12-2020

Improving Breeding Program Efficiency and Genetic Gain Through the Implementation of Genomic Selection in Diverse Wheat Germplasm

Dylan Larkin
University of Arkansas, Fayetteville

Improving Breeding Program Efficiency and Genetic Gain Through the Implementation of Genomic Selection in Diverse Wheat Germplasm

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Crop, Soil, and Environmental Sciences

by

Dylan Larkin
Oregon State University
Bachelor of Science in Crop and Soil Science, 2015
Washington State University
Master of Science in Crop Science, 2017

December 2020
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Richard Esten Mason, Ph.D.
Dissertation Director

Leandro Mozzoni, Ph.D.
Committee Member

David Miller, Ph.D.
Committee Member

Margaret Worthington, Ph.D.
Committee Member

Qingyang Zhang, Ph.D.
Committee Member

ABSTRACT

Genomic selection (GS) is an important tool for increasing genetic gain for economically important traits in breeding programs. Genomic selection uses molecular markers across the entire genome in order to predict the performance of breeding lines for a trait of interest prior to phenotyping. A training population (TP) of elite germplasm, representative of the University of Arkansas wheat breeding program, was developed in order to predict important agronomic and Fusarium head blight (FHB) resistance traits within the University of Arkansas wheat breeding program through cross-validation and forward prediction.

A genome-wide association study (GWAS) was performed on the TP to identify novel FHB resistance loci for deoxynivalenol (DON) accumulation, Fusarium damaged kernels (FDK), incidence (INC), and severity (SEV). Significantly loci were used as fixed effects in a GS model (GS+GWAS) and compared to a naïve GS (NGS) model, where the NGS models had significantly higher prediction accuracies (PA) than the GS+GWAS models for all four FHB traits. The GWAS identified novel loci for all four FHB traits, most notably on chromosomes 3BL and 4BL. Multivariate GS (MVGS) models using correlated traits as covariates were also compared to NGS models and the MVGS models significantly outperformed the NGS models for all four traits.

The same TP was also evaluated for five agronomic traits, including grain yield (GY), heading date (HD), maturity date (MD), plant height (PH), and test weight (TW), where MVGS models were compared to NGS models. Again, MVGS models significantly outperformed NGS models for all five agronomic traits, especially when there were strong genetic correlations between predicted traits and covariates. Additionally, MVGS models were tested using GY data for genotypes only grown in some environments to predict said genotypes in missing

environments. This method significantly improved PA for GY between 6% and 21% for four of six tested environments.

The abovementioned TP was then used for forward prediction to predict GY for untested $F_{4:6}$ breeding lines and DON, FDK, and SEV for $F_{4:7}$ breeding lines. The MVGS models were comparable to phenotypic selection and had higher selection accuracies for two of three breeding cycles for GY, both cycles for DON, and at least one cycle for FDK and SEV. The MVGS model also had higher PAs for all four traits compared with the NGS models.

These results show that GS, and MVGS, can be successfully implemented in a wheat breeding program over multiple breeding cycles and can be effective alongside phenotypic selection for economically important traits. The MVGS models are particularly effective when predicted traits share strong genetic correlations with covariate traits, and covariate traits have a higher heritability than the predicted traits.

ACKNOWLEDGMENTS

I would first like to thank my family: Jim, Stephanie, and Kyle, and friends for their love, support, and patience through my last ten years as an undergraduate and graduate student. They were an invaluable support group even when I was living on the other side of the country. I would also like to thank Dr. Esten Mason for his excellent advising and support, as well as my committee members Dr. Leandro Mozzoni, Dr. Margaret Worthington, Dr. David Miller, and Dr. Qinyang Zhang. Additionally, I want to thank Dr. Pat Hayes and Dr. Kim Garland-Campbell for their mentoring and for inspiring and encouraging me to further pursue my education in plant breeding. I also want to thank Reverends Shane Moore and Terry Gosnell for their spiritual guidance throughout this entire process. Finally, I want to thank the University of Arkansas wheat breeding staff: David Moon, Randy Miller, Peter Rohman, Dr. Dennis Lozada, Dr. Andrea Acuna, Amanda Holder, Zach Winn, Jamison Murry, Mikayla Hammers, and the many undergraduate student workers for their help with conducting my research and their constant support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

LIST OF FIGURES

CHAPTER II

CHAPTER III

CHAPTER IV

Figure 3. Pirate plot comparing the mean prediction accuracies between multivariate genomic selection (MVGS) models with naïve genomic selection (NGS) models for grain yield (GY) in soft red winter wheat across three different training populations (TP). The *x* axis represents the combination of TP and GS model used to predict GY. The *y* axis represents the mean prediction accuracy across 100 iterations of fivefold cross-validation in the form of a Pearson correlation coefficient (*r*) between the predicted genome estimated breeding value (GEBV) and the actual phenotypic value for the validation populations. Individual points represent the Pearson correlation from each fold of each iteration of cross-validation for a total of 500 datapoints. The lines within each plot represent the mean and 95% confidence intervals for prediction accuracy. The curves represent the smoothed densities of the data 214

Figure 4. Pirate plots comparing the mean prediction accuracies between multivariate genomic selection (MVGS) models with naïve genomic selection (NGS) models for three Fusarium head blight resistance traits (FHB), deoxynivalenol (DON) concentration, Fusarium damaged kernels (FDK), and severity (SEV) in soft red winter wheat across two training populations (TPs): (a) TP18_FHB, TP used to predict three FHB resistance traits for the 2018 advanced $F_{4:7}$ generation (ADV18); (b) TP19_FHB, TP used to predict three FHB resistance traits for the 2019 advanced $F_{4:7}$ generation (ADV19), consisting of all genotypes from TP18_FHB and ADV18. The x axis represents the combination of FHB resistance traits and GS model used to predict each trait. The y axis represents the mean prediction accuracy across 100 iterations of fivefold cross-validation in the form of a Pearson correlation coefficient (*r*) between the predicted genome estimated breeding value (GEBV) and the actual phenotypic value for the validation populations. Individual points represent the Pearson correlation from each fold of each iteration of cross-validation for a total of 500 datapoints. The lines within each plot represent the mean and 95% confidence intervals for prediction accuracy. The curves represent the smoothed densities of the data 215

Figure 5. Scatter plots between genome estimated breeding values (GEBVs) for grain yield (GY) in soft red winter wheat from two different genomic selection models (GS), including naïve models without covariates (NGS) and multivariate GS models with test weight as a covariate (MVGS), and adjusted means for GY across multiple generations of selection, starting with the 2016-2017 growing season through the 2018-2019 growing season: (a) predictions for $F_{4:6}$ observation genotypes from the 2016-2017 growing season (WO17) using a NGS model, (b) predictions for WO17 using a MVGS model, (c) predictions for $F_{4:6}$ observation genotypes from the 2017-2018 growing season (WO18) using a NGS model, (d) predictions for WO18 using a MVGS model, (e) predictions for $F_{4:6}$ observation genotypes from the 2018-2018 growing season (WO18) using a NGS model, (f) predictions for WO18 using a MVGS model. The *x*-axis represents adjusted mean for GY across the WO, $F_{4:7}$ advanced (ADV), and $F_{4:8}$ elite (ARE) generations. Adjusted means for GY were only calculated across WO19 and ADV20 for (e) and (f). The *y*-axis represents the GEBVs calculated for GY from the NGS or MVGS models. Different colored datapoints represent genotypes that were advanced to the next generation. The solid vertical line represents the mean of the adjusted means for GY from the WO generation, while the vertical dashed line represents the mean of the adjusted means for GY from the ADV generations, and the vertical dot-dash line represents the mean of the adjusted means for GY from the ARE generation. The solid horizontal line represents the mean of GEBVs for GY

LIST OF PUBLISHED PAPERS

Larkin, D. L., Lozada, D. N., & Mason, R. E. (2019). Genomic Selection-Considerations for Successful Implementation in Wheat Breeding Programs. Agronomy-Basel, 9(9), 18. doi:10.3390/agronomy9090479 (Chapter I)

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

ABBREVIATIONS

ADV	$F_{4:7}$ advanced nursery
AFLP	amplified fragment length polymorphism
AMP	association mapping panel
ARE	$F_{4:8}$ Arkansas elite nursery
AUG	augmented design
BL	Bayesian least absolute shrinkage and selection operator
BLINK	Bayesian-information and linkage-disequilibrium iteratively nested keyway
BLUP	best linear unbiased prediction
BRR	Bayesian ridge regression
CD	coefficient of determination
CMLM	compressed mixed linear model
CV1	genomic selection cross validation with genotypes missing from all environments
CV2	sparse testing genomic selection cross validation with genotypes missing from some environments
DArT	diversity arrays technology
DH	doubled haploid
DON	deoxynivalenol
E-Bayes	empirical Bayes
EGBLUP	extended genomic best linear unbiased predictor
EM	expectation maximization
EMMA	efficient mixed-model association
ENV	environment

FarmCPU	fixed and random model circulating probability unification
FAY	Milo J. Shult Agricultural Research and Extension Center
FDK	Fusarium damaged kernels
FDR	false discovery rate
FHB	Fusarium head blight
GAPIT	Genomic Association and Prediction Integrated Tool
GBLUP	genomic best linear unbiased predictor
GBS	genotyping by sequencing
GEBV	genome estimated breeding value
GEI	genotype by environment interaction
GLM	general linear model
GLMM	generalized linear mixed model
GS	genomic selection
GS+GWAS	genomic selection plus de novo genome-wide association study
GS+HD	genomic selection plus de novo genome-wide association study for heading date
GS+PH	genomic selection plus de novo genome-wide association study for plant height
GWAS	genome-wide association study
GxE	genotype by environment
GY	grain yield
HD	heading date
HTP	high-throughput phenotyping
INC	incidence
iPat	Intelligent Prediction and Association Tool

ISK	kernel quality index
KSR	Northeast Research and Extension Center
LASSO	least absolute shrinkage and selection operator
LD	linkage disequilibrium
LOD	log odds difference
LSU	Macon Ridge Research Station
MAR	Lon Mann Cotton Research Center
MAS	marker assisted selection
MCMC	Markov chain Monte Carlo
MD	maturity date
ME	mega-environment
MET	multiple-environment trial
mha	million hectares harvested
MLM	mixed linear model
MLMM	multiple locus mixed linear model
MMT	million metric tons
MTA	marker trait association
MVGS	multivariate genomic selection
NAM	nested association mapping
NDVI	normalized difference vegetative index
NGS	naïve genomic selection
NNET	artificial neural network
NPT	Newport Extension Center

PA	prediction accuracy
PCA	principal component analysis
PEV	predictive error variance
PH	plant height
PYT	preliminary yield trial
QQ	quantile-quantile
QTL	quantitative trait loci
RCBD	randomized complete block design
RFLP	restriction fragment length polymorphism
RKHS	reproducing kernel Hilbert space
RR-BLUP	ridge regression best linear unbiased predictor
SEV	severity
SNP	single nucleotide polymorphism
SRWW	soft red winter wheat
SS	stratified sampling
SSR	simple sequence length polymorphism
STPGA	selection of training populations by genetic algorithm
Sungrains	Southeastern University Grains Cooperative
TP	training population
TW	test weight
UE	USDA-ARS Uniform Eastern Nursery
US	USDA-ARS Uniform Southern Nursery
VP	validation population

wBSR	weighted Bayesian shrinkage regression
WO	$F_{4:6}$ wheat observation nursery

INTRODUCTION

OVERALL INTRODUCTION

Genomic selection (GS) has rapidly become one of the more valuable tools within the plant breeder's toolbox. As a result, many plant breeding programs throughout the world, in a diverse number of crops, have worked to find the best way to implement GS into their own breeding programs. Likewise, this has been the goal for the University of Arkansas soft red winter wheat (SRWW) (*Triticum aestivum* L.) breeding program.

Genomic selection is a modified form of marker assisted selection; however instead of using only a few select markers that are significantly associated with a trait, GS takes all summed marker and locus effects across the entire genome into consideration to calculate genome estimated breeding values (GEBV). At least some of the markers are assumed to be in linkage disequilibrium (LD) with quantitative trait loci (QTL) associated with a trait of interest. Genomic selection can be beneficial for a breeding program because it can accelerate genetic gain by reducing the amount of time within a breeding cycle, allowing for two generations of selection compared to one generation with traditional phenotypic selection. Additionally, GS can help to reduce the amount of time and resources needed for phenotyping, as well as reducing the impact of genotype by environment interactions for traits that are difficult to phenotype. Previous research has shown that GS mixed models using individual markers as fixed effects, as determined through genome-wide association studies (GWAS) (GS+GWAS) could significantly improve prediction accuracies for a trait of interest if the markers accounted for a large amount of variation in the trait of interest. Multivariate GS (MVGS) models using secondary traits as covariates could also significantly improve prediction accuracies for a trait of interest, especially when the secondary traits shared strong genetic correlations with the trait of interest.

The goal for this research was to determine if the abovementioned GS models could be effectively used to predict the performance of SRWW germplasm within the University of Arkansas wheat breeding program for economically important traits, such as grain yield (GY), agronomic performance, and disease resistance related to Fusarium head blight (FHB). In order to test this, a training population of elite germplasm, used as parental materials for the University of Arkansas wheat breeding program, was developed and grown over multiple environments, and then phenotyped for GY and test weight (TW), in addition to three agronomic traits- heading date (HD), maturity date (MD), and plant height (PH). The same population was also phenotyped for four traits related to FHB resistance- deoxynivalenol (DON) accumulation, Fusarium damaged kernels (FDK), incidence (INC), and severity (SEV). Cross-validation analyses were then performed for each trait between a naïve GS model without fixed effects or covariates (NGS) and the GS+GWAS and MVGS models in order to compare the mean prediction accuracies between each model. The study describing the cross-validation analyses for the FHB resistance traits was in Chapter II, while the study describing the cross-validation analyses for GY, TW, and agronomic traits was in Chapter III.

While cross-validation is effective for evaluating model performance within a training population, it is also important to test model performance for forward prediction of new breeding materials. Once the best models for each trait were identified in the previous two chapters, the training population and best models were used to compare the performance of GS over multiple generations of selection compared with phenotypic selection in Chapter IV. The training population from Chapters II and III was used to train NGS models and the best MVGS models for GY, DON, FDK, and SEV. The GEBVs from the NGS and MVGS models were then compared with phenotypic data collected from $F_{4:6}$ breeding lines for GY and $F_{4:7}$ breeding lines

for FHB resistance traits based on prediction accuracy, selection accuracy, and response to selection in the $F_{4:8}$ generation. These comparisons were performed over three breeding cycles for GY and two cycles for the FHB resistance traits.

The results from the three research chapters in this dissertation will serve as the foundation for the GS program within the University of Arkansas wheat breeding program based on the procedures and methodology developed for each experiment. The primary takeaways from this research are that GS can be used to effectively predict the performance of breeding lines for economically important traits of interest, from GY to FHB resistance traits. This research also shows that MVGS models have higher prediction accuracies and can have stronger selection accuracies than NGS models for both agronomic and FHB resistance traits. Finally, GS using both NGS and MVGS models can be comparable, if not better, than phenotypic selection when using forward prediction, based on prediction accuracy, selection accuracy, and response to selection.

CHAPTER I

LITERATURE REVIEW

LITERATURE REVIEW

Wheat

Wheat (*Triticum aestivum* L.) has a rich history as one of the first cultivated grain crops, domesticated nearly 10,000 years ago in the Fertile Crescent of central Asia (Eckardt, 2010). Today it is one of the most widely grown crops in the world, grown in a wide range of soils and climates (Curtis & Halford, 2014). Wheat also serves as an important and nutritious food source for much of the world, contributing nearly 20% of calories to the Western diet (Shewry & Hey, 2015).

Bread wheat (*Triticum aestivum* L., 2n = 6x = 42, genomes AABBDD) is a self-pollinated allohexaploid cereal crop in the family *Poaceae* (Shewry, 2009). The wheat genome is approximately 17 Gb and it consists of three closely related genomes, A, B, and D, that came together through natural hybridization events (Marcussen et al., 2014). The A genome originated from a species related to *Triticum urartu* Thumanjan ex Gandilyan (2n = 2x = 14, AA) and hybridized with a species related to *Aegilops speltoides* Tausch (2n = 2x = 14, SS), to form the allotetraploid *Triticum turgidum* Desf. (2n = 4x = 28, AABB) (Dvorak, Akhunov, Akhunov, Deal, & Luo, 2006; Dvorak, Diterlizzi, Zhang, & Resta, 1993; Dvorak, McGuire, & Cassidy, 1988). *T. turgidum* then hybridized with the D genome donor, *Aegilops tauschii* Coss. (2n = 2x = 14, DD), to form modern hexaploid wheat (Dubcovsky & Dvorak, 2007; Kihara, 1944; McFadden & Sears, 1946).

Approximately 765 million metric tons (MMT) of wheat was produced globally in 2020. Global wheat production has increased by nearly 60% over the past decade due to rising demands from Southeast Asia, Sub-Saharan Africa, and the Middle East. This is largely due to

growing populations and changing consumer preferences. In 2020, the five largest wheat producing countries were the collective European Union (155 MMT), China (134 MMT), India (104 MMT), Russia (74 MMT), and the United States (53 MMT) (USDA Foreign Agricultural Service, 2020). Wheat is the second most produced grain crop in the world behind coarse grains, which include any grain crop that is not wheat or rice (*Oryza sativa* L.) (1,411 MMT), and above milled rice (496 MMT) (USDA Foreign Agricultural Service, 2020).

Wheat is the fourth most widely produced crop in the United States at 14.9 million ha harvested (mha) in 2020, behind maize for grain (*Zea mays* L.) (33.4 mha) and soybean (*Glycine max* (L.) Merr.) (33.3 mha) (USDA National Agricultural Statistics Service, 2020a). In 2020, North Dakota produced the largest area of wheat (2.7 mha), followed by Kansas (2.5 mha) and Montana (2.2 mha) (USDA National Agricultural Statistics Service, 2020b).

There are six main market classes in the U.S. wheat market, which include hard red winter wheat, hard red spring wheat, soft red winter wheat, soft white wheat, hard white wheat and durum wheat (*Triticum durum* Desf.) (Bond, 2020; Chao, Zhang, Dubcovsky, & Sorrells, 2007). Winter wheat requires a vernalization period in order to flower and produce seed, whereas spring wheat does not. Approximately 40% of the total U.S. wheat crop is hard red winter wheat, which is primarily grown in the Great Plains, extending from northern Texas to Montana. Hard red winter wheat is used to produce bread flour. The second most common market class, hard red spring wheat, contributes to roughly 20% of the U.S. wheat market. It is mostly grown in the Northern Plains states, which include Montana, Minnesota, North Dakota, and South Dakota. Hard red spring wheat is known for its high protein content, ideal for specialty bread flower and blending with lower protein wheat. Soft red winter wheat contributes to between 15% and 20% of U.S. wheat production. It is mostly grown along the Mississippi River and the South Atlantic

coast of the Eastern U.S. Soft red winter wheat flour is mostly used for cakes, cookies, and crackers. Between 10% and 15% of U.S. wheat production is from white wheat. Soft white wheat consists of three classes of wheat, soft white winter, soft white spring, and club wheat. It is mostly grown in Washington, Oregon, Idaho, Michigan, and New York. Its flour is mostly used for noodles, crackers, cereals, and pastries. Hard white wheat is the newest class of wheat, it is mostly grown in the drier climates of Eastern Colorado, Western Kansas, Western Oklahoma, and the Northern Plain states. Hard white wheat flour is used for the production of noodles, flat breads, and leavened bread (Bond, 2020).

Wheat is the sixth largest crop in Arkansas, based on area harvested at 50,000 ha. It falls behind soybean (1,056,229 ha), hay (507,071 ha), rice (455,676 ha), maize (297,444 ha), and cotton (*Gossypium hirsutum* L.) (246,858 ha). Approximately 71 million kg of wheat was produced in Arkansas in 2019 for a crop value of 12.6 million USD (USDA National Agricultural Statistics Service, 2019). Most Arkansas wheat production occurs in the Eastern portion of the state along the Mississippi River delta. The most common wheat market class grown in Arkansas is soft red winter wheat (Bond, 2020).

Increasing Wheat Yield for a Growing Population

According to the United Nations Department of Economic and Social Affairs, the global population is currently expected to grow from 7.6 to approximately 9.8 billion people by 2050 and 11.2 billion by 2100 (United Nations Department of Economic and Social Affairs: Population Division, 2019). This rapid population growth is expected to exceed the rate of global food production, posing the risk of a potential food crisis in the next few decades. While wheat yield has increased rapidly since the 1950s, largely due to improved breeding practices, the introduction of improved pesticides, and synthetic fertilizers, the current rate is insufficient in

meeting future caloric demands. In fact, the rate of yield increase has been stagnant since the 1980s (Graybosch & Peterson, 2010). It has been estimated that an approximate yield improvement of 2.4% per year is needed in order to double production in wheat by 2050, however the current rate of increase is merely .9% (Ray, Mueller, West, & Foley, 2013).

Many factors limit gains in wheat grain yield (GY), including biotic and abiotic stressors. The selection of pesticide-resistant weeds, insects, and pathogens can be devastating for crop production and they have become a greater concern due to over-application in crop production systems (Ray et al., 2013). Recently, the reduction in arable land, due to urban expansion and soil degradation, have also reduced opportunities for increased production and have even reduced crop production in certain regions (d'Amour et al., 2017). The growing threat of climate change is also a major concern for wheat production, largely due to more extreme temperatures and increased water stress (Jaggard, Qi, & Ober, 2010; Parry et al., 2011; Tilman et al., 2001).

In order to meet the goal of doubling wheat yield by 2050, breeders must work to improve the efficiency of their breeding programs while also implementing new and improved technologies in order to increase genetic gain. Many have done this with new technologies, such as marker-assisted selection (MAS), phenomics, trans and cisgenic approaches, and the implementation of genomic selection (GS) (Heffner, Sorrells, & Jannink, 2009; Houle, Govindaraju, & Omholt, 2010; Larkin, Lozada, & Mason, 2019).

The Breeder's Equation and Genetic Gain

The breeder's equation ($R = ir\sigma_A$) has long been used to show how genetic response (R) changes in response to selection intensity (i), the square root of the additive genetic variance (σ_A), and selection accuracy, which is equivalent to narrow sense heritability in phenotypic selection (r). If

plant breeders want to increase the genetic gain of their breeding program, they must increase at least one of the three components of the breeder's equation within a given the amount of time (t) in a breeding cycle or year ($R = ir\sigma_A/t$) (Falconer & MacKay, 1996; Y. B. Xu et al., 2017).

It can be difficult to maximize genetic gains using phenotypic selection. While phenotypic selection can be effective for traits with high heritability, selection for low heritability traits are often relegated to the later stages of a breeding cycle, particularly with inbred crops (Collard & Mackill, 2008). Another problem with phenotypic selection is that environmental interactions can mask the expression of certain traits making selection difficult. The time between generations can also be long as some traits require a full season in order to be expressed. The limitations of phenotypic selection can hamper the increase in genetic gain in a breeding program (Y. B. Xu et al., 2017). One of the ways to maximize genetic gain within a given breeding cycle is to optimize selection accuracy through the use of molecular genetics, through approaches such as MAS or GS (Collard & Mackill, 2008; Heffner et al., 2009; Meuwissen, Hayes, & Goddard, 2001).

Marker Assisted Selection

One of the technologies breeders have used to increase genetic gain is MAS. In a MAS program, molecular markers closely linked or co-segregating with a desired trait can be used to differentiate breeding lines based on the allelic variation underlying a trait (Collard & Mackill, 2008; Mohan et al., 1997). Practitioners of MAS have been able to report genetic gains twice that of phenotypic selection (Collard, Jahufer, Brouwer, & Pang, 2005; Knapp, 1998; Lande & Thompson, 1990). MAS has several advantages compared to phenotypic selection, some of which include; (1) the ability to select for traits that are difficult to phenotype, especially if they are expensive or time consuming to phenotype, (2) the ability to select for traits that have low

heritability and low expression and are therefore difficult to phenotype, (3) the ability to select for traits where phenotyping is dependent on specific environmental conditions or growth stages (Y. B. Xu & Crouch, 2008), and (4) MAS is most effective when selecting for monogenic or qualitative traits, making it an ideal selection strategy for introgression and backcrossing of single genes into germplasm along with pyramiding of disease resistance genes (Y. B. Xu & Crouch, 2008). Strategies have also been developed so that multiple traits can be selected simultaneously via MAS (Ragot, Gay, Muller, & Durovray, 2000; Y. B. Xu & Crouch, 2008).

While MAS has helped to revolutionize plant breeding while also providing a source for increasing genetic gain, there are also downsides. One of the primary downsides to MAS is that it is less effective when screening for more complex or multi-genic quantitative traits (R. Bernardo & Yu, 2007; Heffner et al., 2009; Hospital, 2009). In order to discover new quantitative trait loci (QTL) associated with a trait of interest, biparental mapping populations must be developed. The problem with biparental mapping populations is that they rarely account for the allelic diversity and genetic background of a full breeding program and their effects tend to vary across environments, particularly for quantitative traits. Therefore, multiple mapping populations must be developed for specific environments and breeding programs in order to validate the position and allelic effects of new QTL which is both expensive and time consuming (Heffner et al., 2009; Y. B. Xu & Crouch, 2008). If a breeder attempts to adapt small effect QTL into their breeding program through MAS prior to validating said QTL using local germplasm, they could achieve genetic gains that are lower than conventional phenotypic selection (Rex Bernardo, 2001). Cost can especially limit the wide-spread use of MAS in a breeding program, especially when considering a genome as complex as wheat (William, Trethowan, & Crosby-Galvan, 2007).

Introduction to Genomic Selection

In response to the limitations of traditional MAS, GS was proposed as an alternative. Genomic selection is a modified form of MAS, however instead of using only a few select markers that are significantly associated with a trait, GS takes all summed marker and locus effects across the entire genome into consideration to calculate genome estimated breeding values (GEBV) (Heffner et al., 2009; Meuwissen et al., 2001). The assumption is that at least some of the single nucleotide polymorphism (SNP) markers are in linkage disequilibrium (LD) with QTL associated with a trait of interest. The use of genome-wide markers and loci allows for more efficient and effective selection of complex quantitative traits compared to MAS (Meuwissen et al., 2001). The GS process begins with a panel of genotypes, referred to as a training population, that have been genotyped and phenotyped. The training population is then used to train a model to calculate GEBVs for a panel of genotypes, the validation population, that have only been genotyped. The breeder can then use the calculated GEBVs to make selections from the validation population without the need for phenotyping (Heffner et al., 2009; Meuwissen et al., 2001). While some have proposed making selections based on genotypes with the best GEBVs, others have proposed using GS in order to eliminate poor performing genotypes from the breeding program in order to maintain genetic diversity and larger population sizes (J. Spindel et al., 2015).

The implementation of molecular markers linked to known QTL in order to predict estimated breeding values for quantitative traits was first described for animal breeding (Fernando & Grossman, 1989). This early use of GS was also described in plants by using restriction fragment length polymorphism (RFLP) markers to predict the hybrid performance in maize (R. Bernardo, 1994). The seminal introduction of GS in its modern form came with the

introduction of high-density SNP genotyping where several prediction models for GS were introduced for animal breeding (Meuwissen et al., 2001).

Impact of Genomic Selection on the Breeder's Equation

When considering the breeder's equation, GS greatly reduces the length of a breeding cycle, allowing for greater genetic gain in one cycle of GS compared with one cycle of phenotypic selection. Two rounds of GS were performed in the time it took to perform one round of phenotypic selection when selecting for resistance to stem rust in wheat (J. Rutkoski et al., 2015). When GS was compared with MAS in maize, genetic response was between 18% and 43% greater in the GS population compared to the MAS population (R. Bernardo & Yu, 2007). Differences in response to selection per cycle between GS, MAS, and phenotypic selection were minimal in oats (*Avena sativa* L.), however the ability for GS to make two generations of selection in the time it took to make one generation of selections for phenotypic and MAS made it favorable for maximizing genetic gain (Asoro et al., 2013).

There are many ways to maximize genetic gain through GS; some of these methods include prediction model selection, marker density, heritability of training and validation sets, LD, the relationships between training and validation sets, population structure, and training set optimization methods (Table 1). The best way to determine prediction accuracy is by performing a Pearson correlation between the GEBVs and true breeding values or phenotypic values of the training population. In the context of the breeder's equation, the prediction accuracy is proportional to genetic gain (R), where the prediction accuracy is equivalent to the selection accuracy (r) (Falconer & MacKay, 1996).

Different Factors Affecting the Accuracy of Genomic Selection

Marker Density

Since GS requires dense marker coverage across the entire genome, a proper genotyping platform must be considered (Meuwissen et al., 2001). Several genome-wide platforms have been implemented in wheat. The first of these were RFLP markers (J. M. Chen & Gustafson, 1995), followed by amplified fragment length polymorphism (AFLP) markers (Barrett & Kidwell, 1998), simple sequence repeat (SSR) or microsatellite markers (Roder et al., 1998), diversity arrays technology (DArT) (Akbari et al., 2006), and SNP maker arrays, such as the 9K and 90K Infinium arrays (Chao et al., 2010; S. C. Wang et al., 2014). Most recently genotyping has been performed using SNP markers within genotyping by sequencing (GBS) platforms (Elshire et al., 2011; Poland et al., 2012). In the case of GS, GBS has become the dominant genotyping platform in use for wheat, particularly due to its low cost, high coverage, and reduced sampling bias compared to SNP arrays (Heslot, Rutkoski, Poland, Jannink, & Sorrells, 2013; Poland et al., 2012).

There is evidence that marker density only marginally impacts GS prediction accuracy and that the use of low to medium density marker datasets could be more cost-effective for GS for Fusarium head blight (FHB) resistance in wheat (Y. Jiang et al., 2015). Large numbers of markers can also result in model overfitting, resulting in lower prediction accuracies when predicting independent datasets (Jannink, Lorenz, & Iwata, 2010).

Higher marker density can also be favorable for GS, as a larger amount of markers can increase the probability that markers will be in LD with QTL (Heffner et al., 2009). Higher marker density can also result in lower LD, significantly improving prediction accuracies

(Norman, Taylor, Edwards, & Kuchel, 2018). In most cases lower LD can result in higher recombination frequency and more accurate estimates of QTL effects (B. E. Huang et al., 2012). Low LD combined with higher marker densities and larger training population sizes can strongly improve prediction accuracy as well (Combs & Bernardo, 2013).

In a study predicting performance of agronomic traits in biparental maize and barley (*Hordeum vulgare* L.) populations, increasing the number of markers improved prediction accuracy. However, once the genome was sufficiently saturated with one marker for every 12.5 cM, the gain in prediction accuracy plateaued. This was also observed in mixed wheat and barley populations, where prediction accuracy plateaued after reaching a moderate marker density (2.0 cM in barley and 4.5 cM in wheat) (Combs & Bernardo, 2013). While lower marker density might be adequate for GS, the cost of genotyping and phenotyping large training populations is not always sustainable for breeding programs. As a result, smaller training populations with higher marker densities are appropriate, especially in the case of biparental populations (Heffner, Jannink, Iwata, Souza, & Sorrells, 2011).

Genomic Prediction Models

One of the key parts of GS is the use of a prediction model that can appropriately incorporate genotypic data in order to produce GEBVs (Broman & Speed, 2002; de los Campos et al., 2009). An appropriate model must be able to estimate a large number of marker effects from the small number of phenotyped individuals of a training population, otherwise known as the 'large *p*, small *n*' dilemma. This situation results in a lack of degrees of freedom and low statistical power, so a model that can account for this dilemma is very important (Heffner et al., 2009; Jannink et al., 2010; Meuwissen et al., 2001). A summary of the genomic prediction models is shown in

Table 2, while a summary of programs that can be used to run the aforementioned models is shown in Table 3.

Stepwise Regression

A model that has been traditionally used for MAS is the stepwise regression model. Marker effects are considered fixed, requiring a stepwise approach where markers are selected individually or in groups in order to avoid a lack of degrees of freedom. Markers are added or removed from the model based on a significance threshold, where insignificant markers are given an effect of zero, while significant marker effects are simultaneously tested in order to determine their effects (Lande & Thompson, 1990). Since only significant marker effects are retained in the model, the model only accounts for a small amount of the total variance and model effects can be overestimated (Heffner et al., 2009). When a GS simulation was performed using stepwise regression, the model had a low prediction accuracy as the model was unable to detect enough of the QTL effects (Meuwissen et al., 2001). The general stepwise model is as follows:

$$y = \mu 1_n + \boldsymbol{X_i} g_i + \varepsilon$$

where y is the vector of outputs; μ is the population mean; $\boldsymbol{X_i}$ is the design matrix for the i^{th} marker; g_i represents the genetic effects of haplotypes at the i^{th} marker; and ε is the residual error.

The log likelihood is calculated as:

$$-0.5[n \ln(\sigma_\varepsilon^2) + \varepsilon' \varepsilon / \sigma_\varepsilon^2]$$

where n is the number of records; and ε and σ_ε^2 denote the estimates of error deviations and error variance, where $\sigma_\varepsilon^2 = \varepsilon' \varepsilon / \big(n - Rank([1_n \boldsymbol{X_i}])\big)$. A log-likelihood is calculated for every marker

using this function. The log-likelihood of each marker is then plotted against the marker position and peaks exceeding a log odds difference (LOD) score of 6.0 are identified, implying the presence of a segregating QTL. Haplotype effects at the respective QTL positions are calculated simultaneously using the model:

$$y = \mu 1_n + \sum_i \boldsymbol{X_i} g_i + \varepsilon$$

where the summation is over all QTL positions that correspond to a likelihood peak, while g_i was calculated at the peak. All haplotype effects that fall below the 6.0 LOD score threshold are given a marker effect of zero (Lande & Thompson, 1990; Meuwissen et al., 2001).

Ridge Regression Best Linear Unbiased Predictor

Ridge regression, sometimes referred to as random regression (Heslot, Yang, Sorrells, & Jannink, 2012), best linear unbiased predictor (RR-BLUP) allows for the simultaneous estimation of all marker effects for GS instead of a select few (Heffner et al., 2009; Meuwissen et al., 2001). The RR-BLUP model also assumes that all markers have a common variance and a normal distribution of marker effects that are shrunken towards zero (Whittaker, Thompson, & Denham, 2000). The downside of RR-BLUP is that its assumption of a common variance between all markers is incorrect, since it could underestimate large-effect QTL (R. Bernardo & Yu, 2007). Even so, RR-BLUP is superior to the antiquated stepwise regression model (Habier, Fernando, & Dekkers, 2007; Meuwissen et al., 2001). The standard form for the RR-BLUP model is as follows:

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

where $u \sim N(0, \boldsymbol{I}\sigma_u^2)$ is the vector of marker effects, which is assumed to have a normal distribution, where $\boldsymbol{I}$ is the identity matrix and σ_u^2 is the variance of the individual marker effects; β is the vector of fixed effects; $\boldsymbol{X}$ is the design matrix of fixed effects; $\boldsymbol{Z}$ is the design matrix relating genotypes to phenotypic observations (y), with m markers in columns and n phenotypes in rows; and $\varepsilon_i \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$ is the residual error at the i^{th} locus, which is assumed to have a normal distribution. The GEBV is the sum of all allele effects of a line (Endelman, 2011).

When compared with ten other genomic selection models in wheat, maize, and barley, the RR-BLUP model was identified as the best for use in an applied breeding program as its performance was comparable to more complex Bayesian and machine learning models, largely due to its lower computational demand (Heslot et al., 2012).

Genomic Best Linear Unbiased Predictor

Genomic best linear unbiased predictor (GBLUP) is essentially equivalent to the RR-BLUP model, however while RR-BLUP calculates marker effects that must be summed to obtain GEBVs, GBLUP uses a genome relationship matrix (*G*-Matrix) instead of an identity matrix (*I*-Matrix) allowing for direct calculations of GEBVs (Legarra, Christensen, Aguilar, & Misztal, 2014; VanRaden, 2008; X. Wang, Xu, Hu, & Xu, 2018). The GBLUP model runs into the same problems as RR-BLUP in that its assumption of equal marker variance across all markers is still inaccurate (R. Bernardo & Yu, 2007). The standard form for the GBLUP model is as follows:

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

where $u \sim N(0, \boldsymbol{G}\sigma_u^2)$ is the vector of additive genetic values for all individuals in the model, where $\boldsymbol{G}$ is the genome relationship matrix and σ_u^2 is the additive genetic variance, which is assumed to have a normal distribution; β is the vector of fixed effects; $\boldsymbol{X}$ is the design matrix of

fixed effects; $\boldsymbol{Z}$ is the design matrix relating genotypes to phenotypic observations (y), with m markers in columns and n phenotypes in rows; and $\varepsilon_i \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$ is the residual error at the i^{th} locus, which is assumed to have a normal distribution (Zhang et al., 2015). The G-Matrix is calculated based on the method proposed by Van Raden (2008) as follows:

$$\boldsymbol{G} = \frac{\boldsymbol{Z}\boldsymbol{Z}'}{2\sum_{i=1}^{m} p_i(1-p_i)}$$

where $\boldsymbol{Z}$ is the adjusted design matrix, with m markers in columns and n phenotypes in rows and p_i is the allele frequency of the second allele at the i^{th} locus in the base population (VanRaden, 2008).

Extended Genomic Best Linear Unbiased Predictor

The extended genomic BLUP (EGBLUP) model is an extension of the GBLUP model, only epistasis is accounted for through the addition of marker-based epistatic relationship matrices. The goal of EGBLUP is to outperform GBLUP by accounting for epistatic effects while also being more computationally efficient than Bayesian epistatic models, which will be covered below. The downside with the EGBLUP model is that, much like RR-BLUP and GBLUP, it assumes equal variances for all markers. The standard form for the EGBLUP model is as follows (He et al., 2016; Y. Jiang & Reif, 2015):

$$y = \boldsymbol{X}\beta + \boldsymbol{Z}u_1 + \boldsymbol{Z}u_2 + \varepsilon_i$$

where $u_1 \sim N(0, \boldsymbol{G}\sigma_1^2)$ is the vector of additive genetic values for all individuals in the model, where $\boldsymbol{G}$ is the genome relationship matrix and σ_1^2 is the additive genetic variance, which is assumed to have a normal distribution; $u_2 \sim N(0, \boldsymbol{H}\sigma_2^2)$ is the vector of additive by additive epistatic genotypic values, where $\boldsymbol{H}$ is the epistatic relationship matrix obtained from the

Hadamard product of the genome relationship matrix by itself (***H***=***G***#***G***); β is the vector of fixed effects; ***X*** is the design matrix of fixed effects; ***Z*** is the design matrix relating genotypes to phenotypic observations (y), with m markers in columns and n phenotypes in rows; and $\varepsilon_i \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$ is the residual error at the i^{th} locus, which is assumed to have a normal distribution. The *G*-matrix is calculated using the same method as in the GBLUP model (VanRaden, 2008).

Bayes-A

The Bayesian models help to reduce the over-shrinking of large marker effects, as seen in RR-BLUP and GBLUP, by better accounting for marker effects of different sizes. In the case of Bayesian models, separate variances are estimated for each marker and are assumed to follow a specific prior distribution (Meuwissen et al., 2001). The Bayes-A method, originally described by Meuwissen et al. (2001), uses an inverted chi-square distribution and a hierarchal three-stage model. Model degrees of freedom and scale parameters are chosen so that the mean and variance of the distribution match the expected mean and variance of marker variances (Heffner et al., 2009; Meuwissen et al., 2001). The Bayes-A model operates on three stages- stage one is a normal regression:

$$y = \boldsymbol{X}\beta + \varepsilon_i$$

where ***X*** is a matrix of marker codes, in which case, the rows are the number of phenotypes and columns are the number of markers; β is the vector of allelic substitution for each marker; and ε_i is the vector of residuals with a normal distribution $\varepsilon_i \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$. In stage two a normal conditional prior to each marker effect with a specific variance for each marker (σ_j^2) is assigned whereas in stage three, the same scaled inverse chi-square distribution $\sigma_j^2 \sim X^{-2}(v, S_\beta^2)$ with a

known scale (S_β^2) and degrees of freedom (v) to each marker variance is designated (Gianola, 2013; Karkkainen & Sillanpaa, 2012; Meuwissen et al., 2001; X. Wang et al., 2018).

The downside of Bayes-A is that it uses Markov chain Monte Carlo (MCMC) estimation, which can be very computationally cumbersome limiting its use for common breeding programs. Also, while Bayes-A is more flexible than RR-BLUP, the additional parameters could adversely impact the inference of marker effects (Gianola, 2013; Karkkainen & Sillanpaa, 2012).

Bayes-B

Bayes-B is different from Bayes-A, in that it uses a mixture distribution of marker effects that peaks around zero, which would make sense, since there are many loci that do not segregate and would therefore have no genetic variance. Bayes-A does not allow for marker variances to actually equal zero. Therefore, Bayes-B is more realistic in that some marker variances are actually allowed to be equal to zero (Heffner et al., 2009). The priors for the Bayes-B hierarchal model are:

$$y_i|\beta, \sigma_\varepsilon^2 \sim N(\boldsymbol{X_i}\beta, \sigma_\varepsilon^2) \text{ and}$$

$$\beta_j|S_\beta^2, v, \pi \sim IID \begin{cases} 0 \; with \; probability \; \pi \\ t(0, S_\beta^2, v) \; with \; probability \; (1-\pi) \end{cases}$$

In this case, the prior is a mixture of the point mass at zero with a *t*-distribution with the mixing probabilities being π (portion of markers that have no effect) and (1-π), where π is assumed to be known and arbitrarily specified (Meuwissen et al., 2001). When π = 0, Bayes-B is equivalent to Bayes-A (Heffner, Jannink, & Sorrells, 2011). While some have identified that Bayes-B has higher predictive ability compared to RR-BLUP and GBLUP, others have found a lack of significant differences between the BLUP and Bayes-B models, in spite of different genetic

architectures and traits. This means that Bayes-B is only effective in learning about genetic architecture when the number of genotypes (*n*) is less than the number of markers (*p*) when the number of true nonzero marker effects is also less than *n* (Gianola, 2013).

Bayesian Ridge Regression

Bayesian Ridge Regression (BRR) is similar to RR-BLUP in that they both assume that all markers have a common variance and shrink equally for each marker effect. However, BRR estimates the level of shrinkage through the use of a Bayesian hierarchal model (Heslot et al., 2012; Perez, de los Campos, Crossa, & Gianola, 2010). In simulation studies, marker-specific shrinkage models like Bayes-A, Bayes-B, or Bayesian Least Absolute Shrinkage and Selection Operator (LASSO) (described below) do a better job at capturing LD between markers and QTLs compared to BRR (Habier et al., 2007; Perez et al., 2010). The standard model for BRR is as follows:

$$y = \mu + \boldsymbol{X}\beta + \varepsilon$$

where y is the response vector; μ is the population mean or intercept; $\boldsymbol{X}$ is the marker design matrix; β is the vector of marker effects, expressed as:

$$\underset{\beta}{\arg min} \left(\frac{\|\gamma - \boldsymbol{X}\beta\|_2^2}{2\sigma^2} + \lambda \|\beta\|_2^2 \right)$$

where the notation $\|.\|_2$ represents the Euclidian norm of $\|\beta\|_2 = \left(\sum_i \beta_i^2\right)^{1/2}$; the notation $\underset{\beta}{\arg min}$ refers to the determination of coefficients β minimizing the expression within the brackets; the symbol λ represents the shrinkage parameter; and γ represents the response term. The error term ε is assumed to have a normal distribution with a mean equal to zero and a variance equal to σ^2 (Heslot et al., 2012).

Bayesian Least Absolute Shrinkage and Selection Operator

Bayesian LASSO (BL) pulls weaker marker effects closer to zero much faster than in the case of RR-BLUP as the BL model uses a Laplace or double-exponential distribution as a prior instead of the Gaussian, normal distribution prior used for the Bayes-A, Bayes-B, and BRR models. However, it shrinks large effect markers less than that of RR-BLUP. In this case, all marker effects are assumed to be independently and identically distributed as Laplace, with the prior density as:

$$p(\beta|\lambda) = \frac{\lambda}{2}\exp(-\lambda|\beta|)$$

where $E(\beta|\lambda) = 0$ is the mean prior of the Laplace distribution and $Var(\beta|\lambda) = 2/\lambda^2$ is the variance for all markers. As the shrinkage parameter (λ) increases, the variance of the Laplace distribution decreases and gains a sharper density. β represents the individual marker effects (Gianola, 2013; Heslot et al., 2012; Park & Casella, 2008).

Weighted Bayesian Shrinkage Regression

Weighted Bayesian Shrinkage Regression (wBSR) is a variation of Bayes-B where an expectation maximization (EM) algorithm is used instead of an MCMC algorithm for model construction, as the MCMC algorithm can be very time-consuming when used with datasets with large sample sizes and marker sets. The EM algorithm is much faster than MCMC in that it uses the analytic form of posterior means of marker effects in order to identify the posterior mode of each marker effect instead of a posterior expectation, like in MCMC (Hayashi & Iwata, 2010). The standard equation for wBSR is essentially the same as the Bayes-A however, it also includes a parameter (γ_l) that accounts for the inclusion ($\gamma_l = 1$) or exclusion ($\gamma_l = 0$) of the l^{th} marker from

the model with prior probabilities of p and 1-p respectively, essentially making the model similar to Bayes-B:

$$y = \boldsymbol{X}\beta + \gamma u\boldsymbol{Z} + \varepsilon$$

where y is the numeric vector of phenotypic values for a trait; β is the vector of fixed, non-genetic effects; γ is the marker inclusion effect, the priors for marker effects (g_l) and the variance of marker effects ($\sigma^2_{g_l}$) are not impacted by the presence or absence of a marker in the model; $\boldsymbol{Z}$ is the marker matrix assigning marker genotypes to phenotypes; and ε is the residual error.

The EM algorithm used for wBSR consists of two primary steps, the *E*-step and the *M*-step. Both steps are repeated until the values of the model parameters reach convergence. The *E*-step is where the marker effect variance for each individual ($\sigma^2_{g_l}$) marker is estimated as ($\hat{\sigma}^2_{g_l}$) using the following formula:

$$\hat{\sigma}^2_{g_l} = \frac{g_l^2 + S}{v + 1}(l = 1,2, \dots, N)$$

where g_l is the marker effect of the l^{th} marker; S is the scale parameter; and v is the degrees of freedom. This equation serves as a conditional expectation given a current value of a maker (g_l) effect which is the estimated marker effect ($\hat{g}_l$) for l = 1, 2, …, N. In the *M*-step, values for individual marker effects, fixed non-genetic effects, and residual error variances that maximize the log posterior distribution of the estimates of the three aforementioned parameters are given, , where the value for each parameter is updated by replacing the other parameters with their current values. The *E* and *M* steps are repeated until the parameter values converge. The final parameter values are then used for the final model (Hayashi & Iwata, 2010). The main problem

with this method of parameter selection is that it is sometimes inadequate for certain datasets, resulting in poor results, therefore parameters set by the user are necessary (Heslot et al., 2012).

The wBSR model was compared with the GBLUP, Bayesian LASSO, Bayes-C, reproducing kernel Hilbert space (RKHS), and Random Forest models for predicting soluble solids content and total fruit weight of F_1 tomato (*Solanum lycopersicum* L.) plants, based on the parental combination that resulted in the best progeny and phenotypes of the F_1 progeny. A ten-fold cross-validation of the training population, consisting of 96 genotypes, found that the wBSR model had a significantly lower prediction accuracy than the GBLUP and Bayesian LASSO models, but was not significantly different from the Bayes-C and RKHS models, while significantly out-performing the Random Forest model for soluble solids content. The same cross-validation study for total fruit weight, with lower overall prediction accuracies than soluble solids content, found that the wBSR model had a significant lower prediction accuracy than the Random Forest, RKHS, GBLUP, and Bayes-C models, while the wBSR model had a significantly higher prediction accuracy than the Bayesian LASSO model. When looking at the overall prediction accuracies of parental combinations, the wBSR model had the second highest prediction accuracy for soluble solids content (r = .68), below the Bayes C model (r = .69). For total fruit size, all models had low prediction accuracies below .30, however, the wBSR model still had the second highest prediction accuracy (r = .27), below the Random Forest model (r = .28) (Yamamoto et al., 2017).

Bayes-Cπ

The Bayes-Cπ model was constructed in order to account for the lack of Bayesian learning in the Bayes-A and Bayes-B models, as they assume marker-specific variances and their priors will have too much influence on the marker variances (Gianola, de los Campos, Hill, Manfredi, &

Fernando, 2009; Habier, Fernando, Kizilkaya, & Garrick, 2011). Bayes-Cπ assumes that there is a common marker effect variance for all markers with a nonzero effect and estimates the probability of marker absence (π) or presence (1-π) instead of using a fixed value for π. The MCMC algorithm is used to calculate the posterior parameters for the model. The model for Bayes-Cπ is as follows:

$$y = \mu + \sum_{\{j=1\}}^{K} \boldsymbol{Z_j} a_{j\delta_j} + \varepsilon$$

where *y* is the vector of individual phenotypes; μ is the population mean; *K* is the total number of markers; $\boldsymbol{Z}$ is the incidence matrix linking marker *j* genotypes to individuals; *a* is the effect of marker *j*; δ is the indicator variable indicating the absence or presence of marker *j* in the model; and ε represents the model residual error. It is assumed that marker *j* has an effect of zero and a probability of π when $\delta_j = 0$. When $\delta_j = 1$, marker *j* has an effect of $a_j \sim N(0, \sigma_a^2)$ with a probability of 1-π. In the case of Bayes-Cπ, the value of π is unknown and is estimated from the training data (Habier et al., 2011; Heslot et al., 2012; Sallam, Endelman, Jannink, & Smith, 2015). When Bayes-Cπ was compared with ten other models using wheat, maize, and barley datasets, Bayes-Cπ performed very similarly to RR-BLUP. As a result, it was recommended that plant breeders use RR-BLUP instead, since Bayes-Cπ had an increased computing time in comparison (Heslot et al., 2012). When compared with RR-BLUP and the Gaussian and Exponential kernels, Bayes-Cπ was not significantly different in prediction accuracy from the other models when applied in a barley breeding program (Sallam et al., 2015).

Bayes-Dπ

The Bayes-Dπ model was originally described in the same publication that initially described the Bayes-Cπ model, however it has rarely been used for empirical studies compared to the latter (Habier et al., 2011). The Bayes-Dπ model, much like the Bayes-Cπ model, treats the probability of marker presence or absence as unknown instead of using a fixed value for π. The other main difference between Bayes-Dπ and Bayes-Cπ is that Bayes-Dπ assumes that each marker has its own separate variance, much like the Bayes-A and Bayes-B models, instead of a common marker effect variance for all markers. The general statistical model for Bayes-Dπ is the same as the statistical model for Bayes-Cπ, as discussed above. Since each marker has its own separate variance, the single marker effect prior is calculated as (a_j) with a distribution of $\sim N(0, \sigma^2_{a_j})$, where $\sigma^2_{a_j}$ denotes an individual marker effect variance as opposed to σ^2_a, where a common marker effect variance is assumed for all markers. Each marker variance has an inverse chi-square prior with degrees of freedom of v_a and a scale of S^2_a. When the probability for the presence of a marker is (1-π) the marginal prior of $a_j|v_a$, S^2_a is a univariate students-*t* distribution of $t(0, v_a, S^2_a)$. The degrees of freedom for the inverse chi-squared prior for locus specific variances are known while the scale parameter is treated as unknown with a gamma(1,1) prior in Bayes Dπ. If a marker is fitted with a probability of 1-π, its effect will come from a uniform student-*t* distribution. Therefore, the mixture is due to treating S^2_a as unknown with a gamma prior (Habier et al., 2011).

Bayes-Dπ has a significantly longer computing time compared to Bayes-Cπ, however it is still less computationally intensive than Bayes-A or Bayes-B. The primary reason the three latter models have longer computing times is due to the multiple sampling iterations required to improve mixing, while Bayes-Cπ only requires one iteration. The estimate for π in Bayes-Cπ is

more sensitive to training population size and marker density, providing more information about the genetic architecture of quantitative traits. The prediction accuracies of Bayes-Dπ and the other three Bayesian models were not significantly different from each other when used in a simulation study (Habier et al., 2011).

Empirical Bayes

The Empirical Bayes (E-Bayes) model was developed in order to estimate epistatic effects in an oversaturated model. The prior for each individual marker effect follows a normal distribution with a mean of zero and a marker effect variance of σ_j^2, which is assumed to have an inverse chi-squared distribution with a degree of freedom τ and a scale parameter ω. The E-Bayes model also does not use MCMC sampling in order to estimate variance parameters for σ_j^2, instead using an EM algorithm to reduce computation time (Heslot et al., 2012; Lorenzana & Bernardo, 2009; S. Z. Xu, 2007). The E-Bayes model is as follows:

$$y = \sum_{i=1}^{p} \boldsymbol{X_i}\, \beta_i + \sum_{j=1}^{q} Z_j\, \gamma_j + \varepsilon$$

where $\boldsymbol{X_i}$ is the design matrix of individuals (i); β_i is the vector of fixed effects of individuals (i); Z_j is the vector of genotype indicators for loci (j); γ_j is the main effect vector for loci (j); and ε is the residual error (S. Z. Xu, 2007).

When predicting the performance of biparental maize populations, the E-Bayes model did not show any significant advantages by including pairwise epistatic effects and prediction accuracies were no better than those calculated using RR-BLUP (Lorenzana & Bernardo, 2009). When comparing E-Bayes with ten other genomic selection models, overfitting was consistently observed with E-Bayes due to marker collinearity (Heslot et al., 2012).

Elastic Network

The elastic net is a penalized method that was designed to use the best features of both ridge regression and LASSO. The resulting model consists of a penalty that is a weighted average of the L1 and L2 norms. The L1 norm was derived from LASSO while the L2 norm was derived from ridge regression. In this case, the L1 norm has a penalty weight of α=1, while the L2 norm has a penalty weight of α=0. The L1 norm generates a sparse model while the L2 norm removes the limitation on the number of selected variables The elastic net model performs much like LASSO in that it simultaneously does automatic marker selection along with continuous shrinkage while also selecting groups of correlated variables, therefore selecting large effect markers (Heslot et al., 2012; Zou & Hastie, 2005). The elastic net model starts as a standard linear model:

$$y = \mu + \boldsymbol{X}\beta + \varepsilon$$

where *y* is the trait value, μ is the population mean, *X* is the marker design matrix, β is the vector of marker effects, and ε is the residual error. The elastic net uses an estimator formula to calculate β:

$$J(\beta) = (1 + \lambda_2) \underset{\beta}{\arg min} \left(\frac{\|\gamma - \boldsymbol{X}\beta\|_2^2}{2\sigma^2} + \lambda_2 \|\beta\|_2^2 + \lambda_1 \|\beta\|_1 \right)$$

where λ_1 and λ_2 are shrinkage parameters, the *arg min* notation refers to the determination of coefficients β minimizing the expression inside the brackets, while the notation $\|.\|_2$ represents the Euclidian norm of $\|\beta\|_2 = (\sum_i \beta_i^2)^{1/2}$ (Friedman, Hastie, & Tibshirani, 2010; Heslot et al., 2012; Zou & Hastie, 2005).

