Genomic Selection of Forage Quality Traits in Winter Wheat

Frank Maulana, Ki-Seung Kim, Joshua D. Anderson, Mark E. Sorrells, Twain J. Butler, Shuyu Liu, P. Stephen Baenziger, Patrick F. Byrne, and Xue-Feng Ma*

ABSTRACT
Phenotyping forage quality traits is time-consuming in forage wheat breeding. In this study, prediction accuracies of three genomic selection (GS) models (ridge regression best linear unbiased prediction [RRBLUP], Gaussian kernel [GAUSS], and Bayesian LASSO [BL, where LASSO stands for least absolute shrinkage and selection operator]) for forage quality traits of winter wheat (Triticum aestivum L.) were compared using two genotype sampling methods. In addition, the impact of training population (TP) size and marker density on prediction accuracy was explored. The study was done using a diversity panel (n = 298) that was genotyped using 90K single nucleotide polymorphisms (SNPs) and phenotyped for forage quality traits including crude protein, acid detergent fiber, neutral detergent fiber, sugars, lignin content, and in vitro true dry matter digestibility. Generally, the three models produced similar prediction accuracies, which ranged from 0.34 to 0.61, for all traits. The sampling method had little effect on accuracy. Crude protein was one of the traits with the highest prediction accuracy, and it required only 1000 markers to attain its highest prediction accuracy value. Increasing TP size and marker density increased accuracies of all traits, and increasing the TP size was more effective than increasing marker density. For this panel, the optimal TP size (nTP) was 150, at which point prediction accuracies of all traits, except for sugars, reached over 90% of the highest value at nTP = 250. However, the sampling method for marker density had no effect on accuracy. The results suggest that GS can be an alternative approach to facilitate selection of forage quality traits during forage wheat breeding.

F. Maulana, J.D. Anderson, T.J. Butler, and X.-F. Ma, Noble Research Institute, Ardmore, OK 73401, USA; K.-S. Kim, LG Chem–FarmHannong, Daejeon 34115, Korea; M.E. Sorrells, School of Integrative Plant Science, Cornell Univ., Ithaca, NY 14853-1902, USA; S. Liu, Texas A&M AgriLife Research, Amarillo, TX 79106, USA; P.S. Baenziger, Dep. of Agronomy and Horticulture, Univ. of Nebraska, Lincoln NE 68583-0915, USA; P. F. Byrne, Dep. of Soil and Crop Sciences, Colorado State Univ., Fort Collins CO, 80523-1170, USA. Frank Maulana and Ki-Seung Kim contributed equally to this work. Received 31 Oct. 2018. Accepted 1 July 2019. *Corresponding author (xma@noble.org). Assigned to Associate Editor Heathcliffe Riday.

Abbreviations: ADF, acid detergent fiber; ADL, acid detergent lignin; BL, Bayesian least absolute shrinkage and selection operator; CP, crude protein; ESM, evenly sampling method; GAUSS, Gaussian kernel; GEBV, genomic estimated breeding value; GS, genomic selection; IVTDM, in vitro true dry matter digestibility; LC, lignin content; LD, linkage disequilibrium; MAF, minor allele frequency; MAS, marker-assisted selection; NDF, neutral detergent fiber; NDFD, neutral detergent fiber digestibility; NIRs, near-infrared reflectance spectroscopy; NJ, neighbor-joining; PEBV, phenotypically estimated breeding value; QTL, quantitative trait locus/loci; RRBLUP, ridge regression best linear unbiased prediction; RSM, random sampling method; SNP, single nucleotide polymorphism; SSM, stratified sampling method; SUG, sugars; TCAP, Triticeae Coordinated Agricultural Project; TP, training population; VP, validation population.

