Crop Science

ORIGINAL ARTICLE

Crop Breeding & Genetics

Multivariate genomic selection models improve prediction accuracy of agronomic traits in soft red winter wheat

Zachary J. Winn[1,2] | **Dylan L. Larkin**[1] | **Dennis N. Lozada**[1] | **Noah DeWitt**[2,3] | **Gina Brown-Guedira**[2,4] | **Richard Esten Mason**[1]

[1]Department of Crop, Soil, and Environmental Sciences, University of Arkansas, Fayetteville, Arkansas, USA

[2]Department of Crop and Soil Sciences, North Carolina State University, Raleigh, North Carolina, USA

[3]School of Plant, Environmental, and Soil Sciences, Louisiana State University, Baton Rouge, Louisiana, USA

[4]USDA-ARS Plant Science Research Unit, Raleigh, North Carolina, USA

Correspondence
Zachary J. Winn and Richard Esten Mason, Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO 80521, USA.
Email: zachary.winn@colostate.edu; esten.mason@colostate.edu

Assigned to Associate Editor Collins Kimbeng.

Present addresses

Zachary J. Winn and R. Esten Mason, Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO 80523, USA.

Dylan L. Larkin, Limagrain Cereal Seeds, LLC, Walla Walla, WA 99362, USA.

Dennis N. Lozada, Department of Plant and Environmental Sciences, New Mexico State University, Las Cruces, NM 88003, USA; Chile Pepper Institute, New Mexico State University, Las Cruces, NM 88003, USA.

Abstract

Univariate genomic selection (UVGS) is an important tool for increasing genetic gain and multivariate GS (MVGS), where correlated traits are included in genomic selection, which can improve genomic prediction accuracy. The objectives for this study were to evaluate MVGS approaches to improve prediction accuracy for four agronomic traits using a training population of 351 soft red winter wheat (*Triticum aestivum* L.) genotypes, evaluated over six site-years in Arkansas from 2014 to 2017. Genotypes were phenotyped for grain yield, heading date, plant height, and test weight in both the training and test populations. In cross-validations, various combinations of traits in MVGS models significantly improved prediction accuracy for test weight in comparison to a UVGS model. Marginal increases in predictive accuracy were also observed for grain yield, plant height, and heading date. Multivariate models which were identified as superior to the univariate case in cross-validations were forward validated by predicting the advanced breeding nurseries of 2018 and 2020. In forward validation, consistent increases in accuracy were observed for test weight, plant height, and heading date using MVGS instead of UVGS. Overall, MVGS models improved prediction accuracies when correlated traits were included with the

predicted response. The methods outlined in this study may be used to achieve higher prediction accuracies in unbalanced datasets over multiple environments.

1 | INTRODUCTION

The global population is projected to increase from 7.7 to approximately 9.7 billion by 2050 (United Nations, 2019). While global food production has rapidly expanded over the last 70 years, demand is still expected to surpass food supply in the coming decades (Graybosch & Peterson, 2010). This trend of demand outpacing supply is likewise observed in wheat (*Triticum aestivum* L.), where the genetic gain in grain yield is 0.9% annually, far below the projected 2.4% needed to double wheat production by 2050 (Ray et al., 2013).

To improve future food security, wheat breeders are implementing new technologies to increase genetic gain. These technologies include methods such as marker-assisted selection (MAS), high-throughput phenomics, and genomic selection (GS) (Heffner et al., 2009; Larkin et al., 2019). Genomic selection is the process of using genetic relationships and historical phenotypic data to calculate genomic predictions of unobserved lines' genotypic values (Heffner et al., 2009; Meuwissen et al., 2001). The main assumption of GS is that markers are in linkage disequilibrium (LD) with quantitative trait loci (QTL) associated with a trait of interest, an assumption that allows for the estimation of phenotypic values from genetic information (Meuwissen et al., 2001). If implemented properly, GS may increase genetic gain per cycle compared to phenotypic selection (Asoro et al., 2013; Bernardo & Yu, 2007; Heffner et al., 2009; Rutkoski et al., 2015).

There are many methods used to optimize the prediction ability of GS models (Jannink et al., 2010; Larkin et al., 2019). These methods have largely focused on training population optimization to increase prediction accuracy in the validation population (Akdemir et al., 2015; Combs & Bernardo, 2013; Habier et al., 2007; Isidro et al., 2015; Jannink et al., 2010). Model selection also plays a role in accuracy, whether it is through the use of ridge-regression or Bayesian parametric models, or through the use of semi- or non-parametric models (Heslot et al., 2012).

An alternative to univariate GS (UVGS), where only one response is evaluated and predicted at a time, is multivariate GS (MVGS), which is a method that includes multiple variables in the response of a GS model (Covarrubias-Pazaran et al., 2018; Watson et al., 2019). Prediction accuracies for low-heritability traits of interest can also be improved when high-heritability secondary traits are used as covariates (Calus & Veerkamp, 2011; Guo et al., 2014; Jia & Jannink, 2012). Grain yield has been the primary trait of interest in studies focused on MVGS for agronomic traits in wheat, with normalized difference vegetative index (NDVI) and canopy temperature having been used as secondary traits in previous experiments (Crain et al., 2018; Lozada & Carter, 2019; Rutkoski et al., 2016; Sun et al., 2017). Wheat breeders record several traits during the growing season, all of which can impact grain yield (Chen et al., 2012; Liu et al., 2015; Tshikunde et al., 2019). These traits can be incorporated into MVGS models, along with grain yield, to improve prediction accuracies.

In the current study we sought to accomplish three main objectives: (i) to evaluate a training population of soft red winter wheat (SRWW) genotypes for grain yield, test weight, plant height, and heading date, (ii) perform a cross-validation analysis to compare prediction accuracies between UVGS models and MVGS models for all four agronomic traits; and (iii) assess improvements in prediction accuracy in MVGS models as compared to UVGS in forward validation of new germplasm.

2 | MATERIALS AND METHODS

2.1 | Germplasm

A population of 351 diverse, inbred SRWW genotypes was utilized in this study as the training population. The population consisted of genotypes developed at the University of Arkansas (237), University of Georgia (39), North Carolina State University (38), and Louisiana State University (37). These genotypes were selected during the advanced stage of their respective breeding programs and is representative of the source germplasm of the University of Arkansas' Winter Wheat Breeding Program.

The University of Arkansas 2018 and 2020 advanced nurseries were used for forward validation of models indicated as superior through cross-validation. The 2018 and 2020 advanced breeding nursery consisted of 123 and 125 unique lines, respectively. The total dataset consisted of 284 unique lines across growing seasons. The germplasm of the University of Arkansas breeding program is related to the training population used in this study and is representative of lines which are well adapted to the eastern region of Arkansas.

2.2 | Experimental design and trait measurements

Winter wheat is planted in the fall and harvested during the late spring in the Southern region of the United States; therefore, each growing season spans 2 years. For the training

population, a subset of the total population consisting of 175 entries were evaluated at the Lon Mann Cotton Research Center near Marianna, AR, USA (MAR) and the Northeast Research and Extension Center near Keiser, AR, USA (KAR) during the 2013–2014 growing season (M14 and K14). The remaining 176 entries were evaluated in MAR during the 2014–2015 growing season (M15). All 351 entries of the population were evaluated in MAR during the 2015–2016 growing season (M16) and at both locations during the 2016–2017 growing season (M17 and K17).

The M14 and M15 environments were planted in an unbalanced randomized block design with two replications and repeated checks. In M14, the first replication consisted of 175 genotypes, while the second replication consisted of 117 genotypes from the first replication; whereas in M15, the first replication consisted of 177 genotypes, while the second replication consisted of 120 genotypes from the first replication. The K14, M16, K17, and M17 environments were planted in an unreplicated augmented design with repeated checks.

For the forward validation population, the 125 lines in the 2018 advanced breeding nursery were planted at KAR; MAR; Newport, AR (NAR); and Rohwer, AR (RAR) in two replications of a randomized complete block design during the 2017–2018 growing season. The 125 lines in the 2020 advanced breeding nursery were planted at KAR, MAR, and RAR in two replications of a randomized complete block design during the 2019–2020 growing season.

In all growing seasons regardless of nursery or population, lines were drill seeded at a rate of 118 kg ha^{-1} to establish seven row plots that were 6 m long and 1 m wide with 18 cm spacing between rows. Plots were end-trimmed in the spring to leave a harvestable area of approximately 7 m^2. Plots were managed according to the recommendations for wheat in Arkansas (Kelley, 2018). Each year, plots in MAR received two applications of urea fertilizer at 101 and 67 kg ha^{-1}, along with 27 kg ha^{-1} of ammonium sulfate fertilizer. Plots in KAR received two applications of 78 kg ha^{-1} of urea fertilizer per year. A combination of herbicides, including Axial XL (Syngenta AG), Finesse (DuPont de Nemours, Inc.), Harmony Extra (DuPont de Nemours, Inc.), Osprey (Bayer AG), and Prowl H$_2$O (BASF SE) were used each year for weed control. Fungicides were also applied as needed.