When the elastic net was compared to ten other models in predicting wheat, maize, and barley traits, it was significantly less accurate than the wBSR and RKHS models. The elastic net was also slightly less than RR-BLUP and the Bayesian LASSO models. Even so, the elastic net performed well overall, however it produced an extremely sparse model. This is not ideal since plant breeders generally want models with more markers present (Heslot et al., 2012). When tested on 12 different malting quality traits in barley against RR-BLUP and Bayesian LASSO, the elastic net did not perform any differently from the other models with regards to GEBVs or prediction accuracy (Schmidt et al., 2016).

Semi-parametric and Kernel Methods

Kernel methods, such as the Gaussian and exponential kernel, are not supposed to partition the total genetic variance into additive and non-additive variances, instead they are supposed to capture the additive and non-additive effects (Endelman, 2011; Sallam et al., 2015). The primary kernel model is as follows:

$$y = 1\mu + \boldsymbol{Z}g + \varepsilon$$

where μ is the population mean; $\boldsymbol{Z}$ is the incidence matrix linking markers to individuals; g is the vector of genotypic values with a distribution of $g \sim N(0, \boldsymbol{K}\sigma_g^2)$, where $\boldsymbol{K}$ is the kernel similarity matrix; and ε is the residual error (Endelman, 2011; Sallam et al., 2015).

The Gaussian kernel model is calculated as:

$$K_{ij} = \exp\left[-\left(D_{ij}/\theta\right)^2\right]$$

where θ represents the scale parameter influencing how quickly genetic covariance decays with distance and D_{ij} represents the Euclidian distance between genotypes *i* and *j*, normalized to the interval [0,1] which is calculated as: $D_{ij} = \left[(1/4\,M)\sum_{k=1}^{M}(\boldsymbol{G}_{ik} - \boldsymbol{G}_{jk})\right]^{1/2}$.

where $\boldsymbol{G}$ represents the genotypic matrix for genotypes *i* and *j*, respectively; *k* represents the k^{th} locus; and *M* represents the total number of markers (Endelman, 2011; Gianola & van Kaam, 2008).

The exponential kernel model is calculated as:

$$K_{ij} = exp(-D_{ij}/\theta)$$

where D_{ij} represents the Euclidian distance between genotypes *i* and *j*, normalized to the interval [0,1], which is calculated the same as above; and θ represents the scale parameter influencing how quickly genetic covariance decays with distance (Endelman, 2011; Piepho, 2009).

When comparing the Gaussian and exponential kernel methods with each other, there was little evidence that either result in significantly better prediction accuracies than the other (Endelman, 2011; Piepho, 2009). However, when compared to models such as RR-BLUP, the Gaussian kernel was between 6% and 7% more accurate when predicting wheat GY, possibly due to overfitting (Crossa et al., 2010; Endelman, 2011; Piepho, 2009).

Reproducing Kernel Hilbert Space

The RKHS model is a semi-parametric model that can capture both additive and non-additive interactions among loci by creating a kernel matrix that includes interactions among marker covariates (Gianola & van Kaam, 2008). The RKHS model essentially converts the marker dataset into a set of distances between pairs of observations resulting in a square matrix that can

be used in a linear model. RKHS does not assume linearity, therefore it can capture non-additive effects (Heslot et al., 2012). The standard model for RKHS is essentially a mixed linear model and is presented as follows:

$$y = \boldsymbol{W}\mu + \boldsymbol{K_h}\alpha + \varepsilon$$

where μ represents the vector of fixed effects; $\boldsymbol{W}$ is the design matrix for fixed effects; ε is the vector of random residuals with the independent prior distribution $\varepsilon \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$; α is the unknown coefficients with the independent prior distribution $\alpha \sim N(0, \boldsymbol{K_h}, \sigma_\alpha^2)$; $\boldsymbol{K_h}$ is the matrix that depends on the reproducing kernel function with the smoothing parameter *h*, measuring genomic distance between genotypes; and *h* is the smoothing parameter, which controls the rate of decay of the correlations between genotypes. The $\boldsymbol{K_h}$ kernel matrix uses a Gaussian kernel function, expressed as:

$$\boldsymbol{K_h}(X_i, X_j) = \exp(-hD_{ij}) = exp(-\theta D_{ij}/k)$$

where D_{ij} is the squared Euclidian or Manhattan distance between individuals *i* and *j*; h, the smoothing parameter, is calculated as $h = 2/D^*$, where D^* is the mean of D_{ij} (Gianola, Fernando, & Stella, 2006; Gianola & van Kaam, 2008; Heslot et al., 2012).

When compared with ten other Bayesian and BLUP genomic selection models using wheat, maize, and barley datasets, the RKHS often had much higher accuracy than other models and it performed significantly better than E-Bayes, elastic net, and the neural network. However, in the same analysis, RKHS overfitted more than other models; nevertheless the model was still highly accurate (Heslot et al., 2012). Another study in maize showed that RKHS outperformed Bayesian LASSO for traits where epistasis was more relevant, whereas it was outperformed for more additive traits (Crossa et al., 2010). In another study, the RKHS model performed similarly

to the EGBLUP model, but had a 5% higher prediction accuracy than the Bayes-Cπ and RR-BLUP models while also have a 17% lower standard deviation of prediction accuracies compared to the latter two models (He et al., 2016).

Random Forest

The random forest model is a machine learning method that uses a collection of regression trees that are grown on bootstrap samples of observations using a random subset of predictors. The random subset of predictors is then used to define the best split at each node. Different variables are used at each split in different trees. The random forest prediction for a given observations is computed from the average predictions over trees, where the given observation is not used to build the tree (Breiman, 2001; X. Chen & Ishwaran, 2012; Heslot et al., 2012).

Random forest consistently had the lowest prediction accuracies when compared with ridge regression, LASSO, elastic net, Bayes-A, Bayes-B, and RKHS models for seven drought tolerance traits in maize (Shikha et al., 2017). In another study, the random forest model was compared with ten other models for wheat, maize, and barley predictions and was observed to perform fairly well with higher prediction accuracies and low computing time (Heslot et al., 2012). The random forest model was also compared with Bayesian LASSO, Bayes-B, Bayes-Cπ, and RR-BLUP while predicting the performance of four agronomic traits in chickpea (*Cicer arietinum* L.). The model was observed to be the best fitting model across three location-years but was not significantly different from the others in terms of prediction accuracies (Roorkiwal et al., 2016).

Artificial Neural Networks

Artificial neural networks (NNET) are another form of machine learning model inspired by biological neural networks. NNETs are constructed of a multilayer perceptron, which is a system of simple interconnected nodes. The NNET can be trained to approximate nearly any smooth, measurable function. The general NNET formula is as follows:

$$\gamma_k = f_k\left[\alpha_k + \sum_j^U w_{jk} f_j\left(\alpha_j + \sum_i^N w_{ij} x_i\right)\right]$$

where w_{ij} is the weights linking the i and j nodes; α_j is the bias of a node specific constant; f_i is either a linear or nonlinear activation function that produces an output; x_i is an input variable for each marker (N), the input layer is defined as one neuron per input, data is then sent to hidden intermediate layers, and through an output layer made of one neuron per output variable; γ_k is the output variable in the form of a GEBV that receives output from the hidden intermediate layers; k is the output variable index; and U is the number of nodes within the hidden layers (Heslot et al., 2012).

The fitting of a NNET model is defined by the number of hidden intermediate layers, the number of neurons per layer, the type of activation function, and the connection weights. The best fitting model is determined by identifying the model with the lowest error from a multidimensional surface (Heslot et al., 2012; Hornik, 1993). When compared with ten other models in wheat, maize, and barley, the NNET had prediction accuracies close to the best models; however, it also had one of the highest average mean squared errors, likely attributed to the higher variance of its cross-validation GEBVs. The NNET was also one of the most over fitted models (Heslot et al., 2012).

Training Population Size

One of the important factors for implementing GS in a breeding program is determining the appropriate size of a training population. The goal for a breeding program would be to develop a training population that can maximize prediction accuracy while limiting population size as much as possible, largely due the expense of genotyping and phenotyping a large number of genotypes for the training population (Jannink et al., 2010).

When training population size was tested for four different yield traits in wheat, prediction accuracy increased as the number of genotypes increased from 250 to 2000; glaucousness increased from .78 to .92, GY increased from .72 to .85, thousand-kernel weight increased from .66 to .86, and relative maturity increased from .59 to .82. Even though prediction accuracy increased as the training population size increased, the gain in accuracy plateaued after 2000 genotypes (Norman et al., 2018). When evaluating prediction accuracies for nine different end-use quality traits in biparental wheat populations, increasing population size from 24 to 96 genotypes increased prediction accuracies as well (Heffner, Jannink, Iwata, et al., 2011). It has also been observed that deviations between prediction accuracies for different traits are larger under smaller training population sizes, likely due to the effect of relatedness between the training population and validation set (Poland et al., 2012). Smaller training populations also run the risk of overestimating the genotypic effect when predicting larger validation sets (Jannink et al., 2010). Training population sizes between 25 and 300 were evaluated for both wheat and rice for five and four agronomic traits, respectively. In wheat, prediction accuracies were highest at 300 genotypes for all five traits; however, there was evidence of a plateau after 300. In rice, prediction accuracies plateaued after 175 genotypes for florets per panicle and protein content, while there was actually a decrease in prediction accuracy for plant height and flowering time

after 150 genotypes (Isidro et al., 2015). More diverse training populations usually need to be larger in order to account for the larger genetic diversity, particularly with low heritability traits (Mujibi et al., 2011).

Trait Heritability

Generally, as heritability increases for a trait, so does the predictive ability for the trait (Jannink et al., 2010). The impact of heritability on prediction accuracy for training and validation sets was evaluated for 30 agronomic, quality, and stress tolerance traits in rice and three quality traits in maize. Overall, prediction accuracy improved as heritability increased in both the training and validations sets, with a greater impact in training sets (Z. G. Guo et al., 2014). Heritability was also evaluated for mixed populations of barley and wheat, along with biparental populations of maize and barley. In all cases, prediction accuracy increased as heritability in the training populations increased (Combs & Bernardo, 2013). A strong correlation between high heritability and stronger prediction accuracies was also observed when predicting quality traits in biparental wheat populations (Heffner, Jannink, Iwata, et al., 2011). Even with higher heritability, certain traits may have lower prediction accuracies as the trait could be controlled by a larger number of small-effect genes, such as grain yield, as opposed to a small number of large-effect genes, such as grain dry weight, as evidenced when evaluating a maize test-cross population (Albrecht et al., 2011).

Genetic Relationship between Training and Validation Sets

The genetic similarity between training populations and their validation sets can significantly influence prediction accuracy (Clark, Hickey, Daetwyler, & van der Werf, 2012; Jannink et al., 2010). More closely related individuals will often share a common ancestor in a relatively small

number of generations prior, allowing for less recombination events, preserving QTL and marker linkage-phases. Closely related individuals are also more likely to share polymorphic loci that generate genetic variation, leaving fewer opportunities for genetic drift or mutations (A. J. Lorenz & Smith, 2015). There are also interactions that can occur between QTL and genetic backgrounds, as closely related individuals are more likely to share a portion of their genetic backgrounds compared to distant individuals (A. J. Lorenz & Smith, 2015; K. Lorenz & Cohen, 2012; Mohammadi et al., 2015). When evaluating training populations for FHB resistance in barley, it was observed that prediction accuracy is lower when a training population is predicting the performance of barley genotypes from outside breeding programs (A. J. Lorenz, Smith, & Jannink, 2012). The same has also been observed when predicting across different full-sib families in biparental maize populations (Riedelsheimer et al., 2013). Training populations do not need to be as large when they are predicting the performance of closely related genotypes in the validation population. Along with that, genotypes in the training population do not need to have as large of a marker density when predicting the performance of closely related genotypes (Meuwissen, 2009).

Population Structure

Population structure is the composition of a population that are divided by genetic background, geography, or natural selection. As a result, it plays an important role in the estimation of GEBVs. Changes in allele frequencies between sub-populations can result in erroneous associations between markers and traits, leading to biased prediction accuracies for GS. A common indicator of population structure interference is when small training populations have high prediction accuracies (Windhausen et al., 2012). Several methods can be used to reduce the influence of population structure. These include the separation of breeding lines based on their

origins or traits, inclusion of kinship or genotype matrices, inclusion of marker fixed effects, or including principle components in an analysis based on subpopulations (Clark et al., 2012; Z. G. Guo et al., 2014; A. J. Lorenz et al., 2012). Another way to control for population structure is with epistatic genomic prediction models that can account for both additive and non-additive effects, such as the RKHS, Gaussian kernel, or exponential kernel models (Sallam et al., 2015). Another method for population structure partitioning is the K-means algorithm, where partitioning is based on genetic similarity (Norman et al., 2018). In the case of large training populations, another common source of population structure is LD. Particularly when two loci with differences in allele frequencies are in LD across subpopulations. This can result in spurious associations with multiple QTL (Pritchard, Stephens, Rosenberg, & Donnelly, 2000). When the effects of population structure were evaluated in a historic winter wheat nursery, two major subpopulations emerged based on the presence or absence of the *t2BS:2GS·2GL:2BL* translocation derived from *Triticum timopheevii*. However, when this translocation was used to design an optimal training population, it did not have significantly higher predictive ability than a randomly selected training population (Sarinelli et al., 2019).

Retraining and Training Population Optimization

Retraining of the training population is important because as multiple generations of selections are made, the generational difference between the training population and the validation set grows larger. Since the training population and the validation set become more genetically distant, the prediction accuracy is reduced. Therefore, it is important to add new genotypes to the training population in order to maintain a genetic relationship to the validation set as new germplasm is introduced to the breeding program and more recombination events occur (A. J. Lorenz & Smith, 2015). New germplasm must be added to breeding programs implementing GS

because the increased selection intensity from GS reduces the effective population size in a breeding program and thusly leading to a loss in genetic variability (Heffner et al., 2009).

Several different methods are used to optimize training populations in order to maximize prediction accuracy while also minimizing training population size, allowing breeders to better allocate resources. Among these are the PEVmean, CDmean, Gmean, Stratified Sampling (SS), and Selection of Training Populations by Genetic Algorithm (STPGA) algorithms which are based on genetic relations to the validation set and along with other candidates within the training population (Tiede & Smith, 2018).

The PEVmean algorithm identifies the optimal subset of *n* individuals from a total of *N* candidates for a training population by minimizing the predictive error variance (PEV) of contrasts between each of the entries in a training population subsample and the subsample mean (Tiede & Smith, 2018). The expected PEV for each individual is expressed as:

$$\begin{bmatrix} \boldsymbol{X'X} & \boldsymbol{X'Z} \\ \boldsymbol{Z'X} & \boldsymbol{Z'Z} + \lambda A^{-1} \end{bmatrix}^{-1} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

where $PEV(\hat{u}) = Var(\hat{u} - u) = diag(C_{22}) \times \sigma_e^2$. The PEV of any contrast *c* of the predicted performances is calculated as:

$$PEV = diag\left[\frac{\boldsymbol{c'}(\boldsymbol{Z'}M\boldsymbol{Z} + \lambda A^{-1})^{-1}\boldsymbol{c}}{\boldsymbol{c'c}}\right] \times \sigma_e^2$$

where $\boldsymbol{Z}$ is the design matrix; c is the matrix of contrasts comparing each column to the mean of all training population candidates; *M* is the orthogonal projector of the subspace spanned by the columns of $\boldsymbol{X}$, which is the design matrix of fixed effects: $M = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X'X})^{-}\boldsymbol{X'}$ (Rincent et al., 2012). The PEVmean is calculated using a genomic relationship matrix $\boldsymbol{G} = \frac{WW'}{f}$ where $W_{ik} =$

$X_{ik} - 2p_k$ is the mean centered marker *k* for individual *i, p_k* is the frequency of the one allele at marker *k* for the entire population and X_{ik} denotes the number of minor alleles for the i^{th} individual at marker *k*. The mean of the diagonal elements is 1+*f* using the normalization constant $f = 2\sum_k p_k(1 - p_k)$ (Isidro et al., 2015).

The other algorithm proposed by Rincent et al. (2012) was the CDmean algorithm, where the coefficient of determination (CD) is maximized, resulting in more reliable predictions than the PEVmean algorithm (Isidro et al., 2015; Rincent et al., 2012; Tiede & Smith, 2018). CDmean looks to maximize the variance explained by the training population while only using phenotypic and genotypic information from the training population candidates while identifying the ideal training population subset. A random subset of *n* lines is selected from the training population and the algorithm exchanges existing genotypes in the subset with candidates not currently in the subset. If the exchange between the subset and the training population increases the trace of the following equation:

$$CD(\boldsymbol{c}) = diag\left[\frac{\boldsymbol{c}'(A - \lambda(\boldsymbol{Z}'M\boldsymbol{Z} + \lambda A^{-1})^{-1})\boldsymbol{c}}{\boldsymbol{c}'A\boldsymbol{c}}\right]$$

The new candidate is then retained in the subset and the other genotype will remain in the un-sampled training population. In the above equation, $\boldsymbol{Z}$ is the design matrix; **c** is the matrix of contrasts comparing each column to the mean of all training population candidates; *M* is an orthogonal projector on the subspace spanned by the columns of **X**, which is the design matrix of fixed effects: $M = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$; *A* is calculated as $A = \frac{(C^* s^{-1})C\prime}{m}$, where $\boldsymbol{C} = M_{TP} - \boldsymbol{P}$, in which M_{TP} is the training population genotype matrix and $\boldsymbol{P}$ is the $nTP \times m$ matrix, $s = p(1 - p)$, in which *p* is a vector of length *m* (Isidro et al., 2015; Rincent et al., 2012; Tiede & Smith, 2018).

The SS algorithm randomly samples proportional to a subpopulation's size within the training population. Stratified sampling seeks to maximize variance explained by the training population, requiring genotypic and phenotypic data, much like CDmean (Isidro et al., 2015; Tiede & Smith, 2018). Training population candidates are assigned to one of three subpopulations defined by passing a Euclidean distance matrix calculated from marker genotypes of training population candidates to a hierarchal clustering algorithm. The training populations are actually built by combining individuals (n_c) from each of the aforementioned clusters (c) where $n_c \alpha \frac{nTP_c}{nTP}$, where nTP_c is the number of genotypes from the training population assigned to the cluster in which $n = \sum_1^c n_c$ (Tiede & Smith, 2018).

The Gmean algorithm uses only genotypic data in order to identify an ideal training population subset. The goal is to select *n* individuals from the training population that have the highest average genetic relationship and add them to the pool of selection candidates. More candidates are then added, starting with the most related candidates where $r_{G,\hat{G}}$ increased with each addition of a new, less-related, candidate. Eventually, the addition of new candidates becomes detrimental and the process stops (A. J. Lorenz et al., 2012; Tiede & Smith, 2018).

The STPGA method is a more predictive training subset of size *n* in that it minimizes ridge regression PEV among the selection candidates by using phenotypic and genotypic data from the training population. The formula for STPGA is as follows:

$$PEV^{Ridge}(\boldsymbol{M_{VP}}) = \boldsymbol{M_{VP}}(\boldsymbol{M'_{TP}M_{TP}} + \lambda_{STPGA} I)^{-1}\boldsymbol{M'_{VP}}$$

where $\boldsymbol{M_{TP}}$ is the genotype matrix of the training population; $\boldsymbol{M_{VP}}$ is the genotype matrix of the validation population; and $\lambda=1/m$ which is roughly equal to a trait heritability of .50. The $\boldsymbol{M_{TP}}$

and $\boldsymbol{M_{VP}}$ matrices can also be replaced with principle component matrices in order to reduce computing time (Akdemir, Sanchez, & Jannink, 2015; Tiede & Smith, 2018).

Previous studies have compared these different training population optimization algorithms across different traits and crops. When the CDmean, SS, Gmean, and STPGA algorithms were tested on a barley training population in order to compare prediction accuracies for GY and deoxynivalenol (DON) accumulation, all four significantly improved prediction accuracy for both traits compared to a randomly selected training set. The Gmean algorithm outperformed the other algorithms while predicting grain yield, while the SS algorithm was the top performing algorithm for predicting DON accumulation (Tiede & Smith, 2018). PEVmean and CDmean were also compared with each other in two empirical studies using two different maize diversity panels. CDmean was the more reliable algorithm as it considered the reduction of variance due to relatedness between individuals (Rincent et al., 2012). In a recent study in winter wheat, an overall increase in prediction accuracy was observed for evaluated traits under small population sizes (between 50 and 150 individuals) when a PEVmean algorithm was implemented for selection, compared to random and clustering selection approaches (Sarinelli et al., 2019).

Genomic Selection in Multiple Environments

A vast majority of the applications for GS is in single environments and most genomic prediction models do not have the predictive power to make selections across multiple environments or take genotype by environment interactions into consideration (Burgueno, de los Campos, Weigel, & Crossa, 2012). Even so, genotype by environment interactions play a large role in plant breeding (Burgueno, Crossa, Cotes, San Vicente, & Das, 2011). The first multiple environment models incorporated a factor analytic structure in order to model for genotype by environment

interactions. Multiple environment models with pedigree and marker information performed better than models without pedigree or with only one of the components when making predictions for multiple environments in wheat (Burgueno et al., 2012). Incorporation of weather data and weather covariates has also been beneficial for multiple environment GS as well, improving prediction accuracy by nearly 11% in winter wheat (Heslot, Akdemir, Sorrells, & Jannink, 2014). The problem with incorporating marker and environmental covariate interactions in a multiple environment model is that the number of interactions becomes so large that the ability to model such interactions is impossible, especially when using a large number of markers and environmental covariates. The use of variance components can help to reduce the computational demand of modeling marker and environmental component interactions (Jarquin et al., 2014). When comparing marker by environment GBLUP models with models that ignored a genotype by environment interaction, the marker by environment GBLUP model significantly outperformed the naïve model over three wheat datasets in seven different irrigated and dryland environments (Lopez-Cruz et al., 2015). In another study, a genotype by environment (GxE) GBLUP model and a standard GBLUP model were tested over 35 location years. The standard GBLUP model outperformed the GxE model when predicting the performance of new environments, however modeling for genotype by environment interactions improved the overall predictive ability (Lado, Barrios, Quincke, Silva, & Gutierrez, 2016).

Multivariate Genomic Selection

Most GS models only predict for single traits, however predicting for multiple traits can be advantageous, especially when looking at multiple yield components, quality traits, disease resistance traits, and abiotic stress tolerance traits. In the past, multivariate methods have been developed for QTL mapping (Banerjee, Yandell, & Yi, 2008; C. J. Jiang & Zeng, 1995; C. W.

Xu, Wang, Li, & Xu, 2009). Genomic prediction models have also been developed to predict multiple traits in dairy bulls. The three models are variations on the GBLUP, Bayes SVSS, and Bayes-Cπ models, where the Bayes SVSS outperformed the Bayes-Cπ and GBLUP models (Calus & Veerkamp, 2011). When comparing predictive ability of GBLUP, Bayes-A, and Bayes-Cπ multivariate models on a pine (*Pinus* L.) breeding dataset, the Bayesian models outperformed GBLUP under a major QTL structure and the multivariate models strongly outperformed the single trait models. Under a polygenic genetic architecture, the three multi-trait models performed roughly the same and barely outperformed the single trait models. Multivariate models tend to work better when predicting traits that are genetically correlated with each other (Jia & Jannink, 2012). When predicting cassava (*Manihot esculenta* Crantz) performance, multiple-trait models outperformed single trait models by nearly 40% in prediction accuracy (Okeke, Akdemir, Rabbi, Kulakow, & Jannink, 2017). In American cranberry (*Vaccinium macrocarpon* Ait), a multivariate GBLUP model outperformed a standard, single trait GBLUP model in scenarios with medium to high genetic correlation between traits, however there was little difference between models when genetic correlation between traits was low (Covarrubias-Pazaran et al., 2018). A training population of 557 wheat genotypes were evaluated for GY and three proximal or remote sensing traits and genomic prediction models were tested using a univariate model for only grain yield and a multivariate model using all four traits. Multivariate model prediction accuracies for grain yield outperformed univariate models by 70%, on average. This indicated that the use of proximal and remote sensing data as secondary traits in genomic prediction models could improve prediction accuracy for GY (J. Rutkoski et al., 2016).

Use of Markers as Fixed Effects for Genomic Selection

Traditionally, QTL mapping was performed using QTL studies, which consisted of biparental mapping populations consisting of recombinant inbred lines, F_2 lines, or backcrosses. All of the individuals within the mapping population were then genotyped and phenotyped for a trait of interest, where the resulting data would be analyzed using linkage mapping in order to identify QTL associated with the trait of interest. The problem with QTL mapping is that the resolution is often low, due to low recombination within the population and small population sizes. Then, when QTL have successfully been identified within a single biparental mapping population, markers within LD of the QTL could be used to select for the trait of interest. The problem is that it is difficult to validate QTL across mapping populations, often making it nearly impossible to implement MAS using markers developed for another breeding program (Flint-Garcia et al., 2005; Heffner et al., 2009).

An alternative to QTL mapping is the genome-wide association study (GWAS), which can be used to identify significant marker trait associations (MTA) between genome wide SNP markers and individual traits. This method relies on LD and ancestral recombination events in natural populations (Myles et al., 2009; Rafalski, 2002). Genome-wide association studies can be performed on a wide array of diverse genotypes already existing in a population (Breseghello & Sorrells, 2006; Flint-Garcia et al., 2005). The problem with some GWAS models is that they can often fail to account for false positives due to population structure and relationships between individuals in the association mapping panel (Finno, Aleman, Higgins, Madigan, & Bannasch, 2014). Using population structure and other relationship parameters as covariates can reduce these false positives; however, this can result in more false negatives and longer computational times (M. Huang, Liu, Zhou, Summers, & Zhang, 2019). In that regard, many GWAS models

that can help control for false positives or negatives when identifying MTAs have been implemented. Some of these models include the general linear model (GLM), mixed linear model (MLM), efficient mixed-model association (EMMA), compressed mixed linear model (CMLM), multiple locus mixed linear model (MLMM), fixed and random model circulating probability unification (FarmCPU), and Bayesian-information and linkage-disequilibrium iteratively nested keyway (BLINK) models (M. Huang et al., 2019; Lipka et al., 2012; Liu, Huang, Fan, Buckler, & Zhang, 2016).

Most QTL discovered through QTL mapping and GWAS related to quantitative traits, such as GY, disease resistance, and quality, have not been successful for traditional MAS. This is largely due the fact that many small-effect genes are controlling variation in these traits (R. Bernardo, 2008). While significant MTAs identified through GWAS for these traits may not be helpful regarding MAS, they can still be applied for GS (Bian & Holland, 2017; Cericola et al., 2017; J. E. Spindel et al., 2016). Most simple genomic prediction models, such as RR-BLUP or GBLUP, assume that most markers have small genetic effects; as opposed to GWAS, which assumes there are multiple markers contributing to a larger amount of genetic variation. Both methods are incorrect in their assumptions (Bian & Holland, 2017; Gianola & van Kaam, 2008). As discussed above, more complex Bayesian and non-parametric models were developed in order to account for major-effect QTL using shrinkage algorithms, and variable selection using multiple distributions for marker effects and tuning parameters (Habier et al., 2011; Meuwissen et al., 2001). The problem with more complex models is that they are often more computationally intensive and do not work well with larger marker datasets. Parameter estimates can also be overly sensitive to priors and often do not translate to other studies (Gianola et al., 2009).

An alternative is to use markers linked to major genes associated with traits of interest as fixed effects, particularly when the marker controls for greater than 10% of the genetic variation for the trait of interest (R. Bernardo, 2014). Several GS experiments have shown that this strategy can improve prediction accuracies in wheat and maize (R. Bernardo, 2014; Mason et al., 2018; Owens et al., 2014; Rice & Lipka, 2019; J. E. Rutkoski et al., 2014; Sarinelli et al., 2019). Another proposed method is using only GWAS data from previously published GWAS studies as marker covariates, where the inclusion of GWAS data with a GBLUP model improved prediction accuracies for two out of three traits in cattle and nine out of 11 traits in rice over a Bayes-B model (Zhang et al., 2014). Other studies have included GWAS data from the training populations themselves, otherwise referred to as GS + de novo GWAS (GS+GWAS) (J. E. Spindel et al., 2016).

A GS+GWAS analysis was performed in a tropical rice breeding program using the RR-BLUP model with significant markers from a GWAS as fixed effects. The GS+GWAS model outperformed the RR-BLUP with historic GWAS data, RR-BLUP, Bayesian LASSO, RKHS, Random Forest, and multiple linear regression models for all three traits; however, the GS+GWAS model had a significantly higher prediction accuracy for plant height (J. E. Spindel et al., 2016). A GS+GWAS model was used on a maize nested association mapping (NAM) population, making predictions for three traits; one was a highly polygenic trait, plant height, whereas the other two were moderately polygenic disease resistance traits. The GS+GWAS model performed significantly better for the moderately polygenic disease resistance traits, whereas there was no significant difference in prediction accuracy between the GS+GWAS models and the standard models for plant height (Bian & Holland, 2017). When predicting powdery mildew resistance in winter wheat, seven different training populations were developed

ranging in size from 50 to 350 genotypes, A GWAS was performed on each training population and only the most significant MTA was selected as a fixed effect for the RR-BLUP model. The GS+GWAS model significantly outperformed the standard RR-BLUP model for prediction accuracy at each training population size (Sarinelli et al., 2019). The inclusion of significant markers in a model, however, does not always result to improved predictions. In a recent simulation study in maize and sorghum (*Sorghum bicolor* (L.) Moench), for instance, no significant increase or a decrease in prediction accuracy was observed for a majority of the traits evaluated when GWAS-derived SNPs were included as fixed effects in the GS model. Model performance should therefore be explored on a trait-by-trait basis before its implementation in the breeding program (Rice & Lipka, 2019).

Genomic Selection for Wheat Grain Yield

Improving GY is a major goal for any wheat breeder, as it is an economic trait directly affecting farmers while also contributing to global food security. As genetic gains in wheat GY have stagnated recently, wheat breeders have looked at GS in order to increase genetic gain (Ray et al., 2013). Previously, GS was implemented on 2325 European winter wheat genotypes wherein four genomic prediction models, RR-BLUP, Bayes Cπ, RKHS, and EGBLUP were compared for prediction accuracy. The epistatic models (RKHS) outperformed the additive models, indicating that accounting for epistasis improved prediction accuracy, particularly in larger populations. Forward selection was also successfully implemented with limited decrease in prediction accuracy from one year to the next (He et al., 2016). In another study, prediction accuracy was evaluated for grain yield using RR-BLUP over five independent breeding cycles in a winter wheat breeding program using 659 inbred lines. Over the five cycles, GY prediction accuracy was $r = .38$, however after outlier cycles were removed, accuracy improved to $r = .41$. Removal

of outlier environments did not have a significant impact on prediction accuracy (Michel et al., 2016). Genomic selection was also implemented into the preliminary yield trial (PYT) stage of a wheat breeding program and compared with traditional phenotypic selection and genomic assisted selection (GAS), where breeding values from the PYT stage were combined with GEBVs using a heritability index. The resulting prediction accuracies for all three methods for GY were .39, .33, and .48 respectively. This indicated that the use of heritability indices in order to supplement GS can increase prediction accuracy for GY in wheat (Michel et al., 2017). Different training population sizes were evaluated for the prediction of GY in winter wheat. It was observed that there was a high predictive ability using RR-BLUP when a training population was selected using PEVmean for GY. There was no significant improvement in prediction accuracy as training population size increased above 50 genotypes, with the exception of a reduction in prediction accuracy at 100 genotypes. Prediction accuracies for GY averaged .64 with a training population size of 350 genotypes, selected using PEVmean (Sarinelli et al., 2019).

Genomic Selection for Fusarium Head Blight Resistance

Fusarium head blight is a destructive fungal disease that negatively affects wheat production worldwide, resulting in heavy losses in GY and quality. In the United States, the most common causal pathogen is *Fusarium graminearum*, which is part of the phylum *Ascomycota* (Milus & Parsons, 1994; X. M. Xu & Nicholson, 2009). Symptoms on infected wheat plants include premature bleaching of individual spikelets shortly after flowering, eventually progressing through the entire spike, resulting in a fully bleached spike. Signs of *F. graminearum* are usually present during warm and moist conditions; they appear as salmon colored sporodochia on the rachis and glumes of the spikelets. Blue-black spherical perithecia, which serve as sexual structures for the pathogen, will often appear later in the growing season. As symptoms progress,

the fungal mycelium will colonize the wheat kernels as they develop, resulting in shriveled grain with a pink to light-pink coloration. These damaged kernels are often referred to as tombstone-kernels or Fusarium damaged kernels (FDK) (Goswami & Kistler, 2004; Wegulo, Baenziger, Nopsa, Bockus, & Hallen-Adams, 2015).

The *F. graminearum* pathogen produces the mycotoxin, DON, in order to disable natural plant defenses. Deoxynivalenol is considered a vomitoxin because it disrupts the digestive function of humans and animals that consume infected grain, resulting in nausea, headaches, vomiting, and in extreme cases, death. Levels of DON in human food should not exceed 1 ppm, however grain infected with FHB can exceed 20 ppm (Pestka, 2010). Any grain that does not meet the meet the DON limits set by either the USDA or individual buyers can have the value of their grain docked or rejected entirely, causing significant economic damages for farmers (Dahl & Wilson, 2018).

Wheat breeders have worked to develop varieties that have genetic resistance to FHB. Fusarium head blight resistance is a quantitative trait that is significantly influenced by genotype by environment interactions (Steiner et al., 2017). There are generally two major sources for FHB resistance in the United States, either through 'exotic' or 'native' germplasm. One of the most influential exotic genotypes, 'Sumai-3' was identified in China (Buerstmayr, Ban, & Anderson, 2009). There are several major effect QTL that have been validated for FHB resistance. The first three identified include *Fhb1*, *Fhb2*, and *Qfhs.ifa-5A*, found in Sumai-3 (Anderson et al., 2001; Buerstmayr et al., 2002; Buerstmayr et al., 2003; Waldron, Moreno-Sevilla, Anderson, Stack, & Frohberg, 1999). The next two major QTL, *Fhb4* and *Fhb5* were found in the genotype 'Wangshiubai' (S. Xue et al., 2010; S. L. Xue et al., 2011). The QTL *Qfhs.nau-2DL* was then found in the breeding line CJ9306 (G. L. Jiang, Dong, Shi, & Ward,

2007; G. L. Jiang, Shi, & Ward, 2007). The most recently discovered QTL was *Fhb7* found in *Thinopyrum ponticum* (Podp.) Z.-W.Liu & R.-C.Wang (J. Guo et al., 2015; H. Wang et al., 2020). The two primary forms of FHB resistance exist in field conditions are Type I resistance, also known as incidence (INC), the resistance to initial infection by *F. graminearum*, and Type II resistance, the resistance to fungal spread within the infected spike, also known as severity (SEV) (Steiner et al., 2017). The *Fhb1* and *Qfhs.nau-2DL* QTL primarily confer Type II resistance while also reducing DON accumulation. Whereas *Qfhs.ifa-5A* primarily confers Type I resistance and reduces DON accumulation (Ban, 2000; Buerstmayr & Lemmens, 2015; Mesterhazy, 1995; Schroeder & Christensen, 1963).

As FHB resistance in wheat is a complex, quantitative trait, it is an ideal candidate for GS. A wheat panel of 322 genotypes was used to predict the performance of six traits associated with FHB resistance. Four separate genomic prediction models were compared, including RR-BLUP, Bayesian LASSO, RKHS, and Random Forest. Two different marker sets were also compared, the whole genome marker set and a marker set targeted to markers associated with FHB resistance. The random forest and RKHS models had the highest prediction accuracies for most of the traits. In the case of DON, the random forest model plus the targeted and whole-genome markers had the highest prediction accuracy (J. Rutkoski et al., 2012). A training population of 470 winter wheat genotypes were assessed for FHB resistance prediction accuracy. The study compared the use of a full marker dataset compared to a subset of markers associated with FHB resistance. The use of marker subsets significantly improved accuracy, as did the use of training population subsets (Hoffstetter, Cabrera, Huang, & Sneller, 2016). Another study for predicting FHB resistance using six traits in wheat tested three models: RR-BLUP, elastic net, and LASSO. Marker densities were also tested between 500 and 4500 SNPs along with training

population sizes between 96 and 2018 genotypes. Overall, there was a high prediction accuracy for FHB resistance. The top performing model was RR-BLUP, while the top marker sets ranged between 1500 and 3000 SNPs, while the optimum training population was less than 192 genotypes, depending on the trait (Arruda et al., 2015). Another study on spring wheat genotypes in the Pacific Northwest United States used a population of 170 genotypes. It was discovered that INC had the highest prediction accuracies and genotypes that were similarly related to the initial training population had a prediction accuracy of 60%. The prediction accuracy was fairly high for FHB resistance overall, suggesting that genomic selection could work well for FHB resistance (Dong et al., 2018).

APPROACH TO THE CURRENT STUDY

The overall objective for this study was to successfully develop and implement a GS strategy for the University of Arkansas wheat breeding program for FHB resistance, agronomic traits, and GY. In order to fulfill this objective, a population was developed to be used as a training population for GS and as a mapping panel for GWAS. A GWAS was performed on the training population to determine significant MTAs for four FHB resistance traits. Cross-validation analyses were also performed to compare naïve GBLUP models (NGS) with GS+GWAS and multivariate GS (MVGS) models for FHB resistance and MVGS models for agronomic traits. Grain yield data from correlated environments were also used as covariates in GS models in order to predict GY for missing environments using a sparse testing approach. Genomic prediction accuracy was determined by calculating GEBVs and comparing them to phenotypic values for each genotype for each trait. The training population was then retrained and used to predict the performance of three generations of breeding lines in the University of Arkansas wheat breeding program. Genome-estimated breeding values calculated from the prediction

model were then compared to the actual phenotypic performance of each breeding line. The specific objectives are as follows:

Objective 1: Identify significant MTAs for four FHB resistance traits using a GWAS, and perform GS cross validation analyses to determine prediction accuracies for each trait using NGS, GS+GWAS, and MVGS models. This was accomplished using a FarmCPU GWAS to identify significant MTAs. The hypothesis was that significant loci controlling for variation in FHB resistance were distributed across multiple chromosomes. The GS+GWAS analyses were performed by randomly generating ten training populations and performing GWAS on them to identify markers significantly associated with each FHB resistance trait. The training populations were then used to train a GBLUP model with significant markers as fixed effects in order to calculate GEBVs for the genotypes in the validation population. The hypothesis was that the GS+GWAS models would have higher mean prediction accuracies than the NGS models for each trait of interest. The MVGS analyses were performed using a five-fold cross validation analysis over 100 iterations while using genetically correlated traits as covariates to predict the FHB resistance traits of interest. The hypothesis was that the MVGS models would have higher mean prediction accuracies than the NGS models for each trait of interest.

Objective 2: Predict GY and four agronomic traits through cross validation between NGS and MVGS models. Predict GY in missing environments through sparse testing using GY data from similar environments as covariates in MVGS models through cross validation. This was accomplished by using GBLUP models to predict a trait of interest while including data for traits sharing strong genetic correlations with the trait of interest as covariates in the MVGS model. Mean prediction accuracies were determined by performing five-fold cross validation analyses over 100 iterations. The hypothesis was that the MVGS models would have higher

mean prediction accuracies than the NGS models for each trait of interest. In order to compare sparse testing models with NGS models, GY data from six environments were clustered based on similarity and used as covariates to predict GY for genotypes not grown in a missing environment. Mean prediction accuracies were calculated using a five-fold cross validation approach over 100 iterations. The hypothesis was that the sparse testing models would have higher mean prediction accuracies for GY at each environment compared to NGS models.

Objective 3: Validate MVGS model performance by predicting GEBVs of three generations of breeding lines for GY and three FHB resistance traits using the same training population from objectives one and two for the first generation, then retraining in subsequent years by adding new genotypes. The GEBVs were then compared to phenotypic results for each genotype based on prediction accuracy, response to selection, and selection accuracy. The same training population used in objectives one and two will be used to predict the performance of breeding lines in the University of Arkansas wheat breeding program over the $F_{4:6}$, $F_{4:7}$, and $F_{4:8}$ generations. The GEBVs for GY and three FHB resistance traits listed in objectives one and two will be calculated for each $F_{4:6}$ breeding line using a NGS BLUP model and MVGS model using optimal trait-covariate combinations determined from the first two objectives. The effects of retraining the training population will also be observed. The training population will be retrained at each generation by adding new genotypes from the previous year to the training population. Phenotypic selection at the $F_{4:6}$ generation will be compared with NGS and MVGS by evaluating prediction accuracy, selection accuracy, and response to selection over three years.

REFERENCES

Akbari, M., Wenzl, P., Caig, V., Carling, J., Xia, L., Yang, S. Y., . . . Kilian, A. (2006). Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theoretical and Applied Genetics, 113*(8), 1409-1420. doi:10.1007/s00122-006-0365-4

Akdemir, D., Sanchez, J. I., & Jannink, J. L. (2015). Optimization of genomic selection training populations with a genetic algorithm. *Genetics Selection Evolution, 47*, 10. doi:10.1186/s12711-015-0116-6

Albrecht, T., Wimmer, V., Auinger, H. J., Erbe, M., Knaak, C., Ouzunova, M., . . . Schon, C. C. (2011). Genome-based prediction of testcross values in maize. *Theoretical and Applied Genetics, 123*(2), 339-350. doi:10.1007/s00122-011-1587-7

Anderson, J. A., Stack, R. W., Liu, S., Waldron, B. L., Fjeld, A. D., Coyne, C., . . . Frohberg, R. C. (2001). DNA markers for Fusarium head blight resistance QTLs in two wheat populations. *Theoretical and Applied Genetics, 102*(8), 1164-1168. doi:10.1007/s001220000509

Arruda, M. P., Brown, P. J., Lipka, A. E., Krill, A. M., Thurber, C., & Kolb, F. L. (2015). Genomic selection for predicting Fusarium head blight resistance in a wheat breeding program. *Plant Genome, 8*(3), 12. doi:10.3835/plantgenome2015.01.0003

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. doi:10.2135/cropsci2012.09.0526

Ban, T. (2000). Analysis of quantitative trait loci associated with resistance to Fusarium head blight caused by *Fusarium graminearum* Schwabe and of resistance mechanisms in wheat (*Triticum aestivum* L.). *Breeding Science, 50*(2), 131-137. doi:10.1270/jsbbs.50.131

Banerjee, S., Yandell, B. S., & Yi, N. J. (2008). Bayesian quantitative trait loci mapping for multiple traits. *Genetics, 179*(4), 2275-2289. doi:10.1534/genetics.108.088427

Barrett, B. A., & Kidwell, K. K. (1998). AFLP-based genetic diversity assessment among wheat cultivars from the Pacific Northwest. *Crop Science, 38*(5), 1261-1271. doi:10.2135/cropsci1998.0011183X003800050025x

Bernardo, R. (1994). PREDICTION OF MAIZE SINGLE-CROSS PERFORMANCE USING RFLPS AND INFORMATION FROM RELATED HYBRIDS. *Crop Science, 34*(1), 20-25. doi:10.2135/cropsci1994.0011183X003400010003x

Bernardo, R. (2001). What If We Knew All the Genes for a Quantitative Trait in Hybrid Crops? *Crop Science, 41*(1), 1-4. doi:10.2135/cropsci2001.4111

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: Learning from the last 20 years. *Crop Science, 48*(5), 1649-1664. doi:10.2135/cropsci2008.03.0131

Bernardo, R. (2014). Genomewide selection when major genes are known. *Crop Science, 54*(1), 68-75. doi:10.2135/cropsci2013.05.0315

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

Bian, Y., & Holland, J. B. (2017). Enhancing genomic prediction with genome-wide association studies in multiparental maize populations. *Heredity, 118*(6), 585-593. doi:10.1038/hdy.2017.4

Bond, J. K. (2020). Wheat Sector at a Glance. Retrieved from https://www.ers.usda.gov/topics/crops/wheat/wheat-sector-at-a-glance/#classes

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32. doi:10.1023/a:1010933404324

Breseghello, F., & Sorrells, M. E. (2006). Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Science, 46*(3), 1323-1330. doi:10.2135/cropsci2005.09-0305

Broman, K. W., & Speed, T. P. (2002). A model selection approach for the identification of quantitative trait loci in experimental crosses. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 64*, 641-656. doi:10.1111/1467-9868.00354

Buerstmayr, H., Ban, T., & Anderson, J. A. (2009). QTL mapping and marker-assisted selection for Fusarium head blight resistance in wheat: a review. *Plant Breeding, 128*(1), 1-26. doi:10.1111/j.1439-0523.2008.01550.x

Buerstmayr, H., & Lemmens, M. (2015). Breeding healthy cereals: genetic improvement of Fusarium resistance and consequences for mycotoxins. *World Mycotoxin Journal, 8*(5), 591-602. doi:10.3920/wmj2015.1889

Buerstmayr, H., Lemmens, M., Hartl, L., Doldi, L., Steiner, B., Stierschneider, M., & Ruckenbauer, P. (2002). Molecular mapping of QTLs for Fusarium head blight resistance in spring wheat. I. Resistance to fungal spread (type II resistance). *Theoretical and Applied Genetics, 104*(1), 84-91. doi:10.1007/s001220200009

Buerstmayr, H., Steiner, B., Hartl, L., Griesser, M., Angerer, N., Lengauer, D., . . . Lemmens, M. (2003). Molecular mapping of QTLs for Fusarium head blight resistance in spring wheat. II. Resistance to fungal penetration and spread. *Theoretical and Applied Genetics, 107*(3), 503-508. doi:10.1007/s00122-003-1272-6

Burgueno, J., Crossa, J., Cotes, J. M., San Vicente, F., & Das, B. (2011). Prediction Assessment of Linear Mixed Models for Multienvironment Trials. *Crop Science, 51*(3), 944-954. doi:10.2135/cropsci2010.07.0403

Burgueno, J., de los Campos, G., Weigel, K., & Crossa, J. (2012). Genomic Prediction of Breeding Values when Modeling Genotype x Environment Interaction using Pedigree and Dense Molecular Markers. *Crop Science, 52*(2), 707-719. doi:10.2135/cropsci2011.06.0299

Calus, M. P. L., & Veerkamp, R. F. (2011). Accuracy of multi-trait genomic selection using different methods. *Genetics Selection Evolution, 43*, 14. doi:10.1186/1297-9686-43-26

Cericola, F., Jahoor, A., Orabi, J., Andersen, J. R., Janss, L. L., & Jensen, J. (2017). Optimizing Training Population Size and Genotyping Strategy for Genomic Prediction Using Association Study Results and Pedigree Information. A Case of Study in Advanced Wheat Breeding Lines. *Plos One, 12*(1), 20. doi:10.1371/journal.pone.0169606

Chao, S. M., Dubcovsky, J., Dvorak, J., Luo, M. C., Baenziger, S. P., Matnyazov, R., . . . Akhunov, E. D. (2010). Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (Triticum aestivum L.). *Bmc Genomics, 11*, 17. doi:10.1186/1471-2164-11-727

Chao, S. M., Zhang, W. J., Dubcovsky, J., & Sorrells, M. (2007). Evaluation of genetic diversity and genome-wide linkage disequilibrium among US wheat (Triticum aestivum L.) germplasm representing different market classes. *Crop Science, 47*(3), 1018-1030. doi:10.2135/cropsci2006.06.0434

Chen, J. M., & Gustafson, J. P. (1995). PHYSICAL MAPPING OF RESTRICTION-FRAGMENT-LENGTH-POLYMORPHISMS (RFLPS) IN HOMOEOLOGOUS GROUP-7 CHROMOSOMES OF WHEAT BY IN-SITU HYBRIDIZATION. *Heredity, 75*, 225-233. doi:10.1038/hdy.1995.130

Chen, X., & Ishwaran, H. (2012). Random forests for genomic data analysis. *Genomics, 99*(6), 323-329. doi:10.1016/j.ygeno.2012.04.003

Clark, S. A., Hickey, J. M., Daetwyler, H. D., & van der Werf, J. H. J. (2012). The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genetics Selection Evolution, 44*, 9. doi:10.1186/1297-9686-44-4

Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., & Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica, 142*(1-2), 169-196. doi:10.1007/s10681-005-1681-5

Collard, B. C. Y., & Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B-Biological Sciences, 363*(1491), 557-572. doi:10.1098/rstb.2007.2170

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. doi:10.3835/plantgenome2012.11.0030

Covarrubias-Pazaran, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., . . . 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*, 13. doi:10.3389/fpls.2018.01310

Crossa, J., de los Campos, G., Perez, P., Gianola, D., Burgueno, J., Araus, J. L., . . . Braun, H. J. (2010). Prediction of Genetic Values of Quantitative Traits in Plant Breeding Using Pedigree and Molecular Markers. *Genetics, 186*(2), 713-U406. doi:10.1534/genetics.110.118521

Curtis, T., & Halford, N. (2014). Food security: the challenge of increasing wheat yield and the importance of not compromising food safety. *Annals of Applied Biology, 164*(3), 354-372. doi:10.1111/aab.12108

d'Amour, C. B., Reitsma, F., Baiocchi, G., Barthel, S., Guneralp, B., Erb, K. H., . . . Seto, K. C. (2017). Future urban land expansion and implications for global croplands. *Proceedings of the National Academy of Sciences of the United States of America, 114*(34), 8939-8944. doi:10.1073/pnas.1606036114

Dahl, B., & Wilson, W. W. (2018). Risk premiums due to Fusarium Head Blight (FHB) in wheat and barley. *Agricultural Systems, 162*, 145-153. doi:10.1016/j.agsy.2018.01.025

de los Campos, G., Naya, H., Gianola, D., Crossa, J., Legarra, A., Manfredi, E., . . . Cotes, J. M. (2009). Predicting Quantitative Traits With Regression Models for Dense Molecular Markers and Pedigree. *Genetics, 182*(1), 375-385. doi:10.1534/genetics.109.101501

Dong, H. X., Wang, R., Yuan, Y. P., Anderson, J., Pumphrey, M., Zhang, Z. W., & Chen, J. L. (2018). Evaluation of the potential for genomic selection to improve spring wheat resistance to Fusarium head blight in the Pacific Northwest. *Frontiers in Plant Science, 9*, 15. doi:10.3389/fpls.2018.00911

Dubcovsky, J., & Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science, 316*(5833), 1862-1866. doi:10.1126/science.1143986

Dvorak, J., Akhunov, E. D., Akhunov, A. R., Deal, K. R., & Luo, M. C. (2006). Molecular characterization of a diagnostic DNA marker for domesticated tetraploid wheat provides evidence for gene flow from wild tetraploid wheat to hexaploid wheat. *Molecular Biology and Evolution, 23*(7), 1386-1396. doi:10.1093/molbev/msl004

Dvorak, J., Diterlizzi, P., Zhang, H. B., & Resta, P. (1993). THE EVOLUTION OF POLYPLOID WHEATS - IDENTIFICATION OF THE A-GENOME DONOR SPECIES. *Genome, 36*(1), 21-31. doi:10.1139/g93-004

Dvorak, J., McGuire, P. E., & Cassidy, B. (1988). APPARENT SOURCES OF THE A GENOMES OF WHEATS INFERRED FROM POLYMORPHISM IN ABUNDANCE AND RESTRICTION FRAGMENT LENGTH OF REPEATED NUCLEOTIDE-SEQUENCES. *Genome, 30*(5), 680-689. doi:10.1139/g88-115

Eckardt, N. A. (2010). Evolution of Domesticated Bread Wheat. *Plant Cell, 22*(4), 993-993. doi:10.1105/tpc.110.220410

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *Plos One, 6*(5), 10. doi:10.1371/journal.pone.0019379

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

Falconer, D. S., & MacKay, T. F. C. (1996). *Introduction to Quantitative Genetics. 4th Edition*. Essex, England: Prentice Hall.

