# Genomic selection for grain yield and quality traits in durum wheat

**Article** *in* Molecular Breeding · May 2018

DOI: 10.1007/s11032-018-0818-x



**8 authors**, including:

Jemanesh Haile
University of Saskatchewan
**60** PUBLICATIONS **334** CITATIONS
SEE PROFILE

Amidou N'Diaye
University of Saskatchewan
**122** PUBLICATIONS **1,257** CITATIONS
SEE PROFILE

Ron Knox
Agriculture and Agri-Food Canada
**247** PUBLICATIONS **5,468** CITATIONS
SEE PROFILE

Jessica Rutkoski
Cornell University
**61** PUBLICATIONS **3,793** CITATIONS
SEE PROFILE

Mol Breeding (2018) 38:75
https://doi.org/10.1007/s11032-018-0818-x

# Genomic selection for grain yield and quality traits in durum wheat

**Jemanesh K. Haile · Amidou N'Diaye · Fran Clarke · John Clarke · Ron Knox · Jessica Rutkoski · Filippo M. Bassi · Curtis J. Pozniak**

Received: 9 November 2017 / Accepted: 3 April 2018


**Abstract** The prediction accuracies of genomic selection depend on several factors, including the genetic architecture of target traits, the number of traits considered at a given time, and the statistical models. Here, we assessed the potential of single-trait (ST) and multi-trait (MT) genomic prediction models for durum wheat on yield and quality traits using a breeding panel (BP) of 170 varieties and advanced breeding lines, and a doubled-haploid (DH) population of 154 lines. The two populations were genotyped with the Infinium iSelect 90K SNP assay and phenotyped for various traits. Six ST-GS models (RR-BLUP, G-BLUP, BayesA, BayesB, Bayesian LASSO, and RKHS) and three MT prediction approaches (MT-BayesA, MT-Matrix, and MT-SI approaches which use economic selection index as a trait value) were applied for predicting yield, protein content, gluten index, and alveograph measures. The ST prediction accuracies ranged from 0.5 to 0.8 for the various traits and models and revealed comparable prediction accuracies for most of the traits in both populations, except BayesA and BayesB, which better predicted gluten index, tenacity, and strength in the DH population. The MT-GS models were more accurate than the ST-GS models only for grain yield in the BP. Using BP as a training set to predict the DH population resulted in poor predictions. Overall, all the six ST-GS models appear to be applicable for GS of yield and gluten strength traits in durum wheat, but we recommend the simple computational models RR-BLUP or G-BLUP for predicating single trait and MT-SI for predicting yield and protein simultaneously.

**Keywords** Genomic selection · Quality traits · GS models · Multi-trait · Selection index · *Triticum turgidum* L. var. *durum*

**Abbreviations**

| | |
|---|---|
| BL | Bayesian Least Absolute Shrinkage, and Selection Operator |
| BLUP | Best linear unbiased prediction |
| BP | Breeding panel |
| CIMMYT | International Maize and Wheat Improvement Center |
| CV | Cross-validation |

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s11032-018-0818-x) contains supplementary material, which is available to authorized users.

Fran Clarke is retired.

J. K. Haile (✉) · A. N'Diaye · J. Clarke · C. J. Pozniak (✉)
Department of Plant Sciences, Crop Development Centre, University of Saskatchewan, Saskatoon, SK, Canada
e-mail: jemanesh.haile@usask.cae-mail: curtis.pozniak@usask.ca

F. Clarke · R. Knox
Swift Current Research and Development Centre, Agriculture and Agri-Food Canada, Swift Current, SK, Canada

J. Rutkoski
International Rice Research Institute (IRRI), 4030 Los Baños, Philippines

F. M. Bassi
International Center for Agricultural Research in the Dry Areas (ICARDA), Av. Mohamed Alaoui, 10000 Rabat, Morocco

Published online: 26 May 2018

| | |
|---|---|
| DAPC | Discriminant analysis of principal components |
| DH | Doubled-haploid |
| G-BLUP | Genomic best linear unbiased prediction |
| G × E | Genotype by environment interaction |
| GEBV | Genomic estimated breeding value |
| GI | Gluten index |
| GPC | Grain protein content |
| GS | Genomic selection |
| GYLD | Grain yield |
| L | Average abscissas at rupture (extensibility) |
| LD | Linkage disequilibrium |
| MAF | Minor allele frequency |
| MT | Multi-traits |
| NIR | Near infra-red |
| P | Dough overpressure (tenacity) |
| QTL | Quantitative trait loci |
| RKHS | Reproducing kernel Hilbert space |
| RR-BLUP | Ridge regression best linear unbiased prediction |
| SI | Selection index |
| SNP | Single-nucleotide polymorphism |
| ST | Single trait |
| ST-GS | Single-trait genomic selection |
| TP | Training population |
| VP | Validation population |
| W | Deformation energy (strength) |

## Introduction

Durum wheat (*Triticum turgidum* L. ssp. *durum* Desf. Husn., 2n = 4× = 28; genome AABB) is an important food source, primarily as pasta (Distelfeld et al. 2006). The main traits targeted in durum wheat breeding include grain yield, resistance to priority diseases, sprouting resistance, and end-use quality, such as grain protein content and gluten strength, milling quality (e.g., semolina yield), and color of the wheat, semolina, and pasta (Clarke et al. 1998). Canada is one of the largest producers and exporters of high-quality durum wheat, which makes pasta quality as a primary objective in Canadian durum breeding programs.

Several traits impact pasta quality, but the composition and quality of grain protein content (GPC) determine pasta cooking quality (Dexter and Matsuo 1977) and durum marketing (Guzman et al. 2016). Therefore, wheat breeders have paid the greatest attention to improve GPC and composition in their lines. However, breeding for high-grain protein is challenging. Because it is negatively correlated with grain yield, genetically complex, and its phenotypic expression is strongly influenced by environmental factors (Clarke et al. 2009). Gluten strength is another crucial factor in pasta manufacturing and cooking quality (Feillet and Dexter 1996) because it affects firmness, texture, integrity, and chewiness of the pasta (Subira et al. 2014). Selection for gluten strength in durum breeding programs still largely relies on physical dough tests, such as SDS-sedimentation volume, gluten index, mixograph, and alveograph, some of which are either expensive or time-consuming (Clarke et al. 2010a; Gaines et al. 2006).

Previous mapping studies for gluten strength in durum wheat identified major genomic regions on chromosome 1BS, which harbors genes for glutenin subunits *Glu-B2* and *Glu-B3* (D'Ovidio and Masci 2004). In addition, various quantitative trait loci (QTL) associated with gluten strength have been reported on most of the durum wheat chromosomes (Blanco et al. 1998; Elouafi et al. 2000; Knox et al. 2004; Kumar et al. 2013; Patil et al. 2009; Conti et al. 2011), reflecting the complex inheritance of this trait. Favorable alleles at *Glu-B2* and *Glu-B3* are fixed in most modern durum cultivars. However, there is still considerable variation in phenotypic expression of gluten strength in modern durum cultivars that harbors the favorable *Glu-B2* and *Glu-B3* alleles, which may be due to additional minor effect QTL, epistatic interactions, and genotype-by-environment interactions (Clarke et al. 2010a).

Genomic selection (GS) is another marker-based strategy that incorporates all available marker information simultaneously into a model to predict the breeding values of breeding progenies for selection (Meuwissen et al. 2001). GS allows predicting breeding value for individuals prior to extensive multi-location field testing (Heslot et al. 2015), which enable breeding programs to cull unacceptable lines before time and resources are invested in evaluating undesirable set of germplasm. This improves testing efficiency and results in a greater response to selection over time (Mackay et al. 2015; Bassi et al. 2016; Crossa et al. 2017).

The prediction accuracy in GS depends on several factors, including training population size, genetic relationship between training population and selection candidates, trait complexity (heritability), and marker density (Jia and Jannink, 2012; Zhao et al. 2012; Spindel

et al. 2015; Battenfield et al. 2016). As well, the prediction accuracies of statistical models can impact prediction accuracies because of different underlying assumptions to estimate marker effects. Ridge regression best linear unbiased prediction (RR-BLUP) treats markers homogenously (Endelman 2011; Meuwissen et al. 2001; Whittaker et al. 2000), and genomic best linear unbiased prediction (G-BLUP) considers the contribution of all markers in construction of the genomic relationship matrix (G) (Piepho 2009; VanRaden 2008). Bayesian methods, such as BayesA and BayesB (Meuwissen et al. 2001), and Bayesian least absolute shrinkage and selection operator (BL, Park and Casella 2008) assume that markers do not have common variance (Daetwyler et al. 2010; Pérez et al. 2010). The reproducing kernel Hilbert spaces regression (RKHS) can capture both the additive effects and non-additive interactions among loci by adding a kernel function into the model that includes interactions among marker covariates (Gianola et al. 2006; Pérez and de los Campos 2014). Additionally, prediction models can vary in performance among traits with different genetic architecture (Sallam et al. 2015). BayesB was determined to be more accurate when a smaller number of loci control the trait, whereas RR-BLUP was less sensitive to genetic architecture (Daetwyler et al. 2010). Therefore, it is important to select appropriate models based on genetic architecture of a trait to maximize the accuracy of GEBV predictions.

Most published GS studies compare different statistical models using a single trait. However, breeding populations are evaluated for multiple traits simultaneously and the best individuals are selected for advancement to the next generation based on the best combinations of traits. Some simulation studies (Calus and Veerkamp 2011; Guo et al. 2014; Jia and Jannink 2012; Jiang et al. 2015; Hayashi and Iwata 2013) have reported higher prediction accuracy using multiple-trait (MT) prediction models than single-trait (ST) methods. Using empirical data, however, studies by Bao et al. (2015), Rutkoski et al. (2012) and Schulthess et al. (2016) suggest little advantage of using MT-GS over ST-GS. In the case of within-family selection, both accuracy and efficiency of MT and ST methods were equivalent (Viana et al. 2010).

The International Maize and Wheat Improvement Center (CIMMYT) bread wheat breeding program favors GS over conventional marker-assisted selection due to better response to selection (Battenfield et al. 2016) and outperformed marker-assisted selection for some quality traits, such as milling and baking (Heffner et al. 2011a). However, the quality traits in durum are different with bread wheat due to differences in its end-uses. Bread wheat is milled to flour (whole grain or refined) to produce a large variety of leavened and flat breads and other baking products whereas durum wheat is milled and used to produce semolina (coarse flour) mainly for pasta making.

