REPURPOSING VARIANCE GENOME-WIDE ASSOCIATION STUDIES (vGWAS) FOR IDENTIFYING GENE-BY-ENVIRONMENT INTERACTIONS

BY

MATTHEW D MURPHY

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Crop Sciences
in the Graduate College of the
University of Illinois Urbana-Champaign, 2023

Urbana, Illinois

Doctoral Committee:

Associate Professor Alexander E. Lipka, Chair
Professor Stephen P. Moose
Assistant Professor Jessica E. Rutkoski
Adjunct Professor Elizabeth A. Ainsworth
Associate Professor Gota Morota

ABSTRACT

Identifying genotype by environment (GxE) interactions is vital in realizing different plant breeding objectives, such as breeding for responsiveness to new stressors or yield stability and uniformity. While GxE is an important contributor to a complex trait's overall genetic architecture, a bottleneck of identifying specific GxE loci for use in plant breeding and quantitative genetics is the heavy multiple testing correction involved with testing every genetic marker and environment interaction. One quantitative genetics phenomenon that may provide a shortcut to identifying GxE loci is through variance quantitative trait loci (vQTLs). While vQTLs have been identified in humans, yeast, animal model organisms, and livestock species, this is a relatively new idea in plant breeding. To further investigate using vQTLs to identify GxE loci, this thesis's three primary objectives are to 1.) conduct a simulation study using publicly available genotypic data from *Arabidopsis* and maize to assess the potential of two vGWAS models to identify GxE loci and other non-additive genetic loci like pure vQTLs and epistatic loci, 2.) apply the findings from objective 1.) to real flowering-time data in maize to assess how well vGWAS can aid in detecting GxE loci, and 3.) examine the merit of incorporating peak-associated vGWAS signals for obtaining genomic estimated breeding values that are stable across environments. This thesis argues that vGWAS can help identify GxE loci, but more work needs to be done to realize this full potential. I discuss this notion by using simulations and flowering-time traits in maize as a proof of concept for vGWAS and genomic selection.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Alexander Lipka. I was offered to join his lab as a Ph.D. student in 2019. With his guidance, I learned about the many different facets of quantitative genetics and how to speak “biology” with quantitative genetics. He helped rebuild my confidence in my scientific abilities and helped craft my passion for quantitative genetics.

My Ph.D. would not have been possible if not for the support of other individuals. I thank my committee members, Dr. Moose, Dr. Rutkoski, Dr. Ainsworth, and Dr. Gota Morota. Thank you for your guidance and mentorship in shaping me as a scientist. To the Department of Crop Science staff, thank you for assisting me with my research and helping me plan important events for the department. To the past and present Lipka Lab members for helping me with statistics, coding, research, and life in general: Dr. Samuel Fernandes, Dr. Brian Rice, Dr. Marcus Olatoye, Sarah Widener, and Talissa Oliveira Floriani Z. Souza. Finally, to the Lipka Lab collaborators: Dr. Ruthie Angelovici, Dr. Marianne Slater, Dr. Vivek Shrestha, Dr. Andrea Eveland, and Dr. Maxwell Braud.

I would also like to dedicate this dissertation to others throughout my educational journey. To my mother, thank you for your sacrifice as a parent in allowing me to pursue my dreams. I would not be where I am at today without your sacrifice. To my friends that I have met from childhood, high school, Illinois College, and Clemson University. Thank you for your endless support and for helping me build my confidence. Finally, to Dr. Zettler of Illinois College. Thank you for always believing in me in pursuing a journey in science.

TABLE OF CONTENTS

CHAPTER 1: Assessment of two statistical approaches for variance genome-wide association studies in plants [1]

1.1 ABSTRACT

Genomic loci that control the variance of agronomically important traits are increasingly important due to the profusion of unpredictable environments arising from climate change. The ability to identify such variance-controlling loci in association studies could assist future breeding efforts. Two statistical approaches that have already been used in the variance genome-wide association study (vGWAS) paradigm are the Brown-Forsythe test (BFT) and the double generalized linear model (DGLM). To ensure that these approaches are deployed as effectively as possible, it is critical to study the factors that influence their power to identify variance-controlling loci. We used genome-wide marker data in maize (*Zea mays* L.) and *Arabidopsis thaliana* to simulate traits controlled by epistasis, genotype by environment (GxE) interactions, and variance quantitative trait nucleotides (vQTNs). We then quantified true and false positive detection rates of the BFT and DGLM across all simulated traits. We also conducted a vGWAS using both the BFT and DGLM on plant height in a maize diversity panel. The observed true positive detection rates at the maximum sample size considered (N = 2,815) suggest that both of these vGWAS approaches are capable of identifying epistasic and GxE loci using for sufficiently large sample sizes. We also noted that the DGLM decisively outperformed the BFT for simulated traits controlled by vQTNs at sample sizes of N = 500. Although we conclude that there are still certain aspects of vGWAS approaches that need further refinement, this study suggests that the BFT and DGLM are capable of identifying variance-controlling loci in current state-of-the-art plant or agronomic data sets.

[1] Matthew D. Murphy, Samuel B. Fernandes, Gota Morota, Alexander E. Lipka
DOI: https://doi.org/10.1038/s41437-022-00541-1

1.2 Introduction

The world's food baskets face an expanding amount of unpredictable growing seasons due to the ongoing threat of climate change (Ziervogel and Erickson, 2010). If a 4°C increase in global temperature is not prevented by 2100, there could be a potential loss of $23 quadrillion to agriculture (Schillaci *et al.,* 2019). Unfortunately, most crops are maladapted to highly variable environments, where optimal growing conditions may never be attained (Mulder *et al.,* 2007). An idea that may lend itself to accelerating the development of crops better suited for such variable environments is canalization. Canalization is the hypothesis that natural selection minimizes variation for certain traits in a way that prevents major loci from being influenced significantly by the environment or background genetic variance like epistasis (Waddington, 1942; Rönnegård and Valdar, 2011). Artificial selection facilitates the decanalization of certain loci, which has allowed domesticated crops to grow in novel environments (Kitano, 2004). However, these decanalized loci are disadvantageous if the environment it was adapted to becomes unpredictable (Waddington, 1942). Collectively, the combination of decanalized loci and unpredictable environments has resulted in such loci controlling the variance of a targeted trait; that is, as a variance quantitative trait locus (vQTL; (Debat & David, 2001). A classic example of such vQTLs are genes that encode heat shock proteins, which are involved with various environmental stressors, including heat stress, ultraviolet radiation, cold tolerance, and biotic stressors (Park and Seo, 2015). Variance-controlling loci also arise from epistatic gene action, where the marginal effects of one of the epistatically interacting genes appear as a vQTL (see Forsberg and Carlborg, 2017 for a review). Recent advances in genome-wide association study (GWAS) approaches, such as that described in Li *et al.* (2022) make it possible to indirectly quantify the effects of variance-controlling loci through the inclusion of genotype by

environment (GxE) and epistatic (GxG) terms in the model. To complement these approaches, variance GWAS (vGWAS) has been proposed to directly quantify the effect of these variance-controlling loci (Shen *et al.,* 2012; Shen *et al.,* 2014; Yadav *et al.,* 2016).

The primary purpose of a vGWAS is to detect genetic loci that alter the variance of a phenotype between different genotypes (Al Kawan *et al.,* 2018). While vGWASs have been conducted in plants and crops, its utilization is still not widespread. To date, vGWASs have been conducted for ionomic traits, including molybdenum content in *Arabidopsis thaliana* and cadmium content in bread wheat (*Triticum aestivum)* (Shen *et al.,* 2012; Forsberg *et al.,* 2015; Hussain *et al.,* 2020), as well as for oil-related traits in maize (*Zea mays* L.) (Li *et al.,* 2020). Unlike those used in a standard GWAS (denoted as a mean GWAS or mGWAS), the statistical models used for a vGWAS specifically assume unequal phenotypic variance at each genotypic state of a given locus, i.e., in the presence of variance heterogeneity (Rönnegård and Valdar, 2011). Variance-controlling loci are connected to many different ideas within quantitative genetics, including epistatic and GxE interactions (Struchalin *et al.,* 2012; Rönnegård and Valdar, 2011). One potentially important advantage of using vGWAS for search for the presence of such interactions is it could prioritize genomic regions likely to harbor epistatic interactions, thereby reducing the severity of multiple testing correction (Struchalin *et al.,* 2012; Petterson and Carlborg, 2015). The markers in these regions could then be directly tested for the presence of epistasis or GxE interactions.

Many statistical analyses have been developed to test for variance heterogeneity. From a biological perspective, the choice of test and model can be divided into whether or not one accounts for population structure, relatedness, and other covariates (Rönnegård *et al.,* 2012). Of

the statistical tests that do not account for such factors, Levene's test and its median modification, the Brown-Forsythe test (BFT), have been the most popular (Brown and Forsythe, 1974; Rönnegård and Valdar, 2012). Although they are useful as a quick diagnostic for identifying variance-controlling loci, they cannot explicitly correct for population structure, familial relatedness, or loci that control the mean of a tested trait (called mean QTLs or mQTLs) (Hong *et al.*, 2017). In contrast, models that allow for the inclusion of these factors as covariates theoretically offer higher power to detect variance-controlling loci. In particular, the double generalized linear model (DGLM) (Lee and Nelder, 1996) adjusts for potential confounding between vQTLs, mQTLs, and population structure through the inclusion of fixed-effect covariates. Excitingly, more sophisticated versions of the DGLM also include random effects to account for confounding due to familial relatedness (Lee and Nelder, 2006; Rönnegård and Valdar, 2012).

Although statistical approaches seeking to estimate the effects of variance-controlling loci have opened up many opportunities for discovering new sources of quantitative trait variation, detecting variance-controlling loci still poses challenges. For example, the statistical power needed to detect a variance-controlling locus often requires five times as many individuals compared to the precision needed to detect a mean-controlling locus (Lee and Nelder, 2006; Ronnegard and Valdar, 2012). This suggests that there is a critical need to systematically study the statistical performance of leading vGWAS approaches. Therefore, the purpose of this study was to explore the factors that influence the ability of the BFT and DGLM to detect vQTLs underlying plant traits. We used publicly available whole-genome resequencing data from the 1,001 genomes diversity panel in *Arabidopsis thaliana* (Alonso-Blanco *et al.*, 2016) and the USDA-ARS North Central Region Plant Introduction Station (NCRPIS) Panel in *Zea mays* L.

(Romay *et al.,* 2013) to simulate traits controlled by epistasis, GxE effects, or variance-controlling loci with various effect sizes. We also explored the ability of these two approaches to find variance-controlling loci associated with the plant height data from Peiffer *et al.* (2014) that was measured in the Goodman maize diversity panel (Flint-Garcia *et al*., 2005).

1.3 Materials and Methods

Genotypic data and filtering procedures

We conducted simulation studies using genotypic data from two plant species with contrasting levels of linkage disequilibrium (LD) decay. The first genotypic data set was a subset of 1,087 accessions from the *Arabidopsis thaliana* 1,001 genomes diversity panel, available at https://1001genomes.org/ (Alonso-Blanco *et al.*, 2016). The 1,001 genomes diversity panel consists of germplasm mostly collected from Eurasia, North America, and Northern Africa. These accessions were genotyped using whole-genome resequencing, which produced 10,707,430 biallelic SNPs (Alonso-Blanco *et al.*, 2016). The second set of genotypic data consisted of 2,815 lines from the NCRPIS diversity panel in maize (Romay *et al.*, 2013). This diversity panel was genotyped for 681,257 SNPs, as described in Romay *et al.* (2013). This genotypic data set is publicly available at cbsusrv04.tc.cornell.edu/users/panzea/download.aspx?filegroupid=6.

Both genotypic data sets were filtered with VCFtools (Danecek *et al.*, 2011) to remove SNPs with more than 10% missing data or minor allele frequency (MAF) below 5%(Danecek et al., 2011). These data sets were then further filtered with LD pruning utilizing PLINK (Purcell *et al.*, 2007). The LD pruning parameters for Arabidopsis were set to r^2 = 0.10, a window size of 200 SNPs, and a step size of 20 SNPs. The LD pruning parameters for maize were loosely based

on the procedure done in Romay *et al*. (2013), which were r^2 = 0.2, a window size of 100 SNPs, and a step size of 25 SNPs. This filtering process described above was conducted independently in both species. The resulting number of SNPs was 41,384 for *Arabidopsis* and 72,359 for maize.

To assess how sample size affects the performance of the tested statistical methodologies, we considered two different sample size scenarios for each species. The first scenario focused on employing all individuals in both panels (i.e., *N* = 1,087 for *Arabidopsis* and *N* = 2,815 for maize). In the second scenario, we randomly selected *N* = 500 individuals from each panel using the sample() function in R (R Core Team 2021).

Simulation of traits controlled by variance- and mean- quantitative trait nucleotides

We developed the approach described below to simulate traits controlled by variance quantitative trait nucleotides (vQTNs) and/or mean QTNs (mQTNs). Each of these simulated traits consisted of a unique configuration of vQTNs, mQTNs, their effect sizes, and narrow-sense heritability. These parameters were used in the following formula derived from Hill and Mulder (2010) to obtain simulated trait values for each individual:

$$P_i = A_{m,i} + \chi_i k\left(\sigma_E + A_{v,SDi}\right), \quad (1.1)$$

where P_i is the simulated phenotypic value of the i^{th} individual, $A_{m,i}$ is the total genetic value from all simulated mean QTNs for the i^{th} individual, χ_i is a standard normal random variable (i.e., $N(\mu = 0, \sigma^2 = 1)$) sampled for the i^{th} individual, k is a constant described two paragraphs below that allows for a certain degree of control over the narrow-sense heritability, σ_E is the population standard deviation determined attributed to non-genetic sources, and $A_{v,SDi}$ is the collective genetic value of all simulated vQTN for the i^{th} individual. The values of $A_{m,i}$ and

$A_{v,SDi}$ are respectively calculated as the sum of the observed numeric genotype value at each mean and variance QTN, multiplied by the (respective) mean and variance QTN effects for the i^{th} individual. These simulations are conducted assuming that the covariance between $A_{m,i}$ and $A_{v,SDi}$ is zero.

