Australian Journal of Crop Science AJCS

AJCS 10(1):127-136 (2016) ISSN:1835-2707

Genomic prediction for yield improvement and safeguarding of genetic diversity in CIMMYT spring wheat (*Triticum aestivum* L.)

Ephrem Habyarimana

Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria – Centro di ricerca per le colture industriali, Via di Corticella 133 – 40128 Bologna, Italy

***Corresponding author: Ephrem.habyarimana@entecra.it**

Abstract

Simulation studies have demonstrated that genomic selection (GS) can produce superior results in comparison with methods currently used to select the quantitative traits in plants. The evaluation of GS predictive ability in wheat traits is fragmentary. On the other hand, assessment of GS strategies for grain yield improvement in this crop is still limited. This work aimed to evaluating the cross-validation accuracy (r_{cv}) of two GS models and their efficiency in grain yield improvement. The impact of models on genetic diversity was also assessed using pedigree and genotyping-by-sequencing of 19 traits in two multifamily wheat populations. The breeding problem modeled was the performance of newly developed, untested genotypes. Overall prediction accuracy was comparable between models, moderate in a highly structured population ($r_{cv} \approx 0.50$) and high ($r_{cv} \approx 0.70$) in less structured population. This prediction level can sustain wheat breeding programs. Selecting for genetic merit produced a substantial within-generation yield improvement. Furthermore, lines with a good level of genetic diversity were produced. These produced lines displayed mean pairwise kinship ranging from second-to-third to fourth degree relative levels with even lower median kinship values typical of unrelated individuals.

Keywords: Bread wheat; genomic kinship; genomic selection; predictive ability; selection intensity.

Abbreviations: BL_Bayesian LASSO; BRR_Bayesian ridge regression; C29SAWSN_candidates 29 semi-arid wheat screening nursery; C30SAWSN_candidates 30 semi-arid wheat screening nursery; CIMMYT_international maize and wheat improvement center; DArT_diversity arrays technology; GBS_genotyping-by-sequencing; GEBV_genomic estimated breeding value; GS_genomic selection; PBL_Bayesian LASSO with pedigree; PBRR_Bayesian ridge regression with pedigree; PRKHS_reproducing kernel Hilbert spaces with pedigree; QTL_quantitative trait locus; RKHS_reproducing kernel Hilbert spaces; WRS_Wilcoxon rank-sum test.

Introduction

Genomic selection is a new paradigm in population genetic improvement that uses a larger number of genome-wide distributed molecular markers to predict individual breeding values. The GS has demonstrated superior performance in comparison with the methods used in breeding for quantitatively inherited characters yet, i.e., phenotypic selection and quantitative trait loci approaches (Bernardo and Yu, 2007; Heffner et al., 2011; Lorenzana and Bernardo, 2009). The superiority of the GS strategy is mostly associated with higher accuracy in predicting the individual's genetic merit and the shortening of a breeding cycle due to intercrosses driven by genetic predictions, which results in higher genetic gain per unit of cost and time (Heffner et al., 2011; Jannink, 2010). These GS attributes are expected to have wide-range implications in plant breeding as the cost of cultivar development is reduced (Heslot et al., 2012). Therefore, farmers can grow a better variety sooner due to rapid variety development and release, making more income (Meredith and Mantel, 2000; Morris et al., 1992; Pandey and Rajatasereekul, 1999).

Several factors affect GS efficiency, including model performance, sample size and genetic relatedness, marker density, trait heritability and genetic architecture, and the extent and distribution of marker-QTL linkage disequilibrium (Hickey et al., 2014). Therefore, genomic prediction can improve when model assumptions for the distribution of markers accommodate trait complexity, the training population is built and its size eventually increased using genetically close subpopulations, and the marker density accounts for training population size and linkage disequilibrium span (Crossa et al., 2013; Windhausen et al., 2012; Würschum, 2012). Models, including both marker and pedigree information, perform better than when either information is used alone (Crossa et al., 2013), but the effects of trait heritability on accuracy show mixed results (Abera Desta and Ortiz, 2014).

Pedigree-based selection has been practiced successfully since the 20th century (Bell, 1997). The intense selection, germplasm exchange and genetic introgressions (e.g., through wide crosses in wheat) have maintained a substantial amount of genetic diversity in breeding populations. For GS, a higher rate of relatedness and more rapid decline in selection response is likely (Goddard and Hayes, 2009). Kinship is a good metric to assess genetic diversity in agricultural populations (Lopes et al., 2013; Saura et al., 2013). Within a population, individual genomic inbreeding represents the probability that two alleles at a randomly chosen locus are identical by state, whereas pairwise kinship measures the relatedness represented by the probability that two alleles, one sampled at random from each individual, are identical by state (Powell et al., 2013). Therefore, kinship predicts the future level of inbreeding, which represents the repository for future genetic diversity. With the advent of whole genome high-throughput sequencing and genotyping, several authors (Saura et al., 2013) encouraged application of marker loci instead of genealogical coancestry data to measure genetic diversity (Lopes et al., 2013).

This work intends to contribute to ongoing efforts to explain GS issues related to breeding applications. Many of the GS models published yet have been using different sets of few traits, each in diverse wheat populations, preventing breeders from getting the whole picture of GS prediction accuracy in wheat traits of breeding interest. For instance, Crossa et al. (2010) evaluated grain yield using pedigree and diversity array technology (DArT) markers. Poland et al. (2012) evaluated days to heading, thousand-kernel weight, and grain yield using DArT and genotyping-by-sequencing (GBS) markers, whereas Heslot et al. (2013) evaluated days to heading, plant height, grain yield, and pre-harvest sprouting using DArT, single nucleotide polymorphisms, and simple sequence repeat markers. It should also be noted that a few works have assessed GS predicting ability for disease resistance (Würschum, 2012) and bread-making quality traits (Heffner et al., 2011; Heslot et al., 2012). Several studies (e.g., Heffner et al., 2011; Crossa et al., 2013) have investigated GS accuracy, but results of using genomic estimated breeding values (GEBVs) for population improvement and genetic diversity assessment under a real breeding scheme are lacking.

The objectives of this work were to (1) evaluate two GS models in two empirical wheat populations with different levels of population structure using pedigree and GBS marker information in a large number of traits commonly scored in bread wheat breeding programs, and (2) derive GS-selections using individuals genetic merit (GEBV) and assess the within-generation grain yield improvement and genetic diversity in those selections. The cross-validated accuracy was evaluated across 19 traits including traits of agronomic and production importance, bread-making and trading quality traits, and disease resistance traits in wheat (Table S1). These traits have distinct heritability and were evaluated across 12 environments and four years (2009 through 2012) in two multifamily spring wheat populations developed at the international maize and wheat improvement center (CIMMYT).