Winter wheat (Triticum aestivum L.) is commonly grown as a dual-purpose crop for forage and grain production in the southern Great Plains of the United States, including Kansas, Oklahoma, and Texas. The crop is grown for forage production during the winter season, when warm-season forage crops are dormant or not able to grow due to cold weather, and is later harvested for grain (Kim et al., 2016; MacKown et al., 2011; MacKown and Northup, 2010). The feed quality of any forage crop is one of the most desirable traits considered for forage

Published in Crop Sci. 59:1–11 (2019).
doi: 10.2135/cropsci2018.10.0655

© 2019 The Author(s). This is an open access article distributed under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
production. Factors such as protein content, digestibility, amount of nutrients, and the length of time to digest determine potential utilization of a crop for forage production. Normally, the nutritive value of forages is determined by analyzing a number of chemical components related to animal productivity. Forage quality traits to be considered when assessing the nutritive value include crude protein (CP), acid detergent fiber (ADF), neutral detergent fiber (NDF), sugars (SUG), lignin content (LC), and in vitro true dry matter digestibility (IVTDMD) after 48 h. Livestock need an adequate amount of proteins plus digestible fibers and many other nutrients in their forage feed. The concentrations of chemical components in forage biomass may have a negative or positive effect on the nutritive value and utilization (Biazzi et al., 2017). Forages with less ADF and LC are desirable for good nutritive digestibility of forage, thus increasing forage intake (Biazzi et al., 2017; Kim et al., 2016). There is a high negative correlation between forage intake of animals and fiber concentration (Biazzi et al., 2017).

Previous studies have reported variability in forage quality traits among forage cultivars of alfalfa (Medicago sativa L.) (Annicchiarico, 2007; Julier et al., 2000). Understanding the composition of these traits in forage cultivars is important for livestock production. Forage breeding programs need to evaluate a large number of genotypes to exploit genetic diversity and improve forage quality. However, quantifying forage quality components is time consuming because forage samples need to be collected and dried for near-infrared reflectance spectroscopy (NIRS) analysis. Given these limitations, establishing an efficient approach to characterize forage quality will facilitate forage wheat breeding.

Phenotypic selection is the conventional method used in most breeding programs for crops including wheat. However, this method is time consuming because it involves screening genetic materials for desirable traits over several generations and multiple environments. There has been a tremendous increase in development of molecular marker associations with traits of interest, mostly through quantitative trait locus (QTL) mapping studies. Marker-assisted selection (MAS) is considered a viable alternative method to phenotypic selection for identifying desirable traits in breeding populations and has been effective for markers associated with large-effect QTL (Castro et al., 2003; Xu and Crouch, 2008). The downside of MAS is that only a limited number of markers with relatively large effects on the trait of interest are selected and used, and thus QTL with small effects are not captured (Bernardo, 2010). As a result, its impact for improving breeding efficiency of complex traits is limited (Bernardo, 2010). Moreover, QTL identified through QTL mapping are often specific to a particular genetic background, further jeopardizing its applicability to quantitatively inherited traits.

With recent advancement in genomic technologies, it is now feasible to scan the entire genome of species, instead of a few selected genomic regions, and capture single nucleotide polymorphisms (SNPs) throughout the genome that potentially contribute to the complex traits. Genomic selection (GS) is an alternative selection approach to MAS. Unlike in MAS, in GS, all markers are used to train a prediction model using a training population (TP), which has been phenotyped and genotyped. Marker effects are estimated using the TP and then used to predict phenotypes, known as genomic estimated breeding values (GEBVs), of breeding lines that have only been genotyped (Hayes and Goddard, 2001). Once GEBVs for unphenotyped materials are estimated, selection of superior breeding candidates can be solely based on GEBVs. More importantly, inclusion of all markers in the prediction model increases the chance of finding markers that are in linkage disequilibrium (LD) with QTL associated with the trait (Goddard and Hayes, 2007). Currently, GS is gradually becoming a more powerful approach than other indirect selection approaches used in the past because of its potential of improving genetic gain, reducing phenotyping costs and accelerating development of new cultivars by reducing the cycles of selection (Heffner et al., 2010).

