111)	1.02 (0.101)	1.24* (0.098)
M5 (ETPHW-H×E)	1.12 (0.089)	1.03 (0.068)	1.38** (0.116)	1.30** (0.104)	1.10 (0.096)	1.25** (0.094)
M6 (ETPHW-H×E-H×W)	1.11 (0.089)	1.04 (0.068)	1.38** (0.114)	1.30** (0.104)	1.11 (0.097)	1.25** (0.095)

*,** Significant at the 0.05 and 0.01 probability levels, respectively, according to Student *t* test.

† Components of variance: E, environments; P, progenies; H, combined relationship matrix; W, matrix of environmental covariates; H × E interaction matrix; H × W interaction matrix.

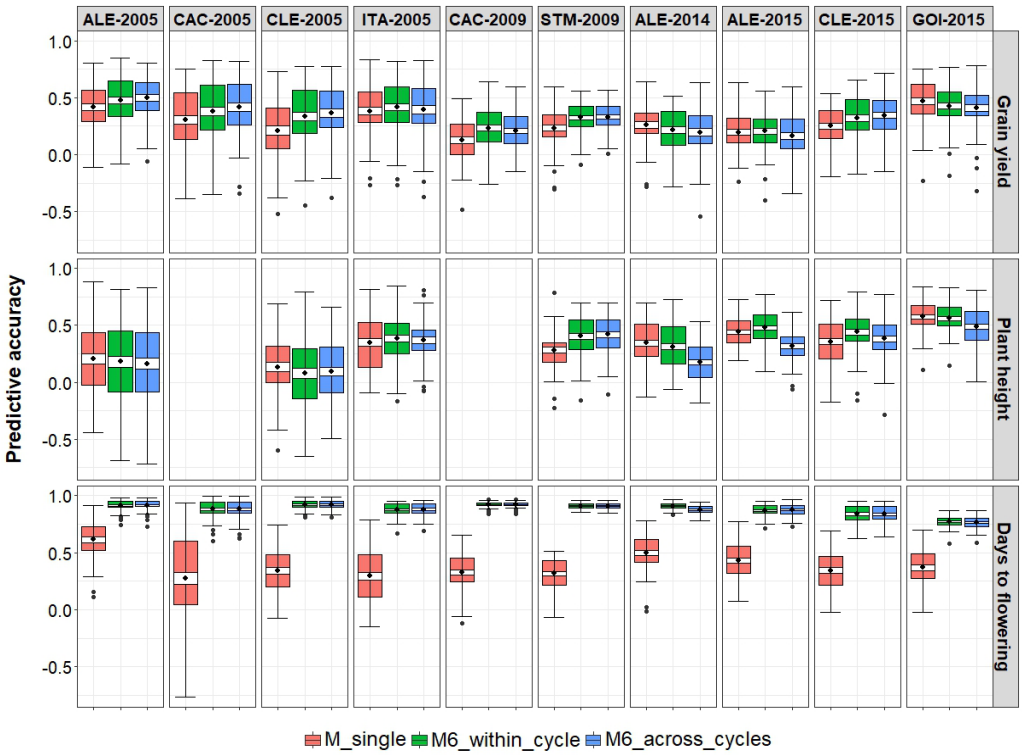

Fig. 4. Predictive accuracy of the single-environment single-step best linear unbiased prediction (HBLUP) model (M_single) and HBLUP-based reaction norm (RN-HBLIP) models (M6_within_cycle, and M6_across_cycles) for traits grain yield, plant height, and days to flowering per environment. M_single considers only the combined relationship matrix, M6 is the full model (ETPHW-H×E-H×W), within_cycle and across_cycles indicate the data used for model adjustment. Black dots in the center of each box indicate the mean, and white bands inside gray boxes represent the range (mean ± SE). See Table 2 for definition of environment abbreviations.

Prediction of Progenies in Untested Environments

The models M6_within_cycle and M6_across_cycles were equivalent in predicting progeny phenotype in most environments, although there were variations in some cases (Fig. 6). Model M6_within_cycle was sufficient to predict progeny phenotypes in untested environments; therefore, it is possible to generate accurate RN-HBLUP models even in the absence of data from previous cycles.