Results

Environmental indices and trait heritability

The environmental index, as the average grain yield of all varieties included in the trial within an environment (Habyarimana et al., 2004), is a good quantitative measure of the crop growing conditions and is used in the evaluation of varietal adaptability (Finlay and Wilkinson, 1963). In this study, indices ranged from to 7.2 to 1.9 t/ha, respectively, under full irrigation bed and drought flat environments in Obregon (Fig. 1). All the main effects (genotype, nursery, environment, and year) and interactions (equation 2) were highly significant ($p \leq 0.01$). Environments performed differently, with decreasing productivity in the following order: full irrigation bed, full irrigation flat, heat bed, drought bed, and drought flat (Fig 1). Overall, C29SAWSN surpassed C30SAWSN (5.1 *vs*. 4.6 t/ha), and the latter performed particularly poorly under severe drought in drought flat plots.
Trait heritability (Fig. 2) was evaluated as it quantifies the precision of the trials (Piepho and Mohring, 2007). The H^2 could not be computed for the Fusarium head blight index because the trait was scored in one replication. Heritability was low (≤ 0.50) for the alveograph ratio between dough elasticity and extensibility, plant height, and yellow rust; medium to high ($0.50 \leq H^2 < 0.80$) for grain yield, grain protein content, leaf rust, bread loaf volume, test weight, flour protein content, spot blotch, mixographic type, and stem rust; and high ($H2 \geq 0.80$) for days to heading, flour sodium dodecyl sulfate sedimentation, grain hardness, mixograph mixing time, sodium dodecyl sulfate sedimentation index, and thousand-kernel weight.

Models' predictive ability and population structures

Subpopulation structure analysis resulted in the assignment of 97% (3 subpopulations and 7 unstructured genotypes) and 62% (11 subpopulations and 101 unstructured genotypes) of the lines in C29SAWSN and C30SAWSN, respectively (Figs S1 through S4). Subpopulation size ranged from 38 to 174 and 6 to 28 in C29SAWSN and C30SAWSN, respectively. On average, the two populations displayed comparable pairwise subpopulation genetic divergence with *D* values of 0.27 and 0.25 in C29SAWSN and C30SAWSN, respectively.
GS accuracies are summarized in Figs 2, 3, and 4. Accuracy was significantly higher ($p<0.001$) in C29SAWSN than in C30SAWSN for all traits common to both populations, with increases ranging from 10% for yellow rust, to more than 200% for grain yield. In C29SAWSN, accuracy was on average 0.66, and all traits, except plant height ($r_{cv}=0.46$), were predicted with an accuracy greater than 0.50, ranging from 0.52 to 0.79 (Fig 2). In C30SAWSN, the mean accuracy was 0.49, and all traits, except leaf rust ($r_{cv} = 0.37$) and grain yield ($r_{cv} = 0.18$), were predicted with an accuracy greater than 0.50, with a range of 0.56 to 0.63 (Fig 2). In C29SAWSN, Spearman's rank correlation coefficient between GS accuracy and trait heritability was positive and moderately high to high ($\rho \approx 0.7$) and significant ($p<0.01$) for all models. In C30SAWSN, this correlation was low, negative ($-0.09 \leq \rho \leq -0.03$) and not significant ($p> 0.9$).

Grain yield improvement and genetic diversity

Grain yield genetic merit was used for this trait as it is critical for wheat line advancement. The genetic merit may guide to intercrosses and determine the phenotypic performance levels in the progeny. Ten percent selection pressure was applied to select 26 lines in each of the two populations. The GS efficiency was gauged using a selection differential (*i*) expressed in fractions of phenotypic standard deviations (Fig 3), and comparing it with phenotypic selection. In the two populations and across models, grain yield (GY) improvement was substantial ($1.1 \leq i \leq 1.3$). Phenotypic selections showed a significant slight grain yield advantage in C29SAWSN (i=1.6, GY=5.55 vs. 5.47 t/ha, $p \leq 0.05$) and C30SAWSN (i=1.8, GY =5.2 vs. 5.0 t/ha, $p \leq 0.001$) relative to GS-based selections, and bootstrap and WRS tests led to similar inferences.

Summary statistics for the genomic pairwise kinship coefficients in the populations and in selections are presented in Fig 4. C29SAWSN and C30SAWSN were not statistically different ($p> 0.9$) in terms of kinship (both had $\Phi= -0.0032$). Negative kinship values indicate greater un-relatedness. PBL- and PRKHS-based selections showed comparable ($p>0.1$) mean kinship at the third-to-fourth-degree relatives level in C29SAWSN ($\Phi = 0.075$ *vs*. 0.107), whereas PRKHS selections were more related ($\Phi = 0.140$ *vs*. 0.100, $p \leq 0.01$) than PBL selections in C30SAWSN. For the two models evaluated, the degree of pairwise relationship (Φ) was statistically ($p \leq 0.01$) higher in selections than in the respective whole populations, and in C30SAWSN-selected lines than in C29SAWSN selections ($p \leq 0.05$).

Discussion

Several simulation studies proved GS can lead to meaningful genetic progress in plant breeding but empirical evidence is still limited (Abera Desta and Ortiz, 2014). Available results in bread wheat are still fragmented and mainly derived from investigations carried out under different conditions, with GEBVs estimated for a few traits at a time. This lacks fair comparisons of the predictive ability for the full set of traits commonly measured and bred for in this crop. This study addresses this issue, modeling the prediction of the genetic merit of newly developed and untested genotypes for a larger set of traits of breeding interest, and assessing the potential for within-generation grain yield improvement and safeguarding genetic diversity in bread wheat.

Adaptability of wheat populations and effects of trait heritability on GS accuracy

The environmental indices quantitatively graded the trials growing conditions, allowing the identification of wheat nurseries specifically adapted to favorable or stressed environments, or those show general adaptability. Overall, and under severe drought, C29SAWSN outperformed C30SAWSN; but under non-water-limiting conditions, the two nurseries performed comparably, suggesting a more general adaptability of C29SAWSN. The CIMMTY germplasm is well-known for its widely adapted genotypes. These findings highlight the need to make selections under both favorable and unfavorable conditions to account for potential genotype × environment interaction and make informed breeding decisions.