The main objective of this study was to evaluate the potential of genomic selection in durum wheat breeding with specific aims to: (a) identify appropriate GS models for grain yield and quality traits in durum wheat, (b) compare the prediction accuracy of single-trait and multi-traist GS approaches for predicting grain yield and protein content, (c) assess the effect of genetic relatedness between training and prediction (selection) sets on the accuracy of predictions in using a breeding panel and bi-parental DH durum wheat populations, and (d) determine the number of single-nucleotide polymorphism (SNP) markers required to obtain a reliable accuracy.

## Materials and methods

### Plant material and phenotyping

The present study was conducted on a doubled-haploid (DH) population of 154 lines and breeding panel (BP) of 170 varieties and advanced breeding lines that were used for genome-wide association analyses in previous studies by our group (Pozniak et al. 2012; N'Diaye et al. 2017). The DH population was developed at the Agriculture and Agri-Food Canada, Swift Current, Saskatchewan from $F_1$ seed obtained using W9262-260D3 and Kofa as described by Knox et al. (2000). W9262-260D3 was developed by the Swift Current breeding program from the cross Kyle*2/Biodur, where Biodur is a semi-dwarf from Germany and Kyle is a cultivar from Canada (Townley-Smith et al. 1987). W9262-260D3 has an intermediate gluten strength whereas Kofa is a strong gluten semi-dwarf cultivar from the USA. A total of 154 DH lines and 17 checks, including the parents, were evaluated at five environments (site × year combinations) in 2000, 2001, and 2002 at Swift Current (50.2851° N, 107.7972° W), and in 2001 and 2002 at Regina (50.4452° N, 104.6189° W), Saskatchewan, Canada, as described by Clarke et al. (2010a). Each field

trial was conducted in a randomized complete block design with two replications.

Grain yield (kg $ha^{-1}$) and grain protein content from ground grain (g $kg^{-1}$) were measured on all plots of each trial as described in Clarke et al. (2009). Gluten index was determined on whole meal samples by Approved Method 38-12 (AACC 2000). Grain samples (500 g) from two environments (Regina 2001 and Swift Current 2002) were milled into semolina at the Grains Research Laboratory, Canadian Grain Commission, Winnipeg, Canada, on an Allis-Chalmers laboratory mill in conjunction with a laboratory purifier using the mill flow as described by Dexter et al. (1990). Alveograph curves for semolina were obtained by Standard No. 121 (ICC 2001) using a constant pressure Alveograph model MA82 (Chopin S.A., Villeneuve-la-Garenne, France). Alveograph measures for dough overpressure (tenacity), average abscissas at rupture (extensibility), and deformation energy (strength) were computed by the instrument.

The breeding panel consisted of 170 advanced breeding lines (Supplemental Table S1) tested for two or more years in the official Canadian durum wheat registration trials (Durum wheat Cooperative Test) between 1999 and 2013. As described by Pozniak et al. (2012), the trials were conducted under the auspices of science/industry groups responsible for recommending cultivars for registration by the Canadian Food Inspection Agency at 10–12 locations annually in western Canada and one location in the United States. The breeding lines and checks were evaluated in lattice designs with four replications. Cultivars Hercules (Leisle 1972), Kyle (Townley-Smith et al. 1987), AC Avonlea (Clarke et al. 1998), AC Morse, AC Navigator (Clarke et al. 2000), and Strongfield (Clarke et al. 2005a) were used as checks from 1999 to 2013, while cultivar Commander (Clarke et al. 2005b) was included as an additional check from 2001 to 2013.

Gluten strength was determined on semolina by gluten index and alveograph on single composite grain samples of each genotype within years at the Canadian Grain Commission Grain Research Laboratory using the techniques described previously. The composites included locations with acceptable physical condition each year, and blended to give a target grain protein concentration of 13%. Grain protein content (g $kg^{-1}$) was estimated by near-infrared spectroscopy (NIR) on whole grain composites of each location within years. Grain samples were equilibrated to constant moisture prior to protein measurement, as described by McCaig et al. (1993). Therefore, all the quality traits were measured on yearly composites, so there was no replication within years.

## Phenotypic data analyses

The phenotypic data of the DH population was analyzed using the Proc Mixed procedure of SAS version 9.3 (Littell et al. 1996) on each environment and over all environments. Genotypes were considered fixed, and blocks, replications, locations, years, and all interactions were random. The historical and unbalanced phenotypic data from the breeding population were analyzed using SAS PROC HPMIXED with three models reflecting the different data structures to calculate best linear unbiased predictions (BLUPs). For grain yield data, year, location, replication, and genotype, and their interactions were considered random; for grain protein concentration, years, locations, and genotypes, and interactions were random; for gluten strength traits, genotypes and years were random. The analyses included all 300 genotypes tested in the registration trials, not just the 170 genotypes used in the present study, to provide a better estimate of random variances and covariances (Clarke et al. 2010b; Pozniak et al. 2012). Subsequently, these BLUPS were used for model development and the calculation of the accuracies of the models.

## Genotyping and population structure analysis

Genomic DNA extraction and genotyping with the 90K iSelect assay were done as described in our previous study (N'Diaye et al. 2017). For the DH population, parents were included on multiple chips to assess repeatability across arrays, and polymorphic SNPs between the parents were filtered to remove SNPs that deviated from the expected 1:1 segregation ratio and resulted in 5153 polymorphic SNPs. In the breeding panel, SNPs were further filtered to remove SNPs having more than 25% missing values and a minor allele frequency of less than 5% which resulted in a total of 9752 polymorphic SNPs for analyses.

The structure of the BP was analyzed using the discriminant analysis of principal components (DAPC) approach as implemented in the Adegenet R package (Jombart et al. 2010). In this study, the narrow-sense heritability ($h^2$) of each trait was estimated using the genetic relationship ($G$) matrix calculated based on

SNPs. The variance components were estimated using the package BGLR (Pérez and de los Campos, 2014). Then, $h^2$ was calculated as the additive genetic variance divided by the total phenotypic variance.

## Genomic predictions

We chose six single-trait-based genomic prediction models for this study based on their use of different assumptions on QTL effect distributions resulting in different marker effect distributions (Heslot et al. 2012): ridge regression best linear unbiased predictor (RR-BLUP, Meuwissen et al. 2001; Whittaker et al. 2000), genomic best linear unbiased predictor (G-BLUP, VanRaden 2008), BayesA and BayesB (Meuwissen et al. 2001), Bayesian least absolute shrinkage and selection operator (BL, Park and Casella 2008), and reproducing kernel Hilbert space (RKHS, Gianola et al. 2006). All statistical modeling was performed in the R 3.1.2 environment (R Development Core Team 2015). All Bayesian approaches and the RKHS model were run as single chains of 10,000 iterations, of which the first 2000 were discarded as burn-in, using the package BGLR (de los Campos and Pérez-Rodríguez 2014).

Genomic estimated breeding values (GEBVs) were calculated by first estimating the effects of allele substitutions at the marker loci with statistical models that used phenotypic and genotypic information from the training population (TP). GEBVs for each entry of the validation set (VP) were then calculated by summing up the effects according to the individual's genotypic makeup.

## Cross-validation

The predictive ability of the models was assessed using five-fold cross-validation. The population was split into five subsets, then four of the subsets (80%) were used for the estimation of marker effects (TP), and the remaining subset (20%) was used as VP. Division into the five subsets was done randomly in the DH population and according to the population structure in the BP which was determined using discriminant analysis of principal components into 4 sub-populations (Supplemental Table S1 and Fig. S1). The subpopulation sizes were 45, 39, 52, and 34 for subpopulations 1, 2, 3, and 4, respectively. Thirty lines (20%) as a VP were formed by randomly selecting 9 lines from subpopulation-3 and 7 lines from each of the other subpopulations. The rest of the lines from all of the subpopulations were used as a TP. Then, the GEBVs for each fold were predicted by training the model on the four remaining folds. The process was repeated five times to include all groups as test (validation) set. For each iteration, the accuracy of the model was assessed by Pearson's correlation between GEBVs and phenotypes of individuals in the fifth (validation) set. The average of accuracies among the five iterations is reported.

To evaluate the accuracy of prediction across populations, the prediction model was trained on the genotypic and phenotypic data of the panel of elite breeding lines (BP) as the TP and used to predict the DH population based only on the marker information. Predictive abilities of the models were estimated by Pearson's correlation between the GEBVs and the actual phenotypic data of the DH population. Analyses were conducted with 3141 SNPs that were common to both populations. Because all the models used in this study produced comparable prediction accuracies, only RR-BLUP was used due to its computational easiness. The predictive abilities were compared with the respective within-population prediction using RR-BLUP.

Additionally, the similarities among GS models were calculated based on the cross-validated GEBVs. The GEBVs for each model were taken, and Pearson correlation was computed to examine the similarities and dissimilarities of the tested GS models.

## Two-traits prediction

Two-trait GS analyses were performed for both populations, considering grain yield and protein content, simultaneously. We chose these traits because protein content is a critical economic trait alongside grain yield, since the price of durum wheat in most wheat markets including Canada is determined by the protein content of the grain. Two of the models used for two-trait GS were multi-trait BayesA (MT-BayesA) (Jia and Jannink 2012) and Bayesian multivariate antedependence (MT-Matrix) model, which considers the antedependence parameter as a matrix (Jiang et al. 2015). In addition, we also used the multi-trait genomic prediction model based on economic selection index (SI), also called the Smith-Hazel index proposed by Smith (1936) for plant breeding and by Hazel (1943) for animal breeding.

The index weight ($b$) is calculated as:

$$b = P^{-1}Gv$$

where $P$ is a phenotypic covariance matrix, $G$ is a genetic covariance matrix between the traits in the index, and $v$ is a vector of economic weights (30:1) for grain yield and protein content, respectively, which was derived from the price ratio provided by the Canadian Wheat Board (http://www.cwb.ca/pricing). The additive genetic covariance matrix and phenotypic covariance matrix were estimated using the EMMREML (emmremlMultivariate) package as developed by Akdemir and Godfrey (2015).

An index value ($SI$) for each individual was calculated as follows:

$$SI = bY$$

where $Y$ is the phenotypic value.

The calculated SIs were used for model training and the GEBVs were compared to the measured SIs based on the phenotypic data. Finally, the prediction accuracies from each two-trait prediction approaches were compared with the best ST-GS model for grain yield and protein content.