Several GS models have been developed and used in GS studies including ridge regression best linear unbiased prediction (RRBLUP) (Endelman, 2011), Gaussian kernel (GAUSS) (Endelman, 2011), and Bayesian LASSO (BL, where LASSO stands for least absolute shrinkage and selection operator) (Park and Casella, 2008). The vari-ances associated with the markers are assumed differently in these models (Lorenz et al., 2011). For example, the RR.BLUP model, one of the parametric and infinitesimal models, assumes that all marker effects are random and contracted toward zero. In addition, it assumes that the marker effects come from a normal distribution with a common variance (Hayes and Goddard, 2001). The RR.BLUP model can be modified to incorporate epistasis effects, thus additive and dominance genetic variances of the markers can be included. In contrast, the GAUSS model is a nonparametric, nonlinear model that captures both additive and nonadditive interaction effects between markers (Endelman, 2011; Sallam et al., 2015). The BL model is somewhat similar to RR.BLUP in the sense that it is also a parametric linear regression model, which captures additive effects of the markers. Unlike RR.BLUP, the marker effects of BL are assumed to have different vari-ances. Moreover, the predictors of BL undergo variable selection and continuous shrinkage simultaneously, and thus the markers with large effect are selected and used for prediction. When the number of predictors (p) is greater than the number of observations or individuals (n), BL selects at most n variables, and then the effects of the remaining predictors are set to zero (Tibshirani, 1996).
Generally, the BL model is suitable when the predictor variables are not highly correlated. Therefore, it is important to exclude markers that are highly correlated prior to GS analysis. These models have been evaluated in simulation and empirical GS studies in wheat for disease resistance (Arruda et al., 2015), processing and end-use quality traits (Battenfield et al., 2016), and optimization of training and validation sets (Hoffstetter et al., 2016). Prediction accuracy is affected by a number of factors, including GS models used, number of markers (marker density) included to train the model, TP size, population structure, heritability of the trait, and genetic relationship of individuals included in the TP with those in the validation population (VP) (Arruda et al., 2015; Rutkoski et al., 2013). For example, the prediction accuracy of R.R.BLUP and BL for Fusarium head blight (caused by a fungus, Fusarium graminearum) resistance traits was assessed by varying TP size and marker density of 5500 SNPs (Arruda et al., 2015). The study showed that the GS models, TP size, and marker density had significant effects on prediction accuracy for all traits.

Few studies have been conducted in forage crops to assess the potential of GS approach to predict forage quality traits in forage breeding. Recently, the potential of using GS to predict alfalfa forage quality traits was reported (Biazzì et al., 2017). The alfalfa study reported the predictive abilities (correlation between phenotypically estimated breeding value [PEBV] and GEBV) of five GS models for forage quality traits, including stem NDF digestibility (NDFD) after 24 h, leaf protein content, and leaf acid detergent lignin (ADL) using 154 lines genotyped using genotyping-by-sequencing (GBS), which generated 11,450 polymorphic SNP markers. Predictive abilities were moderate, ranging from 0.24 to 0.4 for stem NDFD and leaf protein content, while only modest for leaf ADL and NDFD, and low to very low for the other traits (Biazzì et al., 2017).

To the best of our knowledge, no information is available on using GS to facilitate forage quality selection during forage wheat breeding. The objectives of this study were (i) to assess the potential of using GS to predict forage quality traits during forage wheat breeding, (ii) to determine the appropriate TP size, and (iii) to determine the appropriate number of markers for GS on important forage quality traits for forage wheat breeding.

**MATERIALS AND METHODS**

**Genetic Materials and Genotyping**

The materials used in this study were the diversity panel, consisting of 298 hard winter wheat lines from the Triticeae Coordinated Agricultural Project (TCAP) (http://www.triticeapecap.org) (Guttieri et al., 2015a, 2015b). The diversity panel was genotyped using the wheat iSelect 90K SNP array (TCAP Genotype Experiment “TCAP90K_HHWAMP”) (Guttieri et al., 2015a), generating 15,574 SNP markers with >5% minor allele frequency (MAF) and <10% missing data. After filtering for MAF and missing data, all monomorphic SNP markers were also discarded from further analysis. To remove redundant SNPs, the file was further filtered based on LD between marker pairs (Supplemental Fig. S1) using the R package SNPRelate (Zheng et al., 2012). Before GS analysis, the missing data were imputed using population mean with the R package rrBLUP (Endelman, 2011).