One major challenge for simulating traits controlled by vQTNs is the specification of the desired heritability. Because the value of $(\sigma_E + A_{v,SDi})$ changes for every individual, the value of the heritability will also change for every individual. We, therefore, made *ad hoc* adjustments to Equation 1 to ensure at least partial control for a desired narrow-sense heritability (h^2). First, the value of σ_E was also set to 1, and then $A_{m,i}$ was centered and scaled, so its sample mean and standard deviation were respectively 0 and 1. These steps were taken to facilitate the estimation of the k in Equation 1.

We now describe the derivation of the procedure we used to estimate the value of k. Consider the following modified formula for estimating narrow-sense heritability h^2 for traits controlled by vQTNs:

$$\hat{h}^2 = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_A^2 + k^2\left(\hat{\sigma}_E + Mdn\{A_{v,SDi}\}\right)^2}, \quad (1.2)$$

where $\hat{\sigma}_A^2 = Var\{A_{mi}\} = 1$ because $A_{m,i}$ was scaled, $\hat{\sigma}_E$ was set equal to $\sigma_E = 1$ to facilitate calculations, and $Mdn\{A_{v,SDi}\}$ is the median value of $A_{v,SDi}$ across all n individuals (i.e., all individuals in either the *Arabidopsis* or maize data sets used for the simulations). Thus, solving Equation 2 for k yields:

$$k = \sqrt{\frac{\frac{\hat{\sigma}_A^2}{\hat{h}^2} - \hat{\sigma}_A^2}{\left(\hat{\sigma}_E + Mdn\{A_{v,SDi}\}\right)^2}}, \quad (1.3)$$

where all terms are as previously described. Thus, for each simulation setting, the value of k from Equation 3 was used in Equation 1 to obtain simulated trait values for every individual.

Two R functions were used to simulate these traits. Additive mean QTNs, which contribute to $A_{m,i}$ in Equation 1, were simulated using the create_phenotypes() function in the simplePHENOTYPES R package (Fernandes and Lipka, 2020). We then developed our custom R function that was roughly based on the Python code from Dumitrascu *et al.* (2019) to obtain the remaining necessary values in Equations 1, 2, and 3 to simulate the phenotypic values P_i. To facilitate the deployment of our simulation pipeline to future studies, we made it available through simplePHENOTYPES v1.4 (create_phenotypes(… , model = "V")) (https://github.com/samuelbfernandes/simplePHENOTYPES).

Description of all settings considered in simulation study

We conducted a comprehensive study that simulates traits controlled by either i.) no QTN, ii.) epistasis, iii.) GxE, or iv.) a combination of vQTN and mQTN (using the approach described in the previous section). Consequently, our simulation studies were subdivided into four respective scenarios summarized in Table 1. Across all scenarios, a total of 64 unique settings (i.e., combinations of input parameters) of traits were simulated. At each setting, a total of 100 replicate traits were simulated.

To enable a rigorous assessment of false positive rates of the tested vGWAS approaches, the "Null" scenario (as depicted on Table 1) consisted of traits with broad-sense heritability (H^2)

set to H^2= 0 and zero QTNs. Consistent with the hypothesis that epistasis is responsible for vQTLs (Forsberg and Carlborg, 2017), the "Epistasis" scenario simulated traits controlled by three epistatically interacting pairs of loci. For each pair, the epistatic effect was defined as the effect corresponding to the product of additively-encoded explanatory variables at each locus (i.e,. the additive-by-additive effect, i_{aa}, defined in Cordell, 2002). For each individual, the genetic values from each of the epistatically interacting loci were added up, and the resulting simulated trait value was the sum of these genetic values plus a normally distributed random variable with population mean 0 and population variance determined from the broad-sense heritability of the trait. To enable an assessment of the impact of heritability on the results, we kept the effect sizes of each of these epistatic QTN constant at 0.75, and the targeted MAF of all SNPs selected to be QTNs was 0.10. We then simulated traits at two different broad-sense heritabilities, namely H^2 = 0.3 and H^2 = 0.8.

For the "GxE" scenario we used the "partial pleiotropy" setting in simplePHENOTYPES (Fernandes and Lipka, 2020) to simulate one trait in two environments that was controlled by two environment-specific mQTNs. The first of these mQTNs was at the same randomly-selected marker for each environment, but had contrasting additive effect sizes, specifically 0.2 in the first environment (called Environment A) and 0.8 in the second environment (Environment B). The second of these mQTNs were at different randomly-selected markers for each environment, and was assigned an additive effect size of 0.5. All simulated QTNs had MAFs of approximately 0.3. The narrow-sense heritabilities of both traits were set at h^2 = 0.7. Upon completion of simulating this trait in two environments, each individual had two trait values: one from Environment A (Y_A), and one from Environment B (Y_B). However, a single phenotypic value was needed for each

individual for downstream analyses. Thus, for each individual we used the difference between trait values $Y_A - Y_B$ as the response variable in the subsequent statistical analysis.

Finally, we used the findings from previously published vGWAS and vQTL studies conducted in *Arabidopsis* and maize (Shen *et al.*, 2012; Li *et al.*, 2013; Forsberg *et al.*, 2015; Li *et al.*, 2020) as a basis for the "vQTN" scenario. Collectively, the various parameters we explored in this scenario (summarized in Table 1) enabled us to study the impact of narrow-sense heritability, MAF of vQTNs, and the effect sizes of vQTNs on the performance of the various GWAS approaches we explored. Detailed information about the actual SNPs that were randomly selected to be QTNs across all settings are presented in the provided GitHub repository.

Competing GWAS models and tests

We considered two different statistical approaches used in previous plant publications to conduct vGWAS, namely the BFT and the DGLM (Shen *et al.*, 2012; Forsberg *et al.*, 2015; Hussain *et al.*, 2020; Li *et al.*, 2020). In general, the BFT is used in vGWAS to test for variance homogeneity (Brown and Forsythe, 1974; Shen *et al.*, 2012). For each locus, the BFT evaluates:
H_0: $Population\ variances\ of\ traits\ are\ equal\ at\ all\ genotypes$ versus
H_a: $Population\ variances\ of\ traits\ are\ different\ for\ at\ least\ one\ genotype,$

and uses the corresponding test statistic:

$$F = \frac{(N-m)\Sigma_{j=1}^{m} n_j (y_{.j}^{*} - y_{..}^{*})^2}{(m-1)\Sigma_{j=1}^{m}\Sigma_{i=1}^{n_j} (y_{ij}^{*} - y_{.j}^{*})^2}, \quad (1.4)$$

where N is the total number of accessions, n_j is the number of accessions in the j^{th} genotypic group, m is the number of genotypes at the tested genetic marker, and

$$y_{ij}^{*} = |y_{ij} - \tilde{y}_j|. \qquad (1.5)$$

In (5), y_{ij} is the phenotypic value for the i^{th} individual with the j^{th} genotype and $\tilde{y}_j$ is the median phenotypic value of individuals with genotype j. Under H_0, the BFT statistic in (4) follows an *F* distribution with degrees of freedom equal to $m-1, N-m$ (Shen *et al.*, 2012). The BFT was performed using the brown.forsythe.test() function from the vGWAS R package (Shen *et al.*, 2012). Because the BFT does not allow explicit inclusion of covariates to account for false positives arising from population structure and familial relatedness, it often serves as a quick diagnostic test to see if the trait of interest has any underlying vQTLs. Furthermore, the BFT is robust to phenotypic departures from normality (Dumitrascu *et al.*, 2019; Hussain *et al.*, 2020).

The DGLM belongs to a family of generalized linear models, which relaxes the assumption of normality of phenotypic residuals for more flexible modeling. Specifically, the DGLM consists of two linear predictors that model the relationship between a response variable and i) explanatory variables controlling its population mean (Equation 6), and ii) explanatory variables controlling its population variance (Equation 7). The component of the DGLM controlling the population mean is written as follows:

$$Y_i = \mu_m + \Sigma_{k=1}^{q} X_{ik}\beta_k + s_{ij}a_{m_j} + \varepsilon_i \text{ , } (1.6)$$

where Y_i is the observed phenotypic value of the i^{th} individual, μ_m is the intercept; X_{ik} the value of the k^{th} principal component from a principal component analysis (PCA) of the markers (Price *et al.*, 2006) observed in the i^{th} individual (the first q = 4 and q = 3 principal components were included in the models used in *Arabidopsis* and maize, respectively); β_k is the regression

coefficient for the k^{th} principal component; s_{ij} is the value of the j^{th} SNP encoded as 0,1,2 for the i^{th} individual; a_{m_j} is the additive effect size of the j^{th} SNP; and $\varepsilon_i \sim N(0, \sigma^2_{\varepsilon_i})$. In the "vQTN" scenario presented in Table 1, s_j was set equal to the mQTN and a_{m_j} was its effect size; in all other settings, these two terms were omitted from the model because no mQTNs were simulated. The component of the DGLM controlling the population variance of the i^{th} individual $\sigma^2_{\varepsilon_i}$ is written as follows:

$$\log(\sigma^2_{\varepsilon_i}) = \mu_v + s_{ij} a_{v_j} \qquad (1.7)$$

where $\sigma^2_{\varepsilon_i}$ is the residual variance for the i^{th} individual; μ_v is the intercept; s_{ij} is the value of the observed SNP value encoded 0,1 and 2 at the j^{th} marker for the i^{th} individual; and a_{v_j} is the effect size of the j^{th} marker.

To test for a significant association between the j^{th} marker and the variance of the tested trait, we used the Wald test (Agresti, 2003) to test $H_0: a_{v_j} = 0$ versus $H_a: a_{v_j} \neq 0$, which follows an asymptotic χ^2_1 distribution under H_0. Thus, under H_0, the mean component of the DGLM remains as presented in Equation 6, while the component presented in Equation 7 is reduced to:

$$\log(\sigma^2_{\varepsilon_i}) = \mu_v \quad (1.8)$$

where all terms are as previously described.

As described in Corty and Valdar (2018), the DGLM framework is flexible in that it allows one to test for either the presence of a vQTN (i.e., test for $H_0: a_{v_j} = 0$, where a_{v_j} is described in Equation 7), presence of an mQTN (i.e., test for $H_0: a_{m_j} = 0$, where a_{m_j} is described in Equation 6), or for the presence of both (i.e., test for $H_0: a_{v_j} = 0 \; and \; a_{m_j} = 0$) at the j^{th} marker. For the sake of a direct comparison between the ability of the DGLM and the BFT to identify vQTNs, we assume that the user has already ran an *a priori* GWAS scan and that

any peak-associated mQTNs were fitted into the mean component of the DGLM, as presented in Equation (6). Thus, the multiple testing correction, described in detail in the next section, was applied equally to both the BFT and the DGLM. Because our analysis of DGLM is only testing for the presence vQTNs, the stringency of multiple testing will not be as severe as prior applications of the DGLM (e.g., Corty and Valdar, 2018) that tested for the presence of either, vQTNs, mQTNs, or both. To perform DGLM, we used the R code from Hussain *et al.* (2020), which came from the dglm R package (Dunn and Symth, 2020). The PCAs for population structure were obtained using GAPIT version 4.0 (Lipka *et al.*, 2012).

As a counterpoint to both the BFT and DGLM, we also conducted a GWAS at each replicate using a standard GWAS model. Specifically, we used GAPIT version 4.0 (Lipka *et al*., 2012) to fit the unified mixed linear model (MLM; Yu *et al.*, 2006) at each SNP and at each replicate trait considered in this study. Within each species, the same PCs that have been previously described were included in the model to account for subpopulation structure, and the method of VanRaden (2008) was used in the filtered marker sets in each species to obtain additive genetic relatedness (i.e., kinship) matrices to account for familial relatedness.

The ensuing analyses using both of these statistical approaches were conducted on a Dell Precision Tower 3240 with 64.0 GB RAM. While the BFT was ran on a single core, DGLM was ran on four cores using the foreach R package.

QTN detection rates for competing models

To assess whether or not the vGWAS methodologies can correctly identify markers as associated with our simulated traits, we evaluated the true and false positive QTN detection rates using the Benjamini and Hochberg (1995) procedure to control the false-discovery rate (FDR) at 5%. A

statistically significant SNP was labelled as a true positive if it was within a 250 kb window of a simulated vQTN for maize and within 100kb window of a simulated vQTN in *Arabidopsis*. Likewise, a statistically significant SNP was labelled as a false positive if it was outside of these windows. We defined the true positive rate as the proportion of times we detected at least one true positive per replication out of 100 replications. Similarly, the false positive rate is defined as the proportion of times we detected at least one false positive per replication out of 100 replications. True and false positive detection rates were further scrutinized by calculating 95% confidence intervals using the method of the Clopper and Pearson (1934) in the PropCIs R package (Scherer and Scherer, 2018). For each setting, all SNPs selected to be vQTNs were removed prior to calculating the true and false positive detection rates.

We also developed an approach similar to one presented in Gage *et al*. (2018) that used receiver operating characteristic (ROC) curves (Metz, 1978) to evaluate the ability of the three GWAS approaches to differentiate between true and false positives. For all settings except for those under the "Null" scenario, we randomly selected ten replicate traits. For each replicate trait, we used the genome-wide *P*-values from each of the three GWAS approaches to obtain corresponding ROC curves, where cases were considered to be all SNPs within the aforementioned physical windows of each QTN, and the remaining SNPs outside of these windows were considered to be controls. Thus, for a given replicate trait, a separate ROC curve was obtained for each of the three GWAS approaches. For each resulting ROC curve, we calculated the area under the ROC-curve (AUC); values of AUC greater than 0.5 suggest that the corresponding statistical model is capable of discriminating between cases and controls. Finally, for each GWAS approach used in each setting, we reported the median AUC value across the ten replicates.

Analysis of plant height data in a maize diversity panel

We performed a vGWAS using both the BFT and DGLM on plant height best linear unbiased predictors (BLUPs) from Peiffer *et al.* (2014). Briefly, this trait was measured on 279 individuals from the Goodman maize diversity panel (Flint-Garcia *et al.*, 2005) grown in ten different locations. To implicitly control for population structure and familial relatedness, we performed a two-step approach as described in previous vGWAS publications (Shen *et al.*, 2012; Forsberg *et al.*, 2015; Li *et al.*, 2020; Zhang and Qi, 2021). This approach first runs the unified MLM (Yu *et al*., 2006) with PCs (for this analysis, we used the first five PCs) and the VanRaden (2008) kinship matrix in TASSEL 5.0 (Bradbury *et al*., 2007) for the first step. The resulting residuals from this step were used as the response variable in our ensuing analyses. The genotypic data for this analysis consisted of a subset of 48,880 SNPs from the Illumina SNP50 chip (Cook *et al.*, 2012). For both the BFT and DGLM, we used the Benjamini and Hochberg (1995) to control for the genome-wide false discovery rate at 5%. To visualize the loci identified for the BFT, DGLM, and MLM, circular Manhattan plot was created using the Cmplots R package (Yin, 2018).

