

Genomic Prediction of Genetic Values for Resistance to Wheat Rusts

Leonardo Ornella, Sukhwinder-Singh, Paulino Perez, Juan Burgueño, Ravi Singh, Elizabeth Tapia, Sridhar Bhavani, Susanne Dreisigacker, Hans-Joachim Braun, Ky Mathews, and Jose Crossa*

Abstract

Durable resistance to the rust diseases of wheat (*Triticum aestivum* L.) can be achieved by developing lines that have race-nonspecific adult plant resistance conferred by multiple minor slow-rusting genes. Genomic selection (GS) is a promising tool for accumulating favorable alleles of slow-rusting genes. In this study, five CIMMYT wheat populations evaluated for resistance were used to predict resistance to stem rust (*Puccinia graminis*) and yellow rust (*Puccinia striiformis*) using Bayesian least absolute shrinkage and selection operator (LASSO) (BL), ridge regression (RR), and support vector regression with linear or radial basis function kernel models. All parents and populations were genotyped using 1400 Diversity Arrays Technology markers and different prediction problems were assessed. Results show that prediction ability for yellow rust was lower than for stem rust, probably due to differences in the conditions of infection of both diseases. For within population and environment, the correlation between predicted and observed values (Pearson's correlation [ρ]) was greater than 0.50 in 90% of the evaluations whereas for yellow rust, ρ ranged from 0.0637 to 0.6253. The BL and RR models have similar prediction ability, with a slight superiority of the BL confirming reports about the additive nature of rust resistance. When making predictions between environments and/or between populations, including information from another environment or environments or another population or populations improved prediction.

RUST DISEASES are an important cause of wheat production losses worldwide. *Puccinia graminis* (stem rust) and *Puccinia striiformis* (yellow rust) continue to cause major economic losses in various parts of the world and hence receive attention in wheat breeding programs. New races of these fungi have caused yield losses even in areas where the rusts have rarely been detected, and they are more threatening to wheat worldwide than older races (Singh et al., 2011). Regions with vulnerability to yellow rust include, among others, the United States, Asia, and Oceania (Wellings, 2011). Stem rust has also become epidemic in Africa (Singh et al., 2011).

Inheritance of rust resistance in wheat can be either qualitative or quantitative. Quantitative disease resistance is more durable but more difficult to evaluate because it is expressed in mature plants (adult plant resistance) (Rutkoski et al., 2011). Phenotyping adult plant resistance in large populations is expensive and labor intensive. Expertise is needed because expression is affected by, among other factors, the inoculum load and sequential infection (Hickey et al., 2012). Regarding financial costs, in CIMMYT the most basic field assay costs around US\$30 to 40 per genotype (with two replications per location); however, if greenhouse screening is required, the cost increases significantly. In

L. Ornella and E. Tapia, French-Argentine International Center for Information and Systems Sciences (CIFASIS); S. Singh, J. Burgueño, R. Singh, S. Bhavani, S. Dreisigacker, H.-J. Braun, K. Mathews, and J. Crossa, Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, Mexico, D.F., Mexico; P. Perez, Colegio de Postgraduados, Montecillo, Edo. de Mexico, Mexico. Received 5 May 2012. *Corresponding author (j.crossa@cgiar.org).

Abbreviations: ρ , Pearson's correlation; BL, Bayesian LASSO; BLR, Bayesian linear regression; BLUP, best linear unbiased prediction; DArT, Diversity Arrays Technology; GS, genomic selection; H^2 , broad-sense heritability; K-Nyangumi, Kenya Nyangumi; LASSO, least absolute shrinkage and selection operator; RBF, radial basis function; REML, restricted maximum likelihood; RIL, recombinant inbred line; RR, ridge regression; SRM, structural risk minimization; SVR, support vector regression; TRN, training set; TST, testing set.

Published in The Plant Genome 5:136–148.

doi: 10.3835/plantgenome2012.07.0017

© Crop Science Society of America

5585 Guilford Rd., Madison, WI 53711 USA

An open-access publication

All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher.

contrast, according to Lorenz et al. (2012), the current cost of genotyping for genomic selection (GS) is \$20.

The DNA-based strategies such as marker-assisted selection and GS are an alternative for developing high-yielding wheat with high levels of polygenic adult plant resistance (Heffner et al., 2010; Mago et al., 2011). For example, genotyping can be used to select the best candidates before conducting field tests, thus reducing the number of evaluation plots (Heffner et al., 2010). The complex nature of traits, however, makes marker-assisted selection difficult to implement. For example, gene *Sr25* confers a high level of stem rust resistance only when operating in some genetic backgrounds and especially when gene *Sr2* is also present (Singh et al., 2011). Besides, the number of molecular markers linked to resistance genes is insufficient to conduct marker-assisted selection (Singh et al., 2011, 2005). Therefore, GS is a promising alternative for accumulating favorable alleles for rust resistance traits (Rutkoski et al., 2011; Xu et al., 2012).

Genetic values for quantitative traits can be predicted as the sum of all marker effects by regressing phenotypic values on all available markers, as first proposed by Meuwissen et al. (2001). There is a plethora of parametric and nonparametric models for genomic prediction. Ridge regression (RR) and the Bayesian least absolute shrinkage and selection operator (LASSO) (BL) (Park and Casella, 2008) are well-established parametric regression methods for GS. Bayesian LASSO has the advantage of being robust regarding prior distributions of the λ parameter controlling shrinkage of regression coefficients (Crossa et al., 2010; de los Campos et al., 2009; Long et al., 2011b). On the other hand, parametric models may not be flexible enough to handle the complexities of gene-to-gene and gene-to-gene-to-phenotype relations (e.g., epistasis, pleiotropy).

Several nonparametric approaches have been proposed for making phenotypic predictions considering these complexities (de los Campos et al., 2010; Long et al., 2011a; Ober et al., 2011). In such models, no assumption is made regarding the genotype–phenotype relationship. Rather, this relationship is described by a smoothing function driven primarily by the data. Support vector regression (SVR) is considered a state-of-the-art, nonparametric, machine-learning algorithm (Hastie et al., 2011; Smola and Schölkopf, 2004) that is designed to learn from finite training set (TRN) data sets; therefore, it seems to be well suited for GS applications. Furthermore, the incorporation of a nonlinear kernel allows modeling nonlinear relationships between phenotypes and marker genotypes (Long et al., 2011b).

This study aims to evaluate GS for predicting resistance to stem rust and yellow rust in five related wheat populations from CIMMYT's Global Wheat Program. Stem rust was evaluated in two planting seasons, the main and off seasons in Njoro, Kenya, whereas yellow rust was evaluated in Kenya and Mexico (Toluca). Genomic prediction was performed using four different statistical models: BL, RR, and SVR with linear and radial basis function (RBF) kernels.

All five populations were molecularly characterized by a total of 1400 Diversity Arrays Technology (DArT) markers (Triticarte Pty. Ltd., Canberra, Australia). Prediction performance was analyzed using several cross-validation strategies, each assessing four different prediction problems: (i) prediction within populations and environments, (ii) prediction within populations when combining environments, (iii) prediction within populations and between environments, and (iv) prediction between populations.

Materials and Methods

Phenotypic Data

Data sets used in this study come from five wheat populations, PBW343 × Juchi, PBW343 × Pavon76, PBW343 × Muu, PBW343 × Kingbird, and PBW343 × Kenya Nyangumi (K-Nyangumi), evaluated for resistance to both stem and yellow rusts. All populations were derived from crosses between resistant parents (Juchi, Pavon76, Muu, Kingbird, and K-Nyangumi) and PBW343, a moderately susceptible parent. The parents and recombinant inbred lines (RILs) were evaluated for reaction to stem rust at the Kenya Agriculture Research Institute, Njoro, as described by Njau et al. (2010). To analyze resistance to stem race TTKSK (Ug99) using different planting dates (Njau et al., 2010), lines were evaluated in two seasons, the main season (May to October) and the off-season (November to April), in Kenya. The off-season is also known as the wet season whereas the main season is also called the dry season. Percent rust severity for each plot was determined using the modified Cobb Scale (Peterson et al., 1948).