**Phenotyping**

Forage quality traits of this panel were collected during the 2012–2013 and 2013–2014 grazing seasons (Kim et al., 2016). Forage quality analysis was performed using the FOSS 6500 NIRS to estimate forage quality components including CP, ADF, NDF, SUG, LC, and IVTDMD after 48 h. The forage quality components were estimated using the grass equation model developed by the NIRS Forage and Feed Testing Consortium (Hillsboro, WI). Phenotypic data analysis and estimation of variance components were performed using SAS 9.3 (SAS Institute, 2011). Broad-sense heritability ($H^2$) for each trait was estimated on the line mean basis (Kim et al., 2016). To obtain the expected accuracies of phenotypic selection, we also calculated the narrow-sense heritability ($h^2$) estimates using the following formula: $h^2 = \sigma_L^2 / \sigma_{y^2}$, where $h^2$ is the narrow-sense heritability estimate, $\sigma_L^2$ is the additive genetic variance, and $\sigma_{y^2}$ is the total phenotypic variance of the trait.

**Genetic Diversity, Genomic Selection Models, and Genomic Estimated Breeding Values**

The genetic diversity of the panel was assessed using neighbor-joining (NJ) tree analysis performed in TASSEL 5.2.28 (Bradburry et al., 2007). In the present study, we compared prediction accuracies of forage wheat quality traits using the three models: R.R.BLUP, GAUSS, and BL (de los Campos et al., 2009; Endelman, 2011; Park and Casella, 2008). Prior to GS analysis, the SNP marker genotypes were numerically recoded as −1, 0, and 1 using the R package rrBLUP (Endelman, 2011). For the R.R.BLUP and GAUSS models, the R package rrBLUP was used to calculate the GEBVs of the lines. The R.R.BLUP analysis was conducted following the general equation as recommended by Hayes and Goddard (2001): $y = \mu + X\beta + e$, where $y$ is the vector of phenotypic means obtained from phenotypic data analysis, $\mu$ is the overall mean, $X$ is the marker matrix design of the TP, $\beta$ is the matrix of marker effects, and $e$ is the vector of residual effects. The GEBVs of the lines in the VP was calculated by adding the grand mean to the product of genotypic matrix and the vector of mean effect of each SNP marker. The GAUSS model analysis was performed following the equation as suggested by Endelman (2011): $K_j = \exp[-(D_j/\theta)^2]$, where $D_j = \sum_{i=1}^{M}(C_{ik} - G_m)^2$, is the Euclidean distance between individuals $i$ and $j$, normalized to the interval (0, 1), and $\theta$ represents a scale parameter that indicates how quickly the genetic covariance decays with distance, $K_j$ is the linear product between genomic matrices of $i$th and $j$th individuals, $k$ is the number of SNP loci from $k = 1, \ldots, M$, $G_m$ is the SNP genotype code for $m$th individual at SNP locus $k$, and $C_{ik}$ is the SNP genotype code for $j$th individual at SNP locus $k$. The BL model was performed with the R package BLR (Pérez and de los Campos, 2014) using double exponential marginal prior of marker effects, and the parameters recommended by Pérez.
et al. (2010). For BL analysis, the following general model was used: \( Y_{GV} = 1_{n} \beta_0 + X_0 \beta + \varepsilon \), where \( Y_{GV} \) is the genetic value, \( 1_n \) is a vector of ones of \( n \) marker effects, \( \beta_0 \) is the mean, and \( X \) and \( \beta \) are the marker design matrix and a vector of \( n \) marker effects, respectively, while \( \varepsilon \) is a vector of \( n \times 1 \) residuals (Park and Casella, 2008). A total of 50,000 iterations were used during the BL analysis, and the first 20,000 were excluded as burn-in.

