Received: 5 April 2023 | Accepted: 14 August 2023

DOI: 10.1002/tpg2.20385

The Plant Genome

ORIGINAL ARTICLE

Salinity stress tolerance prediction for biomass-related traits in maize (*Zea mays* L.) using genome-wide markers

Vishal Singh[1,4] | **Margaret Krause**[1] | **Devinder Sandhu**[2] | **Rajandeep S. Sekhon**[3] | **Amita Kaundal**[1]

[1]Plants, Soils, and Climate, College of Agricultural and Applied Sciences, Utah State University, Logan, Utah, USA

[2]US Salinity Laboratory (USDA-ARS), Riverside, California, USA

[3]Department of Genetics and Biochemistry, Clemson University, Clemson, South Carolina, USA

[4]ICAR-Indian Institute of Maize Research, Ludhiana, Punjab, India

Correspondence
Amita Kaundal, Plants, Soils, and Climate, College of Agricultural and Applied Sciences, Utah State University, Logan, UT, 84322, USA. Email: amita.kaundal@usu.edu

Assigned to Associate Editor Alexander Lipka.

Funding information
Utah Agriculture Experiment Station (New faculty Start-up); Indian Council of Agricultural Research

Abstract

Maize (*Zea mays* L.) is the third most important cereal crop after rice (*Oryza sativa*) and wheat (*Triticum aestivum*). Salinity stress significantly affects vegetative biomass and grain yield and, therefore, reduces the food and silage productivity of maize. Selecting salt-tolerant genotypes is a cumbersome and time-consuming process that requires meticulous phenotyping. To predict salt tolerance in maize, we estimated breeding values for four biomass-related traits, including shoot length, shoot weight, root length, and root weight under salt-stressed and controlled conditions. A five-fold cross-validation method was used to select the best model among genomic best linear unbiased prediction (GBLUP), ridge-regression BLUP (rrBLUP), extended GBLUP, Bayesian Lasso, Bayesian ridge regression, BayesA, BayesB, and BayesC. Examination of the effect of different marker densities on prediction accuracy revealed that a set of low-density single nucleotide polymorphisms obtained through filtering based on a combination of analysis of variance and linkage disequilibrium provided the best prediction accuracy for all the traits. The average prediction accuracy in cross-validations ranged from 0.46 to 0.77 across the four derived traits. The GBLUP, rrBLUP, and all Bayesian models except BayesB demonstrated comparable levels of prediction accuracy that were superior to the other modeling approaches. These findings provide a roadmap for the deployment and optimization of genomic selection in breeding for salt tolerance in maize.

1 | INTRODUCTION

Climate change projections suggest that environmental variables such as temperature, rainfall, humidity, and solar radiation are expected to change beyond their normal historical limits (U.S. Global Change Research Program, 2017). Fluctuations in these abiotic factors will cause enhanced stress to crop plants and significantly threaten the global food supply. Therefore, there is an urgency to find ways to minimize the impact of climate change on crop production.

Abbreviations: ANOVA, analysis of variance; BL, Bayesian lasso; BLUP, best linear unbiased prediction; BRR, Bayesian ridge regression; EGBLUP, extended GBLUP; GBLUP, genomic best linear unbiased prediction; GS, genomic selection; GWAS, genome-wide association study; LD, linkage disequilibrium; RL, root length; RL_STI, root length salt tolerance index; rrBLUP, ridge-regression BLUP; RW, root weight; RW_STI, root weight salt tolerance index; SL, shoot length; SL_STI, shoot length salt tolerance index; SNP, single nucleotide polymorphism; STI, salt tolerance index; SW, shoot weight; SW_STI, shoot weight salt tolerance index; TRP, training population; TSP, testing population.

Salinity is a major environmental stressor that affects a significant portion of farmlands worldwide. According to estimates, 424 million hectares of topsoil (0–30 cm) and 833 million hectares of subsoil (30–100 cm) in 118 countries around the world are affected by salinity (FAO, 2021). Soil salinity is a dynamic problem affected by many climatic, edaphic, and anthropogenic factors, including temperature, evaporation, evapotranspiration, soil leaching, seawater intrusion, and crop cultivation practices (Corwin, 2021). Efficient irrigation systems and reduced water usage for crop cultivation in salinity-prone regions may inadvertently lead to an escalation of salt concentration in the soil profile (Aragüés et al., 2014). Salinity adversely affects plant growth and development, and most land plants grown in saline soils fail to achieve optimum growth. High sodium ion buildup in the plant tissues damages biological membranes and subcellular organelles (Davenport et al., 2005; Quintero et al., 2007) and interferes with the uptake and availability of other essential nutrient elements, including calcium and potassium (El Bassiouny & Bekheta et al., 2004; M. Hussain et al., 2013; Munns et al., 2006; Sandhu & Kaundal, 2018). The toxic effects of sodium can impair the photosynthetic machinery, resulting in reduced yield and production.

Maize (*Zea mays* L.) is one of the leading cereals, with 1.2 billion metric tonnes of global production during 2021 (FAOStat, 2021). Maize is a multipurpose crop, and 12.8%, 56.3%, 0.7%, 4.7%, and 19.6% of global maize produce is utilized for food, feed, seed, processing, and non-food uses, respectively (Erenstein, 2022). Maize is moderately sensitive to salinity stress (Chinnusamy et al., 2005; Farooq et al., 2015; Maas et al., 1983; Sandhu et al., 2020), which significantly reduces plant height, leaf area, root elongation, root anatomy, photosynthesis, stomatal conductance, and root–shoot biomass (Ali Turan et al., 2009; Dikobe et al., 2021; El Sayed & El Sayed, 2011; D. Hu et al., 2022; Zahra et al., 2020). Decreased photosynthetic activity may adversely affect the grain number per plant (Hiyane et al., 2010; Hütsch et al., 2014; Jung et al., 2017) and grain weight (Barutcular et al., 2005). Although both root and shoot growth are adversely affected by salinity in maize, shoot-related traits experience particularly severe penalties compared to root traits (Farooq et al., 2015; Munns & Sharp, 1993; Shahzad et al., 2012). The yield of maize grain and silage is estimated to decrease by 50% when the soil electrical conductivity reaches 7 dS m^{-1} and 8.6 dS m^{-1}, respectively (Amacher et al., 2000). While global maize production has followed an upward trend (FAO Stat., 2021), rising demands for feed, silage, and ethanol production may potentially expand maize cultivation onto marginal/degraded lands, including those with high salinity. Amid rapidly shrinking agricultural land (Brain et al., 2023), expanding cultivation to areas with degraded soils may be helpful for sustainable agriculture and food security.

The development of salt-tolerant cultivars offers an opportunity to sustain and increase maize yield in saline soils. Several studies have focused on understanding the biology of salt tolerance in plants (Acharya et al., 2022; Hasan et al., 2021; Hussain et al., 2021; Kaundal et al., 2022), and different omics-based approaches have been proposed to improve salt tolerance in food crops (Kumar et al., 2021; Sandhu et al., 2020). Genomic selection (GS) has been successfully used to sustain increased genetic gain in different plant species (Lorenzana & Bernardo, 2009; Singh & Kaundal, 2023). Predicting the breeding value of untested genotypes can reduce the need to test numerous genotypes in the field, ultimately cutting phenotyping costs (Krchov & Bernardo, 2015). Phenotyping for salinity stress tolerance in the field is tedious and time-consuming. Special plots are required for such phenotyping where the salinity of a desired level is attained by applying a saline solution. Maintaining a constant level of salinity in such plots for longer periods is very challenging due to several factors affecting soil salinity, including rainfall and evapotranspiration. Therefore, phenotyping of a large number of field-grown genotypes in a breeding program is a costly, time-consuming, and laborious endeavor. Given the complexity and labor-intensive nature of phenotyping for salinity stress tolerance, GS is an ideal approach for genetically improving this trait. The variability of soil salinity makes large-scale field phenotyping logistically difficult, thereby increasing the practicality of phenotyping a carefully curated selection of lines through GS.