The F_6 generation of RILs was also field tested for yellow rust resistance using artificial epidemics at CIMMYT's research stations in Toluca (Mexico) and natural infection in Njoro (Kenya). The parents and RILs were sown on 75-cm-wide raised beds in paired-row plots, 1 m in length, with 20 cm between rows and a 50-cm pathway. Yellow rust epidemics were initiated about 4 wk after sowing by inoculating spreader rows of susceptible cultivar Morocco planted in hills adjacent to the pathway. To initiate the epidemics, Morocco was sprayed with a suspension of rust urediniospores in Soltrol 170, a lightweight mineral oil (Chevron Phillips Chemical Company, The Woodlands, TX). The yellow rust strains inoculated in Toluca were a pathotype mixture (MEX96.11, MEX07.16, and MEX09.01) whereas in Kenya, yellow rust severity was due to natural infection. The field design for each population was a randomized complete block design with two replicates. The length was of 1 m and the mean of all the plants measured (around 50) were recorded and used for the analyses.

Disease severity of stem rust and yellow rust of the six parents is shown in Table 1. Distribution of the severity for both traits in each population and environment or season is presented in Supplemental Fig. S1; infection levels of the different populations varied in the different environments or seasons.

The number of wheat lines characterized in each population varied depending on the trait and/or the environment where the populations were evaluated (Table 2). Phenotypic data were processed by square root transformation and standardized to mean zero and standard deviation of 1.

Broad-Sense Heritability and Genetic Correlation between Environments

Variance components and broad-sense heritability (H^2) are shown in Table 2. Variance components were estimated for each population separately by restricted maximum likelihood (REML) using Proc Mixed of SAS version 9.0 (SAS Institute, 2004). Data from both trials, that is, Main plus Off Season for stem rust and Njoro plus Toluca for yellow rust, were analyzed using the standard linear mixed model

$$y_{ijk} = \mu + R(E)_{ijk} + E_i + G_j + GE_{ij} + e_{ijk} \quad [1]$$

where μ is the overall mean, $R(E)_{ijk}$ is the effect of the k th replicate of the j th genotype in the i th environment, E_i is the effect of the i th environment, G_j is the effect of the j th genotype, and GE_{ij} is the interaction effect of the j th genotype with the i th environment. All effects were considered random. For both traits the number of replicates per environment was two. Data for yellow rust of population PBW343 \times K-Nyangumi was missing in Njoro.

Broad-sense heritability (repeatability) was estimated as $H^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{ge}^2/e + \sigma_e^2/re)$, where σ_g^2 is the genotypic variance, σ_{ge}^2 is the genotype \times environment interaction, σ_e^2 is the residual variance, e is the number of environments, and r is the number of replicates per environment.

Genetic correlations (ρ_g) between environments and seasons (i th and i' th) were calculated using equations from Cooper et al. (1996): $\rho_g = \sigma_{g(ii')} / \sigma_{g(i)} \sigma_{g(i')}$, where $\sigma_{g(ii')}$ is the arithmetic average of all pairwise genotypic covariances between environments i and i' and $\sigma_{g(i)} \sigma_{g(i')}$ is the arithmetic average of all pairwise geometric means among the genotypic variance components of the environments (Cooper et al., 1996).

Table 1. Range of disease severity (%) on the six parental lines of wheat used in the present study.

Parental line	Stem rust (Kenya) [†]	Yellow rust (Kenya)	Yellow rust (Toluca)
PBW343	60–70	15–20	10–15
Kingbird	0–5	0–5	0–5
Juchi	5–15	20–30	15–20
Muu	5–10	5–10	15–20
Pavon	10–15	10–15	15–20
Kenya Nyangumi	0–5	40–50	0–5

[†]Disease severity in the main and off seasons was not statistically different at the 0.05 level.

Genotypic Data and the Genomic Relationship between Parents and Populations

The six parents and the five wheat populations were molecularly characterized using 1400 DArT markers. Molecular data were scored directly into a spreadsheet as present (1) or absent (0). Before each GS evaluation, noninformative markers (i.e., with an allele frequency of less than 5 or more than 95%) were removed for each individual or combined data sets.

The genetic relationship between lines was examined by means of a heat map of the realized genomic relationship matrix \mathbf{G} . A common notation is $\mathbf{G} =$

$$t^{-1}\mathbf{XX}' \quad (\text{Van Raden, 2008}), \text{ where } t = \sum_{k=1}^m p_k q_k, p_k \text{ is the}$$

frequency of allele A , q_k is the frequency of allele a in the k th marker, \mathbf{X} is the $(n \times k)$ matrix of marker data, and n is the number of individuals. The heat map of the \mathbf{G} matrix of the six parents is depicted in Supplemental Fig. S2a, and the expected genetic relationship between populations using only molecular information of the parents is shown in Supplemental Fig. S2b. Supplemental Fig. S3 depicts the heat map of the \mathbf{G} matrix for lines within a population as well as between populations; four populations are clearly separated in the heat map, but population PBW343 \times Juchi does not seem to have lines that are closely related to one another or to lines from other populations. However, we observed that this relationship is completely consistent with expected genetic relationship between populations using only molecular information of the parents.

Table 2. Numbers of individuals evaluated for rust resistance in a single population \times environment (season) combination and across environments (seasons), variance component estimation, heritability, and genetic correlation between environments.

Population	Stem Rust						Yellow Rust							
	Off season	Main season	σ_g^2	σ_{ge}^2	σ_e^2	H^2	r_g	Njoro (Kenya)	Toluca (Mexico)	σ_g^2	σ_{ge}^2	σ_e^2	H^2	r_g
PBW343 \times Juchi	92	92	96.9	49.7	103.6	0.72	0.47	92	90	100.0	111.1	9.0	0.66	0.52
PBW343 \times Kingbird	90	90	287.3	45.5	88.2	0.90	0.74	90	90	32.1*	74.8	10.0	0.45	0.35
PBW343 \times K-Nyangumi	191	176	309.0	69.4	122.6	0.85	0.66	— [‡]	189	—	—	—	—	—
PBW343 \times Muu	148	148	247.8	31.6	83.9	0.90	0.75	148	148	21.1*	75.5	16.3	0.66	0.21
PBW343 \times Pavon76	180	176	181.23	38.5	95.6	0.83	0.64	147	180	76.8	107.9	18.3	0.56	0.48

^{*} σ_g^2 , genotypic variance; σ_{ge}^2 , genotype \times environment interaction; σ_e^2 , residual variance; H^2 , broad-sense heritability; r_g , genetic correlation between both environments (seasons).

[‡]Yellow rust analysis of population PBW343 \times K-Nyangumi is not presented because there was only one trial.

Statistical Models

Linear Genetic Model

The standard linear genetic model considers that the phenotypic response of the i th individual (y_i) is explained by a factor common to all individuals (μ), a genetic factor specific to that individual (g_i), and the residual comprising all other nongenetic factors (ε_i), for example, environmental effects and effects described by the designs of the experiment. Then, the linear genetic model for n genotypes ($i = 1, \dots, n$) is represented as $y_i = \mu + g_i + \varepsilon_i$. In this standard linear genetic model, the genetic factor g_i can be described by using a summation of molecular marker effects or by using pedigree.

Ridge Regression

Ridge regression, one of the first methods proposed for GS (Meuwissen et al., 2001), is equivalent to best linear unbiased prediction (BLUP) in the context of mixed models (Meuwissen et al., 2001). We used the package rrBLUP (Endelman, 2011), which calculates maximum likelihood (ML or REML) solutions for mixed models of the form

$$\mathbf{y} = \mathbf{X}_F \boldsymbol{\beta}_F + \mathbf{Z} \mathbf{u} + \boldsymbol{\varepsilon}, \quad [2]$$

where \mathbf{X}_F is a full-rank design matrix for the fixed effects $\boldsymbol{\beta}_F$, \mathbf{Z} is the design matrix for the random effects $\mathbf{u} \sim N(0, \mathbf{K}\sigma^2)$, \mathbf{K} is a positive semidefinite matrix, and the residuals are normal with constant variance. The function kinship.BLUP performs genomic prediction based on the kinship between lines; the default method of this function is ridge regression, where $\mathbf{K} = \mathbf{X}\mathbf{X}'$ is the realized genomic relationship matrix computed with the molecular data (Endelman, 2011).

Bayesian Least Absolute Shrinkage and Selection Operator

Bayesian LASSO was first proposed by Park and Casella (2008) and has been successfully applied to GS (e.g., de los Campos et al., 2009).