Genomic Prediction Accuracies of the Three Genomic Selection Models

Genomic prediction accuracies of the three GS models for forage quality traits of the panel were estimated using mean values of raw phenotypic data in this study. The GS analyses were conducted using 238 wheat lines as a TP and 60 lines as a VP, and prediction accuracies of the models were compared. Random and stratified sampling methods were used to assign lines to the TP and VP. For the random sampling method (RSM), wheat lines were randomly selected from the entire panel without replacement and assigned to the TP and VP without considering population structure existing in the panel. For the stratified sampling method (SSM), the lines were first clustered into different groups using NJ tree analysis. Based on the NJ tree results, lines were randomly and proportionately selected from each group to assign to the TP and VP, and thus population structure was controlled. To account for sampling error, the GS analyses were performed using 1000 iterations.

Predictive ability, \( r(\hat{y}_g, y_g) \) or \( r(GEBV:PEBV) \), is the Pearson correlation between GEBV and PEBV of the trait. Prediction accuracy of each model for the trait was calculated following the formula by Dekkers (2007): \( r(GEBV:PEBV)/\sqrt{H^2} \), in which the \( r(GEBV:PEBV) \) correlation is adjusted by the square root of the \( H^2 \) of the trait.

Genomic Prediction Accuracies as Affected by Training Population Size and Marker Density

Prediction accuracy is affected by several factors, including TP size, number of markers (marker density) included to train the GS model, heritability of the trait, and genetic relationship of genotypes included in the TP with those in the VP (Asoro et al., 2011; Jannink et al., 2010; Lorenz et al., 2012). In this study, the effect of TP size and marker density on prediction accuracy was explored. The effect of TP size on prediction accuracy was determined by varying TP sizes (\( n_{TP} = 25, 50, 100, 150, 200, \) and 250) when all SNP markers were used. The effect of marker density on accuracy was investigated using seven subsets of SNP markers (\( p = 250, 500, 1000, 1500, 3000, 6000, \) and 8500 SNPs) selected from the original marker dataset with the same TP and VP, consisted of 238 and 60 lines, respectively. Random (RSM) and evenly sampling (ESM) methods were used to select markers. With RSM, SNP markers were randomly selected without considering their chromosome locations to form different marker subsets. For ESM, the SNP markers were proportionately and evenly sampled according to their chromosome distribution. In each method, sampling of the markers was repeated 1000 times to account for sampling error. Prediction accuracy of each trait was calculated by averaging the accuracies of all 1000 samples for each marker subset.

RESULTS

Wheat Performance, Variance Components, and Heritability

All traits measured were checked for normality before further analysis. Frequency distributions of the forage quality traits for across environments are presented in Fig. 1. Broad-sense and narrow-sense heritability estimates varied greatly among all traits studied (Table 1).

Genetic Diversity Analysis

Genetic diversity analysis was conducted for the purpose of stratified sampling of TP and VP. The NJ tree analysis loosely clustered this panel into four major groups (G1–G4) (Supplemental Fig. S2). With NJ tree analysis, the first major group (G1) consisted of 79 lines, whereas G2, G3 and G4 comprised 43, 69, and 107 lines, respectively. Generally, the clustering of the lines was based on geographic origin. For example, G1 mainly contained lines from Texas, whereas in G2, G3, and G4, the most predominant lines were from Oklahoma, Colorado, and Nebraska, respectively.

Genomic Prediction Accuracy as Affected by Genomic Selection Model

Predictive abilities and prediction accuracies of the three GS models for forage quality traits of wheat using the two sampling methods are presented in Table 1. We will mainly use prediction accuracies for various comparisons throughout the study. For all GS models, prediction accuracies were moderate to high, depending on the trait. With RSM, the prediction accuracy \( r(\hat{y}_g, y_g) \) for the RRBLUP model ranged from 0.38 for SUG to 0.60 for LC, whereas under SSM, the range was from 0.36 for SUG to 0.61 for LC. Moreover, under both sampling methods, LC consistently had the highest prediction accuracy, followed by ADF and NDF. Among all traits, SUG had the lowest prediction accuracy values. For the GAUSS model, the highest prediction accuracy was observed for LC (0.60) and lowest for SUG, also under both sampling methods. A similar pattern was observed for the BL model. For instance, the highest prediction accuracy was observed for LC, followed by CP and ADF, using both sampling methods.