Phenotypic data were collected for four agronomic traits: grain yield, test weight, plant height, and heading date. Individual plots were harvested with a 1994 Hege 140 plot combine (Hege Maschinen GmbH, Baden-Württemberg, Germany) and grain yield was calculated in all environments from the weight of seed from each plot as measured by the HarvestMaster Pro 4100 (Juniper Systems, Inc., UT, USA), expressed in t ha^{-1} (tons per hectare) and adjusted for 13% moisture

Core Ideas

- Multivariate genomic selection models improve prediction accuracy for agronomic traits.
- Multivariate genomic selection models may assist in making genomic predictions using highly unbalanced datasets.
- Stronger correlations between predicted traits and covariates improves genomic prediction accuracy.

content. Test weight was sampled in all environments using a HarvestMaster Pro 4100 and was expressed in kg hL^{-1} (kilograms per hectoliter), adjusted for 13% moisture content.

Heading date for each genotype was collected during all growing seasons in MAR for the training population. In the advanced nursery in 2018, heading date was taken in MAR only. In the advanced nursery in 2020, heading date was taken in MAR and KAR. Heading date was recorded as the day of year after January 1, when 50% of the heads had emerged from the boot. Heading notes were recorded every other day from the onset of heading, continuing until all plots in the nursery had headed.

For the training population, plant height was recorded in cm from the surface of the soil to the tip of the head at physiological maturity in all environments except for K17, excluding awns when present. For the advanced nursery in 2018, plant height was taken in KAR, MAR, and NAR. For the advanced nursery in 2020, plant height was taken in only MAR.

2.3 | Phenotypic data analysis

Phenotypic data was analyzed via the ASReml-R package in R statistical software version 4.2.2 (Butler et al., 2009; R Core Team, 2013). Best linear unbiased estimates (BLUEs) were obtained for all four traits from each environment in the training data via the following model:

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

where y_{ij} is the observed phenotype; μ is the overall mean; τ_i is the fixed effect of the ith genotype; β_j is the random effect of replication for the incomplete randomized block design or sub-block for the augmented design; and ε_{ij} is the residual error term, where $\varepsilon_{ij} \sim N(0, \mathbf{I}\sigma_\varepsilon^2)$, where $\mathbf{I}$ is an identity matrix and σ_ε^2 is the residual error variance.

BLUEs from each environment in the training data were curated into a new dataset where each BLUE, if estimable within the environment, was listed for each genotype in each environment. Multi-environmental BLUEs were calculated

for each trait using the following univariate model:

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

where y_{ij} is the observed phenotype; μ is the overall mean; τ_i is the fixed effect of the ith genotype; β_j is the random effect of replication for the incomplete randomized block design or sub-block for the augmented design; and ε_{ij} is the residual error term, where $\varepsilon_{ij} \sim N(0, \mathbf{I}\sigma_\varepsilon^2)$, where $\mathbf{I}$ is an identity matrix and σ_ε^2 is the residual error variance. Univariate multi-environmental BLUEs of the training population were used to estimate prediction accuracy (r) through correlation with predictions from both UVGS and MVGS models. Fixed effect significance was estimated using a Wald test, and a p-value derived from a chi-squared distribution was used to delineate significance.

For the advanced nurseries in both 2018 and 2020, two separate models were used to estimate BLUEs within a year and across years. To calculate BLUEs across environments within a year or across years and environments, the following mixed linear model was employed:

$$y_{ij} = \mu + \tau_i + \beta\gamma_{jk} + \varepsilon_{ijk}$$

where y_{ij} is the observed phenotype; μ is the overall mean; τ_i is the fixed effect of the ith genotype; $\beta\gamma_{jk}$ is the random environment by replication effect; and ε_{ijk} is the residual error term, where $\varepsilon_{ijk} \sim N(0, \mathbf{I}\sigma_\varepsilon^2)$, where $\mathbf{I}$ is an identity matrix and σ_ε^2 is the residual error variance. An "environment" in the case of the advanced nurseries corresponded to the site-year combination.

Heritability and genetic correlation of each trait in the training population was estimated using a multivariate genomic best linear unbiased prediction (MVgBLUP) model:

$$\begin{bmatrix} y_{11} & \cdots & y_{1m} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nm} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_1 & 0 \\ \vdots & \vdots \\ 0 & \mathbf{B}_m \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_m \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_1 & 0 \\ \vdots & \vdots \\ 0 & \mathbf{Z}_m \end{bmatrix} \begin{bmatrix} \mathbf{g}_1 \\ \vdots \\ \mathbf{e}_m \end{bmatrix} + \varepsilon$$

where the left side is a matrix of n observations by m traits, $\mathbf{B}_m$ is a design matrix relating fixed effects to observations for each trait m, $\boldsymbol{\mu}_m$ is a vector of fixed effects for each trait, $\mathbf{Z}_m$ is a design matrix relating observations to the random effects for each trait m, $\mathbf{g}_1$ to $\mathbf{e}_m$ are the vectors of random genetic and environment effects for each trait m, and ε is the residual error where $\mathbf{I}_n \otimes \sum(\mathbf{n}_m \times \mathbf{n}_m)$. The matrix $\sum(\mathbf{n}_m \times \mathbf{n}_m)$ is an unstructured residual matrix of m traits. The distribution of the random genotypic effect is defined (when using all four traits) as:

$$\begin{bmatrix} \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_4 \end{bmatrix} \sim N(0, \mathbf{G} \otimes \mathbf{T})$$

$$\mathbf{T} = \begin{bmatrix} \sigma_{g11}^2 & \cdots & \sigma_{g14}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{g14}^2 & \cdots & \sigma_{g44}^2 \end{bmatrix}$$

where $\mathbf{G}$ is the realized genomic relationship matrix and $\mathbf{T}$ is the variance–covariance matrix of random genetic effects for each trait; $\mathbf{G} \otimes \mathbf{T}$ is the Kronecker product of $\mathbf{G}$ and $\mathbf{T}$. In practice, this relates estimates of line effects for each trait to the calculated genetic relationship with other lines within each trait. Furthermore, this states that the genetic relatedness and covariance of line effects between those traits are estimated in the model. Genetic correlation among traits was calculated as:

$$r_g = \frac{\sigma_{y1,y2}^2}{\sigma_{y1}^2 \sigma_{y2}^2}$$

where $\sigma_{y1,y2}^2$ is the covariance between two traits, σ_{y1}^2 is the variance of the first trait, and σ_{y2}^2 is the variance of the second trait. Per-plot, narrow-sense, genomic heritability was calculated as follows:

$$h_{\text{gBLUP}}^2 = \frac{\sigma_{g(y)}^2}{\sigma_{g(y)}^2 + \sigma_{\varepsilon(y)}^2}$$

where $\sigma_{g(y)}^2$ is the genetic variance estimated with $\mathbf{G}$ and $\sigma_{\varepsilon(y)}^2$ is the residual error associated with that trait.

2.4 | Genotyping

The 351 genotypes in the training population and the 284 genotypes in the advanced nurseries were genotyped using genotyping-by-sequencing (GBS) (Poland et al., 2012). DNA was extracted using the Mag-Bind Plant DNA Plus kit (Omega Bio-tek, GA, USA), following the manufacturer's instructions. Genomic DNA was quantified using the Quant-iT PicoGreen dsDNA Assay Kit and normalized to 20 ng μL^{-1} (Thermo Fisher Scientific, MA, USA). GBS libraries were created using *Pst1-Msp1* restriction enzyme combinations (Poland et al., 2012). The samples were pooled together at a 192-plex and each pooled library was sequenced on a single lane of an Illumina Hi-Seq 2500 system (Illumina Inc., CA, USA).

Single nucleotide polymorphism (SNP) calling was performed using the TASSEL 5.0 GBSv2 pipeline using 64 base kmer length and a minimum kmer count of five. Reads were aligned to the International Wheat Genome Sequencing Consortium (IWGSC) RefSeq v1.0 "Chinese Spring" wheat reference sequence (Appels et al., 2018) using the Burrows–Wheeler aligner version 0.7.10 (Li & Durbin, 2009). Raw SNP data generated from the TASSEL pipeline were filtered

to remove genotypes with more than 50% missing data and heterozygosity greater than 30%.

Markers were filtered to select for biallelic SNPs with minor allelic frequency of greater than 5%, less than 10% missing data, and heterozygosity of less than or equal to 10%, and all unaligned SNPs were dropped from the dataset prior to imputation. Post filtration, missing data was imputed using the Beagle algorithm (Browning et al., 2018). A principal component analysis (PCA) of the training population was performed using the "prcomp()" function in base R (R Core Team, 2013).

2.5 | Univariate and multivariate genomic selection

UVGS was performed using ASREML-R using the following UVgBLUP model:

$$y = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \varepsilon_i$$

where $\mathbf{u}$ is the vector of genotype effects, which is assumed to have a normal distribution $\mathbf{u} \sim N(0, \mathbf{G}\sigma_u^2)$, where $\mathbf{G}$ is the genomic relationship matrix and σ_u^2 is the variance of the individual genotype effects; $\boldsymbol{\beta}$ is the vector of fixed effects; $\mathbf{X}$ is the design matrix of fixed effects; $\mathbf{Z}$ is the design matrix relating genotypes to phenotypic observations (y), with m markers in columns and n observed genotypes in rows; and ε is the residual error assumed to have a normal distribution $\varepsilon_i \sim N(0, \mathbf{I}\sigma_\varepsilon^2)$, where $\mathbf{I}$ is the identity matrix and σ_ε^2 is the residual error variance. The prediction is the sum of all additive allele effects of a genotype (Chen & Zhang, 2018; Endelman, 2011; VanRaden, 2008).