DISCUSSION

In RGS programs, progenies are tested in multiple environments, where G × E interaction has strong effects on quantitative traits. In this study, we demonstrated that G × E can be modeled using combined relationship (**H**-matrix) and EC information, based on the approach of reaction norm models (Jarquín et al., 2014) associated with the single-step procedure (Legarra et al., 2009). Thus, we evaluated a sequence of RN-HBLUP models for their prediction accuracy of three quantitative traits in rice. The models were applied to MET data in the context of RGS programs.

As demonstrated in earlier studies, the prediction accuracy in GS is associated with heritability of the trait, number of markers, relationship between training and testing sets, size of the training population, and G × E (de los Campos et al., 2009; Wray et al., 2013; Pérez-Rodríguez et al., 2015; Sukumaran et al., 2017). The better the quality of phenotypic data, the greater the accuracy of prediction models and the efficacy of the inference about the breeding values of progenies (Resende and Duarte, 2007). Therefore, evaluation of the experimental precision is a prerequisite to understanding the potential of a dataset for genomic prediction in a plant population. In

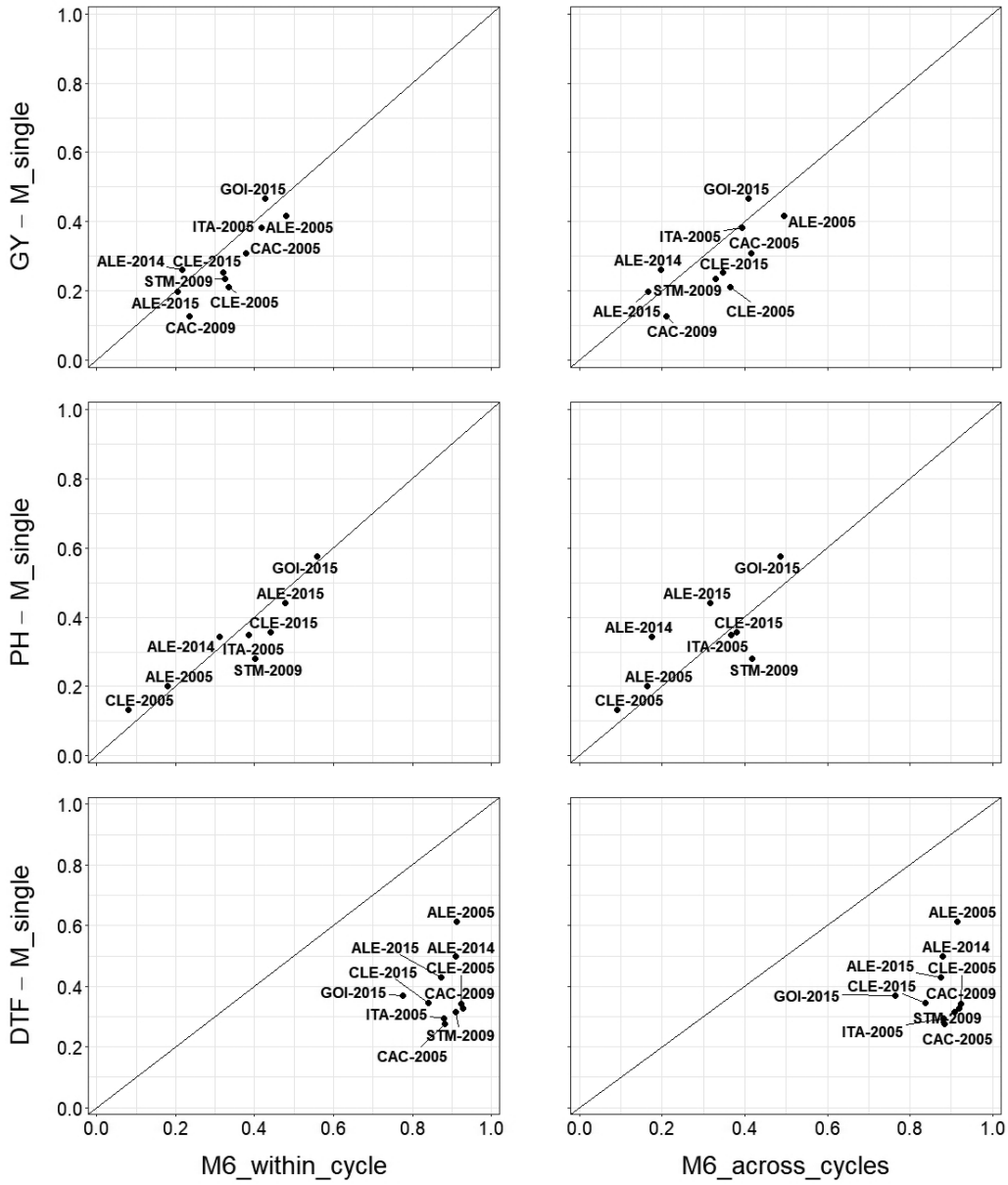

Fig. 5. Predictive accuracy of single-environment single-step best linear unbiased prediction (HBLUP) (M_single) and HBLUP-based reaction norm (RN-HBLIP) models (M6_within_cycle, and M6_across_cycles) for grain yield (GY), plant height (PH), and days to flowering (DTF), evaluated in 10 environments. Each point represents the mean of predictive accuracy. Information on the structure of the models is described in Fig. 5. See Table 2 for definition of environment abbreviations.