Identifying an appropriate model is essential in proceeding with the GS approach for trait improvement. In GS, genomic estimated breeding values (GEBVs) are obtained for the "genotyped only" or testing population (TSP) using a model trained on genotypic and phenotypic data of the training population (TRP). GS studies suffer from the complexity of many markers (P) and small population size (*n*) scenarios. A situation of P > > *n* with correlated predictors makes

Core Ideas

- Genomic selection was evaluated for salt tolerance in maize to have a general idea of expected prediction accuracy.
- Genomic best linear unbiased prediction (GBLUP), ridge-regression BLUP, Bayesian ridge regression, BayesA, and BayesB showed similar performance for all the studied traits when only trait associaterd markers were used.
- Bayesian models showed slightly higher prediction accuracies when higher density of randondomaly selected markers were used.
- A reduced marker set selected based on analysis of variance + linkage disequilibrium showed improved prediction accuracy in cross-validation.

ordinary least squares unfit to estimate marker effects. In such cases, alternative methods have been proposed to address this challenge effectively (Lande & Thompson, 1990; Whittaker et al., 1995). Statistical models employed for tackling the intricacies of GS can broadly be grouped into shrinkage models, for example, ridge-regression best linear unbiased prediction (rrBLUP; Whittaker et al., 2000), Lasso (Usai et al., 2009), Elastic Net (Ogutu et al., 2012); variable selection models, for example, BayesA, BayesB, and BayesC (Habier et al., 2011; Meuwissen et al., 2001); dimension reduction models, for example, principle component analysis (PCA) and partial least square (Macciotta et al., 2010; Solberg et al., 2009); and kernel models, for example, reproducing kernel Hilbert spaces (RKHS) regression and support vector machine (Gianola & Van Kaam, 2008; Ogutu et al., 2012). Genomic BLUP (GBLUP) is one of the most widely implemented GS models employing a marker-based relationship matrix to predict GEBVs (VanRaden, 2008). In contrast to GBLUP, which relies on kinship-based predictions, rrBLUP follows a shrinkage-based methodology to estimate marker effects (Meuwissen et al., 2001; Whittaker et al., 2000). This approach allows greater numerical stability in linear regression, particularly under $P >> n$ scenarios (Hoerl & Kennard, 2000). Due to high computational cost, most GS methods use additive and dominant effects but ignore epistasis, an essential component in explaining complex trait architecture. Extended GBLUP (EGBLUP) reduces computational load by adding a marker-based epistatic relationship matrix to GBLUP (Jiang & Reif, 2015). Bayesian methods are also widely used for variable selection and shrinkage of estimates. The prior densities of marker effects determine the inclusion of variable selection and the extent and type of shrinkage in the Bayesian models for GS (de los Campos et al., 2013). The Bayesian ridge regression (BRR) approach is similar to rrBLUP but uses a Bayesian resolution and a Gaussian prior with mean and variance as hyperparameters. Due to homogenous shrinkage performed by rrBLUP and BRR, these models are not suitable for scenarios where few markers are unlinked to a quantitative trait locus (QTL). Thick-tailed priors are used to resolve this issue where the scaled-t density is used in BayesA (Meuwissen et al., 2001), and a double-exponential prior is used in Bayesian Lasso (BL; Park & Casella, 2008). BayesB (Meuwissen et al., 2001) uses scaled-t density, and BayesC (Habier et al., 2011) uses normal density as the slab.

Here, we performed genomic prediction of salinity stress tolerance in maize to understand the feasibility of using this tool to improve this complex trait. The study compares different GS models, marker densities, and marker selection approaches to provide an efficient approach that is complementary to field-based salinity stress tolerance screenings in maize.

2 | MATERIAL AND METHODS

2.1 | Phenotypic dataset and growth conditions

The dataset used in this study consists of 399 inbred lines from a diversity panel (Mazaheri et al., 2019) that were evaluated in a greenhouse lysimeter system at the US Salinity Laboratory, Riverside, CA. The detailed procedure of phenotyping and data analysis is described in detail elsewhere (Sandhu et al., 2020). In brief, the inbred lines were evaluated in lysimeter sand tanks of size 120 cm (L) × 60 cm (W) × 50 cm (D). Seeds were germinated and grown using a half-strength Hoagland solution with essential macro and micronutrients. Eighteen days after sowing, plants were treated with saline solution ($EC_w = 16$ dS m^{-1}) in a modified half-strength Hoagland solution base with a ratio of ($Ca = 1.25Mg = 0.25Na$) among cations. Another set of plants was treated with half-strength Hoagland solution during this period and considered as control. Plants were harvested 2 weeks after the salt initiation treatment, and data were recorded for shoot length (SL), shoot weight (SW), root length (RL), and root weight (RW) on both the control and treatment plants, thus resulting in eight primary traits. A salt tolerance index (STI) was calculated for each trait by dividing the phenotypic value of a genotype in salt-treated tanks by the phenotypic value in control tanks. BLUPs were calculated for each trait and four different salt tolerance indices, including the SL STI (SL_STI), the SW STI (SW_STI), the RL STI (RW_STI), and the RW STI (RW_STI). All eight primary traits and four derived traits were predicted using GS. Out of a total of 399 inbred lines phenotyped, 358 were used for GS analysis.

2.2 | Genotypic data and its processing

Genome-wide single nucleotide polymorphisms (SNPs) used in this study were derived from an RNA-seq analysis of a panel of 942 diverse inbred lines that provided 899,784 SNPs (Mazaheri et al., 2019). SNPs with a minor allele frequency less than 0.05 and missing values of more than 20% and inbred lines with more than 80% missing values were removed using the R package *snpReady* (Granato et al., 2018). This process resulted in 444,235 SNPs and 358 inbred lines for downstream analysis. Different marker sets with varying marker densities (400K, 300K, 200K, 100K, 50K, 25K, and 10K) were derived by random selection (10 iterations) of SNPs and by testing the association of each marker with a trait using one-way analysis of variance (ANOVA) analysis of linear regression of predictor on the response variable. The response variables were normally distributed (data not shown). In this ANOVA method, the markers showing significant association with trait variation ($p < 0.05$) were retained for GEBV prediction. For the markers selected by ANOVA, linkage disequilibrium (LD)-based pruning was done to filter out other markers that are in LD ($r^2 > 0.80$) with any other markers.

2.3 | Effect of marker density

To preliminarily investigate the effect of different marker densities, the GBLUP model with five-fold cross-validation was implemented using the “*bwgs.cv*” function from the *BWGS* package in R programming language (Charmet et al., 2020). Random selection of markers was done 10 times, and each time the traits were analyzed with 10 iterations of cross-validation. The predictive ability (PA) for all the folds in each iteration was averaged and plotted for comparison. The PA was defined as the value of Pearson’s correlation coefficient between BLUPs of phenotypic values and the GEBV of predicted individuals.

2.4 | Model selection and heritability estimation

We studied GBLUP, EGBLUP, rrBLUP, BRR, BL, BayesA, BayesB, BayesC, and RKHS models for five marker densities (100K, 200K, 300K, 400K, and ANOVA + LD) under five-fold cross-validation. The SNPs were sampled 10 times, and each set of SNPs was analyzed for 10 iterations of cross-validation, resulting in a total of 100 iterations. For Bayesian models and RKHS, the BGLR library of R was used inside the BWGS package (Pérez & De Los Campos, 2014). The PA was calculated as defined in Section 2.3. The best-performing model was used for all further analyses. The standard GBLUP model with a focus on additive genetic effects can be given as follows:

$$\mathbf{y} = \mathbf{1}_n \mu + \mathbf{Z}\mathbf{g} + \mathbf{e}. \tag{1}$$

Here, $\mathbf{y}$ is an n-dimensional vector of phenotypic records (n is the number of genotypes), $\mathbf{1_n}$ is an n-dimensional vector of ones, μ is the population mean, $\mathbf{Z}$ is the design matrix linking phenotypic values to the genotypic values, $\mathbf{g}$ is an n-dimensional vector of additive genotypic value, and $\mathbf{e}$ is a vector for residual terms. The model assumes $\mathbf{g}$ and $\mathbf{e}$ as random parameters with $\mathbf{g} \sim N(0, \mathbf{G}\sigma_g^2)$, and $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$. The G depicts the genomic relationship matrix ($n \times n$) among all individuals. Unlike GBLUP, which utilizes marker derived relationship matrix for predictions, rrBLUP uses marker effects to estimate GEBV. The basic rrBLUP model is given below. Here, $\mathbf{y}$ and $\mathbf{e}$ are as mentioned in Equation (1), and $\mathbf{W}$, $\mathbf{G}$ and U are the design matrix, the genotype matrix of SNP markers, and vector of marker effects, respectively, with $\mathbf{U} \sim N(0, \mathbf{I}\sigma_u^2)$.

$$\mathbf{y} = \mathbf{WGU} + \mathbf{e}. \tag{2}$$

Whereas GBLUP considers only additive genetic effects, EGBLUP considers additive × additive epistasis in addition to the additive genetic effect with the formula:

$$\mathbf{y} = \mathbf{1}_n \mu + \mathbf{g_1} + \mathbf{g_2} + \mathbf{e}, \tag{3}$$

where $\mathbf{y}$, 1_n, μ, and e are the same as described in Equation (1). The g_1 and g_2 are an n-dimensional vector of additive genotypic values and an n-dimensional vector of additive × additive epistatic genotypic values, respectively. The model assumes that $\mathbf{g_1} \sim N(0, \mathbf{G}\sigma_1^2)$ and $\mathbf{g_2} \sim N(0, \mathbf{H}\sigma_2^2)$, where $\mathbf{H}$ is an epistatic relationship matrix obtained as the Hadamard product of the additive relationship matrix (i.e., G) by itself. Another approach for modeling epistasis for GS is using RKHS regression (Gianola & Van Kaam, 2008; Gianola et al., 2006). An RKHS model is described in Equation (4) with the same notations as in Equation (1) and (3).