MVGS was performed using ASReml-R using the same formula used in estimating heritability and genetic correlation. Estimation of prediction accuracies within the training population was performed via cross-validation. Data in the training population were randomly subsetted into a training population (80% of observations) and a validation population (20% of observations), where predictions were made for the selected trait using the training population, and a correlation was made between the predicted value and the calculated BLUEs for the validation population. This process was repeated 100 times to obtain the maximum, minimum, mean, and standard deviation of prediction accuracy for each model. Prediction accuracy was estimated for each trait's UVgBLUP and all combinations of the four traits in MVgBLUP models.

In the case of MVgBLUP, observations in the validation population were only hidden for the predicted trait of interest. For instance, if a multivariate model were conducted to predict grain yield with both grain yield and test weight as responses in the model, only observations in the validation population for grain yield would be hidden. In this scenario,

it is assumed that test weight and not yield data would be available for those individuals. All predictions were correlated with univariate multi-environmental BLUEs to calculate prediction accuracy (r).

To estimate significant differences between the univariate prediction and a multivariate model prediction cross-validated accuracy, the following Z-score was calculated:

$$\hat{Z}_i = \frac{\bar{x}_i - \bar{x}_{\text{univariate}}}{s_{\text{univariate}}}$$

where $\hat{Z}_i$ is the estimated Z-score of training population i, $\bar{x}_i$ is the sample mean of the distribution for multivariate model i, $\bar{x}_{\text{univariate}}$ is the sample mean of the univariate model for a trait, and $s_{\text{univariate}}$ is the sample standard deviation of univariate model. Values where $-1.65 < \hat{Z}_i > 1.65$ were considered significantly different at $\alpha = 0.05$.

Forward validation conducted on the advanced nurseries was performed by comparing univariate BLUEs calculated from observed data to GEBVs produced by UVGS and MVGS models. The adjusted mean either within a year or across years was used as the BLUE and either the UVGS or MVGS prediction was used as the GEBV. Forward prediction accuracies were estimated as the Pearson's correlation coefficient (r) between the BLUE and GEBV for the corresponding year or years. Just as in the cross-validation scheme, only the predicted response was removed from the dataset prior to calculation, meaning that secondary traits were left in MVGS models. This implies that the data for those corresponding traits would be available prior to prediction of the response.

3 | RESULTS

3.1 | Genotypic data

After filtration parameters were applied, 55,402 genome-wide distributed, high-quality SNPs remained. Visualization of the first two PCs of the filtered marker information for the training population identified two distinct clusters of lines (Supporting Information 1). Lines from all breeding programs appeared in both clusters, although there appeared to be some sub-clustering of lines from individual breeding programs within the two main clusters. Such clustering has also been observed in other studies using SRWW populations adapted to the Southeastern United States (Benson et al., 2012; Larkin et al., 2020; Sarinelli et al., 2019). This apparent population structure is attributed to the presence or absence of a large introgression from *Triticum timopheevii* Zhuk. possessing stem rust (*Puccinia graminis* f. sp. *tritici*) and powdery mildew (*Blumeria graminis* f. sp. *tritici*) resistance genes, *Sr36* and *Pm6*, respectively, into wheat chromosome 2B (Nyquist, 1962). This alien introgression is located on

TABLE 1   Table of descriptive statistics.

Trait	Environments[a]	Descriptive statistics				Heritability	
		Mean	Min	Max	SD	h^2_{gBLUP}	SE
Grain yield (t ha^{-1})	K14, K17, M14, M15, M16, M17	3.9	0.1	9.5	1.1	0.19	0.03
Test weight (Kg hL^{-1})	K14, K17, M14, M15, M16, M17	70.4	44.3	82.5	3.8	0.29	0.03
Plant height (cm)	K14, M14, M15, M16, M17	88.9	56.5	120.9	9.8	0.46	0.03
Heading date (days after January 1)	M14, M15, M16, M17	102.5	79.1	128.0	14.1	0.62	0.03

[a]Environments where data for each trait were collected: K14, Keiser 2014; K17, Keiser 2017; M14, Marianna 2014; M15, Marianna 2015; M16, Marianna 2016; M17, Marianna 2017.

chromosome 2B (427 Mb), leading to identification of many polymorphic, tightly linked markers (Walkowiak et al., 2020).

Supporting Information 1. A scatter plot of the first two principal components (PCs) derived from a principal component analysis (PCA) on the filtered genome wide genotyping-by-sequencing (GBS) dataset of the training population showing two distinct clusters. On the x axis is the first principal component (PC1) and on the y axis is the second principal component (PC2). Lines are labeled on the right by shape and color according to which wheat breeding program they originate from; Arkansas (AR), Georgia (GA), Louisiana (LA), and North Carolina (NC).

3.2 | Phenotypic data—Training population

All four agronomic traits had significant differences among genotypes (P (X^2) < 0.05) in each environment. Univariate, single environment BLUEs of grain yield had a range between 0.1 and 9.5 t ha^{-1} across all six environments. Single environment BLUEs for heading date ranged between 79 and 128 days after January 1 across all locations. Univariate, single environment BLUEs across all locations ranged between 56.5 and 120.9 cm for plant height. Test weight for the univariate, single environment BLUEs across all environments ranged between 44.3 and 82.5 kg hL^{-1} across all environments (Table 1).

Significant phenotypic correlations were observed between trait BLUEs derived from the univariate, multi-environmental models (Figure 1). Genetic correlations estimated using multivariate, multi-environmental models ranged from relatively low ($r = -0.01$ for grain yield and heading date) to moderate ($r = 0.53$ for grain yield and test weight) (Figure 2). Traits that had strong phenotypic correlations tended to have strong genetic correlations. The highest per-plot, narrow-sense, genomic heritability among the four agronomic traits was heading date ($h^2_{gBLUP} = 0.62$), followed by plant height

($h^2_{gBLUP} = 0.46$), test weight ($h^2_{gBLUP} = 0.29$), and grain yield ($h^2_{gBLUP} = 0.19$).

3.3 | Univariate versus multivariate genomic selection cross-validated accuracies

Cross-validation accuracy for all combinations of MVgBLUP and UVgBLUP were summarized across all permutations and significant differences (derived by Z-score) in the MVgBLUP from the UVgBLUP were estimated (Table 2). Distributions for each trait UVgBLUP and MVgBLUP results were visualized (Figure 3a–d).

Cross-validated distributions for grain yield compared to all different combinations of MVgBLUP indicated marginal differences of the MVgBLUP cross-validation distribution and the UVgBLUP distribution (Figure 3a, Table 2). The mean of the distribution of prediction accuracies for the UVgBLUP of grain yield was $\bar{r} = 0.53$. An increase in the mean prediction accuracy was observed for the model incorporating test weight to predict grain yield ($\bar{r} = 0.61$). The MVgBLUP models including grain yield, test weight, and heading date produced a marginal improvement in prediction accuracy ($\bar{r} = 0.61$). The combination of grain yield, test weight, and plant height also produced a higher mean than the UVgBLUP for grain yield alone ($\bar{r} = 0.61$). Furthermore, the full MVgBLUP model of all agronomic traits (grain yield, test weight, plant height, and heading date) produced a higher mean accuracy for grain yield than the UVgBLUP model ($\bar{r} = 0.61$).