1.4 Results

False positive detection rates in the "Null" setting suggest BFT and DGLM adequately control for false positives

We ran the "Null" scenario to verify that the observed false positive rates for the BFT and DGLM were similar to what we would expect based on statistical theory (Figure 1.1). We also calculated 95% confidence intervals for these false positive rates using the method described by Clopper and Pearson (1934). All of these CIs contained the targeted FDR of 0.05.

High true positive detection rates were obtained for highly heritable epistatic QTNs, especially at larger sample sizes

When the heritability of the simulated epistatic QTNs were high (H^2 = 0.80), both the BFT and DGLM tended to detect SNP pairs contributing to epistatic QTNs for all combinations of species and sample sizes, although these true positive detection rates notably lower for maize at N = 500 (Figure 1.2). In contrast, both vGWAS approaches yielded extremely low detection rates of the pairs of SNPs contributing to the epistatic QTNs when the broad-sense heritability of the epistatic QTNs was low (H^2= 0.30). In general, both vGWAS approaches tended to detect the epistatic QTNs at similar rates in *Arabidopsis*, while the DGLM tended to yield either similar (at N = 2,815) or higher (at N = 500) true positive detection rates than the BFT in maize. The epistatic QTNs were identified by the MLM at relatively consistent high rates only when they were simulated in maize with sample size N = 2,815 and heritability of H^2 = 0.8. The results from the analysis of the ROC curves and corresponding median AUC values (Table 12.) support the findings presented in Figure 1.2. Thus, these results suggest that the two tested vGWAS approaches are capable of detecting pairwise epistasis, but these epistatic signals need to be highly heritable.

vGWAS approaches yielded high true positive detection rates of GxE signals only at the largest evaluated sample size

The BFT and DGLM could detect true positive signals from simulated GxE effects at non-negligible rates at only the largest sample size we evaluated, namely at N = 2,815 in maize (Figure 1.3). At this sample size, both of these approaches yielded similar true detection rates at the QTN that was simulated at the same genomic position, but with different effect sizes, in both

environments. However, at the two environment-specific QTNs, the BFT yielded higher true positive detection rates than the DGLM (Figure 1.3 D). At this sample size in maize, we also observed that the BFT detected the QTNs at rates either greater than or similar to those from the MLM. The results from the ROC curves and corresponding median AUC values were consistent with the true positive rates presented in Figure 1.3, especially with respect to noticeably higher median AUC values in maize at N =2,815 (Table 1.2.). Collectively, these results suggest that the BFT and DGLM are capable of identifying a GxE signal at reasonably high detection rates, but a large sample size (at least N = 2,815) is needed.

DGLM was capable of identifying vQTNs at smaller sample sizes of N = 500

Across both of the evaluated species and narrow-sense heritabilities, we observed that the true positive detection rates of the DGLM tended to monotonically increase with the effect sizes of vQTNs, particularly for those with MAFs of approximately 0.4 (Figure 1.4). This trend was observed for such vQTNs across all of the evaluated sample sizes. In contrast, the BFT consistently yielded low true positive QTN detection rates at N = 500. Although not as pronounced as the DGLM, we also observed that the true positive detection rates of the BFT tended to monotonically increase with vQTN effect sizes at certain settings. While approximately similar trends in true positive detection rates were observed in maize across the two vQTN MAF settings, notably lower true positive detection rates were observed for both vGWAS approaches in *Arabidopsis* for vQTNs with MAFs of approximately 0.1. We also observed that the DGLM results were more consistent across the two evaluated narrow-sense heritabilities than the BFT. As expected, the simulated vQTLs were not detected by the MLM, which makes sense considering the MLM assumes that the variances between genotypic groups

are equal. Similar trends were noted in the analysis of the ROC curves (Table 1.2) in particular with median AUCs for the BFT and DGLM tending to monotonically increase with sample size, MAF, and vQTN effect size. Overall, these results suggest that the DGLM is capable of outperforming the BFT at sample sizes of $N = 500$.

BFT and DGLM identified significantly associated markers for plant height

The BFT and DGLM both identified statistically significant associations for plant height in the Goodman diversity panel at a genome-wide FDR of 5% (Figure 1.5a), while no statistically significant associations were found using the unified MLM. Interestingly, three statistically significant associations located on chromosomes 1, 2, and 8 were identified by both the BFT and DGLM (Figure 1.5a). The quantile-quantile plots presented Figure 1.5b suggest that the –log(*P*-values) from the DGLM are more inflated than those from the BFT and MLM.

1.5 Discussion

We used both simulated and real traits to evaluate the ability of two vGWAS approaches, namely the BFT and the DGLM, to identify epistasis, GxE, and variance-controlling loci. At the maximum sample size evaluated ($N = 2{,}815$), both vGWAS approaches frequently identified highly heritable epistatic and GxE signals. For simulated traits that were controlled by vQTNs, we observed that the DGLM yielded substantially higher true positive detection rates than the BFT at sample sizes of $N = 500$. Collectively, these results provide a potential benchmark for how the BFT and DGLM are expected to perform when deployed to vGWAS in plants. Such an assessment is essential because the more widespread use of vGWAS in plants could facilitate selection for uniformity of trait values across various environmental conditions.

Prospects on the ability of BFT and DGLM to assist in identifying epistasis and GxE

Because of strong evidence in the literature that both epistasis and GxE could underlie the statistical associations identified in vGWAS (Struchalin *et al.*, 2012; Rönnegård and Valdar 2011), two of our simulation scenarios explicitly simulated these two sources of variability. The results from these two scenarios suggest that the DGLM and BFT are capable of finding highly heritable epistatic and GxE signals at the largest evaluated sample size (N = 2,815 in maize). An even more exciting result was that for the highly heritable epistatic QTNs at sample sizes of N = 500, both the vGWAS approaches yielded high detection rates in *Arabidopsis*, while the DGLM yielded modest to high true positive detection rates in maize.

The fact that we were able to identify these epistatic and GxE loci suggest that vGWAS approaches could assist in the detection of epistasis or GxE effects underlying agronomically important traits. As described in Struchalin *et al.* (2012), the large number of possible interacting loci to be tested when searching for epistasis or GxE results in a heavy multiple testing correction burden. To overcome this, a preliminary vGWAS scan could be conducted to highlight specific genomic markers likely to harbor these sources of genomic variability. Given our results, we expect vGWAS approaches to be successful in identifying highly heritable epistatic and GxE effects for sample sizes of at least N = 2,815. If such loci were to be detected using vGWAS studies, they can then be directly tested for epistasis and/or GxE effects in a follow-up analysis, where the multiple testing correction would be substantially reduced because only the markers identified using vGWAS are analyzed.

Prospects on the ability of the BFT and DGLM to identify vQTN

One consistent result we observed in both species was that the DGLM yielded higher true positive detection rates than the BFT at sample sizes of N =500 and MAF = 0.4. This suggests

that a sample size of 500 could be sufficient for the DGLM to identify common variance-controlling loci. In addition to varying the sample size and species in the "vQTN" scenario, we also evaluated the performance of these two vGWAS approaches across two targeted vQTN MAFs, effect size, and narrow-sense heritabilities. Not surprisingly, we observed that the true positive vQTN detection rates of both models tended to increase monotonically with their simulated effect sizes. We also noted that higher true positive detection rates tended to be observed for vQTNs with the higher targeted MAF = 0.4, which again was consistent with our expectations prior to conducting this study. However, higher than expected true positive detections were observed in maize for vQTNs with targeted MAF = 0.10. Taken together with the less favorable true positive detection rates in *Arabidopsis* for vQTNs with targeted MAF = 0.10, these results suggest that vGWAS could be used to identify genomic regions likely to harbor rare vQTNs under certain circumstances. Therefore, we recommend that future studies investigate the impact of LD decay and marker technologies on the ability to identify rare vQTNs. Although there were certain settings at h^2 = 0.63 where the BFT outperformed the DGLM, the latter approach yielded more stable true positive detection rates across the two evaluated narrow-sense heritabilities. This result suggests that the DGLM is more robust than the BFT for controlling the influence of the simulated mQTN on the overall simulated trait variance, and further underscores our recommendation of the DGLM as the preferred vGWAS approach (Lipka et al., 2015).

Limitations of our simulation studies

Although useful for simulating traits with similar genetic architectures of real traits, the approach we implemented to account for the narrow-sense heritability in the "vQTN" scenario was *ad hoc*. We recommend that future studies focus on accounting for broad-sense heritabilities, as this

would enable more user-control over the total phenotypic variance attributable to genetic effects. Another limitation of our study is that we explored only one configuration of modeling the relationship between vQTNs and a trait. Specifically, the configuration we used in (1) is based on the standard deviation additive model (Hill and Zhang, 2004; Hill and Mulder, 2010). Other vQTN quantitative genetics models, such as reaction norm model from Hill and Mulder (2010) or the other forms of epistasis discussed in Cordell (2002), could be used to simulate more scenarios where vQTNs could arise.

Areas for future research

While the BFT and DGLM are two commonly used statistical methodologies for vGWAS, the results from our analysis of plant height in maize suggest that they may not adequately control for population structure and familial relatedness (Figure 1.5b). Thus, we recommend the consideration of more sophisticated statistical approaches for vGWAS. Two examples are the hierarchical generalized linear model (HGLM) (Lee and Nelder, 1996) and double hierarchical generalized linear model (DHGLM) (Lee and Nelder, 2006). These models account for familial relatedness by including the individuals as a random effect and setting their variance-covariance to be proportional to an additive genetic relatedness matrix. Although the associated computational complexity of fitting these two models rendered them impractical to evaluate in our simulation studies, they have been previously evaluated in wheat (Hussain *et al.*, 2020) and animal breeding (Rönnegård *et al*., 2010). Given that the DGLM and HGLM in Hussain et al. (2020) both identified the same loci associated with cadmium content in wheat, we would expect that the HGLM and DHGLM to yield similar true positive detection rates for traits that are not associated with familial relatedness. We recommend that future work focuses on increasing the

computational efficiency of the HGLM and DHGLM so that their ability to identify vQTNs could be studied in a manner similar to that which is presented in this work.

One noteworthy aspect of several prior vGWAS investigations is that the search for variance-controlling loci was conducted separately from a mean GWAS scan (e.g Hussein *et al.*, 2020; Li *et al.*,2020; Córdova-Palomera *et al.*, 2020). However, Corty and Valdar (2018) demonstrated that models like the DGLM can be used in a single GWAS to test for associations with the mean of a trait, the variance of a trait, or both. Although this results in a 2x- to 3x-increase in the severity of the multiple testing burden (Corty and Valdar, 2018), the use of models like the DGLM to search for both mean and/or variance-controlling loci in a single GWAS scan is advantageous because it reduces the possibility of, for example, not identifying a mean-controlling locus because only a vGWAS scan was conducted. We therefore encourage future vGWAS studies to use models like the DGLM to their fullest extent by testing for mean-controlling loci in addition to variance-controlling loci.

The inbreeding species considered in our simulation studies, *Arabidopsis*, has not been subjected to as much artificial selection compared to what would be expected in crops (Izawa, 2007; Woodward and Bartel, 2018). Thus, future simulations should consider an inbreeding crop species such as sorghum (*Sorghum bicolor* L. Moench) or rice (*Oryza sativa* L). Additionally, our decision to simulate traits similar to those where vQTN have already been identified resulted in our simulated traits resembling metabolic traits with tractable genetic architectures. However, a recent report in maize showed that vQTNs are also present in plant architectural and phenology traits (Zhang and Qi, 2021). Thus, the practicality and utility of the BFT and DGLM to identify variance-controlling loci in crops could be more comprehensively explored if a wider range of genetic architectures were studied in future work. Given that our simulation studies showed that

vGWAS approaches could identify sets of markers likely to harbor epistatic and GxE effects, it is imperative to explore the most effective use of vGWAS to identify the most important candidate loci so that emerging approaches such as that presented in Li *et al.* (2022) can quantify their main-, interaction-, and GxE effects as accurately as possible. Finally, peak-associated markers from a vGWAS could be used for breeding applications. For example, if one wants to constrain the range of possible trait values in a challenging environment, they could potentially select on alleles from peak-associated markers in a vGWAS that reduces the variance of that trait.

1.6 Conclusion

The ability of vGWAS approaches to identify variance-controlling loci needs to be thoroughly scrutinized before they can become more commonplace in quantitative genetics analysis in plants. We conclude that DGLM is preferred over the BFT for practical use in plant vGWAS because of its observed performance at sample sizes of $N = 500$. The ability of both vGWAS approaches to identify epistasis and GxE is encouraging, and future simulation studies should focus on other quantitative genetics parameters and additional statistical models in more plant species. To facilitate such future exploration of vGWAS approaches, the computational approaches we used to simulate traits controlled by vQTN are now publicly available free of charge in the simplePHENOTYPES R package (Fernandes and Lipka, 2020).

Data Archiving

The genotypic data, simulated trait data, ROC curves, and code to simulate traits are available at https://github.com/mdm10-code/vGWAS_arabidopsis_maize .