The model, as implemented in the BLR Package (Pérez et al., 2010), is described as

$$y = \mathbf{1}\mu + \mathbf{X}_F \boldsymbol{\beta}_F + \mathbf{Z} \mathbf{u} + \mathbf{X}_R \boldsymbol{\beta}_R + \mathbf{X}_L \boldsymbol{\beta}_L + \boldsymbol{\varepsilon}, \quad [3]$$

where \mathbf{y} is an $n \times 1$ response vector, μ is an intercept, and \mathbf{X}_F , \mathbf{X}_R , \mathbf{X}_L , and \mathbf{Z} are incidence matrices used to accommodate different types of effects. In our case, \mathbf{X}_F is a matrix of fixed effects indicating populations or environments, and \mathbf{X}_R is a matrix of marker genotypes (coded as 0, 1 for markers). The incidence matrix \mathbf{Z} is used to accommodate random effects. Finally, $\boldsymbol{\varepsilon}$ is a vector of model residuals assumed to be distributed as $N(\mathbf{0}, \text{Diag}(\sigma_\varepsilon^2/\omega_i^2))$, σ_ε^2 is an (unknown) variance parameter, and ω_i^2 are (known) weights that allow for heterogeneous-residual variances. More details on the BLR (Bayesian linear regression) software are given in Pérez et al. (2010).

Support Vector Regression – Linear and Nonlinear Kernels

Conceptually, the support vector machine algorithm is different from BL and RR. It is based on the structural risk minimization (SRM) principle that aims to learn from finite TRN data sets taking into account the complexity of the hypothesis space and the quality of fitting the TRN data (Hastie et al., 2011). We investigated the ε -SVR or “ ε -insensitive” SVR (Smola and Schölkopf, 2004) implemented in Workbench WEKA (Hall et al., 2009).

The general model can be described as $y = \beta_0 + h(\mathbf{x})'\boldsymbol{\beta}$, where y is a real-value response variable, for example, rust resistance, and $h(\mathbf{x})^T$ is a linear (or nonlinear) transformation of one (or more) predictors (x) additively combined with the vector of weights ($\boldsymbol{\beta}$).

For the sake of clarity, we first describe the most basic linear function, which can take the form

$$f(\mathbf{x}) = \langle \boldsymbol{\omega}, \mathbf{x} \rangle + b, \quad \boldsymbol{\omega} \in \mathbb{R}^D \text{ and } b \in \mathbb{R}, \quad [4]$$

where $\langle \boldsymbol{\omega}, \mathbf{x} \rangle$ denotes the dot product in χ , the space of the input patterns.

By the SRM principle, generalization is optimized by the flatness of the regression function; this is done by minimizing

$$1/\|\boldsymbol{\omega}\|^2 \quad [5]$$

subject to

$$\begin{cases} y_i - \langle \boldsymbol{\omega}, \mathbf{x}_i \rangle - b \leq \varepsilon \\ \langle \boldsymbol{\omega}, \mathbf{x}_i \rangle + b - y_i \leq \varepsilon \end{cases} \quad [6]$$

One attractive feature of ε -SVR is that it assigns zero value to residuals smaller in absolute value than the constant (ε) and a linear loss function for larger residuals (ε -tube):

$$L_\varepsilon(y, f(\mathbf{x}, \boldsymbol{\omega})) = \begin{cases} 0 & \text{if } |y - f(\mathbf{x}, \boldsymbol{\omega})| \leq \varepsilon \\ |y - f(\mathbf{x}, \boldsymbol{\omega})| - \varepsilon & \text{otherwise} \end{cases} \quad [7]$$

Usually nonlinear models are required to adequately represent and model the data. The simplest solution is to apply a nonlinear transformation to the data in a high dimensional feature space where linear regression is performed. A computationally efficient way to do this is to perform an implicit mapping using kernel functions (Hastie et al., 2011; Smola and Schölkopf, 2004).

We evaluated the performance of the linear kernel, which can be homogeneous, that is, $k(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$, or nonhomogeneous, $k(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle + 1$, and the RBF kernel, which can be expressed as $k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$. Optimization of parameters was performed by a logarithmic grid search of base 2 over an extensive range of values. Each point on the grid was evaluated by internal fivefold cross-validation on the TRN using Pearson's correlation as a criterion for success. In our preliminary research, we found no significant impact of ε tuning on the success of regression; therefore, given the computational cost involved in a grid search, optimization focused only on the C parameter for the linear kernel and C, γ for the RBF kernel.

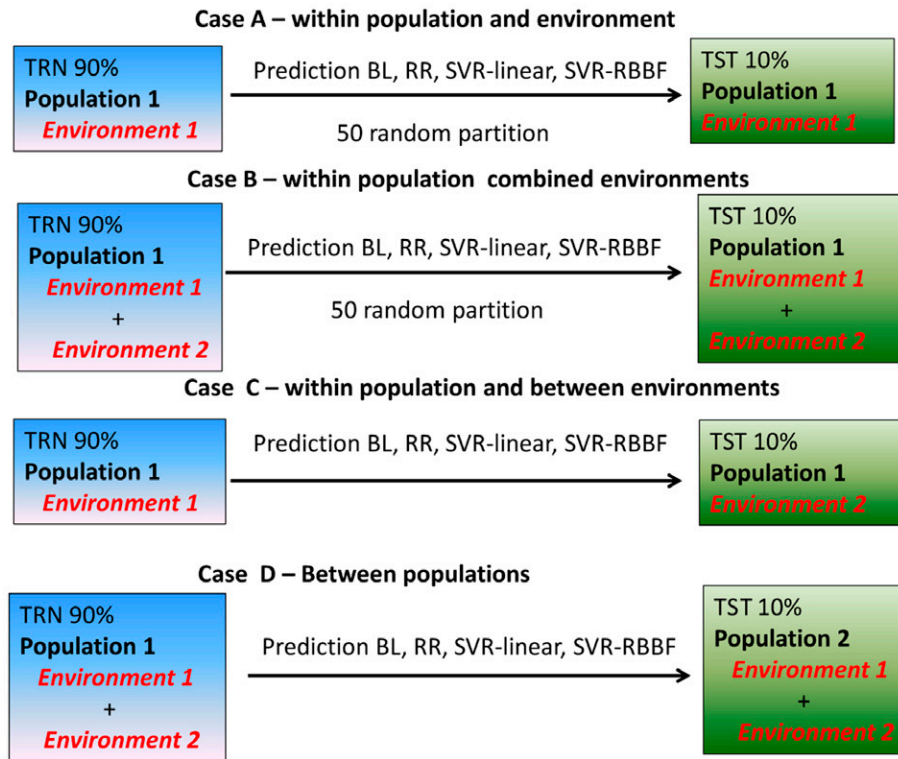


Figure 1. Description of four different prediction designs (Case A, Case B, Case C, and Case D) that split the data into training set (TRN) and testing set (TST). For the four cases, the TRN had 90% of the total number of lines whereas the TST had 10%. BL, Bayesian least absolute shrinkage and selection operator (LASSO); RR, ridge regression; SVR-linear, support vector regression with linear kernel; SVR-RBBF, support vector regression with radial basis function kernel.

Designs for Assessing Different Prediction Problems

Fitting Models Using the Full Data

In this study, we designed four different layouts of the TRN and the testing set (TST), each of them assessing four different prediction cases (Fig. 1). In the first prediction problem (Case A), accuracy is assessed within populations in each environment (season). Data were randomly partitioned 50 times using a TRN containing 90% of the lines and a TST having 10% of the lines. The second prediction problem (Case B) consisted of combining in the TRN the performance of one population evaluated in two environments (seasons or sites) and predicting only one environment (one season or site). For example, the matrix used for the TRN for yellow rust of population PBW343 × Pavon76 in Njoro plus Toluca was generated by combining the matrix used for the yellow rust data of the population evaluated in Njoro and the matrix used on the yellow rust data of the population evaluated in Toluca. This allows assessing the prediction ability of the models using the same TST under different TRN situations (Fig. 1).

The third prediction layout (Case C) involved training the algorithms with the complete information from one environment of each population and evaluating their performance using data from the complementary environment (i.e., the TRN for yellow rust of PBW343 × Pavon76 in Njoro and the TST for yellow rust of PBW343 × Pavon76 in Toluca). Finally, the fourth prediction design (Case D) involved training the algorithms with each of the

populations using information from both environments, for example, Main plus Off Season or Njoro plus Toluca, and evaluating their performance separately in each of the remaining populations. It should be pointed out that all possible pairwise comparisons (direct and reciprocal) were performed; however, for simplicity, only direct predictions are reported.

Fitting Models Using Only Ninety Individuals

To examine the potential effect of different sample sizes in the TRN on the prediction of the TST, we examined three of the above-mentioned prediction problems (Cases A, B, and C) while fitting only 90 randomly selected individuals from the three populations with the largest sample size, that is, PBW343 × K-Nyangumi, PBW343 × Muu, and PBW343 × Pavon76 with 176, 148, and 180 lines, respectively. The procedure consisted of a simple random sampling scheme in which 90 lines were randomly selected 20 times from each of the three populations. Then, for each of the 20 simple random samples of 90 lines each, 50 random partitions were performed to split the data into a TRN (90% of the lines) and a TST (10% of the lines).