our study, selective accuracy values based on phenotypic data were moderate to high, although lower experimental precision for GY was observed in Cycle 2, when trials had a larger population size (N = 316) without replication of progenies (augmented block design).

The predictive accuracy per environment with the single-environment HBLUP model was correlated (r = 0.54, $p < 0.01$) with broad-sense heritability values based on progeny-mean for the three traits (Fig. 7). This was expected, because the heritability of a trait corresponds to the upper limit of the phenotypic variance explained by a linear genetic prediction model (Wray et al., 2013). Traits with higher heritability present higher genetic variance and, hence, are more predictable via a genomic approach. This result highlights the importance of precise phenotypic data for training RN-HBLUP models for accurate genomic prediction.

An increase in predictive accuracy by accounting for additive relationships, whether in single-environment or in multienvironment models, was also observed in other studies (Burgueño et al., 2012; Pérez-Rodríguez et al., 2015, 2017; Jarquín et al., 2014; Ashraf et al., 2016; Sukumaran et al., 2017). In the CV1 scheme (absence of phenotypic information for untested progenies), the influence of the **H**-matrix was even more relevant. Therefore, pedigree or DNA marker information is very useful for RGS programs when part of the phenotypic data is missing or has poor quality, especially in unbalanced designs with a few or no replications of progenies. In addition, marker information allows accounting for genetic variance not explained by pedigree because of the binary sampling variance of the Mendelian segregation. In fact, the use of pedigree data, along with DNA marker data, can enhance the predictive accuracy, regardless of the genetic architecture of the trait (de los