Overall trait heritability was high ($H^2 \approx 0.70$) in the two evaluated populations, reflecting the precision of the trials. Traits were measured in up to 12 environments and four years, which made the estimate of phenotypic means closer approximations to the true genotypic values (Piepho and Mohring, 2007). Heritability is important in breeding. A low heritability of trait indicates that effective selection for the trait of interest would require pedigree information and progeny testing, while high heritability indicates that selection using single plots might be effective (Teich, 1984). Heritability estimates obtained in this study were comparable to findings reported in the literature (Manès et al., 2012; McClung and Cantrell, 1986; Ornella et al., 2012; Poland et al., 2012). The lower plant height H^2 values obtained in this study can be explained by the high levels of soil moisture stress, under which the trait was evaluated; i.e., mainly under water-limiting conditions in 71% of the trials. Alveograph ratio between dough elasticity and dough extensibility was the sole bread-making quality trait displaying lower heritability ($H^2 = 0.47$), which evidences its complex genetic determinism (McClung and Cantrell, 1986). Of the disease resistance traits, yellow rust showed the lowest heritability ($H^2 \leq 0.35$), which can be attributed to pathogenic races differing between experimental locations in Mexico and Ecuador.

The GS accuracy was positively associated with trait heritability in C29SAWSN, but no significant correlation was found in C30SAWSN. A high-heritability in traits and genetic factors may explain most of the differences between individual's performance, while in lower heritable traits phenotypes are less reflective of individual's genetic make-up and are consequently less predictable through genomic selection. The GS is mainly seen as an approach to achieve more genetic gains in traits with lower heritability; however, phenotyping for some highly heritable traits can be very expensive (e.g., some quality traits) and genomic predictions of unphenotyped individuals can therefore be of great advantage. Inconsistent relationship patterns between heritability and GS accuracy have been reported in previous studies (Guo et al., 2014). The present study shows that as the genetic structure of the training population becomes complex due to a larger number of independent haplotypes (e.g., in C30SAWSN), the importance of heritability in GS accuracy declines, yet can be compensated for by using a larger number of observations to estimate marker effects.

Performance of GS models and effects of genetic backgrounds

In this study, we implemented GS models using marker and pedigree information as they proved to be superior to integrate information (Burgueño et al., 2012). The PRKHS and PBL were compared for accuracy of the predicted GEBVs, and the statistical inferences did not show significant differences. This finding is in agreement with Crossa et al. (2010) and Zhong et al. (2009) who described the differences in performance between the two models as scarce and dependent upon breeding scenarios. The PRKHS showed a slight non-significant 8% increase in grain yield prediction accuracy over PBL in C29SAWSN, highlighting the complex inheritance of this trait. The Bayesian LASSO is an additive model, whereas the Kernel model can capture complex non-allelic interactions in addition to allelic effects (Gianola and van Kaam, 2008). The overall comparable predictive ability of these models denotes the greater importance of the additive effects of genetic factors governing the traits evaluated in this work (Ornella et al., 2012).

The C29SAWSN was better predicted than C30SAWSN, and since the two populations were comparable for most of the factors influencing GS accuracy (such as population size, marker density, and trait heritability), it can be inferred that the differences in accuracy originated in the complex C30SAWSN stratification. The effects of population structure on the performance of genomic predictions were also reported in Windhausen et al. (2012), Asoro at al. (2011), and Daetwyler et al. (2010). All empirically

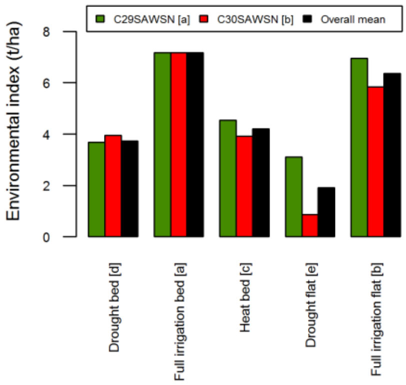

Fig 1. Environmental indice using C29SAWSN and C30SAWSN (29th and 30th CIMMYT semi-arid wheat screening nurseries). Post hoc analysis of environments was carried out on grain yield averaged over nurseries within each environment. Environments or nurseries followed by same letter are not different at the 5% probability level (Tukey's HSD test). Refer to Table S1 for a description of the environments.

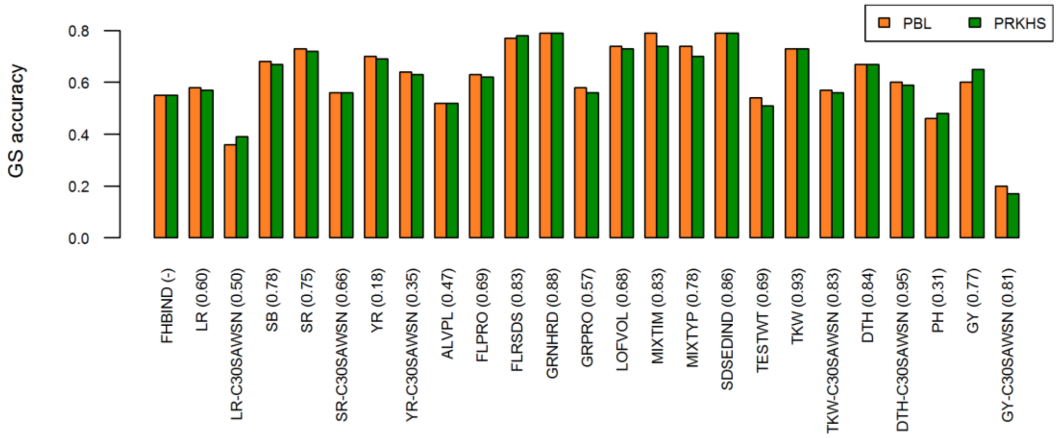

Fig 2. Cross-validation accuracy and heritability (in parentheses) for 25 trait-dataset combinations, using models Bayesian LASSO with pedigree (PBL) and reproducing kernel Hilbert spaces with pedigree (PRKHS) in C29SAWSN and C30SAWSN (29^{th} and 30^{th} CIMMYT semi-arid wheat screening nurseries). FHBIND Fusarium head blight index; LR leaf rust; SB spot blotch; SR stem rust; YR yellow rust; ALVPL alveograph ratio P/L; FLPRO flour protein content; FLRSDS flour sodium dodecyl sulfate sedimentation; GRNHRD grain hardness; GRPRO grain protein content; LOFVOL bread loaf volume; MIXTIM mixograph mixing time; MIXTYP mixographic type; SDSEDIND sodium dodecyl sulfate sedimentation index; TESTWT test weight; TKW thousand kernel weight; DTH days to heading; PH plant height; GY grain yield. Refer to Table S1 for a description of the traits. Traits collected from C30SAWSN are suffixed with this name.