1.7 Figures and Tables

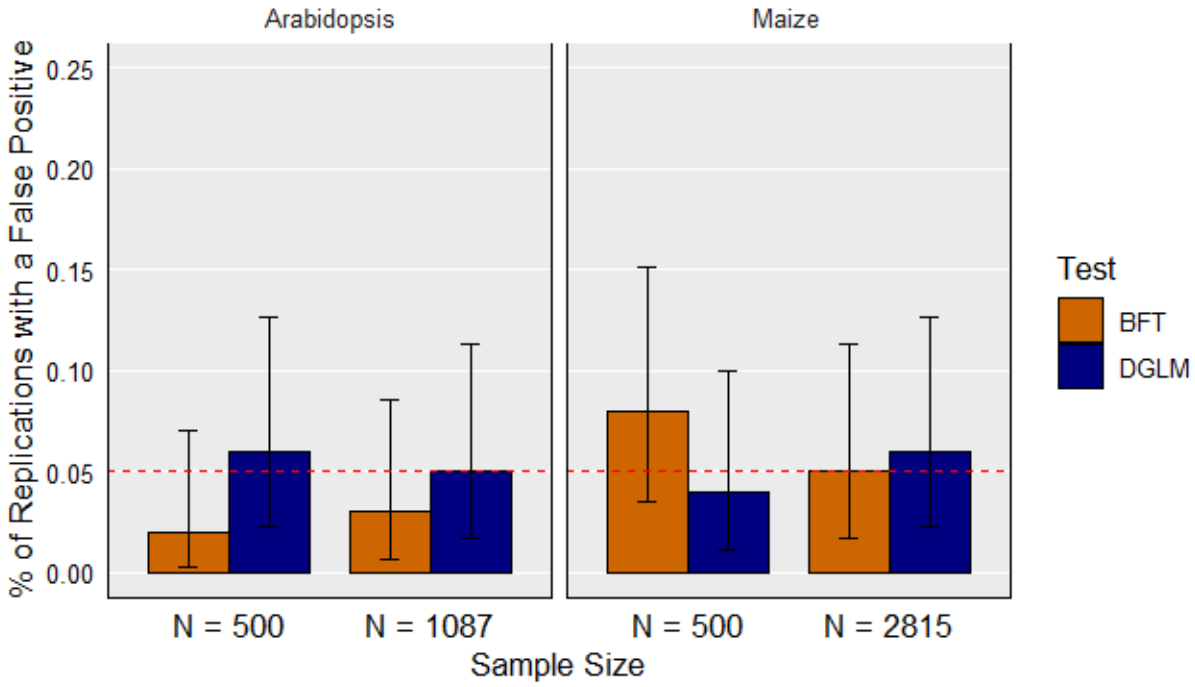

Figure 1.1. False positive detection rates for the null setting at a false discovery rate of 0.05. The X-axis represents the sample size of each diversity panel. The Y-axis is the proportion of replications where a false positive is detected at least once. The error bars depict 95% confidence intervals, calculated using the method from Clopper and Pearson (1934). The dotted red horizontal line depicts the targeted false discovery rate of 0.05. Each panel represents the species indicated in the title. BFT: Brown-Forsythe test; DGLM: double generalized linear model

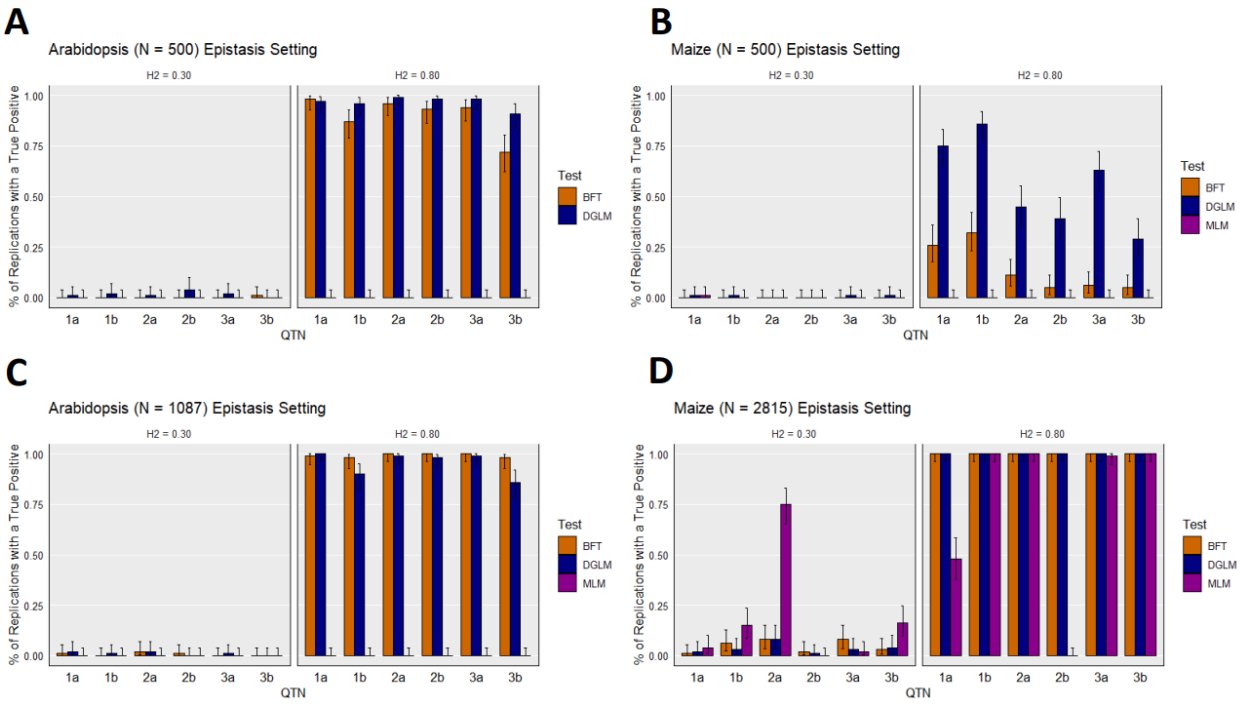

Figure 1.2. True detection rates for the simulated traits under the "epistasis" scenario settings at a false discovery rate of 0.05. (A) *Arabidopsis* at a sample size of N = 500, (B) maize at a sample size of N = 500, (C) *Arabidopsis* at a sample size of N = 1,087, and (D) maize at a sample size of N = 2,815. In each panel, two sets results are presented: one for the simulated trait with broad-sense heritability of set to H^2 = 0.30, and one for those with H^2 = 0.80. On each figure, the X-axis represents the pair of SNPs contributing to the three epistatic quantitative trait nucleotides (QTNs; e.g., "2a" denote the first SNP contributing to the second epistatic QTN). The Y-axis is the proportion of replications where a true positive is detected at least once. The error bars depict 95% confidence intervals, calculated using the method from Clopper and Pearson (1934). BFT: Brown-Forsythe test; DGLM: double generalized linear model; MLM: unified mixed linear model.

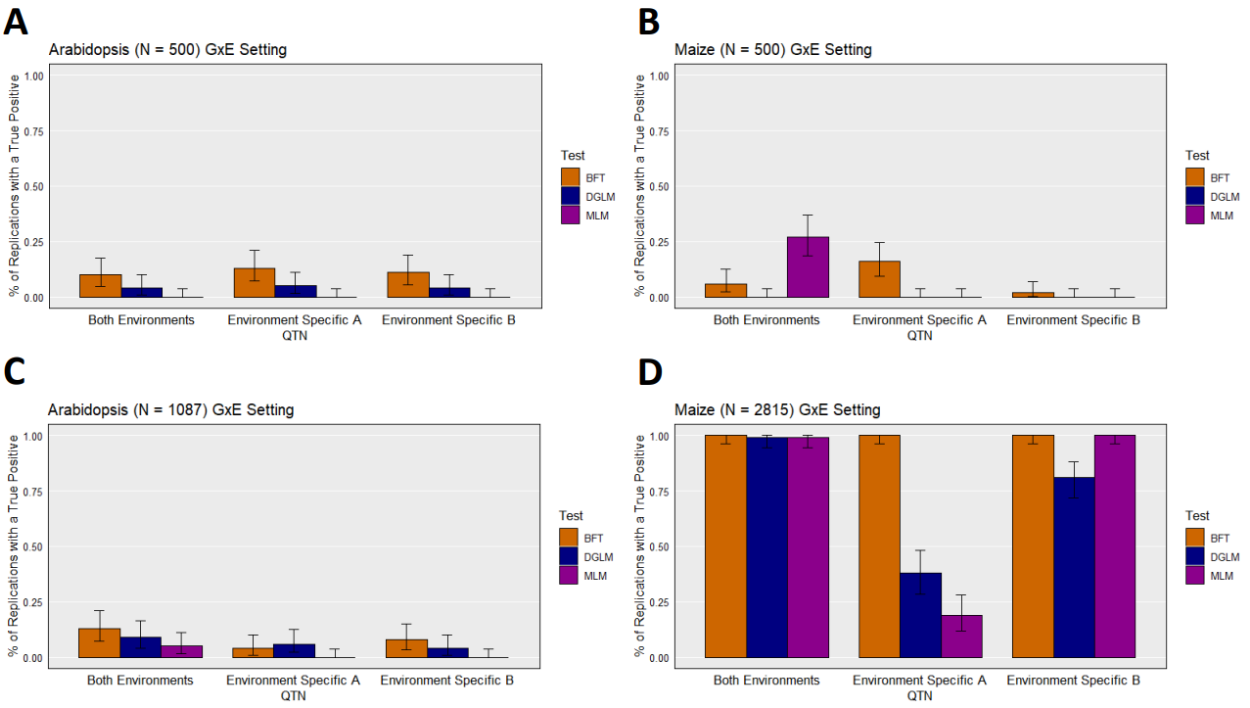

Figure 1.3. True detection rates for the simulated traits under the "GxE" scenario settings at a false discovery rate of 0.05. (A) *Arabidopsis* at a sample size of N = 500, (B) maize at a sample size of N = 500, (C) *Arabidopsis* at a sample size of N = 1,087, and (D) maize at a sample size of N = 2,815. On each figure, the X-axis represents the QTN, and a description of the which environment(s) in which they were simulated are detailed in the corresponding X-coordinate label. The Y-axis is the proportion of replications where a true positive is detected at least once. The error bars depict 95% confidence intervals, calculated using the method from Clopper and Pearson (1934). BFT: Brown-Forsythe test; DGLM: double generalized linear model; MLM: unified mixed linear model.

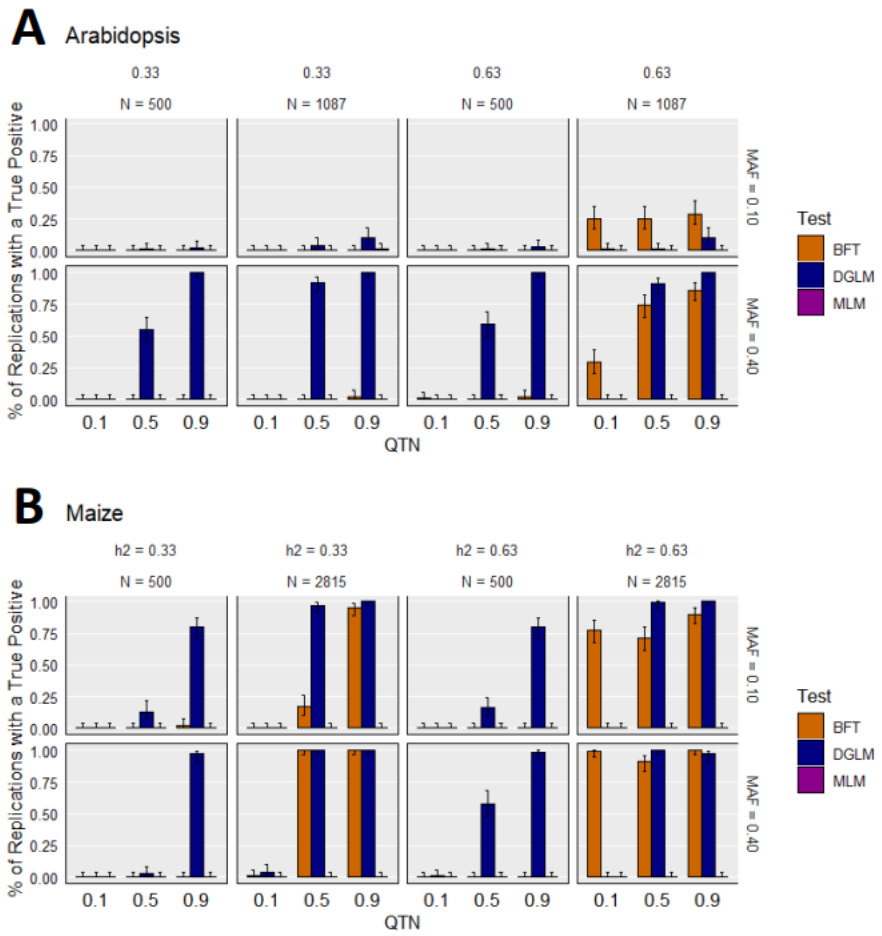

Figure 1.4. True detection rates for the simulated traits under the "vQTN" scenario settings at a false discovery rate of 0.05. (A) *Arabidopsis* and (B) maize. On each panel, the results are subdivided into heritability (h^2), sample size (*N*), and targeted minor allele frequency of the QTN (MAF). On each figure, the X-axis depicts the effect size of the vQTN, and Y-axis is the proportion of replications where a true positive is detected at least once. The error bars depict 95% confidence intervals, calculated using the method from Clopper and Pearson (1934). BFT: Brown-Forsythe test; DGLM: double generalized linear model; MLM: unified mixed linear model.