### Comparison of different marker numbers for prediction accuracy

The effects of marker density on prediction accuracy were performed to assess the minimum number of SNPs required to obtain an accuracy that is equivalent to those obtained using all markers. For such purpose, we randomly sampled a subset of 100, 500, 1000, 2000, 3000, 4000, and 9000 SNPs from the 9752 SNPs in the breeding panel and 100, 500, 1000, 2000, 3000, 4000, and 5000 SNPs from 5153 SNPs in the DH population. For each subset, 1000 replicates were performed. To ensure an accurate comparison, the same genotypes were used as a validation set to calculate GEBVs for all subset of markers. RR-BLUP was used for this analysis.

## Results

### Phenotypic variation

In the breeding population (Supplemental Table S2), grain yield, gluten index, and protein content ranged from 3708 (Kronos) to 4796 kg ha$^{-1}$ (DT772), from 15.2 (DT524) to 90.0% (Commander) and from 13.1 (DT779) to 14.7 g kg$^{-1}$ (DT834), respectively. Alveograph measures of tenacity varied from 35.1 (Plenty) to 134.8 mm (DT760), strength from 117.4 (AC Avonlea) to 316.8 10$^{-4}$ J (DT722) and extensibility from 65.4 (DT760) to 152.0 mm (DT695). The mean values per entry in the DH population are presented in Supplemental Table S3.

Detailed results on population structure in the breeding population have been described in our recent paper (N'Diaye et al. 2017). Briefly, the discriminant analysis of principal component performed on the genotype data clustered the 170 lines of the breeding panel into four subgroups (Supplemental Fig. S1). Group membership of each genotype is shown in Supplemental Table S1. The majority of the lines were of Canadian origin, with additional diversity both in subpopulation-3 and subpopulation-4 due to germplasm introduction from the University of North Dakota and CIMMYT. As presented in Supplemental Fig. S2, in general, the phenotypic variation among the four subpopulations was not statistically significant. However, subpopulation-3 showed the lowest mean gluten index, tenacity, and strength than the other subpopulations.

### Genomic prediction using single-trait and multi-traits GS approaches

The prediction accuracies of RR-BLUP, G-BLUP, BayesA, BayesB, BL, and RKHS models in both the breeding panel and DH population are presented in Fig. 1. The prediction accuracy for extensibility was consistent in both populations irrespective of the models; all other traits showed variable prediction accuracies in the two populations, with grain yield showing the lowest prediction accuracies in the breeding panel irrespective of the models used (Fig. 1). G-BLUP gave the highest prediction accuracy for protein content (0.88) in the breeding panel only. While both BayesA and BayesB models gave better accuracy (0.78–0.84) for gluten index, tenacity extensibility, and strength in the DH population than the breeding panel. Additionally, RKHS model that captures additive and non-additive effects did not outperform the additive models (RR-BLUP, G-BLUP, BayesA, BayesB, and BL) for any of the traits evaluated, regardless of population (Fig. 1). The correlations between prediction accuracies estimated using different models in the breeding panel varied

from 0.72 to 0.99 for grain yield and from 0.87 to 0.99 for gluten index (Supplemental Tables S5a and S5b). The same trend was also observed for the other traits in the DH population as well as for the other traits (data not shown).

There was no significant ($P < 0.05$) difference in prediction accuracy among the three multi-trait GS models (Fig. 2a,b); the multi-trait models were also not statistically different from the single-trait models. However, we found a different pattern of prediction accuracies for grain yield and protein content in the two populations. In the breeding panel, MT-BayesA and MT-Matrix gave higher prediction accuracy for grain yield, but lower for grain protein content than the best ST-GS model; the reverse was observed in the DH population. In general, a marginal improvement in prediction accuracy was found for both traits in the DH population (0.75) using the MT-GS models compared to BP (Fig. 2). This advantage might come from the stronger correlation of GYLD and GPC in the DH population (− 0.35) compared with the BP (− 0.09), as well as better heritability estimates (0.72) of GPC in the DH population (Supplemental Table S4).

### Cross-validation

The fivefold within-population cross-validation values across all six traits varied from 0.47 to 0.72 in the breeding panel and DH population (Fig. 3). To better resemble the activities carried out by a breeding program, we computed cross validation using the breeding panel as a training set for the bi-parental DH population. The prediction accuracies for five of the six traits were positive but less than 0.20, but it was negative for grain yield (Fig. 3), which is likely due to poor relatedness between the training and prediction panels (Fig. 4).

### Effect of marker density on prediction accuracy

The average prediction accuracies for all traits with different marker numbers (100, 500, 1000, 2000, 3000, 4000, 5000, and 9000 SNPs) in the breeding panel and DH population are shown in Fig. 5a and Fig. 5b, respectively. The accuracies were variable among traits and populations with different marker numbers. Accuracies ranged from 0.36 to 0.66 in the breeding panel and from 0.30 to 0.66 in the DH population. In the breeding panel, only 2000 SNPs were needed to get the highest accuracies for gluten index, extensibility, and tenacity as compared with the 3000 SNPs for protein content, 4000 SNPs for grain yield, and 9000 SNPs for strength (Supplemental Table S6a). In the DH population, the number of markers required to get highest accuracies was not consistent depending on trait complexity: 1000 SNPs for extensibility, 2000 SNPs for protein content and tenacity, 3000 SNPs for grain yield and gluten content, and 4000 SNPs for strength (Supplemental Table S6b). In both populations, prediction accuracy for extensibility was less sensitive to varying marker numbers. Overall, at least 500 SNPs in DH population and 1000 SNPs in breeding panel were required to get comparable prediction accuracies with the 9000 SNPs for all the traits. As marker number increased from 100 to 9000, the standard deviation of prediction accuracies was smaller (Fig. 5a and Fig. 5b).

## Discussion

Genomic selection appears promising for accelerating genetic gain per unit time, but it is a relatively new methodology in wheat breeding, with various breeding programs worldwide still working to determine the best strategy for its implementation. Here, we seek to understand the effect of different GS models, trait complexity (relatively simple quality traits vs. complex), genetic background between the training and prediction sets using a bi-parental DH population and breeding panel, and marker density in prediction accuracies. As different GS models differ in their assumptions of genetic architecture, the advantage of using one method over another is still a matter of debate. Some studies reported similar prediction accuracies irrespective of the models used (e.g., Heffner et al. 2011b; Heslot et al. 2015), while other reported different accuracies for different prediction models (e.g., Sallam et al. 2015; Schmidt et al. 2016; Spindel et al. 2015). Overall, prediction accuracies in the present study were comparable across the different models and the two populations, as have been reported several GS studies in diverse species (Clark et al. 2012; Hayes et al. 2009b; Hayes et al. 2009a; Lorenzana and Bernardo 2009; Su et al. 2010; VanRaden et al. 2009) and in wheat (Charmet et al. 2014; Crossa et al. 2010; Heffner et al. 2011b; Lado et al. 2013). However, prediction accuracies for some models were better than others. In the breeding panel, G-BLUP performed better for protein content (0.88) and extensibility (0.62). A similar result was reported

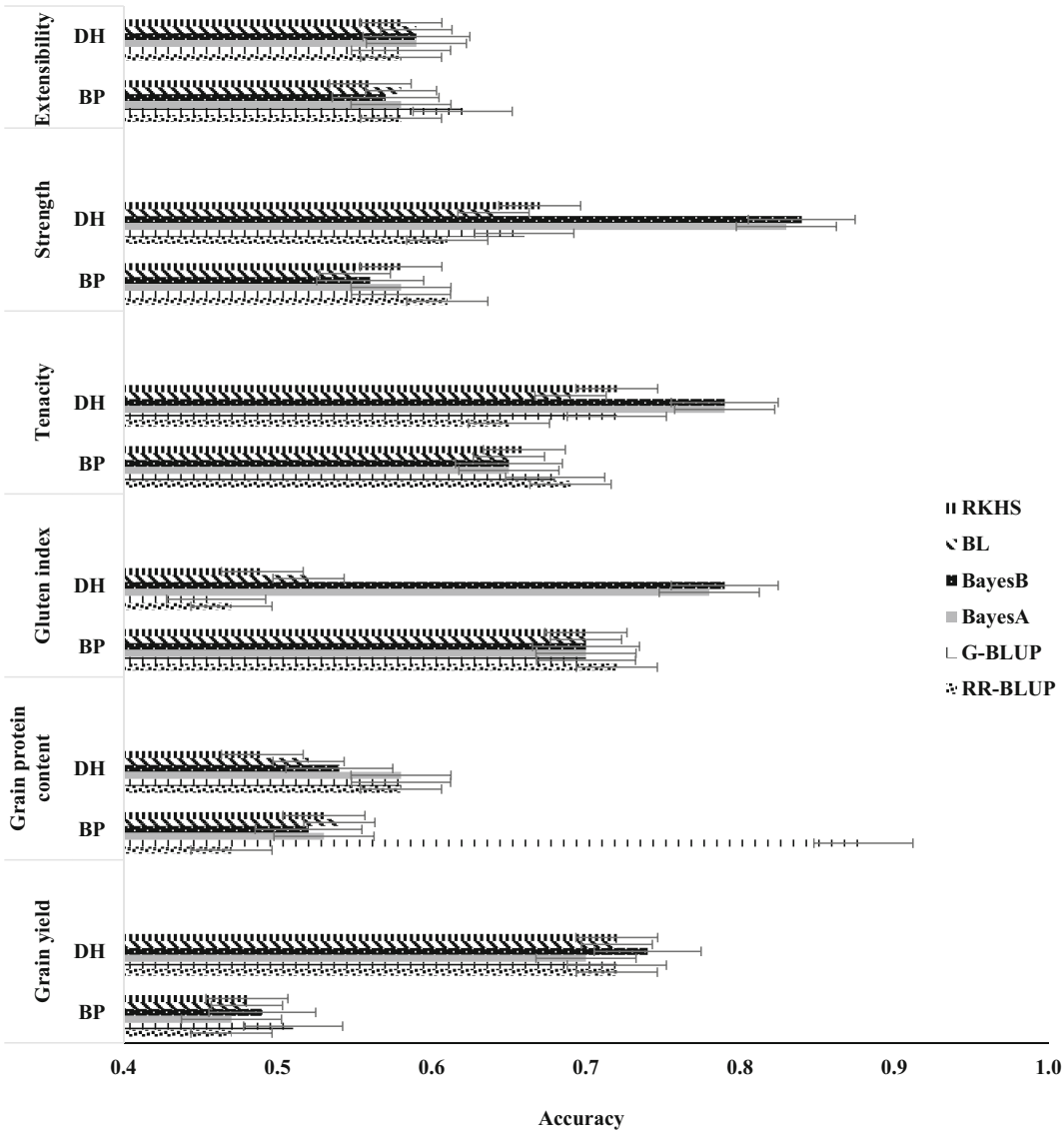



**Fig. 1** Prediction accuracy in a breeding panel (BP) and W9262-260D3/Kofa (DH) population for the six traits using six genomic selection models. Models: RR-BLUP, ridge regression best linear unbiased predictor; G-BLUP, genomic best linear unbiased predictor; BL, Bayesian least absolute shrinkage and selection operator; and RKHS, reproducing kernel Hilbert space

for wheat by Heffner et al. (2011a) for RR-BLUP, which is equivalent to G-BLUP under the assumption that all markers contribute equally to the trait of interest (Habier et al. 2007). Riedelsheimer et al. (2012) did not find any major differences between BL, RR-BLUP, and other models in predicting several traits, including traits with large QTL effects. In the DH population, however, BayesB gave better prediction accuracy for gluten index, tenacity extensibility, which might be partly due to the underlying genetic architecture of the traits.