Distributions of prediction accuracies for test weight compared to all different MVgBLUP combinations indicated significant differences ($P(Z) < 0.05$) of the MVgBLUP distributions compared to the UVgBLUP distribution (Figure 3b, Table 2). The mean prediction accuracy for the UVgBLUP of test weight was $\bar{r} = 0.32$. The combination of test weight, grain yield, and plant height produced a
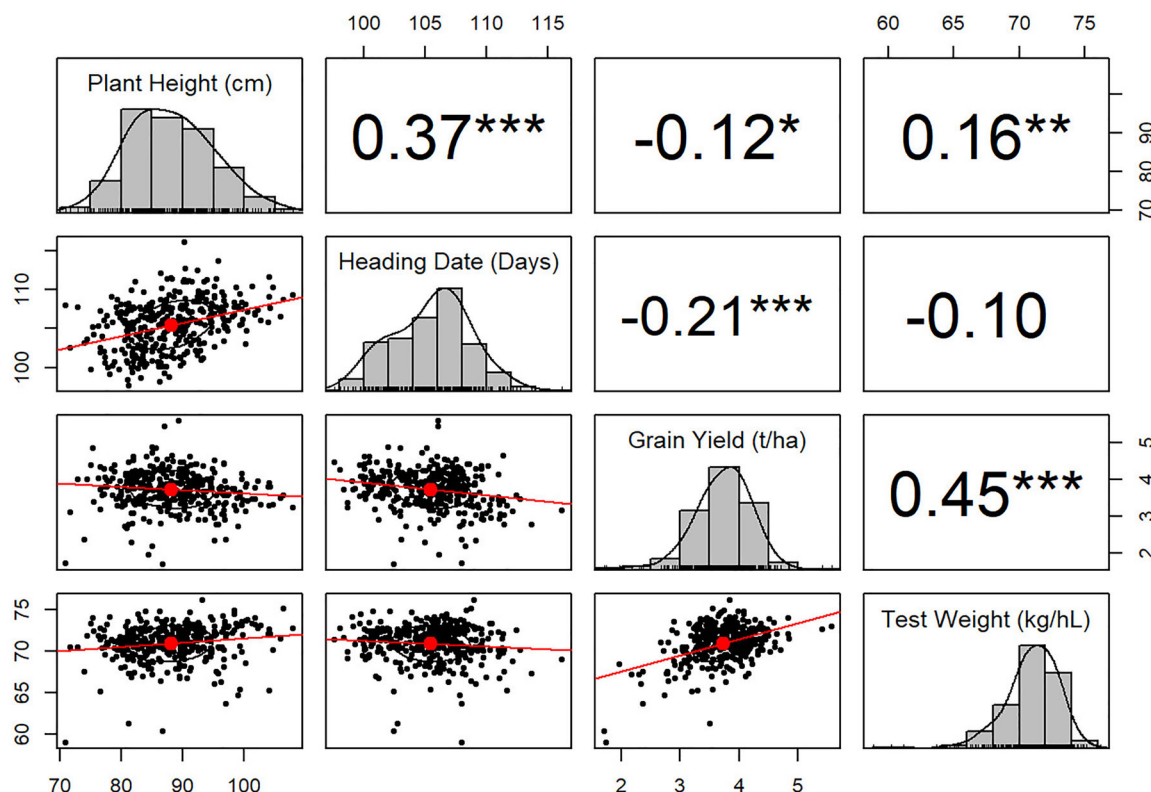
FIGURE 1 Pairs plot of univariate, multi-environmental best linear unbiased estimates of agronomic traits. The diagonal of the matrix displays a density plot and histogram of each trait with their title displayed. The top right triangle of the matrix displays the Pearson's correlation coefficient (R) and the significance of the correlation for corresponding traits where: *Significant at the 0.05 probability level. **Significant at the 0.01 probability level. ***Significant at the 0.001 probability level. The bottom left triangle of the matrix displays a scatterplot of two corresponding traits with a linear line of best fit and centroid labeled by the large red point at the center.

significantly higher mean prediction accuracy than the UVgBLUP for test weight ($\bar{r} = 0.52$). The full MVgBLUP model of all agronomic traits (grain yield, test weight, plant height, and heading date) produced a significantly higher mean prediction accuracy for grain yield than the UVgBLUP model ($\bar{r} = 0.54$).

Unlike test weight, prediction accuracies for plant height were not significantly improved ($P(Z) > 0.05$) when correlated traits were included in a MVgBLUP model (Figure 3c, Table 2). The UVgBLUP mean prediction accuracy for plant height was $\bar{r} = 0.51$. The MVgBLUP combination of plant height and test weight produced an insignificant increase to prediction mean accuracy ($\bar{r} = 0.55$). The MVgBLUP combination of plant height, test weight, and heading date produced an insignificant improvement to prediction mean accuracy ($\bar{r} = 0.58$). The full model MVgBLUP including all agronomic traits also produced a marginal increase in prediction accuracy ($\bar{r} = 0.57$). Including heading date in an MVgBLUP model for prediction of plant height resulted in a significantly lower mean prediction accuracy than the univariate case ($\bar{r} = 0.33$).

Heading date prediction accuracies were only marginally improved through the inclusion of covariates in MVgBLUP

models (Figure 3d, Table 2). The UVgBLUP mean cross-validated accuracy for heading date was $\bar{r} = 0.59$. Models which included plant height resulted in a marginal increase in mean prediction accuracy ($\bar{r} = 0.63$). Furthermore, models which included test weight and plant height increased predictive accuracies ($\bar{r} = 0.64$). The full model MVgBLUP including all agronomic traits (grain yield, test weight, plant height, and heading date) produced a boost in the mean prediction accuracy ($\bar{r} = 0.64$).

3.4 | Comparison of training and test population phenotypic data

For each year and across years, the genotype effect was significant ($P(X^2) < 0.05$). Heritability for each trait within and across years were comparable to those of the training population (Table 3). Heritability values for grain yield ranged from 0.12 in the 2018 advanced to 0.21 in the 2020 advanced. Heritability for test weight ranged from 0.04 in the combined analysis to 0.16 in the 2020 advanced. Plant height was the most heritable trait with heritability values ranging from 0.45 in the 2018 advanced to 0.64 in the 2020 advanced. Heading
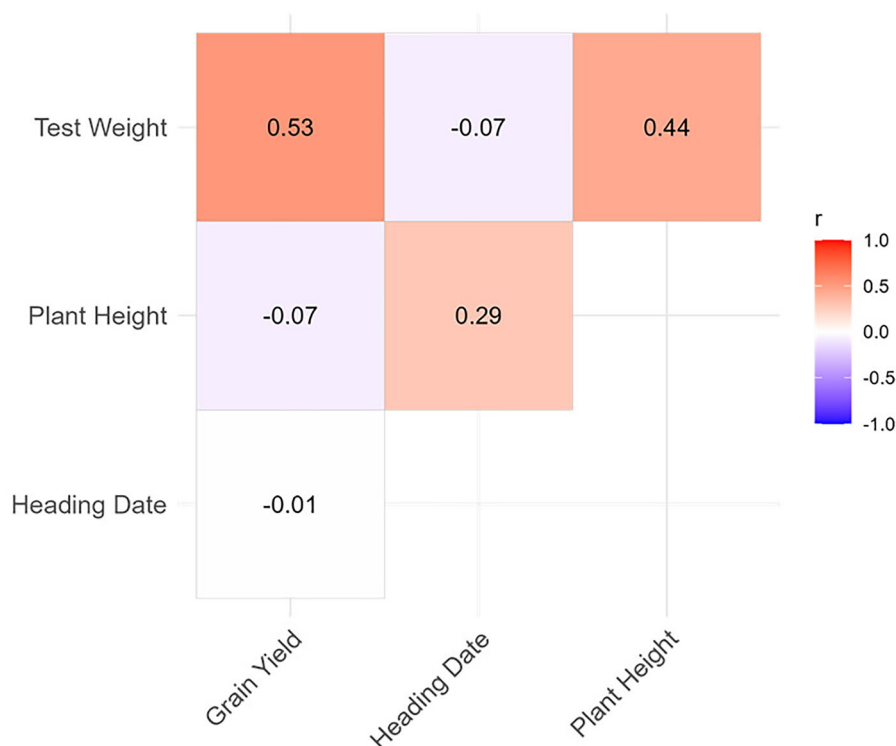
FIGURE 2 Correlation matrix of genetic correlations estimated by a multi-environmental, multivariate model. The names of corresponding agronomic traits are listed on the *y* and *x* axes. Each square in the matrix displays the correlation coefficient and is color coded by the sign and magnitude of the coefficient.

date had the lowest heritability in the 2018 advanced nursery (0.32) and highest in the 2020 advanced nursery (0.50).

Pearson's correlations among BLUEs within years and across years, unlike heritability values, departed from the training population (Figure 4). Where in the training population there were moderate correlations between grain yield and test weight ($\bar{r} = 0.45$), the correlations in the advanced 2018, 2020, and combined analysis were 0.1, −0.03, and 0.08; respectively. Likewise, the correlation between grain yield and heading date in the training population was lower (0.16) than in the advanced 2018 ($\bar{r} = 0.51$), 2020 ($\bar{r} = 0.47$), and combined ($\bar{r} = 0.46$) dataset. This may imply that the environmental conditions were not similar in the 2018 and 2020 test populations as in the training population (2014–2017).

Similar to the Pearson's correlation of the BLUEs across environments, the genetic correlation among traits within and across years in the advanced noticeably departed from the training population (Figure 5). In the training population, strong, positive genetic correlations were seen between test weight and grain yield (0.53; Figure 2). However, this was not reflected in the advanced nurseries where genetic correlations between grain yield and test weight were near zero in the 2018 nursery, −0.39 in the 2020 nursery, and −0.56 in the combined analysis.

Furthermore, in the training population, the genetic correlation between heading date and grain yield was low (−0.01;

Figure 2). However, in the advanced 2018, 2020, and combined dataset, the genetic correlation between grain yield and heading date were 0.67, 0.70, and 0.82; respectively. This implies that later heading lines tended to outperform earlier heading lines in the advanced data than in the training data. Moreover, this may imply that environmental factors, such as a late spring freeze, affected the advanced nurseries in the years 2018 and 2020.