Fernando, R. L., & Grossman, M. (1989). MARKER ASSISTED SELECTION USING BEST LINEAR UNBIASED PREDICTION. *Genetics Selection Evolution, 21*(4), 467-477. doi:10.1051/gse:19890407

Finno, C. J., Aleman, M., Higgins, R. J., Madigan, J. E., & Bannasch, D. L. (2014). Risk of false positive genetic associations in complex traits with underlying population structure: A case study. *Veterinary Journal, 202*(3), 543-549. doi:10.1016/j.tvjl.2014.09.013

Flint-Garcia, S. A., Thuillet, A. C., Yu, J. M., Pressoir, G., Romero, S. M., Mitchell, S. E., . . . Buckler, E. S. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant Journal, 44*(6), 1054-1064. doi:10.1111/j.1365-313X.2005.02591.x

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software, 33*(1), 1-22. doi:10.18637/jss.v033.i01

Gianola, D. (2013). Priors in Whole-Genome Regression: The Bayesian Alphabet Returns. *Genetics, 194*(3), 573-596. doi:10.1534/genetics.113.151753

Gianola, D., de los Campos, G., Hill, W. G., Manfredi, E., & Fernando, R. (2009). Additive Genetic Variability and the Bayesian Alphabet. *Genetics, 183*(1), 347-363. doi:10.1534/genetics.109.103952

Gianola, D., Fernando, R. L., & Stella, A. (2006). Genomic-assisted prediction of genetic value with semiparametric procedures. *Genetics, 173*(3), 1761-1776. doi:10.1534/genetics.105.049510

Gianola, D., & van Kaam, J. (2008). Reproducing kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics, 178*(4), 2289-2303. doi:10.1534/genetics.107.084285

Goswami, R. S., & Kistler, H. C. (2004). Heading for disaster: Fusarium graminearum on cereal crops. *Molecular Plant Pathology, 5*(6), 515-525. doi:10.1111/j.1364-3703.2004.00252.x

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. doi:10.2135/cropsci2009.11.0685

Guo, J., Zhang, X. L., Hou, Y. L., Cai, J. J., Shen, X. R., Zhou, T. T., . . . Kong, L. R. (2015). High-density mapping of the major FHB resistance gene *Fhb7* derived from *Thinopyrum ponticum* and its pyramiding with *Fhb1* by marker-assisted selection. *Theoretical and Applied Genetics, 128*(11), 2301-2316. doi:10.1007/s00122-015-2586-x

Guo, Z. G., Tucker, D. M., Basten, C. J., Gandhi, H., Ersoz, E., Guo, B. H., . . . Gay, G. (2014). The impact of population structure on genomic prediction in stratified populations. *Theoretical and Applied Genetics, 127*(3), 749-762. doi:10.1007/s00122-013-2255-x

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. doi:10.1534/genetics.107.081190

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

Hayashi, T., & Iwata, H. (2010). EM algorithm for Bayesian estimation of genomic breeding values. *BMC Genetics, 11*(1), 3. doi:10.1186/1471-2156-11-3

He, S., Schulthess, A. W., Mirdita, V., Zhao, Y. S., Korzun, V., Bothe, R., . . . Jiang, Y. (2016). Genomic selection in a commercial winter wheat population. *Theoretical and Applied Genetics, 129*(3), 641-651. doi:10.1007/s00122-015-2655-1

Heffner, E. L., Jannink, J. L., Iwata, H., Souza, E., & Sorrells, M. E. (2011). Genomic Selection Accuracy for Grain Quality Traits in Biparental Wheat Populations. *Crop Science, 51*(6), 2597-2606. doi:10.2135/cropsci2011.05.0253

Heffner, E. L., Jannink, J. L., & Sorrells, M. E. (2011). Genomic Selection Accuracy using Multifamily Prediction Models in a Wheat Breeding Program. *Plant Genome, 4*(1), 65-75. doi:10.3835/plantgenome2010.12.0029

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

Heslot, N., Akdemir, D., Sorrells, M. E., & Jannink, J. L. (2014). Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theoretical and Applied Genetics, 127*(2), 463-480. doi:10.1007/s00122-013-2231-5

Heslot, N., Rutkoski, J., Poland, J., Jannink, J. L., & Sorrells, M. E. (2013). Impact of Marker Ascertainment Bias on Genomic Selection Accuracy and Estimates of Genetic Diversity. *Plos One, 8*(9), 8. doi:10.1371/journal.pone.0074612

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. doi:10.2135/cropsci2011.06.0297

Hoffstetter, A., Cabrera, A., Huang, M., & Sneller, C. (2016). Optimizing Training Population Data and Validation of Genomic Selection for Economic Traits in Soft Winter Wheat. *G3-Genes Genomes Genetics, 6*(9), 2919-2928. doi:10.1534/g3.116.032532

Hornik, K. (1993). SOME NEW RESULTS ON NEURAL-NETWORK APPROXIMATION. *Neural Networks, 6*(8), 1069-1072. doi:10.1016/s0893-6080(09)80018-x

Hospital, F. (2009). Challenges for effective marker-assisted selection in plants. *Genetica, 136*(2), 303-310. doi:10.1007/s10709-008-9307-1

Houle, D., Govindaraju, D. R., & Omholt, S. (2010). Phenomics: the next challenge. *Nature Reviews Genetics, 11*(12), 855-866. doi:10.1038/nrg2897

Huang, B. E., George, A. W., Forrest, K. L., Kilian, A., Hayden, M. J., Morell, M. K., & Cavanagh, C. R. (2012). A multiparent advanced generation inter-cross population for genetic analysis in wheat. *Plant Biotechnology Journal, 10*(7), 826-839. doi:10.1111/j.1467-7652.2012.00702.x

Huang, M., Liu, X. L., Zhou, Y., Summers, R. M., & Zhang, Z. W. (2019). BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience, 8*(2), 12. doi:10.1093/gigascience/giy154

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. doi:10.1007/s00122-014-2418-4

Jaggard, K. W., Qi, A. M., & Ober, E. S. (2010). Possible changes to arable crop yields by 2050. *Philosophical Transactions of the Royal Society B-Biological Sciences, 365*(1554), 2835-2851. doi:10.1098/rstb.2010.0153

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. doi:10.1093/bfgp/elq001

Jarquin, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., . . . de los Campos, G. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical and Applied Genetics, 127*(3), 595-607. doi:10.1007/s00122-013-2243-1

Jia, Y., & Jannink, J. L. (2012). Multiple-Trait Genomic Selection Methods Increase Genetic Value Prediction Accuracy. *Genetics, 192*(4), 1513-+. doi:10.1534/genetics.112.144246

Jiang, C. J., & Zeng, Z. B. (1995). MULTIPLE-TRAIT ANALYSIS OF GENETIC-MAPPING FOR QUANTITATIVE TRAIT LOCI. *Genetics, 140*(3), 1111-1127.

Jiang, G. L., Dong, Y., Shi, J., & Ward, R. W. (2007). QTL analysis of resistance to Fusarium head blight in the novel wheat germplasm CJ 9306. II. Resistance to deoxynivalenol accumulation and grain yield loss. *Theoretical and Applied Genetics, 115*(8), 1043-1052. doi:10.1007/s00122-007-0630-1

Jiang, G. L., Shi, J. R., & Ward, R. W. (2007). QTL analysis of resistance to Fusarium head blight in the novel wheat germplasm CJ 9306. I. Resistance to fungal spread. *Theoretical and Applied Genetics, 116*(1), 3-13. doi:10.1007/s00122-007-0641-y

Jiang, Y., & Reif, J. C. (2015). Modeling Epistasis in Genomic Selection. *Genetics, 201*(2), 759-+. doi:10.1534/genetics.115.177907

Jiang, Y., Zhao, Y., Rodemann, B., Plieske, J., Kollers, S., Korzun, V., . . . Reif, J. C. (2015). Potential and limits to unravel the genetic architecture and predict the variation of Fusarium head blight resistance in European winter wheat (Triticum aestivum L.). *Heredity, 114*(3), 318-326. doi:10.1038/hdy.2014.104

Karkkainen, H. P., & Sillanpaa, M. J. (2012). Back to Basics for Bayesian Model Building in Genomic Selection. *Genetics, 191*(3), 969-U568. doi:10.1534/genetics.112.139014

Kihara, H. (1944). Discovery of the DD-analyser, one of the ancestors of Triticum vulgare. *Agric. Hort., 19*, 13-14.

Knapp, S. J. (1998). Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. *Crop Science, 38*(5), 1164-1174. doi:10.2135/cropsci1998.0011183X003800050009x

Lado, B., Barrios, P. G., Quincke, M., Silva, P., & Gutierrez, L. (2016). Modeling Genotype x Environment Interaction for Genomic Selection with Unbalanced Data from a Wheat Breeding Program. *Crop Science, 56*(5), 2165-2179. doi:10.2135/cropsci2015.04.0207

Lande, R., & Thompson, R. (1990). EFFICIENCY OF MARKER-ASSISTED SELECTION IN THE IMPROVEMENT OF QUANTITATIVE TRAITS. *Genetics, 124*(3), 743-756.

Larkin, D. L., Lozada, D. N., & Mason, R. E. (2019). Genomic Selection-Considerations for Successful Implementation in Wheat Breeding Programs. *Agronomy-Basel, 9*(9), 18. doi:10.3390/agronomy9090479

Legarra, A., Christensen, O. F., Aguilar, I., & Misztal, I. (2014). Single Step, a general approach for genomic selection. *Livestock Science, 166*, 54-65. doi:10.1016/j.livsci.2014.04.029

Lipka, A. E., Tian, F., Wang, Q. S., Peiffer, J., Li, M., Bradbury, P. J., . . . Zhang, Z. W. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics, 28*(18), 2397-2399. doi:10.1093/bioinformatics/bts444

Liu, X. L., Huang, M., Fan, B., Buckler, E. S., & Zhang, Z. W. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *Plos Genetics, 12*(2), 24. doi:10.1371/journal.pgen.1005767

Lopez-Cruz, M., Crossa, J., Bonnett, D., Dreisigacker, S., Poland, J., Jannink, J. L., . . . de los Campos, G. (2015). Increased Prediction Accuracy in Wheat Breeding Trials Using a Marker x Environment Interaction Genomic Selection Model. *G3-Genes Genomes Genetics, 5*(4), 569-582. doi:10.1534/g3.114.016097

Lorenz, A. J., & Smith, K. P. (2015). Adding Genetically Distant Individuals to Training Populations Reduces Genomic Prediction Accuracy in Barley. *Crop Science, 55*(6), 2657-2667. doi:10.2135/cropsci2014.12.0827

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. doi:10.2135/cropsci2011.09.0503

Lorenz, K., & Cohen, B. A. (2012). Small- and Large-Effect Quantitative Trait Locus Interactions Underlie Variation in Yeast Sporulation Efficiency. *Genetics, 192*(3), 1123-+. doi:10.1534/genetics.112.143107

Lorenzana, R. E., & Bernardo, R. (2009). Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theoretical and Applied Genetics, 120*(1), 151-161. doi:10.1007/s00122-009-1166-3

Marcussen, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M., Jakobsen, K. S., . . . Int Wheat Genome, S. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science, 345*(6194), 4. doi:10.1126/science.1250092

Mason, R. E., Addison, C. K., Babar, A., Acuna, A., Lozada, D., Subramanian, N., . . . Johnson, J. (2018). Diagnostic markers for vernalization and photoperiod loci improve genomic selection for grain yield and spectral reflectance in wheat. *Crop Science, 58*(1), 242-252. doi:10.2135/cropsci2017.06.0348

McFadden, E. S., & Sears, E. R. (1946). The origin of Triticum spelta and its free-threshing hexaploid relatives. *Journal of Heredity, 37*(3), 81-89.

Mesterhazy, A. (1995). Types and components of resistance to Fusarium head blight of wheat. *Plant Breeding, 114*(5), 377-386. doi:10.1111/j.1439-0523.1995.tb00816.x

Meuwissen, T. H. E. (2009). Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. *Genetics Selection Evolution, 41*, 9. doi:10.1186/1297-9686-41-35

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.

Michel, S., Ametz, C., Gungor, H., Akgol, B., Epure, D., Grausgruber, H., . . . Buerstmayr, H. (2017). Genomic assisted selection for enhancing line breeding: merging genomic and phenotypic selection in winter wheat breeding programs with preliminary yield trials. *Theoretical and Applied Genetics, 130*(2), 363-376. doi:10.1007/s00122-016-2818-8

Michel, S., Ametz, C., Gungor, H., Epure, D., Grausgruber, H., Loschenberger, F., & Buerstmayr, H. (2016). Genomic selection across multiple breeding cycles in applied bread wheat breeding. *Theoretical and Applied Genetics, 129*(6), 1179-1189. doi:10.1007/s00122-016-2694-2

Milus, E. A., & Parsons, C. E. (1994). EVALUATION OF FOLIAR FUNGICIDES FOR CONTROLLING FUSARIUM HEAD BLIGHT OF WHEAT. *Plant Disease, 78*(7), 697-699. doi:10.1094/pd-78-0697

Mohammadi, M., Blake, T. K., Budde, A. D., Chao, S. M., Hayes, P. M., Horsley, R. D., . . . Smith, K. P. (2015). A genome-wide association study of malting quality across eight US barley breeding programs. *Theoretical and Applied Genetics, 128*(4), 705-721. doi:10.1007/s00122-015-2465-5

Mohan, M., Nair, S., Bhagwat, A., Krishna, T. G., Yano, M., Bhatia, C. R., & Sasaki, T. (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular Breeding, 3*(2), 87-103. doi:10.1023/a:1009651919792

Mujibi, F. D. N., Nkrumah, J. D., Durunna, O. N., Stothard, P., Mah, J., Wang, Z., . . . Moore, S. S. (2011). Accuracy of genomic breeding values for residual feed intake in crossbred beef cattle. *Journal of Animal Science, 89*(11), 3353-3361. doi:10.2527/jas.2010-3361

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z. W., Costich, D. E., & Buckler, E. S. (2009). Association Mapping: Critical considerations shift from genotyping to experimental design. *Plant Cell, 21*(8), 2194-2202. doi:10.1105/tpc.109.068437

Norman, A., Taylor, J., Edwards, J., & Kuchel, H. (2018). Optimising Genomic Selection in Wheat: Effect of Marker Density, Population Size and Population Structure on Prediction Accuracy. *G3-Genes Genomes Genetics, 8*(9), 2889-2899. doi:10.1534/g3.118.200311

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*, 10. doi:10.1186/s12711-017-0361-y

Owens, B. F., Lipka, A. E., Magallanes-Lundback, M., Tiede, T., Diepenbrock, C. H., Kandianis, C. B., . . . Rocheford, T. (2014). A foundation for provitamin A biofortification of maize: Genome-wide association and genomic prediction models of carotenoid levels. *Genetics, 198*(4), 1699-+. doi:10.1534/genetics.114.169979

Park, T., & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association, 103*(482), 681-686. doi:10.1198/016214508000000337

Parry, M. A. J., Reynolds, M., Salvucci, M. E., Raines, C., Andralojc, P. J., Zhu, X. G., . . . Furbank, R. T. (2011). Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. *Journal of Experimental Botany, 62*(2), 453-467. doi:10.1093/jxb/erq304

Perez, P., de los Campos, G., Crossa, J., & Gianola, D. (2010). Genomic-Enabled Prediction Based on Molecular Markers and Pedigree Using the Bayesian Linear Regression Package in R. *Plant Genome, 3*(2), 106-116. doi:10.3835/plantgenome2010.04.0005

Pestka, J. J. (2010). Deoxynivalenol: mechanisms of action, human exposure, and toxicological relevance. *Archives of Toxicology, 84*(9), 663-679. doi:10.1007/s00204-010-0579-8

Piepho, H. P. (2009). Ridge Regression and Extensions for Genomewide Selection in Maize. *Crop Science, 49*(4), 1165-1176. doi:10.2135/cropsci2008.10.0595

Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S. Y., Manes, Y., . . . Jannink, J. L. (2012). Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *Plant Genome, 5*(3), 103-113. doi:10.3835/plantgenome2012.06.0006

Pritchard, J. K., Stephens, M., Rosenberg, N. A., & Donnelly, P. (2000). Association mapping in structured populations. *American Journal of Human Genetics, 67*(1), 170-181. doi:10.1086/302959

Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology, 5*(2), 94-100. doi:10.1016/s1369-5266(02)00240-6

Ragot, M., Gay, G., Muller, J.-P., & Durovray, J. (2000). Efficient selection for the adaptation to the environment through QTL mapping and manipulation in maize. *Molecular approaches for the genetic improvement of cereals for stable production in water-limited environments*, 128-130.

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), 8. doi:10.1371/journal.pone.0066428

Rice, B., & Lipka, A. E. (2019). Evaluation of RR-BLUP Genomic Selection Models that Incorporate Peak Genome-Wide Association Study Signals in Maize and Sorghum. *Plant Genome, 12*(1), 14. doi:10.3835/plantgenome2018.07.0052

Riedelsheimer, C., Endelman, J. B., Stange, M., Sorrells, M. E., Jannink, J. L., & Melchinger, A. E. (2013). Genomic Predictability of Interconnected Biparental Maize Populations. *Genetics, 194*(2), 493-+. doi:10.1534/genetics.113.150227

Rincent, R., Laloe, D., Nicolas, S., Altmann, T., Brunel, D., Revilla, P., . . . Moreau, L. (2012). Maximizing the Reliability of Genomic Selection by Optimizing the Calibration Set of Reference Individuals: Comparison of Methods in Two Diverse Groups of Maize Inbreds (Zea mays L.). *Genetics, 192*(2), 715-+. doi:10.1534/genetics.112.141473

Roder, M. S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M. H., Leroy, P., & Ganal, M. W. (1998). A microsatellite map of wheat. *Genetics, 149*(4), 2007-2023.

Roorkiwal, M., Rathore, A., Das, R. R., Singh, M. K., Jain, A., Srinivasan, S., . . . Varshney, R. K. (2016). Genome-Enabled Prediction Models for Yield Related Traits in Chickpea. *Frontiers in Plant Science, 7*, 13. doi:10.3389/fpls.2016.01666

Rutkoski, J., Benson, J., Jia, Y., Brown-Guedira, G., Jannink, J. L., & Sorrells, M. (2012). Evaluation of genomic prediction methods for Fusarium head blight resistance in wheat. *Plant Genome, 5*(2), 51-61. doi:10.3835/plantgenome2012.02.0001

Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Perez, L. G., Crossa, J., . . . 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. doi: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. doi:10.3835/plantgenome2014.10.0074

Rutkoski, J. E., Poland, J. A., Singh, R. P., Huerta-Espino, J., Bhavani, S., Barbier, H., . . . Sorrells, M. E. (2014). Genomic selection for quantitative adult plant stem rust resistance in wheat. *Plant Genome, 7*(3), 10. doi:10.3835/plantgenome2014.02.0006

Sallam, A. H., Endelman, J. B., Jannink, J. L., & Smith, K. P. (2015). Assessing Genomic Selection Prediction Accuracy in a Dynamic Barley Breeding Population. *Plant Genome, 8*(1), 15. doi:10.3835/plantgenome2014.05.0020

Sarinelli, J. M., Murphy, J. P., Tyagi, P., Holland, J. B., Johnson, J. W., Mergoum, M., . . . 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. doi:10.1007/s00122-019-03276-6

Schmidt, M., Kollers, S., Maasberg-Prelle, A., Grosser, J., Schinkel, B., Tomerius, A., . . . Korzun, V. (2016). Prediction of malting quality traits in barley based on genome-wide marker data to assess the potential of genomic selection. *Theoretical and Applied Genetics, 129*(2), 203-213. doi:10.1007/s00122-015-2639-1

Schroeder, H. W., & Christensen, J. J. (1963). FACTORS AFFECTING RESISTANCE OF WHEAT TO SCAB CAUSED BY GIBBERELLA ZEAE. *Phytopathology, 53*(7), 831-&.

Shewry, P. R. (2009). Wheat. *Journal of Experimental Botany, 60*(6), 1537-1553. doi:10.1093/jxb/erp058

Shewry, P. R., & Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food and Energy Security, 4*(3), 178-202. doi:10.1002/fes3.64

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*, 12. doi:10.3389/fpls.2017.00550

Spindel, J., Begum, H., Akdemir, D., Virk, P., Collard, B., Redona, E., . . . McCouch, S. R. (2015). Genomic Selection and Association Mapping in Rice (Oryza sativa): Effect of Trait Genetic Architecture, Training Population Composition, Marker Number and Statistical Model on Accuracy of Rice Genomic Selection in Elite, Tropical Rice Breeding Lines. *Plos Genetics, 11*(2), 25. doi:10.1371/journal.pgen.1004982

Spindel, J. E., Begum, H., Akdemir, D., Collard, B., Redona, E., Jannink, J. L., & McCouch, S. (2016). Genome-wide prediction models that incorporate *de novo* GWAS are a powerful new tool for tropical rice improvement. *Heredity, 116*(4), 395-408. doi:10.1038/hdy.2015.113

Steiner, B., Buerstmayr, M., Michel, S., Schweiger, W., Lemmens, M., & Buerstmayr, H. (2017). Breeding strategies and advances in line selection for Fusarium head blight resistance in wheat. *Tropical Plant Pathology, 42*(3), 165-174. doi:10.1007/s40858-017-0127-7

Tiede, T., & Smith, K. P. (2018). Evaluation and retrospective optimization of genomic selection for yield and disease resistance in spring barley. *Molecular Breeding, 38*(5), 16. doi:10.1007/s11032-018-0820-3

Tilman, D., Fargione, J., Wolff, B., D'Antonio, C., Dobson, A., Howarth, R., . . . Swackhamer, D. (2001). Forecasting agriculturally driven global environmental change. *Science, 292*(5515), 281-284. doi:10.1126/science.1057544

United Nations Department of Economic and Social Affairs: Population Division. (2019). *World Population Prospects 2019: Highlights*. Retrieved from

USDA Foreign Agricultural Service. (2020). World Agricultural Production. Retrieved from https://apps.fas.usda.gov/psdonline/circulars/production.pdf

USDA National Agricultural Statistics Service. (2019). 2019 State Agriculture Overview: Arkansas. Retrieved from https://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=ARKANSAS

USDA National Agricultural Statistics Service. (2020a). Crop Production. Retrieved from https://downloads.usda.library.cornell.edu/usda-esmis/files/tm70mv177/1j92h0351/hm50vh540/crop1120.pdf

USDA National Agricultural Statistics Service. (2020b). *Small Grains 2020 Summary*. Retrieved from https://downloads.usda.library.cornell.edu/usda-esmis/files/5t34sj573/7p88d6171/bn999x131/smgr0920.pdf

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

Waldron, B. L., Moreno-Sevilla, B., Anderson, J. A., Stack, R. W., & Frohberg, R. C. (1999). RFLP mapping of QTL for fusarium head blight resistance in wheat. *Crop Science, 39*(3), 805-811. doi:10.2135/cropsci1999.0011183X003900030032x

Wang, H., Sun, S., Ge, W., Zhao, L., Hou, B., Wang, K., . . . Guo, J. (2020). Horizontal gene transfer of Fhb7 from fungus underlies Fusarium head blight resistance in wheat. *Science*. doi:10.1126/science.aba5435

Wang, S. C., Wong, D. B., Forrest, K., Allen, A., Chao, S. M., Huang, B. E., . . . Int Wheat Genome, S. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnology Journal, 12*(6), 787-796. doi:10.1111/pbi.12183

Wang, X., Xu, Y., Hu, Z. L., & Xu, C. W. (2018). Genomic selection methods for crop improvement: Current status and prospects. *Crop Journal, 6*(4), 330-340. doi:10.1016/j.cj.2018.03.001

Wegulo, S. N., Baenziger, P. S., Nopsa, J. H., Bockus, W. W., & Hallen-Adams, H. (2015). Management of Fusarium head blight of wheat and barley. *Crop Protection, 73*, 100-107. doi:10.1016/j.cropro.2015.02.025

Whittaker, J. C., Thompson, R., & Denham, M. C. (2000). Marker-assisted selection using ridge regression. *Genetical Research, 75*(2), 249-252. doi:10.1017/s0016672399004462

William, H. M., Trethowan, R., & Crosby-Galvan, E. M. (2007). Wheat breeding assisted by markers: CIMMYT's experience. *Euphytica, 157*(3), 307-319. doi:10.1007/s10681-007-9405-7

Windhausen, V. S., Atlin, G. N., Hickey, J. M., Crossa, J., Jannink, J. L., Sorrells, M. E., . . . Melchinger, A. E. (2012). Effectiveness of Genomic Prediction of Maize Hybrid Performance in Different Breeding Populations and Environments. *G3-Genes Genomes Genetics, 2*(11), 1427-1436. doi:10.1534/g3.112.003699

Xu, C. W., Wang, X. F., Li, Z. K., & Xu, S. Z. (2009). Mapping QTL for multiple traits using Bayesian statistics. *Genetics Research, 91*(1), 23-37. doi:10.1017/s0016672308009956

Xu, S. Z. (2007). An empirical Bayes method for estimating epistatic effects of quantitative trait loci. *Biometrics, 63*(2), 513-521. doi:10.1111/j.1541-0420.2006.00711.x

Xu, X. M., & Nicholson, P. (2009). Community Ecology of Fungal Pathogens Causing Wheat Head Blight. In *Annual Review of Phytopathology* (Vol. 47, pp. 83-103). Palo Alto: Annual Reviews.

Xu, Y. B., & Crouch, J. H. (2008). Marker-assisted selection in plant breeding: From publications to practice. *Crop Science, 48*(2), 391-407. doi:10.2135/cropsci2007.04.0191

Xu, Y. B., Li, P., Zou, C., Lu, Y. L., Xie, C. X., Zhang, X. C., . . . Olsen, M. S. (2017). Enhancing genetic gain in the era of molecular breeding. *Journal of Experimental Botany, 68*(11), 2641-2666. doi:10.1093/jxb/erx135

Xue, S., Li, G. Q., Jia, H. Y., Xu, F., Lin, F., Tang, M. Z., . . . Ma, Z. Q. (2010). Fine mapping *Fhb4*, a major QTL conditioning resistance to Fusarium infection in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics, 121*(1), 147-156. doi:10.1007/s00122-010-1298-5

Xue, S. L., Xu, F., Tang, M. Z., Zhou, Y., Li, G. Q., An, X., . . . Ma, Z. Q. (2011). Precise mapping Fhb5, a major QTL conditioning resistance to Fusarium infection in bread wheat (Triticum aestivum L.). *Theoretical and Applied Genetics, 123*(6), 1055-1063. doi:10.1007/s00122-011-1647-z

Yamamoto, E., Matsunaga, H., Onogi, A., Ohyama, A., Miyatake, K., Yamaguchi, H., . . . Fukuoka, H. (2017). Efficiency of genomic selection for breeding population design and phenotype prediction in tomato. *Heredity, 118*(2), 202-209. doi:10.1038/hdy.2016.84

Zhang, Z., Erbe, M., He, J. L., Ober, U., Gao, N., Zhang, H., . . . Li, J. Q. (2015). Accuracy of Whole-Genome Prediction Using a Genetic Architecture-Enhanced Variance-Covariance Matrix. *G3-Genes Genomes Genetics, 5*(4), 615-627. doi:10.1534/g3.114.016261

Zhang, Z., Ober, U., Erbe, M., Zhang, H., Gao, N., He, J. L., . . . Simianer, H. (2014). Improving the Accuracy of Whole Genome Prediction for Complex Traits Using the Results of Genome Wide Association Studies. *Plos One, 9*(3), 12. doi:10.1371/journal.pone.0093017

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 67*, 301-320. doi:10.1111/j.1467-9868.2005.00503.x

Table 1. Modifications to genomic selection and impact on the breeder's equation and accuracy.

Factor	Factor on Breeder's Equation Impacted	Effect on selection accuracy (*r*)	Consideration for breeding
Marker density	Selection accuracy	Varies	When LD is high a small number of markers are needed
Choice of prediction model		Varies	Models vary in their ability to account for larger effect QTL and epistatic interactions
Training population size	Genetic variance	Increased r for increase in training population size	Greater diversity requires a larger training population size
Trait heritability	Selection accuracy	Varies	In the absence of large effect QTL, accuracy generally increases with trait heritability
Genetic relationship between training and validation	Selection accuracy	Increased r as more genetically related populations or families are used	Should use genetically related training and validation populations
Population structure	Selection accuracy	Varies	Sub-populations can lead to erroneous predictions
Training population composition	Selection accuracy	Varies	Should use a training population that captures genetic relationships with the validation population
Genotype by environment interaction	Additive genetic variance	Increased r when GxE effects fitted in the model	GxE together with pedigree and marker information could help improve accuracy
Multiple traits	Additive genetic variance	Increased r when multiple traits are fitted in the model	Accounts for additional genetic variance
Use of GWAS results	Additive genetic variance	Increased r for incorporation of GWAS results in the model (for most empirical studies)	Accounts for large effect QTL
Multiple generations per year	Breeding cycle time	Varies; r affected by one or a combination of factors above	Increases gain per time by reducing cycle time
Early generation parent selection	Breeding cycle time, selection intensity	Varies; r affected by one or a combination of factors above	Increases gain per time by expediting the recombination of favorable alleles

Table 1 (Cont.)

Factor	**Factor on Breeder's Equation Impacted**	**Effect on selection accuracy (*r*)**	**Consideration for breeding**
Selection intensity	Selection accuracy	Varies; r affected by one or a combination of factors above	Genomic selection allows more lines to be characterized and thus increased in selection intensity
Additive genetic variance	Selection accuracy	Varies; r affected by one or a combination of factors above	High additive genetic variance for increased gains

Table 2. Summary of models used for genomic selection in plant breeding.

Model	Prior	Estimation	Marker Variances	Parametric/ Nonparametric	Type	Reference
RR-BLUP	Normal		Equal	Parametric	BLUP	Meuwissen et al. (2001)
GBLUP	Normal		Equal	Parametric	BLUP	Zhang et al. (2015)
EGBLUP	Normal		Equal	Parametric	BLUP (Epistatic)	Jiang & Reif (2015)
Bayes A	Normal	MCMC	Independent (Inverse Chi-square)	Parametric	Bayesian	Meuwissen et al. (2001)
Bayes B	Normal	MCMC	Independent (Some = 0)	Parametric	Bayesian	Meuwissen et al. (2001)
Bayesian RR	Normal	MCMC	Equal	Parametric	Bayesian	Perez et al. (2010)
Bayesian LASSO	Laplace	MCMC	Independent (Inverse Chi-square)	Parametric	Bayesian	de los Campos et al. (2009)
wBSR	Normal	EM	Independent (Some = 0)	Parametric	Bayesian	Hayashi & Iwata (2010)
Bayes Cπ	Normal	MCMC	Equal (Some = 0)	Parametric	Bayesian	Habier et al. (2011)
Bayes Dπ	Normal	MCMC	Independent (Inverse Chi-square)	Parametric	Bayesian	Habier et al. (2011)
E-Bayes	Normal	EM	Independent (Inverse Chi-square)	Parametric	Bayesian (Epistatic)	Xu (2007)
Elastic Network	Normal/Laplace	MCMC		Parametric	Sparse Model	Zou & Hastie (2005)
Exponential Kernel	Normal		Equal	Semi-parametric	Kernel (Epistatic)	Endelman (2011)

Table 2 (Cont.)

Model	Prior	Estimation	Marker Variances	Parametric/ Nonparametric	Type	Reference
Gaussian Kernel	Normal		Equal	Semi-parametric	Kernel (Epistatic)	Endelman (2011)
RKHS	Normal			Semi-parametric	Kernel (Epistatic)	Gianola & van Kaam (2008)
Random Forest				Non-parametric	Machine Learning	Breiman (2001)
Neural Networks				Non-parametric	Machine Learning	Hornik (1993)

Table 3. Some programs and the models used for genomic selection.

Program	Software	Model(s)	Reference
AlphaBayes	Command Line	Bayes A	Hickey & Tier (2009)
		Bayes B	
		Bayes Cπ	
		Bayesian LASSO	
ASreml	Command Line | R	GBLUP	Butler et al. (2009)
BayZ	Command Line | R	GBLUP	Janss (2010)
		Bayesian LASSO	
		Bayes Cπ	
BGLR	R	GBLUP	Perez & de los Campos (2014)
		Bayes A	
		Bayes B	
		Bayes Cπ	
		Bayesian RR	
		Bayesian LASSO	
		RKHS	
BLR	R	Bayesian RR	Perez et al. (2010)
		Bayesian LASSO	
bWGR	R	GBLUP	Xavier et al. (2017)
		Bayes A	
		Bayes B	
		Bayes Cπ	
		Bayes Dπ	
		Bayesian LASSO	
glmnet	R	Elastic Network	Hastie & Qian (2016)
GenSel	Command Line	Bayes A	Fernando & Garrick (2008)
		Bayes B	
		Bayes Cπ	
GS3	Command Line	GBLUP	Legarra et al. (2012)
		Bayes Cπ	
		Bayesian LASSO	
MXNET	Python	Neural Network	Chen et al. (2015)
randomForest	R	Random Forest	Liaw & Wiener (2002)
rrBLUP	R	GBLUP	Endelman (2011)
		RR-BLUP	
		Exponential Kernel	
		Gaussian Kernel	
PROC IML	SAS/IML Software	E-Bayes	Xu (2007)
VIGoR	Command Line | R	Bayes B	Onogi & Iwata (2016)
		Bayes Cπ	
		Bayesian LASSO	
		wBSR	

CHAPTER II

GENOME-WIDE ANALYSIS OF FUSARIUM HEAD BLIGHT RESISTANCE IN SOFT RED WINTER WHEAT

ABSTRACT

Fusarium head blight (FHB) is a disease in wheat (*Triticum aestivum* L.) caused by the fungal pathogen *Fusarium graminearum* Schwabe. Fusarium head blight poses potential economic losses and health risks due to the accumulation of the mycotoxin deoxynivalenol (DON) on infected seed heads. The objectives of this study were to identify novel FHB resistance loci using a genome-wide association (GWAS) approach and evaluate two genomic selection (GS) approaches to improve prediction accuracies for FHB traits in a population of 354 soft red winter wheat (SRWW) genotypes. The GS approaches included GS+GWAS, where markers associated with a trait were used as fixed effects, and multivariate GS (MVGS), where correlated traits were used as fixed effects. The population was evaluated in FHB nurseries in Fayetteville and Newport, AR and Winnsboro, LA from 2014-2017. Genotypes were phenotyped for DON, Fusarium damaged kernels (FDK), incidence (INC), and severity (SEV). Forty-two single nucleotide polymorphism (SNP) markers were significantly (false discovery rate, q [FDR_q] ≤ .10) associated with resistance traits across 17 chromosomes. Ten significant SNPs were identified for DON, notably on chromosomes 2BL and 3BL. Eleven were identified for FDK, notably on chromosomes 4BL, 3AL, 1BL, 5BL, and 5DL. Nine were identified for INC, notably on chromosomes 2BS, 2BL, 7BL, 5DL, 6AS, and 5DS. Twelve were identified for SEV, notably on chromosomes 3BL, 4AL, and 4BL. The naïve GS models significantly outperformed the GS+GWAS model for all traits, while MVGS models significantly outperformed the naïve GS models for all traits. Results from this study will facilitate the development of SRWW cultivars with improved FHB resistance.

INTRODUCTION

Fusarium head blight (FHB) is a disease of small grains caused by fungal pathogens of the *Fusarium* genus. *Fusarium graminearum* is the main causal organism in the United States. Wheat (*Triticum aestivum* L.) producers in the United States incurred an estimated economic loss of nearly US$4.2 billion during the 2015-2016 growing season due to FHB (Wilson, McKee, Nganje, Dahl, & Bangsund, 2017). Economic losses may be due to losses in grain yield, mycotoxin accumulation, loss of grain quality, or failed preventative efforts such as late fungicide application. The *F. graminearum* pathogen also produces the mycotoxin deoxynivalenol (DON) in order to disable natural plant defenses. Consumption of grain high in DON has several adverse effects on the health of both humans and animals, resulting in symptoms such as diarrhea, nausea, vomiting, and weight loss (FDA, 2010; Sobrova et al., 2010).

Four primary forms of FHB resistance (Type I – IV) exist in wheat. Type I is resistance to initial infection by *F. graminearum*, otherwise known as resistance to incidence (INC). Type II, also known as severity (SEV), is the resistance to fungal spread within the infected head (Steiner et al., 2017). Type III is resistance to Fusarium damaged kernels (FDK), whereas Type IV resistance is resistance to the accumulation of DON.

As FHB resistance in wheat is a complex, quantitative trait, it is an ideal candidate for genomic selection (GS) (H. Buerstmayr & Lemmens, 2015; Hoffstetter, Cabrera, & Sneller, 2016; Larkin, Lozada, & Mason, 2019; Mirdita et al., 2015; Steiner et al., 2017). Genomic selection is a modified form of marker assisted selection (MAS) where all summed marker and locus effects across the entire genome are taken into consideration to calculate genome estimated breeding values (GEBVs) (Meuwissen, Hayes, & Goddard, 2001). The assumption is that at least

some of the markers are in linkage disequilibrium (LD) with quantitative trait loci (QTL) associated with a trait of interest. This allows for more efficient and effective selection of complex quantitative traits compared with MAS (Meuwissen et al., 2001). Two cycles of GS can be performed in the time it takes to perform one round of phenotypic selection, allowing for greater genetic gain per cycle of GS compared with phenotypic selection (Asoro et al., 2013; Bernardo & Yu, 2007; Heffner, Sorrells, & Jannink, 2009; J. Rutkoski et al., 2015).

In wheat, a panel of 322 genotypes adapted to the eastern United States, was used to predict the performance of FHB traits, where prediction accuracies ranged between .11 and .64 for SEV, .01 and .56 for INC, .01 and .46 for FDK, and .16 and .58 for DON (J. Rutkoski et al., 2012). Another study was conducted using soft red winter wheat (SRWW) genotypes adapted to the midwestern United States, where prediction accuracies for INC ranged between .51 and .63, .35 and .48 for SEV, .67 and .82 for FDK, and .48 and .64 for DON (Arruda et al., 2015). A study on 170 soft white spring wheat genotypes in the U.S. Pacific Northwest found that INC had the highest prediction accuracy (.87), followed by SEV (.43) and DON (.42) (Dong et al., 2018).

Genome-wide association studies (GWAS) rely on LD and natural ancestral recombination events in diverse populations to identify significant QTL between genome-wide markers and individual traits (Breseghello & Sorrells, 2006; Flint-Garcia et al., 2005; Myles et al., 2009; Rafalski, 2002). Several GWAS were performed in wheat to identify resistance to FHB. A panel of 455 European SRWW genotypes were evaluated for FHB resistance and nine significant QTL were identified with two unique genomic regions on chromosomes 1D and 3A (Miedaner et al., 2011). Another study identified significant QTL on all chromosomes except for 6B for INC and SEV, which also coincided with QTL highlighted by Buerstmayr, Ban, and

Anderson (2009) (H. Buerstmayr, Ban, & Anderson, 2009; Kollers et al., 2013). Arruda et al. (2016) identified significant QTL on chromosomes 1D, 3B, 4A, 4D, 6A, 7A, and 7D. Several of the single nucleotide polymorphisms (SNP) found on chromosome 3B were associated with the FHB resistance QTL *Fhb1* (Arruda, Brown, et al., 2016).

One of the methods used for improving GS prediction accuracies uses markers identified through GWAS, linked to major genes associated with traits of interest, as model fixed effects. This method is referred to as GS + *de novo* GWAS (GS+GWAS) (Bian & Holland, 2017; Rice & Lipka, 2019; Spindel et al., 2016). Several GS experiments have shown that this strategy can improve prediction accuracies of agronomic traits in wheat, rice (*Oryza sativa* L.) and maize (*Zea mays* L.) (Bernardo, 2014; Mason et al., 2018; Owens et al., 2014; J. E. Rutkoski et al., 2014; Sarinelli et al., 2019; Spindel et al., 2016). The GS+GWAS method has also improved prediction accuracy for disease resistance traits in wheat, such as for powdery mildew [*Blumeria graminis* (DC) Speer f. sp. *tritici* emend. E.J. Marchal] and FHB (Arruda, Lipka, et al., 2016; Sarinelli et al., 2019). Arruda et al. (2016) found that GS+GWAS models had 32% higher prediction accuracies than a naïve GS (NGS) model when averaged across six FHB traits, including SEV, INC, FDK, and DON. The inclusion of markers as fixed effects significantly improved prediction accuracy for all traits except for DON (Arruda, Lipka, et al., 2016).

Another method used to improve GS prediction accuracies is multivariate GS (MVGS), which uses secondary traits, phenotyped in both the training and validation populations, that are genetically correlated with a trait of interest as a covariate in the GS model. Multivariate GS models using high-heritability traits as covariates have the ability to improve prediction accuracies of traits that have low heritability (Covarrubias-Pazaran et al., 2018; Y. Jia & Jannink, 2012). A MVGS model was compared with a univariate GS model for SEV in a hybrid wheat

population. Three MVGS models were compared, two with one covariate, including plant height (PH) and heading date (HD), and one with two covariates, HD and PH. The MVGS models with PH as a covariate had significantly higher prediction accuracies than a NGS model or a MVGS model with HD as a covariate (Schulthess, Zhao, Longin, & Reif, 2018). Another study with elite durum wheat (*Triticum durum* Desf.) found that prediction accuracies for SEV improved from .39 to .41 when PH was used as a covariate (Steiner et al., 2019). A kernel quality index (ISK), incorporating INC, SEV, and FDK, was used as a covariate to predict DON in a panel of 322 winter wheat genotypes where prediction accuracy between the NGS model and the MVGS model increased from .44 to .53 (J. Rutkoski et al., 2012). However, in another elite durum wheat population, a MVGS model where HD and PH were used as covariates did not significantly improve prediction accuracy for SEV (Moreno-Amores, Michel, Miedaner, Longin, & Buerstmayr, 2020). Based on previous literature, MVGS has not been performed on any FHB traits outside of SEV or DON, and traits outside of HD and PH have not been used as covariates for MVGS models used to predict FHB resistance, outside of the ISK index used by Rutkoski et al. (2012).

Although there are many studies that focus on breeding for resistance to FHB in wheat using GS, few focus on the implementation of GWAS results or multiple traits as fixed effects in GS models for the southeastern United States. This study accomplishes four objectives: (a) to evaluate SRWW genotypes for resistance to FHB in terms of resistance to initial infection (INC); resistance to spread within the head (SEV); resistance to DON; and resistance to FDK, (b) to identify novel resistance loci using a GWAS approach, (c) to perform a cross-validation analysis to compare prediction accuracies between a NGS model and a GS model with fixed effects obtained from a GS+GWAS analysis of the training population for all four FHB traits; and (d) to

perform a cross-validation analysis to compare prediction accuracies between a NGS model and MVGS models for all four FHB traits.

MATERIALS AND METHODS

Germplasm

An association mapping panel (AMP) of 354 SRWW genotypes was used for this study, consisting of 238 genotypes from the University of Arkansas, 40 from the University of Georgia, and 38 each from Louisiana State University and North Carolina State University. This panel represents the majority of source germplasm in the University of Arkansas wheat breeding program. Two checks were grown alongside the AMP, 'Bess' (PI 642794) was FHB resistant and 'Coker 9835' (PI 584846) was FHB susceptible. Bess shows native resistance to FHB in the form of Type II resistance (McKendry, Tague, Wright, & Tremain, 2007). Coker 9835 shows susceptibility in terms of INC and SEV, but to a lesser extent for DON (Horevaj, Gale, & Milus, 2011).

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 yr. The AMP was grown for a total of nine location-years over four growing seasons spanning from 2013 to 2017. During the 2013-2014 growing season, the first 120 genotypes of the full AMP, exclusively from the University of Arkansas, were grown at two locations, the Milo J. Shult Agricultural Research and Extension Center in Fayetteville, AR (FAY14) and the Newport Extension Center in Newport, AR (NPT14). During the 2014-2015 growing season, the remaining 118 genotypes, exclusively from the University of Arkansas, were grown at both locations (FAY15, NPT15). All 354 genotypes

from the AMP were then grown at both locations during the 2015-2016 (FAY16, NPT16) and 2016-2017 (FAY17, NPT17) growing seasons, with the addition of data from the Macon Ridge Research Station near Winnsboro, LA in 2016-2017 (LSU17).

Genotypes were drill-seeded at a rate of 6.0 g m^{-2} for a seeding rate of 65 kg ha^{-1} with 38 cm row spacing in two-row plots in a randomized complete block design with two replications. Plots were managed according to the recommendations for wheat in Arkansas (Kelley, 2018). The plots in FAY received 100 kg ha^{-1} of urea, whereas plots in NPT received 77 kg ha^{-1} of urea and were also supplemented with 24 kg ha^{-1} of ammonium sulfate. A combination of herbicides including Axial XL (Syngenta AG), Harmony Extra (DuPont de Nemours), and Osprey (Bayer AG) were used each year to control weeds.

The FHB disease nurseries in FAY and NPT were inoculated with maize kernels infected with *F. graminearum*. The inoculum consisted of seven different *F. graminearum* isolates collected at several research stations in Arkansas and grown each year on sterilized maize kernels. The maize inoculum was prepared based on the methods described in Horevaj, Milus, and Bluhm (2011) (Horevaj, Milus, & Bluhm, 2011). The LSU17 location was inoculated with maize kernels infected with a *F. graminearum* isolate derived from Fusarium damaged seed from the same location from the previous season. When the wheat reached a growth stage between 6 and 8 on the Feekes scale, inoculum was spread by hand in the field at a rate of ~65 kernels m^{-2}, allowing for colonization of maize kernels in the field and production of black perithecia before the wheat began to head at Feekes 10.1 (Gilbert & Woods, 2006).

After the spread of inoculum, mist irrigation was set up every sixth plot throughout the disease nursery to provide complete coverage. Mist irrigation commenced at the time perithecia were observed on the maize inoculum to provide optimal conditions for FHB infection and

spread throughout the months of April and May. Duration of misting was adjusted for each location based on the available precipitation and dew point. In order for the fungus to spread, it was important for the young wheat heads to remain moist. Timing of misting, as well as duration, varied between locations and years. During a particularly dry spring, plots may be misted longer or more often in comparison to a wet spring. In general, FAY required less total misting time than NPT. In the 2014-2015 season, FAY received a total of 720 min of misting while NPT received 784 min. Fayetteville received 480 min for both the 2014-2015 and 2015-2016 seasons. During those same seasons, NPT received 720 and 520 min, respectively. During the 2016-2017 season, FAY required 544 min of misting while NPT totaled 704 min of misting. The LSU17 location was misted nightly for 15 min, every 2 h, from 12:00 AM through 6:00 AM.

Data was collected for four FHB traits: INC, SEV, FDK, and DON. Data was also collected for HD and PH. Ratings for FHB resistance were recorded beginning at 1 wk after the mean HD (Feekes 10.5), allowing time for infection and spread of the disease. At each location, INC and SEV ratings were collected on the same day, first for INC and then SEV. The first set of field ratings were generally collected beginning in mid-May at the NPT location with the second set of ratings collected ~10 d later than the first. In comparison, field ratings in FAY typically began in late May. Incidence was recorded as a percentage of the total number of heads in a plot that showed any sign of infection regardless of how severe or contained the infection was. Severity was estimated as a percentage of total infected spikelets within each head within the plot. Both INC and SEV were recorded on a scale of 0-9% in 1% increments in plots with <10% infection and on a scale of 10-100% with increments of 5% in plots with >10% infection. In FAY, two sets of SEV and INC ratings were recorded during all 4 yr. In NPT one set of INC and SEV ratings were recorded for each plot with a second, incomplete, rating recorded between 10 d

to 2 wk later. Generally, the second set of ratings were used for further analysis, excluding cases when the first produced a higher disease rating. At LSU17, INC and SEV were recorded on a scale of 0-100% with increments of 5% on 2 May and 8 May 2017.

At maturity, each genotype was hand harvested and threshed using a Vogel Thresher (Bill's Welding). In order to retain as many damaged and shrunken kernels as possible, the thresher was set to a very low speed as seeds were collected. Seeds were stored in labeled envelopes at room temperature before being evaluated for DON and FDK.

The total FDK percentage was evaluated after harvest. Samples from each genotype were compared to a set of standards to determine the percentage of kernels damaged by FHB. The standards ranged from 0 to 75% in 10% increments. Standards were created by counting and combining damaged kernels with healthy kernels (total=300) for each FDK increment.

Analysis of DON was conducted by the Mycotoxin Diagnostic Laboratory in the Department of Plant Pathology at the University of Minnesota (St. Paul, MN). Postharvest, 50 g samples of grain from each of the 354 genotypes and both checks were sent to the University of Minnesota where DON analysis was conducted by gas chromatography-mass spectrometry.

Heading date was recorded in Julian days after 1 January, when 50% of the heads were 50% emerged from the flag leaf. As there was variation in HD between genotypes in the AMP, heading notes were recorded every other day from the onset of heading and continuing until all plots in the nursery were headed. At maturity, PH was recorded in inches from the surface of the soil to the tip of the head.

Phenotypic Data Analysis

Data was analyzed using procedures in SAS 9.4 (SAS Institute). Best linear unbiased predictions (BLUPs) were calculated for all four FHB traits using the following model in PROC MIXED with all effects and interaction treated as random:

$$Y_{ijk} = \mu + env_i + rep(env)_{ij} + genotype_k + (env\ x\ genotype)_{ik} + \varepsilon_{ijk}$$

where Y_{ijk} is the observed phenotype, μ is the overall mean, *env* is the random effect of i^{th} location-year; $rep(env)_{ij}$ is the random effect of j^{th} replication nested within the i^{th} location-year; $genotype_k$ is the random effect of the k^{th} genotype; $(env\ x\ genotype)_{ik}$ is the interaction between genotype and environment; and ε_{ijk} is the residual error term, where $\varepsilon_{ijk} \sim N(0, \boldsymbol{I}\sigma^2_\varepsilon)$, where $\boldsymbol{I}$ is an identity matrix and σ^2_ε is the residual error variance.

Phenotypic correlations were determined between all four FHB traits, as well as HD and PH using the multivariate function in JMP Pro 14.1.0 software (SAS Institute). The plot mean-based broad-sense heritability (H^2) was calculated for each trait across location-years using variance components estimated from the following equation:

$$H^2_{genotype} = \frac{\sigma^2_{genotype}}{\sigma^2_{genotype} + \frac{\sigma^2_{genotype\ x\ env}}{n_{env}} + \frac{\sigma^2_{error}}{n_{env}\ x\ n_{rep}}}$$

where $\sigma^2_{genotype}$ is the genotypic variance, $\sigma^2_{genotype\ x\ env}$ is the variance of the interaction between genotype and environment, n_{env} is the number of environments where the trait was evaluated, σ^2_{error} is the residual error variance, and n_{rep} is the number of replications within each environment. Variance components and BLUPs for each genotype for each trait were obtained from the PROC MIXED procedure in SAS 9.4 (SAS Institute).

Genotyping by Sequencing

All 354 genotypes in the AMP were genotyped using genotyping by sequencing (GBS). DNA was extracted using the Mag-Bind Plant DNA Plus kit (Omega Bio-tek), following the manufacturer's instructions. Genomic DNA was quantified using the Quant-iT PicoGreen dsDNA Assay Kit and normalized to 20 ng μL^{-1} (ThermoFisher Scientific). Genotyping by sequencing libraries were created using *Pst1-Msp1* and/or the *Pst1-Mse1* restriction enzyme combinations (Poland, Brown, Sorrells, & Jannink, 2012). The samples were pooled together at 192-plex and each pooled library was sequenced on a single lane of an Illumina Hi-Seq 2500 system (Illumina). Single nucleotide polymorphism 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 using the alignment method of Burrows-Wheeler aligner (BWA) version 0.7.10 (Appels et al., 2018; H. Li & Durbin, 2009). Raw SNP data generated from the TASSEL pipeline were filtered to remove taxa with >85% missing data and heterozygosity >30%. Genotypic data were then filtered to select for biallelic SNPs with minor allelic frequency of >5%, < 50% missing data, and heterozygosity of ≤10%. Missing data were imputed using the linkage disequilibrium - k^{th} nearest neighbor imputation (LD-*k*NNi) function in TASSEL 5.0 in order to create a final SNP panel consisting of 72,634 SNPs across the entire genome (Bradbury et al., 2007; Money et al., 2015).

Genome-Wide Association Study

A GWAS was performed using the 354 genotypes from the AMP and 72,634 GBS SNPs in order to obtain associations between SNP markers and all four disease resistance traits using a Fixed and random model Circulating Probability Unification (FarmCPU) model (X. L. Liu, Huang,

Fan, Buckler, & Zhang, 2016). The FarmCPU model is a multi-locus mixed linear model divided into a fixed effect model and a random effect model. The fixed effect model performs single marker tests where multiple associated markers were used as covariates to control for false positives. The random effect model estimated associated markers, which were used to define kinship in order to prevent model overfitting. Both models were run until there was no change in associated markers. *P* values were calculated for each marker and associated marker at each iteration (X. L. Liu et al., 2016).

The FarmCPU model was implemented using the Genomic Association and Prediction Integrated Tool (GAPIT) Version 3.0 in R version 3.6.3 software (Lipka et al., 2012; X. L. Liu et al., 2016; Tang et al., 2016). Population structure was controlled by selecting the optimal number of principal components for each trait and evaluating the respective quantile-quantile (QQ) plot. The model with the best fitting QQ plot was used for analysis. The principal component analysis (PCA) was performed in GAPIT. The ideal number of principal components for INC and DON was one, whereas the ideal number for FDK was three, and four for SEV. The GAPIT tool was also used to calculate a kinship matrix, using the method described by VanRaden (2008) in order to control for relationships between genotypes. Significant QTL were identified using an adjusted *p*-value (q) based on the Benjamini-Hochberg false discovery rate (FDR) method, where QTL with $q \leq .10$ were considered significant (Benjamini & Hochberg, 1995). Significant QTL occurring on the same chromosomes across all six traits were tested for LD using a pairwise analysis in TASSEL 5.0. Any pairwise combination with an $R^2 > .2$ were considered to be in significant LD.

GWAS Assisted Genomic Selection

A GS cross-validation analysis was performed using the rrBLUP package within the Intelligent Prediction and Association Tool (iPat), using source code from R version 3.6.2 in order to determine the prediction accuracies for all four FHB traits using a genomic BLUP model (GBLUP), which is described as follows:

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

where u is the vector of marker effects, which is assumed to have a normal distribution $u \sim N(0, \boldsymbol{G}\sigma_u^2)$, where $\boldsymbol{G}$ is the genomic relationship matrix and σ_u^2 is the variance of the individual marker effects; β is the vector of fixed effects; $\boldsymbol{X}$ is the design matrix of fixed effects, equal to one; $\boldsymbol{Z}$ is the design matrix relating genotypes to phenotypic observations (y), with m markers in columns and n phenotypes in rows; and ε_i is the residual error at the i^{th} locus, which is assumed to have a normal distribution $\varepsilon_i \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$, where $\boldsymbol{I}$ is the identity matrix and σ^2_ε is the residual error variance. The GEBV is the sum of all allele effects of a genotype (Chen & Zhang, 2018; Endelman, 2011; VanRaden, 2008).

Four separate models were compared: a naïve model with no fixed effects (NGS), a model using significant QTL for HD, from a GWAS of the training population, as fixed effects to predict the four FHB traits (GS+HD), a model using significant QTL for PH, from a GWAS of the training population, to predict the four FHB traits (GS+PH), and a model using significant QTL for each FHB resistance trait of interest, from the GWAS of the training population, as fixed effects (GS+GWAS) (Bernardo, 2014; Rice & Lipka, 2019; Spindel et al., 2016; Zhang et al., 2014). Significant QTL for all four FHB traits, HD, and PH used for the GS+HD, GS+PH, and GS+GWAS models were identified using a FarmCPU model, where the top three QTL with

a FDR of <.10 were used as fixed effects (Benjamini & Hochberg, 1995; Chen & Zhang, 2018; X. L. Liu et al., 2016; Spindel et al., 2016).