Software and Statistical Analyses

Scripts for SVR evaluation were implemented in Java code using Eclipse platform version 3.7 (Dexter, 2007). Bayesian LASSO predictions were made with the BLR package for R, version 1.2 (R Development Core Team, 2008; Pérez et al., 2010), and hyperparameters were

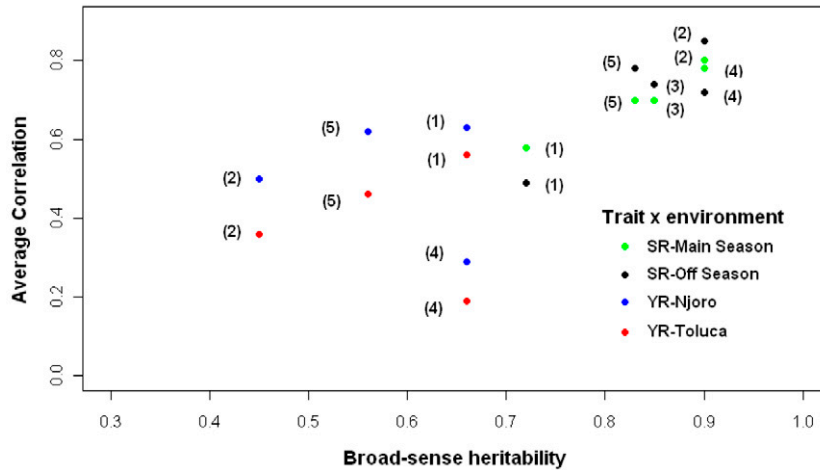


Figure 2. Scatterplot of the average correlations for each population \times environment combination obtained with the Bayesian least absolute shrinkage and selection operator (LASSO) model versus broad-sense heritability when the algorithm was trained with information from both main season plus off-season for stem rust (SR) and from both Njoro plus Toluca for yellow rust (YR). Populations are (1) PBW343 \times Juchi, (2) PBW343 \times Kingbird, (3) PBW343 \times K-Nyangumi, (4) PBW343 \times Muu, and (5) PBW343 \times Pavon76.

chosen based on the guidelines of the author. Ridge regression was performed with the package rrBLUP, version 3.8, also for R (Endelman, 2011).

The Shapiro-Wilk test (Shapiro and Wilk, 1965) for normality of the correlations from the 50 random partitions was rejected for Cases A and B; the Wilcoxon test for paired samples was then performed. Statistical differences on the correlations for the cases presented in Supplemental Tables S1 and S2 were performed using the Mann-Whitney U test (Mann and Whitney, 1947). R project (R Development Core Team, 2008) and R Commander (Fox, 2005) were used to perform statistical or routine analyses.

Algorithms were evaluated in a Sun Fire X2250 Server, and the nodes were Quad Core Intel Xeon Processors, Model L5420 (2.50 GHz, 1333 MHz, and 50 W), all with 8 GB of RAM memory. Routine tasks and preliminary analyses were performed using a standard desktop computer, with an Intel Core Duo 2.80 GHz processor and 4 GB of RAM memory.

Results

The total number of lines varied per population; PBW343 \times Juchi and PBW343 \times Kingbird had only half of the sample size as the other three populations (Table 2). The distribution of stem rust and yellow rust resistance of the five populations in environments varied (Supplemental Fig. S1). Incidence of stem rust ranged from 0 to 85% in the main season and from 0 to 90% in the off-season. The yellow rust incidence ranged from 1 to 75% in Toluca and from 0 to 70% in Njoro.

Variance components and H^2 for both traits in each population are given in Table 2. Broad-sense heritability was higher for stem rust (ranging from 0.72 to 0.90) than for yellow rust (ranging from 0.45 to 0.88). Genetic correlations between the off and main seasons for stem rust were relatively high, varying across populations from 0.47 to 0.75, whereas genetic correlations between Njoro and Toluca for yellow rust ranged from 0.21 to 0.52.

In the next two sections, prediction ability results for the four different prediction problems are presented for stem rust and yellow rust.

Stem Rust

The positive association between the heritability of the trait and prediction ability (correlation between predicted and observed values) of the BL model for the off and main seasons is depicted in Fig. 2 for the five populations. This result indicates the clear relationship between heritability and predictability of stem rust.

Predicting Resistance within Populations and within and between Environments – Fitting Models Using Full Data (Cases A, B, and C)

The ability of BL, RR, and SVR with linear or RBF kernels to predict stem rust incidence in five populations of wheat when evaluated in two seasons of in Kenya (Main plus Off Season) is given in Table 3. The performance of the algorithms was evaluated under three different prediction cases: Case A, Case B, and Case C.

Concerning Case A, for resistance measured in the off-season, the Pearson's correlation (ρ) ranged from 0.26 (PBW343 \times Juchi trained with SVR with linear kernel) to 0.75 (PBW343 \times Kingbird trained with BL). For the main season, correlations ranged from $\rho = 0.32$ (PBW343 \times K-Nyangumi and SVR with RBF kernel) to $\rho = 0.69$ (PBW343 \times Kingbird with BL). In most populations, in both the main and off seasons, best predicted values were obtained with BL. This superiority was statistically significant at 0.05 or 0.01 (Wilcoxon test for paired samples) when compared with the other models.

For Case B, the lowest correlation when predicting the off-season was given by SVR with linear kernel for population PBW343 \times Juchi ($\rho = 0.39$), and the highest value was obtained with BL when predicting the resistance of population PBW343 \times Kingbird ($\rho = 0.85$). When

Table 3. Pearson's correlation (ρ) between observed and predicted stem rust values using four genomic selection models trained under three different conditions in five wheat populations evaluated in two seasons.

Population	Model [†]							
	BL	SVR-linear	SVR-RBF	RR	BL	SVR-linear	SVR-RBF	RR
Case A [‡]	TRN = TST = Off				TRN = TST = Main			
PBW343 × Juchi	0.38	0.26	0.41	0.41	0.55	0.55	0.52	0.56
PBW343 × Kingbird	0.75**	0.66	0.64	0.68	0.69**	0.61	0.59	0.60
PBW343 × K-Nyangumi	0.59**	0.52	0.55	0.56	0.39	0.39	0.32	0.37
PBW343 × Muu	0.57*	0.50	0.53	0.56	0.62*	0.53	0.57	0.61
PBW343 × Pavon76	0.68**	0.55	0.58	0.60	0.60**	0.46	0.50	0.52
Case B [§]	TRN = Main + Off TST = Off				TRN = Main + Off TST = Main			
PBW343 × Juchi	0.49 (28.9)	0.39 (50.0)	0.39 (−0.01)	0.53** (29.3)	0.58** (5.5)	0.45 (−18.2)	0.45 (−13.5)	0.47 (−16.1)
PBW343 × Kingbird	0.85** (13.3)	0.78 (18.2)	0.69 (7.8)	0.82 (20.6)	0.80 (15.9)	0.76 (24.6)	0.59 (0.0)	0.81* (35.0)
PBW343 × K-Nyangumi	0.74 (25.4)	0.66 (26.9)	0.78** (41.8)	0.65 (16.1)	0.70** (79.5)	0.67 (71.8)	0.62 (93.8)	0.62 (67.6)
PBW343 × Muu	0.72 (26.3)	0.72 (44.0)	0.55 (3.8)	0.75** (33.9)	0.78** (25.8)	0.75 (41.5)	0.60 (5.3)	0.72 (18.0)
PBW343 × Pavon76	0.78** (8.8)	0.66 (20.0)	0.71 (22.4)	0.72 (20.0)	0.70 (16.7)	0.67 (45.7)	0.64 (28.0)	0.70 (34.6)
Case C [¶]	TRN = Main TST = Off				TRN = Off TST = Main			
PBW343 × Juchi	0.47	0.47	0.47	0.5	0.5	0.47	0.48	0.47
PBW343 × Kingbird	0.81	0.77	0.77	0.79	0.80	0.77	0.77	0.80
PBW343 × K-Nyangumi	0.61	0.64	0.62	0.62	0.62	0.67	0.67	0.61
PBW343 × Muu	0.77	0.77	0.77	0.80	0.80	0.77	0.80	0.77
PBW343 × Pavon76	0.70	0.68	0.68	0.68	0.70	0.69	0.63	0.68

*Best result for each environment × population combination is statistically significant at the 0.05 probability levels.