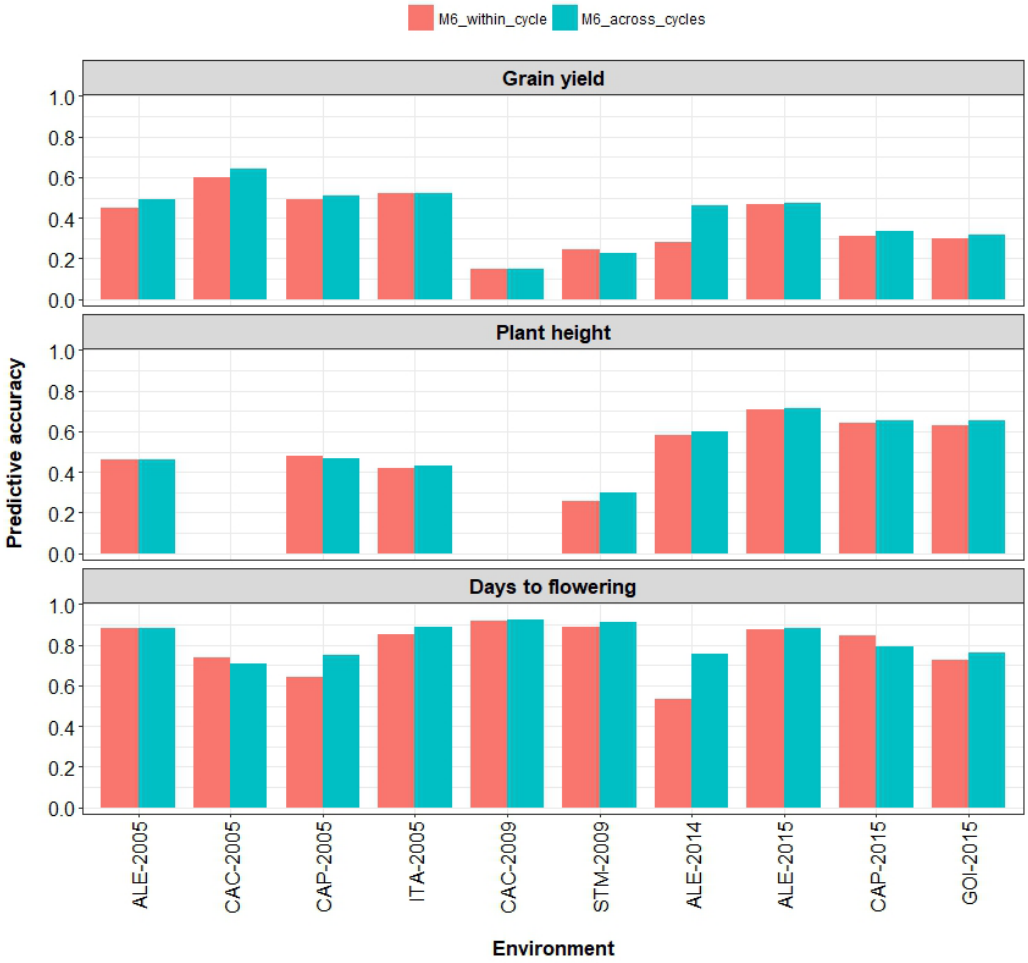

Fig. 6. Predictive accuracy of single-step best linear unbiased prediction-based reaction norm (RN-HBLUP) models M6 (within_cycle and across_cycles) for grain yield, plant height, and days to flowering to predict the phenotype in each untested environment. For details of models, see Materials and Methods. See Table 2 for definition of environment abbreviations.

Campos et al., 2009; Ashraf et al., 2016; Sukumaran et al., 2017).

The reduction in residual variance by inclusion of H × E and H × W effects was evident in model M6 (Appendices B, C, and D in the supplemental material). This result indicated that a considerable portion of the variation among progenies and among environments could be captured by the interaction terms, justifying their inclusion in RN-HBLUP models, even if a portion of the variance explained by the interaction terms was small. The variance attributable to the main effects of progenies was similar to or lower than that attributable to combined relationship effects (**H**-matrix). However, the inclusion of both effects in the model increased the predictive accuracy (Fig. 3) and reduced the residual variance, in comparison with the model with only the **H**-matrix. This result favors the inclusion of progeny effects in more complete models, as done by Sukumaran et al. (2017), which, unlike other studies (Jarquín et al., 2014; Pérez-Rodríguez et al., 2015), used only additive relationships (**G**-matrix or **A**-matrix).

For all the traits evaluated, ECs and interactions (H × E and H × W) captured a significant portion of the phenotypic variation within environments. The EC component was the most relevant for increasing accuracy in both CV schemes (CV1 and CV2), possibly by capturing the variation