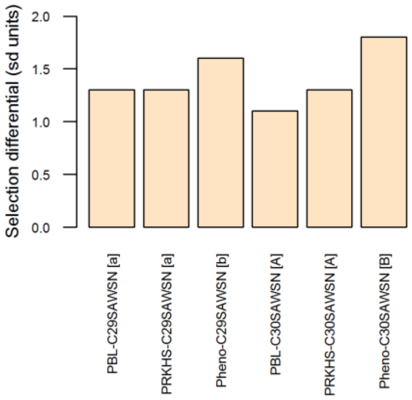

Fig 3. Selection intensity for grain yield in C29SAWSN and C30SAWSN (29th & 30th CIMMYT semi-arid wheat screening nurseries), using phenotypic selection (pheno), and Bayesian LASSO with pedigree (PBL) and reproducing kernel Hilbert Spaces with pedigree (PRKHS) models with a design matrix based on genotyping-by-sequencing markers. sd: standard deviation. For each population, grain yields in selections (labels) followed by same letter are not different at the 5% probability level using bootstrap and Wilcoxon rank-sum tests.

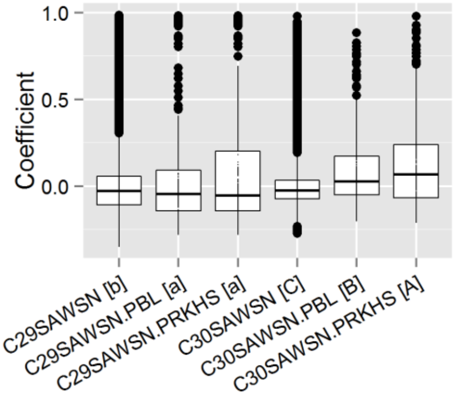

Fig 4. Box plot for pairwise kinship coefficients in the whole population and GS selected lines in C29SAWSN and C30SAWSN (29th & 30th CIMMYT semi-arid wheat screening nurseries), using the Bayesian LASSO with pedigree (PBL) and reproducing kernel Hilbert spaces with pedigree (PRKHS) models with a design matrix based on genotyping-by-sequencing markers. The dots, lower and upper whiskers, and middle line represent outliers, lower and upper quartiles, and the median value, respectively. For each population, groups (labels) followed by same letter are not different at the 5% probability level (bootstrap). Means were ranked in decreasing order.

demonstrated that increasing the number of genetically diverging population strata decreases GS accuracy. Thus, the C29SAWSN was expected to show higher accuracies as it had fewer and larger subpopulations (Fig. S3), meaning that individuals in these subpopulations were better represented in the training sets and hence, better predicted. This was not the case with C30SAWSN, which displayed a greater number of smaller and genetically divergent clusters (Fig S4). These results underline that population structure needs to be taken into account when applying GS in a breeding program, because it could lead to biased assessments of prediction accuracy. In recent studies, several algorithms were tested to optimize training populations and the allocation of resources in GS, but their efficiency was not consistent across populations (Rincent et al., 2012; Isidro et al., 2014). According to Crossa et al. (2013), trait prediction under severe drought stress is more problematic than under well-watered conditions, which suggests that the greater susceptibility to severe drought observed (Fig 1) could have also contributed to lower predictions in C30SAWSN.

Available empirical evidence for GS efficiency in plant breeding set to 0.5 the threshold for GS prediction accuracy for net merit (i.e., overall performance) useful for breeding purposes (Heffner et al., 2010, 2011). The results of this study; therefore, indicate that GS can sustain wheat breeding and allow important genetic gains per unit of time and costs for most of the evaluated traits, except grain yield in the population (C30SAWSN) with complex genetic structure, and plant height in C29SAWSN. This situation suggests that grain yield's complex inheritance might require increasing the size of the training population to improve the likelihood of sampling useful marker-QTL linkages of interest. For the plant height, the low prediction ability can be explained by the possible lack of genotypic variation, as the major plant height gene in CIMMYT germplasm is mainly *Rht1*. Interestingly, the high level of accuracy by which the GEBVs for end-use quality traits were predicted, means that unlike conventional breeding, GS could allow early selection for genetic merit for these traits (flour protein content, flour sodium dodecyl sulfate sedimentation, grain hardness, grain protein content, mixograph mixing time, mixographic type, sodium dodecyl sulfate sedimentation index, loaf volume, thousand-kernel weight, and test weight) without requiring time consuming and costly biochemical quality testing.

GS potential for wheat improvement

GS is an attractive breeding approach as it allows the prediction of genetic merit, upon which a superior parental line is selected for use in intercrosses and the success for a potential new cultivar is assessed before it is field tested. This contributes to the shortening of a breeding cycle, rapid variety development and release, and higher genetic gain per unit of cost and time (Heffner et al., 2011; Jannink 2010). In this study, the GS was applied to improve wheat grain yields in C29SAWSN and C30SAWSN, mimicking breeding for grain yield potential, which was the original priority in the evaluated nurseries. The superiority of individuals was measured at the level of the selected breeding populations using a selection differential, a key tool breeders use to predict the response to selection (Piepho and Mohring, 2007). Overall, GS models were effective in selecting for grain yield (Fig 3). The C29SAWSN and C30SAWSN showed very small phenotypic variance for this trait ($\sigma^2_{GY} \leq 0.07$), indicating that individual performances tended to be very close to each other, which might have limited grain yield selection differential. The slightly higher yield performance observed in phenotypic selections can be attributed to this strategy accounting for non-additive effects of Mendelian factors and non-genetic (environmental factors) effects, which nonetheless are not inheritable and; hence, in contrast to GS selections, cannot dependably drive intercrosses in breeding programs. The GS-based yield improvement reached 7% and 6-7%, respectively in C29SAWSN and C30SAWSN populations, which exceeds the yearly overall progress reported in a recent study (Manès et al., 2012) that was based on phenotypic trials related to ours. Thus, at the within-generation stage of the breeding cycle, GS shows competitive potential relative to conventional breeding and remains an attractive strategy, particularly as it obviates the need for phenotyping the selection candidates, which shortens the length of the breeding cycle (Heffner et al., 2010). The lower prediction accuracy achieved in C30SAWSN did not impact yield improvement relative to the better predicted C29SAWSN nursery.