**Best result for each environment × population combination is statistically significant at the 0.01 probability levels.

[†]BL, Bayesian least absolute shrinkage and selection operator (LASSO); SVR-linear, support vector regression with linear kernel; SVR-RBF, support vector regression with radial basis function kernel; RR, ridge regression; TRN training set; TST, testing set; Off, off season; Main, main season.

[‡]The first two upper sections show the mean Pearson's correlation (ρ) between observed and predicted stem rust values of 50 random partitions of the data for four models in five wheat populations evaluated in two seasons, off and main seasons. Models were trained and tested using three different combinations of training set (TRN) and testing set (TST). Case A: training and testing in the same environment. Case B: training in main plus off season and testing only in off season or only in main season. Case C: training in one season and predicting in the other.

[§]Numbers in parenthesis denote the percentage increase (or decrease) in the models' prediction ability for Case B compared to their respective models in Case A.

[¶]Nonsignificant differences are not indicated.

predicting the resistance of the individuals in the main season, the worst results were obtained with SVR with linear and RBF kernels for population PBW343 × Juchi ($\rho = 0.45$) whereas the best result was obtained with RR in population PBW343 × Kingbird ($\rho = 0.81$). Results indicated an important improvement in prediction accuracy in most of the populations for Case B over Case A (Table 3), except for the population derived from PBW343 × Juchi evaluated in the main season. Prediction ability of the different models differed when trained with data from both seasons as compared with training in one season. For example, BL increased its performance just 5.5% when predicting individuals derived from PBW343 × Juchi in the main season but showed 79.5% improvement when predicting individuals of PBW343 × K-Nyangumi, also in the main season. Support vector regression with RBF showed 93.8% improvement when predicting the same population in the same season (PBW343 × K-Nyangumi and the main season) but its performance decreased 13.5% when predicting lines derived from Juchi in the main season.

Finally, when algorithms were trained exclusively with data from one season and evaluated with data from the other season (Case C), success of prediction on main season (i.e., trained with off-season data) ranged from 0.43 (RR and PBW343 × Pavon76) to 0.80 (BL and RR for PBW343

× Kingbird and SVR with RBF kernel for PBW343 × Muu). When the off-season was predicted, correlations ranged from 0.47 (BL, SVR with linear kernel, and SVR with RBF kernel with PBW343 × Juchi) to 0.81 (BL and PBW343 × Kingbird). In general, training in one season and predicting in the other showed very good prediction ability. Overall, a strong consistency among the three training cases was observed. Population PBW343 × Kingbird always had, on average, the best predictions in both seasons and for the three prediction cases whereas population PBW343 × Juchi gave the lowest correlations in all prediction cases, except for Case A in the main season (the worst was PBW343 × Pavon76). These results can be explained by the higher genomic relationships among the lines of population PBW343 × Kingbird as compared with those comprising population PBW343 × Juchi (Supplemental Fig. S3). Regarding the prediction ability of the models, BL presented the best (mean) correlation in five of eight situations, but RR showed also superior prediction ability in several cases.

Predicting Resistance between Populations (Case D)

The four prediction problems consisted of training the models within each of the five populations using full data from both seasons that had been tested in each of the other populations. Results for the between-populations prediction

Table 4. Pairwise Pearson's correlation (without reciprocals) between observed and predicted stem rust values of four different models. Algorithms were trained in one population (across both environments) and evaluated on the other population.

		Testing or training [†]					
		PBW343 × Juchi	PBW343 × Kingbird	PBW343 × K-Nyangumi	PBW343 × Muu	PBW343 × Pavon76	
Training or testing [‡]	PBW343 × Juchi	–	0.48	0.14	0.28	0.31	Bayes LASSO
	PBW343 × Kingbird	0.53	–	0.29	0.25	0.54	
	PBW343 × K-Nyangumi	0.14	0.30	–	0.28	0.28	
	PBW343 × Muu	0.18	0.30	0.33	–	0.29	
	PBW343 × Pavon76	0.37	0.51	0.22	0.33	–	
		Ridge regression					
		Testing or training [§]					
		PBW343 × Juchi	PBW343 × Kingbird	PBW343 × K-Nyangumi	PBW343 × Muu	PBW343 × Pavon76	
Training or testing	PBW343 × Juchi	–	0.28	0.32	0.48	0.32	SVR-linear
	PBW343 × Kingbird	0.28	–	0.45	0.53	0.59	
	PBW343 × K-Nyangumi	0.35	0.31	–	0.25	0.39	
	PBW343 × Muu	0.14	0.58	0.26	–	0.55	
	PBW343 × Pavon76	0.35	0.64	0.41	0.67	–	
		SVR-RBF					

[†]The upper-right triangle of the first section of the table has the results of Bayes least absolute shrinkage and selection operator (LASSO), with the rows indicating the training population and the columns the testing population.

[‡]The lower-left triangle gives the results of ridge regression, with the columns indicating the training population and the rows the test population.

[§]The second section of the table shows similar results in the upper-right and lower-left triangles but considers support vector regression with linear kernel (SVR-linear) and support vector regression with radial basis function kernel (SVR-RBF).

problem are presented in Table 4 (for simplicity we present only one-way results because reciprocal prediction was, in all cases, very similar to the direct prediction presented here). Over all populations, the SVR models gave the best results, with $\rho = 0.42$ for the linear and $\rho = 0.40$ for the RBF kernel. Correlation for BL was 0.31 and for RR was 0.32. It seems that nonparametric models with linear and nonlinear kernels can better capture the relationship between some populations, while parametric models (RR and BL) can capture the relationship in others (Table 4). Interestingly, for stem rust a clear trend was observed between the relatedness of populations and the average correlation (Fig. 3) with the four models.

Predicting Resistance within Populations and within and between Environments – Fitting Models Using Ninety Randomly Selected Lines

To determine whether, besides its heritability, the size of the TRN had any effect on the predictive ability of the algorithms, we performed a second round of analyses similar to those presented in Table 3 but with the number of individuals of each population fitted to 90 lines. Results are presented in Supplemental Table S1, where the numbers in parentheses and italics represent the decrease in correlation as compared to the case where the full data were used (presented in Table 3). Populations PBW343 × Juchi and PBW343 × Kingbird were not included because they had only 92 and 90 individuals, respectively.

Clearly there is a decreasing effect on the prediction ability of the models for populations PBW343 × K-Nyangumi, PBW343 × Muu, and PBW343 × Pavon76 for the three prediction cases and for most of the models. For Case A, prediction ability decreases 0 to 42%, for

Case B, correlations decrease 0 to 38%, and for Case C, the decrease ranged from 0 to 40%.

Yellow Rust

The association between heritability and prediction ability of the BL model for Njoro and Toluca is depicted in Fig. 2 for four populations (PBW343 × K-Nyangumi was not included). Note that the association is not as clear as in the case of stem rust.

Predicting Resistance within Populations and within and between Environments – Fitting Models Using Full Data

The average accuracy of BL, RR and SVR with linear and RBF kernels for predicting yellow rust severity in five populations of wheat in two environments, Njoro and Toluca, is given in Table 5. In general, prediction ability for yellow rust is lower than that for stem rust, reflecting the generally lower estimated heritability of yellow rust resistance compared to the estimated heritability of stem rust resistance (Table 2). For Case A, ρ values ranged from 0.12 (PBW343 × Muu and RR) to 0.42 (PBW343 × Juchi and SVR-linear) in Njoro whereas in Toluca, ρ ranged from 0.15 (PBW343 × K-Nyangumi and RR) to 0.63 (PBW343 × Pavon76 and BL). In Toluca, BL had the best correlations in two populations, PBW343 × K-Nyangumi and PBW343 × Pavon76, and the same prediction as RR in the other three populations. In Njoro, for PBW343 × Kingbird and PBW343 × Muu, BL gave the best prediction whereas for the other populations both RR and BL gave similar predictions.