The cross-validation analysis for both models was repeated over 10 iterations, where the 354 genotypes were randomly divided into a training population of 283 genotypes and a validation population of 71 genotypes. New training and validation populations were randomly created for each iteration. A GWAS was then performed on each training population in order to obtain significant QTL that were used as fixed effects in the GS+HD, GS+PH, and GS+GWAS models. The calculated GEBVs were then compared with the phenotypic BLUPs in order to determine prediction accuracy, using a Pearson correlation from the multivariate function in JMP Pro 14.1.0 software (SAS Institute). A mean prediction accuracy was obtained from the 10 iterations for all four traits. In order to determine if one model had higher prediction accuracies compared to the other for all four FHB traits, mean prediction accuracies were compared between the four models using a generalized linear mixed linear model (GLMM) and Fisher's LSD with an α of .05, implemented in PROC GLIMMIX in SAS 9.4 (SAS Institute).

Multivariate Genomic Selection

Genetic correlations between the six traits were obtained using a genetic analysis algorithm incorporating genome-wide SNPs in a mixed linear model framework within the sommer package in R software version 3.6.3 (Covarrubias-Pazaran, 2016; Lee & van der Werf, 2016). In order to determine if a MVGS GBLUP model significantly improved prediction accuracy compared with a NGS model, a cross-validation analysis was performed using the rrBLUP package within iPat, using source code from R version 3.6.2 for all four FHB traits. For each FHB trait of interest, phenotypic data from each of the other FHB traits, as well as HD and PH, were used as covariates. Each covariate trait was first tested individually in a single-covariate

model using a five-fold cross-validation approach, where the population was randomly divided into five groups of 71 genotypes. Four of the five groups were then used as the training population to predict the fifth group, serving as the validation population, where the phenotype was set as missing, while the phenotypic data for the covariate trait were used as a fixed effect in the model. The GEBVs for the validation population were compared to the actual phenotypic values using a Pearson correlation. The process was performed over 100 iterations for a total of five single-covariate trait combinations for each FHB resistance trait of interest.

The mean prediction accuracies from the five single-covariate MVGS models were compared with a NGS model to determine if any of the covariate traits significantly improved prediction accuracies using a GLMM and Fisher's LSD with an α of .05, implemented in PROC GLIMMIX in SAS 9.4 (SAS Institute).

Covariates that produced significantly higher prediction accuracies than the NGS models were then included as covariates for multiple-covariate models to predict each FHB resistance trait. If an individual covariate did not significantly increase the prediction accuracy of the FHB resistance trait of interest in the single-covariate MVGS model, it was not included in the multiple-covariate MVGS models. Significant covariates were tested in all possible combinations in multiple-covariate MVGS models using the cross-validation method described above. In total, the mean prediction accuracies from 17 models- including the NGS, single-covariate, and multiple-covariate models- were compared for each FHB resistance trait of interest using a GLMM and Fisher's LSD with an α of .05, implemented in PROC GLIMMIX in SAS 9.4 (SAS Institute).

RESULTS

Variation in Fusarium Head Blight Traits

All four FHB traits had significant variation within the AMP with a range between 0 and 100% for INC and SEV at the FAY16 and NPT17 location-years and for FDK at the FAY16 and NPT16 location-years. The largest range in DON of .38-26.90 μg g^{-1} occurred in FAY16. The highest mean value for DON was found in FAY15 at 29.5 μg g^{-1}. Significant ($p \leq .0001$) genotypic, environmental, and genotype x environment effects were observed in an ANOVA for all four FHB traits, as well as for HD and PH (Table 1).

Significant phenotypic correlations ($p \leq .0001$) were observed between all four FHB traits. Severity and DON, were significantly correlated with HD, whereas all FHB traits but DON were significantly correlated with PH. Traits that had strong phenotypic correlations also had strong genetic correlations (Table 2). The highest heritability among the four FHB traits was for FDK ($H^2 = .82$), followed by DON ($H^2 = .79$), SEV ($H^2 = .78$), and INC ($H^2 = .38$) (Table 1). Evaluating the AMP across nine location-years allowed for moderately high heritability values, while also limiting the significant interaction between genotype and environment (Table 1).

Population Structure

Genotyping by sequencing identified 72,634 SNPs across the entire wheat genome after filtering. There was an uneven distribution of SNPs between the three genomes, with the B genome having the largest number of SNPs (34,268), followed by the A (27,078) and D genomes (9,513), with 1,775 SNPs that were unaffiliated with any chromosome, which was consistent with other studies using GBS SNPs (Arruda, Brown, et al., 2016). Chromosome 2B had the largest number of SNPs with 7,502 whereas chromosome 4D had the smallest number with 653. The percentage

heterozygosity of the dataset was 3.5%, whereas the average minor allele frequency was 19.6%. The allelic distributions were 25.7%, 25.2%, 23.0%, and 22.5% for the C, G, A, and T alleles, respectively.

A PCA found two primary clusters within the AMP, where genotypes from all four breeding programs appeared in both clusters. Although some sub-clustering between individual breeding programs was also observed within the two primary clusters, it is hypothesized that the main effect was due to the presence or absence of stem rust (*Puccinia graminis* f. sp. *tritici*) and powdery mildew resistance genes, *Sr36/Pm6*, located on the translocation from *Triticum timopheevii* Zhuk. (Benson, Brown-Guedira, Murphy, & Sneller, 2012; Nyquist, 1962; Sarinelli et al., 2019). When the region associated with the *Sr36/Pm6* translocation on chromosome 2B was removed from the dataset, the deviation between clusters disappeared, indicating that the translocation was contributing to population structure. Overall, the population structure for the AMP was low, with the first three principal components accounting for only 5.9, 5.0, and 3.5% of the total genetic variation, respectively (Figure 1). For the purpose of the GWAS, the optimal number of principal components used for each individual trait, was determined based on the deviation of expected *p* values from observed *p* values based on a QQ plot for each individual trait. The ideal number of principal components for INC and DON was one, whereas the ideal number for was three for FDK and four for SEV.

Quantitative Trait Loci

The multi-locus FarmCPU model identified 42 significant ($q \leq .10$) QTL for all four FHB traits, with 10 significant QTL for DON, 11 for FDK, nine for INC, and 12 for SEV across 17 of the 21 wheat chromosomes (Appendix 1). The GWAS for DON was performed with a limited set of location-years (FAY15, NPT15, NPT16, and LSU17) because of low accumulation in the other

location-years. The summary of significant QTL showing the chromosome name, physical position, allelic variation, minor allele frequency, and allelic effect is shown in Table 3. Based on the pairwise LD test, none of the significant QTL had a pairwise LD $R^2 > .2$. One significant marker, S4B_577008759 was identified for the traits FDK and SEV on the long arm of chromosome 4B at 577.0 Mb (Table 3, Appendices 1b and 1d).

GWAS-Assisted Genomic Selection

The top three significant QTL exceeding the FDR ($q \leq .10$) threshold were selected using the FarmCPU model to use as fixed effects for the 10 randomly selected training population subsets for all four FHB traits for the GS+GWAS analysis with the GBLUP model. The top three significant QTL exceeding the FDR ($q \leq .10$) threshold for HD and PH, from each of the training population subsets, were also selected to use as fixed effects to predict the four FHB traits for the GS+HD and GS+PH GBLUP models (Appendix 2).

The NGS model had the highest prediction accuracies for all four FHB traits. When comparing the GS+GWAS model with the NGS model, the NGS model had significantly higher mean prediction accuracies for all FHB traits. The NGS models also had higher prediction accuracies than the GS+HD and GS+PH models; however, these differences were not significant (Figure 2). The GS+PH model had the second highest prediction accuracy for all FHB traits except for SEV (Figure 2).

Multivariate Genomic Selection

Naïve GS models were compared with 16 single-covariate and multiple-covariate MVGS models using a five-fold cross-validation approach over 500 iterations to predict four FHB traits. Traits were selected as covariates for the multiple-covariate models if their use in single-covariate

models significantly improved prediction accuracy compared to the NGS models. The MVGS models significantly improved prediction accuracies for all four FHB traits (Figure 3).

Incidence saw the largest increase in prediction accuracy with a 69.0% improvement between the NGS model (r = .40) and the seven-best performing MVGS models (r = .67). All of the single-covariate models significantly improved prediction accuracy compared to the NGS model except for HD. Incidence was also the only trait where a single-covariate MVGS model had the highest prediction accuracy, although it was statistically equal to the multiple-covariate MVGS models (Figure 3c).

The trait with the next largest improvement in prediction accuracy was SEV, where the top four MVGS models (r = .81) improved prediction accuracy by 40.5% compared to the NGS model (r = .58). All of the single covariate models except for HD significantly improved prediction accuracy for SEV compared with the NGS model; therefore, HD was not included as a covariate in the multiple-covariate models. The PH+INC+FDK, PH+DON+INC+FDK, INC+FDK, and DON+INC+FDK MVGS models each had the highest prediction accuracy for SEV (Figure 3d).

Fusarium damaged kernels had the third largest increase in prediction accuracy, where the MVGS model with PH, DON, and SEV covariates (r = .73) had a 37.5% improvement over the NGS model (r = .53). All single-covariate models except for HD significantly improved prediction accuracy for FDK compared with the NGS model; therefore, HD was not included as a covariate in the multiple-covariate models. The PH+INC+FDK MVGS model did not have a significantly higher prediction accuracy than the PH+DON+INC+SEV, DON+SEV, or DON+INC+SEV MVGS models (Figure 3b).

The trait with the smallest percent improvement in prediction accuracy was DON, where the HD+FDK MVGS model (r = .74) improved prediction accuracy by 23.7% compared to the NGS model (r = .60). Single-covariate models with all four FHB traits and HD significantly improved prediction accuracy for DON. The single-covariate model with PH was the only model that did not significantly improve prediction accuracy, therefore PH was not included in the multiple-covariate MVGS models for DON. DON was the only trait that did not have a significant improvement in prediction accuracy from PH, and it was the only trait that had a significant improvement in prediction accuracy from HD. The prediction accuracy for the HD+FDK MVGS model was not significantly higher than the HD+FDK+SEV, HD+INC+FDK, or HD+INC+SEV+FDK MVGS models (Figure 3a).

DISCUSSION

Fusarium head blight is a major fungal disease that causes significant yield losses for farmers, largely due to shrunken FDK, as well as economic losses due to the accumulation of the mycotoxin DON. Since the use of fungicides and cultural control practices provide only marginal control of FHB in susceptible cultivars, it is important to breed for genetic resistance. This can often prove difficult, as genetic resistance to FHB is highly quantitative. The quantitative nature of FHB resistance provides breeders an opportunity to include all possible SNPs using GS (Ban, 2000; H. Buerstmayr & Lemmens, 2015; Hoffstetter et al., 2016; Larkin et al., 2019; Mirdita et al., 2015; Steiner et al., 2017).

Genome-Wide Association Study

Numerous QTL mapping and GWAS analyses have been performed on multiple FHB traits in wheat over the last couple decades (H. Buerstmayr et al., 2009; M. Buerstmayr, Steiner, &

Buerstmayr, 2019). Applications for QTL identified through GWAS include validation of QTL identified through other GWAS and QTL mapping studies (M. Buerstmayr et al., 2019). Another application for QTL identified in this study could be as targets for fine mapping and positional cloning. As of now, positional cloning for FHB resistance genes have been dedicated to exotic germplasm (H. Y. Jia et al., 2018; G. Q. Li et al., 2019; Rawat et al., 2016; Su, Jin, Zhang, & Bai, 2018; Wang et al., 2020).

Deoxynivalenol Accumulation

For DON, the GWAS identified 10 significant QTL across eight different chromosomes. The most notable of these occurred on chromosome arms 2BL and 3BL. The resistance QTL *Qfhs.don-3BL* was identified in the SRWW cultivar 'Ernie', which led to a reduction in DON concentration in three different QTL mapping studies (Abate, Liu, & McKendry, 2008; S. Liu et al., 2007; S. Y. Liu et al., 2013). A QTL on chromosome 2B was identified in the mapping population Ernie/'MO 94-317' related to DON reduction in the SRWW cultivar Ernie, although it could only be identified in one of the 2 yr of the study (Abate et al., 2008; S. Liu et al., 2007). A QTL on 2BL was also identified using an FHB index in the SRWW mapping population 'Becker'/'Massey' (S. Y. Liu et al., 2013).

Fusarium Damaged Kernels

Eleven QTL were significantly associated with FDK across seven chromosomes, most notably on chromosome arms 4BL, 3AL, 1BL, 5BL and 5DL. The SNP associated with the QTL on chromosome 4BL was at the same physical location as a SNP associated with SEV in the AMP. The QTL also occurred in the same 1.5 logarithm of the odds (LOD) region identified in two

other studies in hard red winter wheat, even though the region was associated with SEV, as discussed below (Clinesmith et al., 2019; da Silva et al., 2019).

Fusarium head blight resistance has been identified on the long arm of chromosome 3A primarily for SEV in two spring wheat QTL mapping populations from North Dakota ('ND2603'/'Butte') and South America ('Frontana'/'SeriM82') (Anderson et al., 2001; Mardi et al., 2006). It had also been identified in the European winter wheat mapping population 'Arina'/'Forno' (Paillard et al., 2004). The QTL had also been identified in an Italian hexaploid by tetraploid mapping population ('02-5B-318'/'Saragolla') (Giancaspro, Giove, Zito, Blanco, & Gadaleta, 2016). However, it appears that FDK resistance has not been identified on 3AL until this study.

The FHB resistance QTL *Qfhb.nc□1B* was identified in the SRWW mapping population, 'NC-Neuse'/'AGS2000'. This QTL was associated with NC-Neuse and the resistant allele significantly reduced INC, FDK, and DON (Petersen et al., 2016). Another QTL, *Qfhs.vt□1BL*, for FHB index was identified in the SRWW line, 'VA00W□38' based on the mapping population (VA00W□38/'Pioneer 26R46') (S. Y. Liu et al., 2012).

A QTL associated with FDK resistance on chromosome 5B was identified in an AMP of 256 SRWW breeding lines across 2 yr of data, reducing FDK by .64% of the population mean (Tessmann, Dong, & Van Sanford, 2019). A QTL associated with the reduction of INC and DON was identified on the long arm of chromosome 5B, in the NC-Neuse/AGS2000 mapping population, where AGS2000 contributed the resistance allele, however the QTL region also contained the *vrn-B1* vernalization locus (Petersen et al., 2016).

According to a meta-analysis of sources of FHB resistance, the only QTL associated with FHB resistance on chromosome 5D occurred in Asian germplasm, however the chromosome was associated with INC, SEV, and FDK. The QTL was identified in the mapping population 'Wangshuibai'/'Wheaton' (S. Y. Liu, Hall, Griffey, & McKendry, 2009; Yu, Bai, Zhou, Dong, & Kolb, 2008). Another QTL on 5DL, associated with SEV, was identified in a diverse AMP containing germplasm from China, Italy, Japan, and Mexico, where the QTL occurred in the 546.1 to 547.3 Mb region, whereas the QTL we identified occurred at 548.3 Mb. The QTL identified in the study was also associated with resistance to *Soil-borne wheat mosaic virus* in wheat (Hu et al., 2020).

Incidence

Nine significant QTL were identified for INC across seven chromosomes, most notably on chromosome arms 2BS, 7BL, 6AS, and 5DS. Resistance QTL on 2BS for INC had been discovered in the SRWW mapping populations 'Patterson'/'Goldfield' and Ernie/MO 94-317 (Gilsinger, Kong, Shen, & Ohm, 2005; S. Liu et al., 2007).

FHB resistance QTL have been identified on the long arm of chromosome 7B in the doubled haploid (DH) spring wheat mapping population, 'DH181'/'AC Foremost', as well as a tetraploid wheat AMP (Ghavami et al., 2011; Yang, Gilbert, Fedak, & Somers, 2005). A significant QTL associated with INC and SEV on 7BL was also identified in a European AMP of winter and spring wheat genotypes (Kollers et al., 2013).

A QTL for INC on the short arm of chromosome 6A was identified in the SRWW mapping population NC-Neuse/Bess, where Bess was the source of resistance (Petersen et al.,

2017). The QTL identified in Petersen et al. (2017) overlapped a QTL that was identified in the NC-Neuse/AGS2000 mapping population as well (Petersen et al., 2016).

Most QTL identified for FHB resistance on chromosome 5B were identified on the long arm, particularly for INC (Hu et al., 2020; S. Y. Liu et al., 2009; Yu et al., 2008). The only novel QTL for FHB resistance on 5BS was identified in an AMP of Chinese and Japanese landraces, and it was associated with SEV (T. Li et al., 2016).

Severity

The GWAS identified 12 QTL associated with SEV across 12 chromosomes. The most notable of these occurred on chromosome arms 3BL, 4AL, and 4BL. Several other QTL mapping and GWAS studies have also identified FHB resistance in the form of SEV on 3BL. Native resistance QTL on 3BL associated with SEV have been found in several SRWW cultivars, such as Massey and Ernie, as identified in two mapping populations, Becker/Massey and Ernie/MO 94-317 (S. Y. Liu et al., 2013). The QTL, *Qfhb.nc-3B.2*, in Bess was identified in the NC-Neuse/Bess mapping population (Petersen et al., 2017). A similar QTL was also identified in 'Truman' from the Truman/NC-Neuse DH population (Islam et al., 2016). The long arm of chromosome 3B has also been associated with SEV in European winter wheat germplasm, particularly in the Swiss cultivar Arina, as found in the mapping populations Arina/Forno and 'Capo'/Arina, and the French cultivar, 'Apache' (M. Buerstmayr & Buerstmayr, 2015; Holzapfel et al., 2008; Paillard et al., 2004).

Fusarium head blight resistance QTL related to SEV have also been found on the long arm of chromosome 4B. In SRWW, the QTL *Qfhs.umc-4BL* was identified in Ernie in the Ernie/MO 94-317 mapping population (S. Liu et al., 2007). In the hard red winter wheat

population, ‘Art’/’Everest’, Clinesmith et al. found the QTL *Qfhb-4B.1* on 4BL for SEV, where S4B_577008759 occurs within the 1.5 LOD interval. It was suggested that *Qfhb-4B.1* was associated with the major resistance QTL *Fhb4*, which was found in the Chinese landrace Wangshuibai (Clinesmith et al., 2019; da Silva et al., 2019). This indicates that further research should be performed on Arkansas SRWW germplasm in order to further map the QTL on 4BL, particularly since a QTL for FDK also appeared in the same region.

The QTL *QFhs.fal-4AL* was identified in the Swiss mapping population Arina/Forno, conferring FHB resistance through SEV (Paillard et al., 2004). Another QTL on the long arm of chromosome 4A was also described in the spring wheat mapping population, DH181/AC Foremost, which was also for SEV (Yang et al., 2005). Two meta-studies of FHB resistance QTL have also shown QTL identified on 4AL conferring resistance to FHB through SEV (H. Buerstmayr et al., 2009; S. Y. Liu et al., 2009).

GWAS-Assisted Genomic Selection for Fusarium Head Blight Traits

The general assumption of the GBLUP model is that all SNP effects have an equal variance (Endelman, 2011). In reality, most traits have at least some genes that have larger effects on a trait than others do. Theoretically, if large-effect markers were used as fixed effects in the GBLUP model, prediction accuracies would improve. In the case of FHB traits, Arruda et al. (2016) found that a ridge regression BLUP (RR-BLUP) model with marker fixed effects significantly improved prediction accuracies for SEV, INC, an FHB index, FDK, and an INC-SEV-kernel index, but there was not a significant difference for DON (Arruda, Lipka, et al., 2016).

We observed that the use of significant markers identified through GWAS actually significantly decreased mean prediction accuracy for all four traits. One of the possibilities could be that the significant markers implemented for GS did not account for enough variation in the four FHB traits. It has been noted that using individual markers as fixed effects in a GS model was never disadvantageous, except when the amount of variance explained by the marker was <10% (Bernardo, 2014). In our case, none of the markers used for our GWAS exceeded 5% of the variation for any of the traits tested. Whereas at least one of the SNPs used by Arruda et al. (2016) accounted for >10% of the variation in FHB resistance (Arruda, Lipka, et al., 2016; Bernardo, 2014). Approximately 5.5% of the 273 genotypes in their population also contained resistance alleles for the major FHB resistance QTL, *Fhb1*, whereas our population did not have any genotypes containing resistance alleles for *Fhb1* (Arruda, Lipka, et al., 2016). As a result, nearly all of the resistance in our population came in the form of small-effect QTL, which contribute a lower level of resistance compared with sources of resistance, such as *Fhb1* (Steiner et al., 2017). Bernardo also stated that the use of marker fixed effects in a GS model generally did not improve prediction accuracy if the trait heritability exceeded 50%, which was the case for all of our FHB traits, with the exception of INC (Table 1) (Bernardo, 2014).

Another study found that the use of marker fixed effects with multiple GS models did not significantly improve prediction accuracies for FHB traits. Prediction accuracies were actually lower for FDK, INC, and SEV when marker fixed effects were implemented, aligning with what we observed in our study (J. Rutkoski et al., 2012).

Multivariate Genomic Selection for Fusarium Head Blight Traits

Multivariate GS uses secondary traits that are genetically correlated with a trait of interest as covariates in a GS model, which can increase prediction accuracy, especially when the trait of

interest has a low heritability and the covariate trait has a high heritability (Calus & Veerkamp, 2011; Guo et al., 2014; Y. Jia & Jannink, 2012). Multiple studies have evaluated the use of MVGS for FHB resistance in wheat, however they have primarily focused on SEV as a trait of interest and only used HD and PH as covariates (Moreno-Amores et al., 2020; Schulthess et al., 2018; Steiner et al., 2019). An MVGS model was also used to predict DON using an INC, SEV, and FDK index (J. Rutkoski et al., 2012).

The use of HD and PH as covariates in MVGS models are ideal as both traits are associated with FHB resistance and are easy to phenotype (Moreno-Amores et al., 2020). Wheat genotypes with earlier HD often have higher FHB infections compared to later heading genotypes (Gervais et al., 2003; Paillard et al., 2004; Schmolke et al., 2005). A negative correlation is also typically observed between FHB resistance and PH, where taller genotypes are more likely to escape infection compared to shorter genotypes (Gervais et al., 2003; Jenkinson & Parry, 1994). We found that HD had no significant impact on prediction accuracy as a covariate for all traits except for DON. The low impact from HD on MVGS for FHB resistance was also observed in other studies when predicting SEV (Moreno-Amores et al., 2020; Schulthess et al., 2018; Steiner et al., 2019). The inclusion of PH as a covariate significantly increased prediction accuracies for all traits except for DON, as observed with SEV in other studies (Schulthess et al., 2018; Steiner et al., 2019)

Our experiment is unique in that it uses multiple FHB traits as covariates to predict multiple forms of FHB resistance. This can be particularly helpful for predicting traits that are evaluated post-harvest, such as DON or FDK, while using traits collected during the season, such as SEV or INC, as covariates (J. Rutkoski et al., 2012; Steiner et al., 2019).

In our population, we found that MVGS significantly improved prediction accuracies for all four FHB traits. Increases in prediction accuracy were especially evident when the covariate traits were strongly correlated with the trait of interest (Table 2, Figure 3). The largest improvements between single-covariate MVGS models and NGS models were observed when SEV was used as a covariate to predict FDK, SEV was used to predict INC, and vice-versa, where strong correlations between the abovementioned traits were present (Figure 3). Previous studies using MVGS have showed similar results between traits with strong genetic correlations, especially when they exceeded .5 (Calus & Veerkamp, 2011). However, we observed significant increases in prediction accuracy for genetic correlations as low as .24.

The heritability of all evaluated traits was relatively high, with the exception of INC (Table 1). Previous research has shown that the prediction accuracy of low-heritability traits can significantly increase with the use of high heritability covariates (Calus & Veerkamp, 2011; Guo et al., 2014; Y. Jia & Jannink, 2012). We also observed this when using INC as the trait of interest, where the inclusion of other high-heritability traits as covariates (DON, FDK, PH, and SEV) resulted in significant increases in prediction accuracy, even when there was a lower genetic correlation between the traits.

CONCLUSIONS

As FHB resistance is a quantitative trait in wheat, breeding for resistance has been a complicated task for wheat breeders. In this study, QTL for FHB resistance were identified for four different traits, including DON, FDK, INC, and SEV in regions previously reported in QTL mapping and GWAS studies, particularly on the long arms of chromosomes 3B and 4B. The identification of QTLs in our panel of 354 SRWW breeding genotypes adapted to the southeastern United States

indicates that previously identified QTL are contributing to genetic resistance in this population. The SNPs identified in this study can be implemented for MAS in wheat breeding programs.

This study also found that the use of a NGS model and a GS+GWAS model could provide high prediction accuracies for all four FHB traits. However, the use of a GS+GWAS model did not improve prediction accuracies for any of the four FHB traits compared to the NGS model. This is likely due to the small amount of variance covered by each of the markers and the high heritability of the training population for each of the traits. However, MVGS models did significantly improve prediction accuracy for FHB traits compared to NGS models, especially when there was a strong correlation between the covariate traits and the predicted trait. This indicates that MVGS could be successfully used when breeding for FHB resistance. Plant breeders must consider these factors before using GS for FHB resistance in their wheat breeding programs, as results can vary across populations.

REFERENCES

Abate, Z. A., Liu, S., & McKendry, A. L. (2008). Quantitative trait loci associated with deoxynivalenol content and kernel quality in the soft red winter wheat 'Ernie'. *Crop Science, 48*(4), 1408-1418. doi:10.2135/cropsci2007.07.0411

Anderson, J. A., Stack, R. W., Liu, S., Waldron, B. L., Fjeld, A. D., Coyne, C., . . . Frohberg, R. C. (2001). DNA markers for Fusarium head blight resistance QTLs in two wheat populations. *Theoretical and Applied Genetics, 102*(8), 1164-1168. doi:10.1007/s001220000509

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., . . . Manuscript Writing, T. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science, 361*(6403), 661-+. doi:10.1126/science.aar7191

Arruda, M. P., Brown, P., Brown-Guedira, G., Krill, A. M., Thurber, C., Merrill, K. R., . . . Kolb, F. L. (2016). Genome-wide association mapping of Fusarium head blight resistance in wheat using genotyping-by-sequencing. *Plant Genome, 9*(1), 14. doi:10.3835/plantgenome2015.04.0028

Arruda, M. P., Brown, P. J., Lipka, A. E., Krill, A. M., Thurber, C., & Kolb, F. L. (2015). Genomic Selection for Predicting Fusarium Head Blight Resistance in a Wheat Breeding Program. *Plant Genome, 8*(3), 12. doi:10.3835/plantgenome2015.01.0003

Arruda, M. P., Lipka, A. E., Brown, P. J., Krill, A. M., Thurber, C., Brown-Guedira, G., . . . Kolb, F. L. (2016). Comparing genomic selection and marker-assisted selection for Fusarium head blight resistance in wheat (Triticum aestivum L.). *Molecular Breeding, 36*(7), 11. doi:10.1007/s11032-016-0508-5

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. doi:10.2135/cropsci2012.09.0526

Ban, T. (2000). Analysis of quantitative trait loci associated with resistance to fusarium head blight caused by Fusarium graminearum Schwabe and of resistance mechanisms in wheat (Triticum aestivum L.). *Breeding Science, 50*(2), 131-137. doi:10.1270/jsbbs.50.131

Benjamini, Y., & Hochberg, Y. (1995). CONTROLLING THE FALSE DISCOVERY RATE - A PRACTICAL AND POWERFUL APPROACH TO MULTIPLE TESTING. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 57*(1), 289-300. doi:10.1111/j.2517-6161.1995.tb02031.x

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. doi:10.3835/plantgenome2011.11.0027

Bernardo, R. (2014). Genomewide Selection when Major Genes Are Known. *Crop Science, 54*(1), 68-75. doi:10.2135/cropsci2013.05.0315

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

Bian, Y., & Holland, J. B. (2017). Enhancing genomic prediction with genome-wide association studies in multiparental maize populations. *Heredity, 118*(6), 585-593. doi:10.1038/hdy.2017.4

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics, 23*(19), 2633-2635. doi:10.1093/bioinformatics/btm308

Breseghello, F., & Sorrells, M. E. (2006). Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Science, 46*(3), 1323-1330. doi:10.2135/cropsci2005.09-0305

Buerstmayr, H., Ban, T., & Anderson, J. A. (2009). QTL mapping and marker-assisted selection for Fusarium head blight resistance in wheat: a review. *Plant Breeding, 128*(1), 1-26. doi:10.1111/j.1439-0523.2008.01550.x

Buerstmayr, H., & Lemmens, M. (2015). Breeding healthy cereals: genetic improvement of Fusarium resistance and consequences for mycotoxins. *World Mycotoxin Journal, 8*(5), 591-602. doi:10.3920/wmj2015.1889

Buerstmayr, M., & Buerstmayr, H. (2015). Comparative mapping of quantitative trait loci for Fusarium head blight resistance and anther retention in the winter wheat population Capo x Arina. *Theoretical and Applied Genetics, 128*(8), 1519-1530. doi:10.1007/s00122-015-2527-8

Buerstmayr, M., Steiner, B., & Buerstmayr, H. (2019). Breeding for Fusarium head blight resistance in wheat-Progress and challenges. *Plant Breeding*, 26. doi:10.1111/pbr.12797

Calus, M. P. L., & Veerkamp, R. F. (2011). Accuracy of multi-trait genomic selection using different methods. *Genetics Selection Evolution, 43*, 14. doi:10.1186/1297-9686-43-26

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

Clinesmith, M. A., Fritz, A. K., da Silva, C. L., Bockus, W. W., Poland, J. A., Dowell, F. E., & Peiris, K. H. S. (2019). QTL Mapping of Fusarium Head Blight Resistance in Winter Wheat Cultivars 'Art' and 'Everest'. *Crop Science, 59*(3), 911-924. doi:10.2135/cropsci2018.04.0276

Covarrubias-Pazaran, G. (2016). Genome-Assisted Prediction of Quantitative Traits Using the R Package sommer. *Plos One, 11*(6), 15. doi:10.1371/journal.pone.0156744

Covarrubias-Pazaran, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., . . . 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*, 13. doi:10.3389/fpls.2018.01310

da Silva, C. L., Fritz, A., Clinesmith, M., Poland, J., Dowell, F., & Peiris, K. (2019). QTL mapping of Fusarium head blight resistance and deoxynivalenol accumulation in the Kansas wheat variety "Everest'. *Molecular Breeding, 39*(3), 21. doi:10.1007/s11032-019-0937-z

Dong, H. X., Wang, R., Yuan, Y. P., Anderson, J., Pumphrey, M., Zhang, Z. W., & Chen, J. L. (2018). Evaluation of the Potential for Genomic Selection to Improve Spring Wheat Resistance to Fusarium Head Blight in the Pacific Northwest. *Frontiers in Plant Science, 9*, 15. doi:10.3389/fpls.2018.00911

Endelman, J. B. (2011). Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *Plant Genome, 4*(3), 250-255. doi:10.3835/plantgenome2011.08.0024

FDA. (2010). *Guidance for Industry and FDA: Advisory Levels for Deoxynivalenol (DON) in Finished Wheat Products for Human Consumption and Grains and Grain By-Products used for Animal Feed*. Retrieved from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-industry-and-fda-advisory-levels-deoxynivalenol-don-finished-wheat-products-human

Flint-Garcia, S. A., Thuillet, A. C., Yu, J. M., Pressoir, G., Romero, S. M., Mitchell, S. E., . . . Buckler, E. S. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant Journal, 44*(6), 1054-1064. doi:10.1111/j.1365-313X.2005.02591.x

Gervais, L., Dedryver, F., Morlais, J. Y., Bodusseau, V., Negre, S., Bilous, M., . . . Trottet, M. (2003). Mapping of quantitative trait loci for field resistance to Fusarium head blight in an European winter wheat. *Theoretical and Applied Genetics, 106*(6), 961-970. doi:10.1007/s00122-002-1160-5

Ghavami, F., Elias, E. M., Mamidi, S., Ansari, O., Sargolzaei, M., Adhikari, T., . . . Kianian, S. F. (2011). Mixed Model Association Mapping for Fusarium Head Blight Resistance in Tunisian-Derived Durum Wheat Populations. *G3-Genes Genomes Genetics, 1*(3), 209-218. doi:10.1534/g3.111.000489

Giancaspro, A., Giove, S. L., Zito, D., Blanco, A., & Gadaleta, A. (2016). Mapping QTLs for Fusarium Head Blight Resistance in an Interspecific Wheat Population. *Frontiers in Plant Science, 7*, 13. doi:10.3389/fpls.2016.01381

Gilbert, J., & Woods, S. (2006). *Strategies and considerations for multi-location FHB screening nurseries.* Paper presented at the The global Fusarium initiative for international collaboration: a strategic planning workshop. CIMMYT, El Batàn, Mexico.

Gilsinger, J., Kong, L., Shen, X., & Ohm, H. (2005). DNA markers associated with low Fusarium head blight incidence and narrow flower opening in wheat. *Theoretical and Applied Genetics, 110*(7), 1218-1225. doi:10.1007/s00122-005-1953-4

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*, 7. doi:10.1186/1471-2156-15-30

Heffner, E. L., Sorrells, M. E., & Jannink, J. L. (2009). Genomic Selection for Crop Improvement. *Crop Science, 49*(1), 1-12. doi:10.2135/cropsci2008.08.0512

Hoffstetter, A., Cabrera, A., & Sneller, C. (2016). Identifying Quantitative Trait Loci for Economic Traits in an Elite Soft Red Winter Wheat Population. *Crop Science, 56*(2), 547-558. doi:10.2135/cropsci2015.06.0332

Holzapfel, J., Voss, H. H., Miedaner, T., Korzun, V., Haberle, J., Schweizer, G., . . . Hartl, L. (2008). Inheritance of resistance to Fusarium head blight in three European winter wheat populations. *Theoretical and Applied Genetics, 117*(7), 1119-1128. doi:10.1007/s00122-008-0850-z

Horevaj, P., Gale, L. R., & Milus, E. A. (2011). Resistance in Winter Wheat Lines to Initial Infection and Spread Within Spikes by Deoxynivalenol and Nivalenol Chemotypes of Fusarium graminearum. *Plant Disease, 95*(1), 31-37. doi:10.1094/pdis-03-10-0167

Horevaj, P., Milus, E. A., & Bluhm, B. H. (2011). A real-time qPCR assay to quantify Fusarium graminearum biomass in wheat kernels. *Journal of Applied Microbiology, 111*(2), 396-406. doi:10.1111/j.1365-2672.2011.05049.x

Hu, W. J., Gao, D. R., Wu, H. Y., Liu, J., Zhang, C. M., Wang, J. C., . . . Lu, C. B. (2020). Genome-wide association mapping revealed syntenic loci QFhb-4AL and QFhb-5DL for Fusarium head blight resistance in common wheat (Triticum aestivum L.). *Bmc Plant Biology, 20*(1), 13. doi:10.1186/s12870-019-2177-0

Islam, M. S., Brown-Guedira, G., Van Sanford, D., Ohm, H., Dong, Y. H., & McKendry, A. L. (2016). Novel QTL associated with the Fusarium head blight resistance in Truman soft red winter wheat. *Euphytica, 207*(3), 571-592. doi:10.1007/s10681-015-1550-9

Jenkinson, P., & Parry, D. W. (1994). SPLASH DISPERSAL OF CONIDIA OF FUSARIUM-CULMORUM AND FUSARIUM-AVENACEUM. *Mycological Research, 98*, 506-510. doi:10.1016/s0953-7562(09)80468-1

Jia, H. Y., Zhou, J. Y., Xue, S. L., Li, G. Q., Yan, H. S., Ran, C. F., . . . Ma, Z. Q. (2018). A journey to understand wheat Fusarium head blight resistance in the Chinese wheat landrace Wangshuibai. *Crop Journal, 6*(1), 48-59. doi:10.1016/j.cj.2017.09.006

Jia, Y., & Jannink, J. L. (2012). Multiple-Trait Genomic Selection Methods Increase Genetic Value Prediction Accuracy. *Genetics, 192*(4), 1513-+. doi:10.1534/genetics.112.144246

Kelley, J. (2018). 2018 Arkansas Wheat Quick Facts. In: University of Arkansas Division of Agriculture Research and Extension.

Kollers, S., Rodemann, B., Ling, J., Korzun, V., Ebmeyer, E., Argillier, O., . . . Roder, M. S. (2013). Whole Genome Association Mapping of Fusarium Head Blight Resistance in European Winter Wheat (Triticum aestivum L.). *Plos One, 8*(2), 10. doi:10.1371/journal.pone.0057500

Larkin, D. L., Lozada, D. N., & Mason, R. E. (2019). Genomic Selection-Considerations for Successful Implementation in Wheat Breeding Programs. *Agronomy-Basel, 9*(9), 18. doi:10.3390/agronomy9090479

Lee, S. H., & van der Werf, J. H. J. (2016). MTG2: an efficient algorithm for multivariate linear mixed model analysis based on genomic information. *Bioinformatics, 32*(9), 1420-1422. doi:10.1093/bioinformatics/btw012

Li, G. Q., Zhou, J. Y., Jia, H. Y., Gao, Z. X., Fan, M., Luo, Y. J., . . . Ma, Z. Q. (2019). Mutation of a histidine-rich calcium-binding-protein gene in wheat confers resistance to Fusarium head blight. *Nature Genetics, 51*(7), 1106-+. doi:10.1038/s41588-019-0426-7

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

Li, T., Zhang, D. D., Zhou, X. L., Bai, G. H., Li, L., & Gu, S. L. (2016). Fusarium head blight resistance loci in a stratified population of wheat landraces and varieties. *Euphytica, 207*(3), 551-561. doi:10.1007/s10681-015-1539-4

Lipka, A. E., Tian, F., Wang, Q. S., Peiffer, J., Li, M., Bradbury, P. J., . . . Zhang, Z. W. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics, 28*(18), 2397-2399. doi:10.1093/bioinformatics/bts444

Liu, S., Abate, Z. A., Lu, H., Musket, T., Davis, G. L., & McKendry, A. L. (2007). QTL associated with Fusarium head blight resistance in the soft red winter wheat Ernie. *Theoretical and Applied Genetics, 115*(3), 417-427. doi:10.1007/s00122-007-0577-2

Liu, S. Y., Christopher, M. D., Griffey, C. A., Hall, M. D., Gundrum, P. G., & Brooks, W. S. (2012). Molecular Characterization of Resistance to Fusarium Head Blight in U.S. Soft Red Winter Wheat Breeding Line VA00W-38. *Crop Science, 52*(5), 2283-2292. doi:10.2135/cropsci2012.03.0144

Liu, S. Y., Griffey, C. A., Hall, M. D., McKendry, A. L., Chen, J. L., Brooks, W. S., . . . Schmale, D. G. (2013). Molecular characterization of field resistance to Fusarium head blight in two US soft red winter wheat cultivars. *Theoretical and Applied Genetics, 126*(10), 2485-2498. doi:10.1007/s00122-013-2149-y

Liu, S. Y., Hall, M. D., Griffey, C. A., & McKendry, A. L. (2009). Meta-analysis of QTL associated with Fusarium head blight resistance in wheat. *Crop Science, 49*(6), 1955-1968. doi:10.2135/cropsci2009.03.0115

Liu, X. L., Huang, M., Fan, B., Buckler, E. S., & Zhang, Z. W. (2016). Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLoS Genetics, 12*(2), 24. doi:10.1371/journal.pgen.1005767

Mardi, M., Pazouki, L., Delavar, H., Kazemi, M. B., Ghareyazie, B., Steiner, B., . . . Buerstmayr, H. (2006). QTL analysis of resistance to Fusarium head blight in wheat using a 'Frontana'-derived population. *Plant Breeding, 125*(4), 313-317. doi:10.1111/j.1439-0523.2006.01228.x

Mason, R. E., Addison, C. K., Babar, A., Acuna, A., Lozada, D., Subramanian, N., . . . Johnson, J. (2018). Diagnostic Markers for Vernalization and Photoperiod Loci Improve Genomic Selection for Grain Yield and Spectral Reflectance in Wheat. *Crop Science, 58*(1), 242-252. doi:10.2135/cropsci2017.06.0348

McKendry, A. L., Tague, D. N., Wright, R. L., & Tremain, J. A. (2007). Registration of 'Bess' Wheat. *Journal of Plant Registrations, 1*(1), 21-23. doi:10.3198/jpr2006.06.0405crc

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.

Miedaner, T., Wurschum, T., Maurer, H. P., Korzun, V., Ebmeyer, E., & Reif, J. C. (2011). Association mapping for Fusarium head blight resistance in European soft winter wheat. *Molecular Breeding, 28*(4), 647-655. doi:10.1007/s11032-010-9516-z

Mirdita, V., He, S., Zhao, Y. S., Korzun, V., Bothe, R., Ebmeyer, E., . . . Jiang, Y. (2015). Potential and limits of whole genome prediction of resistance to Fusarium head blight and Septoria tritici blotch in a vast Central European elite winter wheat population. *Theoretical and Applied Genetics, 128*(12), 2471-2481. doi:10.1007/s00122-015-2602-1

Money, D., Gardner, K., Migicovsky, Z., Schwaninger, H., Zhong, G. Y., & Myles, S. (2015). LinkImpute: Fast and Accurate Genotype Imputation for Nonmodel Organisms. *G3-Genes Genomes Genetics, 5*(11), 2383-2390. doi:10.1534/g3.115.021667

Moreno-Amores, J., Michel, S., Miedaner, T., Longin, C. F. H., & Buerstmayr, H. (2020). Genomic predictions for Fusarium head blight resistance in a diverse durum wheat panel: an effective incorporation of plant height and heading date as covariates. *Euphytica, 216*(2), 19. doi:10.1007/s10681-019-2551-x

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z. W., Costich, D. E., & Buckler, E. S. (2009). Association Mapping: Critical Considerations Shift from Genotyping to Experimental Design. *Plant Cell, 21*(8), 2194-2202. doi:10.1105/tpc.109.068437

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.

Owens, B. F., Lipka, A. E., Magallanes-Lundback, M., Tiede, T., Diepenbrock, C. H., Kandianis, C. B., . . . Rocheford, T. (2014). A Foundation for Provitamin A Biofortification of

Maize: Genome-Wide Association and Genomic Prediction Models of Carotenoid Levels. *Genetics, 198*(4), 1699-+. doi:10.1534/genetics.114.169979

Paillard, S., Schnurbusch, T., Tiwari, R., Messmer, M., Winzeler, M., Keller, B., & Schachermayr, G. (2004). QTL analysis of resistance to Fusarium head blight in Swiss winter wheat (Triticum aestivum L.). *Theoretical and Applied Genetics, 109*(2), 323-332. doi:10.1007/s00122-004-1628-6

Petersen, S., Lyerly, J. H., Maloney, P. V., Brown-Guedira, G., Cowger, C., Costa, J. M., . . . Murphy, J. P. (2016). Mapping of Fusarium Head Blight Resistance Quantitative Trait Loci in Winter Wheat Cultivar NC-Neuse. *Crop Science, 56*(4), 1473-1483. doi:10.2135/cropsci2015.05.0312

Petersen, S., Lyerly, J. H., McKendry, A. L., Islam, M. S., Brown-Guedira, G., Cowger, C., . . . Murphy, J. P. (2017). Validation of Fusarium head blight resistance QTL in US winter wheat. *Crop Science, 57*(1), 1-12. doi:10.2135/cropsci2015.07.0415

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), 8. doi:10.1371/journal.pone.0032253

Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology, 5*(2), 94-100. doi:10.1016/s1369-5266(02)00240-6

Rawat, N., Pumphrey, M. O., Liu, S. X., Zhang, X. F., Tiwari, V. K., Ando, K., . . . Gill, B. S. (2016). Wheat Fhb1 encodes a chimeric lectin with agglutinin domains and a pore-forming toxin-like domain conferring resistance to Fusarium head blight. *Nature Genetics, 48*(12), 1576-1580. doi:10.1038/ng.3706

Rice, B., & Lipka, A. E. (2019). Evaluation of RR-BLUP Genomic Selection Models that Incorporate Peak Genome-Wide Association Study Signals in Maize and Sorghum. *Plant Genome, 12*(1), 14. doi:10.3835/plantgenome2018.07.0052

Rutkoski, J., Benson, J., Jia, Y., Brown-Guedira, G., Jannink, J. L., & Sorrells, M. (2012). Evaluation of Genomic Prediction Methods for Fusarium Head Blight Resistance in Wheat. *Plant Genome, 5*(2), 51-61. doi:10.3835/plantgenome2012.02.0001

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. doi:10.3835/plantgenome2014.10.0074

Rutkoski, J. E., Poland, J. A., Singh, R. P., Huerta-Espino, J., Bhavani, S., Barbier, H., . . . Sorrells, M. E. (2014). Genomic Selection for Quantitative Adult Plant Stem Rust Resistance in Wheat. *Plant Genome, 7*(3), 10. doi:10.3835/plantgenome2014.02.0006

Sarinelli, J. M., Murphy, J. P., Tyagi, P., Holland, J. B., Johnson, J. W., Mergoum, M., . . . 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. doi:10.1007/s00122-019-03276-6

Schmolke, M., Zimmermann, G., Buerstmayr, H., Schweizer, G., Miedaner, T., Korzun, V., . . . Hartl, L. (2005). Molecular mapping of Fusarium head blight resistance in the winter wheat population Dream/Lynx. *Theoretical and Applied Genetics, 111*(4), 747-756. doi:10.1007/s00122-005-2060-2

Schulthess, A. W., Zhao, Y. S., Longin, C. F. H., & Reif, J. C. (2018). Advantages and limitations of multiple-trait genomic prediction for Fusarium head blight severity in hybrid wheat (Triticum aestivum L.). *Theoretical and Applied Genetics, 131*(3), 685-701. doi:10.1007/s00122-017-3029-7

Sobrova, P., Adam, V., Vasatkova, A., Beklova, M., Zeman, L., & Kizek, R. (2010). Deoxynivalenol and its toxicity. *Interdisciplinary Toxicology, 3*(3), 94-99. doi:10.2478/v10102-010-0019-x

Spindel, J. E., Begum, H., Akdemir, D., Collard, B., Redona, E., Jannink, J. L., & McCouch, S. (2016). Genome-wide prediction models that incorporate de novo GWAS are a powerful new tool for tropical rice improvement. *Heredity, 116*(4), 395-408. doi:10.1038/hdy.2015.113

Steiner, B., Buerstmayr, M., Michel, S., Schweiger, W., Lemmens, M., & Buerstmayr, H. (2017). Breeding strategies and advances in line selection for Fusarium head blight resistance in wheat. *Tropical Plant Pathology, 42*(3), 165-174. doi:10.1007/s40858-017-0127-7

Steiner, B., Michel, S., Maccaferri, M., Lemmens, M., Tuberosa, R., & Buerstmayr, H. (2019). Exploring and exploiting the genetic variation of Fusarium head blight resistance for genomic-assisted breeding in the elite durum wheat gene pool. *Theoretical and Applied Genetics, 132*(4), 969-988. doi:10.1007/s00122-018-3253-9

Su, Z. Q., Jin, S. J., Zhang, D. D., & Bai, G. H. (2018). Development and validation of diagnostic markers for Fhb1 region, a major QTL for Fusarium head blight resistance in wheat. *Theoretical and Applied Genetics, 131*(11), 2371-2380. doi:10.1007/s00122-018-3159-6

Tang, Y., Liu, X. L., Wang, J. B., Li, M., Wang, Q. S., Tian, F., . . . Zhang, Z. W. (2016). GAPIT Version 2: An Enhanced Integrated Tool for Genomic Association and Prediction. *Plant Genome, 9*(2), 9. doi:10.3835/plantgenome2015.11.0120

Tessmann, E. W., Dong, Y. H., & Van Sanford, D. A. (2019). GWAS for Fusarium Head Blight Traits in a Soft Red Winter Wheat Mapping Panel. *Crop Science, 59*(5), 1823-1837. doi:10.2135/cropsci2018.08.0492

VanRaden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science, 91*(11), 4414-4423. doi:10.3168/jds.2007-0980

Wang, H., Sun, S., Ge, W., Zhao, L., Hou, B., Wang, K., . . . Guo, J. (2020). Horizontal gene transfer of Fhb7 from fungus underlies Fusarium head blight resistance in wheat. *Science*. doi:10.1126/science.aba5435

Wilson, W. W., McKee, G., Nganje, W., Dahl, B., & Bangsund, D. (2017). Economic Impact of USWBSI's Scab Initiative to Reduce FHB. In (Vol. 774). Agribusiness and Applied Economics.

Yang, Z. P., Gilbert, J., Fedak, G., & Somers, D. J. (2005). Genetic characterization of QTL associated with resistance to Fusarium head blight in a doubled-haploid spring wheat population. *Genome, 48*(2), 187-196. doi:10.1139/g04-104

Yu, J. B., Bai, G. H., Zhou, W. C., Dong, Y. H., & Kolb, F. L. (2008). Quantitative trait loci for Fusarium head blight resistance in a recombinant inbred population of Wangshuibai/Wheaton. *Phytopathology, 98*(1), 87-94. doi:10.1094/phyto-98-1-0087

Zhang, Z., Ober, U., Erbe, M., Zhang, H., Gao, N., He, J. L., . . . Simianer, H. (2014). Improving the Accuracy of Whole Genome Prediction for Complex Traits Using the Results of Genome Wide Association Studies. *Plos One, 9*(3), 12. doi:10.1371/journal.pone.0093017

Table 1. Descriptive statistics and analysis of variance of four Fusarium head blight (FHB) resistance traits, including deoxynivalenol (DON) accumulation, Fusarium damaged kernels (FDK), incidence (INC), and severity (SEV), as well as heading date (HD) and plant height (PH) for 354 soft red winter wheat genotypes.

Trait	Mean	Minimum	Maximum	Standard Deviation	H^{2a}	Mean Squares		
						Genotype	Environment	GxE[b]
DON[c]	10.1	1.7	27.6	11.4	.79	13.05***	8.81***	3.90***
FDK	31.5	3.2	83.0	29.7	.82	15.93***	11.37***	4.18***
INC	28.9	.8	80.0	35.4	.38	4.26***	106.21***	3.19***
SEV	24.8	0.0	70.0	24.8	.78	8.38***	26.76***	2.90***
HD[d]	105.6	78.7	129.0	13.7	.92	27,194.00***	9,546.40***	3.36***
PH[e]	90.3	56.5	121.9	10.0	.89	27.24***	1,128.85***	4.57***

[a] Broad sense heritability (H^2) values calculated on an entry-mean basis for each of the four FHB traits.

[b] The interaction between genotype and environment.

[c] DON was measured in μg g^{-1}, whereas the other three traits were measured in percentage.

[d] Heading date (HD) was recorded in Julian days after 1 January, when 50% of the heads were 50% emerged from the flag leaf.

[e] Plant height (PH) was recorded in inches from the surface of the soil to the tip of the awn, but reported in centimeters here.

*** Significant at the .001 probability level.

Table 2. Correlation coefficients between heading date (HD), plant height (PH), and four Fusarium head blight (FHB) traits: deoxynivalenol (DON) accumulation, percent Fusarium damaged kernels (FDK), percent incidence (INC), and percent severity (SEV). Genetic correlations are in the top half, while phenotypic correlations are in the bottom half of the table.

Trait	INC	SEV	FDK	DON	HD	PH
INC	-	.80***	.73***	.25***	-.11*	-.24***
SEV	.65***	-	.80***	.41***	-.26***	-.35***
FDK	.48***	.66***	-	.42***	.08	-.30***
DON	.19***	.32***	.40***	-	.28***	.00
HD	-.06	-.23***	-.02	.28***	-	.28***
PH	-.27***	-.41***	-.28***	.00	.35***	-
* Significant at the 0.05 probability level.						
*** Significant at the 0.001 probability level.						

Table 3. List of significant quantitative trait loci (QTL) associated with four Fusarium head blight (FHB) traits, including deoxynivalenol (DON) accumulation, percent Fusarium damaged kernels (FDK), percent incidence (INC), and percent severity (SEV) across nine location-years using a fixed and random model circulating probability unification (FarmCPU) model with a false discovery rate (FDR) threshold of $q \leq .10$.

Trait[a]	Chr.[b]	Position	Allele	-log$_{10}$(p)	MAF[c]	Allelic Effect[d]
		bp				
DON	2B	574,225,916	A/G	8.80	0.240	-0.73
	3B	741,517,608	T/G	7.02	0.079	1.07
	3B	749,054,930	C/T	6.80	0.418	0.50
	3D	498,442,075	A/C	6.01	0.058	0.93
	3D	521,400,798	G/T	5.99	0.127	0.65
	4A	39,713,449	T/C	5.90	0.343	0.46
	5B	590,614,717	G/A	5.47	0.203	-0.53
	7B	230,436,017	C/T	5.34	0.199	0.53
	7D	457,776,058	A/T	5.13	0.059	0.87
	7A	50,110,754	T/C	4.99	0.404	-0.39
FDK	4B	577,008,759	C/G	14.35	0.417	4.16
	3A	714,306,573	A/G	11.65	0.225	-3.17
	1B	298,191,278	G/T	9.02	0.081	-5.27
	5B	673,624,297	G/A	8.31	0.051	-5.98
	5D	548,345,563	C/T	8.09	0.395	-2.19
	5D	424,046,598	C/A	6.08	0.379	-1.77
	2D	61,772,585	C/G	5.80	0.097	-2.87
	3B	378,832,973	C/T	5.70	0.112	-3.08
	3B	212,432,145	A/T	5.49	0.093	-2.93
	6A	11,159,506	G/T	5.18	0.106	2.64
	3B	51,321,489	T/A	4.87	0.202	-2.13
INC	2B	35,068,159	G/A	10.02	0.116	0.64
	2B	189,749,199	A/G	9.07	0.175	-1.22
	7B	716,844,038	C/A	8.19	0.064	-0.80
	5D	133,059,272	C/T	8.07	0.184	-0.45
	6A	5,113,561	T/A	6.79	0.203	-0.44
	5D	74,734,312	T/G	6.48	0.081	-0.76
	4B	21,368,968	A/G	6.00	0.264	-0.34
	7A	709,861,127	C/A	5.24	0.182	0.34
	1A	263,596,817	T/A	5.19	0.121	-0.43
SEV	3B	783,125,543	T/A	11.68	0.381	-1.83
	4A	562,577,189	T/C	7.87	0.366	-1.38
	4B	577,008,759	C/G	7.74	0.417	1.55
	5A	365,630,402	A/C	6.69	0.370	1.40
	2B	653,352,399	A/G	6.49	0.059	2.19
	7A	325,554,142	G/A	6.36	0.066	-2.38

Table 3 (Cont.)