$$\mathbf{y} = \mathbf{1_n} \mu + \mathbf{g} + \mathbf{e}. \tag{4}$$

The model assumes that $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$, and $\mathbf{g} \sim N(0, \mathbf{K}\sigma_g^2)$, where $\mathbf{K} = (k(x_i, x_j))$ depicts a kernel matrix ($n \times n$) whose entries are functions of the marker profiles of genotype i and j. Bayesian methods of GS may lead to higher prediction accuracy (PA) by incorporating prior information and capturing complex relationships. The basic equation is represented as

$$y_i = \eta_i + e_i. \tag{5}$$

Here, $(i = 1, .., n)$, and η is a linear predictor which represents a conditional expectation function as

$$\eta = 1\mu + \sum_{j=1}^{J} \mathbf{X_j}\beta_j + \sum_{l=1}^{L} \mathbf{u_l}, \tag{6}$$

where μ is the intercept, $\mathbf{X}_j$ are design matrices for predictors, $\boldsymbol{\beta}_j$ are vectors of the effects associated to the columns of $\mathbf{X}_j$, and $\mathbf{u}_l$ are vectors of random effects. Based on the above criteria, a conditional distribution equation of the data is defined, and prior densities are assigned to different unknown coefficients in the equation. Based on the choice of prior densities of the regression coefficients ($\boldsymbol{\beta}_j$), the models are differentiated into BRR, BL, BayesA, BayesB, and BayesC. The Gaussian prior is used in BRR, shrinking all the effects to a similar extent (Gianola, 2013). The Laplace prior is used in the BL (Park & Casella, 2008), whereas the scaled-t density prior is used in BayesA (Meuwissen et al., 2001). For BayesB and BayesC, two different finite mixture priors are used. In BayesB, these are a mixture of a point of mass at zero and a scaled-t slab (Meuwissen et al., 2001), and in BayesC, they are a mixture of a point of mass at zero and a Gaussian slab (Habier et al., 2011).

Narrow-sense heritability (h^2) was estimated for eight primary traits and four derived traits using ASReml-R software (Gilmour et al., 2021). Variance components were estimated by using model (1) in ASReml. The additive relationship matrix was calculated using SNPs by the “A.mat” function of the rrBLUP package in R, and narrow-sense heritability was calculated as follows:

$$h^2 = \frac{\sigma_g^2}{\sigma_p^2},$$

where σ_g^2 is additive genetic variance, and σ_p^2 is the phenotypic variance.

3 | RESULTS

3.1 | Low-density markers selected using the ANOVA + LD approach showed the best prediction accuracy

Random selection of markers resulted in very low PA for all four traits under GBLUP, and changing the marker density did not improve PA (Figure 1). PA significantly improved after selecting markers through ANOVA. The numerical values of PAs for markers selected through ANOVA and ANOVA+LD are given in supplementary data (Table S2). The distribution of these markers across the studied traits is displayed in Venn diagrams (Figure S1). The chromosome-wise distribution of markers is presented in the supplementary data (Figure S2). The ANOVA + LD approach showed the highest average PA among all maker selection criteria. For further analysis, markers selected through ANOVA + LD approach were used. The response of different models under different densities of randomly selected markers (100K, 200K, 300K, and 400K) is shown in Figure 2. All other models performed better than GBLUP for SL_STI and RL_STI. BayesB showed increasing PA upon increasing the marker densities for SL_STI, RL_STI, and RW_STI. Interestingly, BRR gave negative PA values for SW_STI when 300K marker sets were used. RKHS showed significantly lower PA than EGBLUP for SL_STI, RL_STI, and SW_STI and higher PA than EGBLUP for RW_STI for all the four marker densities. BayesC and BRR gave an increasing trend of PAs upon increasing marker densities for RL_STI.

3.2 | Model selection

Various GS models were fitted on four derived salinity indices to identify the best model. The analysis of a set of 358 inbred lines and markers selected through the ANOVA + LD approach showed that GBLUP, EGBLUP, and rrBLUP and all Bayesian models performed similarly (Figure 3). The rrBLUP model performed slightly better than GBLUP for RW_SI. The differences between different PA models were practically negligible for all the traits (Figure 3). All Bayesian models, with the exception of BayesB, were on par with GBLUP for all four traits. BayesB showed the lowest average PA values for all the traits, followed by EGBLUP (Figure 3).

3.3 | Traits under the control treatment showed better prediction accuracy than those under the salinity treatment

GBLUP was selected to fit the GS model on eight primary traits: SL and SW, RL, and RW under normal and saline conditions. The primary traits were evaluated to compare the control versus the salt treatment. The average PA of GBLUP for these traits varied from 0.59 to 0.73 (Figure 4). PA for SL, RL, and RW was slightly higher for the control than for the salt treatment (Figure 4) and the control and treatment for SW. Among the salt indices, SL_STI and RL_STI showed considerably higher PA than SW_STI and RW_STI, and the average PA for SW_STI was below 0.50 (Figure 4). The exact values of average PA for each trait with their standard deviations are given in Table S1.

3.4 | Heritability was positively correlated with PA

Narrow-sense heritability ranged between 0.35 and 0.54 for the eight primary traits and between 0.31 and 0.46 for the four derived traits (Table 1). Salt tolerance indices of all four traits showed lower heritability than their respective primary (both control and salt treatments) traits. A significant positive correlation between heritability and PA ($p = 0.05$) was found with an R^2 value of 0.43 (Figure S3).

4 | DISCUSSION

GS has been implemented in several cereal breeding programs for predicting various traits related to biotic and abiotic stress tolerance (Cao et al., 2021; Dias et al., 2018; dos Santos et al., 2016; Lorenz et al., 2012; Shikha et al., 2017). GS on biotic stress tolerance traits in maize has shown PA as high as 0.706 for northern corn leaf blight resistance (Lorenz et al., 2012),

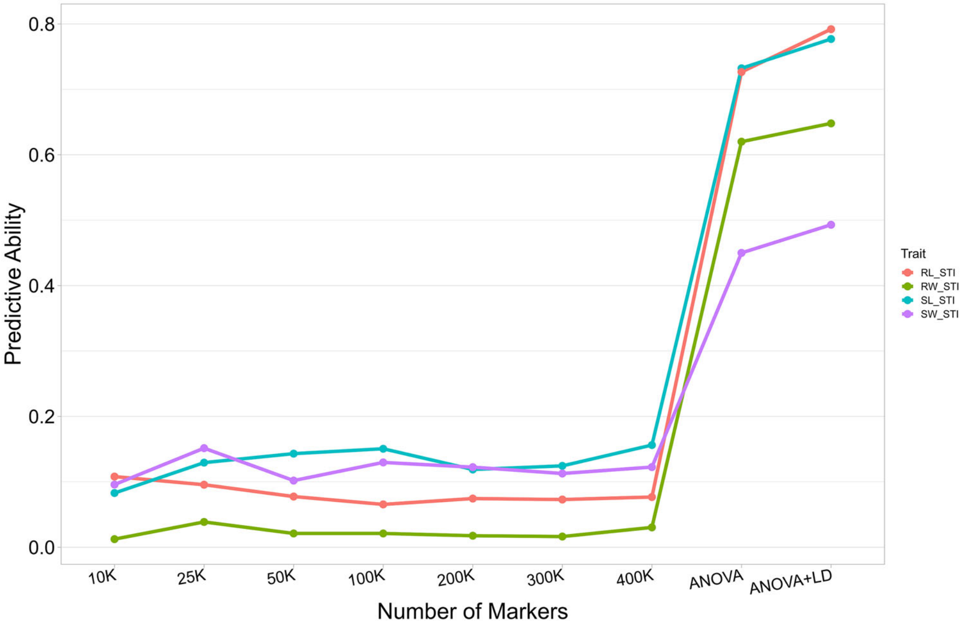

FIGURE 1 Effect of the number of markers on cross-validation predictive ability of the genomic best linear unbiased prediction (GBLUP) model for salt indices of different traits. RL, root length; RW, root weight; SL, shoot length; STI, salt tolerance index; SW, shoot weight.