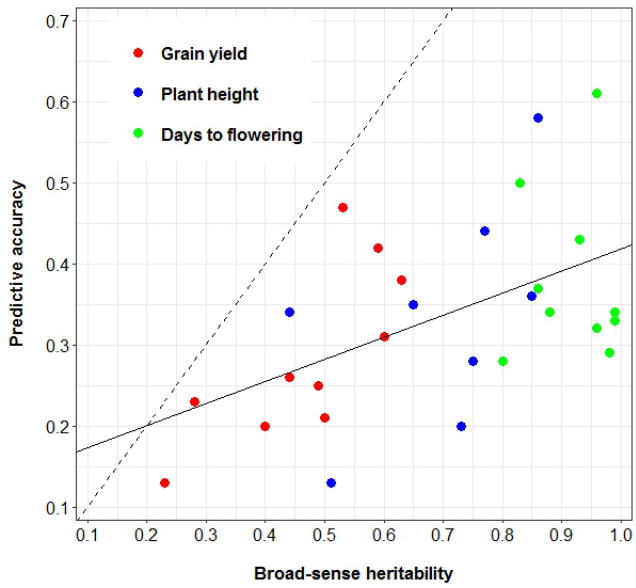

Fig. 7. Relationship between the predictive accuracy of the single-environment single-step best linear unbiased prediction model obtained per environment and the broad-sense heritability based on progeny-mean for grain yield, plant height, and days to flowering. The solid line is the trend line of the regression across traits, and the dashed line indicates the angular coefficient equal to one. Pearson's correlation coefficient (r) = 0.54 (p = 0.003).

among maturity groups. The main benefit of considering these effects (EC, H × E, and H × W) appears to be the reduction in residual variance, with a consequent increase in accuracy and a decrease in bias of the EBV prediction. Hence, it is advantageous to incorporate EC effects in mixed linear-bilinear models based on reaction norm whenever they are available for target environments (Jarquín et al., 2014). These models allow dealing with complex structures of unbalanced data and different covariance structures (Pérez-Rodríguez et al., 2017).

The gains in prediction accuracy obtained here by modeling G × E effects, in comparison with other models that do not consider these effects, agree with reports from other studies dealing with genomic prediction (Burgueño et al., 2012; Jarquín et al., 2014; Pérez-Rodríguez et al., 2015, 2017). Jarquín et al. (2014) suggested that considering EC effects, as well as their interactions in genetic relationships, was a promising strategy for increasing accuracy of phenotypic prediction with important implications for plant breeding. These authors highlighted that reaction norm models admitted only one mode of interaction, based on covariance functions (Kempthorne, 1954). Therefore, when interactions among genetic factors and environments assume different forms, this approach may have limitations in explaining the G × E component. In addition to not inducing variable selection, a limitation of this approach occurs when markers or ECs have large effects, a situation in which the shrinkage caused by Gaussian prior density may not be appropriate (Cuevas et al., 2016). Thus, additional research on modeling interactions between genetic relationships and EC is needed to improve the efficiency of multienvironment models for genomic prediction.

Multienvironment models allow prediction of progeny performance for all target environments simultaneously, in addition to exploring G × E to select potential progenies for wide adaptability and high stability (Burgueño et al., 2012). In this study, we found greater accuracy in multienvironment models than in single-environment models to predict GY and DTF of untested progenies in target environments. The RN-HBLUP model is notoriously superior to a single-environment model, even when both have similar accuracy, by allowing the phenotype prediction for all target environments simultaneously. The greater relevance of this model for GY was observed in Cycles 1 and 2, when the population was evaluated without replication of progenies. Therefore, this model has the potential to mitigate limitations attributable to less robust experimental designs with large numbers of segregating progenies with few or no replications, such as those normally used in preliminary yield trials. This potential of increasing accuracy to predict phenotypes comes from the information sharing across progenies and across correlated environments (Burgueño et al., 2012; Zhang et al., 2014; Lado et al., 2016). Therefore, the RN-HBLUP model with comprehensive structure (ETPHW-H×E-H×W) can predict phenotypes in environments where no progenies have been tested.

For phenotype prediction of progenies in untested environments, the use of data from all cycles did not increase accuracy when compared with use of data derived from a single cycle. The same was observed for phenotype prediction of untested progenies in targeted environments. In addition, when the genomic prediction is performed in population after recombining progenies selected from the training population, there is a reduction in accuracy, because the prediction must be based as much as possible on marker-quantitative trait locus linkage disequilibrium, rather than on genetic relationship (Habier et al., 2013). Therefore, within-cycle information, whether genetic (**H**-matrix) or EC data, is sufficient for training RN-HBLUP models. Lado et al. (2016) also obtained similar results, showing that the phenotype prediction of genotypes in untested environments was better performed within years. However, with greater volume of data generated through cycles, regarding the number of genotyped progenies and sampled environments, the predictive accuracy may increase because population size is a relevant factor to adjustment of genomic-enabled prediction models (Grattapaglia and Resende, 2011).