In a QTL analysis using the same DH population, we identified a major effect QTL on chromosome 1B that explained up to 58 and 48% of the phenotypic variation for gluten index and strength (data not shown). All the six GS models gave very similar prediction accuracies ($r$ = 0.87–1.00) for gluten index than all other traits, which may be due to the major effect QTL on chromosome 1B. BayesB has been reported in identifying a subset of markers with large phenotypic effects that increase prediction accuracy of traits influenced by a few large QTL (Daetwyler et al. 2010; Jannink et al.

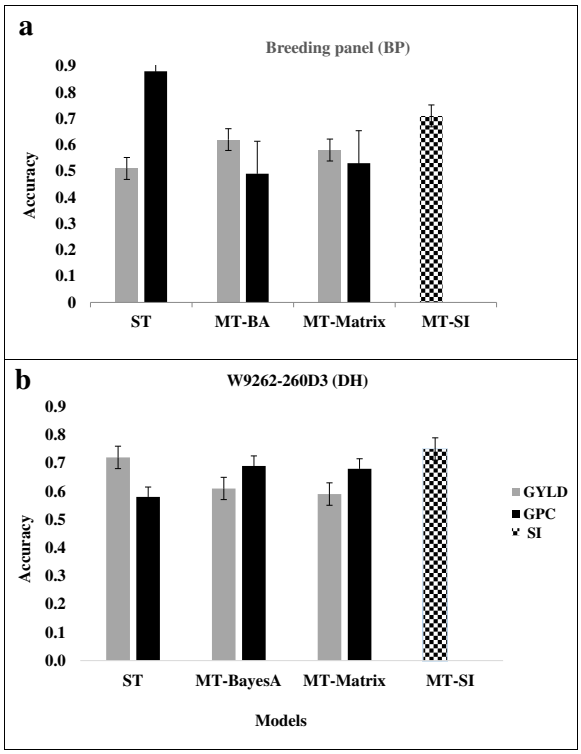



**Fig. 2** Comparison of single-trait (ST) and multiple-trait (MT) prediction accuracies for grain yield (GYLD) and grain protein content (GPC) in the breeding panel (**a**) and W9262-260D3/Kofa (**b**) population. MT-Matrix, two-traits prediction using a matrix antedependence parameter; MT-SI, two-traits predcition using economic selection index. Genomic best linear unbiased predictor (G-BLUP) and BayesA were used for single-trait analyses in the BP and DH population, respectively

2010; Jia and Jannink 2012; Shikha et al. 2017; Thavamanikumar et al. 2015). In the DH population, both BayesA and BayesB gave better prediction accuracies for gluten index, tenacity, and strength, which may be due to preferential in identifying SNPs associated with the above large-effect QTL on chromosome 1B.

Grain yield is a complex trait with low to moderate heritability, which is often affected by G × E interactions than highly heritable qualitative traits (Pozniak et al. 2012). As a result, we expected RKHS model (that incorporates both additive and non-additive effects) to show better prediction on grain yield than all other models that depend only on additive effects. The lack of improved prediction accuracy for grain yield from the RKHS model over all other models used in the present study suggests that most loci may be fixed even in the highly diverse breeding panel. A similar result was reported by Sallam et al. (2015) for wheat and Charmet et al. (2014) for barley.

Regarding grain yield and gluten index predictions using different models, there are two striking results of the different methods. Firstly, the median predictions of all models except G-BLUP are similar (Supplemental Fig. S3), and secondly, the predictions from the four models were highly correlated (Supplemental Tables S5a and S5b). Therefore, we suggest using one or more of the four models for applying GS for grain yield and gluten strength traits in durum wheat.

## Genomic prediction using single-trait and two-traits GS approaches

Durum wheat breeders evaluate advanced breeding lines for a wide range of traits (e.g., agronomic traits, resistance to diseases and pests, quality traits) and advance the best individuals to the next generation based on selection indices that often incorporate multiple traits

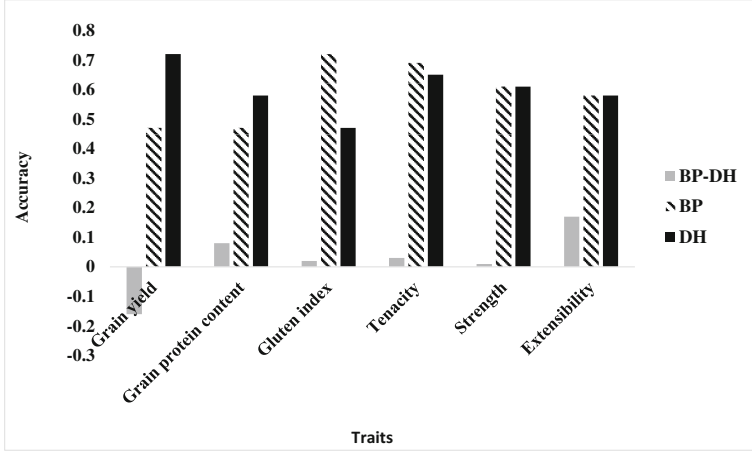



**Fig. 3** Within-population prediction accuracies versus across-population cross-validation prediction accuracies for the six traits based on breeding panel as training population and W9262-260D3/Kofa as validation population (BP-DH) using ridge regression best linear unbiased predictor (RR-BLUP) model

simultaneously. Two-trait selection indices, however, may be sometimes complicated due to biases in assigning appropriate weights for traits of high economic value.

The accuracy of two-traits prediction in durum wheat may be superior over single-trait prediction, but the MT-GS models have not yet been evaluated in this crop based on experimental data. In our results, MT-GS models were more accurate than single-trait models only for grain yield in the breeding panel. The ranking of the models is as follows: MT-SI > MT-BayesA > MT-Matrix > ST (Fig. 2a). Similar results were reported by Jia and Jannink (2012), Jiang et al. (2015), and Schulthess et al. (2016). Selection indices are used to maximize the economic value of animals or plants and are expected to give the most rapid improvement of economic value (Falconer and Mackay 1996), and, therefore, the application of GS to predict SI could potentially bridge the gap between GS and multi-trait improvement (Schulthess et al. 2016).

## Effect of population type on GS

Genomic models have been studied to assess the feasibility of GS for plant breeding to select among sets of elites and cultivars with limited genetic relatedness (Asoro et al. 2011; Charmet et al. 2014; Crossa et al. 2010; Fiedler et al. 2017; Heffner et al. 2011b; Poland et al. 2012; Sallam et al. 2015; Storlie and Charmet 2013; Wang et al. 2014; Würschum et al. 2013), or among narrow-based bi-parental populations (Heffner et al. 2011a; Krchov et al. 2015; Lorenzana and Bernardo 2009), or extremely diverse panels of landraces (Crossa et al. 2010; Crossa et al. 2016; Daetwyler et al. 2014).

Genomic selection for bread wheat grain quality traits in bi-parental populations was shown by Heffner et al. (2011a) to achieve higher prediction accuracies than in multifamily populations. On the other hand, the benefit of using multifamily GS to shorten the breeding

**Fig. 5 a** Effect of marker density on genomic prediction accuracies for grain yield (GYLD), grain protein content (GPC), gluten strength (GI), and alveograph measures [tenacity (P), extensibility (L), and strength (W)]. The x-axis shows the subset of markers (100, 500, 1000, 2000, 3000, 4000, and 9000 SNPs) that were randomly selected from 9752 SNPs from the breeding population (BP). **b** Effect of marker density on genomic prediction accuracies for grain yield (GYLD), grain protein content (GPC), gluten strength (GI), and alveograph measures [tenacity (P), extensibility (L), and strength (W)]. The x-axis shows the subset of markers (100, 500, 1000, 2000, 3000, 4000, and 5000 SNPs) that were randomly selected from 5153 SNPs from W9262-260D3/Kofa (DH) population