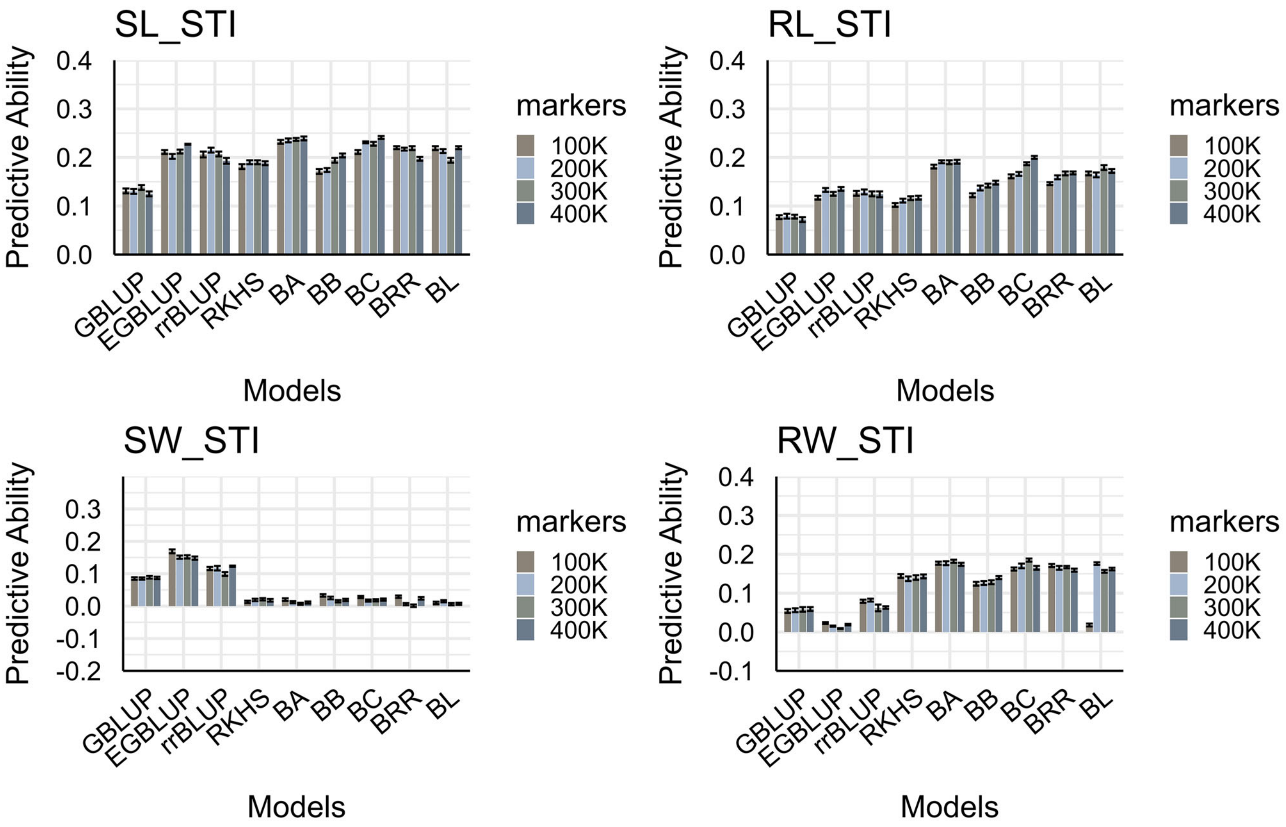

FIGURE 2 Effect of the number of markers on cross-validation predictive ability of different models for salt tolerance indices of different traits. RL, root length; RW, root weight; SL, shoot length; STI, salt tolerance index; SW, shoot weight.

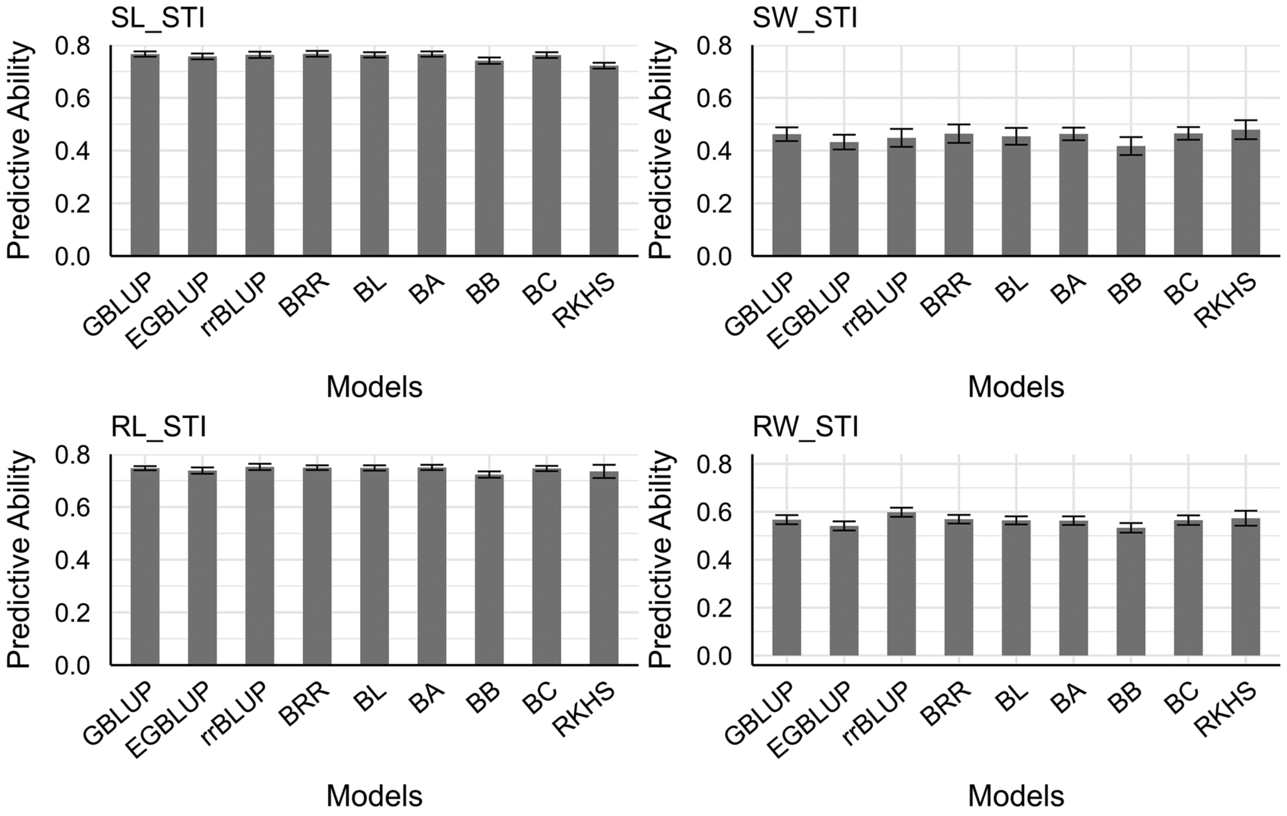

FIGURE 3 Predictive abilities of four derived traits for different genomic selection models. GBLUP, genomic best linear unbiased prediction; EGBLUP, extended EGBLUP; rrBLUP, ridge regression BLUP; BRR, Bayesian ridge regression; BA, BayesA; BB, BayesB; BC, BayesC; BL, Bayesian Lasso; RKHS, reproducing kernel Hillbert spaces; RL, root length; RW, root weight; SL, shoot length; STI, salt tolerance index; SW, shoot weight.

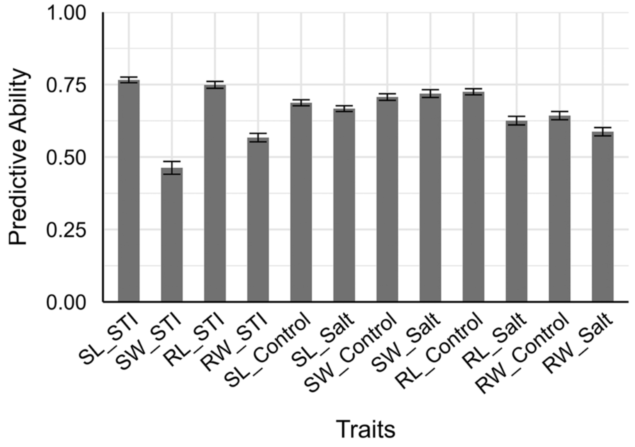

FIGURE 4 Predictive ability of GBLUP for different traits in fivefold cross-validation. RL, root length; RW, root weight; SL, shoot length; STI, salt tolerance index; SW, shoot weight.

0.878 for ear rot resistance (dos Santos et al., 2016), and 0.87 for tar spot resistance (Cao et al., 2021). GS has been successfully implemented for the improvement of drought tolerance in maize (Dias et al., 2018; Shikha et al., 2017). Response to GS varies among crop species and depends on the genetic architecture of the trait (Morgante et al., 2018), marker density (Q. Wang et al., 2017), size of the TRP (A. Zhang et al., 2017), and level of relatedness between the TRP and TSP (Fraslin et al., 2022; Lorenz et al., 2015). Therefore, it is very important to investigate the potential level of PA for the target trait before deploying GS on a larger scale.

The number of markers used for the analysis is an important and widely studied aspect of GS. Very low PA values were observed in cross-validation with a higher density of markers in our study (Figure 2). However, a substantial gain in the PA was observed upon selecting markers based on linear regression of each marker with trait values followed by ANOVA analysis, and a further increase in the PA was observed after trimming the markers based on LD criteria (ANOVA + LD). Being fundamentally a regression of marker data to trait variation, this ANOVA + LD method is partly similar to genome-wide association study (GWAS) conducted without consideration of population structure and adjustment of *p*-values for multiple comparisons. GWAS-based marker prioritization has been explored in previous studies (Ling et al., 2022; Sehgal et al., 2020; Veerkamp et al., 2016). One study reported a decrease in PA for protein yield in Holstein-Friesian bulls (Veerkamp et al., 2016), while

TABLE 1 Heritability for different traits.