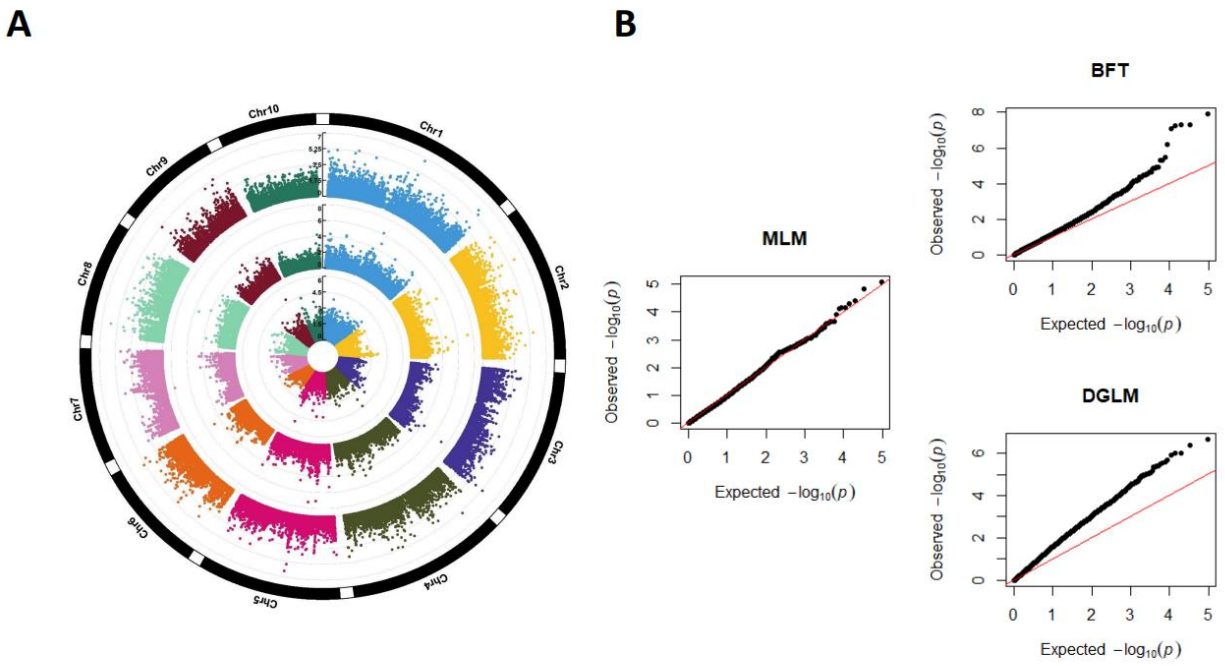

Figure 1.5. Genome-wide association study (GWAS) of plant height in the Goodman maize diversity panel. (A) Circular Manhattan plots summarizing the results from the unified mixed linear model (innermost circle), the Brown-Forsythe test (middle circle), and the double generalized linear model (outermost circle). The X-axis is the physical position of the SNPs along the maize genome, and the Y-axis is the –log(*P*-values) from each of the three GWAS models. (B) Quantile-quantile (Q-Q) plots showing the expected –log(*P*-values) under H_0: No association at tested marker on the X-axis and the observed –log(*P*-values) on the Y-axis. MLM: unified mixed linear model (middle column); BFT: Brown-Forsythe test (top-right column); DGLM: double generalized linear model (bottom-right column).

Table 1.1. Parameters considered in the simulation study. Across all scenarios, a total of 64 unique configurations of these parameters were simulated.

dScenario	nQTN[a]	Species	Heritability	Targeted MAF[b]	Sample Size	Allelic Effect
Null	0	Maize; *Arabidopsis*	0.00	N/A	500; Maximum	N/A
Epistasis[c]	3 epistatic QTNs[c]	Maize; *Arabidopsis*	0.30; 0.80	0.10	500; Maximum	0.75 for all 3 QTNs;
GxE[d]	2 mQTNs[e] in 2 environments	Maize; *Arabidopsis*	0.70	0.3	500; Maximum	Env.1: 0.2[f], 0.5[g] Env.2: 0.8[f], 0.5[g]
vQTN	1 mQTN 1 vQTN[h]	Maize; *Arabidopsis*	0.63; 0.33	0.10; 0.40	500; Maximum	mQTN: 0.25 vQTN: 0.1; 0.5; 0.9

[a]QTN – quantitative trait nucleotides; [b]MAF – minor allele frequency; [c]Each epistatic effect consists of the additive-by-additive effect of two randomly selected single nucleotide polymorphisms; [d]GxE – genotype by environment; [e]mQTN – mean QTNs; [f]This environment-specific mQTN was simulated at the SNP in both environments, but the effect sizes differed in each environment; [g]This environment-specific mQTN was simulated at different SNPs in each environment; [h]vQTN – variance QTNs.

Table 1.2. Median AUC values of all the 32 simulation settings

Species	Setting	Sample Size	MAF	Heritability	Effect Size	SNP Codename	BFT	DGLM	MLM
Arabidopsis	vQTN	500	0.1	0.33	0.1	NA	52.57297	52.08513	52.29831
Arabidopsis	vQTN	FULL	0.1	0.33	0.1	NA	52.30901	50.4743	52.13734
Arabidopsis	vQTN	500	0.1	0.33	0.5	NA	51.31866	53.40312	52.81646
Arabidopsis	vQTN	FULL	0.1	0.33	0.5	NA	52.25515	53.35654	52.04732
Arabidopsis	vQTN	500	0.1	0.33	0.9	NA	52.51251	53.89955	53.20752
Arabidopsis	vQTN	FULL	0.1	0.33	0.9	NA	52.7063	51.95754	51.86568
Arabidopsis	vQTN	500	0.1	0.63	0.1	NA	51.86509	52.31931	51.27516
Arabidopsis	vQTN	FULL	0.1	0.63	0.1	NA	52.12818	52.48105	52.12531
Arabidopsis	vQTN	500	0.1	0.63	0.5	NA	49.99569	53.76465	51.08654
Arabidopsis	vQTN	FULL	0.1	0.63	0.5	NA	52.20029	51.6923	51.82426
Arabidopsis	VQTN	500	0.1	0.63	0.9	NA	53.21822	53.43396	52.01721
Arabidopsis	VQTN	FULL	0.1	0.63	0.9	NA	52.63831	53.44069	51.50445
Arabidopsis	VQTN	500	0.4	0.33	0.1	NA	52.14024	52.34366	53.80448
Arabidopsis	VQTN	FULL	0.4	0.33	0.1	NA	52.18126	52.5446	52.27187

Table 1.2 (cont)

Arabidopsis	VQTN	500	0.4	0.33	0.5	NA	52.85383	52.72398	50.73699
Arabidopsis	VQTN	FULL	0.4	0.33	0.5	NA	53.68162	56.09933	52.21413
Arabidopsis	VQTN	500	0.4	0.33	0.9	NA	52.53524	53.24144	51.94996
Arabidopsis	VQTN	FULL	0.4	0.33	0.9	NA	53.90186	57.81029	50.09359
Arabidopsis	VQTN	500	0.4	0.63	0.1	NA	52.56871	51.52208	51.99103
Arabidopsis	VQTN	FULL	0.4	0.63	0.1	NA	52.15106	52.94764	52.21669
Arabidopsis	VQTN	500	0.4	0.63	0.5	NA	53.36255	55.17999	53.46859
Arabidopsis	VQTN	FULL	0.4	0.63	0.5	NA	52.85554	54.28068	51.78579
Arabidopsis	VQTN	500	0.4	0.63	0.9	NA	51.80207	55.11323	54.00815
Arabidopsis	VQTN	FULL	0.4	0.63	0.9	NA	51.83172	57.73794	53.0169
Maize	VQTN	500	0.1	0.33	0.1	NA	59.73439	60.74553	56.60283
Maize	VQTN	FULL	0.1	0.33	0.1	NA	55.58641	58.10417	55.20203
Maize	VQTN	500	0.1	0.33	0.5	NA	52.87468	50.14665	58.99188
Maize	VQTN	FULL	0.1	0.33	0.5	NA	57.36674	58.13135	52.31872
Maize	VQTN	500	0.1	0.33	0.9	NA	54.82983	56.12777	57.72731
Maize	VQTN	FULL	0.1	0.33	0.9	NA	56.4764	59.4778	58.01369
Maize	VQTN	500	0.1	0.63	0.1	NA	59.41097	53.78255	63.22778
Maize	VQTN	500	0.1	0.63	0.5	NA	53.14765	57.44081	58.23357
Maize	VQTN	FULL	0.1	0.63	0.5	NA	61.19516	54.58347	54.93808
Maize	VQTN	500	0.1	0.63	0.9	NA	57.86849	53.73408	59.62115
Maize	VQTN	FULL	0.1	0.63	0.9	NA	58.19977	60.23577	55.08327
Maize	VQTN	500	0.4	0.33	0.1	NA	51.83962	53.05837	52.16444
Maize	VQTN	FULL	0.4	0.33	0.1	NA	52.58255	59.40607	53.10769
Maize	VQTN	500	0.4	0.33	0.5	NA	53.80868	53.0206	51.36311
Maize	VQTN	FULL	0.4	0.33	0.5	NA	72.29017	68.92994	54.30279
Maize	VQTN	500	0.4	0.33	0.9	NA	52.50509	57.68923	52.09022

Table 1.2 (cont)

Maize	VQTN	FULL	0.4	0.33	0.9	NA	77.05214	67.68449	51.97274
Maize	VQTN	500	0.4	0.63	0.1	NA	52.36527	50.73433	52.57057
Maize	VQTN	500	0.4	0.63	0.5	NA	51.03006	54.57586	52.57745
Maize	VQTN	FULL	0.4	0.63	0.5	NA	58.83756	68.79571	52.67703
Maize	VQTN	500	0.4	0.63	0.9	NA	52.69184	56.04058	53.09614
Maize	VQTN	FULL	0.4	0.63	0.9	NA	70.89146	71.534	53.4196
Maize	VQTN	FULL	0.1	0.63	0.1	NA	58.3733	56.31001	54.94959
Maize	VQTN	FULL	0.4	0.63	0.1	NA	53.40596	55.4274	55.10461
Arabidopsis	Epistasis	500	0.1	0.3	0.75	1a	50.1331	50.45479	50.21667
Arabidopsis	Epistasis	500	0.1	0.3	0.75	1b	50.72499	50.4195	50.22026
Arabidopsis	Epistasis	500	0.1	0.3	0.75	2a	50.33601	50.56981	50.44165
Arabidopsis	Epistasis	500	0.1	0.3	0.75	2b	50.36014	50.31238	50.18411
Arabidopsis	Epistasis	500	0.1	0.3	0.75	3a	50.88781	50.37521	50.18854
Arabidopsis	Epistasis	500	0.1	0.3	0.75	3b	50.36022	50.70047	50.20294
Arabidopsis	Epistasis	FULL	0.1	0.3	0.75	1a	50.14821	50.44895	50.39826
Arabidopsis	Epistasis	FULL	0.1	0.3	0.75	1b	50.59707	50.56744	50.67394
Arabidopsis	Epistasis	FULL	0.1	0.3	0.75	2a	50.52063	50.7176	50.80671
Arabidopsis	Epistasis	FULL	0.1	0.3	0.75	2b	50.40066	50.26271	50.73977
Arabidopsis	Epistasis	FULL	0.1	0.3	0.75	3a	50.19144	50.67595	50.36948
Arabidopsis	Epistasis	FULL	0.1	0.3	0.75	3b	50.08748	50.50466	50.80671
Arabidopsis	Epistasis	500	0.1	0.8	0.75	1a	50.54266	50.54514	50.17167
Arabidopsis	Epistasis	500	0.1	0.8	0.75	1b	50.40582	50.43273	50.45957
Arabidopsis	Epistasis	500	0.1	0.8	0.75	2a	50.3749	50.2822	50.57618
Arabidopsis	Epistasis	500	0.1	0.8	0.75	2b	50.29373	50.53696	50.1615
Arabidopsis	Epistasis	500	0.1	0.8	0.75	3a	50.67487	50.40837	50.21739
Arabidopsis	Epistasis	500	0.1	0.8	0.75	3b	50.29226	50.30035	50.24398

Table 1.2 (cont)

Arabidopsis	Epistasis	FULL	0.1	0.8	0.75	1a	50.52361	50.17984	50.47854
Arabidopsis	Epistasis	FULL	0.1	0.8	0.75	1b	50.40726	50.22014	50.47854
Arabidopsis	Epistasis	FULL	0.1	0.8	0.75	2a	50.11095	50.91324	50.28623
Arabidopsis	Epistasis	FULL	0.1	0.8	0.75	2b	50.41943	50.55049	50.28623
Arabidopsis	Epistasis	FULL	0.1	0.8	0.75	3a	50.39265	50.75887	50.28623
Arabidopsis	Epistasis	FULL	0.1	0.8	0.75	3b	50.52202	50.51846	50.4247
Arabidopsis	GxE	500	0.3	0.7	NA	Both Environments	50.09205	50.3165	50.35181
Arabidopsis	GxE	500	0.3	0.7	NA	Environment Specific A	50.18458	50.36237	50.26638
Arabidopsis	GxE	500	0.3	0.7	NA	Environment Specific B	50.20492	50.39801	50.26673
Arabidopsis	GxE	FULL	0.3	0.7	NA	Both Environments	50.4035	50.61249	50.99902
Arabidopsis	GxE	FULL	0.3	0.7	NA	Environment Specific A	50.62795	50.45034	50.82828
Arabidopsis	GxE	FULL	0.3	0.7	NA	Environment Specific B	50.81268	50.43116	51.30795
Maize	Epistasis	FULL	0.1	0.3	0.75	1a	52.32864	50.25778	50.65048
Maize	Epistasis	FULL	0.1	0.3	0.75	1b	52.2307	50.57947	50.84308
Maize	Epistasis	FULL	0.1	0.3	0.75	2a	52.57656	50.74501	50.52855
Maize	Epistasis	FULL	0.1	0.3	0.75	2b	52.2667	50.53097	50.65048
Maize	Epistasis	FULL	0.1	0.3	0.75	3a	52.49987	50.46516	50.86274
Maize	Epistasis	FULL	0.1	0.3	0.75	3b	52.06507	50.65302	50.84308
Maize	Epistasis	FULL	0.1	0.8	0.75	1a	54.30695	51.62179	51.10159
Maize	Epistasis	FULL	0.1	0.8	0.75	1b	53.84306	52.023	51.40015
Maize	Epistasis	FULL	0.1	0.8	0.75	2a	53.71757	51.59016	51.40015
Maize	Epistasis	FULL	0.1	0.8	0.75	2b	53.78033	52.00053	51.10159
Maize	Epistasis	FULL	0.1	0.8	0.75	3a	53.732	51.97243	51.32507
Maize	Epistasis	FULL	0.1	0.8	0.75	3b	53.72608	51.70764	51.40015

Table 1.2 (cont)

Maize	GxE	FULL	0.3	0.7	NA	Both Environments	56.459	50.35084	53.08873
Maize	GxE	FULL	0.3	0.7	NA	Environment Specific A	55.98505	50.66707	52.9836
Maize	GxE	FULL	0.3	0.7	NA	Environment Specific B	56.90893	50.36751	53.04125
Maize	GxE	500	0.3	0.7	NA	Both Environments	50.29954	51.15671	51.5471
Maize	GxE	500	0.3	0.7	NA	Environment Specific A	50.63374	50.97805	52.24543
Maize	GxE	500	0.3	0.7	NA	Environment Specific B	50.63556	50.91966	51.67165
Maize	Epistasis	500	0.1	0.3	0.75	1a	50.58567	50.43186	50.37853
Maize	Epistasis	500	0.1	0.3	0.75	1b	50.85551	50.05869	50.81581
Maize	Epistasis	500	0.1	0.3	0.75	2a	50.42984	50.31396	51.00688
Maize	Epistasis	500	0.1	0.3	0.75	2b	50.66408	50.69367	50.31918
Maize	Epistasis	500	0.1	0.3	0.75	3a	50.48336	50.21711	50.47427
Maize	Epistasis	500	0.1	0.3	0.75	3b	50.38972	50.63208	50.48528
Maize	Epistasis	500	0.1	0.8	0.75	1a	51.22772	50.8902	50.70673
Maize	Epistasis	500	0.1	0.8	0.75	1b	50.98016	50.56267	50.67238
Maize	Epistasis	500	0.1	0.8	0.75	2a	50.75863	50.69719	50.67238
Maize	Epistasis	500	0.1	0.8	0.75	2b	50.65302	50.35186	50.70673
Maize	Epistasis	500	0.1	0.8	0.75	3a	51.35779	50.95579	50.95794
Maize	Epistasis	500	0.1	0.8	0.75	3b	52.54604	50.61178	51.15892