Generally, the three GS models performed similarly with regard to prediction accuracies for all traits studied. Furthermore, the sampling method of assigning lines to TP for cross-validation had little effect on prediction accuracies for all traits, except for SSM, which produced higher accuracy than RSM on IVTDM after 48 h with the BL model.

Genomic Prediction Accuracy as Affected by Training Population Size

Prediction accuracies of wheat forage quality traits as affected by TP size were estimated using the RRBLUP model.
model by varying TP size using SSM according to genetic diversity of the lines (Fig. 2). In this study, prediction accuracy increased with an increase in TP size for all the traits. For example, prediction accuracy of CP increased from 0.34 to 0.59 when TP size was increased from 25 to 250, on average, representing a 74% increase in accuracy. Similarly, when TP was increased from 25 to 250 lines, the prediction accuracies of ADF and NDF increased by almost 61 and 67%, respectively. For SUG, the prediction accuracy increased from 0.15 to 0.37, representing a 147% increase in accuracy. The prediction accuracy of LC and IVTDMD after 48 h increased by 49 and 100%, respectively, when the TP size was increased from 25 to 250. Generally, prediction accuracy showed an almost linear increase up to the TP size of 100 for all the traits, then tended to flatten with less improvement from 150 to 200 for most of the traits investigated. The two traits, SUG and IVTDMD, with the lowest prediction accuracies showed a relatively steady increase until the TP size of 250, indicating the importance of increasing TP size for traits with low prediction accuracy in general. For this panel, the optimal TP size is 150, at which point prediction accuracies reached over 90% of the highest value at the TP size of 250 for all the traits, except for SUG.
Table 1. Predictive abilities and prediction accuracies of three genomic selection models for forage wheat quality traits using two genotype sampling methods. Values ± SE are shown.

<table>
<thead>
<tr>
<th>Traits‡</th>
<th>h²</th>
<th>H²</th>
<th>RSM</th>
<th>SSM</th>
<th>RSM</th>
<th>SSM</th>
<th>RSM</th>
<th>SSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>0.07</td>
<td>0.25</td>
<td>0.30 ± 0.004</td>
<td>0.59 ± 0.005</td>
<td>0.29 ± 0.005</td>
<td>0.58 ± 0.006</td>
<td>0.28 ± 0.003</td>
<td>0.56 ± 0.004</td>
</tr>
<tr>
<td>ADF</td>
<td>0.22</td>
<td>0.57</td>
<td>0.40 ± 0.003</td>
<td>0.53 ± 0.004</td>
<td>0.40 ± 0.004</td>
<td>0.53 ± 0.005</td>
<td>0.39 ± 0.003</td>
<td>0.52 ± 0.004</td>
</tr>
<tr>
<td>NDF</td>
<td>0.17</td>
<td>0.55</td>
<td>0.37 ± 0.003</td>
<td>0.50 ± 0.005</td>
<td>0.37 ± 0.004</td>
<td>0.50 ± 0.005</td>
<td>0.38 ± 0.005</td>
<td>0.51 ± 0.003</td>
</tr>
<tr>
<td>SUG</td>
<td>0.07</td>
<td>0.44</td>
<td>0.25 ± 0.005</td>
<td>0.38 ± 0.004</td>
<td>0.24 ± 0.006</td>
<td>0.36 ± 0.005</td>
<td>0.24 ± 0.003</td>
<td>0.36 ± 0.005</td>
</tr>
<tr>
<td>LC</td>
<td>0.20</td>
<td>0.55</td>
<td>0.44 ± 0.005</td>
<td>0.60 ± 0.004</td>
<td>0.45 ± 0.004</td>
<td>0.61 ± 0.004</td>
<td>0.44 ± 0.003</td>
<td>0.60 ± 0.004</td>
</tr>
<tr>
<td>IVTDMD</td>
<td>0.02</td>
<td>0.56</td>
<td>0.34 ± 0.003</td>
<td>0.45 ± 0.004</td>
<td>0.34 ± 0.003</td>
<td>0.45 ± 0.004</td>
<td>0.28 ± 0.003</td>
<td>0.38 ± 0.004</td>
</tr>
</tbody>
</table>

† Models: RRBLUP, ridge regression best linear unbiased prediction; GAUSS, Gaussian kernel; BL, Bayesian LASSO (where LASSO stands for least absolute shrinkage and selection operator). Sampling methods: RSM, random sampling method, population structure not controlled; SSM, stratified sampling method, population structure controlled.