Concerning Case B in Toluca, the lowest correlation was obtained with SVR with RBF kernel PBW343 × Muu population ($\rho = 0.06$) and the highest value was obtained

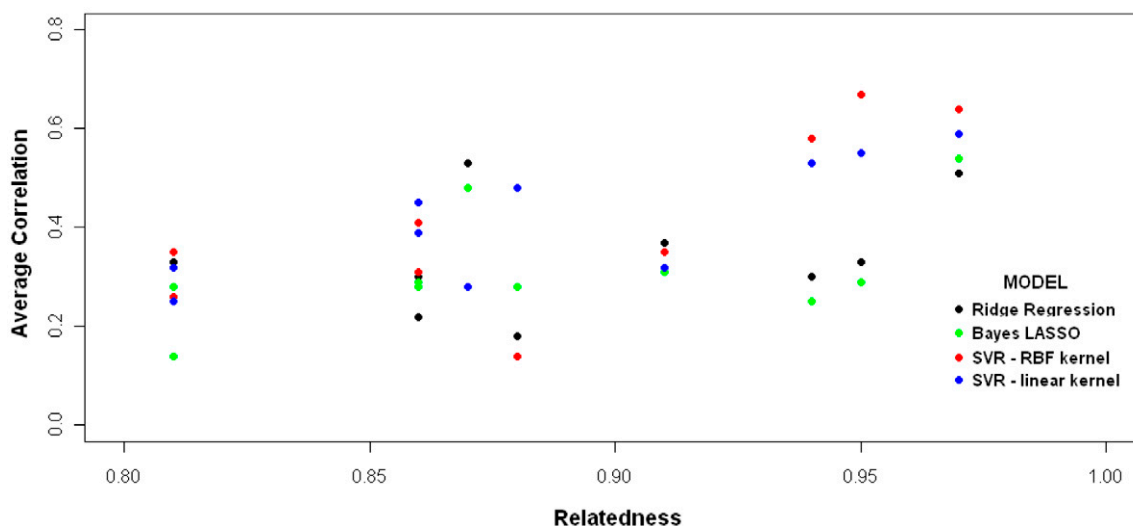


Figure 3. Average correlation of stem rust for four genomic selection models (ridge regression, Bayesian least absolute shrinkage and selection operator (LASSO), support vector regression (SVR) with radial basis function (RBF) and SVR with linear kernel vs. the relatedness between populations as estimated from the **G** matrix (see Supplemental Fig. S3).

Table 5. Pearson's correlation (ρ) between observed and predicted yellow rust values for four models trained under three different conditions in five wheat populations evaluated in two sites Njoro (Kenya) and Toluca (Mexico).

Population	Model [†]							
	BL	SVR-linear	SVR-RBF	RR	BL	SVR-linear	SVR-RBF	RR
Case A ^{‡¶}	TRN = TST = Njoro				TRN = TST = Toluca			
PBW343 × Juchi	0.33	0.42*	0.41	0.33	0.37	0.37	0.34	0.37
PBW343 × Kingbird	0.23**	0.16	0.14	0.20	0.49	0.42	0.48	0.49
PBW343 × K Nyangumi	—	—	—	—	0.30**	0.24	0.23	0.15
PBW343 × Muu	0.17	0.19	0.20*	0.12	0.49	0.45	0.46	0.49
PBW343 × Pavon76	0.26	0.20	0.33**	0.25	0.63**	0.54	0.58	0.59
Case B ^{§¶}	TRN = Njoro + Toluca TST = Njoro				TRN = Njoro + Toluca TST = Toluca			
PBW343 × Juchi	0.56 (69.7)	0.53 (26.2)	0.52 (26.9)	0.56 (69.7)	0.63 (70.3)	0.60 (62.2)	0.60 (44.1)	0.52 (40.5)
PBW343 × Kingbird	0.36** (60.9)	0.30 (100)	0.33 (135.7)	0.32 (60.0)	0.50* (2.0)	0.31 (−26.1)	0.47 (−14.6)	0.45 (−8.2)
PBW343 × Muu	0.19 (11.8)	0.15 (−21.1)	0.20 (0.0)	0.14 (16.7)	0.29 (−40.8)	0.34 (−24.0)	0.06 (−87.0)	0.26 (−47.0)
PBW343 × Pavon76	0.46 (76.9)	0.46 (130.0)	0.46 (39.4)	0.45 (80.0)	0.62** (−1.6)	0.47 (−13.0)	0.56 (−3.45)	0.60 (1.69)
Case C [¶]	TRN = Toluca TST = Njoro				TRN = Njoro TST = Toluca			
PBW343 × Juchi	0.52	0.52	0.52	0.54	0.52	0.52	0.52	0.56
PBW343 × Kingbird	−0.09	0.30	0.31	−0.10	−0.16	0.34	0.42	0.10
PBW343 × Muu	0.18	0.13	0.13	0.18	0.17	0.15	0.15	0.17
PBW343 × Pavon76	0.48	0.47	0.47	0.48	0.47	0.44	0.41	0.45

*Best result for each environment × population combination is statistically significant at the 0.05 probability levels.

**Best result for each environment × population combination is statistically significant at the 0.01 probability levels.

[†]BL, Bayesian least absolute shrinkage and selection operator (LASSO); SVR-linear, support vector regression with linear kernel; SVR-RBF, support vector regression with radial basis function kernel; RR, ridge regression; TRN, training set; TST, testing set.

[‡]The two upper sections show the mean Pearson's correlation (ρ) between observed and predicted yellow rust values of 50 random partitions of the data for four models in five wheat populations evaluated in two sites Njoro (Kenya) and Toluca (Mexico). Models were trained and tested using three different combinations of training set (TRN) and testing set (TST). Case A: training and testing in the same environment. Case B: training in Njoro plus Toluca and testing only in Njoro or only in Toluca. Case C: training in one environment and predicting in the other.

[§]Numbers in parentheses denote the percentage increase (or decrease) in the models' predictive ability for Case B compared to their respective models in Case A.

[¶]Nonsignificant differences are not indicated.

with BL when predicting the resistance of population PBW343 × Juchi ($\rho = 0.63$). When predicting the resistance of the individuals in Njoro, the worst results were also obtained with the PBW343 × Muu-derived population but with RR (0.14) followed by the two SVR models (0.15) whereas the best result was obtained with BL and RR and population PBW343 × Juchi ($\rho = 0.56$).

Prediction results when using both sites, Njoro plus Toluca, to predict one of them indicate a substantial increase with respect to Case A, with the algorithms showing a differential response in prediction accuracy between most populations (Table 5). For example, SVR with linear kernel improved 100 (PBW343 × Kingbird in Njoro) and 130% (PBW343 × Pavon76 in Njoro) and

Table 6. Pairwise Pearson's correlation (without reciprocals) between observed and predicted yellow rust values for four different models. The algorithms were trained in one population (across both environments) and evaluated in the other population.

		Testing or training				
		PBW343 × Juchi	PBW343 × Kingbird	PBW343 × Muu	PBW343 × Pavon76	
Training or testing	PBW343 × Juchi	–	0.12	0.06	0.09	Bayesian LASSO†
	PBW343 × Kingbird	0.06	–	0.16	0.14	
	PBW343 × Muu	0.05	0.16	–	0.07	
	PBW343 × Pavon76	0.10	0.14	0.09	–	
Ridge regression‡						
		Testing or training				
		PBW343 × Juchi	PBW343 × Kingbird	PBW343 × Muu	PBW343 × Pavon76	
Training or testing	PBW343 × Juchi	–	0.09	0.05	0.07	SVR-linear§
	PBW343 × Kingbird	0.09	–	0.04	0.21	
	PBW343 × Muu	0.09	0.11	–	0.09	
	PBW343 × Pavon76	0.12	0.18	0.12	–	
SVR-RBF						

†The upper-left triangle in the first section of the table shows the results of Bayesian least absolute shrinkage and selection operator (LASSO), with the rows indicating the training population and the columns the testing population.

‡The lower-left triangle shows the results of ridge regression, with the columns indicating the training population and the rows the test population.

§The second section of the table shows similar results in the upper and lower triangles but considering support vector regression with linear kernel (SVR-linear) and support vector regression with radial basis function (SVR-RBF) kernel.

decreased 57.8% when predicting lines of PBW343 × Muu in Toluca. On most of the population, this response depended on the algorithms, that is, the performance of some algorithms increased while that of others decreased. However, population PBW343 × Muu showed a decrease in the performance of all the algorithms when predicting lines in Toluca using information from both trials.

For Case C, when algorithms were trained exclusively with information from Toluca to predict Njoro, correlations ranged from -0.09 (or -0.10) (RR or BL and population PBW343 × Kingbird) to 0.54 (RR and PBW343 × Juchi) whereas in the reciprocal, correlations ranged from -0.16 (BL and population PBW343 × Kingbird) to 0.56 (RR and PBW343 × Juchi). In general, training in one environment and predicting in the other gave good prediction accuracies for populations PBW343 × Juchi and PBW343 × Pavon76 but low correlations for PBW343 × Muu. Correlations for population PBW343 × Kingbird were negligible.