3.5 | Forward validated accuracies in the advanced breeding nurseries

Multivariate models identified in cross-validation as either producing significant or marginal increases in prediction accuracy were compared to the univariate case in forward prediction of the 2018, 2020, and combined university of Arkansas advanced nurseries (Figure 6). For grain yield (Figure 6a), predictions were relatively non-informative in the 2018 advanced as the majority of models produced a slightly negative prediction accuracy. In 2020, prediction accuracy was highest in the univariate model for grain yield ($r = 0.17$). The combined analysis univariate accuracy was $r = 0.14$ and the multivariate model which incorporated all agronomic traits (grain yield, test weight, plant height, and heading date) had an accuracy of $r = 0.16$.

TABLE 2 Minimum, mean, maximum, standard deviation, Z-score, and significance of accuracies for all models utilized in 100 permutations of cross-validation for grain yield, test weight, plant height, and heading date.

Model[a]	Min[b]	Mean[c]	Max[d]	SD[e]	Z-score[f]	Significance[g]
Univariate GY	0.36	0.53	0.70	0.07	–	–
GY + TW	0.44	0.61	0.76	0.07	1.07	–
GY + PH	0.36	0.52	0.69	0.07	−0.12	–
GY + HD	0.35	0.53	0.69	0.07	0.01	–
GY + TW + PH	0.45	0.61	0.76	0.07	1.07	–
GY + TW + HD	0.42	0.61	0.80	0.07	1.16	–
GY + PH + HD	0.37	0.51	0.68	0.07	−0.26	–
GY + TW + PH + HD	0.44	0.61	0.75	0.06	1.16	–
Univariate TW	−0.01	0.32	0.60	0.11	–	–
TW + GY	0.29	0.45	0.65	0.08	1.12	–
TW + PH	0.15	0.40	0.62	0.09	0.65	–
TW + HD	0.08	0.33	0.56	0.10	0.04	–
TW + GY + PH	0.30	0.52	0.72	0.08	1.73	*
TW + GY + HD	0.26	0.44	0.66	0.08	1.08	–
TW + PH + HD	0.15	0.41	0.63	0.09	0.75	–
TW + GY + PH + HD	0.31	0.54	0.69	0.08	1.94	*
Univariate PH	0.30	0.51	0.69	0.08	–	–
PH + GY	0.30	0.50	0.68	0.08	−0.15	–
PH + TW	0.35	0.55	0.75	0.08	0.44	–
PH + HD	0.00	0.33	0.60	0.11	-2.35	*
PH + GY + TW	0.37	0.54	0.73	0.08	0.36	–
PH + GY + HD	0.35	0.54	0.73	0.07	0.34	–
PH + TW + HD	0.33	0.58	0.79	0.08	0.92	–
PH + GY + TW + HD	0.38	0.57	0.72	0.07	0.80	–
Univariate HD	0.39	0.59	0.75	0.07	–	–
HD + GY	0.40	0.60	0.75	0.07	0.10	–
HD + TW	0.45	0.59	0.72	0.06	0.02	–
HD + PH	0.41	0.63	0.79	0.07	0.50	–
HD + GY + TW	0.42	0.61	0.74	0.08	0.23	–
HD + GY + PH	0.45	0.62	0.78	0.07	0.39	–
HD + TW + PH	0.42	0.64	0.78	0.07	0.61	–
HD + GY + TW + PH	0.49	0.64	0.80	0.07	0.65	–

[a]Trait abbreviations are as follows: grain yield (GY), test weight (TW), plant height (PH), and heading date (HD). Models noted as "Univariate" are distributions of cross-validation accuracies for the univariate case of the following trait. All other combinations, signified by "+" are predicting the first listed trait while using the following traits in the multivariate model.

[b]Minimum.

[c]Maximum.

[d]Standard deviation.

[e]Z-scores are expressed as deviation away from the expected mean of the univariate case for each trait.

[f]Significance is denoted as such: "-" denotes P(Z) ≥ 0.95 and "*" denotes P(Z) < 0.05.

[g]Significance derived by Z-score. " *"<0.05 and "-">0.05

For test weight (Figure 6b), forward validated accuracies in 2018 ranged from −0.13 to 0.10. The models which achieved the highest forward validated accuracies were the multivariate models that incorporated grain yield and plant height as well as grain yield, plant height, and heading date ($r = 0.10$). In 2020, the highest achieved accuracy was $r = 0.18$ in the mul-

tivariate model which incorporated all agronomic traits. In the combined analysis, the highest achieved accuracy ($r = 0.17$) was found for the multivariate model which incorporated all traits.

For plant height (Figure 6c), forward validated accuracies in 2018 ranged from 0.09 to 0.43. The highest achieved
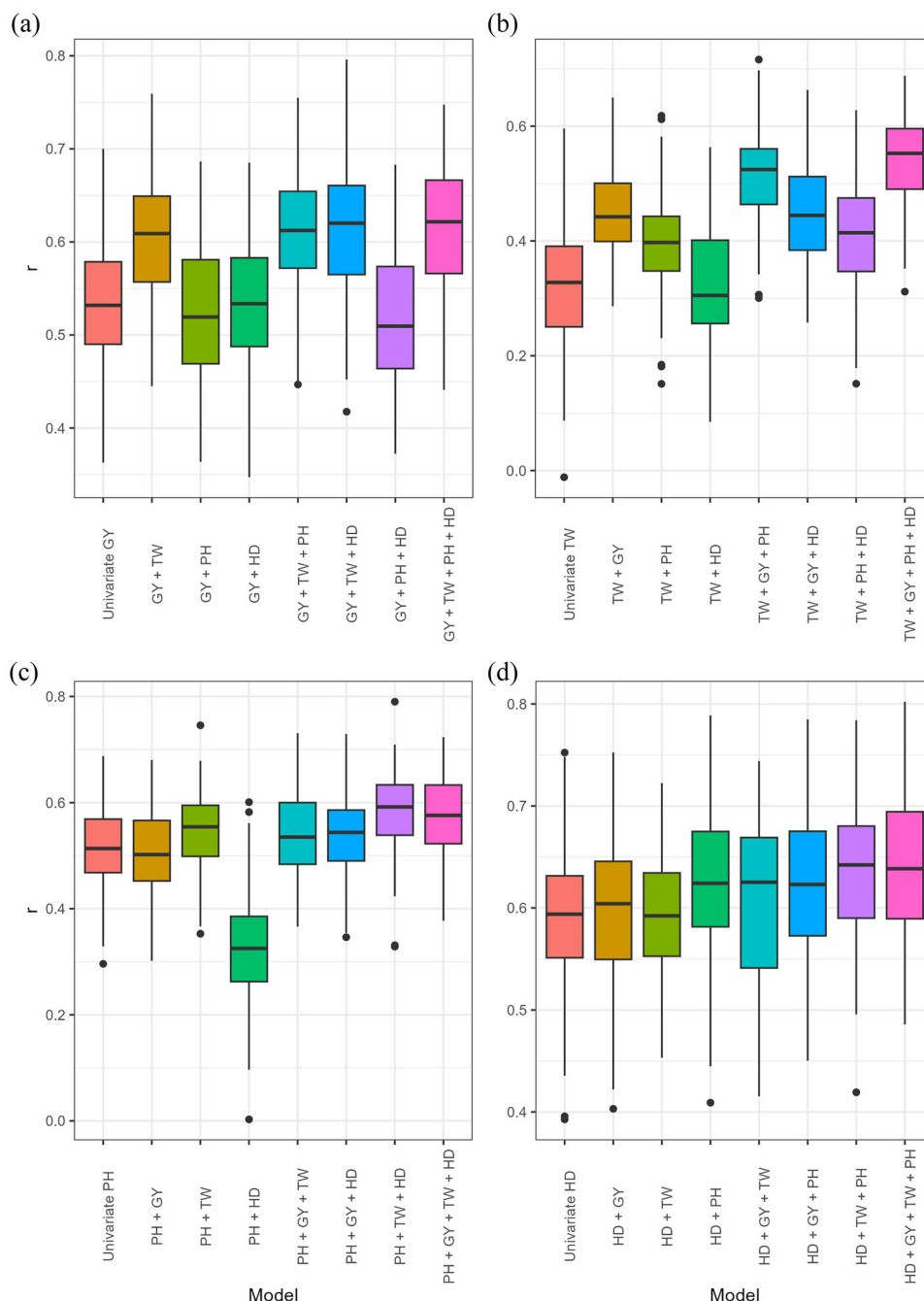
FIGURE 3 Box-and-whisker plots of cross-validation accuracies for (a) grain yield (GY), (b) test weight (TW), (c) plant height (PH), and (d) heading date (HD) univariate and multivariate genomic best linear unbiased prediction models. In all panels, displayed on the y axis is the Pearson's correlation coefficient (r) between the multi-environmental univariate best linear unbiased estimated value of the lines in the validation population and the genomic estimated breeding value of the validation population derived from labeled univariate or multivariate genomic best linear unbiased prediction models. Displayed on the x axis is the model to which the above boxplot belongs. Models labeled "univariate" display distributions derived from correlations on univariate genomic estimated breeding values. All other models featuring "+" symbols and multiple short-hand trait names feature the multivariate model cross-validation distributions where those listed traits were used in the model.

accuracy ($r = 0.43$) was found in the multivariate model which incorporated all agronomic traits. In 2020, the highest performing model ($r = 0.28$) was the univariate model. In combined analysis, the highest performing model was the multivariate model which incorporated all agronomic traits, which produced an accuracy of $r = 0.35$.