‡ CP, crude protein; ADF, acid detergent fiber; NDF, neutral detergent fiber; SUG, sugars; LC, lignin content; IVTDMD, in vitro true dry matter digestibility after 48 h.
using 279 maize lines genotyped with 2953 SNPs to predict flowering time, ear height, and diameter using the RRBLUP and GAUSS models. The authors found no significant difference in accuracies of the RRBLUP and GAUSS models for all traits studied, similar to what we found in the current study. The current results also agree with other GS studies done in forage quality traits of alfalfa (Biazzi et al., 2017), agronomic traits of wheat (Heffner et al., 2011b; Heslot et al., 2012), processing and end-use quality traits of wheat (Battenfield et al., 2016), fruit quality traits of apple (Malus × domestica Borkh.) (Kumar et al., 2012), and agronomic traits of barley (Hordeum vulgare L.) (Sallam et al., 2015).

The prediction accuracies of the GS statistical models used in this study ranged from 0.34 to 0.61, which correspond to predictive abilities of 0.23 to 0.45. The results generally agree with those reported in alfalfa (predictive abilities in fact) (Biazzi et al., 2017). In contrast, the expected accuracies of selecting these quality traits using phenotypic selection as estimated using narrow-sense heritability estimates were very low in this study, ranging from 0.07 to 0.22 (Table 1), suggesting that GS may potentially enhance the selection of the traits. In the present study, we also assessed the effect of two sampling methods used to assign individuals to TP and VP on

---

**Fig. 2. Genomic prediction accuracies of forage quality traits as affected by training population (TP) size assessed using ridge regression best linear unbiased prediction (RRBLUP). CP, crude protein (g kg⁻¹); ADF, acid detergent fiber (g kg⁻¹); NDF, neutral detergent fiber (g kg⁻¹); SUG, sugars (g kg⁻¹); LC, lignin content (g kg⁻¹); IVTDMD, in vitro true dry matter digestibility after 48 h (g kg⁻¹).**
prediction accuracy. However, no significant differences in accuracy for all traits were observed between RSM and SSM. This result suggests that sampling method did not influence prediction accuracies of the models for all forage quality traits studied in this panel. It also indicates that both sampling methods selected representative TP and VP because the number of iterations, 1000, used in the study was large enough to minimize the sampling errors.

Determining and optimizing TP size is one of the areas of focus in GS studies as the overall goal of GS is to reduce phenotyping costs. Therefore, it is important to define the minimum representative number of lines to be included in the TP that can give higher prediction accuracy to reduce expenses associated with phenotyping. A previous empirical GS study showed that depending on the trait and model used, prediction accuracy increases...
with an increase in TP size until it reaches an optimal TP size beyond which no significant increase in accuracy is observed (Cericola et al., 2017). Defining the optimal TP size can help to balance prediction accuracy with costs associated with genotyping and phenotyping of TP. Because the three GS models evaluated in the present study had generally similar prediction accuracies, the assessment of the impact of TP size, and marker density on accuracy was investigated using only the RRBLUP model. Compared with the other GS models tested, the RRBLUP requires less computational time and is relatively stable because the RRBLUP model gives high prediction accuracies across traits irrespective of their genetic architecture. In this study, prediction accuracies of all traits significantly increased with an increase in TP size. Increasing TP size enhances estimates of marker effects, thereby improving prediction accuracies (Arruda et al., 2015). In this study, accuracies for most traits increased substantially up to TP size of 100, then the increases started to slow down. Similar results have been reported in previous GS studies conducted in different crop species (Zhong et al., 2009; Norman et al., 2018), where prediction accuracy increased with increasing TP size up to a certain point before starting to slow down steadily. Our findings corroborate results reported in other GS studies done in wheat (Norman et al., 2018), maize (Crossa et al., 2014), and oat (Avena sativa L.) (Asoro et al., 2011). However, we observed that the extent of increase was markedly different among the traits studied. Besides TP size, factors that have been reported to affect prediction accuracy include population type (biparental vs. multifamily), trait architecture, family structure, and LD between the markers and QTL associated with the trait of interest (Li et al., 2015; Rutkoski et al., 2015).