The predictive accuracy in our study was equal or superior to that observed in other populations of rice (Grenier et al., 2015; Onogi et al., 2015). Factors such as phenotypic data with higher heritability, larger population size, and higher linkage disequilibrium could explain the differences among studies (Grattapaglia and Resende, 2011). As this is the first study using RN-HBLUP models in rice, additional research with other populations and environments is needed. More research is also needed on the potential of RN-HBLUP models to predict genotype performance in untested environments in an entire geographic region that is the target of a breeding program. Genotype adaptability maps could be generated using estimated genotypic coefficients, which describe the sensitivities of genotypes to specific environmental conditions. With the increased capacity in high-throughput phenotyping, genotyping, and envirotyping, a large amount of high-quality data will be available, thus allowing the achievement of accurate RN-HBLUP models, which in turn could improve the efficiency of selection and release of cultivars.

CONCLUSIONS

Modeling progeny relationship from DNA marker and pedigree information, environmental covariates, and interactions between main effects increased the accuracy and decreased the bias of prediction by reaction norm models. Within-cycle data were sufficient for accurate phenotype predictions of untested progenies, even in untested environments. In this study, we have demonstrated that comprehensive reaction norm-based HBLUP models are a

promising tool to improve the predictive accuracy relative to quantitative traits in RGS programs in rice.

Conflict of Interest

The authors declare that there is no conflict of interest.

Supplemental Material Available

Supplemental material for this article is available online.

Author Contribution Statement

O.P. Morais Júnior contributed to experimental design, conducted some field trials, did statistical analyses, and drafted the manuscript. J.B. Duarte and F. Breseghello collaborated on the preparation and review of the manuscript. A.S.G. Coelho collaborated in some statistical analyses. O.P. Morais and A.M.M. Júnior contributed to experimental design and conducted some field trials. All authors have approved the final manuscript.

Acknowledgments

The authors dedicate this article to the memory of Dr. Orlando Peixoto de Morais, a close friend, wonderful person, outstanding scientist, and recognized mentor of the rice breeding program at Embrapa. We thank the entire rice breeding staff at Embrapa, especially the research assistants and field workers who performed the field trials, and the Capes/Brazilian Ministry of Education for the doctoral scholarship granted to the first author.

References

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

Allard, R.W., and A.D. Bradshaw. 1964. Implications of genotype-environmental interactions in applied plant breeding. Crop Sci. 4:503–508. doi:10.2135/cropsci1964.0011183X000400050021x

Ashraf, B., V. Edriss, D. Akdemir, E. Autrique, D. Bonnett, J. Crossa, et al. 2016. Genomic prediction using phenotypes from pedigreed lines with no marker data. Crop Sci. 56:957–964. doi:10.2135/cropsci2015.02.0111

Bates, D., and A. Vazquez. 2009. pedigreemm: Pedigree-based mixed-effects models. R Found. Stat. Comput. http://CRAN.R-project.org/package=pedigreemm (accessed 20 Apr. 2017)

Bernardo, R. 2010. Breeding for quantitative traits in plants. Stemma Press, Woodbury, MN.

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

Bustos-Korts, D., M. Malosetti, S. Chapman, and F. van Eeuwijk. 2016. Modelling of genotype by environment interaction and prediction of complex traits across multiple environments as a synthesis of crop growth modelling, genetics and statistics. In: X. Yin and P.C. Struik, editors, Crop systems biology, narrowing the gaps between crop modelling and genetics. Springer, Cham, Switzerland. p. 55–82. doi:10.1007/978-3-319-20562-5_3

Christensen, O.F., and M.S. Lund. 2010. Genomic prediction when some animals are not genotyped. Genet. Sel. Evol. 42:2–9. doi:10.1186/1297-9686-42-2

Counce, P., T.C. Keisling, and A.J. Mitchell. 2000. A uniform, objective, and adaptive system for expressing rice development. Crop Sci. 40:436–443. doi:10.2135/cropsci2000.402436x

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

Cuevas, J., J. Crossa, O.A. Montesinos-López, J. Burgueño, P. Pérez-Rodríguez, and G. de los Campos. 2016. Bayesian genomic prediction with genotype × environment interaction kernel models. G3 (Bethesda) 7:41–53. doi:10.1534/g3.116.035584

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

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

Eberhart, S.A., and W.A. Russell. 1966. Stability parameters for comparing varieties. Crop Sci. 6:36–40. doi:10.2135/cropsci1966.0011183X000600010011x

Federer, W.T. 1956. Augmented (hoonuiaku) designs. Hawaii. Plant. Rec. 55:191–208.