Trait[a]	Chr.[b]	Position	Allele	-log$_{10}$(p)	MAF[c]	Allelic Effect[d]
SEV	2D	11,171,031	T/C	6.02	0.222	-1.62
	5B	548,333,940	G/A	5.94	0.338	-1.17
	3A	347,926,467	C/G	5.85	0.123	-1.85
	7B	103,622,350	C/A	5.56	0.110	2.38
	6B	18,241,620	G/A	5.42	0.052	-3.02
	1B	77,853,663	G/A	5.13	0.048	-2.90

[a] DON, deoxynivalenol accumulation; FDK, Fusarium damaged kernels; INC, incidence; SEV, severity.

[b] Chr, Wheat (*Triticum aestivum* L.) chromosome number.

[c] MAF, minor allele frequency.

[d] Allelic effect was calculated in FarmCPU and reported as the difference between the mean best linear unbiased prediction (BLUP) between genotypes with the major and minor allele for each FHB resistance. A positive sign means that the minor allele causes a percent increase in the trait of interest, a negative sign means that the minor allele causes a percent decrease in the trait of interest.

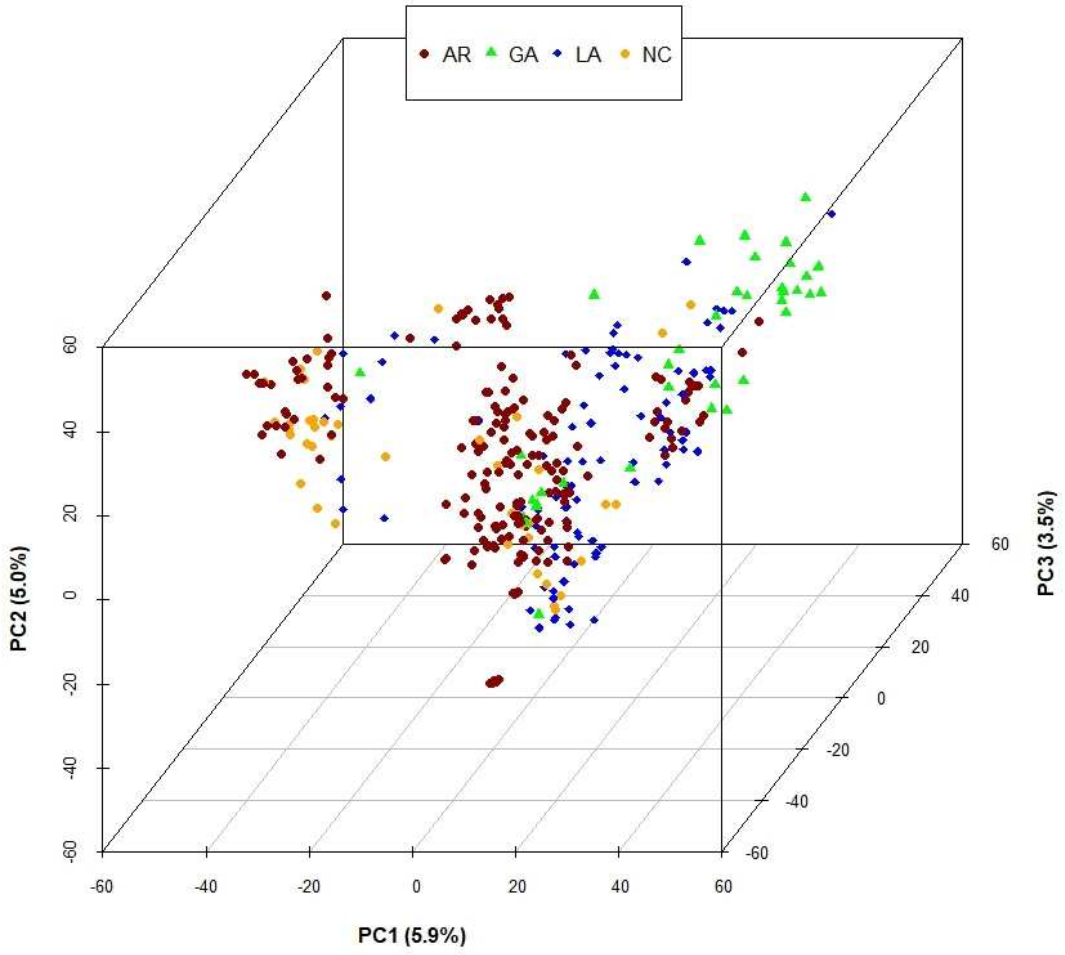


Figure 1. Population structure of 354 soft red winter wheat genotypes using 72,634 single nucleotide polymorphisms (SNPs). Colors represent the origin of the genotypes. AR, developed at the University of Arkansas, Fayetteville, AR; GA, developed at the University of Georgia, Athens, GA; LA, developed at Louisiana State University, Baton Rouge, LA; NC, developed at North Carolina State University, Raleigh, NC; PC, principal component.

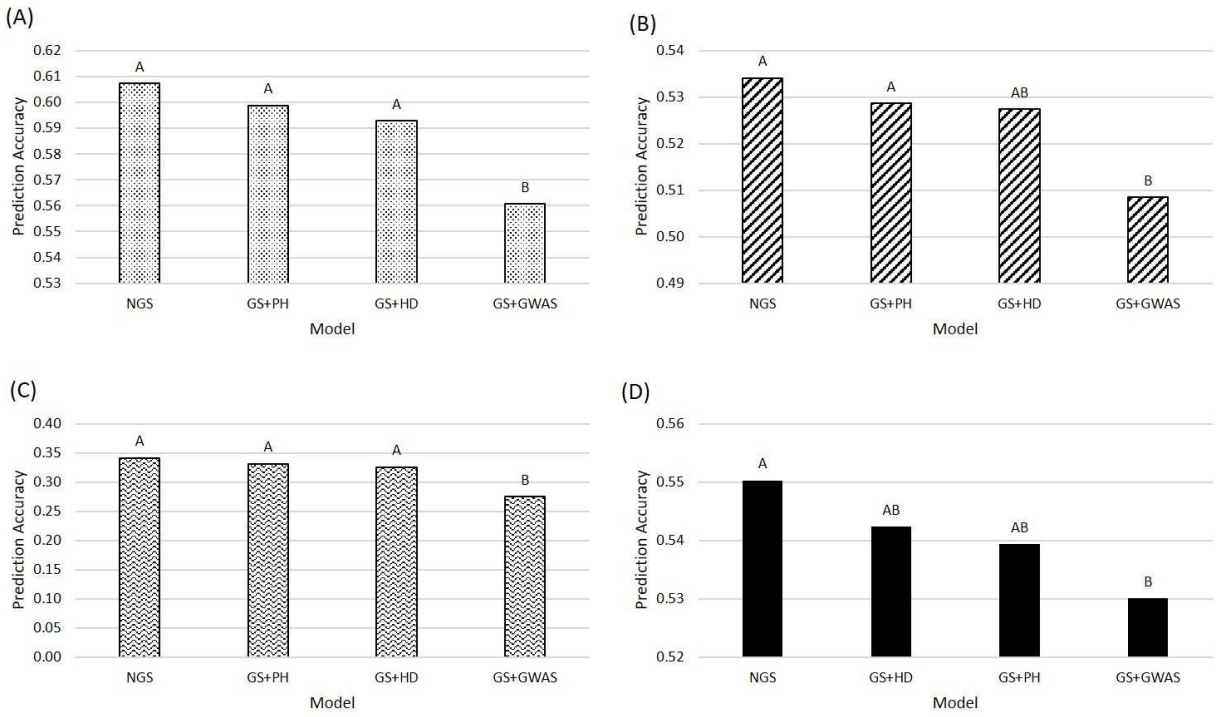


Figure 2. Bar charts comparing the mean genomic prediction accuracies (*y* axes) between four genomic selection (GS) models used to predict four Fusarium head blight (FHB) resistance traits: (a) deoxynivalenol accumulation (DON), (b) Fusarium damaged kernels (FDK), (c) incidence (INC), and (d) severity (SEV). The *x* axis represents the four models being compared: a naïve genomic best linear unbiased prediction (GBLUP) GS model with no fixed effects (NGS), and three *denovo* genome-wide association study (GWAS) assisted GBLUP models with significant marker fixed effects for heading date (GS+HD), plant height (GS+PH), and the respective FHB traits (GS+GWAS). The *y* axis represents the mean prediction accuracy across 10 iterations of cross-validation in the form of a Pearson correlation coefficient (*r*) between the predicted genome estimated breeding value (GEBV) and the actual phenotypic value for the validation populations. Mean prediction accuracies of each model not sharing any letter above each bar are significantly different based on Fisher's LSD separation at the 5% level of significance.

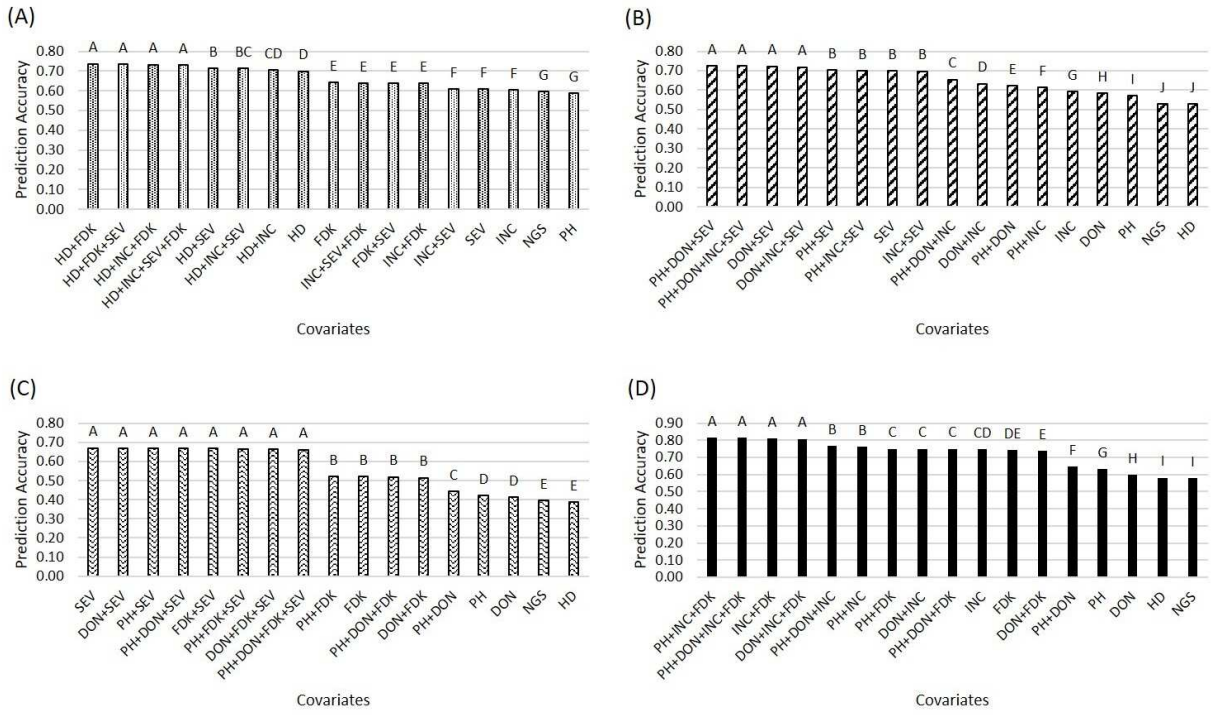


Figure 3. Bar charts comparing the mean prediction accuracies between 16 multivariate genomic selection (MVGS) models with a naïve genomic selection (NGS) model for four different Fusarium head blight (FHB) resistance traits: (a) deoxynivalenol accumulation (DON), (b) Fusarium damaged kernels (FDK), (c) incidence (INC), and (d) severity (SEV). The *x* axis represents the combination of covariates used for each model, including the abovementioned FHB traits, as well as heading date (HD) and plant height (PH). The *y* axis represents the mean prediction accuracy across 100 iterations of five-fold cross-validation in the form of a Pearson correlation coefficient (*r*) between the predicted genome estimated breeding value (GEBV) and the actual phenotypic value for the validation populations. Mean prediction accuracies of each model not sharing any letter above each bar are significantly different based on Fisher's LSD separation at the 5% level of significance.

CHAPTER III

IMPLEMENTATION OF MULTIVARIATE GENOMIC SELECTION METHODS FOR AGRONOMIC TRAITS IN SOFT RED WINTER WHEAT

ABSTRACT

Genomic selection (GS) is an important tool for increasing genetic gain and multivariate GS (MVGS) can improve prediction accuracy for a trait using correlated secondary traits as covariates. Likewise, data for breeding lines from other environments can be used as covariates to predict performance in untested environments using sparse-testing (CV2). The objectives for this study were to evaluate the abovementioned GS approaches to improve prediction accuracy for five agronomic traits in a population of 351 soft red winter wheat genotypes, evaluated over six site-years in Arkansas from 2014 to 2017. Genotypes were phenotyped for grain yield (GY), heading date, maturity date (MD), plant height, and test weight. The MVGS models significantly improved prediction accuracy for all five traits compared to a naïve GS model without covariates by between 10 and 30%, where the lowest heritability trait, MD (H^2 = .67), saw the greatest increase in prediction accuracy. Genetic correlations exceeding r = .30, between predicted traits and covariates, also resulted in significantly higher prediction accuracies ($p < .05$). The CV2 method significantly improved prediction accuracy ($p < .05$) for GY between 6% and 21% in four of six tested environments, especially when covariate environments and predicted environments were highly correlated and closely related. Overall, MVGS models improved prediction accuracies when there was a strong genetic correlation between the predicted trait and covariates. The CV2 method was also effective for predicting traits for genotypes in new environments, when covariate environments were correlated with the predicted environment of interest.

INTRODUCTION

The global population is expected to increase from 7.7 to approximately 9.7 billion by 2050, according to the United Nations Department of Economic and Global Affairs (United Nations, 2019). While the global capacity for food production has rapidly expanded over the last 70 years, largely due to improved agronomic practices and genetic improvements in crop species, this rapid population growth is still expected to surpass the rate of global food production in the coming decades (Graybosch & Peterson, 2010). This trend can also be observed in wheat (*Triticum aestivum* L.), which contributes to over 20% of caloric and protein intake for humans. The current genetic gain in grain yield (GY) is .9%, far below the 2.4% needed to double wheat production by 2050 (Ray, Mueller, West, & Foley, 2013).

To meet these future demands, wheat breeders have looked toward new technologies to improve the efficiency of their breeding programs by implementing new and improved technologies to increase genetic gain. These technologies include methods such as marker-assisted selection (MAS), phenomics, and genomic selection (GS) (Heffner, Sorrells, & Jannink, 2009; Larkin, Lozada, & Mason, 2019). Genomic selection is a modified form of MAS; however instead of using only a few select markers that are significantly associated with a trait, GS takes all summed marker and locus effects across the entire genome into consideration to calculate genome estimated breeding values (GEBV) (Heffner et al., 2009; Meuwissen, Hayes, & Goddard, 2001). At least some of the markers are assumed to be in linkage disequilibrium (LD) with quantitative trait loci (QTL) associated with a trait of interest (Meuwissen et al., 2001). One of the primary advantages of GS is more genetic gain per cycle of GS 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 improve the predictive ability of GS models (Jannink, Lorenz, & Iwata, 2010; Larkin et al., 2019). These methods largely focus on training population (TP) optimization, based on size, population structure, composition, and genetic relationships between the TP and validation population (VP) (Akdemir, Sanchez, & Jannink, 2015; Combs & Bernardo, 2013; Habier, Fernando, & Dekkers, 2007; Isidro et al., 2015; Jannink et al., 2010). The type of prediction model also plays a valuable role in improving prediction accuracy, whether it is through the use of ridge-regression or Bayesian parametric models, or through the use of semi- or non-parametric models (N. Heslot, Yang, Sorrells, & Jannink, 2012). An alternative to this is multivariate GS (MVGS), which uses mixed models with secondary traits, genetically correlated with a trait of interest, as covariates (Calus & Veerkamp, 2011; Covarrubias-Pazaran et al., 2018; Jia & Jannink, 2012). 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). In studies using MVGS models focused on agronomic traits of interest in wheat, GY has been the primary trait of interest, while high-throughput phenotyping traits, such as normalized difference vegetative index (NDVI) and canopy temperature, were used as secondary traits (Crain, Mondal, Rutkoski, Singh, & Poland, 2018; Lozada & Carter, 2019; Rutkoski et al., 2016; Sun et al., 2017). However, wheat breeders collect data on multiple important agronomic traits of interest throughout the growing season including heading date (HD), maturity date (MD), plant height (PH), and test weight (TW) that can impact GY (X. J. Chen, Min, Yasir, & Hu, 2012; Liu, Searle, Mather, Able, & Able, 2015; Tshikunde, Mashilo, Shimelis, & Odindo, 2019). Therefore, these traits, along with GY can reasonably be incorporated as covariates in MVGS models to improve prediction accuracies for each of these important agronomic traits of interest.

Another application for MVGS can be for predicting the performance of genotypes in environments (ENVs) where they have not been tested while data from other ENVs, where the genotype was tested, are used as covariates in the MVGS model. Plant breeders often grow genotypes in multiple-environment trials (METs) in order to evaluate their performance under diverse environmental conditions at multiple locations over multiple years (Basford & Cooper, 1998). The estimation of the genotype-by-environment interaction (GEI) is valuable so that the breeder can determine that a genotype's phenotypic performance is stable in multiple ENVs over time (Allard & Bradshaw, 1964; de Leon, Jannink, Edwards, & Kaeppler, 2016; Eberhart & Russell, 1966; Lin, Binns, & Lefkovitch, 1986; Rosielle & Hamblin, 1981). Ideally, each genotype would be evaluated in every testing ENV; however, due to limitations of field space, seed availability, time, or cost this is not always possible. These limitations are especially evident during early generations of the breeding cycle (Jarquin et al., 2020). In terms of GS, this scenario fits into the cross-validation 2 (CV2) sparse testing cross-validation design, where the performance of a genotype is predicted in a specific ENV using phenotypic data of the same genotype from a different ENV in the TP (Burgueno, de los Campos, Weigel, & Crossa, 2012).

The CV2 scheme was compared with a cross-validation scheme where the genotypes were untested in any ENVs (CV1) while using pedigree, markers, and pedigree + marker data in a wheat population of 599 genotypes in order to predict GY. The CV2 scheme had higher prediction accuracies across all three model types (Burgueno et al., 2012). This was also observed in a diverse wheat panel grown across five ENVs when comparing the CV1 and CV2 schemes using seven different GS models, where the CV2 scheme had higher prediction accuracies than the CV1 scheme across all ENVs and models (Saint Pierre et al., 2016). The CV2 scheme was again the highest performing scheme when using a panel of 320 chickpea

(*Cicer arietinum* L.) genotypes evaluated for eight traits across six ENVs using 13 different GS models and two different genotyping platforms (Roorkiwal et al., 2018).

While the abovementioned sparse cross-validation studies used phenotypic data from correlated ENVs in the TP, we wanted to take a MVGS approach to the CV2 scheme and use phenotypic data for missing genotypes grown in ENVs correlated with an ENV of interest as a covariate, while only using data from the ENV of interest in the TP, in order to see if the CV2 scheme could improve prediction accuracy compared to a CV1 scheme. This study accomplishes three main objectives: (a) to evaluate soft red winter wheat (SRWW) genotypes for five agronomically important traits, including GY, HD, MD, PH, and TW; (b) to perform a cross-validation analysis to compare prediction accuracies between a naïve GS (NGS) model and MVGS models for all five agronomic traits; and (c) to perform a CV2 analysis in order to predict GY for genotypes that were grown in some ENVs but not others using GY data from correlated ENVs as model covariates in a MVGS approach.

MATERIALS AND METHODS

Germplasm

A population of 351 SRWW genotypes was used for this study, consisting of 237 genotypes from the University of Arkansas, 39 from the University of Georgia, 38 from North Carolina State University, and 37 from Louisiana State University. This population represents the majority of source germplasm in the University of Arkansas wheat breeding program.

Experimental Design and Trait Measurements

Winter wheat is planted in the fall and harvested during the late spring in the southern region of the U.S., therefore each growing season spans two years. The first 175 entries of the population

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 (KSR) during the 2013-2014 growing season (M14 and K14), while the remaining 176 entries were evaluated at MAR during the 2014-2015 growing season (M15). All 351 entries of the population were only evaluated at MAR during the 2015-2016 growing season (M16) and at both locations during the 2016-2017 growing season (M17 and K17) (Table 1).

Genotypes 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 between rows. The plots were then end-trimmed in the spring. Plots were managed according to the recommendations for wheat in Arkansas (Kelley, 2018). Each year, plots in MAR received two applications of urea fertilizer of 101 kg ha^{-1} and 67 kg ha^{-1}, along with 27 kg ha^{-1} of ammonium sulfate fertilizer. Plots in KSR 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), Harmony Extra (DuPont de Nemours, Inc.), Osprey (Bayer AG), and Prowl H_2O (BASF SE), were used each year for weed control.

The M14 and M15 ENVs were planted in an unbalanced randomized complete block design (RCBD) with two replications with repeated checks. In M14, the first replication consisted of 175 genotypes, while the second replication consisted of the first 117 genotypes from the first replication, while in M15, the first replication consisted of 177 genotypes, while the second replication consisted of the first 120 genotypes from the first replication. The K14, M16, M17, and M17 ENVs were planted in an unreplicated augmented design with repeated checks. During the 2013-2014 growing season, the repeated checks were ‘Branson’ (PI 639227), ‘Croplan Genetics 514W’, ‘Jamestown’ (PI 653751), ‘Pioneer 26R20’ (PI 658150), ‘Armor Ricochet’, ‘Shirley’ (PI 656753), ‘SS 8461’, and ‘UniSouth Genetics 3120’ (PI 672163). The

2014-2015 growing season used the same eight checks, with the addition of 'SY Harrison' (PI 664946), 'Pioneer 26R41', and 'Pat' (PI 631466). Only four checks, Branson, Jamestown, Pioneer 26R41, and SS 8641, were used during the 2015-2016 and 2016-2017 growing seasons.

Phenotypic data were collected for five agronomic traits, including GY, TW, HD, MD, and PH. Individual plots were harvested with a 1994 Hege 140 plot combine (Hege Maschinen GmbH) and GY was calculated at all six ENVs from the weight of seed from each plot as measured by the HarvestMaster Pro 4100 (Juniper Systems) and was expressed in t ha^{-1} and adjusted for 13% moisture content. Test weight was sampled from each plot for all six ENVs using a HarvestMaster Pro 4100 (Juniper Systems), and was expressed in kg hl^{-1} then adjusted for 13% moisture content. Heading dates for each genotype were collected during all four growing seasons from MAR. Heading date was recorded in Julian days after 1 January, when 50% of the heads were 50% emerged from the flag leaf. As there was variation in HD between genotypes in the population, heading notes were recorded every other day from the onset of heading and continuing until all plots in the nursery were headed. Maturity date was evaluated for all genotypes at M14, K14, and M15. Maturity date was reported as the Julian days after 1 January, when an estimated 90% of culms in an individual plot were senesced or yellow. At maturity, PH was recorded in cm from the surface of the soil to the tip of the head at all ENVs, except for K17.

Phenotypic Data Analysis

Data was analyzed using procedures in SAS 9.4 software (SAS Institute), using a two-stage MET analysis (Mohring & Piepho, 2009; Piepho, Mohring, Melchinger, & Buchse, 2008). Adjusted means were obtained for all five traits from each ENV for the first stage of the analysis using the following model in PROC MIXED:

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

where y_{ij} is the observed phenotype; μ is the overall mean; τ_i is the fixed effect of the i^{th} genotype; β_j is the random effect of replication for the RCBD or incomplete block for the augmented design; and ε_{ij} is the residual error term, where $\varepsilon_{ij} \sim N(0, \boldsymbol{I}\sigma^2_\varepsilon)$, where $\boldsymbol{I}$ is an identity matrix and σ^2_ε is the residual error variance.

Adjusted means from each ENV for each trait were then compiled into a single dataset for the second stage of the analysis, where best linear unbiased predictions (BLUPs) were calculated for each trait using the following model in PROC MIXED with all effects and interaction treated as random:

$$y_{ij} = \mu + \tau_i + \delta_j + (\tau \times \delta)_{ij} + \varepsilon_{ij}$$

where y_{ij} is the observed phenotype; μ is the overall mean; τ_i is the random effect of the i^{th} genotype; δ_j is the random effect of ENV; $(\tau \times \delta)_{ij}$ is the interaction between genotype and ENV; and ε_{ij} is the residual error term, where $\varepsilon_{ij} \sim N(0, \boldsymbol{I}\sigma^2_\varepsilon)$, where $\boldsymbol{I}$ is an identity matrix and σ^2_ε is the residual error variance.

Phenotypic correlations were determined between all five traits using the multivariate function in JMP Pro 15.1.0 software (SAS Institute). Due to the unbalanced nature of the design for each of the five traits, broad-sense heritability $\bar{H}^2_C$ for each of the five traits was calculated using the following formula proposed by Cullis et al. (2006), which is ideal for unbalanced METs:

$$\bar{H}^2_C = 1 - \frac{\bar{v}_{BLUP}}{2\sigma^2_G}$$

where $\bar{v}_{BLUP}$ is the mean variance between two BLUPs and σ_G^2 is the genotypic variance of the trait of interest, obtained from the second stage of the MET model (Cullis, Smith, & Coombes, 2006; Piepho & Mohring, 2007).

Genotyping by Sequencing

All 351 genotypes in the population were genotyped using genotyping by sequencing (GBS). DNA was extracted using the Mag-Bind Plant DNA Plus kit (Omega Bio-tek), following the manufacturer's instructions. Genomic DNA was quantified using the Quant-iT PicoGreen dsDNA Assay Kit and normalized to 20 ng μL^{-1} (ThermoFisher Scientific). Genotyping by sequencing libraries were created using *Pst1-Msp1* and/or the *Pst1-Mse1* restriction enzyme combinations (Poland, Brown, Sorrells, & Jannink, 2012). The samples were pooled together at 192-plex and each pooled library was sequenced on a single lane of an Illumina Hi-Seq 2500 system (Illumina). Single nucleotide polymorphism calling was performed using the TASSEL 5.0 GBSv2 pipeline using 64 base kmer length and a minimum kmer count of five (Bradbury et al., 2007). Reads were aligned to the International Wheat Genome Sequencing Consortium (IWGSC) RefSeq v1.0 'Chinese Spring' wheat reference sequence using the alignment method of Burrows-Wheeler aligner version 0.7.10 (Appels et al., 2018; Li & Durbin, 2009). Raw single nucleotide polymorphism (SNP) data generated from the TASSEL pipeline were filtered to remove taxa with more than 85% missing data and heterozygosity greater than 30%. Genotypic data were then filtered to select for biallelic SNPs with minor allelic frequency of greater than five percent, less than 50% missing data, and heterozygosity of less than or equal to 10%. Missing data were imputed using the linkage disequilibrium - k^{th} nearest neighbor imputation (LD-*k*NNi) function in TASSEL 5.0 in order to create a final SNP panel consisting of 73,618 SNPs across the entire genome (Bradbury et al., 2007; Money et al., 2015). A principal

component analysis (PCA) was performed using TASSEL 5.0 using the PCA function and abovementioned genotypic dataset in order to identify the population structure within the population (Bradbury et al., 2007).

Multivariate Genomic Selection

Genotypic correlations between the five traits were obtained using a genetic analysis algorithm incorporating genome-wide SNPs in a mixed linear model framework within the sommer package in R v3.6.3 software (Covarrubias-Pazaran, 2016; Lee & van der Werf, 2016; R, 2020).

In order to determine if a MVGS genomic BLUP (GBLUP) model significantly improved prediction accuracy compared to a NGS model, a cross-validation analysis was performed using the rrBLUP package within the Intelligent Prediction and Association Tool (iPat), using source code from R version 3.6.3 for all five agronomic traits (C. P. J. Chen & Zhang, 2018; Endelman, 2011; R, 2020). The GBLUP model used for the analysis is described as follows:

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

where u is the vector of marker effects, which is assumed to have a normal distribution $u \sim N(0, \boldsymbol{G}\sigma_u^2)$, where $\boldsymbol{G}$ is the genomic relationship matrix and σ_u^2 is the variance of the individual marker effects; β is the vector of fixed effects; $\boldsymbol{X}$ is the design matrix of fixed effects; $\boldsymbol{Z}$ is the design matrix relating genotypes to phenotypic observations (y), with m markers in columns and n phenotypes in rows; and ε_i is the residual error at the i^{th} locus, which is assumed to have a normal distribution $\varepsilon_i \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$, where $\boldsymbol{I}$ is the identity matrix and σ^2_ε is the residual error variance. The GEBV is the sum of all allele effects of a genotype (C. P. J. Chen & Zhang, 2018; Endelman, 2011; VanRaden, 2008).

Phenotypic data from each of the other four traits were used as covariates. Each covariate trait was first tested individually in a single-covariate model using a five-fold cross-validation approach, where the population was randomly divided into five groups of 71 genotypes. Four of the five groups were then used as the TP to predict the fifth group, serving as the validation population (VP), where the phenotype was set as missing, while the phenotypic data for the covariate trait were used as a fixed effect in the model. The GEBVs for the VP were compared to the actual phenotypic values using a Pearson correlation. The process was performed over 100 iterations for a total of four single-covariate trait combinations for each agronomic trait of interest (Larkin et al., 2020). The mean prediction accuracies from the five single-covariate, MVGS models were compared with a NGS model to determine if any of the covariate traits significantly improved prediction accuracies using a generalized linear mixed model (GLMM) and Fisher's LSD with an α of .05, implemented in PROC GLIMMIX in SAS 9.4 (SAS Institute).

Covariates that produced significantly higher prediction accuracies than the NGS models were then included as covariates for multiple-covariate models to predict each agronomic trait. If an individual covariate did not significantly increase the prediction accuracy of the agronomic trait of interest in the single-covariate MVGS model, it was not included in the multiple-covariate MVGS models. Significant covariates were tested in all possible combinations in multiple-covariate MVGS models using the cross-validation method described above. Mean prediction accuracies from the NGS, single-covariate, and multiple-covariate models were compared for each agronomic trait of interest using a GLMM and Fisher's LSD with an α of .05, implemented in PROC GLIMMIX in SAS 9.4 (SAS Institute). A total of nine models were tested for TW, while six were tested for HD, MD, and PH, and five were tested for GY. Figures

comparing mean prediction accuracies between models were created using the yarrr package in R version 4.0.2 (Phillips, 2017; R, 2020).

Sparse Testing Genomic Selection

A PCA was performed for mean GY between all six ENVs in order to identify the relationship between ENVs using the Principal Components tool in JMP Pro 15.1.0 software (SAS Institute). Pearson correlation coefficients were also obtained between the mean GY of each ENV using the Multivariate Methods tool in JMP Pro 15.1.0 software (SAS Institute).

The Cluster Variables procedure within the Principal Components tool was then used to identify the optimal clustering between ENVs by using a variable clustering algorithm in JMP Pro 15.1.0 software (SAS Institute). The algorithm started with all six ENVs in a single cluster, then they were split into separate clusters, based on their principal components, until new splits were no longer possible. At each iteration, the cluster with the second largest eigenvalue was chosen to be split into two new clusters. Member ENVs were assigned to new clusters based on the first two orthoblique rotated principal components of the cluster being split. After splitting, the other ENVs were examined and re-assigned to another cluster if it had a higher correlation with a different cluster. The clustering process was terminated once the second eigenvalue of all clusters was less than one (Anderberg, 2014; Sarle, 1990). After the analysis, two distinct clusters, each containing three ENVs, were identified. The first cluster, named Mega-environment 1 (ME1), consisted of M17, K17, and M15, while the second cluster, named Mega-environment 2 (ME2), consisted of K14, M14, and M16.

Best linear unbiased predictions for GY were obtained for each Mega-environment using the two-stage model described above using PROC MIXED in SAS 9.4 software (SAS Institute).

Pearson correlation coefficients were then obtained between the GY BLUPs of each mega-environment and the GY BLUPs calculated across all ENVs using the Multivariate Methods tool in JMP Pro 15.1.0 software (SAS Institute). Broad-sense heritability was also calculated for each mega-environment using the formula proposed by Cullis et al. (2006), as described above.

A five-fold NGS cross-validation analysis was performed for GY in each individual ENV using the rrBLUP package within iPat, using source code from R v3.6.3 (C. P. J. Chen & Zhang, 2018; Endelman, 2011; R, 2020). An NGS model and five different single-covariate sparse-testing GS (CV2) models were compared for each ENV. The first and second models used adjusted means for GY from the other two ENVs from the same mega-environment as covariates. The third model used BLUPs for GY from the associated mega-environment, minus the ENV being tested, as a covariate. The fourth model used BLUPs for GY from all ENVs, minus the ENV being tested, as a covariate. The fifth model used BLUPs for GY from the opposing mega-environment as a covariate; for example, if the ENV of interest is in ME1, the covariate will be the BLUPs from ME2. The TP and VP sampling methods used for the five-fold cross-validation analysis was the same as the methods for the MVGS models. The GEBVs for the VP were compared to the actual phenotypic values for GY from the same ENV using a Pearson correlation in order to obtain a prediction accuracy.

The mean prediction accuracies from the NGS model and five CV2 models were compared to determine if the STGS models significantly improved prediction accuracies compared to an NGS model for each ENV using a GLMM and Fisher's LSD with an α of .05, implemented in PROC GLIMMIX in SAS 9.4 (SAS Institute Inc.). Figures comparing mean prediction accuracies between models were created using the yarrr package in R v4.0.2 (Phillips, 2017; R, 2020).

RESULTS

Genotypic Data

The population used for this study was also used for Larkin et al. (2020). The primary differences were that fewer genotypes were included in the population and more SNP markers were used (Larkin et al., 2020). Genotyping by sequencing identified 73,618 bi-allelic SNP markers across the entire wheat genome after filtering. Between the three genomes, there was an uneven distribution, where the B genome had the largest number of SNP markers (38,086), followed by the A (27,382) and D (9,682) genomes, which was consistent with other studies of similar populations using GBS marker data (Arruda et al., 2016; Larkin et al., 2020). Out of the total genotypic dataset, 1,795 SNP markers were unaffiliated with any genome. The chromosome with the largest number of SNP markers was chromosome 2B (7,669), whereas chromosome 4D had the smallest number (665). The percentage heterozygosity for the dataset was 3.5%, whereas the average minor allele frequency was 19.6%. The allelic distributions were 25.8% for C, 25.2% for G, 23.0% for A, and 22.5% for T.

The PCA of the population identified two distinct clusters within the population. Genotypes between all four breeding programs appeared in both clusters, although there appeared to be some sub-clustering between individual breeding programs within the clusters. Such clustering has also been observed in other studies using SRWW populations adapted to the Southeastern US and it is hypothesized that this is due to the presence or absence of a stem rust (*Puccinia graminis* f. sp. *tritici*) and powdery mildew (*Blumeria graminis* f. sp. *tritici*) resistance genes, *Sr36/Pm6*, located on a translocation from *Triticum timopheevii* Zhuk. (Benson, Brown-Guedira, Murphy, & Sneller, 2012; Larkin et al., 2020; Nyquist, 1962; Sarinelli et al., 2019). The

population structure was generally low, where the first three principal components only accounted for 5.9, 5.0, and 3.5% of the total genetic variation (Figure 1).

Multivariate Genomic Selection

Phenotypic Data

All five agronomic traits had significant variation within the population. Grain yield had a range between .10 and 9.54 t ha^{-1} across all six ENVs. Data for HD was collected all four years from MAR for a total of four ENVs, where there was a range between 79 and 129 Julian days after 1 Jan. Data for MD was only collected during the 2014 (K14 and M14) and 2015 (M15) seasons for a total of three ENVs, where there was a range between 138 and 159 Julian days after 1 Jan. Data for PH was collected at all ENVs, excluding K17, where there was a range in PH between 56.46 and 121.74 cm. The only trait, outside of GY, that was collected in all six ENVs was TW, where there was a range between 45.57 and 85.20 kg hl^{-1} (Table 1).

Significant phenotypic and genetic correlations were observed between all agronomic traits except for the correlations between HD and TW. Traits that had strong phenotypic correlations also had strong genetic correlations (Table 1). The highest heritability among the five agronomic traits was for HD (.94), followed by PH (.91), GY (.85), TW (.84), and MD (.67) (Table 1). Evaluating the population between three and six ENVs, depending on the trait, allowed for high heritability values.

Model Comparisons

Multiple MVGS models were compared with an NGS model using a five-fold cross-validation approach over 100 cycles to predict five agronomic traits. If a trait significantly improved predication accuracy compared to a NGS in a single-covariate model, the trait was further tested

in multiple-covariate models. The MVGS models significantly improved prediction accuracies for all five traits (Table 2, Figure 2).

Maturity date saw the largest increase in prediction accuracy compared to the NGS model (r = .44) at 29.6%, where the single-covariate model using HD as a covariate, and the multiple-covariate models using HD and TW as covariates were the top performing models (r = .73). Only two single-covariate models, HD and TW, significantly improved prediction accuracy, therefore there was only one multiple-covariate model containing HD and TW as covariates. Maturity date was also one of two traits of interest that had a single-covariate model as the top performing model, however it was not significantly better than the multiple-covariate model (Table 2, Figure 2c).

The trait with the second largest increase in prediction accuracy with an MVGS model was TW, where there was a 25.0% increase between the NGS model (r = .36) and the multiple-trait MVGS model containing GY, MD, and PH as covariates (r = .61). All single-covariate models, except for HD, had significantly higher prediction accuracies compared to the NGS model, therefore it was not included as a covariate in the multiple-covariate models. The top performing GY+MD+PH model performed significantly better than all other MVGS models as well. Test weight was the only trait of interest that had a MVGS model with more than two covariates (Table 2, Figure 2e).

Heading date had the third largest increase in prediction accuracy where there was a 20.3% increase between the multiple-covariate MD+PH model (r = .80) and the NGS model (r = .60). The single-covariate models with GY and TW did not significantly improve prediction accuracy compared to the NGS model, therefore they were not included in the multiple-covariate

models. The only multiple-covariate model, MD+PH, was also significantly the top performing model (Table 2, Figure 2b).

Grain yield had the second lowest level of improvement in prediction accuracy between the single-covariate MVGS model, with TW as a covariate (r = .63), and the NGS model (r = .53) with a 10.2% increase. No single-covariate models significantly improved prediction accuracy for GY outside of the model with TW, therefore no multiple-covariate models were tested. Grain yield was the only trait of interest where a single-covariate model was significantly the top performing model (Table 2, Figure 2a).

The trait with the smallest level of improvement in prediction accuracy was PH, where the top performing model HD+TW (r = .80) had an 8.4% increase in prediction accuracy compared to the NGS model (r = .56). The single-covariate models with GY and MD did not significantly improve prediction accuracy compared to the NGS model, therefore they were not included as covariates in the multiple-covariate MVGS models. The only multiple-covariate model, HD+TW, was also significantly the top performing model (Table 2, Figure 2d).

Sparse Testing Genomic Selection

Environmental Clustering

A PCA was performed for mean GY between all six ENVs in order to identify the relationship between ENVs. The analysis found that the six ENVs could be divided into two mega-environments, each containing three ENVs (Appendix 3). The first mega-environment, ME1, consisted of K17, M15, and M17, while the second mega-environment, ME2, consisted of K14, M14, and M16. The first principal component in the analysis accounted for 38.5% of phenotypic variation in the population, while the second principal component accounted for 20.7% of

phenotypic variation. The first mega-environment accounted for 34.9% of the phenotypic variation, while ME2 accounted for 22.7% of the phenotypic variation. Overall, 57.7% of the phenotypic variation was explained by the clustering analysis. The most representative ENV for ME1 was M17, while the most representative ENV for ME2 was K14.

Phenotypic Data

Adjusted means of GY were obtained for all six individual ENVs to use as covariates for the CV2 GS models. These adjusted means were also used to obtain BLUPs for GY for three other types of covariate datasets. The first of these consisted of BLUPs calculated across five ENVs per covariate, including all ENVs except for the ENV of interest (All-*ENV*), for a total of six covariates. The second type of covariate dataset consisted of BLUPs for GY across the two other ENVs within the mega-environment from which the ENV of interest is associated (ME*n*-*ENV*), for a total of six covariates. The third type of covariate dataset contained BLUPs for GY obtained across all three ENVs for each mega-environment. These full mega-environment datasets were intended to be used as covariates for the opposite mega-environment. In total, 20 different datasets were created for each type of covariate (Table 3). The covariate dataset with the highest mean GY was the individual ENV of M14 (4.88 t ha^{-1}), while the covariate dataset with the lowest mean GY was the individual ENV of K14 (2.57 t ha^{-1}). Between the two mega-environments, ME2 had the highest mean GY (4.08 t ha^{-1}) compared to ME1 (3.68 t ha^{-1}). However, ME1 (H^2 = .88) had a higher broad-sense heritability than ME2 (H^2 = .66). The covariate dataset with the highest broad-sense heritability was the individual ENV of M17 (H^2 = .89) while the dataset with the lowest broad-sense heritability was the individual ENV of M14 (H^2 = .42) (Table 3).

Phenotypic correlations were also obtained between all six individual ENVs of interest and their respective covariate datasets (Table 4). Generally, individual ENVs from ME1 had stronger correlations with the other ENVs from ME1 as well as the covariate datasets containing all ENVs except for themselves. The three individual ENVs within ME1 even had significant correlations with the ME2 covariate dataset. The correlations between the individual ENVs in ME2 were not as strong as those from ME1. Keiser 2014 only shared a significant correlation with the ME2-K14 dataset, but did not have significant correlations between the other two ENVs individually. Marianna 2014 had significant correlations with M16, ME2-M14, and All-M14 datasets. Marianna 2016 had the strongest correlations, where it was significantly correlated with all datasets affiliated with ME2, All-M16, and the ME1 dataset. There was also a significant correlation between ME1 and ME2 (r = .25).

Model Comparisons

Sparse testing GS (CV2) models, using GY data from other ENVs as covariates and GY data from only the ENV of interest in the TP, were compared to NGS models using GY from only the ENV of interest in the TP and no covariates using a five-fold cross-validation approach over 100 cycles for six different ENVs.

Between the three ME1 individual ENVs, CV2 models had significantly higher prediction accuracies than the NGS models. When ME2 was used as a covariate for all three ENVs, prediction accuracy was significantly decreased (Table 4, Figure 3).

The ENV with the greatest increase in prediction accuracy with a CV2 model compared to an NGS model (r = .43) was K17, where the model using M17 as a covariate (r = .64) significantly improved prediction accuracy by 21%. All other CV2 models also significantly

improve prediction accuracy except for the model using ME2 as a covariate. The M17 covariate model was not significantly different than the model using ME1-K17 as a covariate (r = .63) (Table 4, Figure 3a).

The ENV from ME1 with the second largest increase in prediction accuracy was M15, where two models, M17 (r = .50) and ME1-M15 (r = .50), significantly improved prediction accuracy compared to the NGS model (r = .30) with a 20% increase. The two top performing models were not significantly better than the All-M15 model (r = .49) (Table 4, Fig. 3b).

Marianna 2017 had the smallest improvement in prediction accuracy compared to the NGS model (r = .55), where the top performing model ME1-M17 (r = .66) significantly improved prediction accuracy by 11%. Unlike the other two ENVs within ME1, not all ME1 covariates significantly improved prediction accuracy for M17, where there was no change in prediction accuracy when M15 (r = .55) was used as a covariate (Table 4, Figure 3c).

The three ENVs in ME2 had far less improvement in prediction accuracy compared to the ENVs of ME1, where CV2 models only significantly improved prediction accuracy for K14 (Table 4, Figure 3). The top performing CV2 model for K14 was the model using ME2-K14 as a covariate, where the model significantly improved prediction accuracy by 6% compared to the NGS model (Table 4, Figure 3d). The CV2 models actually reduced prediction accuracy for the other two ENVs in ME2 (Table 4, Figures 3e and 3f).

DISCUSSION

Genomic selection has proven to be a valuable tool within the wheat breeder's toolbox, as shown in multiple studies throughout the past decade, for many different economically important traits (Heffner et al., 2009; Nicolas Heslot, Jannink, & Sorrells, 2015; Larkin et al., 2019; Sorrells,

2015). One of the primary goals of GS is to increase genetic gain for economically important traits within the breeding program by reducing the time within a breeding cycle and by increasing the accuracy of selection (Asoro et al., 2013; Bernardo & Yu, 2007; Heffner et al., 2009; Rutkoski et al., 2015; Schaeffer, 2006). There are many methods that can be used to improve the prediction accuracy of GS models, among these is the type of GS model used (Heffner et al., 2009; N. Heslot et al., 2012). We chose to focus on how we could use MVGS models to improve prediction accuracy for agronomic traits in SRWW by including secondary traits or ENVs correlated with our traits or ENVs of interest as covariates, much like a selection index (Calus & Veerkamp, 2011; Jia & Jannink, 2012).

Multivariate Genomic Selection for Agronomic Traits

Our study found that the MVGS models significantly increased prediction accuracies for all five agronomic traits, suggesting that the use of highly correlated secondary traits as covariates can be effective for increasing selection accuracy in a wheat breeding program. Previous studies with MVGS models have also observed that, in general, stronger genetic correlations between the trait of interest and the secondary trait increased prediction accuracy (Jia & Jannink, 2012; Lozada & Carter, 2019; Schulthess et al., 2016; Ward et al., 2019). A simulation study found that when predicting a low heritability trait, stronger genetic correlations with a secondary trait increased prediction accuracy. However, when the predicted trait had a high heritability, there was not a significant change in prediction accuracy as the genetic correlation with the secondary trait increased (Jia & Jannink, 2012). A study of four resistance traits to sudden death syndrome in soybean [*Glycine max* (L.) Merr.] found that MVGS models did not significantly improve prediction accuracy for the four traits, due to the fact that the four traits were weakly correlated with each other, which we also observed between weakly correlated traits (Bao, Kurle,

Anderson, & Young, 2015). This was also a common observation with all five of our agronomic traits. In fact, we observed an increase in prediction accuracy for any genetic correlation exceeding .30, regardless if it was positive or negative. This is far less conservative than a simulation study that found that prediction accuracies increased only if genetic correlations exceeded .50 (Calus & Veerkamp, 2011).

As with several different studies, we also found that the prediction accuracies for lower heritability traits could be improved through the use of higher heritability traits as covariates, and that the level of improvement in prediction accuracy increased as the heritability of the trait of interest decreased (Calus & Veerkamp, 2011; Guo et al., 2014; Jia & Jannink, 2012). However, unlike the results observed by Jia & Jannink (2012), we observed significant increases in prediction accuracy even when the predicted trait had a high heritability, particularly for GY, HD, PH, and TW. Regardless, our lowest heritability trait, MD, saw the largest gain in prediction accuracy compared to the other four traits.

Prior MVGS studies related to agronomic traits in wheat have primarily focused on GY as the trait of interest, while they have used high-throughput phenotyping traits, such as NDVI and canopy temperature as model covariates (Crain et al., 2018; Lozada & Carter, 2019; Rutkoski et al., 2016; Sun et al., 2017). These studies observed similar results to ours, in that the inclusion of these high-throughput phenotyping traits significantly improved prediction accuracy for GY compared to NGS models, especially when strong genetic correlations were observed between the secondary traits and GY. Two studies actually observed increases in prediction accuracy of up to 70% through the use of MVGS models compared to NGS models for GY (Rutkoski et al., 2016; Sun et al., 2017). Another study found increases in prediction accuracy for GY in low heritability datasets by 43 to 64% by including NDVI and normalized water index as

covariates. These were also conditional on a strong genetic correlation between the secondary traits and GY (Lozada & Carter, 2019). These large improvements were not always observed, however, as another study did not see as large of an improvement in prediction accuracy between NGS models and MVGS models when using NDVI and canopy temperature as covariates, where MVGS models improved prediction accuracy by only 7% at most and in some cases reduced prediction accuracy by as much as 33%, with an average increase in prediction accuracy of r = .14 compared to the NGS model (Crain et al., 2018).

Our study was unique in that, in addition to GY, we also focused on four other economically important traits that are typically collected throughout the growing season within a wheat breeding program. Since these traits can influence GY, it was worth investigating their influence on prediction accuracy for GY, as well as the other traits of interest, as covariates in MVGS models (X. J. Chen et al., 2012; Liu et al., 2015; Tshikunde et al., 2019).

Sparse Testing Genomic Selection for Grain Yield

Our study found that four out of our six ENVs of interest had significantly higher prediction accuracies for GY with models using GY data from correlated ENVs as covariates compared to NGS models without covariates. The cross-validation method using data from correlated ENVs in order to predict the performance of genotypes from a missing ENV is known as the CV2 method (Burgueno et al., 2012). Our procedure assumed that our TP had been grown in the ENV of interest, so unlike the traditional CV2 method, we only used data from the ENV of interest as phenotypic data for the TP, then we used phenotypic data from a closely related ENV, mega-ENV, or all other ENVs as a model covariate.

Much like the abovementioned MVGS models, we found that there was a greater increase in prediction accuracy when the covariate ENV dataset had a stronger correlation with the ENV of interest, as reflected in other MVGS studies (Calus & Veerkamp, 2011; Jia & Jannink, 2012). This was consistently evident in all three ENVs within ME1, where the top CV2 models also had the strongest correlations.

Unlike ME1, ME2 had results that were less consistent. There were significant correlations between M16 and all of its associated covariates, however, none of the CV2 models significantly improved prediction accuracies compared to the NGS model. There were also no CV2 models that significantly improved prediction accuracy for GY in M14 either, in fact they actually reduced prediction accuracy. The only ENV in ME2 that saw a significant increase in prediction accuracy due to the use of CV2 models was K14, where the models with M14, M16, and ME2-K14 covariates significantly improved prediction accuracy, this was in spite of the low correlations between the ENV of interest and the covariates. One explanation for the inconsistency in ME2 could come from the clustering algorithm itself. The ENVs in ME1 were strongly associated with each other, whereas the ENVs in ME2 were more loosely associated, where there was a smaller proportion of variation explained by the cluster compared to that of ME1. It is also notable that K14 was the most representative ENV for ME2, which could explain why K14 benefitted the most from the CV2 models compared to the other two ENVs in ME2.

Overall, our results are comparable to other studies where multi-environment models using the CV2 method outperformed the CV1 method, where the performance of newly developed lines is predicted with no prior testing data. The CV1 method was otherwise referred to as the NGS model in our study. One study found that models using the CV2 method outperformed models using the CV1 method when using pedigree information (.42 vs. .56),

marker data (.47 vs. .56), or both (.49 vs. .59) (Burgueno et al., 2012). Maximum prediction accuracies were also observed with the CV2 method in chickpea when compared with the CV1 method and CV0 method, where all lines are predicted for one ENV between nine ENVs and 13 models (Roorkiwal et al., 2018). Prediction accuracies were also higher for models using the CV2 method compared to the CV1 method when comparing seven models across five ENVs in Mexico (Saint Pierre et al., 2016).

CONCLUSIONS

Since plant breeders phenotype for many traits throughout the growing season, it makes sense that some of these traits would be genetically correlated with each other. Therefore, some of these traits that are collected earlier in the growing season could be used to improve the prediction accuracy of later season traits as part of MVGS models. Our study found that MVGS models significantly improved the prediction accuracy of five different agronomic traits of interest when the secondary traits used as covariates had strong genetic correlations with the predicted trait of interest. We also found that the percent increase in prediction accuracy between an NGS model and the top performing MVGS model increase as the heritability of the trait of interest decreased. This indicates that MVGS can be successfully used when breeding for agronomic traits. Before implementing MVGS in their breeding programs, breeders must consider the genetic correlations between their traits as well as the heritability of their traits as results can vary across populations. Breeders must also be careful not to unintentionally select for undesirable traits if they have strong genetic correlations with a trait of interest. An example would be unintentionally selecting for taller plants while using PH as a covariate to predict TW, as both traits are positively correlated.

This study also found that the use of MVGS models for a CV2 cross-validation scheme significantly improved the prediction accuracy of GY for individual ENVs when the GY data from a covariate ENV was strongly correlated with the individual ENV of interest. This can especially be helpful when predicting the performance of breeding lines in new ENVs, especially when those breeding lines have been grown in similar ENVs. When implementing this form of CV2 cross-validation, breeders must still consider the genetic relationships between the TP and VP while also considering the similarity between ENVs within their breeding program.

REFERENCES

Akdemir, D., Sanchez, J. I., & Jannink, J. L. (2015). Optimization of genomic selection training populations with a genetic algorithm. *Genetics Selection Evolution, 47*, 10. doi:10.1186/s12711-015-0116-6

Allard, R. W., & Bradshaw, A. D. (1964). Implications of Genotype-Environmental Interactions in Applied Plant Breeding1. *Crop Science, 4*(5), cropsci1964.0011183X000400050021x. doi:10.2135/cropsci1964.0011183X000400050021x

Anderberg, M. R. (2014). *Cluster analysis for applications: probability and mathematical statistics: a series of monographs and textbooks* (Vol. 19): Academic press.