For heading date (Figure 6d), the highest performing models in 2018 were both the multivariate models which

TABLE 3 Heritability of the University of Arkansas' advanced nurseries for 2018, 2020, and the combined dataset.

Trait	Year	h^2_{gBLUP}	SE[a]
Grain yield (t ha^{-1})	2018	0.12	0.03
	2020	0.21	0.04
	Combined	0.15	0.03
Test weight (Kg hL^{-1})	2018	0.12	0.03
	2020	0.16	0.03
	Combined	0.04	0.02
Plant height (cm)	2018	0.45	0.04
	2020	0.64	0.05
	Combined	0.51	0.05
Heading date (days after January 1)	2018	0.32	0.10
	2020	0.50	0.06
	Combined	0.35	0.08

[a]Standard error of heritability estimate.
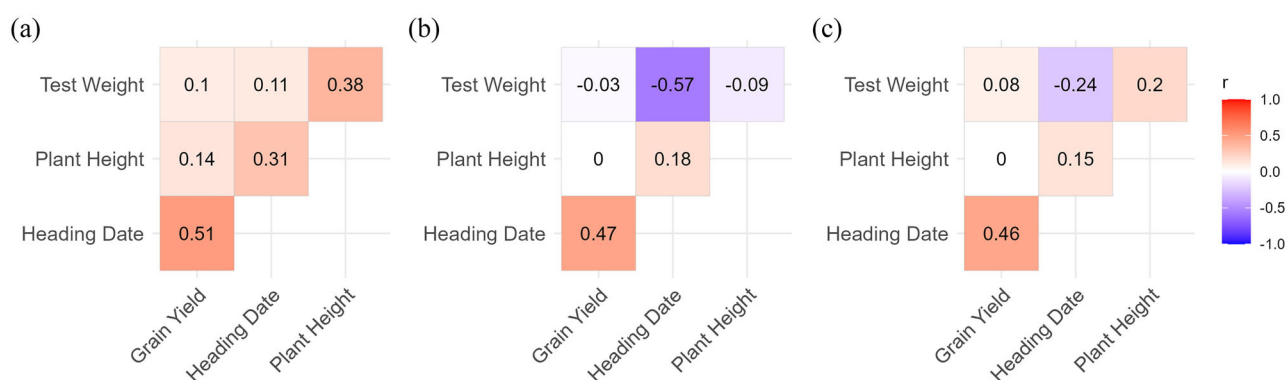
FIGURE 4 Pearson's correlation of best linear unbiased estimations in (a) the 2018 advanced nursery, (b) 2020 advanced nursery, and (c) the combined analysis across years. The *y* and *x* axes display the trait name. Boxes within each subplot display the Pearson's correlation coefficient. Boxes are color coded to indicate strength of correlation between traits. The legend on the right displays the scale for the heatmap associated with correlation strength and sign.
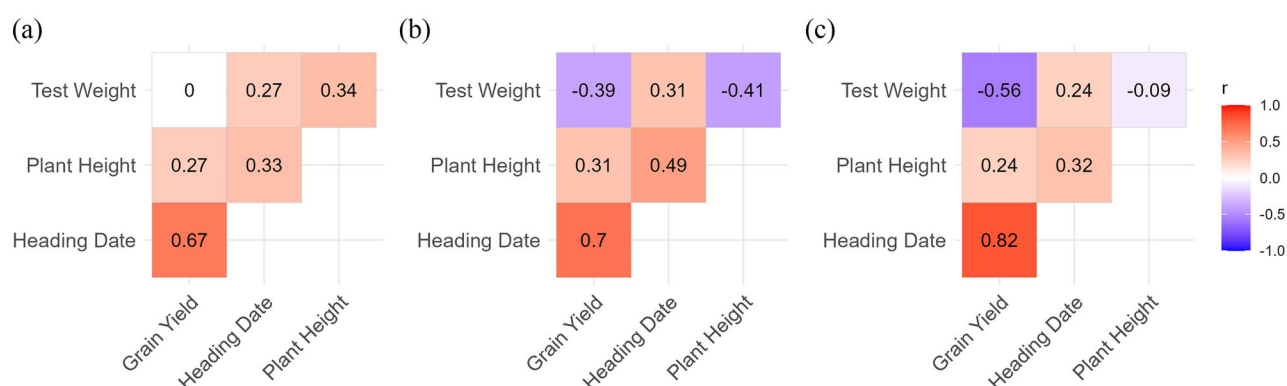
FIGURE 5 Genetic correlations derived from multivariate models in (a) the 2018 advanced nursery, (b) 2020 advanced nursery, and (c) the combined analysis across years. The *y* and *x* axes display the trait name. Boxes within each subplot display the genetic correlation coefficient. Boxes are color coded to indicate strength of correlation between traits. The legend on the right displays the scale for the heatmap associated with correlation strength and sign.
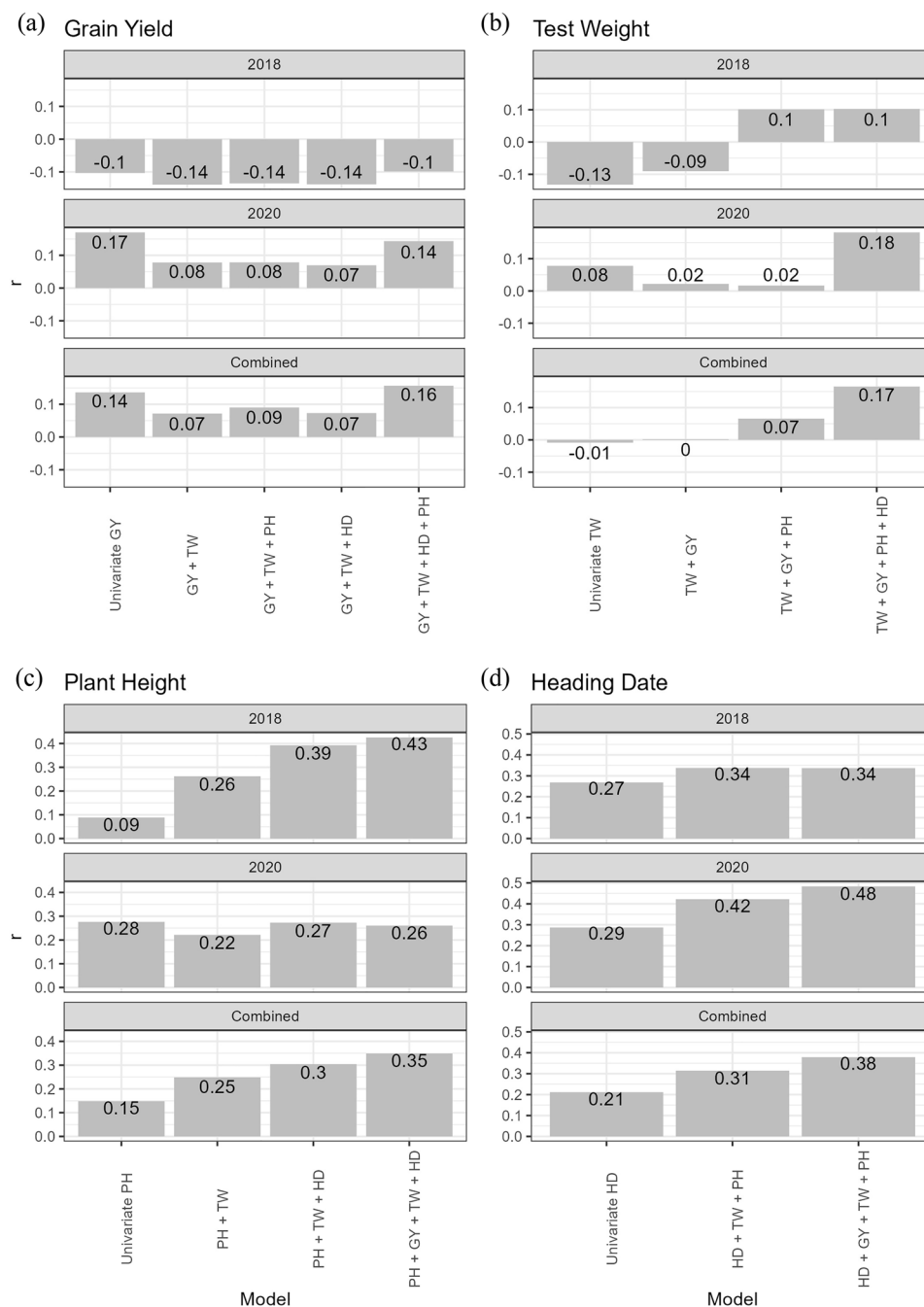
FIGURE 6 Bar charts of forward validated accuracies for (a) grain yield (GY), (b) test weight (TW), (c) plant height (PH), and (d) heading date (HD). Each subplot contains subsections for the 2018, 2020, and combined analysis of the University of Arkansas' advanced nurseries found in the header of each subsection. The prediction accuracy presented as a Pearson's correlation coefficient (r) is found on the y axis and displayed in the corresponding bar. The x axis displays the names of the models. Any model labeled as "Univariate" presents the univariate genomic selection prediction accuracy for the trait following the name. All other x axis labels identify multivariate genomic selection models. The trait to be predicted leads all other traits used as covariates.

incorporated test weight and heading date as well as the model which utilized all agronomic traits ($r = 0.34$). The highest achieved accuracy in 2020 was through the multivariate model which incorporated all agronomic traits ($r = 0.48$). In the combined analysis, the model which achieved the highest prediction accuracy was the multivariate model which incorporated all agronomic traits ($r = 0.38$).