Predicting Resistance between Populations

The other prediction problem is concerned with predicting yellow rust resistance in one population with the aim of predicting the resistance of another. Correlations shown in Table 6 were lower than those presented in Table 3 for stem rust. This was expected because stem rust was generated with only one pathotype. Results show that BL and RR gave relatively higher predictions than the SVR models.

Predicting Resistance within Populations and within and between Environments – Fitting Models Using Ninety Randomly Selected Lines

Results are presented in Supplemental Table S2. Regarding the evaluation of individual data sets (Case A), prediction ability using adjusted data sets decreased between 5 and 61%. On average, algorithms decreased their performance

by 28%. In the combined data sets (Case B), performance decreased 0 to 39%, averaging $-21%$ in performance variation. Finally, using only information from one environment to predict the other one the prediction ability of the models decreased, on average, 27% on data sets fitted to 90 individuals compared to the case where the full data were used.

Discussion

Comparing Model Prediction and Different Prediction Problems

We compared the performance of RR, BL and SVR, with linear or RBF kernel, for predicting stem rust (race TTKSK) and yellow rust resistance of five wheat populations. For stem rust resistance, we analyzed the performance of the algorithms on 10 individual data sets (representing 10 population × environment combinations), each including from 90 to 180 individuals. Performance was evaluated under different training–testing scenarios.

The first scenario, consisting of the random partitioning of the populations in each environment, was proposed considering the problems associated with rust phenotyping—for example, if a program has to evaluate 200 lines in one environment but only has enough resources to phenotype 100 individuals, the genetic value of the remaining candidates can be inferred with relatively high prediction ability as long as the lines in the environment where the model is trained are related to the unobserved lines to be predicted. The second scenario was proposed considering the multienvironments trials commonly used at CIMMYT for evaluating grain yield and diseases (Singh et al., 2012). As shown in Burgueño et al. (2012), combining information from correlated environments can improve prediction accuracy.

The third case allows testing prediction accuracy when the TRN and TST data sets are independent. The algorithm is trained with data from one environment and evaluated using information from the other environment, that is, TRN (Population 1 in Environment A) and TST (Population 1 in Environment B). This situation gives an idea of how to perform GS in other environments (site-year combinations). Finally, in the fourth prediction problem, the TRN and TST data sets are also independent, but now allele effects are obtained by training in one population and used to predict the allele effects in another population, that is, TRN (Population 1 in Environment A) and TST (Population 2 in Environment A).

Predictions of stem rust in all prediction cases, populations, and models were higher than those for yellow rust. This may be because these populations were originally designed for stem rust quantitative trait loci mapping (Bhavani et al., 2011) and also because only stem rust race TTKSK was used to inoculate and cause stem rust infection in both seasons whereas yellow rust infection was achieved using natural populations in Kenya or a mixture of pathotypes in Mexico. These and other unknown genotype \times environment effects resulted in a much lower genetic variance for yellow rust.

The scenarios presented here attempted to reproduce real situations faced by plant breeders. As reported by Heffner et al. (2010), if net merit (i.e., overall performance) exceeds 0.50, GS could greatly outperform conventional marker-assisted selection in terms of gain per unit time and cost. The results of the three prediction problems (Case A, Case B, and Case C) examined in Table 3 for stem rust may be due to the high heritability of the trait, the high genetic correlation between the off and main seasons (inoculation was performed using only race TTKSK), and the high genomic relationship among the lines comprising the populations, except PBW343 \times Juchi. Regarding this population, the low genomic relationship observed among the lines can be attributed to the distance between the parents, that is, PBW343 and Juchi. This observation was confirmed by calculating the expected matrix **G** based on parent data (Supplemental Fig. S2) and also the **G** matrix of the five populations (Supplemental Fig. S3).

Additive effects due to multiple minor slow-rusting genes reported in the literature (Singh et al., 2008) may also explain the good results obtained with BL and RR, which in most data sets outperformed the two SVR models. This is not the case for more complex traits; for example, Long et al. (2011b) found that SVR with RBF kernel compared favorably to BL when evaluating a data set for wheat grain yield.

Clearly, when environments were correlated (as the off and main seasons), this favored the prediction of one environment while training in the other (Case B; Tables 3 and 5). In relation to this, differential prediction ability of the models is observed when trained with data from both seasons as compared with training in one season. For example, BL increased its performance just 5.5% when predicting individuals derived from PBW343 \times Juchi in the main season (Table 3) but showed 79.5% improvement when

predicting individuals of PBW343 \times K-Nyangumi, also in the main season. These differences can be attributed to the size of the populations, that is, Juchi has 92 individuals and K-Nyangumi, 186. However, when we fitted the individuals of K-Nyangumi with 90 individuals, a 27.8% improvement in prediction ability was obtained (Supplemental Table S1). Support vector regression with RBF showed a 93.8% improvement when predicting the same population in the main season, but SVR with linear kernel decreased 13.5% when predicting lines derived from Juchi in the main season.

Finally, the same algorithm (SVR with RBF kernel) showed a 20.3% increase in its prediction ability when trained with population PBW343 \times K-Nyangumi population but fitted to 90 individuals (Supplemental Table S1). As previously pointed out by Jannink et al. (2010), different models capture different aspects of the relationship between marker genotype and resistance, and thus they could complement each other.

Reducing the number of individuals to 90 lines significantly influenced the success of the prediction, indicating the importance of having large TRN, especially for traits with low heritability.

Prediction between Populations, the Size of the Training Population, and the Two Traits

A difficult prediction problem arises when attempting to predict one population using another. However, as shown in Table 4, some populations are excellent predictors of others when the SVR model is used. For example, data from PBW343 \times Kingbird were used to predict PBW343 \times Muu (0.73) and also PBW343 \times Pavon76 (0.70). Figure 2 shows that the cross products of these pairs of populations in the **G** matrix had average values of 0.97, 0.95, and 0.84, respectively. Figure 3 shows a strong association between the relatedness of the populations and average accuracy.

Results show that the decrease in size of the TRN population produced a decrease in the models' prediction ability (Supplemental Table S1) compared with prediction using full data (Table 2), indicating the importance of having large TRN. We should point out that the size of all the populations was only adjusted to 90 to 92 individuals to compare the different populations without the effect of TRN size.

We evaluated nine individual data sets for yellow rust resistance, and predictions were, in general, not as high as those for stem rust resistance. This lower prediction ability may be due to the lower heritability of yellow rust resistance as compared to stem rust resistance. Natural infection of yellow rust in Njoro could contribute to the lower heritability and thus a decrease in GS accuracies. With natural infections there is less uniformity of disease pressure throughout the field. Lorenz et al. (2012) obtained similar results ($\rho = 0.72$) when predicting Fusarium head blight of barley (*Hordeum vulgare* L.) using GS. In Rutkoski et al. (2012), ρ reached almost 0.55 when predicting incidence, severity, and kernel quality index of Fusarium head blight in wheat using a similar number of individuals (170).

Although the yellow rust races operating in the two international environments (Njoro and Toluca) were different and heritability was low, the results for yellow rust improved when the algorithms were trained using simultaneous information from Njoro plus Toluca. Correlations of 0.63 for BL were reached when predicting population PBW343 × Juchi in Toluca (Table 5). Also for PBW343 × Juchi and PBW343 × Pavon76, prediction in Njoro while TRN in Toluca or vice versa produced good prediction ability for most models, indicating a good genetic correlation. These results, especially those obtained with stem rust, indicate that it is possible to use GS to predict the performance of rust resistance in future environments.

When predicting one population based on another, results show that, in general, most populations were well predicted by any other and with any of the models used. As shown in Table 4, some populations (for example, PBW343 × Kingbird and PBW343 × Juchi) were excellent predictors of others when BL or RR models were used.

Conclusions

In summary, the main conclusions drawn from this study may serve to delineate the implementation of genomics in stem rust and yellow rust resistance work in wheat breeding programs. The additive effects of the two traits resulted in the superior prediction ability of the two linear models RR and BL over the SVR with RBF kernel model. Despite the fact that rust resistance is not governed by as many genes as grain yield, correlation values show that it is possible and practical to implement a GS strategy in a rust resistance breeding program.

Results indicate that using data from one environment to predict another correlated environment can be useful in the implementation of a practical GS breeding program and that, even under population structure, good prediction accuracy between populations may be achieved when the TRN and the TST are genetically related. This prediction ability increases when the heritability of the trait increases. Finally, for both stem and yellow rust traits, prediction ability decreased when the sample size of the TRN population decreased. Further research is needed to study genomic-enabled prediction of rust diseases in the presence of genotype × environment interaction and determine whether prediction patterns follow trends similar to those observed for grain yield (Burgueño et al., 2012).