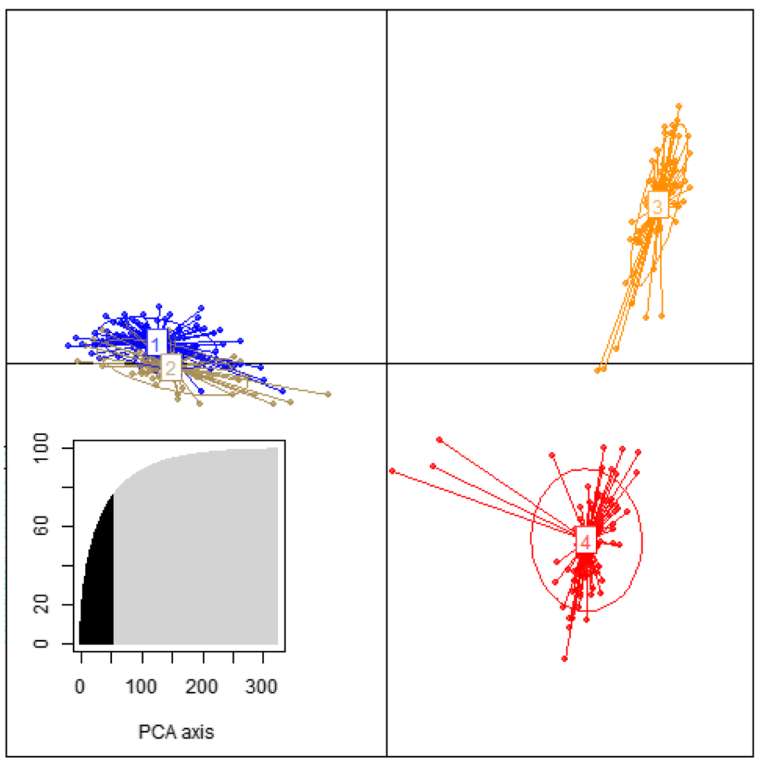
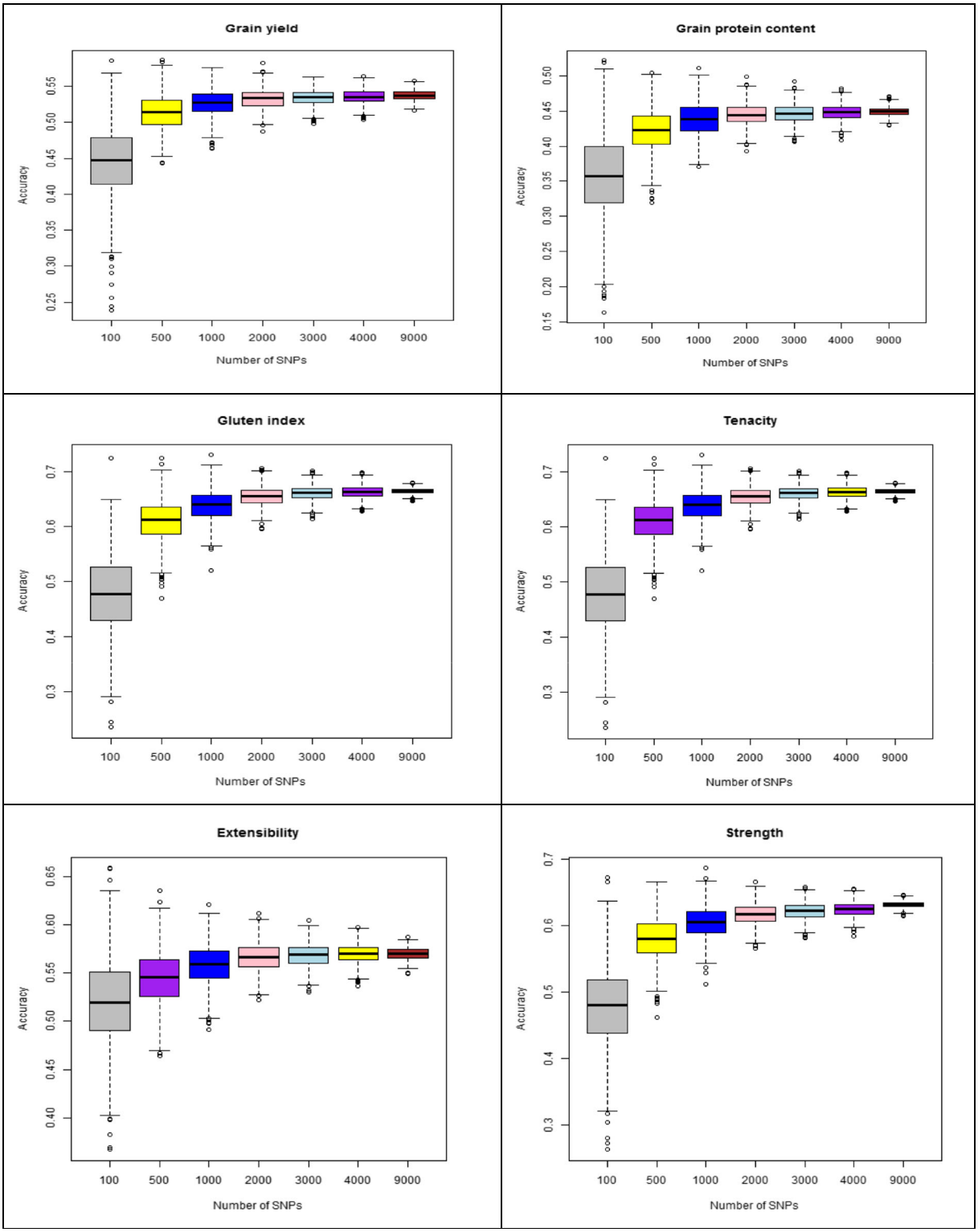



**Fig. 4** Population structure of the breeding panel (BP) and W9262-260D3/Kofa (DH) population as revealed by discriminant analysis of principal components using 3141 SNPs that were common to both populations. The first 56 axes explained 80% of the total variance. Each color represents a sub-population. Where 77 individuals from DH belongs to subpopulation 1; 77 individuals from DH; and 3 from BP to subpopulation 2; 61 and 105 individuals from BP formed subpopulations 3 and 4, respectively

**a**


Grain yield
Accuracy
Number of SNPs
100
500
1000
2000
3000
4000
9000
Grain protein content
Gluten index
Tenacity
Extensibility
Strength

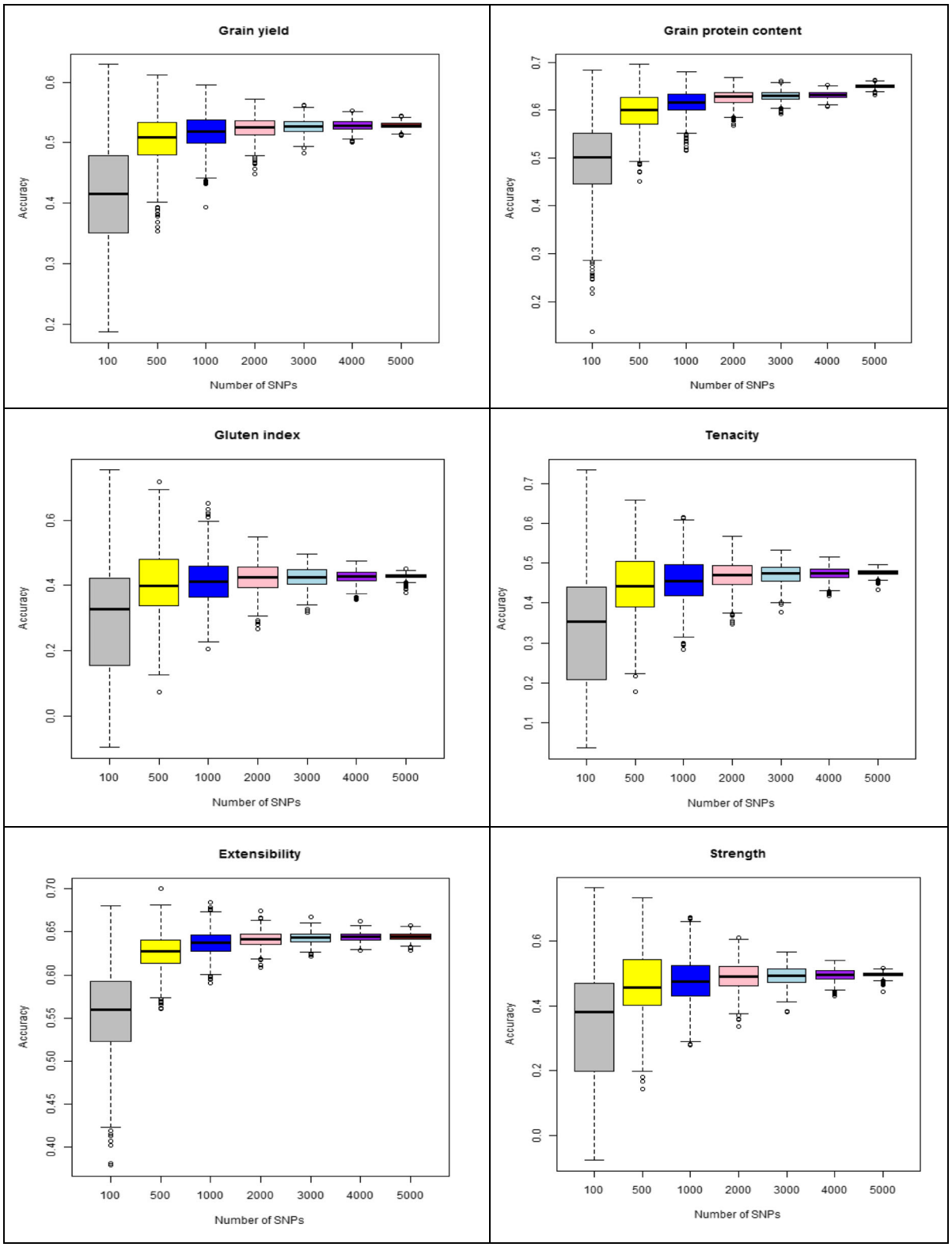



**Fig. 5** (continued)

cycle was shown by Heffner et al. (2010) in a simulation study for wheat and maize. Here, most of the GS models predicted yield values with more accuracy in the bi-parental DH population (24% of advantage), while the prediction of GI resulted better in the panel of cultivars and elites. For GPC, P, W, and L, most of the GS models showed non-significant differences among the two population types. Positive and significant within-population cross-validation accuracies were obtained in both types of populations for all the traits (Fig. 3). As in this study, high accuracy values were reported for predictions within narrow-based bi-parental bread wheat populations (Heffner et al. 2011a).

Most breeding programs would prefer to apply GS methodologies that rely on already available dataset to predict the performances of new progenies (Tayeh et al. 2015). In that sense, using the set of cultivars and elite breeding lines with historical data available (BP) would be cost-effective because expensive agronomic and end-use quality tests are measured extensively in the advanced yield trials (Pozniak et al. 2012). Similarly, the DH is a good example of a newly generated bi-parental population that can be used to simulate what breeders would require selecting among. It is therefore of interest to assess the feasibility of selecting among the DH progenies, using the BP as TP. However, accuracies of prediction were consistently lower compared with within-population prediction. In the case of grain yield, the GEBVs were negatively correlated with the actual phenotypic values (Fig. 3). Similarly, negative accuracy in cross-population validation for grain yield has been reported for prediction among less related bi-parental families in maize (Riedelsheimer et al. 2013) and in sugar beet (Würschum et al. 2013). This suggests that there may be opposite linkage phases between SNPs and important QTL for grain yield in the BP and in the DH population. Additionally, the DAPC results (Fig. 4) support allelic un-relatedness between the two sets, and this can cause the dropping in accuracies as compared to the within panel. In fact, within-population validation methods are generally considered to overestimate the potential of genomic prediction (Hofheinz et al. 2012).