Conventional genetic improvement is frequently blamed for the genetic bottleneck in plants (Bai and Lindhout, 2007). Kinship is a good metric for assessing genetic diversity in populations. Kinship values were comparable between C29SAWSN and C30SAWSN. PBL- and PRKHS-based selections, showing comparable mean kinship in C29SAWSN, whereas PRKHS selections were more related (higher Φ) than PBL selections in C30SAWSN. Across models, the degree of pairwise relationship was statistically higher in selections than in the respective whole populations. Since most of the predictive ability is due to linkage disequilibrium between QTL and marker loci (Long et al., 2011), the similar performance of PRKHS and PBL in C29SAWSN can be explained by the broader linkage disequilibrium span (Riedelsheimer et al., 2012) common in wheat (Dreisigacker et al., 2012). In the highly structured C30SAWSN population, the higher kinship values in PRKHS-selected lines relative to PBL suggest an increased effect of population structure. RKHS can capture population structure and substructures, as well as within and between family means (de los Campos et al., 2010) due to the nontrivial correlation structure introduced by the reproducing kernel matrix K_h (equation 5) (Gianola and van Kaam, 2008). This is analogous to utilizing information regarding similarities between individuals (kinship) for prediction purposes (Gianola and van Kaam, 2008) and can increase relatedness in selections (Heslot et al., 2012). The use of markers in linkage disequilibrium-based models was expected to provide information on the Mendelian sampling terms and to add the impact of pedigree information, combining between- and within-subpopulation selection, resulting in lower pairwise genomic kinship in the selected lines. The observed higher levels of kinship in the selections relative to the base populations can therefore be explained by (1) the high selection pressure applied in this work, and (2) several individuals being selected from each of the few better predicted subpopulations just as in conventional breeding, where between-family selection saves a few families to make up the future breeding population (Hallander and Waldmann, 2009; Hallauer et al., 1988). The higher kinship values observed in C30SAWSN selections (second-to-third *vs.* third-to-fourth degree relatives level) relative to selections in C29SAWSN imply that in the highly structured C30SAWSN population, individuals were mostly selected from fewer better predicted subpopulations. Higher pairwise relatedness in selections would lead to increased levels of faster inbreeding in a breeding program, which would jeopardize the genetic diversity. Our work evaluated the within-generation genetic diversity and recommends that GS long term effects be further investigated.

Materials and Methods

Plant materials

Plant materials consisted of 264 and 268 elite lines that were the CIMMYT candidates for the 29th (C29SAWSN) and 30th (C30SAWSN) semi-arid bread wheat screening nurseries, respectively. CIMMYT developed the semi-arid wheat nurseries in response to the needs expressed by farmers in drought-prone areas (Manès et al., 2012). The crossing history showed that C30SAWSN was produced from 125 unique crosses with 1 to 15 sister lines per cross, whereas C29SAWSN was derived from 121 unique crosses with 1 to 12 sister lines per cross. The pedigrees of two C30SAWSN lines could not be accessed. A total of 25 dataset-trait combinations were analyzed, 19 and 6 of which were collected from C29SAWSN and C30SAWSN, respectively (Table S1). Traits were measured throughout the 2009, 2010, 2011, and 2012 cropping seasons at six CIMMYT test sites: 1) El Batan, State of Mexico, Mexico; 2) CENEB-Obregon, Sonora, Mexico; 3) Toluca and Boximo, State of Mexico, Mexico; 4) Agua Fria, Puebla, Mexico; 5) Njoro, Kenya, and 6) Santa Catalina, Ecuador. Variance components and trait broad sense heritability (H^2), hereafter referred to as heritability) were estimated by fitting the linear mixed model equation $y_{ij} = \mu + g_i + e_{ij}$ (1)
i = 1,…., s genotypes, j =1,….., n_i replicates for genotype i, y_{ij} is the response variable for genotype i in replicate j (Ausemus et al., 1967; Crossa et al., 2011; de los Campos et al., 2010). It was assumed that $g_i \sim N(0,\sigma_u^2)$ and $e_i \sim N(0,\sigma_e^2)$; furthermore, g_i and e_{ij} are independent. The model was fitted with restricted maximum likelihood using the R package lme4 (R Core Team, 2013). Yield environmental indices were used to characterize crop growing conditions and adaptability. Environmental indices were defined as the average grain yield of all varieties included in the trial within an environment (Habyarimana et al., 2004) and derived solving the linear model

$$y_{ijkl} = \mu + y_i + s_j + ys_{ij} + g_k + gy_{ki} + gs_{kj} + gsy_{kij} + e_{ijkl} \qquad (2)$$

Where y_{ijkl} is the measurement on plot l, in environment j, year i, containing genotype k. μ, y_i, s_j, ys_{ij}, g_k, gy_{ki}, gs_{kj}, gsy_{kij}, and e_{ijkl} are the overall mean, the effects of year i, environment j, interaction of year i with environment j, genotype k, interaction of genotype k with year i, interaction of genotype k with environment j, interaction of genotype k with environment j and year i, and the plot residual. The effects were considered fixed, and the Tukey's test was used for mean comparisons. Computations were executed using the R software (R Core Team, 2013).

Whole genome genotyping and model-based clustering

GBS markers (Poland et al., 2012) were used to genotype the C29SAWSN and C30SAWSN populations. GBS markers are co-dominant in nature and were coded as -1, 0, 1 for the first allele of homozygote, the heterozygote and the homozygote for the second allele, respectively. The molecular marker matrix was cleaned for non-informative markers, i.e., markers with an allele frequency of less than 5 or more than 95%. In the GBS data set, we observed that 2013 GBS markers with less than 28.8% missing data and the total number of GBS markers (33,762) with up to 80% missing data revealed a comparable predicting power across six molecular marker-based genomic selection algorithms, i.e., Bayesian LASSO (BL), Bayesian ridge regression (BRR), and Reproducing kernel Hilbert spaces (RKHS), each with and without using pedigree as a covariance matrix (Table S2). The cross-validated accuracy is shown for the C29SAWSN dataset using data from trials conducted in Obregon for traits grain yield, days to heading and plant height. The results in Table S2 prompted us to use the smaller marker data set (2,013 loci), which reduced computing time. Residual missing data were imputed using simple marker means (Poland and Rife, 2012).