Trait	STI				Control				Salt			
	SL	SW	RL	RW	SL	SW	RL	RW	SL	SW	RL	RW
Heritability	0.424	0.313	0.456	0.320	0.505	0.531	0.543	0.365	0.476	0.379	0.494	0.348

Abbreviations: RL, root length; RW, root weight; SL, shoot length; STI, salt tolerance index; SW, shoot weight.

another study reported no increase in PA upon implementing marker selection strategies (Ober et al., 2012). In a simulation study, increased PA has been reported upon filtering markers based on F_{ST}, a measure of allele frequency variation among subpopulations (Chang et al., 2018). In another study focused on trait-associated markers in GS (Sehgal et al., 2020), up to 10% of PA was observed for grain yield in wheat (*Triticum aestivum*) after using loci identified by GWAS as fixed effects in the model (Sehgal et al., 2020). Similar results were reported in rice (*Oryza sativa*; Anilkumar et al., 2023), where PA reduced significantly after dropping GWAS-detected markers (Anilkumar et al., 2023). These reports suggest the importance of marker prioritization for GS. The ANOVA + LD approach used in our study selects markers based on the data from the training population alone, so it would be interesting to validate this approach on untested genotypes. The GS models showed different patterns for PA under higher densities of markers in comparison to when ANOVA + LD method was used. This observation suggests a need for further investigation into the behavior of GS models with respect to the trait under study and the marker set used. The SNP data used in this study were exclusively from the expressed regions of the genome. Including SNPs from non-genic regions, as reported earlier (Liang et al., 2020; Tan et al., 2017), may improve the PA, considering that non-genic regions of the genome have critical regulatory functions (Song et al., 2021).

When implementing GS for trait improvement, selecting a predictive model suitable for a specific crop-trait scenario is important. Different GS models have advantages and disadvantages (Crossa et al., 2017), and the genetic architecture of the target trait partially determines their effectiveness (Daetwyler et al., 2010). Genetic architecture is defined by the number of QTL controlling a trait, their effect sizes, and their intra-locus and inter-loci interactions. Model assumptions about genetic architecture as well as the treatment of marker effects within models can influence the PA of GS (J. Wang et al., 2018). A GWAS on the population evaluated in this study previously reported that all the identified salt indices-associated SNPs had minor effects with R^2 values less than 0.13 (Sandhu et al., 2020). This finding indicates the complex genetic architecture of salt tolerance in maize, which suggests the suitability of GS for improving these traits. The lack of significant differences between Bayesian models, rrBLUP, and GBLUP is consistent with the results previously reported in predicting complex traits (Daetwyler et al., 2010; Endelman, 2011; Heslot et al., 2012; Lorenz et al., 2012; Merrick & Carter, 2021; Rolling et al., 2020). The similar PA for GBLUP and rrBLUP for all the studied traits may be due to their theoretical equivalence (Goddard, 2009; Hayes et al., 2009), and their similar performance has been reported previously (Jacquin et al., 2016; Zhu et al., 2021). A slightly higher PA of rrBLUP in comparison to GBLUP for RW_STI may be attributed to the method of calculation of the genetic relationship matrix in GBLUP (Endelman, 2011). The EGBLUP model performed poorer than most of the models for all four salt tolerance indices. The lower PA of EGBLUP suggests either the low importance of epistasis for explaining these traits or the inability of EGBLUP to model the epistatic effects appropriately. Some studies reported an increase (Z. Hu et al., 2011; Jiang et al., 2018; Martini et al., 2017; Miller et al., 2023; Vojgani et al., 2021), while others reported a decrease (Jiang & Reif, 2015; Lorenzana & Bernardo, 2009) in PA upon modeling epistasis, indicating trait specific response. The difference in PA between rrBLUP and EGBLUP for RW_STI was 0.05, suggesting that these models would likely provide similar levels of realized genetic gain when applied.

BayesB model assumes that most markers do not affect the trait and show higher PA when a few QTLs have a large effect size (Meuwissen et al., 2001). None of the SNPs in the previously reported GWAS study were found to have major effects on the traits tested (Sandhu et al., 2020), which may explain the relatively poor performance of BayesB. Another probable reason for the lower PAs of BayesB may be the lower heritability of the studied traits, as BayesB assumes a more realistic distribution in marker effect for high heritability traits (Zhu et al., 2021). Higher PAs have been reported for traits with higher heritability (Bolormaa et al., 2013; Daetwyler et al., 2010). Therefore, one of the probable explanations for the low PA of SW_STI and RW_STI, compared to SL_STI and RL_STI may be their lower heritability.

GS may be used for narrowing down selection candidates from a large breeding population to a smaller subset for field evaluation, thereby improving the cost-efficiency and feasibility of breeding for salinity tolerance. Higher genetic gains have been reported in maize for abiotic stress using GS over phenotypic selection (Vivek et al., 2017). Rapid cycle GS (RCGS) has recently been used in maize to improve drought and waterlogging tolerance with a substantial genetic gain and without any yield penalty under optimum conditions (Das et al., 2020). Several studies have reported a minimal loss of genetic diversity in the advanced breeding cycles under RCGS schemes (Das et al., 2020; Dreisigacker et al., 2023; Zhang et al., 2017). Implementing RCGS for salinity stress tolerance in maize would be helpful to increase genetic gains for this trait. The next critical step following the results presented herein is to empirically measure the realized gain from the selection when applying GS to improve salinity tolerance in maize.

5 | CONCLUSION

Several research programs focused on genetic, physiological, molecular, and genomic aspects of salt tolerance in crops have expanded our mechanistic understanding of this complex trait. However, these efforts have not been fully translated into applied breeding programs to improve salt tolerance in maize. With progressively reduced genotyping costs, phenotyping has become a rate-limiting step in crop improvement, especially for biotic and abiotic stress-related traits influenced by the environment. Selecting markers based on ANOVA and LD criteria may be useful for improving PA and the cost-effectiveness of genotyping. The present study concludes that GS is a promising approach for developing salt-tolerant maize cultivars and provides a roadmap for implementing GS in salinity-tolerance breeding programs.

AUTHOR CONTRIBUTIONS

Vishal Singh: conceptualization; formal analysis; investigation; methodology; visualization; writing—original draft. **Margaret Krause**: formal analysis; supervision; writing—review and editing. **Devinder Sandhu**: formal analysis; methodology; writing—review and editing. **Rajandeep S. Sekhon**: formal analysis; methodology; writing—review and editing. **Amita Kaundal**: conceptualization; data curation; funding acquisition; project administration; resources; software; supervision; validation; writing—review and editing.

ACKNOWLEDGMENTS

We acknowledge the Indian Council of Agricultural Research (ICAR) for Netaji Subhash ICAR International Fellowships to Vishal Singh. This study was supported by New Faculty Startup funding provided by the Department of Plants, Soils, and Climate, College of Agriculture and Applied Science, USU, and Utah Agriculture Experiment Station (UAES) office to Amita Kaundal and approved as journal manuscript number 9675. We acknowledge CHPC at the University of Utah for providing computational help.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

All the relevant data have been provided as tables, figures in the main text, and supplemental data.

ORCID

Vishal Singh https://orcid.org/0000-0003-4926-7411
Margaret Krause https://orcid.org/0000-0003-1154-337X
Devinder Sandhu https://orcid.org/0000-0003-4193-3408
Amita Kaundal https://orcid.org/0000-0002-9154-1173

REFERENCES

Acharya, B. R., Sandhu, D., Dueñas, C., Dueñas, M., Pudussery, M., Kaundal, A., Ferreira, J. F. S., Suarez, D. L., & Skaggs, T. H. (2022). Morphological, physiological, biochemical, and transcriptome studies reveal the importance of transporters and stress signaling pathways during salinity stress in *Prunus*. *Scientific Reports*, *12*, 1274. https://doi.org/10.1038/s41598-022-05202-1

Ali Turan, M., Hassan Awad Elkarim, A., Taban, N., & Taban, S. (2009). Effect of salt stress on growth, stomatal resistance, proline and chlorophyll concentrations on maize plant. *African Journal of Agricultural Research*, *4*(9), 893–897.

Amacher, J., Koenig, R., & Kitchen, B. (2000). Salinity and plant tolerance. *AG-SO*, *3*, 1–8. http://digitalcommons.usu.edu/extension_histall/43

Anilkumar, C., Muhammed Azharudheen, T. P., Sah, R. P., Sunitha, N. C., Devanna, B. N., Marndi, B. C., & Patra, B. C. (2023). Gene based markers improve precision of genome-wide association studies and accuracy of genomic predictions in rice breeding. *Heredity*, *130*, 335–345. https://doi.org/10.1038/S41437-023-00599-5

Aragüés, R., Medina, E. T., Zribi, W., Clavería, I., Álvaro-Fuentes, J., & Faci, J. (2014). Soil salinization as a threat to the sustainability of deficit irrigation under present and expected climate change scenarios. *Irrigation Science*, *33*(1), 67–79. https://doi.org/10.1007/s00271-014-0449-x

Barutcular, C., Koc, M., Unlu, M., Kanber, R., Ozekici, B., Yano, T., & Watanabe, T. (2005). Maize photosynthesis and productivity in relation to sowing time in a Mediterranean environment, Cukurova. *Asian Journal of Plant Sciences*, *4*, 550–554. https://doi.org/10.3923/AJPS.2005.550.554

Bassiouny, E., Hala, M. S., & Bekheta, M. A. (2004). Role of putrescine on growth, regulation of stomatal aperture, ionic contents and yield by two wheat cultivars under salinity stress. *Egyptian Journal of Physiological Sciences*, *26*, 95–114.