Models trained on one wheat population were not found to predict breeding values in a different population or if so, low accuracies were noted in DH wheat populations (Charmet et al. 2014; Thavamanikumar et al. 2015). Cross-validation between different panels of sugarcane yielded promising correlations ranging from 0.13 to 0.55 (Gouy et al. 2013) whereas close-to-zero prediction accuracies were obtained in eucalyptus (Resende et al. 2012) and maize DH populations (Riedelsheimer et al. 2013). Moderate value of cross-population predictions in alfalfa was observed by Annicchiarico et al. (2015), and lower prediction accuracy for across prediction for two half-sib populations of hybrid rye was observed by Wang et al. (2014). Prediction accuracy values of 0.27 to 0.66 were reported for grain and semolina quality traits by applying forward GS approach in $F_{4:7}$ durum breeding lines (Fiedler et al. 2017). All these results confirm that the degree of relationship between the training and validation populations affects the accuracies of GS more than the statistical model deployed or the number of markers (Akdemir et al. 2015; Asoro et al. 2011; Lorenz et al. 2012; Sallam et al. 2015; Wang et al. 2014; Zhong et al. 2009; Bassi et al. 2016; Clark et al. 2012; Habier et al. 2007; Hayes et al. 2009b; Rincent et al. 2017). Therefore, using "historical" dataset to conduct GS to predict performances unrelated bi-parental population will result in selection strategies that do not result in better progenies. Rather, half-sibs or full-sibs should be used or as Rutkoski et al. (2015) recommended use historical data for initializing a GS-based breeding program where the selection candidates are founded by historical individuals but once GS model updating can occur, it may be best to discard historical data and use the more recent data for model training.

## Effects of marker number on prediction accuracy

Several marker densities were assessed to determine the lowest number that could be used to obtain significant accuracies when deploying GS in durum wheat breeding. It was observed that the use of 100 markers could generate the highest recorded accuracy, as well as the lowest, and the lowest average performances. This indicates that if a strategy could be derived to adequately select the set of markers to be used, maximum accuracies could be obtained with as little as 100 markers, in spite of the complexity of the huge durum wheat genome. As the number of markers was increased, the total level of variation in prediction accuracies decreased. When predicting strength in the breeding panel and protein in the DH population, a 7% improvement in accuracy was observed with 9000 and 5000 SNPs, respectively, in all other cases increasing marker number beyond 2000 did not improve accuracy. For

1000 SNPs, the total level of variation was higher than what was observed by doubling this number, but the average accuracies were not statistically different from each other. A similar value was also reported by Heffner et al. (2011a) for grain quality traits in two bread wheat bi-parental populations.

On the other hand, the benefit of increasing marker densities was supported by Heffner et al. (2011b) in which the highest accuracy was observed at their maximum marker density of 1158 DArT markers for grain quality traits in a multifamily wheat population. Interestingly, this number is about the same number of markers that we found to be at the inflection point towards a plateau in accuracy. Similar results were detected in other empirical studies (Lorenz et al. 2012; Lorenzana and Bernardo 2009). Using simulated data on maize, Hickey et al. (2014) found that increasing markers from 10,000 to 100,000 did not increase accuracy. These results may appear to be in contrast with the main assumption of genomic selection that all QTL are in LD with at least one marker (Goddard and Hayes 2009) and that increasing marker density will improve prediction accuracy by increasing the number of QTL that are in LD with markers and thus capturing more of the haplotype variation (Asoro et al. 2011; Heffner et al. 2011b; Sallam et al. 2015; Spindel et al. 2015; Zhao et al. 2012). For its practical implementation, the use of GS in durum wheat breeding programs depends on the results of a cost–benefit analysis. Based on the result of our study, testing a large number of markers may not be cost-effective, although genotyping costs are decreasing rapidly. Thus, scaling of marker number to the level of variation within the breeding material will be important for breeders to capture the benefits of affordable, high-density genotyping (Heffner et al. 2011b).

## Implications

The obtained accuracies of GEBVs for various traits in durum wheat in this and previous studies (Fiedler et al. 2017) support the feasibility of applying genomic selection as a cost-effective means to enhance genetic gain for complex and expensive-to-measure traits in durum wheat. The prediction of these traits from DNA analysis means that genetic evaluations can occur at an early stage, shortening the generation cycle and increasing selection intensities for traits not normally measured on all breeding materials.

Hence, we recommend that GS in durum wheat breeding could be applied successfully for within-population selection. All the six ST-GS models appear to be applicable for GS of grain yield and gluten strength traits in durum wheat. But due to ease of computation, either RR-BLUP or G-BLUP is recommended. For predicting grain yield and protein content simultaneously, as in conventional breeding, the use of an economic selection index proved advantageous and hence should also be kept in consideration in practical applications of multi-trait GS. Further, the number of markers to be deployed can be effectively scaled down to reduce costs. Notably, the use of historical breeding lines (BP) failed to provide any meaningful prediction when applied to a bi-parental population. Hence, durum breeders should be extremely careful in designing their GS schemes to ensure good genetic similarity between the lines used as TP and the progenies to be selected. Whenever possible, full-sibs or half-sibs should be used for this scope. However, practical implementation of genomic selection is still a major challenge for breeding programs and genomic selection remains largely unexplored in durum wheat (Tuberosa and Pozniak 2014). Lastly, the availability of high-density consensus maps for tetraploid wheat (Marone et al. 2012; Maccaferri et al. 2015) coupled with the less complexity of durum wheat genome compared with bread wheat will facilitate future investigations of GS for other important complex traits.

**Acknowledgements** We gratefully acknowledge funding of this research by the Genome Canada and Genome Prairie as part of the Canadian *Triticum* Applied Genomics (CTAG) project. Additional funding provided by the Saskatchewan Ministry of Agriculture, Western Grains Research Foundation, Agriculture and Agri-Food Canada, and the Natural Sciences and Engineering Research Council of Canada is also gratefully acknowledged. We also acknowledge the technical assistance of Shawn Yates at the Agriculture and Agri-Food Canada Research and Development Centre, Swift Current, for curation of the database from which the breeding panel data were drawn. The data from these trials were conducted under the auspices of the Prairie Recommending Committee for Wheat, Rye, and Triticale, and we acknowledge them for permission to use the data presented for the breeding panel, and the many people who conducted the field trials. We acknowledge the coordinators of these trials, including Dr. D. Leisle, Agriculture and Agri-Food Canada, Winnipeg (retired), from 1961 to 1994, Dr. J. M. Clarke, from 1995 to 2007, Dr. A.K. Singh from 2008 to 2012, and Dr. R.M. DePauw in 2013, all from Agriculture and Agri-Food Canada. We also acknowledge those responsible for end-use quality analysis of these trials from the Grain Research Laboratory of the Canadian Grain Commission. Lastly, we acknowledge Krysta Wiebe, Jennifer Ens, and Justin Coulson for support in generation of the molecular data used in this project.

## References


AACC (2000) Approved methods of American Association of Cereal Chemists, 10th edn. The Association, St. Paul, USA

Akdemir D, Godfrey O (2015) EMMREML: fitting mixed models with known covariance structures. R package version, vol 3, p 01

Akdemir D, Sanchez JI, Jannink J-L (2015) Optimization of genomic selection training populations with a genetic algorithm. Genet Sel Evol 47(1):38. https://doi.org/10.1186/s12711-015-0116-6

Annicchiarico P, Nazzicari N, Li X, Wei Y, Pecetti L, Brummer EC (2015) Accuracy of genomic selection for alfalfa biomass yield in different reference populations. BMC Genomics 16:1020. https://doi.org/10.1186/s12864-015-2212-y

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(2):132–144. https://doi.org/10.3835/plantgenome2011.02.0007

Bao Y, Kurle JE, Anderson G, Young ND (2015) Association mapping and genomic prediction for resistance to sudden death syndrome in early maturing soybean germplasm. Mol Breed 35(6):128. https://doi.org/10.1007/s11032-015-0324-3

Bassi FM, Bentley AR, Charmet G, Ortiz R, Crossa J (2016) Breeding schemes for the implementation of genomic selection in wheat (Triticum spp.). Plant Sci 242(Supplement C):23–36. https://doi.org/10.1016/j.plantsci.2015.08.021

Battenfield SD, Guzman C, Gaynor RC, Singh RP, Pena RJ, Dreisigacker S, Fritz AK, Poland JA (2016) Genomic selection for processing and end-use quality traits in the CIMMYT spring bread wheat breeding program. Plant Genome 9(2). https://doi.org/10.3835/plantgenome2016.01.0005

Blanco A, Bellomo M, Lotti C, Pasqualone A (1998) Mapping of quantitative trait loci for grain quality using recombinant inbred lines of durum wheat. EWAC Newsletter:106–110

Calus MP, Veerkamp RF (2011) Accuracy of multi-trait genomic selection using different methods. Genet Sel Evol 43(1):26. https://doi.org/10.1186/1297-9686-43-26

Charmet G, Storlie E, Oury FX, Laurent V, Beghin D, Chevarin L, Lapierre A, Perretant MR, Rolland B, Heumez E, Duchalais L, Goudemand E, Bordes J, Robert O (2014) Genome-wide prediction of three important traits in bread wheat. Mol Breed 34(4):1843–1852. https://doi.org/10.1007/s11032-014-0143-y

Clark SA, Hickey JM, Daetwyler HD, van der Werf JH (2012) The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. Genet Sel Evol 44:4. https://doi.org/10.1186/1297-9686-44-4

Clarke FR, Clarke JM, Ames NA, Knox RE, Ross RJ (2010a) Gluten index compared with SDS-sedimentation volume for early generation selection for gluten strength in durum wheat. Can J Plant Sci 90(1):1–11. https://doi.org/10.4141/CJPS09035

Clarke JM, Clarke FR, Pozniak CJ (2010b) Forty-six years of genetic improvement in Canadian durum wheat cultivars. Can J Plant Sci 90(6):791–809. https://doi.org/10.4141/cjps10091

Clarke FR, Clarke JM, Pozniak CJ, Knox RE, McCaig TN (2009) Protein concentration inheritance and selection in durum wheat. Can J Plant Sci 89(4):601–612. https://doi.org/10.4141/CJPS08191