Genetic diversity was measured with genomic pairwise kinship coefficients (Φ_g) using a GBS marker-derived genomic relationship matrix built according to the first method described in van Raden (2008). Population genetic structure was estimated using a model-based Bayesian clustering method implemented in STRUCTURE software, version 2.3.4 (Pritchard et al., 2000), with 1,666 and 1,664 GBS markers in C29SAWSN and C30SAWSN, respectively. The size of the marker matrices was determined to be comparable between the two populations by setting the ceiling for missing loci at 20.5% and 29% for C29SAWSN and C30SAWSN, respectively. The increased number of markers represented a computational burden but allowed for better resolution of population genetic structures (Inghelandt et al., 2010; Bouchet et al., 2012). The number of subpopulations (K) was determined using 3.10^4 burn-in periods and 7.10^4 iterations, with five independent runs per each K value (from 2 to 20), applying the admixture model and assuming allele frequencies were uncorrelated. Both the posterior probability of the data for a given K (Pritchard et al., 2000) and the Evanno method (Evanno et al., 2005) were used to estimate K. Membership coefficients (Q) were estimated for each individual based on allele frequencies to depict the percentage of its genome that derived from each subpopulation. Wheat lines were highly admixed (Figs S1 and S2); hence, to assign individuals to unique genetic groups (Figs S3 and S4). The threshold for Q was set to 0.5, as suggested by Royo et al. (2010). Pairwise genetic divergence between subpopulations was measured using the net nucleotide distance (*D*) approach, as described in Pritchard et al. (2000). This metric represents the average probability that a pair of alleles (one from each of two subpopulations) is different and takes on a value close to zero in the case of two similar subpopulations.

Genomic selection algorithms

Two of the models commonly used in plant breeding (Crossa et al., 2011; Heffner et al., 2011; Heslot et al., 2012; Perez et al., 2012) namely, Bayesian LASSO (PBL) and reproducing kernel Hilbert spaces (PRKHS), each with pedigree and markers, were evaluated based on the predictive ability for individual genetic merit, grain yield improvement, and genetic diversity in candidates for selection. Pedigree information was included in the models as a covariance matrix for polygenic effects, and represented by the coefficient of parentage. The general structure of the linear model PBL was:

$$y = 1_n\mu + X\beta + Zu + e \qquad (3)$$

Where; $\boldsymbol{y}$ is the vector of the phenotypes with n lines, $\boldsymbol{1_n}$ is a vector of n ones, μ is the population mean, $\boldsymbol{X}$ is the design matrix allocating records to marker effects, $\boldsymbol{\beta}$ is the vector of marker effects, $\boldsymbol{e}$ is a vector of random deviates, $\boldsymbol{u}$ is a vector of polygenic breeding values (polygenic effects) in the model, and $\boldsymbol{Z}$ is a design matrix allocating plant materials to records. The covariance structure of $\boldsymbol{e}$ and $\boldsymbol{u}$ takes on the forms $u \sim N(0, \sigma_u^2 A)$ and $e \sim N(0, \sigma_e^2 I)$, respectively, where $\boldsymbol{A}$ is the additive relationship matrix built from the pedigree of the population, $\boldsymbol{I}$ is the identity matrix, σ_u^2 is the additive genetic variance, and σ_e^2 is the error variance. The corresponding genomic estimated breeding values (GEBV) were derived as:

$$GEBV = \hat{u} + X\hat{\beta} \qquad (4)$$

The polygenic covariate was found to be effective at reducing spurious marker effects and improving the accuracy of breeding value prediction (MacLeod et al., 2010). The use of coefficient of parentage in genomic selection models was detailed in Crossa et al. (2007, 2011). The relationship between genetic covariance of relatives i and i' due to their additive genetic effects, and the coefficient of parentage ($f_{ii'}$) can be described by the equation $2 f_{ii'}\sigma_u^2 = A\sigma_u^2$. The relationship matrix was built using the Browse application of the International Crop Information System (ICIS) as described at http://cropwiki.irri.org/icis/index.php/TDM_GMS_Browse (McLaren et al., 2005).

In Bayesian LASSO (de los Campos et al., 2009b; Park and Casella 2008), marker effects are assigned a double exponential (DE) distribution conditionally on the regularization parameter λ. The effects are centered at zero with marker-specific variance namely, $p(\beta_j \mid \lambda, \sigma_e^2) = DE(\beta_j \mid 0, \frac{\lambda}{\sigma_e^2})$.

The PBL is easily implemented as the double exponential distribution can be represented as a mixture of scaled normal densities (Perez et al., 2012). The priors used in Bayesian LASSO were valorized as suggested in González-Camacho et al. (2012). The reproducing kernel Hilbert space is a semi-parametric nonlinear algorithm that was first heralded by Gianola and van Kaam (2008). Besides the allelic effects, the model can capture allelic and epistatic genetic interactions (Gianola et al., 2006), which may explain its superior predictive accuracy relative to linear models in some situations (Crossa et al., 2011; de los Campos et al., 2009a, 2010). The model was implemented as suggested in Gianola and van Kaam (2008), and González-Recio et al. (2009), as follows:

$$y = 1\mu + K_h\alpha + e \qquad (5)$$

Where $\boldsymbol{y}$ is a vector of phenotypes, $\boldsymbol{\mu}$ is a scalar standing for the overall mean, $\boldsymbol{1}$ is a vector of ones, and K_h is a positive definite matrix, called a reproducing kernel matrix, which is dependent on the bandwidth parameter h:

$$K(x_i, x_{i'}) = \exp(-h\|x_i - x_{i'}\|^2),$$

$x_i = (x_{i1}, \ldots\ldots, x_{ip})'$ and $x_{i'} = (x_{i'1}, \ldots\ldots, x_{i'p})'$ being input vectors whose elements are marker genotypes for individuals i and i', and $\|x_i - x_{i'}\|$ being the Euclidean norm between each pair of input vectors (Lopes et al., 2013); α is a vector of coefficients with distribution $\alpha \sim N(0, K_h^{-1}\sigma_\alpha^2)$, whereas e is a vector of random residuals with distribution $e \sim N(0, I\sigma_e^2)$. To optimize the bandwidth parameter, we implemented the kernel averaging strategy by applying the Bayesian approach, as suggested in de los Campos et al. (2010).

Samples from posterior distributions for PBL and PRKHS were produced using a Gibbs sampler/Metropolis-Hastings algorithm (de los Campos et al., 2009a, 2010). We based our statistical inferences upon 35000 samples obtained after the first 5000 iterations were discarded as burn-in. The models were run using R software, version 3.0.2 (R Core Team, 2013).

Cross-validation prediction accuracy

Phenotypic data were Box-Cox (Box and Cox, 1964) transformed before they were inputted into the models to improve model fit (Kim et al., 2014). The criterion for model prediction accuracy (r_{cv}) was the Pearson correlation coefficient between the phenotypic data and the cross-validated GEBVs. The GEBVs were estimated using a 10-

fold cross-validation by randomly dividing the phenotypic dataset into 10 folds, and alternately removing one fold to be used as a testing set, and training the model on the remaining nine folds (Crossa et al., 2013; Heslot et al., 2012).