Supplemental Information Available

Supplemental material is included with this article.

Acknowledgments

We acknowledge financial resources from the Durable Rust Resistant Wheat Project led by Cornell University and supported by the Bill & Melinda Gates Foundation. We also thank HPC-Cluster CCT-Rosario for computing time and technical assistance. Leonardo Ornella is a postdoctoral fellow from the National Council for Scientific and Technical Research of Argentina. The authors greatly appreciate the constructive comments made by the associate editor and an anonymous reviewer, which significantly improved the quality of the article.

References

- Bhavani, S., R.P. Singh, O. Argillier, J. Huerta-Espino, S. Singh, P. Njau, S. Brun, S. Lacam, and N. Desmouceaux. 2011. Mapping durable adult plant stem rust resistance to the race Ug99 group in six CIMMYT wheats. Borlaug Global Rust Initiative (BGRI) Technical Workshop, St. Paul, MN. 13–16 June 2011. University of Minnesota and USDA Cereal Disease Lab, St. Paul, MN. <http://www.globalrust.org/db/attachments/knowledge/122/2/Bhavani-revised.pdf> (accessed 6 Jan. 2012).
- 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
- Cooper, M., I.H. DeLacy, and K.E. Basford. 1996. Relationships among analytical methods used to analyze genotypic adaptation in multi-environment trials. In: M. Cooper and G.L. Hammer, editors, *Plant adaptation and crop improvement*. CAB Int., Wallingford, Oxon, UK. p. 193–224.
- 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
- de los Campos, G., D. Gianola, G. Rosa, K. Weigel, and J. Crossa. 2010. Semi-parametric genomic-enabled prediction of genetic values using reproducing kernel Hilbert spaces methods. *Genet. Res.* 92:295–308. doi:10.1017/S0016672310000285
- 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
- Dexter, M. 2007. Eclipse and Java for total beginners companion tutorial document. Eclipse, New York, NY. <http://www.ecliptutorial.sourceforge.net> (accessed 6 Apr. 2012).
- Endelman, J.B. 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Gen.* 4:250–255. doi:10.3835/plantgenome2011.08.0024
- Fox, J. 2005. The R Commander: A basic-statistics graphical user interface to R. *J. Stat. Software* 14:1–42.
- Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten. 2009. The WEKA data mining software: An update. *SIGKDD Explor. Newsl.* 11.
- Hastie, T., R. Tibshirani, and J. Friedman. 2011. *The elements of statistical learning: Data mining, inference, and prediction*. Springer, Springer, New York, NY.
- Heffner, E.L., A.J. Lorenz, J.-L. Jannink, and M.E. Sorrells. 2010. Plant breeding with genomic selection: Gain per unit time and cost. *Crop Sci.* 50:1681–1690. doi:10.2135/cropsci2009.11.0662
- Hickey, L.T., P.M. Wilkinson, C.R. Knight, I.D. Godwin, O.Y. Kravchuk, E.A.B. Aitken, U.K. Bansal, H.S. Bariana, I.H. De Lacy, and M.J. Dieters. 2012. Rapid phenotyping for adult-plant resistance to stripe rust in wheat. *Plant Breed.* 131:54–61. doi:10.1111/j.1439-0523.2011.01925.x
- Jannink, J.J., A.J. Lorenz, and H. Iwata. 2010. Genomic selection in plant breeding: From theory to practice. *Briefings Funct. Genomics* 9:166–177. doi:10.1093/bfgp/elq001
- Long, N., D. Gianola, G. Rosa, and K. Weigel. 2011a. Marker-assisted prediction of non-additive genetic values. *Genetica (The Hague)* 139:843–854. doi:10.1007/s10709-011-9588-7
- Long, N., D. Gianola, G.J. Rosa, and K.A. Weigel. 2011b. Application of support vector regression to genome-assisted prediction of quantitative traits. *Theor. Appl. Genet.* 123:1065–1074. doi:10.1007/s00122-011-1648-y
- Lorenz, A.J., K.P. Smith, and J.-L. Jannink. 2012. Potential and optimization of genomic selection for Fusarium head blight resistance in six-row barley. *Crop Sci.* 52:1609–1621. doi:10.2135/cropsci2011.09.0503
- Mago, R., G. Brown-Guedira, S. Dreisigacker, J. Breen, Y. Jin, R. Singh, R. Appels, E. Lagudah, J. Ellis, and W. Spielmeier. 2011. An accurate

- DNA marker assay for stem rust resistance gene *Sr2* in wheat. *Theor. Appl. Genet* 122:735–744. doi:10.1007/s00122-010-1482-7
- Mann, H. B., and D.R. Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* 18:50–60. doi:10.1214/aoms/1177730491
- 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.
- Njau, P.N., Y. Jin, J. Huerta-Espino, B. Keller, and R.P. Singh. 2010. Identification and evaluation of sources of resistance to stem rust race Ug99 in wheat. *Plant Dis.* 94:413–419. doi:10.1094/PDIS-94-4-0413
- Ober, U., M. Erbe, N. Long, E. Porcu, M. Schlather, and H. Simianer. 2011. Predicting genetic values: A kernel-based best linear unbiased prediction with genomic data. *Genetics* 188:695–708. doi:10.1534/genetics.111.128694
- Park, T., and G. Casella. 2008. The Bayesian Lasso. *J. Am. Stat. Assoc.* 103:681–686. doi:10.1198/016214508000000337
- Pérez, P., G. de los Campos, J. Crossa, and D. Gianola. 2010. Genomic-enabled prediction based on molecular markers and pedigree using the Bayesian linear regression package in R. *Plant Gen.* 3:106–116. doi:10.3835/plantgenome2010.04.0005
- Peterson, R.F., A.B. Campbell, and A.E. Hannah. 1948. A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. *Can. J. Res.* 26:496–500. doi:10.1139/cjr48c-033
- R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org> (accessed 30 June 2012).
- Rutkoski, J., J. Benson, Y. Jia, G. Brown-Guedira, J.-J. Jannink, and M. Sorrells. 2012. Evaluation of genomic prediction methods for Fusarium head blight resistance in wheat. *Plant Gen.* 5:51–61. doi:10.3835/plantgenome2012.02.0001
- Rutkoski, J., E. Heffner, and M. Sorrells. 2011. Genomic selection for durable stem rust resistance in wheat. *Euphytica* 179:161–173. doi:10.1007/s10681-010-0301-1
- SAS Institute. 2004. The SAS system for Windows. Release 9.0. SAS Inst., Cary, NC.
- Shapiro, S.S., and M.B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52:591–611.
- Singh, R.P., D.P. Hodson, J. Huerta-Espino, Y. Jin, S. Bhavani, P. Njau, S. Herrera-Foessel, P.K. Singh, S. Singh, and V. Govindan. 2011. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. *Annu. Rev. Phytopathol.* 49:465–481. doi:10.1146/annurev-phyto-072910-095423
- Singh, R.P., D.P. Hodson, J. Huerta-Espino, Y. Jin, P. Njau, R. Wanyera, S. Herrera-Foessel, and R. Ward. 2008. Will stem rust destroy the world's wheat crop? *Adv. Agron.* 98:271–309. doi:10.1016/S0065-2113(08)00205-8
- Singh, R.P., J. Huerta Espino, and H.M. William. 2005. Genetics and breeding for durable resistance to leaf and stripe rusts in wheat. *Turk. J. Agric. For.* 29:1–7.
- Singh, S., M.V. Hernandez, J. Crossa, P.K. Singh, N.S. Bains, K. Singh, and I. Sharma. 2012. Multi-trait and multi-environment QTL analyses for resistance to wheat diseases. *PLoS ONE* 7(6):e38008. doi:10.1371/journal.pone.0038008
- Smola, A.J., and B. Schölkopf. 2004. A tutorial on support vector regression. *Stat. Comput.* 14:199–222. doi:10.1023/B:STCO.0000035301.49549.88
- Van Raden, P.M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. doi:10.3168/jds.2007-0980
- Wellings, C. 2011. Global status of stripe rust: A review of historical and current threats. *Euphytica* 179:129–141. doi:10.1007/s10681-011-0360-y
- Xu, Y., Y. Lu, C. Xie, S. Gao, J. Wan, and B. Prasanna. 2012. Whole-genome strategies for marker-assisted plant breeding. *Mol. Breed.* 29:833–854. doi:10.1007/s11032-012-9699-6