Bolormaa, S., Pryce, J. E., Kemper, K., Savin, K., Hayes, B. J., Barendse, W., Zhang, Y., Reich, C. M., Mason, B. A., Bunch, R. J., Harrison, B. E., Reverter, A., Herd, R. M., Tier, B., Graser, H. U., & Goddard, M. E. (2013). Accuracy of prediction of genomic breeding values for residual feed intake and carcass and meat quality traits in *Bos taurus*, *Bos indicus*, and composite beef cattle. *Journal of Animal Science*, *91*, 3088–3104. https://doi.org/10.2527/JAS.2012-5827

Brain, R., Perkins, D., Ghebremichael, L., White, M., Goodwin, G., & Aerts, M. (2023). The shrinking land challenge. *ACS Agricultural Science and Technology*, *3*, 152–157. https://doi.org/10.1021/acsagscitech.2c00250

Cao, S., Song, J., Yuan, Y., Zhang, A., Ren, J., Liu, Y., Qu, J., Hu, G., Zhang, J., Wang, C., Cao, J., Olsen, M., Prasanna, B. M., San Vicente, F., & Zhang, X. (2021). Genomic prediction of resistance to tar spot complex of maize in multiple populations using genotyping-by-sequencing SNPs. *Frontiers in Plant Science*, *12*, Article 672525. https://doi.org/10.3389/FPLS.2021.672525

Chang, L. Y., Toghiani, S., Ling, A., Aggrey, S. E., & Rekaya, R. (2018). High density marker panels, SNPs prioritizing and accuracy of genomic selection. *BMC Genetics*, *19*, Article 4. https://doi.org/10.1186/S12863-017-0595-2

Charmet, G., Tran, L. G., Auzanneau, J., Rincent, R., & Bouchet, S. (2020). BWGS: A R package for genomic selection and its application to a wheat breeding programme. *PLoS One*, *15*, e0222733. https://doi.org/10.1371/JOURNAL.PONE.0222733

Chinnusamy, V., Jagendorf, A., & Zhu, J. K. (2005). Understanding and improving salt tolerance in plants. *Crop Science*, *45*, 437–448. https://doi.org/10.2135/CROPSCI2005.0437

Corwin, D. L. (2021). Climate change impacts on soil salinity in agricultural areas. *European Journal of Soil Science*, *72*, 842–862. https://doi.org/10.1111/EJSS.13010

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

Daetwyler, H. D., Pong-Wong, R., Villanueva, B., & Woolliams, J. A. (2010). The impact of genetic architecture on genome-wide evaluation methods. *Genetics*, *185*, 1021–1031. https://doi.org/10.1534/GENETICS.110.116855

Das, R. R., Vinayan, M. T., Patel, M. B., Phagna, R. K., Singh, S. B., Shahi, J. P., Sarma, A., Barua, N. S., Babu, R., Seetharam, K., Burgueño, J. A., & Zaidi, P. H. (2020). Genetic gains with rapid-cycle genomic selection for combined drought and waterlogging tolerance in tropical maize (*Zea mays* L.). *The Plant Genome*, *13*(3), e20035. https://doi.org/10.1002/TPG2.20035

Davenport, R., James, R. A., Zakrisson-Plogander, A., Tester, M., & Munns, R. (2005). Control of sodium transport in durum wheat. *Plant Physiology*, *137*, 807–818. https://doi.org/10.1104/pp.104.057307

de los Campos, G., Hickey, J. M., Pong-Wong, R., Daetwyler, H. D., & Calus, M. P. L. (2013). Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics*, *193*, 327–345. https://doi.org/10.1534/genetics.112.143313

Dias, K. O. D. G., Gezan, S. A., Guimarães, C. T., Nazarian, A., Da, C. E., Silva, L., Parentoni, S. N., De Oliveira Guimarães, P. E., De Oliveira Anoni, C., Pádua, J. M. V., De Oliveira Pinto, M., Noda, R. W., Ribeiro, C. A. G., De Magalhães, J. V., Garcia, A. A. F., De Souza, J. C., Guimarães, L. J. M., & Pastina, M. M. (2018). Improving accuracies of genomic predictions for drought tolerance in maize by joint modeling of additive and dominance effects in multi-environment trials. *Heredity*, *121*, 24–37. https://doi.org/10.1038/s41437-018-0053-6

Dikobe, T. B., Mashile, B., Sinthumule, R. R., Ruzvidzo, O., Dikobe, T. B., Mashile, B., Sinthumule, R. R., & Ruzvidzo, O. (2021). Distinct morpho-physiological responses of maize to salinity stress. *American Journal of Plant Sciences*, *12*, 946–959. https://doi.org/10.4236/AJPS.2021.126064

dos Santos, J. P. R., Pires, L. P. M., de Castro Vasconcellos, R. C., Pereira, G. S., Von Pinho, R. G., & Balestre, M. (2016). Genomic selection to resistance to *Stenocarpella maydis* in maize lines using DArTseq markers. *BMC Genetics*, *17*, Article 86. https://doi.org/10.1186/S12863-016-0392-3

Dreisigacker, S., Pérez-Rodríguez, P., Crespo-Herrera, L., Bentley, A. R., & Crossa, J. (2023). Results from rapid-cycle recurrent genomic selection in spring bread wheat. *G3: Genes|Genomes|Genetics*, *13*, jkad025. https://doi.org/10.1093/G3JOURNAL/JKAD025

El Sayed, H., & El Sayed, A. (2011). Influence of salinity stress on growth parameters, photosynthetic activity and cytological studies of *Zea mays*, L. plant using hydrogel polymer. *Agriculture and Biology Journal of North America*, *2*, 907–920. https://doi.org/10.5251/ABJNA.2011.2.6.907.920

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome*, *4*, 250–255. https://doi.org/10.3835/PLANTGENOME2011.08.0024

Erenstein, O., Jaleta, M., Sonder, K., Mottaleb, K., & Prasanna, B. M. (2022). Global maize production, consumption and trade: Trends and R&D implications. *Food Security*, *14*(5), 1295–1319. https://doi.org/10.1007/s12571-022-01288-7

FAO Stat. (2021). FAOSTAT. https://www.fao.org/faostat/en/#home

Farooq, M., Hussain, M., Wakeel, A., & Siddique, K. H. M. (2015). Salt stress in maize: Effects, resistance mechanisms, and management. A review. *Agronomy for Sustainable Development*, *35*, 461–481. https://doi.org/10.1007/s13593-015-0287-0

Fraslin, C., Yáñez, J. M., Robledo, D., & Houston, R. D. (2022). The impact of genetic relationship between training and validation populations on genomic prediction accuracy in Atlantic salmon. *Aquaculture Reports*, *23*, 101033. https://doi.org/10.1016/J.AQREP.2022.101033

Gianola, D. (2013). Priors in whole-genome regression: The Bayesian alphabet returns. *Genetics*, *194*, 573. https://doi.org/10.1534/GENETICS.113.151753

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

Gianola, D., & Van Kaam, J. B. C. H. M. (2008). Reproducing kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics*, *178*, 2289–2303. https://doi.org/10.1534/GENETICS.107.084285

Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J., & Thompson, R. (2021). *ASReml user guide release 4.2 functional specification*.

Goddard, M. (2009). Genomic selection: Prediction of accuracy and maximisation of long term response. *Genetica*, *136*, 245–257. https://doi.org/10.1007/S10709-008-9308-0

Granato, I. S. C., Galli, G., de Oliveira Couto, E. G., e Souza, M. B., Mendonça, L. F., & Fritsche-Neto, R. (2018). *snpReady*: A tool to assist breeders in genomic analysis. *Molecular Breeding*, *38*, Article102. https://doi.org/10.1007/S11032-018-0844-8

Habier, D., Fernando, R. L., Kizilkaya, K., & Garrick, D. J. (2011). Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics*, *12*, Article 186. https://doi.org/10.1186/1471-2105-12-186

Hasan, M. S., Singh, V., Islam, S., Islam, M. S., Ahsan, R., Kaundal, A., Islam, T., & Ghosh, A. (2021). Genome-wide identification and expression profiling of glutathione S-transferase family under multiple abiotic and biotic stresses in *Medicago truncatula* L. *PLoS One*, *16*, e0247170. https://doi.org/10.1371/JOURNAL.PONE.0247170

Hayes, B. J., Bowman, P. J., Chamberlain, A. J., & Goddard, M. E. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science*, *92*, 433–443. https://doi.org/10.3168/JDS.2008-1646

Heslot, N., Yang, H.-P., Sorrells, M. E., Jannink, J.-L., Heslot, N., Sorrells, M. E., Yang, H.-P., & Jannink, J.-L. (2012). Genomic selection in plant breeding: A comparison of models. *Crop Science*, *52*, 146–160. https://doi.org/10.2135/CROPSCI2011.06.0297