Clarke JM, McCaig TN, DePauw RM, Knox RE, Ames NP, Clarke FR, Fernandez MR, Marchylo BA, Dexter JE (2005b) Commander durum wheat. Can J Plant Sci 85(4):901–904. https://doi.org/10.4141/P04-189

Clarke JM, McCaig TN, DePauw RM, Knox RE, Clarke FR, Fernandez MR, Ames NP (2005a) Strongfield durum wheat. Can J Plant Sci 85(3):651–654. https://doi.org/10.4141/P04-119

Clarke JM, McLeod JG, DePauw RM, Marchylo BA, McCaig TN, Knox RE, Fernandez MR, Ames N (2000) AC navigator durum wheat. Can J Plant Sci 80(2):343–345. https://doi.org/10.4141/P99-108

Clarke JM, Marchylo BA, Kovacs MIP, Noll JS, McCaig TN, Howes NK (1998) Breeding durum wheat for pasta quality in Canada. Euphytica 100:163–170. https://doi.org/10.1023/A:1018313603344

Conti V, Roncallo PF, Beaufort V, Cervigni GL, Miranda R, Jensen CA, Echenique VC (2011) Mapping of main and epistatic effect QTLs associated to grain protein and gluten strength using a RIL population of durum wheat. J Appl Genet 52(3):287–298. https://doi.org/10.1007/s13353-011-0045-1

Crossa J, Campos G, Pérez P, Gianola D, Burgueño J, Araus JL, Makumbi D, Singh RP, Dreisigacker S, Yan J, Arief V, Banziger M, Braun H-J (2010) Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics 186(2):713–724. https://doi.org/10.1534/genetics.110.118521

Crossa J, Jarquin D, Franco J, Perez-Rodriguez P, Burgueno J, Saint-Pierre C, Vikram P, Sansaloni C, Petroli C, Akdemir D, Sneller C, Reynolds M, Tattaris M, Payne T, Guzman C, Pena RJ, Wenzl P, Singh S (2016) Genomic prediction of Gene Bank wheat landraces. G3 (Bethesda) 6(7):1819–1834. https://doi.org/10.1534/g3.116.029637

Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, de los Campos G, Burgueño J, González-Camacho JM, Pérez-Elizalde S, Beyene Y, Dreisigacker S, Singh R, Zhang X, Gowda M, Roorkiwal M, Rutkoski J, Varshney RK (2017) Genomic selection in plant breeding: methods, models, and perspectives. Trends Plant Sci 22(11):961–975. https://doi.org/10.1016/j.tplants.2017.08.011

D'Ovidio R, Masci S (2004) The low-molecular-weight glutenin subunits of wheat gluten. Journal of Cereal Science 39(3):321–339. https://doi.org/10.1016/j.jcs.2003.12.002

Daetwyler HD, Bansal UK, Bariana HS, Hayden MJ, Hayes BJ (2014) Genomic prediction for rust resistance in diverse wheat landraces. Theor Appl Genet 127(8):1795–1803. https://doi.org/10.1007/s00122-014-2341-8

Daetwyler HD, Pong-Wong R, Villanueva B, Woolliams JA (2010) The impact of genetic architecture on genome-wide evaluation methods. Genetics 185(3):1021–1031. https://doi.org/10.1534/genetics.110.116855

de los Campos G, Pérez-Rodríguez P (2014) Bayesian generalized linear regression. R package version 1.0.4.

Dexter J, Matsuo R, Kruger J (1990) The spaghetti-making quality of commercial durum wheat samples with variable alpha-amylase activity. Cereal Chem 67(5):405–412

Dexter JE, Matsuo RR (1977) Influence of protein content on some durum wheat quality parameters. Can J Plant Sci 57(3):717–727. https://doi.org/10.4141/cjps77-105

Distelfeld A, Uauy C, Fahima T, Dubcovsky J (2006) Physical map of the wheat high-grain protein content gene Gpc-B1 and development of a high-throughput molecular marker. New Phytol 169(4):753–763. https://doi.org/10.1111/j.1469-8137.2005.01627.x

Elouafi I, Nachit MM, Elsaleh A, Asbati A, Mather DE (2000) QTL-mapping of genomic regions controlling gluten strength in durum (*Triticum turgidum* L. var. durum). Options Méditerranéennes Série A, Séminaires Méditerranéens (No. 40):505–509

Endelman JB (2011) Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome 4. doi: https://doi.org/10.3835/plantgenome2011.08.0024

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Longman Group, Essex

Feillet P, Dexter J (1996) Quality requirements of durum wheat for semolina milling and pasta production. Monograph on pasta and noodle. Technology:95–131

Fiedler JD, Salsman E, Liu Y, Michalak de Jiménez M, Hegstad JB, Chen B, Manthey FA, Chao S, Xu S, Elias EM, Li X (2017) Genome-wide association and prediction of grain and semolina quality traits in durum wheat breeding populations. Plant Genome 10. https://doi.org/10.3835/plantgenome2017.05.0038

Gaines CS, Reid JF, Vander Kant C, Morris CF (2006) Comparison of methods for gluten strength assessment. Cereal Chem J 83(3):284–286. https://doi.org/10.1094/CC-83-0284

Gianola D, Fernando RL, Stella A (2006) Genomic-assisted prediction of genetic value with semiparametric procedures. Genetics 173(3):1761–1776. https://doi.org/10.1534/genetics.105.049510

Goddard ME, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet 10(6):381–391. https://doi.org/10.1038/nrg2575

Gouy M, Rousselle Y, Bastianelli D, Lecomte P, Bonnal L, Roques D, Efile JC, Rocher S, Daugrois J, Toubi L, Nabeneza S, Hervouet C, Telismart H, Denis M, Thong-Chane A, Glaszmann JC, Hoarau JY, Nibouche S, Costet L (2013) Experimental assessment of the accuracy of genomic selection in sugarcane. Theor Appl Genet 126(10):2575–2586. https://doi.org/10.1007/s00122-013-2156-z

Guo G, Zhao F, Wang Y, Zhang Y, Du L, Su G (2014) Comparison of single-trait and multiple-trait genomic prediction models. BMC Genet 15(1):30. https://doi.org/10.1186/1471-2156-15-30

Guzman C, Peña RJ, Singh R, Autrique E, Dreisigacker S, Crossa J, Rutkoski J, Poland J, Battenfield S (2016) Wheat quality improvement at CIMMYT and the use of genomic selection on it. Appl Transl Genomics 11:3–8. https://doi.org/10.1016/j.atg.2016.10.004

Habier D, Fernando R, Dekkers J (2007) The impact of genetic relationship information on genome-assisted breeding values. Genetics 177

Hayashi T, Iwata H (2013) A Bayesian method and its variational approximation for prediction of genomic breeding values in multiple traits. BMC Bioinformatics 14:34–34. https://doi.org/10.1186/1471-2105-14-34

Hayes BJ, Bowman PJ, Chamberlain AC, Verbyla K, Goddard ME (2009b) Accuracy of genomic breeding values in multi-breed dairy cattle populations. Genet Sel Evol 41(1):51. https://doi.org/10.1186/1297-9686-41-51

Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009a) Invited review: genomic selection in dairy cattle: progress and challenges. J Dairy Sci 92:433–443. https://doi.org/10.3168/jds.2008-1646

Hazel LN (1943) The genetic basis for constructing selection indexes. Genetics 28(6):476–490

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

Heffner EL, Jannink J-L, Iwata H, Souza E, Sorrells ME (2011a) Genomic selection accuracy for grain quality traits in biparental wheat populations. Crop Sci 51:2597–2606. https://doi.org/10.2135/cropsci2011.05.0253

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. https://doi.org/10.2135/cropsci2009.11.0662

Heslot N, Jannink J-L, Sorrells ME (2015) Perspectives for genomic selection applications and research in plants. Crop Sci 55:1–12. https://doi.org/10.2135/cropsci2014.03.0249

Heslot N, Yang HP, Sorrells ME, Jannink JL (2012) Genomic selection in plant breeding: a comparison of models. Crop Sci 52. doi:https://doi.org/10.2135/cropsci2011.06.0297

Hickey JM, Dreisigacker S, Crossa J, Hearne S, Babu R, Prasanna BM, Grondona M, Zambelli A, Windhausen VS, Mathews K, Gorjanc G (2014) Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation. Crop Sci 54(4):1476–1488. https://doi.org/10.2135/cropsci2013.03.0195

Hofheinz N, Borchardt D, Weissleder K, Frisch M (2012) Genome-based prediction of test cross performance in two subsequent breeding cycles. Theor Appl Genet 125(8):1639–1645. https://doi.org/10.1007/s00122-012-1940-5

ICC (2001) Standard methods of the International Association for Cereal Science and Technology. Standard No.121 and 158

Jannink JL, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding: from theory to practice. Brief Funct Genomics 9:166–177. https://doi.org/10.1093/bfgp/elq001

Jia Y, Jannink JL (2012) Multiple-trait genomic selection methods increase genetic value prediction accuracy. Genetics 192(4):1513–1522. https://doi.org/10.1534/genetics.112.144246

Jiang J, Zhang Q, Ma L, Li J, Wang Z, Liu JF (2015) Joint prediction of multiple quantitative traits using a Bayesian multivariate antedependence model. Heredity (Edinb) 115(1):29–36. https://doi.org/10.1038/hdy.2015.9

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11(1):94. https://doi.org/10.1186/1471-2156-11-94

Knox R, Clarke J, DePauw R (2000) Dicamba and growth condition effects on doubled haploid production in durum wheat crossed with maize. Plant Breed 119:289–298. https://doi.org/10.1046/j.1439-0523.2000.00498.x

Knox RE, Houshmand S, Clarke FR, Clarke JM, Ames NA (2004) QTL association with measures of gluten strength across

environments in durum wheat. In: The gluten proteins. The Royal Society of Chemistry, pp 148–151. doi:https://doi.org/10.1039/9781847552099-00148

Krchov L-M, Gordillo GA, Bernardo R (2015) Multienvironment validation of the effectiveness of phenotypic and genomewide selection within biparental maize populations. Crop Sci 55:1068–1075. https://doi.org/10.2135/cropsci2014.09.0608

Kumar A, Elias EM, Ghavami F, Xu X, Jain S, Manthey FA, Mergoum M, Alamri MS, Kianian PMA, Kianian SF (2013) A major QTL for gluten strength in durum wheat (*Triticum turgidum* L. var. durum). J Cereal Sci 57(1):21–29. https://doi.org/10.1016/j.jcs.2012.09.006