Forni, S., I. Aguilar, and I. Misztal. 2011. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. Genet. Sel. Evol. 43:1–7. doi:10.1186/1297-9686-43-1

Freitas, T.F.S., P.R.F. Silva, M.L. Strieder, and A.A. Silva. 2006. Validation of the development scale for Brazilian flooded rice cultivars. (In Portuguese, with English abstract.). Cienc. Rural 36:404–410. doi:10.1590/S0103-84782006000200008

Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: J.M. Bernardo, J. Berger, A.P. Dawid, and A.F.M. Smith, editors, Bayesian statistics. 4th ed. Oxford Univ. Press, Oxford, UK. p. 69–193.

Grattapaglia, D., and M.D.V. Resende. 2011. Genomic selection in forest tree breeding. Tree Genet. Genomes 7:241–255. doi:10.1007/s11295-010-0328-4

Grenier, C., T.V. Cao, Y. Ospina, C. Quintero, M.H. Châtel, J. Tohme, et al. 2015. Predictive ability of genomic selection in a rice synthetic population developed for recurrent selection breeding. PLoS One 10:e0136594. doi:10.1371/journal.pone.0136594 [erratum: 11:e0154976].

Habier, D., R.L. Fernando, and D.J. Garrick. 2013. Genomic BLUP decoded: A look into the black box of genomic prediction. Genetics 194:597–607. doi:10.1534/genetics.113.152207

Hasanuzzaman, M., K. Nahar, M.M. Alam, R. Roychowdhury, and M. Fujita. 2013. Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. Int. J. Mol. Sci. 14:9643–9684. doi:10.3390/ijms14059643

Henderson, C.R. 1976. A simple method for computing the inverse of the numerator relationship matrix used in prediction of breeding values. Biometrics 32:69–83. doi:10.2307/2529339

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

Jarquín, D., J. Crossa, X. Lacaze, P. Du Cheyron, J. Daucourt, J. Lorgeou, et al. 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

Kang, M.S. 1997. Using genotype-by-environment interaction for crop cultivar development. Adv. Agron. 62:199–252. doi:10.1016/S0065-2113(08)60569-6

Kang, M.S., and H.G. Gauch. 1996. Genotype-by-environment interaction. CRC Press, Boca Raton, FL.

Kang, M.S., and R. Magari. 1996. New developments in selecting for phenotypic stability in crop breeding. In: M.S. Kang and H.G. Gauch, editors, Genotype-by-environment interaction. CRC Press, Boca Raton, FL. p. 1–14.

Kempthorne, O. 1954. The correlation between relatives in a random mating population. Proc. R. Soc. Lond. B Biol. Sci. 143:103–113. doi:10.1098/rspb.1954.0056

Lado, B., P.G. Barrios, M. Quincke, P. Silva, and L. Gutiérrez. 2016. Modeling genotype × environment interaction for genomic selection with unbalanced data from a wheat breeding program. Crop Sci. 56:2165–2179. doi:10.2135/cropsci2015.04.0207

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

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

Liu, X.F., and M.J. Daniels. 2006. A new algorithm for simulating a correlation matrix based on parameter expansion and reparameterization. J. Comput. Graph. Stat. 15:897–914. doi:10.1198/106186006X160681

López-Cruz, M.A., J. Crossa, D. Bonnet, S. Dreisigacker, J. Poland, J. Jannink, et al. 2015. Increased prediction accuracy in wheat breeding trials using a markers x environment interaction genomic selection model. G3 (Bethesda) 5:569–582. doi:10.1534/g3.114.016097

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.