Hiyane, R., Hiyane, S., Tang, A. C., & Boyer, J. S. (2010). Sucrose feeding reverses shade-induced kernel losses in maize. *Annals of Botany*, *106*, 395–403. https://doi.org/10.1093/AOB/MCQ132

Hoerl, A. E., & Kennard, R. W. (2000). Ridge regression: Biased estimation for non-orthogonal problems. *Technometrics*, *42*, 80–86. https://doi.org/10.2307/1271436

Hu, D., Li, R., Dong, S., Zhang, J., Zhao, B., Ren, B., Ren, H., Yao, H., Wang, Z., & Liu, P. (2022). Maize (*Zea mays* L.) responses to salt stress in terms of root anatomy, respiration and antioxidative enzyme activity. *BMC Plant Biology*, *22*, 602. https://doi.org/10.1186/S12870-022-03972-4

Hu, Z., Li, Y., Song, X., Han, Y., Cai, X., Xu, S., & Li, W. (2011). Genomic value prediction for quantitative traits under the epistatic model. *BMC Genetics*, *12*, 15. https://doi.org/10.1186/1471-2156-12-15

Hussain, M., Park, H.-W., Farooq, M., Jabran, K., & Lee, D.-J. (2013). Morphological and physiological basis of salt resistance in different rice genotypes. *International Journal of Agriculture and Biology*, *15*, 113–118.

Hussain, S., Hussain, S., Ali, B., Ren, X., Chen, X., Li, Q., Saqib, M., & Ahmad, N. (2021). Recent progress in understanding salinity tolerance in plants: Story of Na^+/K^+ balance and beyond. *Plant Physiology and Biochemistry*, *160*, 239–256. https://doi.org/10.1016/j.plaphy.2021.01.029

Hütsch, B. W., Saqib, M., Osthushenrich, T., & Schubert, S. (2014). Invertase activity limits grain yield of maize under salt stress. *Journal of Plant Nutrition and Soil Science*, *177*, 278–286. https://doi.org/10.1002/JPLN.201300345

Jacquin, L., Cao, T.-V., & Ahmadi, N. (2016). A unified and comprehensible view of parametric and kernel methods for genomic prediction with application to rice. *Frontiers in Genetics*, *7*, Article 145. https://doi.org/10.3389/fgene.2016.00145

Jiang, Y., & Reif, J. C. (2015). Modeling epistasis in genomic selection. *Genetics*, *201*, 759–768. https://doi.org/10.1534/GENETICS.115.177907

Jiang, Y., Schmidt, R. H., & Reif, J. C. (2018). Haplotype-based genome-wide prediction models exploit local epistatic interactions among markers. *G3 Genes|Genomes|Genetics*, *8*, 1687–1699. https://doi.org/10.1534/g3.117.300548

Jung, S., Hütsch, B. W., & Schubert, S. (2017). Salt stress reduces kernel number of corn by inhibiting plasma membrane H^+-ATPase activity. *Plant Physiology and Biochemistry*, *113*, 198–207. https://doi.org/10.1016/J.PLAPHY.2017.02.009

Kaundal, A., Sandhu, D., Singh, V., Duenas, M., Acharya, B. R., Nelson, B., Ferreira, J. F. S., & Litt, A. (2022). Transgenic expression of *Prunus persica* Salt Overly Sensitive 2 (*PpSOS2*) in the *atsos2* mutant imparts salt tolerance in Arabidopsis. *ACS Agricultural Science and Technology*, *2*, 153–164. https://doi.org/10.1021/acsagscitech.1c00276

Krchov, L. M., & Bernardo, R. (2015). Relative efficiency of genomewide selection for testcross performance of doubled haploid lines in a maize breeding program. *Crop Science*, *55*, 2091–2099. https://doi.org/10.2135/cropsci2015.01.0064

Kumar, P., Choudhary, M., Jat, B. S., Kumar, B., Singh, V., Kumar, V., Singla, D., & Rakshit, S. (2021). Skim sequencing: An advanced NGS technology for crop improvement. *Journal of Genetics*, *100*, Article 38. https://doi.org/10.1007/s12041-021-01285-3

Lande, R., & Thompson, R. (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, *124*, 743–756. https://doi.org/10.1093/genetics/124.3.743

Liang, Z., Tan, C., Prakapenka, D., Ma, L., & Da, Y. (2020). Haplotype analysis of genomic prediction using structural and functional genomic information for seven human phenotypes. *Frontiers in Genetics*, *11*, Article 588907. https://doi.org/10.3389/FGENE.2020.588907

Ling, A., Hay, E. H., Aggrey, S. E., & Rekaya, R. (2022). Fuzzy logic as a strategy for combining marker statistics to optimize pre-selection of high-density and sequence genotype data. *Genes*, *13*(11), 2100. https://doi.org/10.3390/GENES13112100

Lorenz, A. J., Smith, K. P., & Jannink, J. L. (2012). Potential and optimization of genomic selection for Fusarium head blight resistance in six-row barley. *Crop Science*, *52*(4), 1609–1621. https://doi.org/10.2135/CROPSCI2011.09.0503

Lorenz, A. J., Smith, K. P., Lorenz, A. J., & Smith, K. P. (2015). Adding genetically distant individuals to training populations reduces genomic prediction accuracy in barley. *Crop Science*, *55*, 2657–2667. https://doi.org/10.2135/CROPSCI2014.12.0827

Lorenzana, R. E., & Bernardo, R. (2009). Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theoretical and Applied Genetics*, *120*, 151–161. https://doi.org/10.1007/S00122-009-1166-3

Maas, E. V., Hoffman, G. J., Chaba, G. D., Poss, J. A., & Shannon, M. C. (1983). Salt sensitivity of corn at various growth stages. *Irrigation Science*, *4*, 45–57. https://doi.org/10.1007/BF00285556

Macciotta, N. P. P., Gaspa, G., Steri, R., Nicolazzi, E. L., Dimauro, C., Pieramati, C., & Cappio-Borlino, A. (2010). Using eigenvalues as variance priors in the prediction of genomic breeding values by principal component analysis. *Journal of Dairy Science*, *93*, 2765–2774. https://doi.org/10.3168/JDS.2009-3029

Martini, J. W. R., Gao, N., Cardoso, D. F., Wimmer, V., Erbe, M., Cantet, R. J. C., & Simianer, H. (2017). Genomic prediction with epistasis models: On the marker-coding-dependent performance of the extended GBLUP and properties of the categorical epistasis model (CE). *BMC Bioinformatics*, *18*, Article 3. https://doi.org/10.1186/S12859-016-1439-1

Mazaheri, M., Heckwolf, M., Vaillancourt, B., Gage, J. L., Burdo, B., Heckwolf, S., Barry, K., Lipzen, A., Ribeiro, C. B., Kono, T. J. Y., Kaeppler, H. F., Spalding, E. P., Hirsch, C. N., Robin Buell, C., de Leon, N., & Kaeppler, S. M. (2019). Genome-wide association analysis of stalk biomass and anatomical traits in maize. *BMC Plant Biology*, *19*, Article 45. https://doi.org/10.1186/s12870-019-1653-x

Merrick, L. F., & Carter, A. H. (2021). Comparison of genomic selection models for exploring predictive ability of complex traits in breeding programs. *Plant Genome*, *14*(3), e20158. https://doi.org/10.1002/TPG2.20158

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*, 1819–1829. https://doi.org/10.1093/GENETICS/157.4.1819

Miller, M. J., Song, Q., Fallen, B., & Li, Z. (2023). Genomic prediction of optimal cross combinations to accelerate genetic improvement of soybean (*Glycine max*). *Frontiers in Plant Science*, *14*, Article 1171135. https://doi.org/10.3389/FPLS.2023.1171135

Morgante, F., Huang, W., Maltecca, C., & Mackay, T. F. C. (2018). Effect of genetic architecture on the prediction accuracy of quantitative traits in samples of unrelated individuals. *Heredity*, *120*, 500–514. https://doi.org/10.1038/s41437-017-0043-0

Munns, R., James, R. A., & Läuchli, A. (2006). Approaches to increasing the salt tolerance of wheat and other cereals. *Journal of Experimental Botany*, *57*(5), 1025–1043. https://doi.org/10.1093/JXB/ERJ100

Munns, R., & Sharp, R. E. (1993). Involvement of abscisic acid in controlling plant growth in soil of low water potential. *Functional Plant Biology*, *20*, 425–437. https://doi.org/10.1071/PP9930425

Ober, U., Ayroles, J. F., Stone, E. A., Richards, S., Zhu, D., Gibbs, R. A., Stricker, C., Gianola, D., Schlather, M., Mackay, T. F. C., & Simianer, H. (2012). Using whole-genome sequence data to predict quantitative trait phenotypes in *Drosophila melanogaster*. *PLOS Genetics*, *8*, e1002685. https://doi.org/10.1371/JOURNAL.PGEN.1002685