Differences in accuracy between models and populations and in the levels of kinship between GS selections, on the one hand, and GS selections and the respective base populations, on the other, were tested using bootstrap with 1000 resamples and Wilcoxon rank-sum test (WRS). We used non-parametric instead of parametric tests (e.g., t-test) as the latter would violate the normality assumption in the original distributions because of the small size of the evaluated samples.

Selection for grain yield genetic merit

Superior individuals were selected based on grain yield GEBV, mimicking selection for yield potential of unphenotyped candidates as would be carried out under the actual GS regime. Mean performance of GS selected lines was compared with selections based on phenotypic values (phenotypic selection) using bootstrapping approach with 1000 resamples and WRS test. Ten percent directional selection pressure (higher values) was applied. Selections were evaluated based on genomic pairwise kinship and the realized selection differential in standard deviation units (i). The latter is an important criterion for evaluating breeding strategies as it is directly related to selection response (Fehr, 1987).

Conclusions

Genomic selection leverages high-throughput genotyping technology together with efficient algorithms to predict the genetic merit, for which superior individuals are selected. This strategy can significantly reduce phenotypic evaluations and efficiently expedite breeding processes. The objectives of this study were to evaluate the value of GS by implementing two algorithms across 19 traits of wheat breeding interest and two genetic backgrounds using pedigree and GBS marker information. The breeding problem modeled included prediction of the performance of newly developed, untested genotypes that have not been evaluated in open-field trials. The results show that a good level of GS accuracy that can sustain a wheat breeding program can be achieved for most traits commonly bred for in this crop. Population structure plays a significant role for the prediction ability, and this can adversely impact genetic diversity and breeding efficiency in successive generations. Selecting for genetic merit led to a substantial level of within-generation grain yield improvement and genetic diversity, producing a great proportion of unrelated individuals.

Acknowledgements

Financial support was provided by the Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria (C.R.E.A.) and the Bill and Melinda Gates Foundation. The author completed part of this work as a visiting scientist at CIMMYT headquarters in Mexico from January to December 2012, and acknowledges the didactic support received.

References

Abera Desta Z, Ortiz R (2014) Genomic selection: genome-wide prediction in plant improvement. Trends Plant Sci. 19:592-601.

Asoro FG, Newell MA, Beavis WD, Scott MP, Jannink J-L (2011) Accuracy and training population design for genomic selection on quantitative traits in elite North American oats. Plant Genome. 4:132-144.

Ausemus ER, McNeal FH, Schmidt JW (1967) Genetics and inheritance. In: Quisenberry KS, Reitz LP (eds) Wheat and wheat improvement. Am Soc Agron Madison, Wisc.

Bai Y, Lindhout P (2007) Domestication and breeding of tomatoes: what have we gained and what can we gain in the future? Ann Bot-London. 100:1085-1094.

Bell G (1997) Selection: The mechanism of evolution. Chapman & Hill, New York.

Bernardo R, Yu J (2007) Prospects for genomewide selection for quantitative traits in maize. Crop Sci. 47:1082-1090.

Box GEP, Cox DR (1964) An analysis of transformations (with discussion). J R Stat Soc B. 26:211-252.

Bouchet S, Pot D, Deu M et al (2012) Genetic structure, linkage disequilibrium and signature of selection in sorghum: Lessons from physically anchored DArT markers. PLOS ONE. doi:10.1371/journal.pone.0033470.

Burgueño J, de los Campos G, Weigel K, Crossa J (2012) Genomic prediction of breeding values when modeling genotype-environment interaction using pedigree and dense molecular markers. Crop Sci. 52:707-719.

Crossa J, Burgueño J, Dreisigacker S et al (2007) Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. Genetics. 177:1889-1913.

Crossa J, de los Campos G, Perez P et al (2010) Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics. 186:713-724.

Crossa J, Perez P, de los Campos G, Mahuku G, Dreisigacker S, Magorokosho C (2011) Genomic selection and prediction in plant breeding. J Crop Improv. 25:239-261.

Crossa J, Perez P, Hickey J et al (2013) Genomic prediction in CIMMYT maize and wheat breeding programs. Heredity. 1-13.

Daetwyler HD, Hickey JM, Henshall JM, Dominik S, Gredler B, van der Werf JHJ, Hayes BJ (2010) Accuracy of estimated genomic breeding values for wool and meat traits in a multi-breed sheep population. Anim Prod Sci. 50:1004-1010. doi:10.1071/AN10096.

de los Campos G, Gianola D, Rosa GJM (2009a) Reproducing kernel Hilbert spaces regression: a general framework for genetic evaluation. J Anim Sci. 87:1883-1887.

de los Campos G, Naya H, Gianola D et al (2009b) Predicting quantitative traits with regression models for dense molecular markers and pedigrees. Genetics. 182:375-385.

de los Campos G, Gianola D, Rosa GJM, Weigel KA, Crossa J (2010) Semi-parametric genomic-enabled prediction of genetic values using reproducing kernel Hilbert spaces methods. Genet Res. 92:295-308.

Dreisigacker S, Shewayrga H, Crossa C et al (2012) Genetic structures of the CIMMYT international yield trial targeted to irrigated environments. Mol Breeding. 29:529-541.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol Ecol. 14:2611-2620.

Finlay KW, Wilkinson GM (1963) The analysis of adaptation in a plant-breeding programme. Aust J Agr Res. 14:742-754.

Fehr RW (1987) Principles of cultivar development, Vol. 1. Macmillan Pub Co, New York.

Gianola D, Fernando RL, Stella A (2006) Genomic-assisted prediction of genetic value with semiparametric procedures. Genetics. 173:1761-1776.

Gianola D, van Kaam JBCHM (2008) Reproducing kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. Genetics. 178:2289-2303.

Goddard ME, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet. 10:381-391.

González-Camacho J, de los Campos G, Perez P et al (2012) Genome enabled prediction of genetic values using radial basis function neural networks. Theor Appl Genet. 125:759-771.

González-Recio O, Gianola D, Rosa GJM., Weigel KA, Kranis A (2009) Genome assisted prediction of a quantitative trait measured in parents and progeny: application to food conversion rate in chickens. Genet Sel Evol. 41:3.

Guo Z, Tucker DM, Basten CJ et al (2014) The impact of population structure on genomic prediction in stratified populations. Theor Appl Genet. 127:749-762.

Habyarimana E, Laureti D, de Ninno M, Lorenzoni C (2004) Performances of biomass sorghum [*Sorghum bicolor* (L.) Moench] under different water regimes in Mediterranean regions. Ind Crop Prod. 20:23-28.