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., . . . Manuscript Writing, T. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science, 361*(6403), 661-+. doi:10.1126/science.aar7191

Arruda, M. P., Brown, P., Brown-Guedira, G., Krill, A. M., Thurber, C., Merrill, K. R., . . . Kolb, F. L. (2016). Genome-wide association mapping of Fusarium head blight resistance in wheat using genotyping-by-sequencing. *Plant Genome, 9*(1), 14. doi:10.3835/plantgenome2015.04.0028

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. doi:10.2135/cropsci2012.09.0526

Bao, Y., Kurle, J. E., Anderson, G., & Young, N. D. (2015). Association mapping and genomic prediction for resistance to sudden death syndrome in early maturing soybean germplasm. *Molecular Breeding, 35*(6), 14. doi:10.1007/s11032-015-0324-3

Basford, K. E., & Cooper, M. (1998). Genotype x environment interactions and some considerations of their implications for wheat breeding in Australia. *Australian Journal of Agricultural Research, 49*(2), 153-174. doi:10.1071/a97035

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. doi: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. doi:10.2135/cropsci2006.11.0690

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics, 23*(19), 2633-2635. doi:10.1093/bioinformatics/btm308

Burgueno, J., de los Campos, G., Weigel, K., & Crossa, J. (2012). Genomic Prediction of Breeding Values when Modeling Genotype x Environment Interaction using Pedigree and Dense Molecular Markers. *Crop Science, 52*(2), 707-719. doi:10.2135/cropsci2011.06.0299

Calus, M. P. L., & Veerkamp, R. F. (2011). Accuracy of multi-trait genomic selection using different methods. *Genetics Selection Evolution, 43*, 14. doi:10.1186/1297-9686-43-26

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

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. doi:10.1016/j.fcr.2012.09.008

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. doi:10.3835/plantgenome2012.11.0030

Covarrubias-Pazaran, G. (2016). Genome-Assisted Prediction of Quantitative Traits Using the R Package sommer. *Plos One, 11*(6), 15. doi:10.1371/journal.pone.0156744

Covarrubias-Pazaran, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., . . . 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*, 13. doi: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), 14. doi:10.3835/plantgenome2017.05.0043

Cullis, B. R., Smith, A. B., & Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural Biological and Environmental Statistics, 11*(4), 381-393. doi:10.1198/108571106x154443

de Leon, N., Jannink, J. L., Edwards, J. W., & Kaeppler, S. M. (2016). Introduction to a Special Issue on Genotype by Environment Interaction. *Crop Science, 56*(5), 2081-2089. doi:10.2135/cropsci2016.07.0002in

Eberhart, S. A., & Russell, W. A. (1966). Stability Parameters for Comparing Varieties1. *Crop Science, 6*(1), cropsci1966.0011183X000600010011x. doi:10.2135/cropsci1966.0011183X000600010011x

Endelman, J. B. (2011). Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *Plant Genome, 4*(3), 250-255. doi: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. doi: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*, 7. doi: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. doi: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. doi:10.2135/cropsci2008.08.0512

Heslot, N., Jannink, J.-L., & Sorrells, M. E. (2015). Perspectives for genomic selection applications and research in plants. *Crop Science, 55*(1), 1-12.

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. doi: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. doi: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. doi:10.1093/bfgp/elq001

Jarquin, D., Howard, R., Crossa, J., Beyene, Y., Gowda, M., Martini, J. W. R., . . . Prasanna, B. M. (2020). Genomic Prediction Enhanced Sparse Testing for Multi-environment Trials. *G3-Genes Genomes Genetics, 10*(8), 2725-2739. doi:10.1534/g3.120.401349

Jia, Y., & Jannink, J. L. (2012). Multiple-Trait Genomic Selection Methods Increase Genetic Value Prediction Accuracy. *Genetics, 192*(4), 1513-+. doi:10.1534/genetics.112.144246

Kelley, J. (2018). 2018 Arkansas Wheat Quick Facts. In: 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., . . . Dong, Y. (2020). Genome-wide analysis and prediction of fusarium head blight resistance in soft red winter wheat. *Crop Science, n/a*(n/a). doi: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-Basel, 9*(9), 18. doi:10.3390/agronomy9090479

Lee, S. H., & van der Werf, J. H. J. (2016). MTG2: an efficient algorithm for multivariate linear mixed model analysis based on genomic information. *Bioinformatics, 32*(9), 1420-1422. doi:10.1093/bioinformatics/btw012

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

Lin, C. S., Binns, M. R., & Lefkovitch, L. P. (1986). Stability Analysis: Where Do We Stand?1. *Crop Science, 26*(5), cropsci1986.0011183X002600050012x. doi:10.2135/cropsci1986.0011183X002600050012x

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. doi: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. doi: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.

Mohring, J., & Piepho, H. P. (2009). Comparison of Weighting Methods in Two-Stage Analysis of Plant Breeding Trials. *Crop Science, 49*(6), 1977-1988. doi:10.2135/cropsci2009.02.0083

Money, D., Gardner, K., Migicovsky, Z., Schwaninger, H., Zhong, G. Y., & Myles, S. (2015). LinkImpute: Fast and Accurate Genotype Imputation for Nonmodel Organisms. *G3-Genes Genomes Genetics, 5*(11), 2383-2390. doi:10.1534/g3.115.021667

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.

Phillips, N. D. (2017). Yarrr! The pirate's guide to R. *APS Observer, 30*(3).

Piepho, H. P., & Mohring, J. (2007). Computing heritability and selection response from unbalanced plant breeding trials. *Genetics, 177*(3), 1881-1888. doi:10.1534/genetics.107.074229

Piepho, H. P., Mohring, J., Melchinger, A. E., & Buchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica, 161*(1-2), 209-228. doi:10.1007/s10681-007-9449-8

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), 8. doi:10.1371/journal.pone.0032253

R, C. T. (2020). R: A language and environment for statistical computing. In. R Foundation for Statistical Computing, Vienna, Austria.

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), 8. doi:10.1371/journal.pone.0066428

Roorkiwal, M., Jarquin, D., Singh, M. K., Gaur, P. M., Bharadwaj, C., Rathore, A., . . . Varshney, R. K. (2018). Genomic-enabled prediction models using multi-environment trials to estimate the effect of genotype x environment interaction on prediction accuracy in chickpea. *Scientific Reports, 8*, 11. doi:10.1038/s41598-018-30027-2

Rosielle, A. A., & Hamblin, J. (1981). THEORETICAL ASPECTS OF SELECTION FOR YIELD IN STRESS AND NON-STRESS ENVIRONMENTS. *Crop Science, 21*(6), 943-946. doi:10.2135/cropsci1981.0011183X002100060033x

Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Perez, L. G., Crossa, J., . . . 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. doi: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. doi:10.3835/plantgenome2014.10.0074

Saint Pierre, C., Burgueno, J., Crossa, J., Davila, G. F., Lopez, P. F., Moya, E. S., . . . Singh, S. (2016). Genomic prediction models for grain yield of spring bread wheat in diverse agro-ecological zones. *Scientific Reports, 6*, 11. doi:10.1038/srep27312

Sarinelli, J. M., Murphy, J. P., Tyagi, P., Holland, J. B., Johnson, J. W., Mergoum, M., . . . 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. doi:10.1007/s00122-019-03276-6

Sarle, W. (1990). The VARCLUS Procedure. SAS/STAT User's Guide: Cary NC: SAS Institute. In: Inc.

Schaeffer, L. R. (2006). Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics, 123*(4), 218-223. doi:10.1111/j.1439-0388.2006.00595.x

Schulthess, A. W., Wang, Y., Miedaner, T., Wilde, P., Reif, J. C., & Zhao, Y. S. (2016). Multiple-trait- and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes. *Theoretical and Applied Genetics, 129*(2), 273-287. doi:10.1007/s00122-015-2626-6

Sorrells, M. E. (2015). *Genomic Selection in Plants: Empirical Results and Implications for Wheat Breeding*, Tokyo.

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. doi: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*, 18. doi:10.3389/fpls.2019.01428

United Nations, D. o. E. a. S. A., Population Division. (2019). *World Population Prospects 2019: Highlights*. Retrieved from

VanRaden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science, 91*(11), 4414-4423. doi:10.3168/jds.2007-0980

Ward, B. P., Brown-Guedira, G., Tyagi, P., Kolb, F. L., Van Sanford, D. A., Sneller, C. H., & Griffey, C. A. (2019). Multienvironment and Multitrait Genomic Selection Models in Unbalanced Early-Generation Wheat Yield Trials. *Crop Science, 59*(2), 491-507. doi:10.2135/cropsci2018.03.0189

Table 1. Descriptive statistics, Pearson phenotypic correlations (bottom half), genetic correlations (top half), and broad-sense heritabilities (H^2) for adjusted means for 351 soft red winter wheat genotypes for five agronomic traits.

Trait	Environments[a]	Descriptive Statistics					H^{2b}	Correlations				
		Mean	Min.	Max.	Range	SD		GY[c]	HD[d]	MD[e]	PH[f]	TW[g]
GY	K14, K17, M14, M15, M16, M17	3.4	.1	9.5	9.4	1.1	.85	-	-.08[ns†]	-.16**	-.12*	.63***
HD	M14, M15, M16, M17	105.8	79.1	129.0	49.9	13.7	.94	-.17***	-	.80***	.30***	-.14**
MD	K14, M14, M15	149.7	138.2	159.0	20.8	4.2	.67	-.11*	.69***	-	.15**	-.33***
PH	K14, M14, M15, M16, M17	90.4	56.5	121.7	65.3	10.0	.91	-.14**	.35***	.19***	-	.31***
TW	K14, K17, M14, M15, M16, M17	73.1	45.6	85.2	39.8	4.0	.84	.43***	-.10[ns†]	-.20***	.22***	-

[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.

[b] Broad-sense heritability (H^2) for unbalanced, multi-environmental experimental designs calculated using the formula proposed in Cullis et al. (2006) for each of the five agronomic traits.

[c] Grain yield (GY) was recorded in t ha^{-1}.

[d] Heading date (HD) was recorded in Julian days after 1 Jan, when 50% of the heads were 50% emerged from the flag leaf.

[e] Maturity date (MD) was recorded in Julian days after 1 Jan, when an estimated 90% of culms in an individual plot were senesced or yellow.

[f] Plant height (PH) was recorded in cm from the surface of the soil to the tip of the head.

[g] Test weight (TW) was recorded in kg hl^{-1}.

*Significant at the .05 probability level. **Significant at the .01 probability level. ***Significant at the .001 probability level. †ns, nonsignificant at the .05 probability level.

Table 2. A comparison of mean prediction accuracies between multivariate genomic selection (MVGS) models and naïve genomic selection (NGS) for five agronomic traits in soft red winter wheat, including grain yield (GY), heading date (HD), maturity date (MD), plant height (PH), and test weight (TW).

Trait	Covariate(s)[a]	PA
GY	TW	.63a[b]
	NGS	.53b
	HD	.53b
	PH	.53b
	MD	.53b
HD	MD+PH	.80a
	MD	.79b
	PH	.63c
	TW	.61d
	GY	.60d
	NGS	.60d
PH	HD+TW	.65a
	HD	.60b
	TW	.60b
	MD	.57c
	NGS	.56cd
	GY	.56d
TW	GY+PH+MD	.61a
	GY+PH	.58b
	GY+MD	.54c
	GY	.52d
	PH+MD	.47e
	PH	.44f
	MD	.38g
	HD	.36h
	NGS	.36h
MD	HD+TW	.73a
	HD	.73a
	TW	.45b
	PH	.44c
	NGS	.44cd
	GY	.43d

[a] Covariates used as fixed effects in the MVGS models. Covariates listed with (+) indicate a model with more than one covariate.

[b] Mean prediction accuracies of each model not sharing any letter are significantly different based on Fisher's LSD separation at the 5% level of significance.

Table 3. Descriptive Statistics and broad-sense heritabilities (H^2) for grain yield (GY) of covariates used in multivariate, sparse-testing genomic selection (GS) models.

Covariate[a]	Environments[b]	Descriptive Statistics					H^{2c}
		Mean	Min.	Max.	Range	SD	
All-K17	M15, M17, K14, M14, M16	4.17	.65	9.54	8.89	1.02	.77
All-M15	K17, M17, K14, M14, M16	3.81	.10	9.54	9.44	1.16	.79
All-M17	K17, M15, K14, M14, M16	3.73	.10	9.54	9.44	1.20	.74
All-K14	K17, M15, M17, M14, M16	4.03	.10	9.54	9.44	1.09	.82
All-M14	K17, M15, M17, K14, M16	3.72	.10	9.54	9.44	1.12	.82
All-M16	K17, M15, M17, K14, M14	3.70	.10	7.33	7.23	1.15	.83
ME1	K17, M15, M17	3.68	.10	6.11	6.01	1.09	.88
ME1-K17	M15, M17	4.30	.84	6.11	5.26	.78	.80
ME1-M15	K17, M17	3.54	.10	6.11	6.01	1.09	.88
ME1-M17	K17, M15	3.25	.10	5.62	5.52	1.08	.75
ME2	K14, M14, M16	4.08	.65	9.54	8.89	1.17	.66
ME2-K14	M14, M16	4.60	.86	9.54	8.68	.83	.62
ME2-M14	K14, M16	3.78	.65	9.54	8.89	1.18	.62
ME2-M16	K14, M14	3.75	.65	7.33	6.68	1.29	.65
K17	-	2.72	.10	4.33	4.23	.74	.80
M15	-	4.23	.84	5.62	4.77	.90	.58
M17	-	4.33	1.25	6.11	4.86	.71	.89
K14	-	2.57	.65	4.01	3.36	.47	.73
M14	-	4.88	3.58	7.33	3.75	.55	.42
M16	-	4.42	.86	9.54	8.68	.90	.43

[a] Covariates used for multivariate, sparse-testing GS models. The first six covariates consist of adjusted means for GY (t ha^{-1}) across all environments, except for the environment of interest. The next eight covariates consist of adjusted means from two mega-environments (ME1 and ME2), ME1 and ME2 include all three environments within each mega-environment (ME) and they are followed by each ME minus the environment of interest. The last six covariates are individual environments.

[b] Environments within each covariate: K14, Keiser 2014; K17, Keiser 2017; M14, Marianna 2014; M15, Marianna 2015; M16, Marianna 2016; M17, Marianna 2017.

[c] Broad-sense heritability (H^2) for unbalanced, multi-environmental experimental designs calculated using the formula proposed in Cullis et al. (2006) for each covariate consisting of multiple environments. Broad-sense heritabilities for individual environments were calculated on a plot-mean basis.

Table 4. Mean prediction accuracies obtained from multivariate, sparse testing (CV2) genomic selection (GS) models. An environment (ENV) of interest provides grain yield (GY) data from the training population and adjusted means for GY from all ENVs, a mega-environment (ME), or individual ENV, excluding the ENV of interest are used as covariates. Pearson correlations between GY for the ENV of interest and covariates were also obtained.

Environment[a]	Covariate[b]	PA[c]	Correlations[d]
All	-	.54	-
ME1	ME2	.53a[e]	.25***
	-	.53a	-
ME2	ME1	.34a	.25**
	-	.33a	-
K17	M17	.64a	.63***
	ME1-K17	.63a	.62***
	All-K17	.55b	.48***
	M15	.45c	.42***
	-	.43d	-
	ME2	.15e	.19***
M15	M17	.50a	.52***
	ME1-M15	.50a	.52***
	All-M15	.49a	.51***
	K17	.40b	.42***
	-	.30c	-
	ME2	.24d	.24***
M17	ME1-M17	.66a	.66***
	K17	.62b	.63***
	All-M17	.59c	.55***
	-	.55d	-
	M15	.55d	.52***
	ME2	.18e	.24***
K14	ME2-K14	.20a	.20**
	M14	.18b	.14[ns†]
	M16	.16b	.17[ns]
	-	.14c	-
	All-K14	.13c	.11[ns]
	ME1	.02d	-.02[ns]
M14	-	.43a	-
	ME2-M14	.27b	.19*
	All-M14	.26b	.18*
	M16	.21c	.16*
	K14	.20c	.14[ns]
	ME1	-.10d	.10[ns]
M16	-	.30a	-

Table 4 (Cont.)

Environment[a]	Covariate[b]	PA[c]	Correlations[d]
M16	All-M16	.30a	.30***
	ME2-M16	.28b	.21**
	ME1	.27bc	.28***
	M14	.26c	.17*
	K14	.18d	.20*

[a] The ENV of interest is the source of GY data for the training population. The CV2 GS models predicted genome-estimated breeding values (GEBVs) for GY for both mega-environments and all six ENVs of interest: All, consists of all ENVs; ME1, mega-environment 1; mega-environment 2; K14, Keiser 2014; K17, Keiser 2017; M14, Marianna 2014; M15, Marianna 2015; M16, Marianna 2016; M17, Marianna 2017.

[b] The covariate used for the CV2 GS models.

[c] Prediction accuracy calculated as the Pearson correlation (*r*) between the GY and GEBVs calculated for the genotypes in the validation population using the CV2 GS models.

[d] Pearson correlations between the GY of the ENVs of interest and the covariates used for the CV2 GS models.

[e] Mean prediction accuracies of each model not sharing any letter are significantly different based on Fisher's LSD separation at the 5% level of significance.

*Significant at the .05 probability level. **Significant at the .01 probability level. ***Significant at the .001 probability level. †ns, nonsignificant at the .05 probability level.

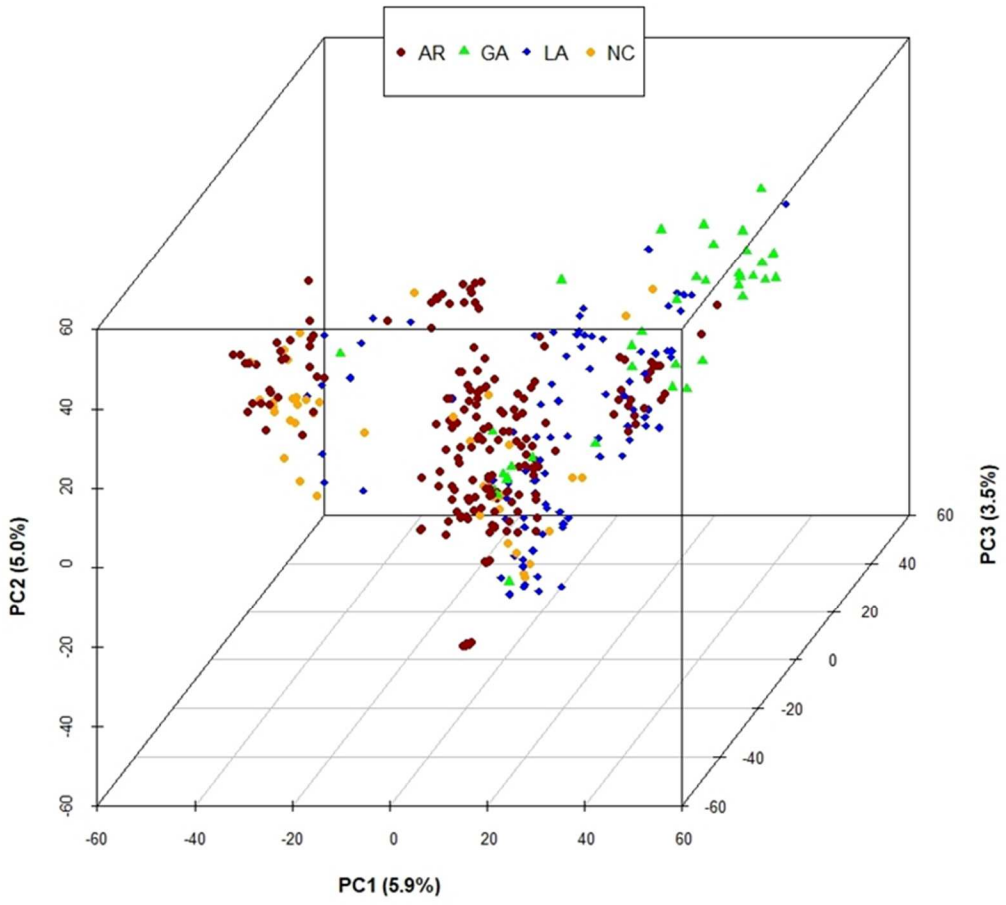


Figure 1. Population structure of 351 soft red winter wheat genotypes using 73,618 single nucleotide polymorphisms (SNPs). Colors represent the origin of the genotypes. AR, developed at the University of Arkansas, Fayetteville; GA, developed at the University of Georgia, Athens; LA, developed at Louisiana State University, Baton Rouge; NC, developed at North Carolina State University, Raleigh; PC, principal component.

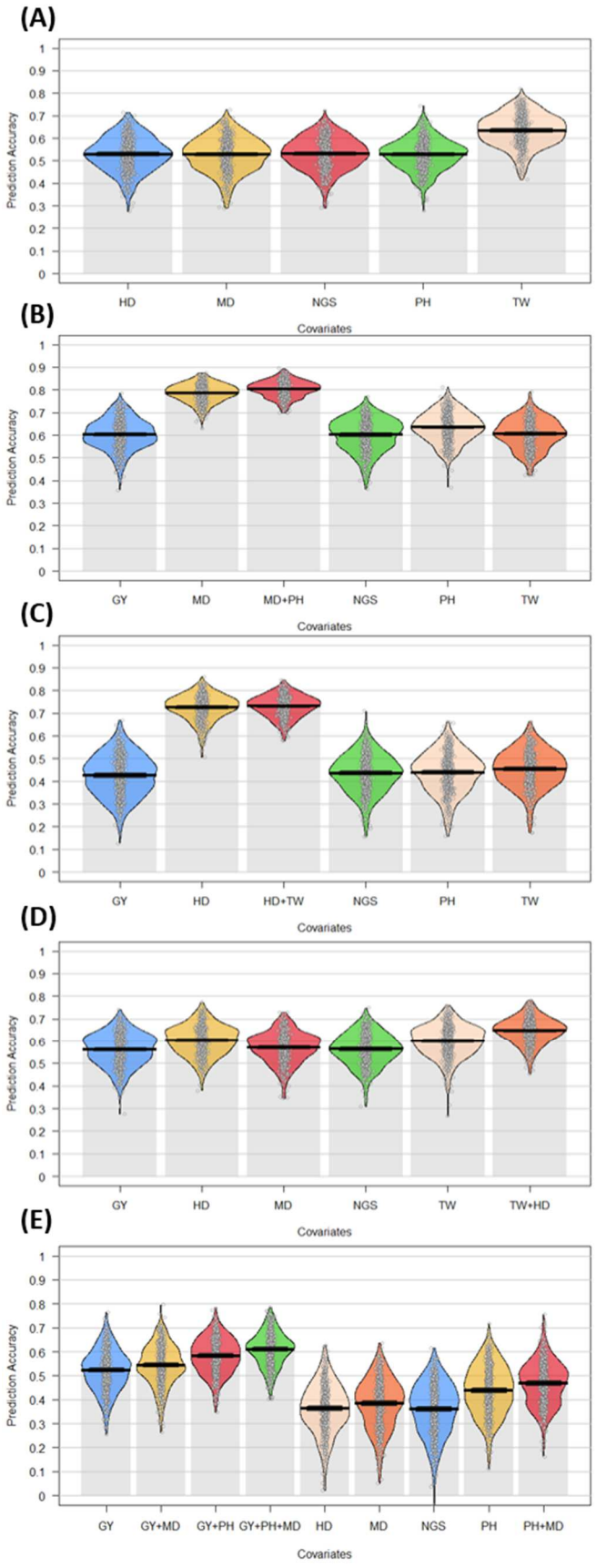

(A)
Prediction Accuracy
HD
MD
NGS
PH
TW
Covariates
(B)
Prediction Accuracy
GY
MD
MD+PH
NGS
PH
TW
Covariates
(C)
Prediction Accuracy
GY
HD
HD+TW
NGS
PH
TW
Covariates
(D)
Prediction Accuracy
GY
HD
MD
NGS
TW
TW+HD
Covariates
(E)
Prediction Accuracy
GY
GY+MD
GY+PH
GY+PH+MD
HD
MD
NGS
PH
PH+MD
Covariates

Figure 2. Pirate plots comparing the mean prediction accuracies between multivariate genomic selection (MVGS) models with a naïve genomic selection model (NGS) for five agronomic traits in soft red winter wheat: (A) grain yield (GY), (B) heading date (HD), (C) maturity date (MD), (D) plant height (PH), (E) and test weight (TW). The *x* axis represents the combination of covariates used for each model, including each of the agronomic traits. The *y* axis represents the mean prediction accuracy across 100 iterations of fivefold cross-validation in the form of a Pearson correlation coefficient (*r*) between the predicted genome estimated breeding value (GEBV) and the actual phenotypic value for the validation populations. Individual points represent the Pearson correlation from each fold of each iteration of cross-validation for a total of 500 datapoints. The lines within each plot represent the mean and 95% confidence intervals for prediction accuracy. The curves represent the smoothed densities of the data.

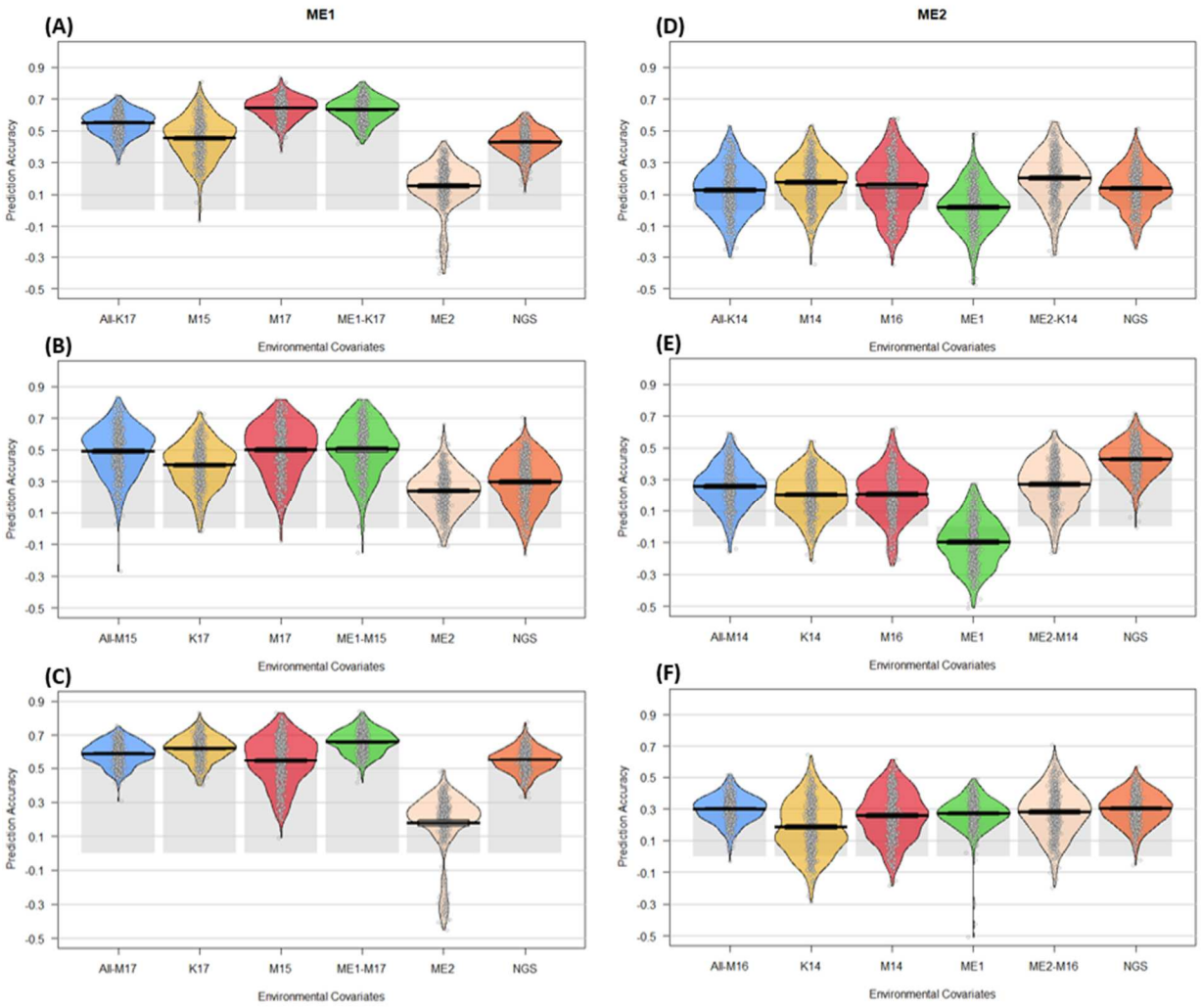


Figure 3. Pirate plots comparing the mean prediction accuracies between multivariate, sparse testing (CV2) genomic selection (GS) models, using grain yield (GY) data from other environments as covariates, with a naïve genomic selection model (NGS) to predict GY for six environments (ENVs) of interest within two mega-environments. Mega-environment 1 (ME1) consists of: (A) Keiser 2017 (K17), (B) Marianna 2015 (M15), and (C) Marianna 2017 (M17). Mega-environment 2 (ME2) consists of: (D) Keiser 2014 (K14), Marianna 2014 (M14), and Marianna 2016 (M16). The adjusted means for GY from the ENV of interest is used as data for the training population. The *x* axis represents the ENV covariates used in each model. The *y* axis represents the mean prediction accuracy across 100 iterations of fivefold cross-validation in the form of a Pearson correlation coefficient (*r*) between the predicted genome estimated breeding value (GEBV) and the actual phenotypic value for the validation populations. Individual points represent the Pearson correlation from each fold of each iteration of cross-validation for a total of 500 datapoints. The lines within each plot represent the mean and 95% confidence intervals for prediction accuracy. The curves represent the smoothed densities of the data.

CHAPTER IV

PREDICTING THE PERFORMANCE OF OBSERVATION TRIALS IN A SOFT RED WINTER WHEAT BREEDING PROGRAM WITH GENOMIC SELECTION

ABSTRACT

Many studies have evaluated the effectiveness of genomic selection (GS) using cross-validation within training populations (TPs), however few have looked at its implementation and performance over multiple generations within a breeding program. The objectives for this study were to compare the performance of naïve GS (NGS) and multivariate GS (MVGS) models by predicting the performance of three years of $F_{4:6}$ breeding lines for grain yield (GY) in soft red winter wheat and comparing the predictions with actual phenotypic performance over three years of selection. We also compared the performance of NGS and MVGS models when predicting three Fusarium head blight resistance traits, deoxynivalenol (DON) accumulation, Fusarium damaged kernels (FDK), and severity (SEV) for two years of $F_{4:7}$ breeding lines and comparing predictions with phenotypic results across two years of selection. For GY, the MVGS model was comparable to phenotypic selection in terms of response to selection ($t = 2.14$, $p = .058$). The MVGS model also had a higher selection accuracy for elite breeding lines during two of the three breeding cycles for GY and both cycles for DON and at least one cycle for FDK and SEV, compared to phenotypic selection. On average, the MVGS model also had higher prediction accuracies for GY, DON, FDK, and SEV across breeding cycles compared to the NGS model. This shows that MVGS can be successfully implemented in a wheat breeding program over multiple breeding cycles and can be effective alongside phenotypic selection for economically important traits.

INTRODUCTION

Wheat (*Triticum aestivum* L.) is one of the most important cereal crops in the world accounting for over 20% of caloric and protein intake for humans, globally (Ray, Mueller, West, & Foley, 2013) With a global population that is expected to increase to over 9.7 billion people by 2050, it is increasingly important that both public and private wheat breeders increase the genetic gain for grain yield (GY) and other economically important traits within their breeding programs. In addition to GY, resistance to the disease Fusarium head blight (FHB) is incredibly important, particularly in the Southeastern US. Fusarium head blight is a fungal disease caused by the *Fusarium graminearum* pathogen and incurs nearly US$4.2 in losses annually (Wilson, McKee, Nganje, Dahl, & Bangsund, 2017). The *F. graminearum* pathogen produces the mycotoxin deoxynivalenol (DON), which is harmful for humans and animals that consume infected grain (Safety & Compliance, 2010; Sobrova et al., 2010).

Traditionally, wheat breeders have primarily relied on phenotypic selection within their breeding programs in order to advance breeding material. However, phenotypic selection has its limitations, especially with low-heritability traits of interest that are difficult to phenotype. Difficulties with phenotyping are also compounded by genotype x environment interactions that can complicate phenotyping by masking the expression of certain traits, reducing the accuracy of selections. Alternatives to phenotypic selection include marker assisted selection (MAS) and genomic selection (GS). Marker assisted selection can be effective for qualitative traits controlled by one or two genes or quantitative traits that are controlled by large-effect quantitative trait loci (QTL) (Xu & Crouch, 2008). However, MAS is less effective for complex quantitative traits controlled by many small effect QTL (Bernardo & Yu, 2007; Heffner, Sorrells, & Jannink, 2009). Genomic selection is an effective alternative to both phenotypic selection and

MAS, in that it incorporates allelic effects across the entire genome, making it ideal for highly quantitative traits. Genomic selection can also reduce the time within a breeding selection, where two rounds of GS can be performed compared to one cycle of phenotypic selection allowing for greater genetic gain over time (Asoro et al., 2013; Bernardo & Yu, 2007; Heffner et al., 2009; Rutkoski et al., 2015).

Genomic selection was first applied to animal breeding, particularly in the dairy industry, but it has since been adapted by plant breeders over the last decade (Heffner et al., 2009; Meuwissen, Hayes, & Goddard, 2001). Genomic selection uses a training population (TP), which is a panel of genotypes that have been phenotyped for a trait of interest and also genotyped using whole-genome sequencing, in order to train a genomic prediction model. The genomic prediction model is then used to assign allelic effects to markers throughout the genome in order to calculate genome-estimated breeding values (GEBVs) for breeding lines, otherwise known as the validation population (VP), that have only been genotyped. The breeder can then make selections based on the GEBVs for a trait of interest (Meuwissen et al., 2001).

Most studies involving GS have focused on increasing prediction accuracy by manipulation and cross-validation of the TP (Akdemir, Sanchez, & Jannink, 2015; Combs & Bernardo, 2013; Habier, Fernando, & Dekkers, 2007; Heffner et al., 2009; Isidro et al., 2015; Jannink, Lorenz, & Iwata, 2010; Larkin, Lozada, & Mason, 2019). Many have also investigated the genomic prediction model used for GS analysis (Heslot, Yang, Sorrells, & Jannink, 2012). While these methods are valuable, few have researched the effectiveness of applying GS in breeding programs for forward prediction of breeding lines (Bernardo, 2016). However, when investigated, many have seen mixed results regarding prediction accuracy of forward prediction, compared to the prediction accuracy of TPs (Asoro et al., 2013; Belamkar et al., 2018; Calvert,

Evers, Wang, Fritz, & Poland, 2020; Combs & Bernardo, 2013; Massman, Jung, & Bernardo, 2013; Michel et al., 2017).

In an evaluation of GS in the Kansas State University wheat breeding program, GS was used to predict GY in a TP where the prediction accuracy was between $r = .31$ and $r = .47$. However, when the TP was used for forward prediction, the highest prediction accuracy between the GEBVs for GY in the preliminary yield trials (PYTs) and the actual phenotypic results for GY was $r = -.16$ (Calvert et al., 2020). This trend was also observed in an evaluation of the University of Nebraska wheat breeding program, where GY data from PYTs from three years were used to predict the performance of a fourth year. When no lines for the fourth year were included in the TP, prediction accuracies for GY were between $r = .22$ and $r = .26$. However, as more lines from the fourth year were included in the TP, the prediction accuracy of GY for the fourth year increased to between $r = .37$ and $r = .52$, when 90% of the lines from the fourth year were included in the TP (Belamkar et al., 2018). Phenotypic selection and GS were also compared in terms of selection accuracy between the PYT and advanced yield trial generations. Genomic selection outperformed phenotypic selection during the 2012 and 2015 seasons, where Nebraska experienced severe drought and disease stress. Even still, prediction accuracies were low, indicating that prediction accuracy is not the best indicator of GS success for forward prediction (Belamkar et al., 2018). Another study using forward prediction for GY in wheat adapted to central Europe found that the use of GS ($r = .39$) to select high performing lines for multiple-environment trials was far better than phenotypic selection ($r = .21$) (Michel et al., 2017).

In addition to traditional GS, some have started to research the efficacy of multivariate GS (MVGS). Multivariate GS uses mixed models that incorporate secondary traits that are

genetically correlated with a trait of interest as covariates in order to improve the prediction accuracy for the trait of interest (Calus & Veerkamp, 2011; Covarrubias-Pazaran et al., 2018; Y. Jia & Jannink, 2012). Additionally, MVGS can improve prediction accuracies for low-heritability traits when, high-heritability secondary traits are used as covariates (Calus & Veerkamp, 2011; Guo et al., 2014; H. Y. Jia et al., 2018). Many studies have evaluated MVGS models for cross-validation, particularly for GY in wheat using high-throughput phenotyping traits (Crain, Mondal, Rutkoski, Singh, & Poland, 2018; Dennis N. Lozada & Carter, 2019; Rutkoski et al., 2016; Sun et al., 2017). Others have evaluated resistance traits related to FHB in wheat using traits such as heading date (HD), plant height (PH), or other FHB resistance traits (Larkin et al., 2020; Moreno-Amores, Michel, Miedaner, Longin, & Buerstmayr, 2020; Rutkoski et al., 2012; Schulthess, Zhao, Longin, & Reif, 2018; Steiner et al., 2019). Few have evaluated the use of multivariate GS (MVGS) for forward prediction. However, one study used high-throughput phenotyping traits as a covariate in a MVGS model for forward prediction of GY in wheat, however the prediction accuracy was unfavorable unless a large TP was used (Calvert et al., 2020). Therefore, our aim is to validate the use of MVGS models compared to naïve GS (NGS) without covariates using secondary traits regularly collected throughout the season within a breeding program.

The University of Arkansas soft red winter wheat (SRWW) breeding program makes over 800 unique crosses per year. Progeny are then tested over the following 10 seasons in order to release a cultivar (Mason et al., 2018). The F_1 progeny are grown in headrows, bulk harvested, and planted into F_2 plots, where 50 heads are selected based on disease resistance, PH, HD, and agronomic performance. The 50 F_2 heads are then bulked and planted into F_3 plots where they are again evaluated for the aforementioned traits. Fifty heads are again selected per bulk and

planted into individual $F_{3:4}$ headrows where rows were selected based on agronomic performance, HD, PH, and disease resistance. Eight heads are then collected from selected headrows and four heads from each selected $F_{3:4}$ headrow are planted as $F_{4:5}$ headrows at two locations. The $F_{4:5}$ headrows are then evaluated and selected based on agronomic performance, HD, PH, and disease resistance and whole rows are bulk harvested and planted into $F_{4:6}$ yield plots in the Wheat Observation nursery (WO) in one location in an unreplicated augmented design with replicated check cultivars. Breeding lines in the WO nursery are genotyped using genotyping by sequencing (GBS) and Kompetitive allele specific polymerase chain reaction (PCR) (KASP) markers associated with major genes associated with PH, vernalization, photoperiod, and disease resistance. The WO nursery is also evaluated for the abovementioned traits and harvested for GY and test weight (TW). The Arkansas Advanced yield trial (ADV; $F_{4:7}$) is the first multiple environment yield trial, planted into yield plots at five locations and replicated twice per location in a randomized complete block design (RCBD) with replicated check cultivars. Breeding lines selected from the ADV yield trials are then advanced to the Arkansas Elite yield trials (ARE; $F_{4:8}$-$F_{4:11}$) where they are grown at five locations in a RCBD with four replications per location over three years. Few elite lines are also grown in regional statewide variety testing trials, as well as the USDA-ARS Uniform Eastern (UE) and Southern nurseries (US), Southeastern University Grains (Sungrains) cooperative nurseries, and foundation seed increases. The UE and US nurseries include approximately 36 elite breeding lines from public and private SRWW breeding programs in the Southern and Eastern US, grown between 22 and 36 locations with between one and three replications per location annually (Boyles, Marshall, & Bockelman, 2019). The Sungrains cooperative consists of Southeastern US SRWW breeding programs that performs regional testing within the Southeastern US (Boyles et

al., 2019; Harrison et al., 2017; Johnson et al., 2017; Mason et al., 2018). Select breeding lines from the WO, ADV, and ARE are grown in these regional Sungrains nurseries. The F_2, $F_{3:4}$, WO, ADV, and ARE breeding lines are also tested in an inoculated stripe rust nursery at one location, unreplicated for the F_2, $F_{3:4}$, and WO breeding lines and in a RCBD with two replications at a single location for the ADV and ARE breeding lines. The $F_{4:5}$, WO, ADV, and ARE breeding lines are also tested in an inoculated FHB nursery, unreplicated at one location for the $F_{4:5}$ and WO breeding lines and two replications at two locations for the ADV and ARE breeding lines.

The ADV and ARE allow for more accurate phenotypic selections for GY due to the greater number of locations and replication. However, selection at the WO generation is less accurate as it is unreplicated and only grown at one environment. In theory, GS can help to improve selection accuracy in the early generations of the breeding program for GY and FHB resistance traits. In this study, we wanted to evaluate the selection accuracy of GS from the WO through ARE generations compared to phenotypic selection through forward prediction using NGS and MVGS models. The three goals for this study were to: (1) compare NGS and MVGS with phenotypic selection for GY and three FHB resistance traits, including DON accumulation, Fusarium damaged kernels (FDK), and severity (SEV) for new breeding lines that have not been phenotyped at the WO and ADV generations; (2) compare the selection accuracy between NGS, MVGS, and phenotypic selection between the ADV and ARE generations of the University of Arkansas SRWW breeding program; and (3) compare the response to selection between NGS, MVGS, and phenotypic selection between the ADV and ARE generations of the University of Arkansas SRWW breeding program.

MATERIALS AND METHODS

Plant Materials

Breeding Materials

In order to predict GY, three generations of the WO nurseries, from the 2016-2017 through 2018-2019 growing seasons, were used as VPs for this study, consisting of $F_{4:6}$ breeding lines from the University of Arkansas wheat breeding program and affiliated breeding programs within the Sungrains cooperative and doubled haploid (DH) lines developed through the Sungrains cooperative. A doubled haploid WO population was also grown during the 2018-2019 season (DHWO19). Whereas, two generations of the ADV trials, 2017-2018 and 2018-2019, consisting of $F_{4:7}$ breeding lines from the University of Arkansas wheat breeding program and DH lines developed through the Sungrains cooperative, were used as VPs to predict three FHB traits, DON, FDK, and SEV since FHB data were not collected for the WO populations. Approximately 75% of the WO17 and WO18 breeding lines were advanced to the ADV yield trials in the following 2017-2018 (ADV18) and 2018-2019 (ADV19) growing seasons, while approximately 33% of the WO19 breeding lines were advanced to the ADV20 yield trial. Subsequently, 20% of breeding lines from the ADV18 and ADV19 yield trials were selected and advanced to the ARE19 and ARE20 yield trials for the 2018-2019 and 2019-2020 growing seasons, respectively. Genotypes were advance based on both GS and phenotypic selection (Table 1).

Training Populations

A population of 355 SRWW genotypes were used as the initial 2017 TP (TP17) for this study in order to predict GEBVs for GY in the WO17 nursery and GEBVs for DON, FDK, and SEV in

the ADV18 trial (TP18_FHB). The population consisted 187 genotypes from the University of Arkansas, 87 from Louisiana State University, 40 from North Carolina State University, 38 from the University of Georgia, and one genotype each from Syngenta AG, Pioneer Hi-Bred International, Inc., and Virginia Polytechnic Institute and State University (Larkin et al., 2020).

The 2018 TP (TP18) for GY consisted of the 355 genotypes from TP17 as well as the 140 breeding lines from WO17 nursery. The 2019 TP (TP19) for GY also consisted of the 355 genotypes from TP17 as well as the 140 genotypes from WO17, data from the 104 genotypes advanced to the ADV18 yield trial, and the 160 genotypes from WO18. The 2019 TP (TP19_FHB) for the three FHB traits consisted of the 355 genotypes from TP18_FHB, as well as the 104 genotypes from the ADV18 trial (Table 2).

Experimental Design and Trait Measurements

Winter wheat is planted during the fall and harvested during the late spring in the southern U.S., therefore the growing season spans two years. The TP18_FHB genotypes were evaluated for three FHB resistance traits, including DON, FDK, and SEV, over four seasons between 2014 and 2017 at two locations, at the Milo J. Shult Agricultural Research and Extension Center in Fayetteville, AR, USA (FAY) and the Newport Research and Extension Center near Newport, AR, USA (NPT). The data collection and experimental design methods were outlined in Larkin et al. (2020), as TP18_FHB was the same population used in their study.

Data were also collected for three agronomic traits including GY, TW, and HD in the TP17 population. The first 175 entries of the population were evaluated at the Lon Cotton Research Station near Marianna, AR, USA (MAR) and the Northeast Research and Extension Center near Keiser, AR, USA (KSR) during the 2013-2014 growing season (M14 and K14),

while the remaining 180 entries were evaluated at MAR during the 2014-2015 growing season (M15). All entries of TP17 were evaluated at MAR during 2015-2016 growing season (M16) and at both locations during the 2016-2017 growing season (M17 and K17).

Genotypes 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 between rows. The plots were then end-trimmed in the spring. Plots were managed according to the recommendations for wheat in Arkansas (Kelley, 2018). Each year, plots in MAR received two applications of urea fertilizer of 101 kg ha^{-1} and 67 kg ha^{-1}, along with 27 kg ha^{-1} of ammonium sulfate fertilizer. Plots in KSR 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), Harmony Extra (DuPont de Nemours, Inc.), Osprey (Bayer AG), and Prowl H_2O (BASF SE), were used each year for weed control.

The M14 and M15 ENVs were planted in an unbalanced randomized complete block design (RCBD) with two replications with repeated checks. In M14, the first replication consisted of 175 genotypes, while the second replication consisted of the first 117 genotypes from the first replication, while in M15, the first replication consisted of 177 genotypes, while the second replication consisted of the first 120 genotypes from the first replication. The K14, M16, M17, and M17 ENVs were planted in an unreplicated augmented design with repeated checks. During the 2013-2014 growing season, the repeated checks were ‘Branson’ (PI 639227), ‘Croplan Genetics 514W’, ‘Jamestown’ (PI 653751), ‘Pioneer 26R20’ (PI 658150), ‘Armor Ricochet’, ‘Shirley’ (PI 656753), ‘SS 8461’, and ‘UniSouth Genetics 3120’ (PI 672163). The 2014-2015 growing season used the same eight checks, with the addition of ‘SY Harrison’ (PI 664946), ‘Pioneer 26R41’, and ‘Pat’ (PI 631466). Only four checks, Branson, Jamestown, Pioneer 26R41, and SS 8641, were used during the 2015-2016 and 2016-2017 growing seasons.

Individual plots were harvested with a 1994 Hege 140 plot combine (Hege Maschinen GmbH) and GY was calculated at all six ENVs from the weight of seed from each plot as measured by the HarvestMaster Pro 4100 (Juniper Systems, Inc.) and was expressed in t ha^{-1} and adjusted for 13% moisture content. Test weight was sampled from each plot for all six ENVs using a HarvestMaster Pro 4100 (Juniper Systems, Inc.), and was expressed in kg hl^{-1} then adjusted for 13% moisture content. Heading dates for each genotype were collected during all four growing seasons from MAR. Heading date was recorded in Julian days after 1 Jan., when 50% of the heads were 50% emerged from the flag leaf. As there was variation in HD between genotypes in the population, heading notes were recorded every other day from the onset of heading and continuing until all plots in the nursery were headed.

Breeding trials were planted and grown using the same methods described for the TP17. Data for GY, HD, and TW were also collected for the WO, ADV, and ARE trials using the same methods described for the TP17. The WO nurseries were grown over three growing seasons, between 2016-2017 and 2018-2019 in MAR using an unreplicated augmented design with two repeated checks, Jamestown and Pioneer 26R41. Two additional repeated checks were used with the DH population in the WO19 nursery, 'AGS 2055' (PI 678970) and AR06146E-1-4.

The ADV yield trials were grown over three years, between the 2017-2018 and 2019-2020 growing seasons with anywhere between two and four locations. The ADV18 location was planted in KSR, NPT, MAR, and the Rohwer Research Station near Rohwer, AR, USA (RWR). The ADV yield trial grown during the 2018-2019 (ADV19) growing season only had data collected from MAR and NPT as the other locations were not planted due to late rainfall. Therefore, the ADV19 genotypes were replanted at KSR, MAR, and NPT during the 2019-2020 growing season (ADV19R). Data for the ADV yield trial grown during the 2019-2020 growing

season was collected from KSR, MAR, and RWR. The ADV yield trials were planted in a RCBD with two replications, with the exception of ADV19R, which was planted in an unreplicated augmented design. Each ADV yield trial had four repeated checks, AGS 2055, 'Hilliard' (PI 676271), Jamestown, and Pioneer 26R41. The ADV19R yield trial only used Jamestown and Pioneer 26R41 as repeated checks.

The ARE yield trials were grown over two years between the 2018-2019 and 2019-2020 growing seasons (ARE19 and ARE20). The ARE19 yield trial was grown at two locations, MAR and NPT, while the ARE20 yield trial was grown at KSR, MAR, NPT, and RWR. Each ARE yield trial was grown in a RCBD with four replications, with the exception of RWR which only had three replications. The ARE yield trials used the same four repeated checks as the ADV yield trials.

The AVD18, ADV19, and ARE19 FHB nurseries for the 2017-2018 and 2018-2019 growing seasons were grown at two locations, FAY and NPT, in a RCBD with two replications using the same methods described with the TP18_FHB and TP19_FHB populations and Larkin et al. (2020). This was also the case for the ARE20 FHB nursery, however it was only grown in NPT during the 2019-2020 season due to poor growing conditions in FAY. Data were also collected for HD, PH, DON, FDK, and SEV for the FHB nurseries using methods described in Larkin et al. (2020).

Phenotypic Data Analyses

Phenotypic data were analyzed using two different mixed linear models within the PROC MIXED procedure in SAS 9.4 (SAS Institute). The first model was a single-stage model that was used to obtain adjusted means for GY, TW, DON, FDK, and SEV for the TP18_FHB, ADV18,

ADV19, ADV20, ARE19, and ARE20 trials, which was described in Larkin et al. (2020) as follows:

$$y_{ijk} = \mu + genotype_i + rep(env)_{jk} + env_k + (genotype \times env)_{ik} + \varepsilon_{ijk}$$

where y_{ijk} was the observed phenotype, μ was the population mean, $genotype_i$ was the fixed effect of the i^{th} genotype, $rep(env)_{jk}$ was the random effect of the j^{th} replication nested within the k^{th} location-year (or location) (*env*), env_k was the random effect of the k^{th} location-year (or location), $(genotype \times env)_{ik}$ was the random effect of the interaction between genotype and location-year (or location), and ε_{ijk} was the residual error term, where $\varepsilon_{ijk} \sim N(0, \boldsymbol{I}\sigma^2_\varepsilon)$, where $\boldsymbol{I}$ was an identity matrix and σ^2_ε was the residual error variance.

The second mixed linear model obtained adjusted means for the five agronomic and FHB traits for TP17, TP18, TP19, ADV19R, and TP19_FHB using a two-stage multiple-environment trial (MET) approach for unbalanced METs with mixed experimental designs, such as between augmented and RCBDs (Mohring & Piepho, 2009). The first stage obtained adjusted means for traits within each individual location-year using the following mixed linear model:

$$y_{ij} = \mu + genotype_i + block_j + \varepsilon_{ij}$$

where y_{ij} was the observed phenotype, μ was the population mean, $genotype_i$ was the fixed effect of the i^{th} genotype, $block_j$ was the random effect of the j^{th} replication for the RCBD design or the j^{th} incomplete block for the augmented design, and ε_{ij} was the residual error term, where $\varepsilon_{ij} \sim N(0, \boldsymbol{I}\sigma^2_\varepsilon)$, where $\boldsymbol{I}$ was an identity matrix and σ^2_ε was the residual error variance.

Adjusted means from each location-year for each trait were then compiled into a single dataset for the second stage of the analysis, where adjusted means were calculated for each trait using the following mixed linear model:

$$y_{ij} = \mu + genotype_i + env_j + (genotype \times env)_{ij} + \varepsilon_{ij}$$

where y_{ij} was the observed phenotype, μ was the population mean, $genotype_i$ was the fixed effect of the i^{th} genotype, env_j was the random effect of the j^{th} location-year, $(genotype \times env)_{ij}$ was the random interaction term between genotype and location-year, and ε_{ij} was the residual error term, where $\varepsilon_{ij} \sim N(0, \boldsymbol{I}\sigma^2_\varepsilon)$, where $\boldsymbol{I}$ was an identity matrix and σ^2_ε was the residual error variance.

Phenotypic Pearson correlations were determined between GY, HD, and TW within TP17, TP18, TP19, the WO nurseries, ADV yield trials, and ARE yield trials using the multivariate function in JMP Pro 15.2.0 software (SAS Institute Inc.). Heading notes were not obtained for the WO19 nursery, therefore HD was not included in TP19. Phenotypic Pearson correlations were also determined between DON, FDK, HD, PH, and SEV within TP18_FHB and TP19_FHB as well as the ADV and ARE FHB nurseries using the multivariate function in JMP Pro 15.2.0 software (SAS Institute Inc.).

Plot mean-based broad-sense heritability (H^2) was calculated for each trait in trials evaluated using the single-stage mixed linear model approach using the following equation:

$$H^2 = \frac{\sigma^2_{genotype}}{\sigma^2_{genotype} + \frac{\sigma^2_{genotype \times env}}{n_{env}} + \frac{\sigma^2_\varepsilon}{n_{env} \times n_{rep}}}$$

where $\sigma^2_{genotype}$ was the genotypic variance, $\sigma^2_{genotype \times env}$ was the variance of the interaction between genotype and location-year, n_{env} was the number of location-years where the trait was evaluated, σ^2_ε was the residual error variance, and n_{rep} was the number of replications within each location-year. Variance components were obtained from the single-stage mixed linear model described above using the PROC MIXED procedure in SAS 9.4 (SAS Institute).

Broad-sense heritability ($\overline{H}_C^2$) for traits from trials evaluated using the two-stage mixed linear model used a formula proposed by Cullis et al. (2006) for unbalanced METs:

$$\overline{H}_C^2 = 1 - \frac{\overline{v}_{BLUP}}{2\sigma^2_{genotype}}$$

where $\bar{v}_{BLUP}$ was the mean variance between two best linear unbiased predictions (BLUPs), obtained from the second stage of the MET model, where genotype was set as a random effect, and $\sigma^2_{genotype}$ was the genotypic variance of the trait of interest, obtained from the second stage of the MET model (Cullis, Smith, & Coombes, 2006; Piepho & Mohring, 2007).

Genotyping by Sequencing

All genotypes were genotyped using genotyping by sequencing (GBS) using methods described in Larkin et al. (2020). Single nucleotide polymorphism calling was performed using the TASSEL 5.0 GBSv2 pipeline using 64 base kmer length and a minimum kmer count of five (Bradbury et al., 2007). Reads were aligned to the International Wheat Genome Sequencing Consortium (IWGSC) RefSeq v1.0 'Chinese Spring' wheat reference sequence using the alignment method of Burrows-Wheeler aligner version 0.7.10 (Appels et al., 2018; Li & Durbin, 2009).

Raw single nucleotide polymorphism (SNP) data generated from the TASSEL pipeline were filtered using PLINK software to remove taxa with more than 85% missing data and heterozygosity greater than 30%. Genotypic data were then filtered to select for biallelic SNPs with minor allelic frequency of greater than five percent, less than 20% missing data, and heterozygosity of less than or equal to 10% (Purcell et al., 2007). Missing marker data were then imputed using BEAGLE software, based on a sliding window length encompassing the entire

chromosome (Browning, Zhou, & Browning, 2018). Markers were again filtered after imputation to remove SNP markers with minor allele frequency greater than five percent and heterozygosity of less than equal to 10% using PLINK software. Markers unaffiliated with any chromosome within the wheat genome were also removed for a final genotypic dataset of 5,202 SNP markers.