Ogutu, J. O., Schulz-Streeck, T., & Piepho, H. P. (2012). Genomic selection using regularized linear regression models: Ridge regression, lasso, elastic net and their extensions. *BMC Proceedings*, *6*, Article S10. https://doi.org/10.1186/1753-6561-6-S2-S10

Park, T., & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, *103*, 681–686. https://doi.org/10.1198/016214508000000337

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

Quintero, J. M., María Fournier, J., & Benlloch, M. (2007). Na^+ accumulation in shoot is related to water transport in K^+-starved sunflower plants but not in plants with a normal K^+ status. *Journal of Plant Physiology*, *164*, 60–67. https://doi.org/10.1016/j.jplph.2005.10.010

Rolling, W. R., Dorrance, A. E., & McHale, L. K. (2020). Testing methods and statistical models of genomic prediction for quantitative disease resistance to *Phytophthora sojae* in soybean [*Glycine max* (L.) Merr] germplasm collections. *Theoretical and Applied Genetics*, *133*, 3441–3454. https://doi.org/10.1007/S00122-020-03679-W

Sandhu, D., & Kaundal, A. (2018). Dynamics of salt tolerance: Molecular perspectives. In S. S. Gosal & S. H. Wani (Eds.), *Biotechnologies of crop improvement, Volume 3: Genomic approaches* (pp. 25–40). Springer International Publishing. https://doi.org/10.1007/978-3-319-94746-4_2

Sandhu, D., Pudussery, M. V., Kumar, R., Pallete, A., Markley, P., Bridges, W. C., & Sekhon, R. S. (2020). Characterization of natural genetic variation identifies multiple genes involved in salt tolerance in maize. *Functional & Integrative Genomics*, *20*, 261–275. https://doi.org/10.1007/S10142-019-00707-X

Sehgal, D., Rosyara, U., Mondal, S., Singh, R., Poland, J., & Dreisigacker, S. (2020). Incorporating genome-wide association mapping results into genomic prediction models for grain yield and yield stability in CIMMYT spring bread wheat. *Frontiers in Plant Science*, *11*, Article 197. https://doi.org/10.3389/FPLS.2020.00197

Shahzad, M., Witzel, K., Zörb, C., & Mühling, K. H. (2012). Growth-related changes in subcellular ion patterns in maize leaves (*Zea mays* L.) under salt stress. *Journal of Agronomy and Crop Science*, *198*, 46–56. https://doi.org/10.1111/J.1439-037X.2011.00487.X

Shikha, M., Kanika, A., Rao, A. R., Mallikarjuna, M. G., Gupta, H. S., & Nepolean, T. (2017). Genomic selection for drought tolerance using genome-wide SNPs in maize. *Frontiers in Plant Science*, *8*, Article 550. https://doi.org/10.3389/FPLS.2017.00550

Singh, V., & Kaundal, A. (2023). Genomic selection in maize breeding. In S. H. Wani, Z. A. Dar, & G. P. Singh (Eds.), *Maize improvement: Current advances in yield, quality, and stress tolerance under changing climatic scenarios* (pp. 99–108). Springer International Publishing.

Solberg, T. R., Sonesson, A. K., Woolliams, J. A., & Meuwissen, T. H. (2009). Reducing dimensionality for prediction of genome-wide breeding values. *Genetics, Selection, Evolution*, *41*, Article 29. https://doi.org/10.1186/1297-9686-41-29

Song, B., Buckler, E. S., Wang, H., Wu, Y., Rees, E., Kellogg, E. A., Gates, D. J., Khaipho-Burch, M., Bradbury, P. J., Ross-Ibarra, J., Hufford, M. B., & Romay, M. C. (2021). Conserved noncoding sequences provide insights into regulatory sequence and loss of gene expression in maize. *Genome Research*, *31*, 1245–1257. https://doi.org/10.1101/gr.266528.120

Tan, B., Grattapaglia, D., Martins, G. S., Ferreira, K. Z., Sundberg, B., & Ingvarsson, P. K. (2017). Evaluating the accuracy of genomic prediction of growth and wood traits in two Eucalyptus species and their F_1 hybrids. *BMC Plant Biology*, *17*, Article 110. https://doi.org/10.1186/S12870-017-1059-6

U.S. Global Change Research Program. (2017). *Climate science special report: Fourth national climate assessment* (Vol. 1). U.S. Global Change Research Program. https://doi.org/10.7930/J0J964J6

Usai, M. G., Goddard, M. E., & Hayes, B. J. (2009). LASSO with cross-validation for genomic selection. *Genetics Research*, *91*, 427–436. https://doi.org/10.1017/S0016672309990334

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*, 4414–4423. https://doi.org/10.3168/JDS.2007-0980

Veerkamp, R. F., Bouwman, A. C., Schrooten, C., & Calus, M. P. L. (2016). Genomic prediction using preselected DNA variants from a GWAS with whole-genome sequence data in Holstein–Friesian cattle. *Genetics Selection Evolution*, *48*, Article 95. https://doi.org/10.1186/S12711-016-0274-1

Vivek, B. S., Krishna, G. K., Vengadessan, V., Babu, R., Zaidi, P. H., Kha, L. Q., Mandal, S. S., Grudloyma, P., Takalkar, S., Krothapalli, K., Singh, I. S., Ocampo, E. T. M., Xingming, F., Burgueño, J., Azrai, M., Singh, R. P., & Crossa, J. (2017). Use of genomic estimated breeding values results in rapid genetic gains for drought tolerance in maize. *The Plant Genome*, *10*(1), https://doi.org/10.3835/PLANTGENOME2016.07.0070

Vojgani, E., Pook, T., Martini, J. W. R., Hölker, A. C., Mayer, M., Schön, C. C., & Simianer, H. (2021). Accounting for epistasis improves genomic prediction of phenotypes with univariate and bivariate models across environments. *Theoretical and Applied Genetics*, *134*, 2913–2930. https://doi.org/10.1007/S00122-021-03868-1

Wang, J., Zhou, Z., Zhang, Z., Li, H., Liu, D., Zhang, Q., Bradbury, P. J., Buckler, E. S., & Zhang, Z. (2018). Expanding the BLUP alphabet for genomic prediction adaptable to the genetic architectures of complex traits. *Heredity*, *121*, 648–662. https://doi.org/10.1038/s41437-018-0075-0

Wang, Q., Yu, Y., Yuan, J., Zhang, X., Huang, H., Li, F., & Xiang, J. (2017). Effects of marker density and population structure on the genomic prediction accuracy for growth trait in Pacific white shrimp *Litopenaeus vannamei*. *BMC Genetics*, *18*, 45. https://doi.org/10.1186/S12863-017-0507-5

Whittaker, J. C., Curnow, R. N., Haley, C. S., & Thompson, R. (1995). Using marker-maps in marker-assisted selection. *Genetics Research*, *66*(3), 255–265. https://doi.org/10.1017/S0016672300034698

Whittaker, J. C., Thompson, R., & Denham, M. C. (2000). Marker-assisted selection using ridge regression. *Genetics Research*, *75*(2), 249–252. https://doi.org/10.1017/S0016672399004462

Zahra, N., Raza, Z. A., & Mahmood, S. (2020). Effect of salinity stress on various growth and physiological attributes of two contrasting maize genotypes. *Brazilian Archives of Biology and Technology*, *63*, e20200072. https://doi.org/10.1590/1678-4324-2020200072

Zhang, A., Wang, H., Beyene, Y., Semagn, K., Liu, Y., Cao, S., Cui, Z., Ruan, Y., Burgueño, J., Vicente, F. S., Olsen, M., Prasanna, B. M., Crossa, J., Yu, H., & Zhang, X. (2017). Effect of trait heritability, training population size and marker density on genomic prediction accuracy estimation in 22 bi-parental tropical maize populations. *Frontiers in Plant Science*, *8*, 1916. https://doi.org/10.3389/FPLS.2017.01916

Zhang, X., Pérez-Rodríguez, P., Burgueño, J., Olsen, M., Buckler, E., Atlin, G., Prasanna, B. M., Vargas, M., Vicente, F. S., & Crossa, J. (2017). Rapid cycling genomic selection in a multiparental tropical maize population. *G3: Genes|Genomes|Genetics*, *7*, 2315–2326. https://doi.org/10.1534/G3.117.043141

Zhu, S., Guo, T., Yuan, C., Liu, J., Li, J., Han, M., Zhao, H., Wu, Y., Sun, W., Wang, X., Wang, T., Liu, J., Tiambo, C. K., Yue, Y., & Yang, B. (2021). Evaluation of Bayesian alphabet and GBLUP based on different marker density for genomic prediction in Alpine Merino sheep. *G3 Genes|Genomes|Genetics*, *11*(11), jkab206. https://doi.org/10.1093/G3JOURNAL/JKAB206

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Singh, V., Krause, M., Sandhu, D., Sekhon, R. S., & Kaundal, A. (2023). Salinity stress tolerance prediction for biomass-related traits in maize (*Zea mays* L.) using genome-wide markers. *The Plant Genome*, *16,* e20385. https://doi.org/10.1002/tpg2.20385