MAF – minor allele frequency; SNP codename – details describing which environment(s) (for the GxE setting) or epistatic QTN a given SNP is contributing to (for the Epistasis setting); BFT – median area under the receiver operating curve (AUC) value for the Brown-Forsythe test across ten randomly selected replicate traits; DGLM – median AUC value for the double generalized linear model across ten randomly selected replicate traits; MLM- median AUC value for the unified mixed linear mode across ten randomly selected replicate traits

CHAPTER 2: An application of vGWAS to differences in flowering time in maize across mega-environments [2]

2.1 ABSTRACT

Genomic regions containing loci with effect sizes that interact with environmental factors are desirable targets for selection because of increasingly unpredictable growing seasons. Although selecting upon such genotype-by-environment (GxE) loci is vital, identifying significantly associated loci is challenging due to the multiple testing correction. Consequently, GxE loci of small- to moderate effect sizes may never be identified via traditional genome-wide association studies (GWAS). Variance genome-wide association studies (vGWAS) have been previously shown to identify GxE loci. Combined with its inherent reduction in the severity of multiple testing, we hypothesized that vGWAS could be successfully used to identify genomic regions likely to contain GxE effects. We used publicly available genotypic and phenotypic data in maize (*Zea mays* L.) to test the ability of two vGWAS approaches to identify GxE loci controlling two flowering traits. We observed high inflation of $-\log_{10}(P-values)$ from both approaches. This suggests that these two vGWAS approaches are not suitable to the task of identifying GxE loci. We advocate that similar future applications of vGWAS use more sophisticated models that can adequately control the inflation of $-\log_{10}(P-values)$. Otherwise, the application of vGWAS to search for GxE effects that are critical for combating the effects of climate change will not reach its full potential.

[2] Matthew D. Murphy, Alexander E. Lipka

2.2 Introduction

Climate change is already undermining many aspects of eukaryotic life, with negative impacts affecting species in vulnerable ecosystems, crops and livestock in breeding populations, and even human health (Cheng et al., 2022; Kliem and Sievers-Glotzbach, 2022). Major abiotic consequences of climate change include more unpredictable precipitation, whether it is too much or too little, as well as generally higher temperatures (Ceccarelli and Grando, 2020). In addition, climate change will make environments more favorable to biotic stressors such as increased weed, insect, and pathogen pressures (Ceccarelli and Grando, 2020; Shahzad et al., 2021). With these stressors becoming more prevalent due to climate change, the need to breed for more yield-stable crops are of utmost importance (Langridge et al., 2021). One approach for achieving these breeding objectives is to exploit gene-by-environment (GxE) interactions (Bernardo, 2010). Under this approach, breeders would select against GxE loci to reduce phenotypic variance across environments and years (Langridge et al., 2021; Reckling et al., 2021). Contrastingly, breeders could also select for GxE loci to increase phenotypic variance and hence increase responsiveness when introducing a crop to a novel environment, new management practices, or new biotic pressures (Kusmec et al., 2018). Thus, the identification of GxE loci could prove to be pivotal for enabling breeders to promptly respond to challenging environments emerging due to climate change.

The availability of approaches seeking to understand GxE (van Eeuwijk et al., 2010; Des Marais et al., 2013; Li et al., 2021) underscore the attention that has been given better understand this critical component of trait variability. Harnessing GxE interactions is vital for plant breeding and there are several important challenges that need to be addressed. First, identification of

putative GxE loci themselves poses a substantial increase in the statistical multiple testing correction because the number of tests for association at each marker is at least doubled (Dempfle et al., 2008). The resulting conservativeness arising from correcting for this could result in a failure to detect GxE loci of small to moderate effect sizes (Bustos-Korts, 2016; Gauderman et al., 2017). To further complicate this issue, the risk of identifying false-positive GxE interactions increases when not all environmental covariates are accounted for (Westerman et al., 2022). Finally, GxE loci that behave non-additively (for example, having different phenotypic variance between genotypic groups) may be missed entirely with the most commonly-used statistical approaches (Ansarifar et al., 2020; Westerman et al., 2022).

Statistical analyses seeking to identify loci that control the variance of a trait, called variance quantitative trait loci (vQTL), have been shown to be capable of identifying GxE loci underlying simulated (Murphy et al., 2022) and real plant traits such as grain weight and water content in maize (Song et al., 2022). Such analyses have been routinely applied to variance genome-wide association studies (vGWAS) in plants and have contributed to the elucidation of the genetic architecture of metabolic plant traits (Shen et al., 2012; Forsberg et al., 2015; Hussain et al., 2020; Li et al., 2020), as well as plant architectural and phenology (Zhang and Qi, 2021). Given these contributions, as well as the lower multiple testing correction burden in vGWAS arising from the need to only conduct one test of association at each marker, it is critical to further study the potential of vGWAS approaches to identify GxE in plants.

There are two commonly-used vGWAS statistical approaches whose potential for identifying GxE have been previously shown in Murphy et al. (2022). The first approach is to conduct the Brown-Forsythe Test (BFT) at each marker (Brown and Forsythe, 1974). Although this test is computationally efficient, it has potential to lose power to detect vQTLs in the presence of other

large-effect loci controlling the studied trait (Hong et al., 2017; Córdova-Palomera et al., 2020). The other approach is the double generalized linear model (DGLM; Smyth, 1989; Ronnegard et al., 2012), and is widely used for vGWAS because of its potential to account for population structure and the presence of loci that control the population mean value of the traits. In Murphy et al. (2022), it was shown that the BFT and DGLM could identify simulated GxE loci at the largest possible tested sample size of N = 2,815.

The purpose of this study was to apply the findings of Murphy et al. (2022) on the potential for vGWAS approaches to identify simulated GxE loci to the analysis of actual traits. We partitioned a subset of maize data collected across multiple environments (Buckler et al., 2009; Tian et al., 2011; Hung et al. 2012a) to Midwestern and southern mega-environments, and then assessed the ability of two different vGWAS approaches to detect GxE loci associated with two flowering time traits. From a biological perspective, there is a substantial amount of diversity in maize flowering time, which allows this species to thrive in both temperate and tropical climatic conditions (Bouchet et al., 2013). Moreover, flowering time traits in maize are highly heritable and controlled by many small-effect additive loci, with GxE making a small but non-zero contribution to total phenotypic variance (Buckler et al. 2009). The latter implies that the effect sizes of GxE loci could be small, and thus we hypothesized that vGWAS could be particularly well-suited to identify these loci. Given the evidence of genes responsible for environmental sensitivity in maize (Li et al. 2016), we also hypothesized that the day lengths of at the Midwestern and southern mega environments were sufficiently different for vGWAS to identify the GxE loci responsible for differences in flowering time across environments.

2.3 Materials and Methods

Phenotypic Data

We analyzed two publicly available maize traits: growing degree days (GDD) to anthesis and GDD to silking. These traits were collected across multiple grow-outs of the 281-member Goodman-Buckler Diversity Panel (Flint-Garcia et al., 2005) at five locations in the USA from 2006-2009. This diversity panel was grown alongside the US maize nested association mapping panel (Yu et al., 2008; Buckler et al., 2009; McMullen et al., 2009), and the field design is described in the Materials and Methods of Hung et al. (2012a). To study GxE, we first collated the GDD phenotypic data into two separate data sets corresponding to two locations in the Midwestern USA (Urbana, IL; and Columbia, MO; called the Midwestern mega-environment in subsequent analyses) and two in the Southern USA (Homestead, FL; and Ponce, PR; called the southern mega-environment). We choose these two mega-environments due to the potential of capturing photoperiod sensitivity differences (Bonhomme et al., 1994; Xu et al., 2012; Chen et al., 2015). With the exception of Ponce, PR, all locations had two years' worth of data. However, we excluded phenotypic data from Urbana, IL, 2006 and Columbia, MO, 2007 due to one of the studied traits being unavailable in the former and extensive missing phenotypic data in the latter. For each of the two traits, best unbiased linear predictions (BLUPs) of the genotype effect were predicted using a mixed linear model fitted at each of these two mega-environments using the lme4 R package (Bates et al., 2015). This mixed linear model fitted is written as follows:

$$Y_{ij} = \mu + G_i + E_j + \varepsilon_{ij}, \quad (2.1)$$

where Y_{ij} represents the observed trait value of the i^{th} genotype grown in the j^{th} environment, μ is the grand mean, G_i represents the random effect of the i^{th} genotype, E_j is the random effect of

the j^{th} environment, and ε_{ij} represents the error term for the i^{th} genotype grown in the j^{th} environment. Because the 5,702 recombinant inbred lines (RILs) from the US NAM panel were grown alongside the Goodman-Buckler diversity panel at all locations, these RILs were included during the model fitting process to improve BLUP accuracy, as previously described (McMullen et al., 2009; B. R. Rice et al., 2020).

The process of configuring the BLUPs for GDD to anthesis and GDD to silking from each mega-environment for vGWAS is visualized in Figure 2.1. Briefly, for each of the two traits, the difference between the BLUPs from each mega-environment was taken. To factor out signals attributable to population structure and familial relatedness, each of these differences were fitted to a unified linear mixed model (MLM) similar to Yu et al., (2006) in TASSEL 5.0 (Bradbury et al., 2007). The Bayesian Information Criterion (BIC) (Schwarz, 1978) criterion option in the GAPIT R package (Lipka et al., 2012) was used to determine the optimal number of fixed-effect covariates (in this case, principal components, or PCs, of genome-wide markers) to include in the unified MLM. The ensuing analysis suggested to include the first five PCs as fixed-effect covariates to account for population structure. Additionally, we used the VanRaden (2008) additive kinship matrix to account for familial relatedness in the unified MLM. Consequently, the residuals from these two fitted unified MLMs (one with the difference in GDD to anthesis as the response variable, and one with the difference in GDD to silking as the response variable) were used as the response variable for vGWAS. We henceforth refer to these residuals as the GxE traits.

All of these steps, along with each vGWAS, are outlined in a flow chart created using the "DiagremmeR" R package (Iannone and Iannone, 2022).

Genotypic Data

We used fully sequenced genotypic data from the Goodman-Buckler Diversity Panel for the GxE traits (Flint-Garcia et al., 2005; Bukowski et al., 2018). Briefly, this marker data set was genotyped for 327,056 SNPs and anchored to the B73 RefGen_v4 reference genome. The filtering and imputation procedures for this dataset are described in Rice et al. (2020). LinkImpute (Money et al. 2015) was used to impute missing marker data.

Competing vGWAS models

We used two common vGWAS approaches to search for GxE, namely the Brown-Forsythe Test (BFT) and the double generalized linear model (DGLM). Both of these approaches are described in detail in Murphy et al. (2022). Briefly, both approaches test for evidence against equality of population variances of the GxE traits across the genotypes at a tested marker. In general, the BFT tests for equality of population variances across different groups by running a standard analysis of variance (ANOVA) on a median-derived transformation of the response variable (Brown and Forsythe, 1974). The test statistic follows an *F*-distribution under the null hypothesis of equal population variances across treatment levels. When applied to vGWAS, the BFT will test for equal trait population variances across the genotypes at a tested marker, as described previously (Shen et al., 2012; Murphy et al., 2022). The CAR R package was used to fit the BFT using the Levene.Test function with the central argument set to "median" (Fox and Weisberg, 2019).

The DGLM is a common vGWAS statistical model that belongs to the family of generalized linear models. The DGLM can model a single trait to a.) explanatory variables

controlling its population mean and b) explanatory variables controlling its population variance. The DGLM can also incorporate fixed effects (e.g., PCs calculated from a genome-wide marker set) to control population structure and markers tagging major-effect genes that control the trait. More detailed descriptions of the DGLM can be found in Corty and Valdar (2018), Hussain et al. (2020), and Murphy et al. (2022). Because we implicitly controlled for population structure prior to running the vGWAS (as described in Phenotypic Data), we did not incorporate PCs in the mean component of DGLM. To perform the DGLM, we used R code from Hussain et al. (2020) and Murphy et al. (2022), which was fitted using the dglm R package (Dunn and Symth, 2020).

To account for multiple testing, we used the Bonferroni procedure to control for the genome-wide type I error rate at $\alpha = 0.05$ for all tests. The CMPlots R package generated qqplots and Manhattan plots for all the tests (Yin, 2018). We used the genomic control (GC) (Devlin and Roeder, 1999) to quantify the degree of inflation of the test statistics from both vGWAS approaches.

Linkage disequilibrium analysis

To assess the local linkage disequilibrium (LD) in candidate genomic regions identified by our vGWAS approaches, we calculated r^2 estimates between each SNP in a given genomic region with the SNP that has the highest $-\log_{10}(P-value)$ in this genomic region using TASSEL version 5 (Bradbury et al., 2007). The resulting r^2 estimates of the candidate genomic region were plotted with the $-\log_{10}(P-values)$ with respect to base pairs in the B73 reference genome (version 4) using an R script previously used in Lipka et al. (2013).

Mixed model fitted across mega-environments

To assess the extent to which the two mega-environments captured GxE underlying GDD to anthesis and GDD to silking, the following mixed model was fitted to each of these two traits:

$$Y_{ijk} = \mu + ME_i + Trial_{k(i)} + G_j + (MExG)_{ij} + \varepsilon_{ijk}, \quad (2.2)$$

where Y_{ijk} represents the observed trait value of the j^{th} genotype grown in the k^{th} trial (i.e., particular year at a particular location) nested within the i^{th} mega-environment; μ represents the grand mean; ME_i represents the random effect of the i^{th} mega-environment; $Trial_{k(i)}$ is the random effect of the k^{th} trial nested within the i^{th} mega environment; G_j is the random effect of the j^{th} genotype; $(MExG)_{ij}$ is the random two-way interaction effect between the i^{th} mega-environment and the j^{th} genotype; and ε_{ijk} represents the error term for the j^{th} genotype grown in the k^{th} trial (i.e., particular year at a particular location) nested within the i^{th} mega-environment. After fitting this model in the "statgenGxE" R package (van Rossum, 2022), we used the variance component estimates to assess the contribution of $(MExG)_{ij}$ variance component (which quantifies GxE) to the overall phenotypic variability of both traits.