4 | DISCUSSION

In the present work, we used an unbalanced, multiple site-year phenotypic dataset and the corresponding genetic marker data to investigate differences in prediction accuracies for both univariate and multivariate genomic prediction models in SRWW. We then utilized multivariate models which appeared

to outperform the univariate models in cross-validation in a forward validation scenario using the University of Arkansas' advanced nurseries. We observed that multivariate models that include moderately correlated traits generally produced either significantly or marginally higher prediction means in comparison to their respective univariate models in cross-validation. Furthermore, we observed that for higher-heritability traits, such as plant height and heading date, the inclusion of other traits does not generally produce a notable increase in prediction accuracy in cross-validation.

In forward validation, we observed that high-heritability traits, like heading date and plant height, did achieve improved accuracies when covariates were included in MVgBLUP models. Moreover, we did not observe differences in forward validated accuracies that we saw in cross-validation grain yield. Despite this, we observed that, in general, inclusion of correlated traits in multivariate prediction increases the forward predictive accuracy for all traits assessed. Furthermore, this increase was apparent irrespective of the notable differences in the genetic correlations of the traits in the advanced nursery in comparison to the training population.

When considering a relaxed critical value threshold of $\alpha = 0.15$, our results support the hypothesis that multivariate models, using correlated traits, can moderately improve genomic prediction accuracies, in cross-validation, for at least grain yield and test weight in this dataset. Other works in cranberry and African cassava have also reported higher cross and forward validated genomic prediction accuracies for multivariate models in comparison to a univariate approach to genomic prediction (Covarrubias-Pazaran et al., 2018; Okeke et al., 2017).

Crain et al. (2018) observed that using multivate models with weather covariates to predict yield resulted in prediction accuracies that were 33% lower to 7% higher than predictions made by univariate models. Increases in accuracy for prediction of traits within environments while using secondary traits as covariates were also reported in wheat by Rutkoski et al. (2016). These similar results indicate that MVGS has potential application in increasing genomic prediction accuracies for expensive or hard to collect phenotypes.

Sandhu et al. (2021) utilized a convolutional feed-forward neural network to perform multivariate genomic prediction with the utilization of GEBVs and spectral information, and found that mean prediction accuracies varied between −4% and 15% better than univariate genomic prediction for grain yield and protein content in spring wheat. Furthermore, Larkin et al. (2020) found that including covariates such as heading date and plant height in MVGS increased predictive accuracies for *Fusarium* head blight resistance traits. Like these previous studies, this current work supports the hypothesis that inclusion of covariates in MVGS can increase predictive accuracy.

In our cross-validation scheme, we included data of correlated traits in MVgBLUP models for the validation population, instead of removing those data points in each permutation like the predicted response. This implies that, in forward validation, we would have data for other traits (other than the predicted response) and use them in the prediction model. This would also require collection of these traits from multiple site-years going forward. Within most wheat breeding programs, however, it is commonplace to collect data on plant height and heading date prior to the evaluation of lines for yield. Collection of test weight data is also feasible in early generations where plot sizes are not large enough to collect informative yield data. Given this, using these traits to improve genomic predictions of grain yield has real utility in applied cultivar development.

Pre-harvest traits, like heading date and plant height, can be taken prior to harvest and could potentially be used for covariates. Likewise, if plant height or heading date was not taken for a harvested and genotyped field, the other agronomic traits could be used to produce a genomic prediction for the missing trait. However, we did not find any case where a MVgBLUP model significantly outperformed the UVgBLUP model for either plant height or heading date in the cross-validations. Perhaps this was due to the unbalance of the training dataset, the genetic architecture of the population, or the relatively high heritability of these traits in the dataset.

For test weight, our implemented methods may be applicable because yield may be taken in fields and test weight data may not be available due to limited workforce or resources. For yield, test weight will have to be taken in correspondence with yield, because grain will have to be harvested for test weight to be measured. Since measurements of test weight often occur during or after harvest, this inclusion of test weight in a MVgBLUP model for forward prediction of grain yield is therefore not a possibility in practical applications. However, for application of these methods in forward prediction, test weight could be sampled in smaller plots in earlier generations and used with heading date and plant height to produce a higher accuracy genomic prediction for yield—which would presumptively not be taken in earlier generations. Furthermore, in fields where mechanized harvest is not possible due to resource limitations or where harvest is not desirable due to poor plot quality, subsamples of yield plots could be taken, test weight could be measured from those samples, and predictions for yield could be subsequently made using the test weight to obtain increased prediction accuracies.

In our forward validation scenario, predictions for grain yield were uninformative in 2018. In 2020, it was observed that inclusion of covariates caused a reduction in prediction accuracy. However, when we analyzed the data across years in the combined dataset, we saw a marginal improvement ($r_{\mathrm{MVgBLUP}} - r_{\mathrm{UVgBLUP}} = 0.02$) in prediction accuracy when

including test weight, heading date, and plant height. When looking at the Pearson's and genetic correlations of the 2018, 2020, and combined data, we see relatively high correlations of grain yield with heading date. This may imply late freeze in both 2018 and 2020, which was not apparent in the training data, which may have led to low forward prediction accuracies for grain yield.

For test weight, it appears that, across years, inclusion of grain yield, plant height and heading date resulted in improvements in prediction accuracies ($r_{\text{MVgBLUP}} - r_{\text{UVgBLUP}} = 0.20$ in 2018, $r_{\text{MVgBLUP}} - r_{\text{UVgBLUP}} = 0.10$ in 2020, and $r_{\text{MVgBLUP}} - r_{\text{UVgBLUP}} = 0.18$ in the combined dataset). These results are encouraging for more accurate predictions of test weight when utilizing other traits as covariates. For plant height and heading date, prediction accuracies consistently increased, except for 2020 plant height, with the addition of other agronomic traits as covariates in the model.

5 | CONCLUSION

Genomic selection has changed, and will continue to change, the field of plant breeding. We demonstrated in the current work that grain yield, test weight, plant height, and heading date can be predicted with higher accuracy when using a multivariate genomic prediction approach in comparison to the univariate method. While cross-validation indicates improved accuracy for grain yield and test weight within this training population, increases in accuracy are more easily observed for test weight, plant height, and heading date in our forward validation scenario. The data we presented in the current work is only a modicum of the total available phenotypic and genotypic data, and further investigation of large-scale implementation of these methods should be conducted. We therefore suggest this method be more broadly applied across multiple years and locations of breeding nurseries to identify if these trends hold true in larger datasets.

AUTHOR CONTRIBUTIONS
Zachary J. Winn: Data curation; formal analysis; methodology; software; visualization; writing—original draft; writing—review and editing. **Dylan L. Larkin**: Conceptualization; data curation; formal analysis; resources; visualization; writing—original draft; writing—review and editing. **Dennis N. Lozada**: Data curation; writing—review and editing. **Noah DeWitt**: Writing—review and editing. **Gina Brown-Guedira**: Data curation; writing—review and editing. **Richard Esten Mason**: Writing—original draft; writing—review and editing.

CONFLICT OF INTEREST STATEMENT
The authors declare no conflict of interest.

ORCID
Zachary J. Winn https://orcid.org/0000-0003-1543-1527
Dylan L. Larkin https://orcid.org/0000-0002-4991-5637
Noah DeWitt https://orcid.org/0000-0001-9055-993X

REFERENCES
Akdemir, D., Sanchez, J. I., & Jannink, J. L. (2015). Optimization of genomic selection training populations with a genetic algorithm. *Genetics Selection Evolution*, 47, 38. https://doi.org/10.1186/s12711-015-0116-6

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., Feuillet, C., Keller, B., Rogers, J., Pozniak, C. J., Choulet, F., Distelfeld, A., Poland, J., Ronen, G., Sharpe, A. G., Barad, O., Baruch, K., Keeble-Gagnère, G., Mascher, M., … Wang, L. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, 361(6403), eaar7191. https://doi.org/10.1126/science.aar7191

Asoro, F. G., Newell, M. A., Beavis, W. D., Scott, M. P., Tinker, N. A., & Jannink, J. L. (2013). Genomic, marker-assisted, and pedigree-BLUP selection methods for beta-glucan concentration in elite oat. *Crop Science*, 53(5), 1894–1906. https://doi.org/10.2135/cropsci2012.09.0526

Benson, J., Brown-Guedira, G., Murphy, J. P., & Sneller, C. (2012). Population structure, linkage disequilibrium, and genetic diversity in soft winter wheat enriched for fusarium head blight resistance. *Plant Genome*, 5(2), 71–80. https://doi.org/10.3835/plantgenome2011.11.0027

Bernardo, R., & Yu, J. M. (2007). Prospects for genomewide selection for quantitative traits in maize. *Crop Science*, 47(3), 1082–1090. https://doi.org/10.2135/cropsci2006.11.0690

Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *The American Journal of Human Genetics*, 103(3), 338–348. https://doi.org/10.1016/j.ajhg.2018.07.015

Butler, D., Cullis, B. R., Gilmour, A., & Gogel, B. (2009). *ASReml-R reference manual*. The State of Queensland, Department of Primary Industries and Fisheries.