Hallander J, Waldmann P (2009) Optimization of selection contribution and mate allocations in monoecious tree breeding program. BMC Genet. 10:70. doi: 10.1186/1471-2156-10-70.

Hallauer AR, Carena MJ, Miranda FJB (1988) Quantitative genetics in maize breeding, 2nd edn.

Springer Science and Business Media LLC2010, New York.

Heffner EL, Jannink J-L, Sorrells ME (2011) Genomic selection accuracy using multifamily prediction models in a wheat breeding program. Plant Genome. 4:65-75.

Heffner EL, Lorenz AJ, Jannink J-L, Sorrells ME (2010) Plant breeding with genomic selection: Gain per unit time and cost. Crop Sci. 50:1681-1690.

Heslot N, Rutkoski J, Poland J, Jannink J-L, Sorrells ME (2013) Impact of marker ascertainment bias on genomic selection accuracy and estimates of genetic diversity. PLOS ONE. doi:10.1371/journal.pone.0074612.

Heslot N, Yang H-P, Sorrells ME, Jannink J-L (2012) Genomic selection in plant breeding: A comparison of models. Crop Sci. 52:146-160.

Hickey JM, Dreisigacker S, Crossa J et al (2014) Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation. Crop Sci. 54:1476-1488.

Inghelandt DV, Melchinger AE, Lebreton C, Stich B (2010) Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. Theor Appl Genet. 120:1289-1299.

Isidro J, Jannink J-L., Akdemir D, Poland J, Heslot N, Sorrells ME (2014) Training set optimization under population structure in genomic selection. Theor Appl Genet. 128:145-58. doi: 10.1007/s00122-014-2418-4.

Jannink J-L (2010) Dynamics of long-term genomic selection. Genet Sel Evol. 42:35. doi: 10.1186/1297-9686-42-35.

Kim ES, Sonstegard TS, Silva MV, Gasbarre LC, Van Tassell CP (2014) Identification of quantitative trait loci affecting gastrointestinal parasite resistance in an experimental Angus population. Anim Genet. 45:117-121. doi: 10.1111/age.12101.

Long N, Gianola D, Guilherme R, Kent W (2011) Application of support vector regression to genome-assisted prediction of quantitative traits. Theor Appl Genet. 123:1065-1074.

Lopes MS, Silva FF, Harlizius B, Duijvesteijn N, Lopes PS et al (2013) Improved estimation of inbreeding and kinship in pigs using optimized SNP panels. BMC Genet. 14:92. doi: 10.1186/1471-2156-14-92.

Lorenzana RE, Bernardo R (2009) Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor Appl Genet. 120:151-161.

MacLeod IM, Hayes BJ, Savin KW, Chamberlain AJ, McPartlan HC, Goddard ME (2010) Power of a genome scan to detect and locate quantitative trait loci in cattle using dense single nucleotide polymorphisms. J Anim Breed Genet. 127:133-42.

Manès Y, Gomez HF, Puhl L, Reynolds M, Braun H-J, Trethowan R (2012) Genetic yield gains of the CIMMYT international semi-arid wheat yield trials from 1994 to 2010. Crop Sci. 52:1543-1552.

McClung AM, Cantrell RG (1986) Inheritance of glume color and gluten strength in durum wheat. Euphytica. 35:885-890.

McLaren CG, Bruskiewich RM, Portugal AM, Cosico AB (2005) The international rice information system. A platform for meta-analysis of rice crop data. Plant Physiol. 139:637-642.

Meredith JR, Mantel SJ (2000) Project management: A managerial approach, 4th edn. John Wiley & Sons Inc, New York.

Morris ML, Dubin HJ, Pokhrel T (1992) Returns to wheat research in Nepal. CIMMYT economics program working paper 92/04, CIMMYT, DF, Mexico.

Ornella L, Singh S, Perez P et al (2012) Genomic prediction of genetic values for resistance to wheat rusts. Plant Genome-US. 5:136-148.

Pandey S, Rajatasereekul S (1999) Economics of plant breeding: the value of shorter breeding cycles for rice in Northeast Thailand. Field Crop Res. 64:187-197.

Park T, Casella G (2008) The Bayesian LASSO. J Am Stat Assoc. 103:681-686.

Perez P, Gianola D, González-Camacho JM, Crossa J, Manès Y, Dreisigacker S (2012) A comparison between linear and non-parametric regression models for genome enabled prediction in wheat. G3-Genes Genom Genet. 2:1595-1605.

Piepho HP, Mohring J (2007) Computing heritability and selection response from unbalanced plant breeding trials. Genetics. 177:1881-1888.

Poland J, Endelman J, Dawson J et al (2012) Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genome. 5:103-113.

Poland JA, Rife TW (2012) Genotyping-by-sequencing for plant breeding and genetics. Plant Genome. 5:92-102.

Powell JE, Henders AK, McRae AF et al (2013) Congruence of additive and non-additive effects on gene expression estimated from pedigree and SNP data. PLOS Genet. doi: 10.1371/journal.pgen.1003502.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics. 155:945-959.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rincent R, Laloë D, Nicolas S et al (2012) Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: Comparison of methods in two diverse groups of maize inbreds (*Zea mays* L.). Genetics. 192:715-728.

Riedelsheimer C, Technow F, Melchinger AE (2012) Comparison of whole-genome prediction models for traits with contrasting genetic architecture in a diversity panel of maize inbred lines. BMC Genomics. 13:452. doi: 10.1186/1471-2164-13-452.

Royo C, Maccaferri M, Alvaro F et al (2010) Understanding the relationships between genetic and phenotypic structures of a collection of elite durum wheat accessions. Field Crop Res. 119:91-105.

Saura M, Fernandez A, Rodriguez MC et al (2013) Genome-wide estimates of coancestry and inbreeding in a closed herd of ancient Iberian pigs. PLOS ONE. doi:10.1371/journal.pone.0078314.

Teich AH (1984) Heritability of grain yield, plant height and test weight of a population of winter wheat adapted to Southwestern Ontario. Theor Appl Genet. 68:21-23. doi: 10.1007/BF00252304.

van Raden PM (2008) Efficient methods to compute genomic predictions. J Dairy Sci. 91:4414-23. doi: 10.3168/jds.2007-0980.

Windhausen VS, Atlin GN, Hickey JM et al (2012) Effectiveness of genomic prediction of maize hybrid performance in different breeding populations and environments. G3-Genes Genom Genet. 2:1427-1436.

Würschum T (2012) Mapping QTL for agronomic traits in breeding populations. Theor Appl Genet. 125:201-210.

Zhong S, Dekkers JCM, Fernando RL, Jannink J-L (2009) Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: A barley case study. Genetics. 182:355-364.