Principal component analyses were performed within each of the TPs to evaluate the genetic relationships between subpopulations within each TP using the PCA function in TASSEL 5.0 (Bradbury et al., 2007). These relationships between the first three principal components were visualized for each TP using the scatterplot3d package in R v4.0.3 software (Ligges, Maechler, Schnackenberg, & Ligges, 2018; R, 2020).

Genomic Selection

Two different models were tested for all five TPs in order to obtain GEBVs for GY in the WO17, WO18, and WO19 nurseries and DON, FDK, and SEV for the ADV18 and ADV19 trials. The first model was a naïve genomic BLUP (GBLUP) model with no covariates (NGS). The second model was a MVGS GBLUP model using TW as a covariate for predicting GY, while DON was predicted using FDK and HD as covariates, FDK was predicted using DON and SEV as covariates, and SEV was predicted using FDK and PH as covariates. The optimal covariate combinations for the MVGS models were determined in Larkin et al. (2020) for the FHB traits, while the optimal covariate of TW was determined by comparing the mean prediction accuracy of GY for TP17 by using TW, HD, PH, and maturity date (MD) as covariates (Larkin et al., Chapter III).

Cross Validation

Mean prediction accuracies between the NGS and MVGS models for each TP were obtained using a five-fold cross-validation analysis performed using the Genomic Selection function in TASSEL 5.0 (Bradbury et al., 2007). The GBLUP model used for the analyses is described as follows:

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

where u is the vector of marker effects, which is assumed to have a normal distribution $u \sim N(0, \boldsymbol{G}\sigma_u^2)$, where $\boldsymbol{G}$ is the genomic relationship matrix, obtained using the Kinship function within TASSEL 5.0, and σ_u^2 is the variance of the individual marker effects; β is the vector of fixed effects; $\boldsymbol{X}$ is the design matrix of fixed effects; $\boldsymbol{Z}$ is the design matrix relating genotypes to phenotypic observations (y), with m markers in columns and n phenotypes in rows; and ε_i is the residual error at the i^{th} locus, which is assumed to have a normal distribution $\varepsilon_i \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$, where $\boldsymbol{I}$ is the identity matrix and σ^2_ε is the residual error variance. The GEBV is the sum of all allele effects of a genotype (Endelman, 2011; VanRaden, 2008).

The five-fold cross-validation approach randomly divided the TP into five equal sized groups. Four of the five groups were then used as the TP to train the GBLUP model to calculate GEBVs for the fifth group, serving as the VP, where the phenotypic values were set as missing. In the case of the MVGS models, the phenotypic data for the covariate traits were used as a fixed effect in the model. The GEBVs calculated for the VP were compared to the actual phenotypic values using a Pearson correlation. The process was repeated over 100 iterations. The mean prediction accuracies between the NGS and MVGS models were compared between all five TPs using a generalized linear mixed model (GLMM) and Fisher's LSD with an α of .05,

implemented in PROC GLIMMIX in SAS 9.4 (SAS Institute Inc.). Mean prediction accuracy comparisons between the NGS and MVGS models for each TP were visualized using the yarrr package in R v4.0.3 (Phillips, 2017; R, 2020).

Forward Prediction

Each of the five TPs were then used to obtain predictions for their respective VPs using the NGS and MVGS GBLUP models associated with each trait. For example, TP17 was used to calculate GEBVs for GY for the WO17 nursery using the NGS and MVGS models, and TP18_FHB was used to calculate GEBVs for DON, FDK, and SEV for the ADV18 trial using the NGS and MVGS models (Table 2).

Once GEBVs for each trait for each model were obtained, GEBVs were compared to the adjusted mean of the trait of interest for each genotype in the following generation using a Pearson correlation using the multivariate function in JMP 15.2.0 software (SAS Institute Inc.). For example, GEBVs for GY obtained for WO17 were compared to the adjusted mean GY for each genotype across the WO17, ADV18, and ARE19 generations using a Pearson correlation. This serves as a form of prediction accuracy for the respective model and TP. A scatterplot visualizing the comparison between GEBVs and adjusted means across years for each genotype, as well as individual genotypes advanced to the next generation, was created using the ggplot2 package in R v4.0.3 for each model for each TP (R, 2020; Wickham, Chang, & Wickham, 2016). Selection accuracy was also determined as the percentage of genotypes advanced to the ARE generation that were in the top 50% of GEBVs from the NGS or MVGS models as well as the top 50% based on phenotypic values.

Response to selection was also compared between the NGS and MVGS models and phenotypic selection, based on the adjusted means from the WO for GY or ADV generations for FHB traits, using a selection pressure of 50%. The response to selection formula is as follows:

$$R = H^2 S$$

where H^2 was the broad-sense heritability calculated as above, and S is the selection differential, calculated as $S = \mu_{Selected} - \mu_{Unselected}$ where $\mu_{Selected}$ is the mean of the phenotypic data for the top 50% of genotypes selected for genotypes in both the ADV and ARE generations using either phenotypic selection, NGS, or MVGS, and $\mu_{Unselected}$ is the mean of the full unselected population of the genotypes in both the ADV and ARE generations of the breeding cycle (Arruda, Lipka, et al., 2016; Falconer & McKay, 1996; D. N. Lozada, Ward, & Carter, 2020).

RESULTS

Variation in Agronomic and Disease Traits

Each TP, WO, ADV, and ARE trial saw significant variation in GY, TW, and HD (Table 3). Mean GY varied between trials and across years, where WO17 had the highest mean GY at 5.29 t ha^{-1} while WO19 had the lowest mean GY at 3.10 t ha^{-1}. The trial with the highest mean TW was ADV20 at 77.88 kg hl^{-1} while the trial with the lowest mean TW was WO19 at 63.88 kg hl^{-1}. Both the WO19 and DHWO19 trials had the lowest mean GY and TW numbers between 2017 and 2020. This is likely due to the fact that there was a late rainfall event that occurred during harvest, resulting in severe lodging across both trials. Heritability for GY, TW, and HD was high throughout nearly all trials, however WO19 had very low heritabilities for GY and TW, this can also be attributed to severe lodging (Table 3).

Significant positive phenotypic correlations were consistently observed between GY and TW in all trials except for TP18, this was likely due to the weaker positive correlation between

GY and TW in the WO17 trial (Table 3). Notably strong positive correlations between GY and HD were observed in the WO18 and ADV18 trials during the 2017-2018 growing season due to a late freeze event that occurred when earlier heading genotypes were flowering, resulting in lower GY for early genotypes. Typically, earlier heading genotypes have higher GY than later heading genotypes, which was observed during more normal years where there were negative or weak positive correlations between GY and HD (Table 3).

Both FHB TPs as well as the ADV and ARE FHB trials had significant variation for all three FHB traits as well as HD and PH. The ADV18 FHB trial had the highest mean DON and FDK, but it also had the lowest mean SEV. The ARE20 FHB trial had the lowest mean DON and FDK, likely due to stronger genetic resistance (Table 4). All trials also had significant correlations between the three FHB traits. Correlations between DON and HD were consistently positive, however the correlations were not significant with smaller population sizes, while DON was significantly correlated with PH only in ADV19. There were generally negative correlations between FDK and PH with the exception of ADV19, however the significance of the correlations between FDK and PH were not significant with smaller population sizes. There were strong negative correlations between SEV and HD and PH for nearly all trials, however they were not significant for smaller populations. High heritability was also observed for all three FHB traits in addition to HD and PH (Table 4).

Population Structure

Genotyping by sequencing identified 5,202 SNPs across the entire wheat genome after filtering and imputation. The number of SNP markers were unevenly distributed between genomes, where the B genome had the largest number of markers (2,315), followed by the A (2,210) and D (677) genomes, which was consistent with other studies using GBS SNPs (Arruda, Brown, et al., 2016;

Larkin et al., 2020). The chromosome with the largest number of SNPs was 3B at 477, while the chromosome with the smallest number was 4D (38). The proportion of heterozygosity within the dataset was 2.5% and the average minor allele frequency was 21.6%.

The PCA of the initial TP17 or TP18_FHB population showed two primary clusters within the population. Genotypes from all breeding programs appeared in both clusters, although there was evidence of sub-clustering by breeding program within the two main clusters. This clustering has also been observed in other studies using SRWW populations adapted to the Southeastern US and it is hypothesized that this is due to the presence or absence of a stem rust (*Puccinia graminis* f. sp. *tritici*) and powdery mildew (*Blumeria graminis* f. sp. *tritici*) resistance genes, *Sr36/Pm6*, located on a translocation from *Triticum timopheevii* Zhuk. (Benson, Brown-Guedira, Murphy, & Sneller, 2012; Larkin et al., 2020; Nyquist, 1962; Sarinelli et al., 2019). The population structure was generally low, where the first three principal components only accounted for 5.23%, 3.99%, and 3.42% of the total genetic variation (Figure 1).

When evaluating the PCAs of the TP17 and TP18 populations with their respective VPs, there did not appear to be a strong deviation between TP and VP (Appendices 4a and 4b). However, there was a deviation between TP19 and WO19 and DHWO19, where a majority of genotypes in the WO19 and DHWO19 contained the FHB resistance alleles for the gene *Fhb1*, whereas a majority of the genotypes in TP19 did not (Rawat et al., 2016) (Figure 2). There was also no noticeable deviation between the TP18_FHB population and ADV18 and the TP19_FHB population and ADV19 (Appendices 4c and 4d).

Cross Validation

Between all three TPs for GY, the MVGS models using TW as a covariate had significantly higher prediction accuracies compared to the NGS models (Figure 3). This reflected the strong

and significant phenotypic correlations between GY and TW within TP17 and TP19. The primary exception to this was TP18, where there was not a significant correlation between GY and TW (r = .05) (Table 3). As a result, the MVGS model had a smaller, albeit significant, increase in prediction accuracy compared to the NGS model. As new genotypes were added to the TP for each year, the prediction accuracy of the TP also increased (Figure 3). The TP19 population had the highest mean prediction accuracy for both NGS (r = .64) and MVGS (r = .69) models. This was followed by TP18, which had the second highest prediction accuracy between NGS models (r = .59) and it was tied for second among the MVGS models (r = .60). The population with the population with the lowest mean prediction accuracy among NGS models was TP17 (r = .49), whereas the mean prediction accuracy of the MVGS model was equal to the MVGS model from TP18 (r = .60) (Figure 3).

This trend was also observed for both FHB TPs, where MVGS models had significantly higher prediction accuracies compared to NGS models for DON, FDK, and SEV (Figure 4). Contrary to the TPs for GY, prediction accuracies between TP18_FHB and TP19_FHB decreased for all three traits and both models. This is likely a result of background population structure within TP19_FHB between genotypes from the TP18_FHB population, which does not contain genotypes with *Fhb1*, and ADV18 which does contain genotypes with *Fhb1* (Appendix 4c). The trait with the highest mean prediction accuracies among the NGS models for both TPs was DON, the mean prediction accuracy for TP18_FHB was .61 and .32 for TP19_FHB. The trait with the second highest mean prediction accuracy among the NGS models was SEV, where TP18_FHB was .54 and TP19_FHB was .17. Fusarium damaged kernels had the lowest mean prediction accuracy among the NGS models where TP18_FHB (r = .45) was more accurate than TP19_FHB (r = .14). The ranking of traits between the MVGS models was not consistent with

the NGS models or between TPs. Severity had the highest prediction accuracy in TP18_FHB (r = .76), followed by FDK (r = .74) and DON (r = .72). With TP19_FHB, DON also had the MVGS model with the lowest mean prediction accuracy (r = .56), while SEV still had the second highest mean prediction accuracy (r = .68), like the NGS model. However, FDK had the highest mean prediction accuracy (r = .68) (Figure 4).

Forward Prediction

When TP17 was used to calculate GEBVs of GY for WO17, there was a significant correlation between the GEBVs for GY from the NGS model and with the actual GY data from WO17 (r = .38, p < .001), however there was a smaller yet significant correlation between the GEBVs for GY from the MVGS model and actual GY data from WO17 (r = .26, p = .002) (Table 5). The strength of the Pearson correlations for both models decreased when GEBVs were compared with the adjusted mean GY across ADV18 and ARE19, where the correlations for all three methods were no longer significant and the MVGS model was the only method that remained positive (r = .14) (Table 5). However, when comparing the accuracy of selecting genotypes advanced to ARE19 that were above average for both GEBVs and adjusted mean GY, both models correctly selected 36.4% of the genotypes in the ARE19, but phenotypic selection based on GY from WO17 had a 50.0% selection accuracy (Table 5, Figures 5a and 5b). In terms of response to selection (R), when the top 50% of genotypes were selected from ARE19 based on phenotypic GY results from WO17, GEBVs from NGS, and GEBVs from MVGS, the genotypes selected based on phenotypic selection had the lowest response in GY (R = -31.71 kg ha^{-1}), followed by NGS (R = 26.13 kg ha^{-1}) and MVGS (R = 61.99 kg ha^{-1}) (Table 5).

Unlike WO17, there was a stronger correlation between the GY data from WO18 and the GEBVs for GY from the MVGS model (r = .35, p < .001) and there was not a significant

correlation between WO18 and the NGS GEBVs (r = .12, p = .131) (Table 5). Correlations were weaker when GEBVs for both models were compared to adjusted means for GY across ADV19, ADV19R, and ARE20, where the correlation with MVGS decreased to r = .11 and NGS decreased to r = .11 (Table 5). The MVGS model had the greatest selection accuracy at 69.6%, based on selecting genotypes that were advanced to ARE20 based on above average GEBVs and above average mean GY across generations, while the NGS model had a selection accuracy of 60.8% and phenotypic selection based on GY from WO18 was the lowest at 52.2% (Table 5, Figures 5c and 5d). Response to selection based on phenotypic selection with GY from WO18 was again the weakest with R = -14.33 kg ha^{-1} followed by the MVGS model (R = -1.49 kg ha^{-1}) and NGS model (R = 4.71 kg ha^{-1}) (Table 5).

The TP19 GEBVs for GY from the NGS (r = .10, p = .037) and MVGS (r = .29, p < .001) models were significantly correlated with the WO19 GY adjusted means (Table 5). The genotypes from WO19 had only been grown in the ADV20 trial, therefore GEBVs could only be compared with the adjusted means from ADV20. Correlations between the GEBVs from the NGS and MVGS models and adjusted means for GY from ADV20 were both negative and not significant (Table 5). Much like TP18, the MVGS model had a better selection accuracy, correctly selecting 71.2% of genotypes advanced to ADV20 based on above average GEBVs and adjusted means for GY, while the NGS model correctly selected 55.5%. The MVGS model also had a higher selection accuracy than phenotypic selection (67.8%) (Table 5, Figures 5e and 5f). The NGS (R = .06 kg ha^{-1}) model had a greater response to selection than the MVGS (R = -4.98 kg ha^{-1}) model, however both were far less than phenotypic selection based on GY from WO19 (R = 11.80 kg ha^{-1}) (Table 5).

When TP18_FHB was used to predict DON, FDK, and SEV for ADV18, there were significant correlations between the NGS and MVGS models for all FHB resistance traits. The strength of both correlations decreased for all methods when compared with phenotypic data from ARE19, with the exception for the MVGS model for SEV, where the correlation increased to r =.60 compared with r = .57 (Table 6). Both NGS and MVGS models had higher selection accuracies compared to phenotypic selection from ADV18 DON data (52.9%), where the NGS model correctly selected 82.4% of genotypes in ARE19, while the MVGS model correctly selected 70.6% (Table 6, Figures 6a and 6b). The NGS (R = -.37 μg g^{-1}) model had the highest response to selection for DON compared to the NGS model (R = -.23 μg g^{-1}) and phenotypic selection (R = .20 μg g^{-1}) (Table 6).

When predicting FDK for ADV18, the MVGS model had the strongest correlations with the ADV18 FDK data as well as the FDK adjusted means from ARE19. The NGS (R = -4.09%) model again had the highest response to selection than the MVGS (R = -2.83%) model and phenotypic selection (R = -1.59%) for FDK (Table 6). The MVGS and NGS models had the same selection accuracy for FDK (70.6%) where both models outperformed phenotypic selection based on adjusted means for FDK from ADV18 (58.8%) (Table 6, Figures 6c and 6d).

The MVGS model had stronger correlations between GEBVs for SEV and adjusted means for SEV from ADV18 and ARE19 than the NGS model (Table 6). The MVGS model also had the strongest response to selection (R = -2.29%) and selection accuracy (47.1%) compared with the NGS model, where R = -.82% and selection accuracy was 41.2%. The NGS model underperformed phenotypic selection for both response to selection (R = -1.49%) and selection accuracy (52.9%), with the MVGS model only underperformed phenotypic selection for selection accuracy (Table 6, Figures 6e and 6f).

When using TP19_FHB to predict FHB resistance traits for ADV19, the correlations between GEBVs from the MVGS models and phenotypic results from AVD19 were stronger than TP18_FHB for all three traits. Correlations between GEBVs from the MVGS models were stronger than TP18_FHB when compared with adjusted means from ARE20 for DON and FDK (Table 6). Response to selection for TP19_FHB was different from TP18_FHB in that phenotypic selection outperformed the GS models for DON and SEV, whereas the MVGS model had a stronger response to selection than the NGS model and phenotypic selection for FDK (Table 6). Selection accuracies did change between TPs, as the MVGS model (69.6%) outperformed both phenotypic selection (13.0%) and the NGS model (56.5%) for DON for TP19_FHB (Table 6, Figures 7a and 7b). Unlike the results for TP18_FHB, both GS models had far lower selection accuracies than phenotypic selection (91.3%), although the MVGS model (60.9%) was better than the NGS model (34.8%) (Table 6; Figures 7c and 7d). Selection accuracy for SEV also changed, where the MVGS model had the same selection accuracy as phenotypic selection (82.6%) while also outperforming the NGS model (60.9%) (Table 6, Figures 7e and 7f).

DISCUSSION

Genomic selection is a valuable tool for plant breeders, and many studies have shown the vast realm of possibilities for its application (Heffner et al., 2009; Larkin et al., 2019; Sorrells, 2015). The primary goal for GS is to increase genetic gain for a trait of interest within a breeding program through the reduction of time within a breeding cycle and by improving selection accuracy (Asoro et al., 2013; Bernardo & Yu, 2007; Heffner et al., 2009; Rutkoski et al., 2015; Schaeffer, 2006). While most research in GS has focused on improving the accuracy of TPs, less have focused on the implementation of GS into breeding programs in the form of forward

selection (Bernardo, 2016). In our study, we chose to focus on forward prediction through the use of NGS and MVGS models and compared their performance, based on selection accuracy and response to selection, to phenotypic selection for economically important traits, such as GY and FHB resistance traits.

Prediction Accuracy of Training Populations

In our study, we saw that MVGS models consistently had significantly higher prediction accuracies for GY, DON, FDK, and SEV in every TP compared to NGS. These results were consistent with previous studies involving MVGS for GY and FHB resistance traits (Crain et al., 2018; Larkin et al., 2020; Dennis N. Lozada & Carter, 2019; Moreno-Amores et al., 2020; Rutkoski et al., 2016; Schulthess et al., 2018; Sun et al., 2017). While previous studies predicting GY using MVGS have primarily focused on high-throughput phenotyping traits as covariates, we used TW as a covariate due to its strong correlation with GY (Crain et al., 2018; Dennis N. Lozada & Carter, 2019; Rutkoski et al., 2016; Sun et al., 2017). This follows the general trend for MVGS, where covariate traits sharing a strong correlation with a trait of interest can improve prediction accuracies for said trait of interest (Calus & Veerkamp, 2011; Y. Jia & Jannink, 2012; Dennis N. Lozada & Carter, 2019; Schulthess et al., 2016; Ward et al., 2019).

We also updated our TPs for each generation by adding phenotypic data for genotypes from the previous generation into the following year's TP. Other studies have found that updating TPs helped to prevent the deviation in genetic relationships between the TP and VP as new germplasm was added and advanced through the breeding program (Clark, Hickey, Daetwyler, & van der Werf, 2012; Lorenz & Smith, 2015; Lorenz, Smith, & Jannink, 2012; Meuwissen, 2009; Neyhart, Tiede, Lorenz, & Smith, 2017). Studies have also shown that larger TP sizes can have higher prediction accuracies as well, particularly when working with more

diverse populations where new germplasm is continually added to the breeding program (Heffner, Jannink, Iwata, Souza, & Sorrells, 2011; Heslot et al., 2012; Isidro et al., 2015; Mujibi et al., 2011; Norman, Taylor, Edwards, & Kuchel, 2018; Poland et al., 2012).

We found that prediction accuracy increased alongside TP size, where TP17 was the smallest with 355 genotypes while also having the lowest prediction accuracy for GY, followed by TP18 with 495 genotypes, and finally TP19 with 655 genotypes and also the highest prediction accuracy. This trend was also observed for GY in wheat where prediction accuracy increased from $r = .72$ to $r = .85$ as the number of genotypes in the TP increased from 250 to 2,000 (Norman et al., 2018). Another study found that increasing the TP size from 50 to 350 genotypes increased prediction accuracy for GY in SRWW from $r = .40$ to $r = .64$ (Sarinelli et al., 2019). It is difficult to say if the increased TP size reduced the impact of genetic variation between the initial genotypes and genotypes that were newly added to each TP, as there was little genetic variation within the TPs based on their respective PCAs.

We did not observe this trend between TP18_FHB and TP19_FHB, where prediction accuracy actually decreased for all three FHB resistance traits when additional genotypes were added from ADV18. This can likely be attributed to less variation in DON, FDK, and SEV within ADV18 as well as lower heritabilities for each trait. Genotypes within ADV18 also had the FHB resistance alleles for *Fhb1*, which could have increased background population structure within TP19_FHB.

Forward Prediction

Much like the results from the cross-validation analyses of the TPs, the MVGS models had stronger correlations between their calculated GEBVs and phenotypic results from their respective VPs for GY and FHB resistance traits, aligning with other studies involving MVGS

models (Y. Jia & Jannink, 2012; Larkin et al., 2020; Dennis N. Lozada & Carter, 2019; Schulthess et al., 2016; Ward et al., 2019). The only instance where this did not occur was for TP17 when it was used to predict GY for WO17. While there was a strong correlation between GY and TW in TP17, there was a very weak correlation between GY and TW in WO17, therefore it is reasonable to suspect that the weak correlation between the two traits in the VP negated the benefit of using a MVGS model, as other studies have noted the lack of benefit of using covariates that were weakly correlated with a trait of interest (Calus & Veerkamp, 2011; Y. Jia & Jannink, 2012). Inversely, while there was little benefit to using MVGS for cross-validation in TP18 since there was a weak correlation between GY and TW, there was a strong correlation between GY and TW in the WO18 VP, therefore the MVGS model performed better than the NGS model for forward prediction even when there was a weak correlation in the TP.

An additional caveat for WO18 was that there was a late freeze event during the spring of 2018 that killed most early-heading plots during flowering, resulting in an abnormally positive correlation between HD and GY, as HD and GY are typically negatively correlated (Whittal, Kaviani, Graf, Humphreys, & Navabi, 2018; Worland et al., 1998). Therefore, we used HD as a covariate in an MVGS model for forward prediction. However, the strong negative correlation between GY and HD in TP18 masked the effect of the strong positive correlation in WO18, resulting in a weak correlation between the GEBVs from the MVGS model and the phenotypic data for GY compared to the MVGS model using TW as a covariate.

Our range in prediction accuracy for NGS models was between $r = .09$ and $r = .29$ while the range of our MVGS models were between $r = .01$ and $r = .45$. In an evaluation of forward prediction in the Kansas State University wheat breeding program, the highest prediction accuracy between the GEBVs for GY in the preliminary yield trials (PYTs) and the actual

phenotypic results for GY was r = -.16 (Calvert et al., 2020). The same study also used high-throughput phenotyping traits as covariates in a MVGS model for forward prediction of GY in wheat, however the prediction accuracy was unfavorable unless a large TP was used (Calvert et al., 2020). This contrasts with our results where the use of another agronomic trait such as GY significantly improved prediction accuracy for every training population but TP17. It is notable that TP17 was the smallest TP in this case, but the lack of effectiveness in the MVGS model could be attributed more to the weak correlation between GY and TW in the VP than to TP size. In another forward prediction study for GY at the University of Nebraska, there was a range in prediction accuracy between r = .22 and r = .26 (Belamkar et al., 2018). While our best performing NGS model was better than their top performing model, we still had a wider range in prediction accuracies and there was certainly a weaker prediction accuracy when using the NGS model for TP19.

Response to selection was measured as the top 50% of breeding lines in the ARE generation, selected based on GEBVs and adjusted means of GY for the WO population. Other studies have shown that GS could not have as high of a response to selection as phenotypic selection; however, our method of excluding phenotypic data from the WO genotypes from the selection dataset allowed for greater independence from bias towards the phenotypic selection method (D. N. Lozada, Mason, Sarinelli, & Brown-Guedira, 2019). In our study, our best performing model compared to phenotypic selection was the MVGS model with TP17, where the MVGS model had the highest genetic gain compared to phenotypic selection and the NGS model. Our responses to selection using NGS and MVGS models for GY were also higher than another study that used NGS models and genome-wide association study (GWAS) assisted GS models (D. N. Lozada et al., 2020). It was notable that when there was a stronger correlation

between the GEBVS or WO phenotypic data and the phenotypic data for the ADV and ARE generations there was also a stronger gain in response to selection. As of yet, there have been no extensive forward prediction studies for FHB resistance traits in wheat. Relative to studies related to GY, we found that the percent change relative to phenotypic selection from the MVGS models for all three FHB resistance traits in both years, were higher than those for GY, where the best performing model was the NGS model for FDK, using TP18_FHB. It is also notable that phenotypic selection did not significantly outperform the MVGS model across years or traits, indicating that MVGS could be a good supplement, if not substitute, for phenotypic selection particularly during years when it is difficult to phenotype.

In terms of selection accuracy, our MVGS models outperformed phenotypic selection based on WO GY data for two out of our three TPs. However, when comparing GS models with phenotypic selection for FHB resistance traits, the NGS and MVGS models had higher selection accuracies for DON using TP18_FHB, and the MVGS model was equal to phenotypic selection using TP19_FHB. Both the MVGS and NGS models were equally more accurate than phenotypic selection for FDK with TP18_FHB. Additionally, the MVGS model was equal in performance with phenotypic selection for SEV in TP19_FHB. It has been mentioned that prediction accuracy does not necessarily correlate with selection accuracy for forward prediction (Belamkar et al., 2018). In our case, phenotypic selection always had the greatest prediction accuracy by default. However, when it came to accurately selecting genotypes that were advanced to the elite generation, GS outperformed, or was equal to, phenotypic selection 67% of the time.

CONCLUSIONS

This study showed that both NGS and MVGS could be successfully implemented into a SRWW breeding program, while using other agronomic and disease traits as covariates with reasonable accuracy compared to phenotypic selection while again showing its value as a tool for plant breeders. We also found that MVGS models performed significantly better than NGS models in terms of both cross-validation within TPs as well as forward prediction of untested genotypes for economically important traits, such as GY or FHB resistance traits. This was particularly evident when there was a strong correlation between the trait of interest and the covariate trait. This is one of the first studies to show that MVGS could be effectively implemented for forward prediction within a wheat breeding program. This is also the first study to extensively investigate the use of forward prediction when breeding for FHB resistance in wheat. We found that GS could serve as a suitable, albeit imperfect, alternative to phenotypic selection when implemented during years where environmental conditions prohibit accurate phenotypic selection, particularly when experiencing late freezing events or extensive lodging.

Prior to implementing GS into their own breeding programs, breeders must consider the genetic relationships between their prospective TPs and the breeding lines they hope to use as their VP. In the case of MVGS, breeders must also consider the correlations between their traits of interest and secondary traits used as covariates, as these correlations can differ between the TP and VP. For example, there could be a strong correlation between GY and TW in the TP but there could be a weak correlation between the two traits in the VP, therefore the MVGS model might not be more accurate than a NGS model. Inversely, there could be a strong correlation between traits in the VP while there is a weak correlation between traits in the TP, therefore MVGS could be more accurate than expected when forward prediction is implemented.

REFERENCES

Akdemir, D., Sanchez, J. I., & Jannink, J. L. (2015). Optimization of genomic selection training populations with a genetic algorithm. *Genetics Selection Evolution, 47*, 10. doi:10.1186/s12711-015-0116-6

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., . . . Manuscript Writing, T. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science, 361*(6403), 661-+. doi:10.1126/science.aar7191

Arruda, M. P., Brown, P., Brown-Guedira, G., Krill, A. M., Thurber, C., Merrill, K. R., . . . Kolb, F. L. (2016). Genome-wide association mapping of Fusarium head blight resistance in wheat using genotyping-by-sequencing. *Plant Genome, 9*(1), 14. doi:10.3835/plantgenome2015.04.0028

Arruda, M. P., Lipka, A. E., Brown, P. J., Krill, A. M., Thurber, C., Brown-Guedira, G., . . . Kolb, F. L. (2016). Comparing genomic selection and marker-assisted selection for Fusarium head blight resistance in wheat (*Triticum aestivum* L.). *Molecular Breeding, 36*(7), 11. doi:10.1007/s11032-016-0508-5

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. doi:10.2135/cropsci2012.09.0526

Belamkar, V., Guttieri, M. J., Hussain, W., Jarquin, D., El-basyoni, I., Poland, J., . . . Baenziger, P. S. (2018). Genomic Selection in Preliminary Yield Trials in a Winter Wheat Breeding Program. *G3-Genes Genomes Genetics, 8*(8), 2735-2747. doi:10.1534/g3.118.200415

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. doi:10.3835/plantgenome2011.11.0027

Bernardo, R. (2016). Bandwagons I, too, have known. *Theoretical and Applied Genetics, 129*(12), 2323-2332. doi:10.1007/s00122-016-2772-5

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

Boyles, R. E., Marshall, D. S., & Bockelman, H. E. (2019). Yield Data from the Uniform Southern Soft Red Winter Wheat Nursery Emphasize Importance of Selection Location and Environment for Cultivar Development. *Crop Science, 59*(5), 1887-1898. doi:10.2135/cropsci2018.11.0685

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics, 23*(19), 2633-2635. doi:10.1093/bioinformatics/btm308

Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A One-Penny Imputed Genome from Next-Generation Reference Panels. *American Journal of Human Genetics, 103*(3), 338-348. doi:10.1016/j.ajhg.2018.07.015

Calus, M. P. L., & Veerkamp, R. F. (2011). Accuracy of multi-trait genomic selection using different methods. *Genetics Selection Evolution, 43*, 14. doi:10.1186/1297-9686-43-26

Calvert, M., Evers, B., Wang, X., Fritz, A., & Poland, J. (2020). Breeding Program Optimization for Genomic Selection in Winter Wheat. *bioRxiv*, 2020.2010.2007.330415. doi:10.1101/2020.10.07.330415

Clark, S. A., Hickey, J. M., Daetwyler, H. D., & van der Werf, J. H. J. (2012). The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genetics Selection Evolution, 44*, 9. doi:10.1186/1297-9686-44-4

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. doi:10.3835/plantgenome2012.11.0030

Covarrubias-Pazaran, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., . . . 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*, 13. doi: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), 14. doi:10.3835/plantgenome2017.05.0043

Cullis, B. R., Smith, A. B., & Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural Biological and Environmental Statistics, 11*(4), 381-393. doi:10.1198/108571106x154443

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

Falconer, D. S., & McKay, T. F. C. (1996). Introduction to quantitative genetics (4th edn): by Douglas S. Falconer and Trudy FC Mackay Longman, 1996.£ 24.99 pbk (xv and 464 pages) ISBN 0582 24302 5. In: Elsevier Current Trends.

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*, 7. doi: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. doi:10.1534/genetics.107.081190

Harrison, S. A., Babar, M. A., Barnett, R. D., Blount, A. R., Johnson, J. W., Mergoum, M., . . . Ibrahim, A. M. H. (2017). 'LA05006', a Dual-Purpose Oat for Louisiana and Other Southeastern Regions of the USA. *Journal of Plant Registrations, 11*(2), 89-94. doi:10.3198/jpr2016.08.0040crc

Heffner, E. L., Jannink, J. L., Iwata, H., Souza, E., & Sorrells, M. E. (2011). Genomic Selection Accuracy for Grain Quality Traits in Biparental Wheat Populations. *Crop Science, 51*(6), 2597-2606. doi:10.2135/cropsci2011.05.0253

Heffner, E. L., Sorrells, M. E., & Jannink, J. L. (2009). Genomic selection for crop improvement. *Crop Science, 49*(1), 1-12. doi: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. doi: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. doi: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. doi:10.1093/bfgp/elq001

Jia, H. Y., Zhou, J. Y., Xue, S. L., Li, G. Q., Yan, H. S., Ran, C. F., . . . Ma, Z. Q. (2018). A journey to understand wheat Fusarium head blight resistance in the Chinese wheat landrace Wangshuibai. *Crop Journal, 6*(1), 48-59. doi:10.1016/j.cj.2017.09.006

Jia, Y., & Jannink, J. L. (2012). Multiple-Trait Genomic Selection Methods Increase Genetic Value Prediction Accuracy. *Genetics, 192*(4), 1513-+. doi:10.1534/genetics.112.144246

Johnson, J. W., Chen, Z., Buck, J. W., Buntin, G. D., Babar, M. A., Mason, R. E., . . . Mergoum, M. (2017). 'GA 03564-12E6': A High-Yielding Soft Red Winter Wheat Cultivar Adapted to Georgia and the Southeastern Regions of the United States. *Journal of Plant Registrations, 11*(2), 159-164. doi:10.3198/jpr2016.07.0036crc

Kelley, J. (2018). 2018 Arkansas Wheat Quick Facts. In: 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., . . . Dong, Y. (2020). Genome-wide analysis and prediction of fusarium head blight resistance in soft red winter wheat. *Crop Science, n/a*(n/a). doi: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-Basel, 9*(9), 18. doi:10.3390/agronomy9090479

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

Ligges, U., Maechler, M., Schnackenberg, S., & Ligges, M. U. (2018). Package 'scatterplot3d'. *Recuperado de https://cran. rproject. org/web/packages/scatterplot3d/scatterplot3d. pdf*.

Lorenz, A. J., & Smith, K. P. (2015). Adding Genetically Distant Individuals to Training Populations Reduces Genomic Prediction Accuracy in Barley. *Crop Science, 55*(6), 2657-2667. doi:10.2135/cropsci2014.12.0827

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. doi:10.2135/cropsci2011.09.0503

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. doi:10.20900/cbgg20190012

Lozada, D. N., Mason, R. E., Sarinelli, J. M., & Brown-Guedira, G. (2019). Accuracy of genomic selection for grain yield and agronomic traits in soft red winter wheat. *Bmc Genetics, 20*(1), 12. doi:10.1186/s12863-019-0785-1

Lozada, D. N., Ward, B. P., & Carter, A. H. (2020). Gains through selection for grain yield in a winter wheat breeding program. *Plos One, 15*(4), 18. doi:10.1371/journal.pone.0221603

Mason, R. E., Johnson, J. W., Mergoum, M., Miller, R. G., Moon, D. E., Carlin, J. F., . . . Blount, A. R. (2018). 'AR11LE24, a Soft Red Winter Wheat Adapted to the Mid-South Region of the USA. *Journal of Plant Registrations, 12*(3), 357-361. doi:10.3198/jpr2017.09.0060crc

Massman, J. M., Jung, H. J. G., & Bernardo, R. (2013). Genomewide Selection versus Marker-assisted Recurrent Selection to Improve Grain Yield and Stover-quality Traits for Cellulosic Ethanol in Maize. *Crop Science, 53*(1), 58-66. doi:10.2135/cropsci2012.02.0112

Meuwissen, T. H. E. (2009). Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. *Genetics Selection Evolution, 41*, 9. doi:10.1186/1297-9686-41-35

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.

Michel, S., Ametz, C., Gungor, H., Akgol, B., Epure, D., Grausgruber, H., . . . Buerstmayr, H. (2017). Genomic assisted selection for enhancing line breeding: merging genomic and phenotypic selection in winter wheat breeding programs with preliminary yield trials. *Theoretical and Applied Genetics, 130*(2), 363-376. doi:10.1007/s00122-016-2818-8

Mohring, J., & Piepho, H. P. (2009). Comparison of Weighting Methods in Two-Stage Analysis of Plant Breeding Trials. *Crop Science, 49*(6), 1977-1988. doi:10.2135/cropsci2009.02.0083

Moreno-Amores, J., Michel, S., Miedaner, T., Longin, C. F. H., & Buerstmayr, H. (2020). Genomic predictions for Fusarium head blight resistance in a diverse durum wheat panel: an effective incorporation of plant height and heading date as covariates. *Euphytica, 216*(2), 19. doi:10.1007/s10681-019-2551-x

Mujibi, F. D. N., Nkrumah, J. D., Durunna, O. N., Stothard, P., Mah, J., Wang, Z., . . . Moore, S. S. (2011). Accuracy of genomic breeding values for residual feed intake in crossbred beef cattle. *Journal of Animal Science, 89*(11), 3353-3361. doi:10.2527/jas.2010-3361

Neyhart, J. L., Tiede, T., Lorenz, A. J., & Smith, K. P. (2017). Evaluating Methods of Updating Training Data in Long-Term Genomewide Selection. *G3-Genes Genomes Genetics, 7*(5), 1499-1510. doi:10.1534/g3.117.040550

Norman, A., Taylor, J., Edwards, J., & Kuchel, H. (2018). Optimising Genomic Selection in Wheat: Effect of Marker Density, Population Size and Population Structure on Prediction Accuracy. *G3-Genes Genomes Genetics, 8*(9), 2889-2899. doi:10.1534/g3.118.200311

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.

Phillips, N. D. (2017). Yarrr! The pirate's guide to R. *APS Observer, 30*(3).

Piepho, H. P., & Mohring, J. (2007). Computing heritability and selection response from unbalanced plant breeding trials. *Genetics, 177*(3), 1881-1888. doi:10.1534/genetics.107.074229

Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S. Y., Manes, Y., . . . Jannink, J. L. (2012). Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *Plant Genome, 5*(3), 103-113. doi:10.3835/plantgenome2012.06.0006

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., . . . Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics, 81*(3), 559-575. doi:10.1086/519795

R, C. T. (2020). R: A language and environment for statistical computing. In. R Foundation for Statistical Computing, Vienna, Austria.

Rawat, N., Pumphrey, M. O., Liu, S. X., Zhang, X. F., Tiwari, V. K., Ando, K., . . . Gill, B. S. (2016). Wheat Fhb1 encodes a chimeric lectin with agglutinin domains and a pore-forming toxin-like domain conferring resistance to Fusarium head blight. *Nature Genetics, 48*(12), 1576-1580. doi:10.1038/ng.3706

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), 8. doi:10.1371/journal.pone.0066428

Rutkoski, J., Benson, J., Jia, Y., Brown-Guedira, G., Jannink, J. L., & Sorrells, M. (2012). Evaluation of genomic prediction methods for Fusarium head blight resistance in wheat. *Plant Genome, 5*(2), 51-61. doi:10.3835/plantgenome2012.02.0001

Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Perez, L. G., Crossa, J., . . . 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. doi: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. doi:10.3835/plantgenome2014.10.0074

Safety, F. O. o. F., & Compliance, F. O. o. S. a. (2010). *Guidance for Industry and FDA: Advisory Levels for Deoxynivalenol (DON) in Finished Wheat Products for Human Consumption and Grains and Grain By-Products used for Animal Feed*. Retrieved from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-industry-and-fda-advisory-levels-deoxynivalenol-don-finished-wheat-products-human

Sarinelli, J. M., Murphy, J. P., Tyagi, P., Holland, J. B., Johnson, J. W., Mergoum, M., . . . 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. doi:10.1007/s00122-019-03276-6

Schaeffer, L. R. (2006). Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics, 123*(4), 218-223. doi:10.1111/j.1439-0388.2006.00595.x

Schulthess, A. W., Wang, Y., Miedaner, T., Wilde, P., Reif, J. C., & Zhao, Y. S. (2016). Multiple-trait- and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes. *Theoretical and Applied Genetics, 129*(2), 273-287. doi:10.1007/s00122-015-2626-6

Schulthess, A. W., Zhao, Y. S., Longin, C. F. H., & Reif, J. C. (2018). Advantages and limitations of multiple-trait genomic prediction for Fusarium head blight severity in hybrid wheat (Triticum aestivum L.). *Theoretical and Applied Genetics, 131*(3), 685-701. doi:10.1007/s00122-017-3029-7

Sobrova, P., Adam, V., Vasatkova, A., Beklova, M., Zeman, L., & Kizek, R. (2010). Deoxynivalenol and its toxicity. *Interdisciplinary Toxicology, 3*(3), 94-99. doi:10.2478/v10102-010-0019-x

Sorrells, M. E. (2015). *Genomic Selection in Plants: Empirical Results and Implications for Wheat Breeding*, Tokyo.

Steiner, B., Michel, S., Maccaferri, M., Lemmens, M., Tuberosa, R., & Buerstmayr, H. (2019). Exploring and exploiting the genetic variation of Fusarium head blight resistance for genomic-assisted breeding in the elite durum wheat gene pool. *Theoretical and Applied Genetics, 132*(4), 969-988. doi:10.1007/s00122-018-3253-9

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. doi:10.3835/plantgenome2016.11.0111

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

Ward, B. P., Brown-Guedira, G., Tyagi, P., Kolb, F. L., Van Sanford, D. A., Sneller, C. H., & Griffey, C. A. (2019). Multienvironment and Multitrait Genomic Selection Models in Unbalanced Early-Generation Wheat Yield Trials. *Crop Science, 59*(2), 491-507. doi:10.2135/cropsci2018.03.0189

Whittal, A., Kaviani, M., Graf, R., Humphreys, G., & Navabi, A. (2018). Allelic variation of vernalization and photoperiod response genes in a diverse set of North American high latitude winter wheat genotypes. *Plos One, 13*(8), 17. doi:10.1371/journal.pone.0203068

Wickham, H., Chang, W., & Wickham, M. H. (2016). Package 'ggplot2'. *Create Elegant Data Visualisations Using the Grammar of Graphics. Version, 2*(1), 1-189.

Wilson, W., McKee, G., Nganje, W., Dahl, B., & Bangsund, D. (2017). Economic Impact of USWBSI's Scab Initiative to Reduce FHB. *Agribusiness and Applied Economics, 774*. doi:10.22004/ag.econ.264672

Worland, A. J., Borner, A., Korzun, V., Li, W. M., Petrovic, S., & Sayers, E. J. (1998). The influence of photoperiod genes on the adaptability of European winter wheats (Reprinted from Wheat: Prospects for global improvement, 1998). *Euphytica, 100*(1-3), 385-394. doi:10.1023/a:1018327700985

Xu, Y. B., & Crouch, J. H. (2008). Marker-assisted selection in plant breeding: From publications to practice. *Crop Science, 48*(2), 391-407. doi:10.2135/cropsci2007.04.0191

Table 1. Description of the number of genotypes, composition, and experimental design of three generations of F4:6 observation nurseries (WO), F4:7 advanced nurseries (ADV), and F4:8 elite nurseries (ARE), as well as the initial training population (TP17/TP18_FHB) used to predict grain yield and three Fusarium head blight resistance traits.

Trial[a]	Generation[b]	Conventional Lines	DH Lines	Total	Location(s)	Rep(s)	Design[c]
TP17	-	355	-	355	6	1-2	AUG/RCBD
TP18_FHB	-	355	-	355	9	2	RCBD
WO17	F4:6/DH	95	45	140	1	1	AUG
WO18	F4:6/DH	73	87	160	1	1	AUG
WO19	F4:6/DH	76	357	433	1	1	AUG
ADV18	F4:7/DH	64	40	104	4	2	RCBD
ADV18_FHB	F4:7/DH	64	40	104	2	2	RCBD
ADV19	F4:7/DH	50	70	120	2	2	RCBD
ADV19_FHB	F4:7/DH	50	70	120	2	2	RCBD
ADV19R	F4:7/DH	50	70	120	3	2	RCBD
ADV20	F4:7/DH	56	90	146	3	2	RCBD
ARE19	F4:8/DH	16	6	22	2	3-4	RCBD
ARE19_FHB	F4:8/DH	16	6	22	2	2	RCBD
ARE20	F4:8/DH	12	11	23	4	3-4	RCBD
ARE20_FHB	F4:8/DH	12	11	23	1	2	RCBD

[a] Trial types and the years each were grow. TP17/TP18_FHB was grown over four years between 2013-2014 and 2016-2017; 17, 2016-2017; 18, 2017-2018; 19, 2018-2019; 20, 2019-2020. FHB, grown in a Fusarium head blight nursery.

[b] Breeding trials consisted of conventionally bred genotypes as well as doubled haploid (DH) genotypes.

[c] AUG, augmented design; RCBD, randomized complete block design.

Table 2. Description of five training populations (TP), including the subpopulations within them, the number of genotypes within them, the validation populations (VP) they are predicting for, and the validation populations genome estimated breeding values (GEBVs) will be compared with.

Training Population	Subpopulations[a]	TP Genotypes	Validation Population[a]	VP Genotypes	Comparisons[a,b]
TP17	-	355	WO17	140	ADV18, ARE19
TP18	TP17, WO17	495	WO18	160	ADV19, ADV19R, ARE20
TP19	TP17, WO17, ADV18, WO18	655	WO19, DHWO19	433	ADV20
TP18_FHB	-	355	ADV18	104	ARE19
TP19_FHB	TP18_FHB, ADV18	459	ADV19	120	ARE20
[a] WO, $F_{4:6}$ observation breeding nursery; ADV, $F_{4:7}$ advanced breeding nursery; ARE, $F_{4:8}$ elite breeding nursery.					
[b] GEBVs were compared with phenotypic results from following generations.					

Table 3. Descriptive statistics, Pearson phenotypic correlations, and broad-sense heritabilites (H^2) for adjusted means for three training populations (TP), four $F_{4:6}$ observation nurseries (WO), four advanced $F_{4:7}$ nurseries (ADV), and three elite $F_{4:8}$ nurseries (ARE) for three agronomic traits including grain yield (GY), heading date (HD), and test weight (TW).

Trial	Trait	Summary Statistics						Correlations	
		Mean	Min	Max	Range	SD	H^{2a}	GY[b]	HD[c]
TP17	GY	3.86	.13	9.54	9.41	1.14	.82	-	
	HD	105.66	78.74	129.00	50.25	13.66	.94	-.17***	-
	TW[d]	74.35	47.27	86.94	41.08	5.33	.82	.44***	-.07[ns]
TP18	GY	4.27	1.92	6.92	5.00	.84	.82	-	
	HD	103.21	82.00	118.38	36.38	10.68	.92	-.73***	-
	TW	74.69	55.24	83.23	29.40	4.65	.94	.05[ns†]	.10*
TP19	GY	4.04	.01	6.92	6.91	.95	.84	-	
	TW	74.53	29.40	83.23	55.26	5.44	.81	.37***	-
WO17	GY	5.29	3.46	6.92	3.46	.71	.37	-	
	HD	86.91	82.00	96.00	14.00	3.37	.71	-.08[ns]	-
	TW	74.86	55.89	84.60	30.13	6.56	.47	.16*	-.25***
WO18	GY	3.55	.20	5.36	5.17	1.15	.86	-	
	HD	106.26	102.00	116.00	14.00	3.51	1.00	.56***	-
	TW	72.32	29.40	78.07	50.09	6.77	.95	.48***	.21**
WO19	GY	3.10	.32	5.47	5.14	1.09	.26	-	
	TW	63.77	51.68	71.38	21.12	5.11	.00	.54***	-
DHWO19	GY	3.10	.38	5.43	5.05	.96	.69	-	
	HD	108.17	101.00	129.00	28.00	7.42	.78	-.28***	-
	TW	65.94	6.15	80.94	76.20	11.12	.96	.31***	-.69***
ADV18	GY	3.57	1.07	6.78	5.70	1.02	.41	-	
	HD	103.11	97.00	110.00	13.00	3.13	.88	.86***	-
	TW	77.29	47.12	105.68	59.99	6.87	.36	.13***	.03[ns]
ADV19	GY	3.68	1.16	5.39	4.23	.72	.56	-	
	HD	102.38	97.00	109.00	12.00	2.35	.81	.28**	-
	TW	70.11	53.78	82.61	30.26	6.70	.85	.56***	.18*
ADV19R	GY	4.30	1.47	7.50	6.03	1.02	.44	-	
	HD	109.92	94.00	121.00	27.00	6.85	.95	-.43***	-
	TW[d]	76.19	45.22	82.21	38.40	5.32	.54	.33***	-.28***
ADV20	GY	3.92	.30	7.61	7.30	1.10	.43	-	
	HD	107.20	91.00	124.00	33.00	8.39	.58	.15**	-
	TW	77.88	16.90	95.21	79.73	10.06	.39	.16***	.20***
ARE19	GY	3.91	1.77	5.59	3.82	.79	.81	-	
	HD	102.20	98.00	108.00	10.00	2.27	.84	.09[ns†]	-
	TW	72.00	50.98	82.66	33.10	6.77	.85	.66***	.18[ns]
ARE20	GY	3.89	0.39	6.41	6.02	.90	.32	-	
	HD	99.49	94.00	111.00	17.00	3.39	.76	-.22**	-

Table 3 (Cont.)

Trial	Trait	Summary Statistics						Correlations	
		Mean	Min	Max	Range	SD	H^{2a}	GY[b]	HD[c]
ARE20	TW	75.39	20.70	94.47	75.19	11.74	.38	.58***	-.59***

[a] Broad-sense heritability (H^2) for unbalanced multi-environmental experimental designs calculated using the formula proposed by Cullis et al. (2006) for TP17, TP18, TP19, ADV19R, and TP19_FHB. Broad-sense heritability for all over populations was calculated using plot-mean based broad-sense heritability.

[b] Grain yield was recorded in t ha^{-1}.

[c] Heading date was recorded in Julian days after 1 Jan, when 50% of the heads were 50% emerged from the flag leaf.

[d] Test weight was recorded in kg hl^{-1}.

*Significant at the .05 probability level. **Significant at the .01 probability level. ***Significant at the .001 probability level. †ns, nonsignificant at the .05 probability level.

Table 4. Descriptive statistics, Pearson phenotypic correlations, and broad-sense heritabilities (H^2) for adjusted means for two training populations, two advanced $F_{4:7}$ nurseries, and two elite $F_{4:8}$ nurseries for three Fusarium head blight (FHB) resistance traits, including deoxynivalenol (DON), Fusarium damaged kernels (FDK), and severity (SEV) as well as heading date (HD) and plant height (PH).

Trial[a]	Trait	Summary Statistics						Correlations			
		Mean	Min	Max	Range	SD	H^{2b}	DON[c]	FDK	SEV	HD[d]
FHB_TP18	DON	10.53	.08	92.80	92.72	11.35	.74	-			
	FDK	32.49	.00	100.00	100.00	29.93	.79	.40***	-		
	SEV	27.88	.00	100.00	100.00	25.78	.82	.32***	.73***	-	
	HD	94.69	74.00	118.00	44.00	10.19	.90	.25***	-.05^{ns†}	-.10*	-
	PH[e]	35.54	22.23	47.98	25.75	3.96	.91	.01^{ns}	-.29***	-.36***	.34***
FHB_TP19	DON	14.26	6.15	37.50	31.35	4.59	.98	-			
	FDK	38.22	6.00	92.12	86.12	14.86	1.00	.45***	-		
	SEV	28.67	3.75	91.71	87.96	12.97	1.00	.12*	.55***	-	
	HD	97.91	86.76	116.50	29.74	8.30	.94	.31***	.02^{ns}	-.54***	-
	PH	35.59	28.00	44.50	16.50	2.74	.94	.00^{ns}	-.29***	-.31***	.16***
ADV18	DON	16.64	3.60	51.50	47.90	7.60	.62	-			
	FDK	39.32	2.00	75.00	73.00	16.29	.77	.62***	-		
	SEV	15.44	.00	85.00	85.00	14.88	.38	.27**	.54***	-	
	HD	112.22	108.00	117.00	9.00	2.16	.90	.28**	-.10^{ns}	-.34***	-
	PH	35.22	27.00	47.00	20.00	3.36	.71	-.05^{ns}	-.22*	-.18*	.25***
ADV19	DON	10.09	.12	74.50	74.38	10.08	.61	-			
	FDK	31.01	.00	98.00	98.00	23.94	.83	.86***	-		
	SEV	25.60	.00	95.00	95.00	25.54	.45	.76***	.86***	-	
	HD	102.38	97.00	109.00	12.00	2.35	.81	.10^{ns}	-.04^{ns}	-.17^{ns}	-
	PH	31.97	25.00	40.00	15.00	2.82	.74	.29***	.09^{ns}	.01^{ns}	.35***
ARE19	DON	8.51	.59	64.10	63.51	8.32	.50	-			
	FDK	27.04	1.00	95.00	94.00	21.06	.71	.84***	-		
	SEV	23.55	.00	90.00	90.00	22.87	.43	.74***	.86***	-	
	HD	102.20	98.00	108.00	10.00	2.27	.84	.01^{ns}	-.28^{ns}	-.28^{ns}	-
	PH	31.52	21.00	37.00	16.00	2.81	.76	.21^{ns}	-.12^{ns}	-.18^{ns}	.41*

Table 4 (Cont.)

Trial[a]	Trait	Summary Statistics						Correlations			
		Mean	Min	Max	Range	SD	H^{2b}	DON[c]	FDK	SEV	HD[d]
ARE20	DON	7.30	.99	19.30	18.31	3.95	.78	-			
	FDK	15.21	2.00	60.00	58.00	12.49	.84	.78***	-		
	SEV	16.83	.00	60.00	60.00	13.11	.76	.65***	.82***	-	
	HD	99.49	94.00	111.00	17.00	3.39	.76	.09^{ns†}	.09^{ns}	-.08^{ns}	-
	PH[e]	89.69	76.20	101.60	25.40	6.20	.87	.01^{ns}	-.05^{ns}	-.11^{ns}	.48**

[a] TP, training population; ADV, $F_{4:7}$ advanced FHB trial; ARE, $F_{4:8}$ elite FHB trial.

[b] Broad-sense heritability (H^2) for unbalanced, multi-environmental designs calculated using the formula proposed in Cullis et al. (2006) for TP19_FHB. Broad-sense heritability for all other trials calculated using plot-mean based heritability.