Calus, M. P. L., & Veerkamp, R. F. (2011). Accuracy of multi-trait genomic selection using different methods. *Genetics Selection Evolution*, 43, 26. https://doi.org/10.1186/1297-9686-43-26

Chen, X. J., Min, D. H., Yasir, T. A., & Hu, Y. G. (2012). Evaluation of 14 morphological, yield-related and physiological traits as indicators of drought tolerance in Chinese winter bread wheat revealed by analysis of the membership function value of drought tolerance (MFVD). *Field Crops Research*, 137, 195–201. https://doi.org/10.1016/j.fcr.2012.09.008

Chen, C. P. J., & Zhang, Z. W. (2018). iPat: Intelligent prediction and association tool for genomic research. *Bioinformatics*, *34*(11), 1925–1927. https://doi.org/10.1093/bioinformatics/bty015

Combs, E., & Bernardo, R. (2013). Accuracy of genomewide selection for different traits with constant population size, heritability, and number of markers. *Plant Genome*, *6*(1), 7. https://doi.org/10.3835/plantgenome2012.11.0030

Covarrubias-Pazaran, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., Vorsa, N., Iorizzo, M., & Zalapa, J. (2018). Multivariate GBLUP improves accuracy of genomic selection for yield and fruit weight in biparental populations of *Vaccinium macrocarpon* Ait. *Frontiers in Plant Science*, *9*, 1310. https://doi.org/10.3389/fpls.2018.01310

Crain, J., Mondal, S., Rutkoski, J., Singh, R. P., & Poland, J. (2018). Combining high-throughput phenotyping and genomic information to increase prediction and selection accuracy in wheat breeding. *Plant Genome*, *11*(1), 170043. https://doi.org/10.3835/plantgenome2017.05.0043

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R Package rrBLUP. *Plant Genome*, *4*(3), 250–255. https://doi.org/10.3835/plantgenome2011.08.0024

Graybosch, R. A., & Peterson, C. J. (2010). Genetic improvement in winter wheat yields in the great plains of North America, 1959–2008. *Crop Science*, *50*(5), 1882–1890. https://doi.org/10.2135/cropsci2009.11.0685

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

Habier, D., Fernando, R. L., & Dekkers, J. C. M. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics*, *177*(4), 2389–2397. https://doi.org/10.1534/genetics.107.081190

Heffner, E. L., Sorrells, M. E., & Jannink, J. L. (2009). Genomic selection for crop improvement. *Crop Science*, *49*(1), 1–12. https://doi.org/10.2135/cropsci2008.08.0512

Heslot, N., Yang, H. P., Sorrells, M. E., & Jannink, J. L. (2012). Genomic selection in plant breeding: A comparison of models. *Crop Science*, *52*(1), 146–160. https://doi.org/10.2135/cropsci2011.06.0297

Isidro, J., Jannink, J. L., Akdemir, D., Poland, J., Heslot, N., & Sorrells, M. E. (2015). Training set optimization under population structure in genomic selection. *Theoretical and Applied Genetics*, *128*(1), 145–158. https://doi.org/10.1007/s00122-014-2418-4

Jannink, J. L., Lorenz, A. J., & Iwata, H. (2010). Genomic selection in plant breeding: From theory to practice. *Briefings in Functional Genomics*, *9*(2), 166–177. https://doi.org/10.1093/bfgp/elq001

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

Kelley, J. (2018). *2018 Arkansas wheat quick facts*. University of Arkansas Division of Agriculture Research and Extension.

Larkin, D. L., Holder, A. L., Mason, R. E., Moon, D. E., Brown-Guedira, G., Price, P. P., Harrison, S. A., & Dong, Y. (2020). Genome-wide analysis and prediction of fusarium head blight resistance in soft red winter wheat. *Crop Science*, *60*(6), 2882–2900. https://doi.org/10.1002/csc2.20273

Larkin, D. L., Lozada, D. N., & Mason, R. E. (2019). Genomic selection-considerations for successful implementation in wheat breeding programs. *Agronomy*, *9*(9), 479. https://doi.org/10.3390/agronomy9090479

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Liu, H. P., Searle, I. R., Mather, D. E., Able, A. J., & Able, J. A. (2015). Morphological, physiological and yield responses of durum wheat to pre-anthesis water-deficit stress are genotype-dependent. *Crop & Pasture Science*, *66*(10), 1024–1038. https://doi.org/10.1071/cp15013

Lozada, D. N., & Carter, A. H. (2019). Accuracy of single and multi-trait genomic prediction models for grain yield in US Pacific Northwest Winter Wheat. *Crop Breeding, Genetics and Genomics*, *1*(2), e190012. https://doi.org/10.20900/cbgg20190012

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Nyquist, W. (1962). Differential fertilization in the inheritance of stem rust resistance in hybrids involving a common wheat strain derived from *Triticum timopheevi*. *Genetics*, *47*(8), 1109. https://doi.org/10.1093/genetics/47.8.1109

Okeke, U. G., Akdemir, D., Rabbi, I., Kulakow, P., & Jannink, J.-L. (2017). Accuracies of univariate and multivariate genomic prediction models in African cassava. *Genetics Selection Evolution*, *49*(1), 88. https://doi.org/10.1186/s12711-017-0361-y

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE*, *7*(2), e32253. https://doi.org/10.1371/journal.pone.0032253

Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS ONE*, *8*(6), e66428. https://doi.org/10.1371/journal.pone.0066428

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Perez, L. G., Crossa, J., Reynolds, M., & Singh, R. (2016). Canopy temperature and vegetation indices from High-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. *G3-Genes Genomes Genetics*, *6*(9), 2799–2808. https://doi.org/10.1534/g3.116.032888

Rutkoski, J., Singh, R. P., Huerta-Espino, J., Bhavani, S., Poland, J., Jannink, J. L., & Sorrells, M. E. (2015). Genetic gain from phenotypic and genomic selection for quantitative resistance to stem rust of wheat. *Plant Genome*, *8*(2), 10. https://doi.org/10.3835/plantgenome2014.10.0074

Sandhu, K., Patil, S. S., Pumphrey, M., & Carter, A. (2021). Multitrait machine- and deep-learning models for genomic selection using spectral information in a wheat breeding program. *The Plant Genome*, *14*(3), e20119. https://doi.org/10.1002/tpg2.20119

Sarinelli, J. M., Murphy, J. P., Tyagi, P., Holland, J. B., Johnson, J. W., Mergoum, M., Mason, R. E., Babar, A., Harrison, S., Sutton, R., Griffey, C. A., & Brown-Guedira, G. (2019). Training population selection and use of fixed effects to optimize genomic predictions in a historical USA winter wheat panel. *Theoretical and Applied Genetics*, *132*(4), 1247–1261. https://doi.org/10.1007/s00122-019-03276-6

Sun, J., Rutkoski, J. E., Poland, J. A., Crossa, J., Jannink, J. L., & Sorrells, M. E. (2017). Multitrait, random regression, or simple repeatability model in high-throughput phenotyping data improve genomic prediction for wheat grain yield. *Plant Genome*, *10*(2), 12. https://doi.org/10.3835/plantgenome2016.11.0111

Tshikunde, N. M., Mashilo, J., Shimelis, H., & Odindo, A. (2019). Agronomic and physiological traits, and associated quantitative trait loci (QTL) affecting yield response in wheat (*Triticum aestivum* L.): A review. *Frontiers in Plant Science*, *10*, 1428. https://doi.org/10.3389/fpls.2019.01428

United Nations Department of Economic and Social Affairs. (2019). *World population prospects 2019: Highlights*. United Nations. https://population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

Walkowiak, S., Gao, L. L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., Ramirez-Gonzalez, R. H., Kolodziej, M. C., Delorean, E., Thambugala, D., Klymiuk, V., Byrns, B., Gundlach, H., Bandi, V., Siri, J. N., Nilsen, K., Aquino, C., Himmelbach, A., … Pozniak,

C. J (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature*, *588*, 277–283. https://doi.org/10.1038/s41586-020-2961-x

Watson, A., Hickey, L. T., Christopher, J., Rutkoski, J., Poland, J., & Hayes, B. J. (2019). Multivariate genomic selection and potential of rapid indirect selection with speed breeding in spring wheat. *Crop Science*, *59*(5), 1945–1959. https://doi.org/10.2135/cropsci2018.12.0757