Misztal, I., A. Legarra, and I. Aguilar. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. J. Dairy Sci. 92:4648–4655. doi:10.3168/jds.2009-2064

Morais Júnior, O.P., F. Breseghello, J.B. Duarte, O.P. Morais, P.H.N. Rangel, and A.S.G. Coelho. 2017. Effectiveness of recurrent selection in irrigated rice breeding. Crop Sci. 57:3043–3058. doi:10.2135/cropsci2017.05.0276

Mota, R.R., P.S. Lopes, R.J. Tempelman, F.F. Silva, I. Aguilar, C.C.G. Gomes, and F.F. Cardoso. 2016. Genome-enabled prediction for tick resistance in Hereford and Braford beef cattle via reaction norm models. J. Anim. Sci. 94:1834–1843. doi:10.2527/jas.2015-0194

Onogi, A., O. Ideta, and Y. Inoshita. 2015. Exploring the areas of applicability of whole-genome prediction methods for Asian rice (*Oryza sativa* L.). Theor. Appl. Genet. 128:41–53. doi:10.1007/s00122-014-2411-y

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

Pérez-Rodríguez, P., and G. de los Campos. 2014. Genome-wide regression and prediction with the BGLR statistical package. Genetics 198:483–495. doi:10.1534/genetics.114.164442

Pérez-Rodríguez, P., J. Crossa, J. Rutkoski, J. Poland, R. Singh, A. Legarra, et al. 2017. Single-step genomic and pedigree genotype × environment interaction models for predicting wheat lines in international environments. Plant Genome 10:1–15. doi:10.3835/plantgenome2016.09.0089

R Core Team. 2017. R: A language and environment for statistical computing. R Found. Stat. Comput., Vienna, Austria.

Resende, M.D.V., and J.B. Duarte. 2007. Precision and quality control in cultivar trials. (In Portuguese, with English abstract.). Pesq. Agrop. Trop. 37:182–194.

Sorensen, D., and D. Gianola. 2002. Likelihood, Bayesian and MCMC methods in quantitative genetics. Springer, New York. doi:10.1007/b98952

Stackhouse Jr., P.W., 2014. Prediction of worldwide energy resource. NASA Langley Res. Ctr., Hampton, VA. http://power.larc.nasa.gov (accessed 20 June 2016).

Sukumaran, S., J. Crossa, D. Jarquín, M. Lopes, and M.P. Reynolds. 2017. Genomic prediction with pedigree and genotype × environment interaction in spring wheat grown in South and West Asia, North Africa, and Mexico. G3 (Bethesda) 7:481–495. doi:10.1534/g3.116.036251

Trenberth, K.E., A. Dai, G.S. Van Der, P.D. Jones, J. Barichivich, K.R. Briffa, et al. 2014. Global warming and changes in drought. Nat. Clim. Chang. 4:17–22. doi:10.1038/nclimate2067

van Eeuwijk, F.A., J.B. Denis, and M.S. Kang. 1996. Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables. In: M.S. Kang and H.G. Gauch, editors, Genotype-by-environment interaction. CRC Press, Boca Raton, FL. p. 15–49. doi:10.1201/9781420049374.ch2

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

Wray, N.R., J. Yang, B.J. Hayes, A.L. Price, M.E. Goddard, and P.M. Visscher. 2013. Pitfalls of predicting complex traits from SNPs. Nat. Rev. Genet. 14:507–515. doi:10.1038/nrg3457

Wright, S. 1922. Coefficients of inbreeding and relationship. Am. Nat. 56:330–338. doi:10.1086/279872

Xu, Y. 2016. Envirotyping for deciphering environmental impacts on crop plants. Theor. Appl. Genet. 129:653–673. doi:10.1007/s00122-016-2691-5

Yamada, Y. 1962. Genotype by environment interaction and genetic correlation of the same trait under different environments. Jpn. J. Genet. 37:498–509. doi:10.1266/jjg.37.498

Yang, R.C. 2007. Mixed model analysis of crossover genotype-environment interactions. Crop Sci. 47:1051–1062. doi:10.2135/cropsci2006.09.0611

Zhang, X., P. Pérez-Rodríguez, K. Semagn, Y. Beyene, R. Babu, M.A. López-Cruz, et al. 2014. Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. Heredity 114:291–299. doi:10.1038/hdy.2014.99