2.4 Results

Midwestern and southern mega-environments capture GxE for both studied traits

We fitted model (2) across the two mega-environments for GDD to anthesis and GDD to silking. The results revealed that the variance component estimate corresponding to the GxE term (i.e., the variance component estimate of the $(MExG)_{ij}$ random effect) accounted for approximately

20% of the total variability of both traits (Tables 2.2 and 2.3). This suggests that the subdivision of the publicly available field trials into Midwestern and Southern mega-environments captures a sufficient amount of GxE for both of the studied traits to justify conducting the ensuing vGWAS.

DGLM identified the greatest number of statistically significantly associated markers

Even after using the conservative Bonferroni procedure to adjust for multiple testing across the genome at α = 0.05, each of the two tested vGWAS approaches identified markers that were statistically significantly associated with both GxE traits (i.e., the difference in GDD to anthesis and GDD to silking across the two mega-environments). The two vGWAS approaches found more markers that significantly associated with the difference in GDD to anthesis than the difference in GDD to silking. This maybe attributed to the fact that some accessions underwent anthesis but not silking in this dataset. Of the two vGWAS approaches, the DGLM identified the greatest number of statistically significant associations. The number of significantly associated markers identified from each vGWAS approach are summarized in Figure 2.2 and presented in detail in Table 2.1.

Neither vGWAS approaches did not sufficiently control for false positives

For each of the tested traits, we noted that both vGWAS approaches yielded $-\log_{10}(P-values)$ that were highly inflated relative to what would be expected under the corresponding null hypotheses tested at each marker (Figure 2.3). The greatest amount of such inflation was observed for the DGLM. Nevertheless, the observed increases in $-\log_{10}(P-values)$ for all the analyses performed suggest that these two vGWAS approaches inadequately control for false positive associations for these data.

Peak-associated markers from vGWAS approaches colocalized to similar genomic regions within traits

We sought to characterize the consistency of which genomic regions were found to contain peak-associated markers across the two vGWAS approaches (Figure 2.4). This task was particularly challenging for analyzing the vGWAS results for the difference in GDD to anthesis because the DGLM identified statistically significant associations on every chromosome. Nevertheless, many of the peak-associated markers identified by the DGLM were located roughly in the same genomic regions as peak-associated markers identified by the BFT. When vGWAS was conducted on the difference in GDD to silking, we noted that the DGLM and BFT both identified markers significantly associated with the difference in GDD to silking in proximal genomic regions located on Chromosome 9.

Chromosome 9 region containing plausible candidate genes consistently identified from both vGWAS approaches for both traits

To illustrate the potential of these vGWAS approaches to highlight potential candidate genes, we present detailed results for GDD to anthesis using the DGLM in the genomic region on Chromosome 9 in Figure 2.5. We also created similar figures for the other vGWAS scans, which can be found in Figures 2.6-2.8. Across all vGWAS approaches and traits, the strongest peak associations (S9_151791144 and S9_151791148; 154942763 and 154942767 bp, respectively, *P*-value $1.19\,x\,10^{-10}$) was identified for the difference in GDD to anthesis by the DGLM. This marker was within 40 Kb of two candidate genes (Zm00001d048358 and Zm00001d048359,

B73 version 4). We also noted that there were SNPs surrounding these two candidate genes that are in moderate LD with this peak-associated SNP.

2.5 Discussion

One of the most important findings of our previous study (Murphy et al., 2022) was that vGWAS approaches could identify quantitative trait nucleotides contributing GxE effects to simulated traits. Therefore, we evaluated the potential for two different vGWAS approaches to identify genomic regions likely to contain loci with GxE effects in real, publicly available flowering time traits in maize. Although all two vGWAS models identified statistically significant marker-trait associations, we also observed that they yielded highly inflated $-\log_{10}(P-values)$. Thus, this analysis did not provide any further insights into the contribution of GxE to maize flowering time, but instead suggested that there is a critical need to explore the genomic sources underlying this severe inflation and to account for them in vGWAS models.

vGWAS identified markers significantly associated with the difference between at least one flowering time trait across two mega-environments

Identifying specific loci with GxE effects is vital because alleles at these loci could be selected for (or against) to achieve specific breeding goals. Selecting for such GxE loci is theoretically favorable for both introducing a crop to a novel environment and increasing responsiveness to favorable environments using new management practices (Kusmec et al., 2018). On the other hand, selecting against GxE loci is preferred when breeding for uniformity and stability in unpredictable growing environments (Bernardo, 2010; Kusmec et al., 2018). The fact that both vGWAS models identified statistically significant associations for the difference in two

flowering time traits across two mega-environments suggests that vGWAS could become a viable approach to help breeders identify GxE loci once the issues of inflated $-\log_{10}(P-values)$ are addressed.

Inflated false positive rates prevent inferences on GxE in flowering time The genetic architecture of flowering time in maize is consistent with the Fisher-Orr model (Fisher, 1918; Orr, 1998) in that it consists of many small-effect genes, as well as a small number of large-effect genes (Bouchet et. 2013, Peiffer et al. 2013). In particular, the analysis conducted by Buckler et al. (2009) in the US Maize NAM population provided overwhelming evidence that flowering time is controlled by many small-effect genes, and that the overarching genetic architecture can be well-approximated by accounting for their additive effects. Analyses conducted in other panels have identified larger-effect loci (e.g., Romay et al. 2013, Bouchet et al. 2013), while also underscoring that flowering time appeared to be controlled by a large number of small-effect genes (Bouchet et al. 2013; Li et al. 2016).

Although not as strong as the collective contribution of the additive effects of these genes, Buckler et al. (2009) found evidence that GxE contributes to the overall genetic architecture of flowering time. It was for this reason that we chose to perform our vGWAS analysis on GDD to anthesis and GDD to silking. Specifically, the potential for vGWAS to identify smaller-effect GxE loci (by reducing the severity of the multiple testing correction) could highlight which particular loci contribute to the small GxE signals detected by Buckler et al. (2009). Unfortunately, our analyses did not provide further elucidation into the genetic architecture of flowering time because both vGWAS approaches yielded substantially inflated type I error rates. Therefore, we urge future research to determine how to reduce the severity of

inflated type I error rates in vGWAS approaches and then reconduct this analysis. Once this issue is resolved, vGWAS has potential to characterize the putatively small-effect GxE loci for flowering time in maize. This could lead to similar, complementary advances that were already made for understanding the role of GxE for flowering time in *Arabidopsis* (Sasaki et al., 2015) , sorghum (Li et al., 2018) , and rice (Guo et al., 2020).

Future studies need to overcome the hurdle of high false positive rates

The most important finding from this study was the severe inflation of $-\log_{10}(P-values)$ from both vGWAS approaches, especially those from the DGLM. This degree of inflation was surprising because both GxE flowering time traits were first fitted to a unified MLM that accounted for population structure and familial relatedness (please see the Materials and Methods for details). The residuals from this fitted model were used in our subsequent vGWAS evaluations. Similar approaches have been implemented in previous studies (Shen et al., 2012; Forsberg et al., 2015; Li et al., 2020), and thus we anticipated that these sources of false positives would have already been accounted for in our vGWAS. Consequently, we empirically demonstrated that the fixed and random effects used by the unified MLM to account for false positives in traditional GWAS approaches are not guaranteed to also account for false positives inflating vGWAS associations. Thus, it is imperative that future research focuses on identifying and characterizing the genomic sources underlying these false positives. Concurrently, the computational bandwidth of statistical models that should account for these sources of false positives, for example the hierarchical generalized linear model (HGLM) and the double hierarchical generalized linear model (DHGLM) (Lee and Nelder, 1996; Lee and Nelder, 2006;

Rönnegård and Valdar, 2012), needs to be reduced so that researchers can feasibly use them in a vGWAS.

Follow-up studies need to address the inflation of $-\log_{10}(P-values)$ before investigating the putative GxE genomic regions identified in our study

Even though the tested vGWAS statistical approaches were prone to severe inflation of $-\log_{10}(P-values)$, both of them identified a genomic region of interest on Chromosome 9 containing peak-associated markers with both traits. This result does offer some promise of using vGWAS as a tool for prioritizing GxE genomic regions. However, follow-up studies using more sophisticated vGWAS models like the HGLM and DHGLM that account for false positives need to be conducted to determine if there are still peak-associated markers in this genomic region after explicitly controlling for population structure and familial relatedness. These follow-up studies should also be complemented with vQTL linkage mapping analyses as described in Corty and Valdar (2018). By using biparental crosses and similar experimental populations, linkage mapping has the potential to confirm peak vGWAS associations in independent data where extraneous sources of genetic variability can be controlled for by the mating design (Nordborg and Weigel, 2008; Korte and Farlow, 2013).

2.6 Conclusion

While our previous work highlighted the potential of the BFT and DGLM to identify GxE loci in simulated traits (Murphy et al., 2022), the analysis conducted here clearly shows these approaches yielded highly inflated $-\log_{10}(P-values)$ when applied to searching for GxE signals associated with flowering time in maize. This inflation highlights a serious weakness in this application of both approaches, and thus we recommend not applying the BFT and DGLM to

search for GxE loci with real trait data. That said, this also needs to be explored on a trait-by-trait basis. There is a critical need for future studies to explore use of more sophisticated vGWAS models to account for false positives. Without such an undertaking, the potential for using vGWAS to highlight specific genomic regions likely to harbor GxE loci for agronomically important traits like flowering time in maize will not be realized.

Data Availability

The genotypic and phenotypic data, results, and scripts used for the analyses are available at https://github.com/mdm10-code/An-application-of-vGWAS-to-differences-in-flowering-time-in-maize-across-mega-environments

2.7 Figures and Tables

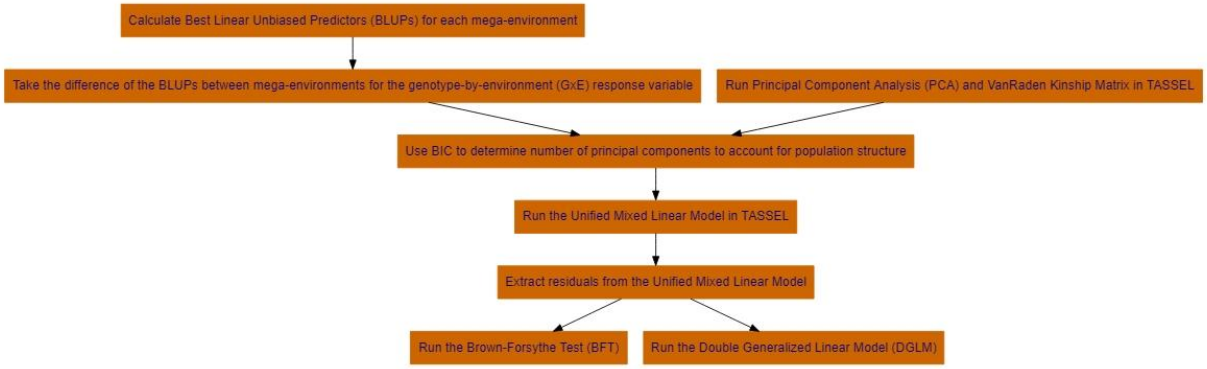

Figure 2.1. Flow chart. Summarizing the steps taken to process the difference in best linear unbiased predictions (BLUPs) for growing degree dates (GDD) to anthesis and silking calculated between the Midwestern and Southern mega-environments, factor out sources of genetic variability attributed to population structure and familial relatedness, and then run a variance genome-wide association study using the Brown-Forsythe test (BFT) and double generalized linear model (DGLM). Orange rectangles depict each step this process, and black arrows depict the how these steps are interconnected with each other. BIC, Bayesian information criterion; PCs, principal components; TASSEL, a software package whose abbreviation is Trait Analysis by aSSociation, Evolution, and Linkage.

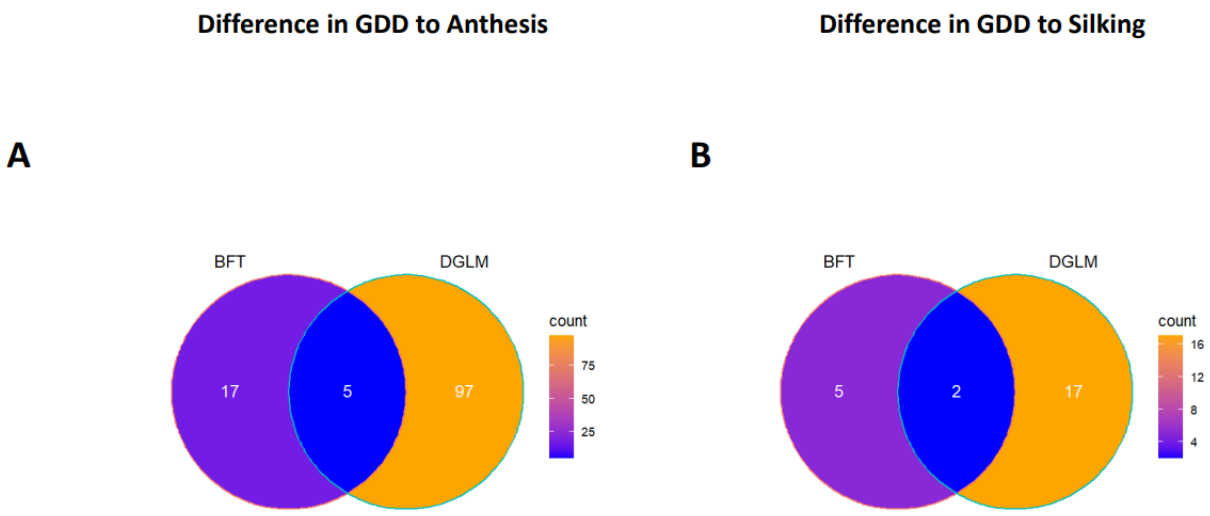

Figure 2.2. Venn diagrams. Showing the number of statistically significant SNPs for each trait identified by two variance genome-wide association studies (vGWAS) approaches, the Brown-Forsythe test (BFT) as the left circle, and the double generalized linear model (DGLM) as the right circle. The overlap between the circles represents the number of statistically significant SNPs that were identified in both of the vGWAS approaches. (A) Venn diagram showing the number of statistically significant SNPs identified by the vGWAS tests and their overlaps for the difference in growing degree days (GDD) to anthesis. (B) Venn diagram showing the number of statistically significant SNPs identified by the vGWAS tests and their overlaps for the difference in GDD to silking.