[c] DON was recorded in μg g^{-1}, whereas FDK and SEV were recorded in percentage.

[d] Heading date was recorded in Julian days after 1 Jan, when 50% of the heads were 50% emerged from the flag leaf.

[e] Plant height was recorded in inches from the surface of the soil to the tip of the head, but reported in centimeters here.

*Significant at the .05 probability level. **Significant at the .01 probability level. ***Significant at the .001 probability level. †ns, nonsignificant at the .05 probability level.

Table 5. Comparison of three selection methods, phenotypic selection (PS) based on grain yield (GY) data from the observation nursery (WO), naïve genomic selection (NGS), and multivariate genomic selection (MVGS), based on correlations between genome estimated breeding values (GEBVs) and the adjusted means from following generations, response to selection, and selection accuracy of genotypes in the final generation.

TP[a]	Method	*r* WO[b]	*r* ADV+ARE[c]	Selection Differential[d]	Response to Selection[e]	Selection Accuracy[f]
				kg ha^{-1}	kg ha^{-1}	%
TP17	PS	-	-.02^{ns†}	-36.87	-31.71	50
	NGS	.38***	-.01^{ns}	30.38	26.13	36.4
	MVGS	.26**	.14^{ns}	72.08	61.99	36.4
TP18	PS	-	-.01^{ns}	-21.77	-14.33	52.2
	NGS	.12^{ns}	.11^{ns}	7.15	4.71	60.8
	MVGS	.35***	.11^{ns}	-2.26	-1.49	69.6
TP19	PS	-	.09^{ns}	27.43	11.80	67.8
	NGS	.10*	-.05^{ns}	.13	.06	55.5
	MVGS	.29***	-.05^{ns}	-11.57	-4.98	71.2

[a] TP, training population used to calculate GEBVs.

[b] Pearson correlation coefficient between GEBVs and phenotypic data from the WO population used as a validation population (VP).

[c] Pearson correlations coefficient between GEBVs and adjusted means for phenotypic data across the advanced (ADV) and elite (ARE) generations.

[d] Calculated as $S = \mu_{selected} - \mu_{unselected}$; $\mu_{selected}$, the mean of the top 50% of genotypes in the ARE (or ADV) selected using the given selection strategy; $\mu_{unselected}$, the mean of the full population without any selection.

[e] Calculated as $R = H^2S$ where H^2 is the heritability for GY based on the adjusted means across ADV and ARE generations.

[f] Percent of lines correctly selected in the ARE generation based using genotypes with above average GEBVs and also above average phenotypic values for GY; ADV generation for TP19.

*Significant at the .05 probability level. **Significant at the .01 probability level. ***Significant at the .001 probability level. †ns, nonsignificant at the .05 probability level.

Table 6. Comparison of three selection methods, phenotypic selection based on three FHB resistance traits using two training populations (TP), deoxynivalenol (DON) concentration, Fusarium damaged kernels (FDK), and severity (SEV) from the advanced trials (ADV), naïve genomic selection (NGS), and multivariate genomic selection (MVGS), based on correlations between genome estimated breeding values and the adjusted means from following generations, response to selection, and selection accuracy of genotypes in the final generation.

TP	Trait	Method	*r* ADV[a]	*r* ARE[b]	Selection Differential[c]	Response to Selection[d]	Selection Accuracy[e]
TP18_FHB	DON	PS	-	-.01^{ns†}	.40	.20	52.9
		NGS	.22*	.19^{ns}	-.73	-.37	82.4
		MVGS	.53***	.10^{ns}	-.46	-.23	70.6
	FDK	PS	-	.14^{ns}	-2.24	-1.59	58.8
		NGS	.41***	.38^{ns}	-5.77	-4.09	70.6
		MVGS	.70***	.42^{ns}	-3.99	-2.83	70.6
	SEV	PS	-	.54*	-3.46	-1.49	52.9
		NGS	.29**	.16^{ns}	-1.90	-.82	41.2
		MVGS	.57***	.60*	-5.33	-2.29	47.1
TP19_FHB	DON	PS	-	.51*	-1.32	-1.03	13.0
		NGS	.17^{ns}	.37^{ns}	-.67	-.53	56.5
		MVGS	.71***	.45*	-.96	-.75	69.6
	FDK	PS	-	.67***	-4.07	-3.42	91.3
		NGS	.18*	.45*	-3.21	-2.70	34.8
		MVGS	.83***	.64**	-4.57	-3.84	60.9
	SEV	PS	-	.78***	-5.86	-4.45	82.6
		NGS	.25**	.08^{ns}	.50	.38	60.9
		MVGS	.67***	.12^{ns}	-.18	-.13	82.6

[a] Pearson correlation coefficient between GEBVs and phenotypic data from the ADV population used as a validation population (VP).

[b] Pearson correlations coefficient between GEBVs and adjusted means for phenotypic data from the elite (ARE) generation.

Table 6 (Cont.)

[c] Calculated as $S = \mu_{selected} - \mu_{unselected}$; $\mu_{selected}$, the mean of the top 50% of genotypes selected using the given selection strategy; $\mu_{unselected}$, the mean of the full population without any selection. DON is presented as $\mu g\ g^{-1}$ and FDK and SEV are presented as percentages.
[d] Calculated as $R = H^2S$ where H^2 is the heritability for a given trait based on the adjusted means across ADV and ARE generations. DON is presented as $\mu g\ g^{-1}$ and FDK and SEV are presented as percentages.
[e] Percent of lines correctly selected in the ARE generation based using genotypes with above average GEBVs and also above average phenotypic values for a given trait.
*Significant at the .05 probability level. **Significant at the .01 probability level. ***Significant at the .001 probability level. †ns, nonsignificant at the .05 probability level.

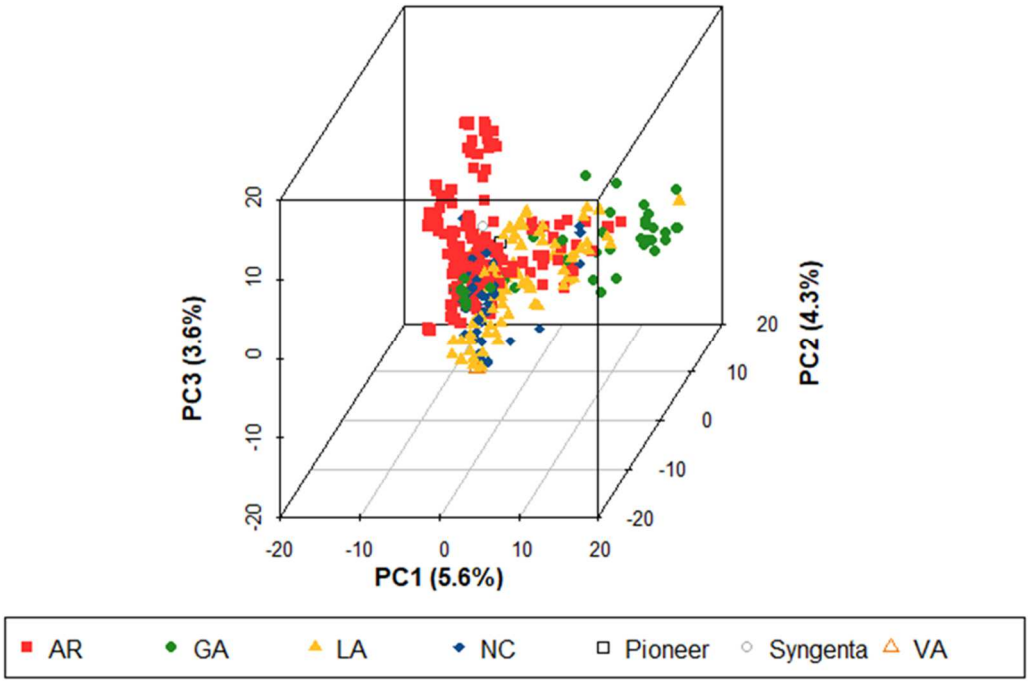


Figure 1. Population structure of 355 soft red winter wheat genotypes using 5,202 single nucleotide polymorphism (SNP) markers. This population represents the 2017 training population (TP17) used to predict grain yield for the 2017 $F_{4:6}$ observation nursery (WO17) and the training population used to predict three Fusarium head blight (FHB) resistance traits including deoxynivalenol (DON) concentration, Fusarium damaged kernels (FDK), and severity (SEV) (TP18_FHB) for the 2018 advanced Fusarium head blight trial (ADV18). Colors represent the origin of the genotypes. AR, developed at the University of Arkansas, Fayetteville; GA, developed at the University of Georgia, Athens; LA, developed at Louisiana State University, Baton Rouge; NC, developed at North Carolina State University, Raleigh; Pioneer, developed by Pioneer Hi-Bred International; Syngenta, developed by Syngenta and AgriPro; and VA, developed by Virginia Polytechnic Institute and State University, Blacksburg; PC, principal component.

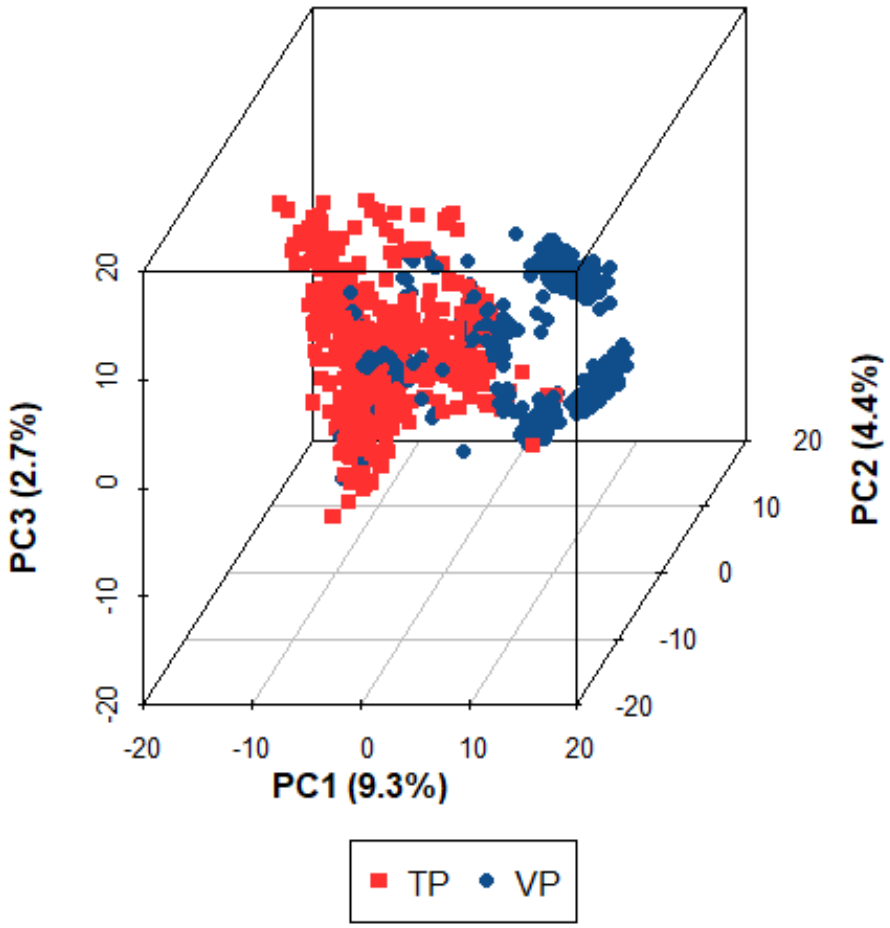


Figure 2. Population structure between the 2019 training population (TP19) used to predict grain yield for the 2019 $F_{4:6}$ observation nursery (WO19) using 5,202 single nucleotide polymorphism (SNP) markers. Colors represent the population type. TP, TP19 population consisting of genotypes from the 2017 training population (TP17), WO17 and WO18 nurseries, and 2018 advanced $F_{4:7}$ nursery (ADV18); VP, validation population consisting of genotypes from the WO19 nursery; PC, principal component.

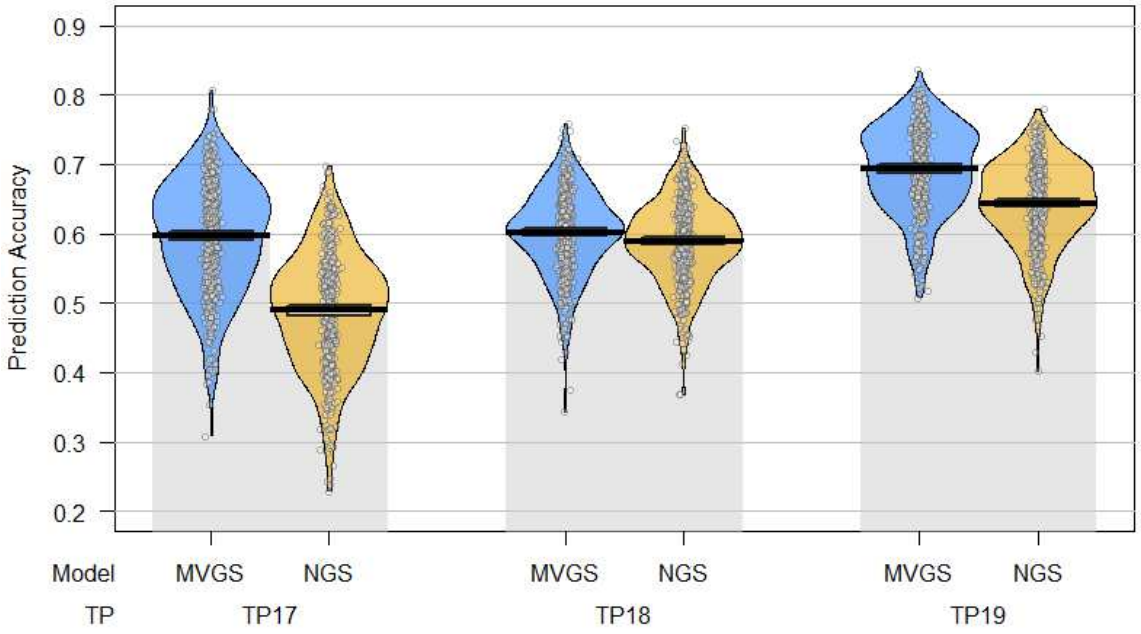


Figure 3. Pirate plot comparing the mean prediction accuracies between multivariate genomic selection (MVGS) models with naïve genomic selection (NGS) models for grain yield (GY) in soft red winter wheat across three different training populations (TP). The *x* axis represents the combination of TP and GS model used to predict GY. The *y* axis represents the mean prediction accuracy across 100 iterations of fivefold cross-validation in the form of a Pearson correlation coefficient (*r*) between the predicted genome estimated breeding value (GEBV) and the actual phenotypic value for the validation populations. Individual points represent the Pearson correlation from each fold of each iteration of cross-validation for a total of 500 datapoints. The lines within each plot represent the mean and 95% confidence intervals for prediction accuracy. The curves represent the smoothed densities of the data.

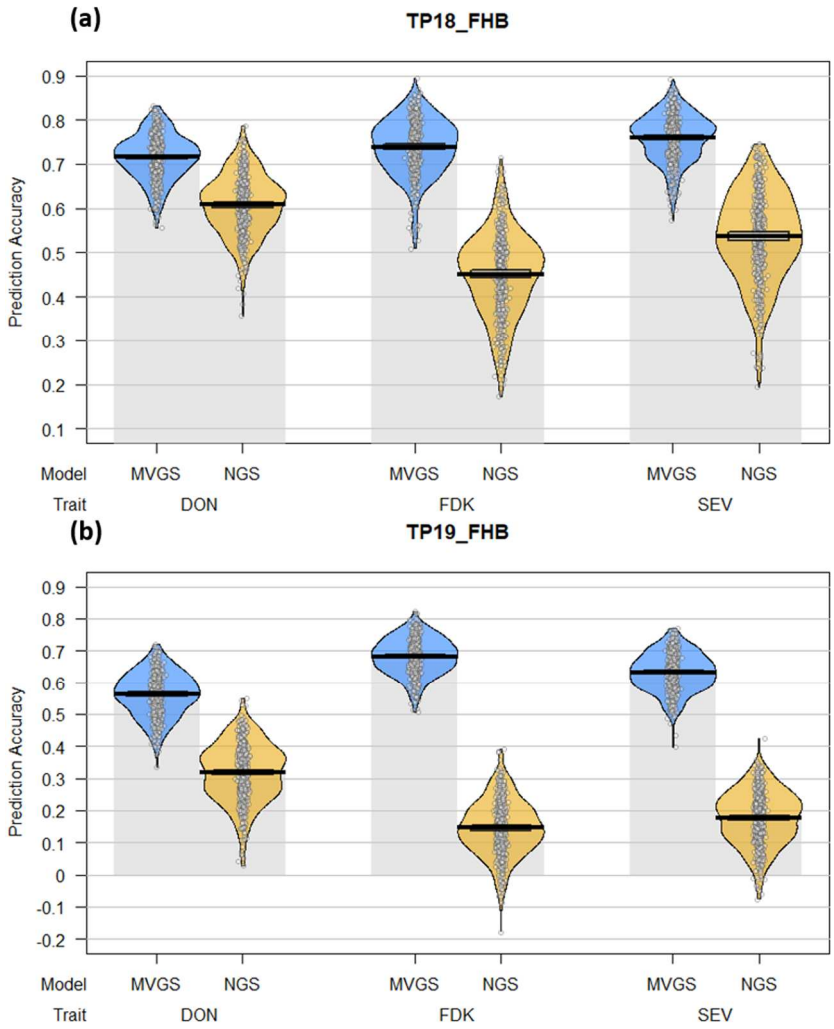


Figure 4. Pirate plots comparing the mean prediction accuracies between multivariate genomic selection (MVGS) models with naïve genomic selection (NGS) models for three Fusarium head blight resistance traits (FHB), deoxynivalenol (DON) concentration, Fusarium damaged kernels (FDK), and severity (SEV) in soft red winter wheat across two training populations (TPs): (a) TP18_FHB, TP used to predict three FHB resistance traits for the 2018 advanced $F_{4:7}$ generation (ADV18); (b) TP19_FHB, TP used to predict three FHB resistance traits for the 2019 advanced $F_{4:7}$ generation (ADV19), consisting of all genotypes from TP18_FHB and ADV18. The x axis represents the combination of FHB resistance traits and GS model used to predict each trait. The y axis represents the mean prediction accuracy across 100 iterations of fivefold cross-validation in the form of a Pearson correlation coefficient (r) between the predicted genome estimated breeding value (GEBV) and the actual phenotypic value for the validation populations. Individual points represent the Pearson correlation from each fold of each iteration of cross-validation for a total of 500 datapoints. The lines within each plot represent the mean and 95% confidence intervals for prediction accuracy. The curves represent the smoothed densities of the data.

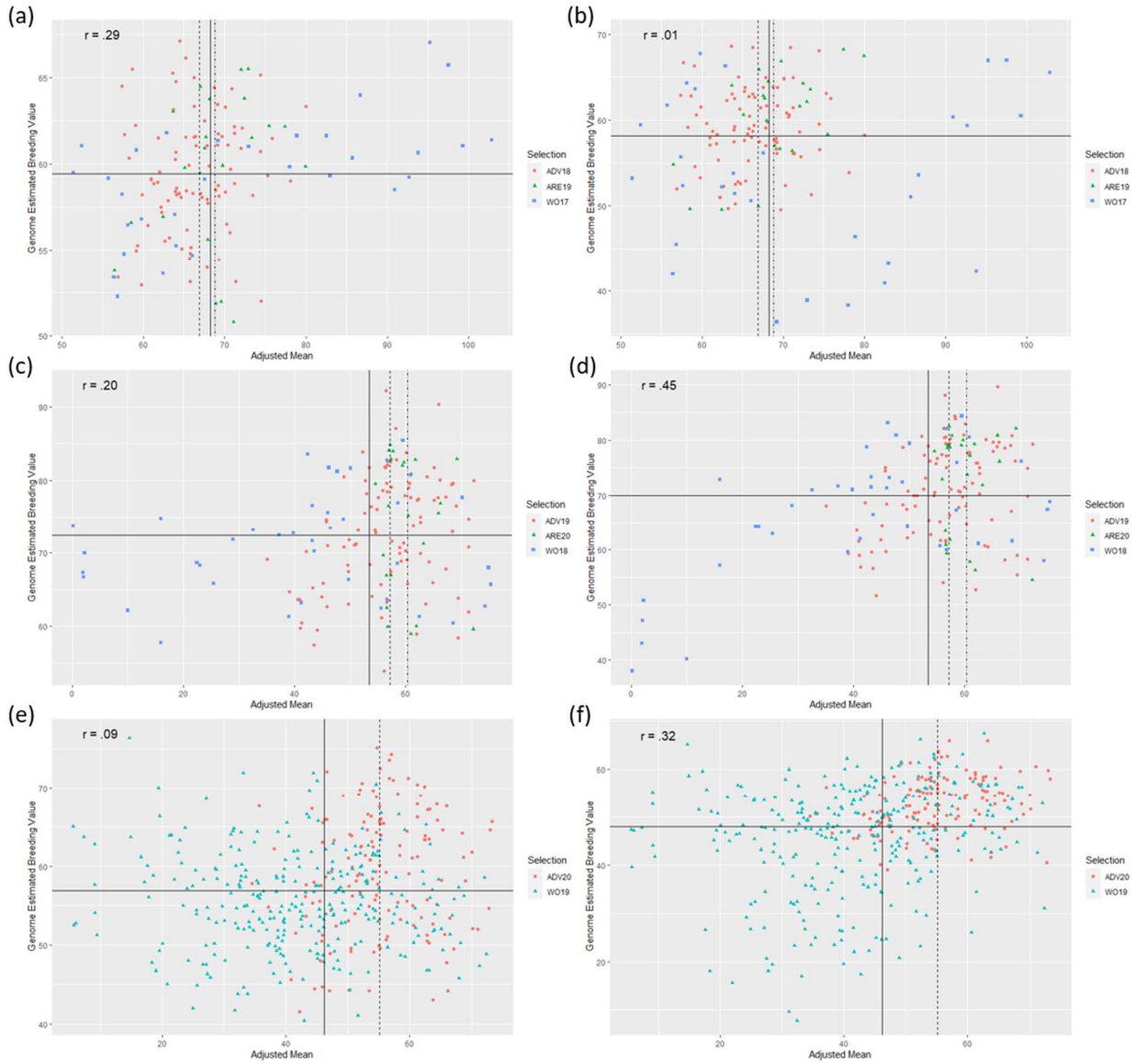


Figure 5. Scatter plots between genome estimated breeding values (GEBVs) for grain yield (GY) in soft red winter wheat from two different genomic selection models (GS), including naïve models without covariates (NGS) and multivariate GS models with test weight as a covariate (MVGS), and adjusted means for GY across multiple generations of selection, starting with the 2016-2017 growing season through the 2018-2019 growing season: (a) predictions for $F_{4:6}$ observation genotypes from the 2016-2017 growing season (WO17) using a NGS model, (b) predictions for WO17 using a MVGS model, (c) predictions for $F_{4:6}$ observation genotypes from the 2017-2018 growing season (WO18) using a NGS model, (d) predictions for WO18 using a MVGS model, (e) predictions for $F_{4:6}$ observation genotypes from the 2018-2018 growing season (WO18) using a NGS model, (f) predictions for WO18 using a MVGS model. The *x*-axis represents adjusted mean for GY across the WO, $F_{4:7}$ advanced (ADV), and $F_{4:8}$ elite (ARE) generations. Adjusted means for GY were only calculated across WO19 and ADV20 for (e) and (f). The *y*-axis represents the GEBVs calculated for GY from the NGS or MVGS models. Different colored datapoints represent genotypes that were advanced to the next generation. The solid vertical line represents the mean of the adjusted means for GY from the WO generation,

while the vertical dashed line represents the mean of the adjusted means for GY from the ADV generations, and the vertical dot-dash line represents the mean of the adjusted means for GY from the ARE generation. The solid horizontal line represents the mean of GEBVs for GY calculated from the NGS or MVGS models. The *r* label represents the Pearson correlation between GEBVs and adjusted means.

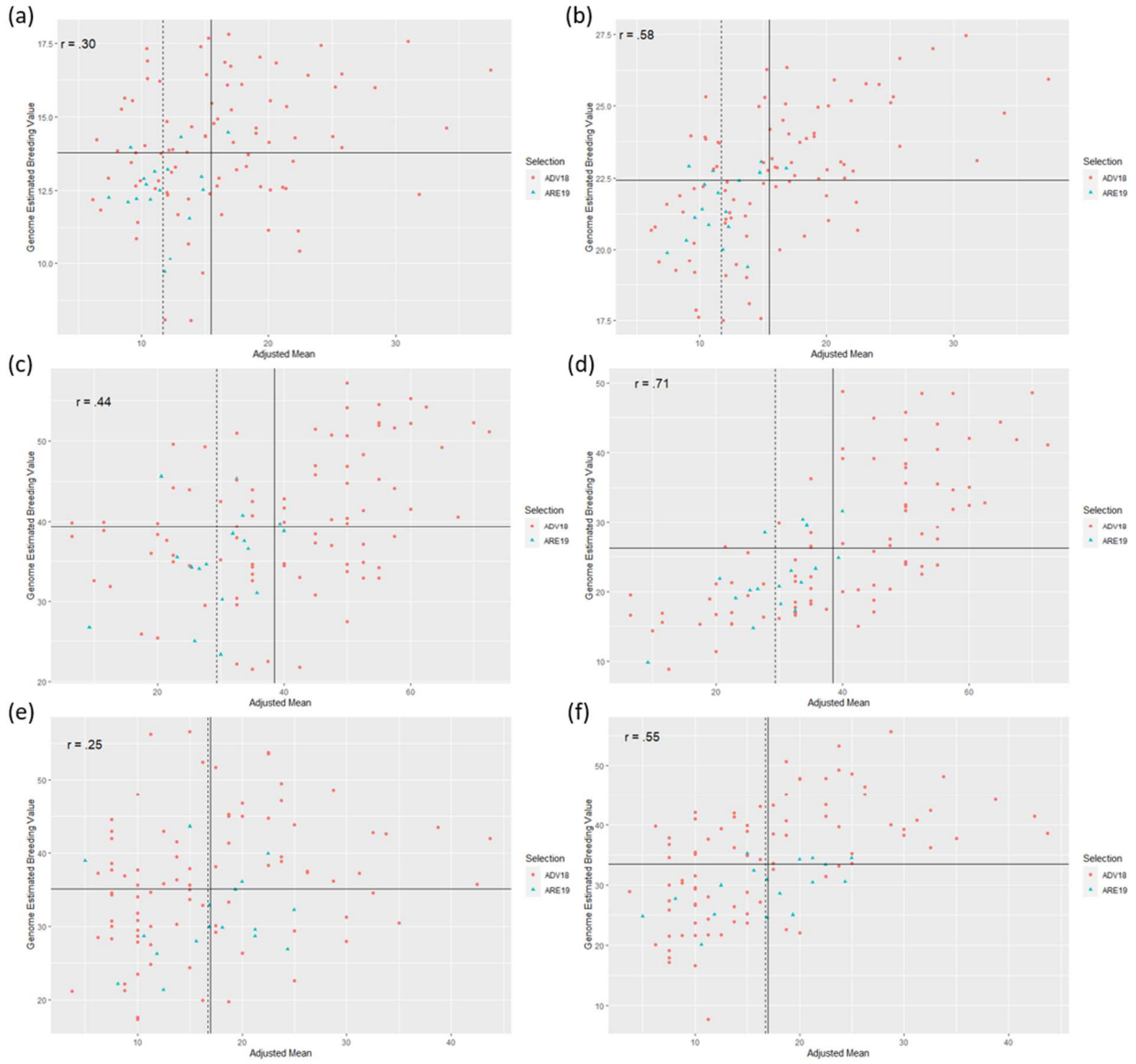


Figure 6. Scatter plots between genome estimated breeding values (GEBVs) for three Fusarium head blight (FHB) resistance traits in soft red winter wheat from two different genomic selection models (GS), including naïve models without covariates (NGS) and multivariate GS models with covariates (MVGS), and adjusted means for deoxynivalenol (DON) concentration, Fusarium damaged kernels (FDK), and severity (SEV) across two generations, $F_{4:7}$ advanced from 2017-2018 (ADV18) and $F_{4:8}$ elite from 2018-2019 (ARE19): (a) predictions for DON in ADV18 using a NGS model, (b) predictions for DON using a MVGS model, (c) predictions for FDK from ADV18 using a NGS model, (d) predictions for FDK using a MVGS model, (e) predictions for SEV in ADV18 using a NGS model, (f) predictions for SEV using a MVGS model. The *x*-axis represents adjusted mean for DON, FDK, or SEV across the ADV and ARE generations. The *y*-axis represents the GEBVs calculated for DON, FDK, or SEV from the NGS or MVGS models. Different colored datapoints represent genotypes that were advanced to the next generation. The solid vertical line represents the mean of the adjusted means for the respective FHB resistance trait from the ADV generation, while the vertical dashed line represents the mean of the adjusted means for the respective FHB resistance trait from the ARE generations. The

solid horizontal line represents the mean of GEBVs for the respective FHB resistance trait calculated from the NGS or MVGS models. The *r* label represents the Pearson correlation between GEBVs and adjusted means.

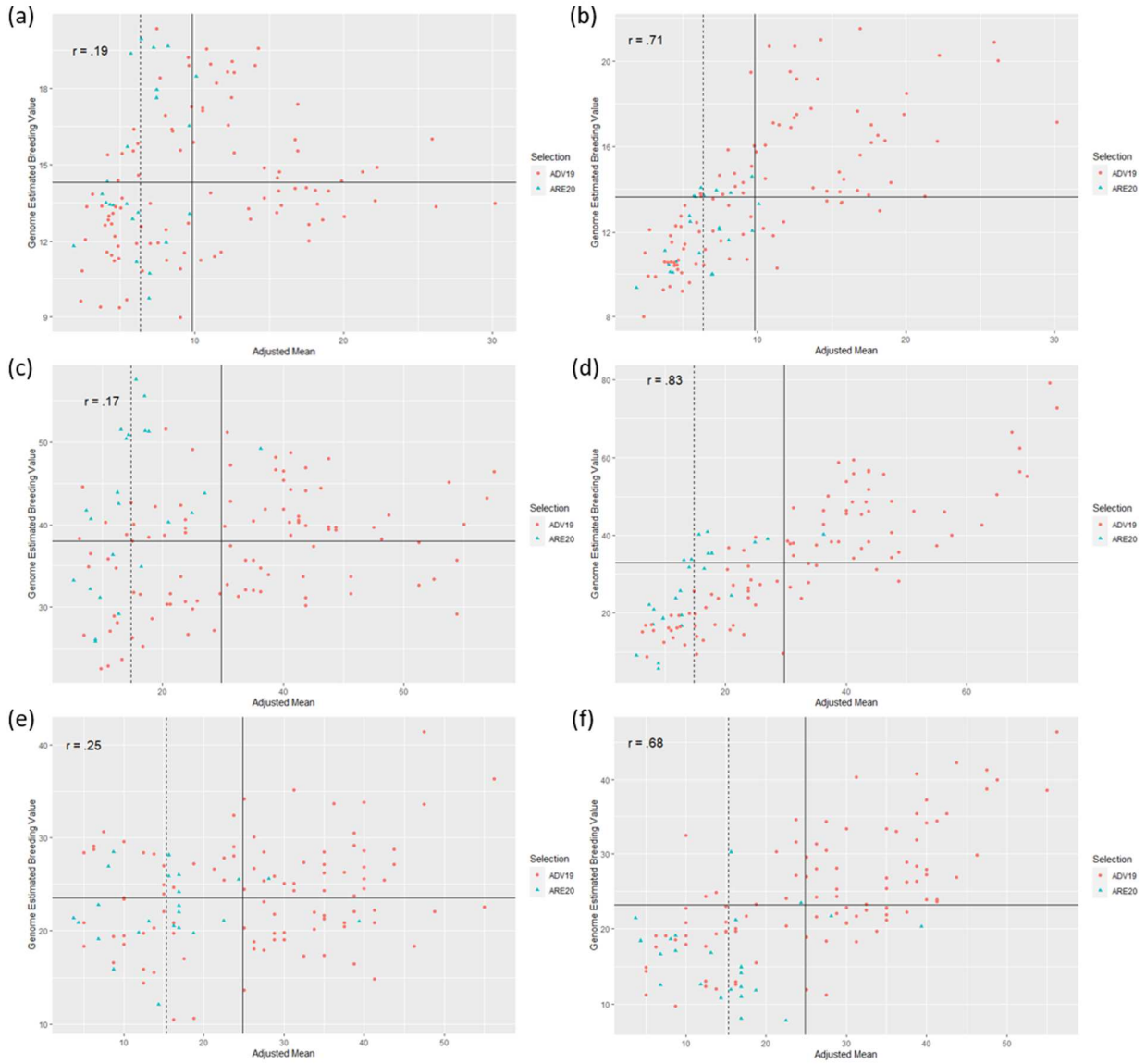


Figure 7. Scatter plots between genome estimated breeding values (GEBVs) for three Fusarium head blight (FHB) resistance traits in soft red winter wheat from two different genomic selection models (GS), including naïve models without covariates (NGS) and multivariate GS models with covariates (MVGS), and adjusted means for deoxynivalenol (DON) concentration, Fusarium damaged kernels (FDK), and severity (SEV) across two generations, $F_{4:7}$ advanced from 2018-2019 (ADV19) and $F_{4:8}$ elite from 2019-2020 (ARE20): (a) predictions for DON in ADV19 using a NGS model, (b) predictions for DON using a MVGS model, (c) predictions for FDK from ADV19 using a NGS model, (d) predictions for FDK using a MVGS model, (e) predictions for SEV in ADV18 using a NGS model, (f) predictions for SEV using a MVGS model. The *x*-axis represents adjusted mean for DON, FDK, or SEV across the ADV and ARE generations. The *y*-axis represents the GEBVs calculated for DON, FDK, or SEV from the NGS or MVGS models. Different colored datapoints represent genotypes that were advanced to the next generation. The solid vertical line represents the mean of the adjusted means for the respective FHB resistance trait from the ADV generation, while the vertical dashed line represents the mean of the adjusted means for the respective FHB resistance trait from the ARE generations. The

solid horizontal line represents the mean of GEBVs for the respective FHB resistance trait calculated from the NGS or MVGS models. The *r* label represents the Pearson correlation between GEBVs and adjusted means.

CONCLUSIONS

OVERALL CONCLUSIONS

Multiple genomic selection (GS) analyses were conducted for five agronomic traits, including grain yield (GY), heading date (HD), maturity date (MD), test weight (TW), and plant height (PH), as well as four Fusarium head blight (FHB) resistance traits, including deoxynivalenol accumulation (DON), Fusarium damaged kernels (FDK), incidence (INC), and severity (SEV) using a training population of soft red winter wheat (SRWW) genotypes. The training population was also used for forward prediction in order to predict the performance of untested $F_{4:6}$ breeding lines for GY and $F_{4:7}$ breeding lines for FHB resistance traits.

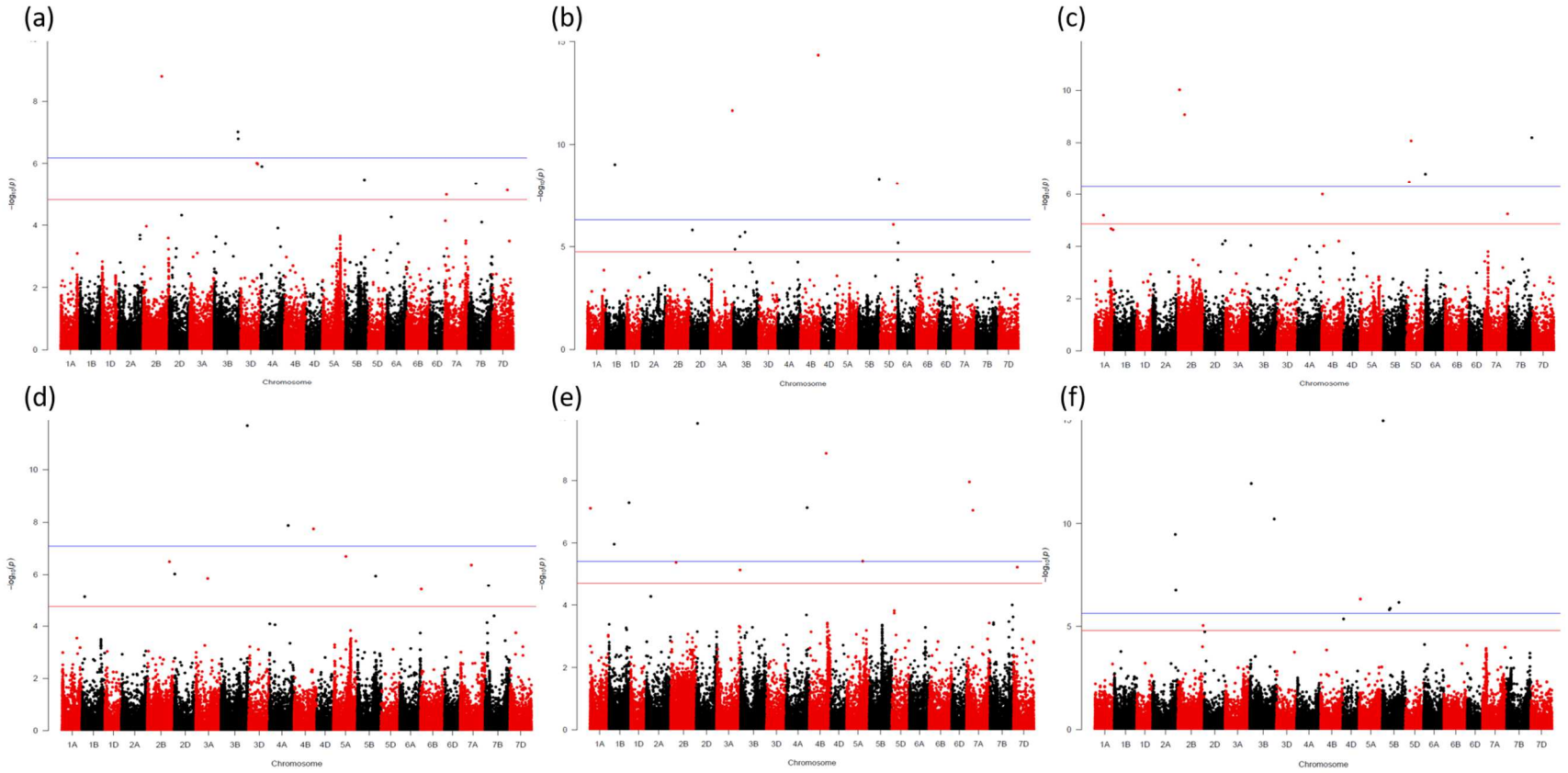
Additionally, a genome-wide association study (GWAS) was performed on the training population to identify novel loci associated with the four abovementioned FHB resistance traits. The GWAS identified QTL in regions previously reported in QTL mapping and GWAS studies, particularly on the long arms of chromosomes 3B and 4B. The identification of QTLs in our panel of 354 SRWW breeding genotypes adapted to the southeastern United States indicates that previously identified QTL are contributing to genetic resistance in this population. The SNPs identified could be implemented for MAS in wheat breeding programs.

The GS analyses for FHB resistance traits also found that naïve GS (NGS) models significantly outperformed GS models using significant markers identified through GWAS as fixed effects. However, multivariate GS (MVGS) models using secondary traits, correlated with the primary trait of interest, as covariates significantly improved prediction accuracies for all FHB and agronomic traits. This was particularly evident when there was a stronger genetic correlation between the trait of interest and covariates and the covariate had a higher heritability than the trait of interest. This indicated that MVGS could be successfully used when breeding for FHB resistance and agronomic traits. This study also found that using GY data from

environments closely related to a missing environment as covariates in a MVGS model to predict GY in the missing environment significantly improved prediction accuracy. Before implementing MVGS in their breeding programs, breeders must consider the genetic correlations between their traits as well as the heritability of their traits as results can vary across populations. Breeders must also be careful not to unintentionally select for undesirable traits if they have strong genetic correlations with a trait of interest. An example would be unintentionally selecting for taller plants while using PH as a covariate to predict TW, as both traits are positively correlated.

We also found that MVGS models performed significantly better than NGS models in terms forward prediction of untested genotypes for economically important traits, such as GY or FHB resistance traits. This is one of the first studies to show that MVGS could be effectively implemented for forward prediction within a wheat breeding program. This is also the first study to extensively investigate the use of forward prediction when breeding for FHB resistance in wheat. We found that GS could serve as a suitable, albeit imperfect, alternative to phenotypic selection when implemented during years where environmental conditions prohibit accurate phenotypic selection, particularly when experiencing late freezing events or extensive lodging.

APPENDICES


(a)
(b)
(c)
(d)
(e)
(f)
Chromosome
1A 1B 1D 2A 2B 2D 3A 3B 3D 4A 4B 4D 5A 5B 5D 6A 6B 6D 7A 7B 7D

Appendix 1. Manhattan plots for four phenotypic traits associated with Fusarium head blight resistance: (a) deoxynivalenol (DON) accumulation, (b) Fusarium damaged kernels (FDK), (c) incidence (INC), (d) severity (SEV), (e) heading date (HD), and (f) plant height (PH). The x-axis represents the 21 wheat chromosomes. The y-axis represents the $-\log_{10}(p\text{-value})$ of the association between a single nucleotide polymorphism (SNP) marker and the trait of interest. The red horizontal lines represent the threshold for declaring a SNP marker as significant, using a Bonjamini-Hochberg false discovery rate of 0.10 (DON: $-\log_{10}(p) = 4.82$, FDK: $-\log_{10}(p) = 4.75$, INC: $-\log_{10}(p) = 4.86$, SEV: $-\log_{10}(p) = 4.75$, HD: $-\log_{10}(p) = 4.72$, PH: $-\log_{10}(p) = 4.78$). The blue horizontal line represents the threshold calculated in FarmCPU using a permutation test for each individual trait (DON: $-\log_{10}(p) = 6.18$, FDK: $-\log_{10}(p) = 6.31$, INC: $-\log_{10}(p) = 6.30$, SEV: $-\log_{10}(p) = 7.08$, HD: $-\log_{10}(p) = 5.42$, PH: $-\log_{10}(p) = 5.63$).

Appendix 2. The top three significant single-nucleotide polymorphisms (SNP) from a fixed and random model circulating probability unification (FarmCPU) model, with a Benjamini-Hochberg false discovery rate threshold of $q \leq 0.10$, used as fixed effects in a genotype **best linear unbiased prediction (**GBLUP**) genomic selection model for four Fusarium head blight (FHB) traits, including deoxynivalenol (DON) accumulation, Fusarium damaged kernels, incidence (INC), and severity (SEV), as well as two phenological traits, heading date (HD) and plant height (PH) over 10 randomly selected training subsets of 283 genotypes.**

	Traits[a]					
Subset[b]	**DON**	**FDK**	**INC**	**SEV**	**HD**	**PH**
TP1	S7A_15248425	S3A_36924697	S7A_706438978	S3B_51321489	S4B_526725227	S5B_21403762
	S7A_14282369	S6B_175593693	S4A_597028557	S3B_20446350	S1B_561386646	S5A_31207940
	S3D_498442075	S4B_534404586	S2B_35068139	S7D_615435411	S7A_82582815	S3B_518277482
TP2	S3A_642133424	S2D_175925167	S2B_57283068	S4B_21368968	S7A_81885452	S3B_752173863
	S6A_5482360	S6A_561979919	S2B_31271227	S1D_222230632	S3A_737441262	S7B_126611056
	S7A_263474297	S2B_789393547	S1D_484428278	S1B_655834997	S7B_590232747	S5B_21403762
TP3	S2A_54864749	S4B_577008759	S7B_633023455	S4A_647100355	S2D_161255383	S7B_189318561
	S7A_263474297	S5D_548469050	S3D_181251255	S3B_783125543	S3B_741065919	S7D_203060
	S3B_779116067	S5D_424046598	S2B_642399809	S5D_548469057	S3A_713731888	S5A_30400811
TP4	S2B_574225916	S4B_577008759	S4B_543165477	S3B_555867387	S4B_526725227	S2B_618048921
	S6B_687241289	S6D_425897815	S1B_525464530	S2B_775018873	S6B_687101754	S3B_752173863
	S3B_225212457	S1B_671293704	S4B_522382631	S6B_94286753	S1A_6496631	S5D_363846822
TP5	S2B_574225916	S4A_718920432	S2A_546944252	S4B_575810081	S2D_40925400	S2A_694862057
	S7A_263474297	S5D_548463464	S7A_606560075	S5D_548930259	S7B_709513949	S2D_561935999
	S3B_749054930	S3B_689868478	S2B_774330791	S7B_706708847	S1B_139627751	S4D_469909453
TP6	S3B_749054930	S7B_712097406	S2B_744317223	S5B_459147174	S3A_739309750	S5B_21403762
	S7A_263474297	S3B_51321534	S1B_331007886	S3B_818987684	S2D_40925400	S5A_31207940
	S6B_346977040	S7A_4580957	S1D_18351199	S6B_687085031	S1A_6496631	S7A_580136731
TP7	S3B_749054930	S5B_270368480	S2B_23375290	S6B_687085031	S4B_526725227	S5B_21403762
	S6A_145739256	S3B_785489583	S4A_148243841	S1B_676778147	S6A_617200734	S1A_6496613
	S2B_18265867	S3B_51321489	S6A_23793725	S5B_548333940	S7A_82582815	S3D_556878443
TP8	S6D_25812397	S4B_534404586	S2D_13433473	S5D_358143404	S1A_6496631	S3B_752173863
	S2A_605208569	S7B_456959818	S5B_548689817	S4A_168927580	S1B_561386646	S3B_49364658

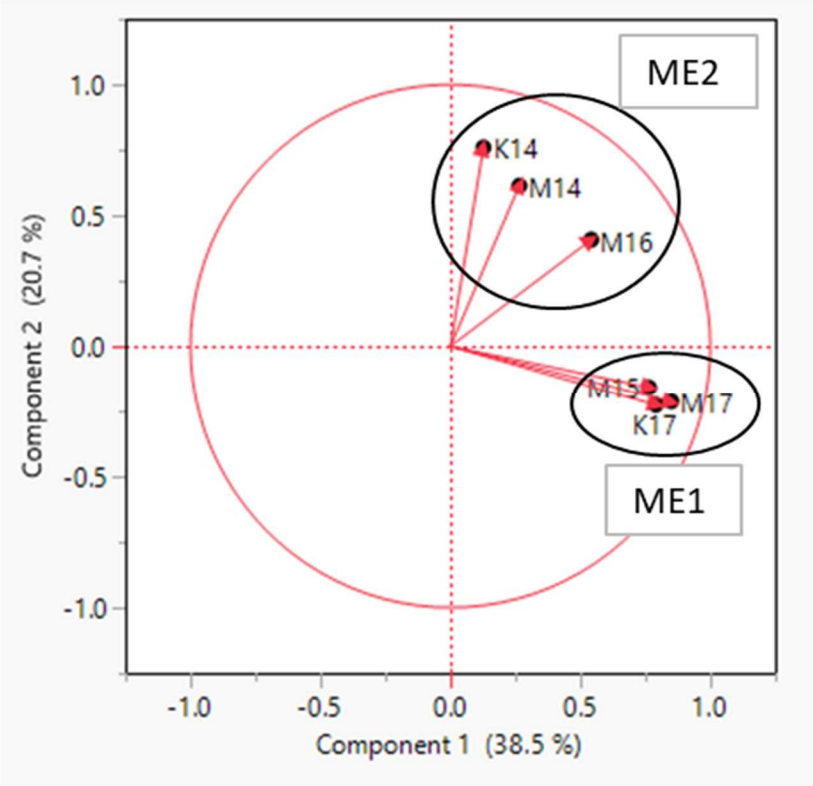
Appendix 2 (Cont.)

Subset[b]	Traits[a]					
	DON	**FDK**	**INC**	**SEV**	**HD**	**PH**
TP8	NA[c]	S4B_538715342	S4A_148243841	S6B_643642305	S4B_656452095	S5B_21403762
TP9	S4A_39713449	S5D_377751712	S4A_597058951	S7B_494497635	S4B_656452095	S5B_21403762
	S3B_749054930	S5D_424046598	S2B_197586765	S3B_818987707	S3B_141635343	S1B_446198866
	S4B_634284495	S4B_577008759	S4B_534795951	S4B_554987684	S1B_165428415	S5B_629129312
TP10	S7A_263474297	S4B_577008759	S5D_109218483	S3B_818987707	S1A_346637882	S4D_23740304
	S2B_574225916	S3A_714306573	S2D_481950465	S7D_169051163	S2D_40925400	S2D_14629280
	S5B_584753871	S1B_298191278	S4A_713631416	S3B_786072341	S7A_83786115	S5A_317247327

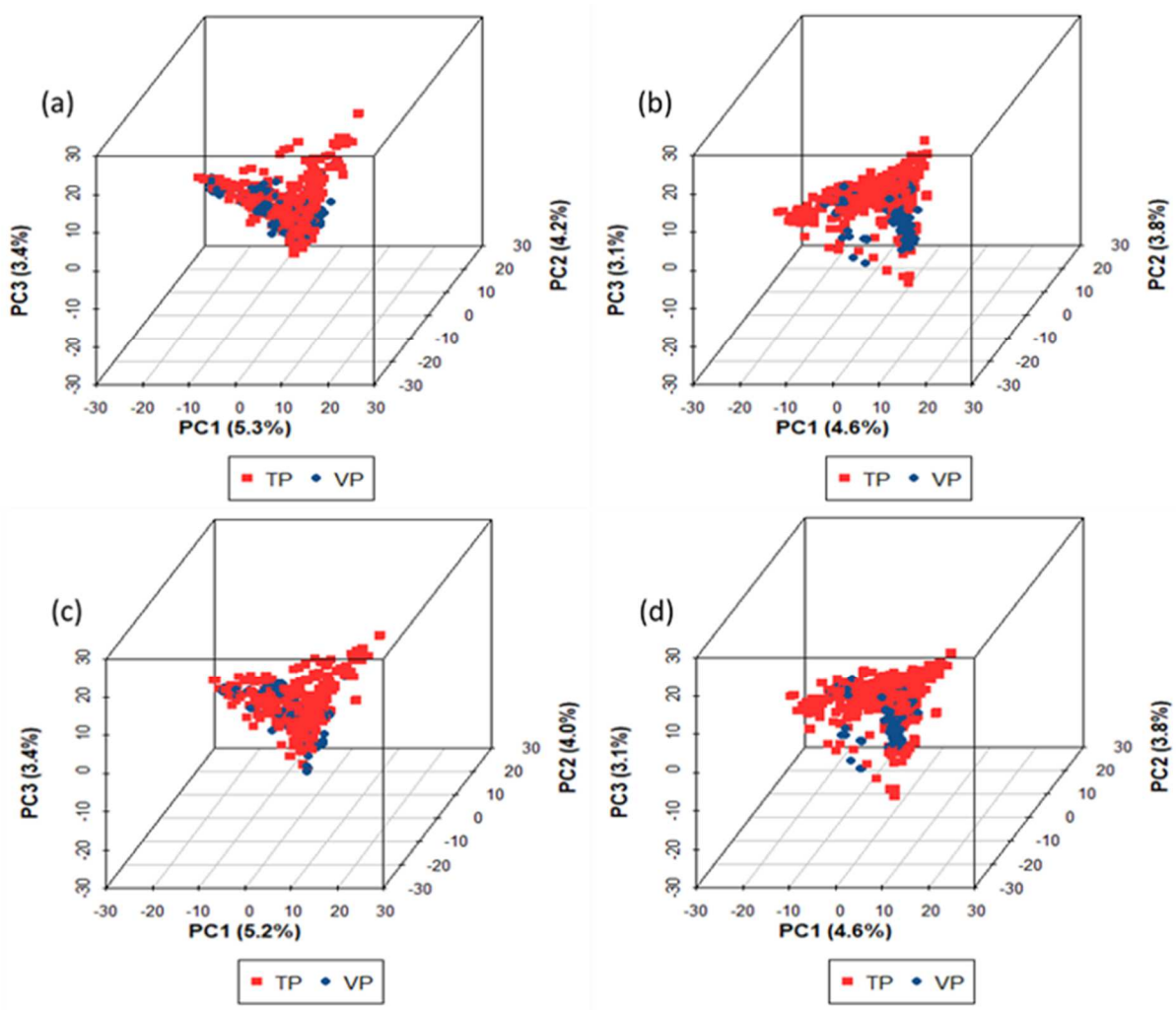
[a] DON, deoxynivalenol accumulation; FDK, Fusarium damaged kernels; INC, incidence; SEV, severity; HD, heading date; PH, plant height.

[b] TP, training population subsets were created by randomly selecting 283 genotypes from the full population of 354 genotypes.

[c] Only two markers exceeded the false discovery rate threshold from the FarmCPU model for the TP8 subset, therefore two markers were used as fixed effects in the G-BLUP model for DON for TP8.



Appendix 3. A biplot of six environments tested for grain yield and their respective clusters based on a principal component analysis-based clustering algorithm. The first cluster, mega-environment 1 (ME1), consisted of three environments, including Marianna 2017 (M17), Keiser 2017 (K17), and Marianna 2015 (M15). The second cluster, mega-environment 2 (ME2), consisted of three environments, including Keiser 2014 (K14), Marianna 2016 (M16), and Marianna 2014 (M14).



Appendix 4. Population structure between training populations (TP) used to predict grain yield or FHB resistance traits for validation populations (VP) using 5,202 single nucleotide polymorphism (SNP) markers. Colors represent the population type. (a) TP, training population consisting of genotypes from the 2017 training population (TP17); VP, validation population consisting of genotypes from the $F_{4:6}$ 2017 wheat observation nursery (WO17); PC, principal component. (b) TP, TP18 population consisting of genotypes from TP17 and WO17; VP, validation population consisting of genotypes from the $F_{4:6}$ 2018 wheat observation nursery (WO18); PC, principal component. (c) TP, training population consising of the same genotypes from TP17 to predict FHB resistance traits (TP18_FHB); VP, validation population consisting of genotypes from the $F_{4:7}$ 2018 advanced nursery (ADV18); PC, principal component. (d) TP, training population consising of the same genotypes from TP18_FHB and ADV18 to predict FHB resistance traits (TP19_FHB); VP, validation population consisting of genotypes from the $F_{4:7}$ 2019 advanced nursery (ADV19); PC, principal component.