Furthermore, prediction accuracy in GS studies is affected by the number of markers used to estimate marker effects in TP during GS model training (Arruda et al., 2015; Tayeh et al., 2015). As a rule of thumb, in GS studies, the marker coverage should be sufficient to explain the genotypic variability of the traits for high prediction accuracy (Goddard and Hayes, 2007) and the markers should be in strong LD with QTL controlling traits of interest. Although marker density that can produce high prediction accuracy in GS studies depends on the population type, it is important to determine the optimal number of markers to use for model training to keep genotyping costs at a minimum. For instance, fewer markers are required with biparental populations compared with multifamily populations, which require high marker density to improve accuracy. In the present study, the effect of marker density in combination with marker sampling strategies on prediction accuracies were compared. In general, prediction accuracies increased with an increase in marker density in all traits regardless of the marker sampling method used. However, the sampling method had minimal or no effect on prediction accuracy. Generally, a minimal increase was observed after including at least 1000 or 1500 SNPs, depending on the trait. For example, the increase of prediction accuracy of CP with both RSM and ESM was minimal with marker density ranging from \( p = 1000 \) (0.57) to \( p = 8500 \) (0.58) SNPs, representing only 2.0% accuracy increase after adding 7.5-fold markers. Overall results from this study suggest that for most traits, \( \sim 1000 \) SNPs were able to capture most of the genetic variance explaining the traits. Therefore, marker density used in this study can be reduced to 1000 well-distributed markers without significant loss of prediction accuracies for most of the forage quality traits of wheat for an economic implementation of GS. Reducing the marker density in GS studies cannot only lower genotyping costs but also reduce the issue of collinearity among markers. These results corroborate previous findings in GS studies done in wheat (Arruda et al., 2015; Heffner et al., 2011a), oat (Asoro et al., 2011), maize (Lorenzana and Bernardo, 2009), and rice (Oryza sativa L.) (Spindel et al., 2015). Wheat is a self-pollinated crop with high LD due to inbreeding; hence, lower marker density might be enough to capture the larger portion of the genetic variance of the trait. In general, breeders should carefully design the GS taking into consideration the GS model, the optimal TP size, and the number of markers to use during cross-validation for improved accuracy while reducing cost.

In summary, the present study has shown that GS can be used to predict forage quality traits of wheat during forage wheat breeding with moderate to high accuracy, depending on the target trait. Although the three models tested produced similar prediction accuracies for most traits, the RRBLUP was generally the best model with regard to stability and computational time. High prediction accuracies were achieved with 150 lines as TP size for most traits and marker density of at least 1000 or 1500 SNPs, depending on the trait. However, sampling methods (RSM and SSM) for selecting lines to constitute the TP and VP did not have significant impact on accuracy for any of the traits. Similarly, RSM and ESM of selecting markers did not result in significant difference in prediction accuracy for the traits studied. Overall, increasing TP size was more effective than increasing marker density. Given field limitation in measuring forage quality traits, GS provides an alternative approach to facilitate selection of forage quality traits during forage wheat breeding.

**Supplemental Material Available**
Supplemental material for this article is available online.

**Conflict of Interest**
The authors declare that there is no conflict of interest.
Acknowledgments
The authors sincerely thank Susie Fagan for critical reading of the manuscript and the Triticeae Coordinated Agricultural Project (TCAP), funded by USDA Agriculture and Food Research Initiative Competitive Grant 2011-68002-30029, for making genotypic data used in this study publicly available. This project was supported by the Noble Research Institute.

References