Lado B, Matus I, Rodríguez A, Inostroza L, Poland J, Belzile F, del Pozo A, Quincke M, Castro M, von Zitzewitz J (2013) Increased genomic prediction accuracy in wheat breeding through spatial adjustment of field trial data. G3: Genes|Genomes|Genetics 3(12):2105–2114. https://doi.org/10.1534/g3.113.007807

Leisle D (1972) Hercules, a new durum wheat. Can J Plant Sci 50: 118–119 1970

Littell R, Milliken G, Stroup W, Wolfinger R (1996) SAS®system for mixed models. .633

Lorenz AJ, Smith KP, Jannink JL (2012) Potential and optimization of genomic selection for Fusarium head blight resistance in six-row barley. Crop Sci 52:1609–1621. https://doi.org/10.2135/cropsci2011.09.0503

Lorenzana RE, Bernardo R (2009) Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor Appl Genet 120(1):151–161. https://doi.org/10.1007/s00122-009-1166-3

Maccaferri M, Ricci A, Salvi S, Milner SG, Noli E, Martelli PL, Casadio R, Akhunov E, Scalabrin S, Vendramin V, Ammar K, Blanco A, Desiderio F, Distelfeld A, Dubcovsky J, Fahima T, Faris J, Korol A, Massi A, Mastrangelo AM, Morgante M, Pozniak C, N'Diaye A, Xu S, Tuberosa R (2015) A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. Plant Biotechnol J 13:648–663. https://doi.org/10.1111/pbi.12288

Mackay I, Ober E, Hickey J (2015) GplusE: beyond genomic selection. Food and Energy Secur 4(1):25–35. https://doi.org/10.1002/fes3.52

Marone D, Laido G, Gadaleta A, Colasuonno P, Ficco DB, Giancaspro A, Giove S, Panio G, Russo MA, De Vita P, Cattivelli L, Papa R, Blanco A, Mastrangelo AM (2012) A high-density consensus map of A and B wheat genomes. Theor App Gen 125:1619–1638. https://doi.org/10.1007/s00122-012-1939-y

McCaig TN, DePauw RM, Williams PC (1993) Assessing seed-coat color in a wheat breeding program with a NIR/VIS instrument. Can J Plant Sci 73(2):535–539. https://doi.org/10.4141/cjps93-073

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157

N'Diaye A, Haile JK, Cory AT, Clarke FR, Clarke JM, Knox RE, Pozniak CJ (2017) Single marker and haplotype-based association analysis of semolina and pasta colour in elite durum wheat breeding lines using a high-density consensus map. PLoS One 12(1):e0170941. https://doi.org/10.1371/journal.pone.0170941

Park T, Casella G (2008) The Bayesian lasso. J Am Stat Assoc 103(482):681–686. https://doi.org/10.1198/016214508000000337

Patil RM, Oak MD, Tamhankar SA, Rao VS (2009) Molecular mapping of QTLs for gluten strength as measured by sedimentation volume and mixograph in durum wheat (*Triticum turgidum* L. ssp durum). J Cereal Sci 49(3):378–386. https://doi.org/10.1016/j.jcs.2009.01.001

Pérez P, de los Campos G (2014) Genome-wide regression and prediction with the BGLR statistical package. Genetics 198(2):483–495. https://doi.org/10.1534/genetics.114.164442

Pérez P, de los Campos G, Crossa J, Gianola D (2010) Genomic-enabled prediction based on molecular markers and pedigree using the Bayesian linear regression package in R. Plant Genome 3(2):106–116. https://doi.org/10.3835/plantgenome2010.04.0005

Piepho HP (2009) Ridge regression and extensions for genomewide selection in maize. Crop Sci 49. doi: https://doi.org/10.2135/cropsci2008.10.0595

Poland J, Endelman J, Dawson J, Rutkoski J, Wu S, Manes Y, Dreisigacker S, Crossa J, Sánchez-Villeda H, Sorrells M, Jannink J-L (2012) Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genome 5:103–113. https://doi.org/10.3835/plantgenome2012.06.0006

Pozniak CJ, Clarke JM, Clarke FR (2012) Potential for detection of marker–trait associations in durum wheat using unbalanced, historical phenotypic datasets. Mol Breed 30(4): 1537–1550. https://doi.org/10.1007/s11032-012-9737-4

R Development Core Team (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Resende MD, Resende MF Jr, Sansaloni CP, Petroli CD, Missiaggia AA, Aguiar AM, Abad JM, Takahashi EK, Rosado AM, Faria DA, Pappas GJ Jr, Kilian A, Grattapaglia D (2012) Genomic selection for growth and wood quality in eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees. New Phytol 194(1):116–128. https://doi.org/10.1111/j.1469-8137.2011.04038.x

Riedelsheimer C, Endelman JB, Stange M, Sorrells ME, Jannink JL, Melchinger AE (2013) Genomic predictability of interconnected biparental maize populations. Genetics 194(2): 493–503. https://doi.org/10.1534/genetics.113.150227

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(1):452. https://doi.org/10.1186/1471-2164-13-452

Rincent R, Charcosset A, Moreau L (2017) Predicting genomic selection efficiency to optimize calibration set and to assess prediction accuracy in highly structured populations. Theor Appl Genet 130(11):2231–2247. https://doi.org/10.1007/s00122-017-2956-7

Rutkoski J, Benson J, Jia Y, Brown-Guedira G, Jannink J-L, Sorrells M (2012) Evaluation of genomic prediction methods for Fusarium head blight resistance in wheat. Plant Genome 5:51–61. https://doi.org/10.3835/plantgenome2012.02.0001

Rutkoski J, Singh RP, Huerta-Espino J, Bhavani S, Poland J, Jannink JL, Sorrells ME (2015) Efficient use of historical data for genomic selection: a case study of stem rust resistance in wheat. Plant Genome 8. https://doi.org/10.3835/plantgenome2014.09.0046

Sallam AH, Endelman JB, Jannink JL, Smith KP (2015) Assessing genomic selection prediction accuracy in a dynamic barley breeding population. Plant Genome 8. https://doi.org/10.3835/plantgenome2014.05.0020

Schmidt M, Kollers S, Maasberg-Prelle A, Großer J, Schinkel B, Tomerius A, Graner A, Korzun V (2016) Prediction of malting quality traits in barley based on genome-wide marker data to assess the potential of genomic selection. Theor Appl Genet 129(2):203–213. https://doi.org/10.1007/s00122-015-2639-1

Schulthess AW, Wang Y, Miedaner T, Wilde P, Reif JC, Zhao Y (2016) Multiple-trait- and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes. Theor Appl Genet 129(2):273–287. https://doi.org/10.1007/s00122-015-2626-6

Shikha M, Kanika A, Rao AR, Mallikarjuna MG, Gupta HS, Nepolean T (2017) Genomic selection for drought tolerance using genome-wide SNPs in maize. Front Plant Sci 8:550. https://doi.org/10.3389/fpls.2017.00550

Smith H (1936) A discriminant function for plant selection. Ann Eugenics 7:240–250

Spindel J, Begum H, Akdemir D, Virk P, Collard B, Redoña E, Atlin G, Jannink J-L, McCouch SR (2015) Genomic selection and association mapping in rice (Oryza sativa): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. PLoS Genet 11(2):e1004982. https://doi.org/10.1371/journal.pgen.1004982

Storlie E, Charmet G (2013) Genomic selection accuracy using historical data generated in a wheat breeding program. Plant Genome 6(1). https://doi.org/10.3835/plantgenome2013.01.0001

Su G, Guldbrandtsen B, Gregersen VR, Lund MS (2010) Preliminary investigation on reliability of genomic estimated breeding values in the Danish Holstein population. J Dairy Sci 93(3):1175–1183. https://doi.org/10.3168/jds.2009-2192

Subira J, Pena R, Álvaro F, Ammar K, Ramdani A, Royo C (2014) Breeding progress in the pasta-making quality of durum wheat cultivars released in Italy and Spain during the 20th century. Crop and Pasture Sci 65(1):16–26

Tayeh N, Aubert G, Pilet-Nayel M-L, Lejeune-Hénaut I, Warkentin TD, Burstin J (2015) Genomic tools in pea breeding programs: status and perspectives. Front Plant Sci 6:1037. https://doi.org/10.3389/fpls.2015.01037

Thavamanikumar S, Dolferus R, Thumma BR (2015) Comparison of genomic selection models to predict flowering time and spike grain number in two hexaploid wheat doubled haploid populations. G3: Genes|Genomes|Genetics 5(10):1991–1998. https://doi.org/10.1534/g3.115.019745

Townley-Smith TF, Patterson LA, Depauw RM, Lendrum CWB, McCrystal GE (1987) Kyle durum wheat. Can J Plant Sci 67(1):225–227. https://doi.org/10.4141/cjps87-026

Tuberosa R, Pozniak C (2014) Durum wheat genomics comes of age. Mol Breed 34(4):1527–1530. https://doi.org/10.1007/s11032-014-0188-y

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

VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, Schenkel FS (2009) Invited review: reliability of genomic predictions for North American Holstein bulls. J Dairy Sci 92(1):16–24. https://doi.org/10.3168/jds.2008-1514

Viana JMS, Sobreira FM, De Resende MDV, Faria VR (2010) Multi-trait BLUP in half-sib selection of annual crops. Plant Breed 129(6):599–604. https://doi.org/10.1111/j.1439-0523.2009.01745.x

Wang Y, Mette MF, Miedaner T, Gottwald M, Wilde P, Reif JC, Zhao Y (2014) The accuracy of prediction of genomic selection in elite hybrid rye populations surpasses the accuracy of marker-assisted selection and is equally augmented by multiple field evaluation locations and test years. BMC Genomics 15:556. https://doi.org/10.1186/1471-2164-15-556

Whittaker JC, Thompson R, Denham MC (2000) Marker-assisted selection using ridge regression. Genet Res 75(2):249–252

Würschum T, Reif JC, Kraft T, Janssen G, Zhao Y (2013) Genomic selection in sugar beet breeding populations. BMC Genet 14(1):85. https://doi.org/10.1186/1471-2156-14-85

Zhao Y, Gowda M, Liu W, Wurschum T, Maurer HP, Longin FH, Ranc N, Reif JC (2012) Accuracy of genomic selection in European maize elite breeding populations. Theor Appl Genet 124(4):769–776. https://doi.org/10.1007/s00122-011-1745-y

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(1):355–364. https://doi.org/10.1534/genetics.108.098277