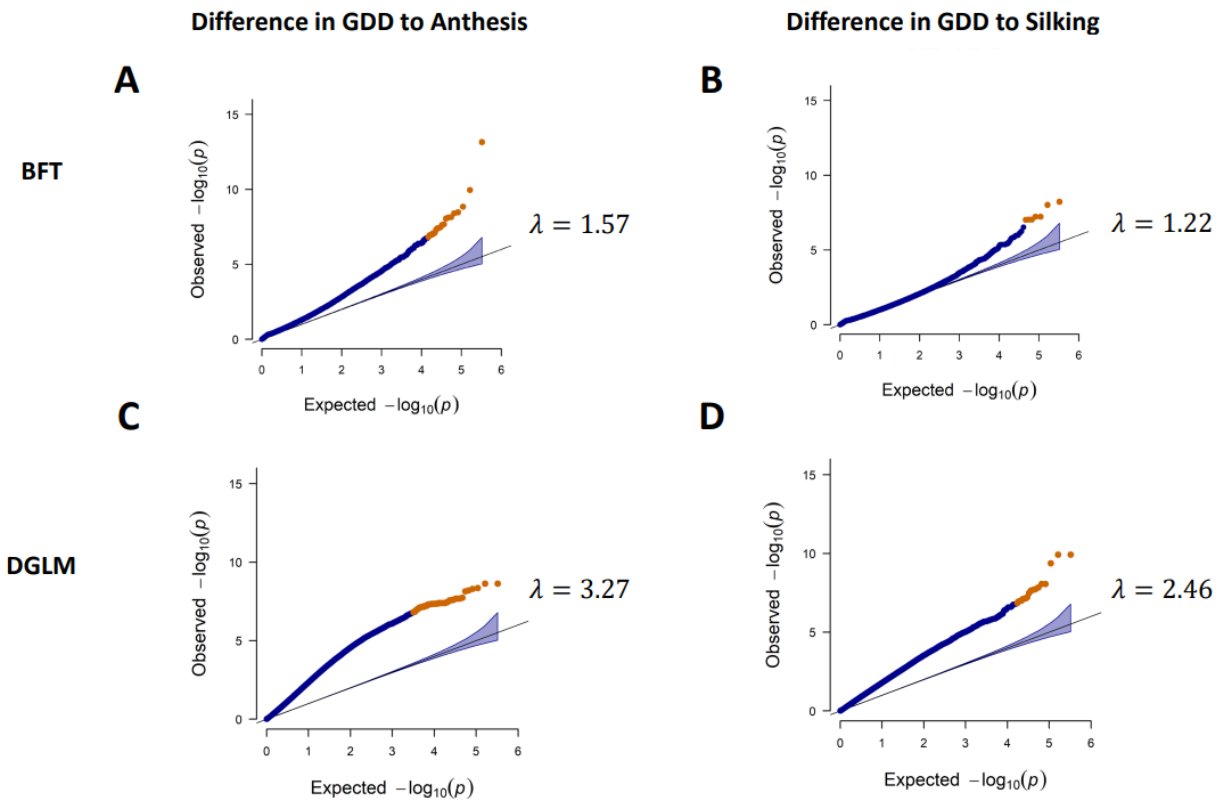

Figure 2.3. Quantile-quantile (QQ)-plots of all the tested variance genome-wide association studies (vGWAS) approaches (rows) tested in all traits (columns). For each plot, the observed $-\log_{10}(P-values)$ from testing each marker is presented on the *Y*-axis, while the expected $-\log_{10}(P-values)$ assuming the corresponding null hypotheses are correct are presented on the *X*-axis. The orange dots correspond to SNPs that are statistically significant after the Bonferroni procedure was used to control for multiple testing across the entire genome at $\alpha = 0.05$, whereas the blue dots correspond to non-statistically significant SNPs. Lambda values (λ) for genomic control are shown to the right of each QQ-plot. (A and C) QQ-plot of the difference in growing degree days (GDD) to anthesis for the Brown-Forsythe test (BFT) and double generalized linear model (DGLM). (B and D) QQ-plot of the difference in GDD to silking for the BFT and DGLM.

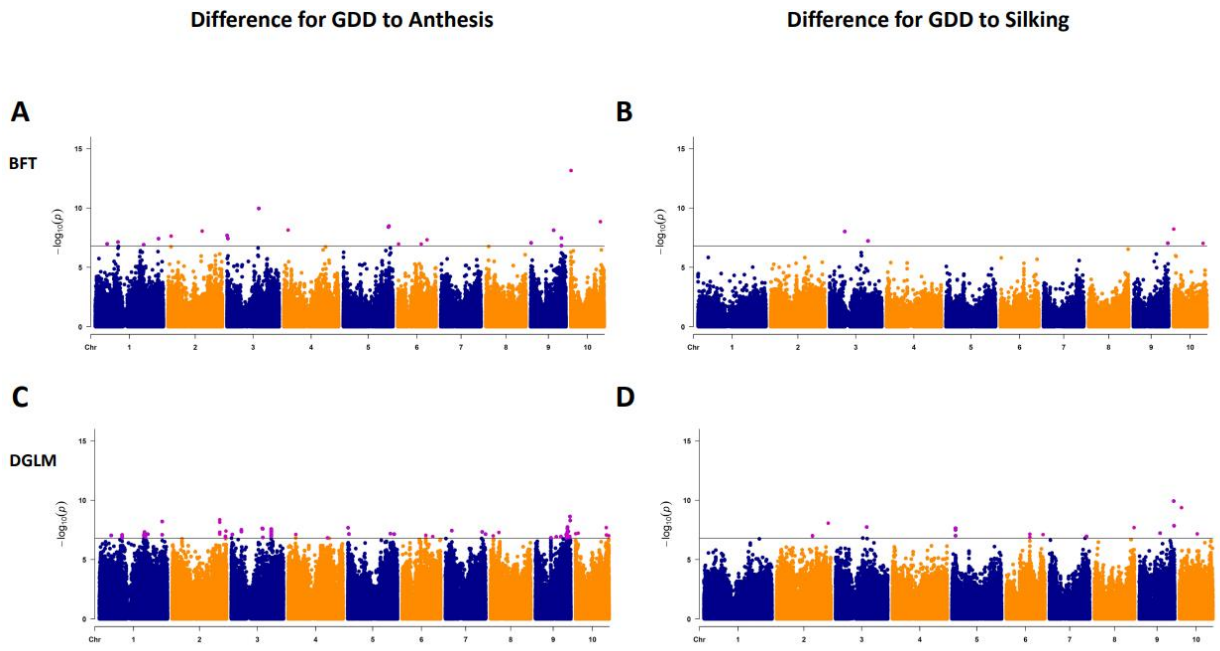

Figure 2.4. Summary of two variance genome-wide association studies (vGWAS) results. (A-D) Manhattan plots of the association results from the two vGWAS approaches across the ten maize chromosomes. The *Y*-axis represents the $-\log_{10}(P-values)$ plotted with respect to B73 RefGen_v4 genome position (*X*-axis). The gray horizontal line represents the threshold from Bonferroni procedure to control for genome-wide type I error rate at $\alpha = 0.05$. Statistically significantly associated SNPs are highlighted in purple. (A and C) Association results from the Brown-Forsythe Test (BFT) and double generalized linear model (DGLM) for difference in growing degree days (GDD) to anthesis. (B-D) Association results from the BFT and DGLM for difference in GDD to silking.

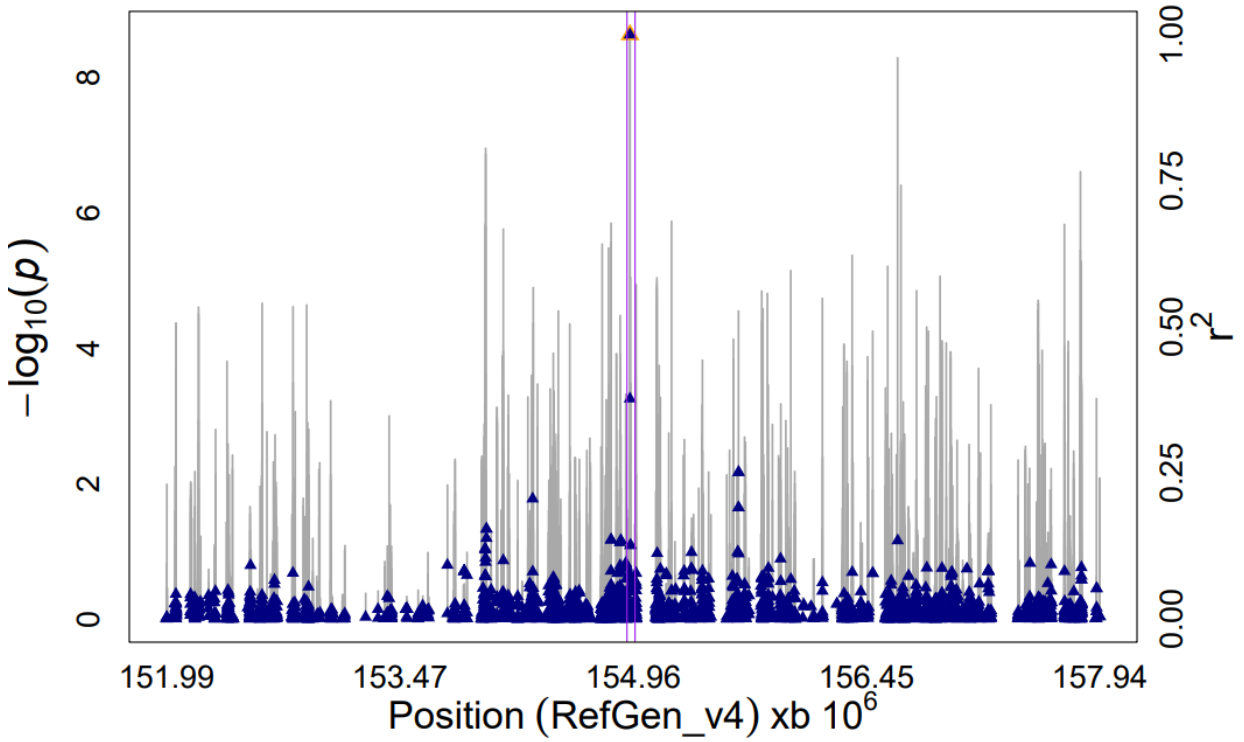

Figure 2.5. Variance genome-wide association studies (vGWAS) results for the difference in growing degree days (GDD) to anthesis. Plot of association results from the double generalized linear model (DGLM) and linkage disequilibrium (r^2) estimates in the genomic region between 152 and 158 Mb on Chromosome 9. The left *Y*-axis presents $-\log_{10}(P-values)$ from the vGWAS, while the right *Y*-axis and plots the r^2 values between each SNP and the peak associated marker, which is indicated by an orange triangle. The *X*-axis is the B73 RefGen_v4 position of this 6-MB region on Chromosome 9. The gray vertical-lines represent the $-\log_{10}(P-values)$ of the SNPs. The blue triangles represent the r^2 values of each SNP relative to the peak associated SNP. The solid purple lines represent the end positions of Zm00001d048358 (154,923,673 bp) and Zm00001d048359 (154,975,225 bp).

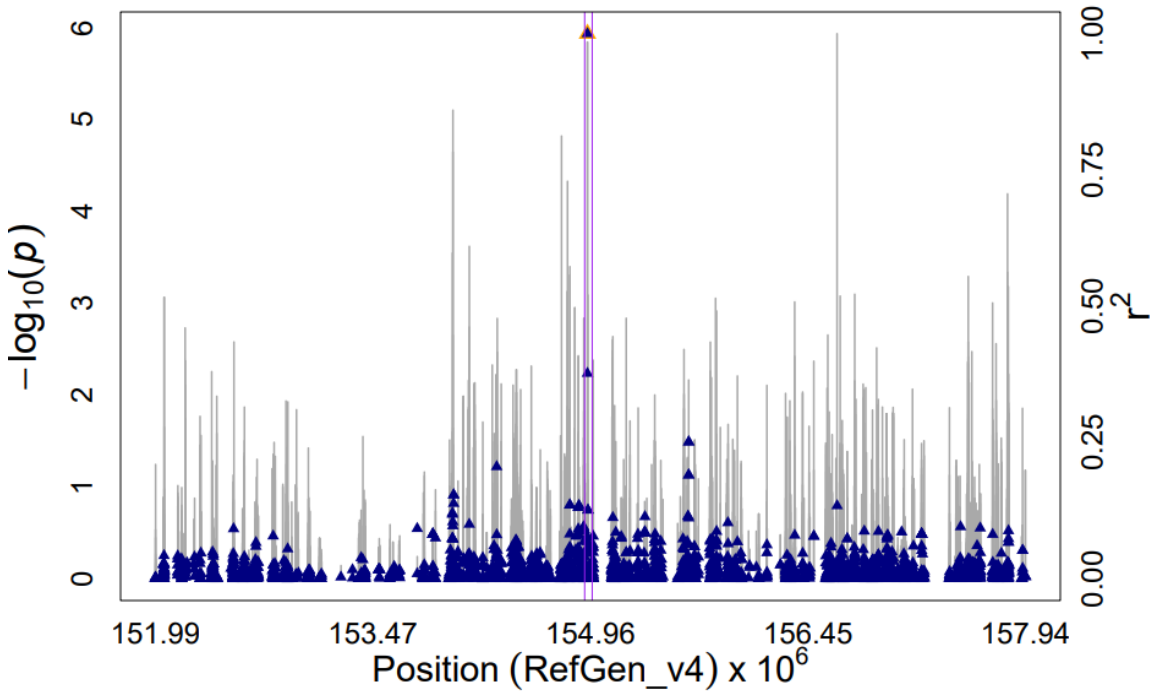

Figure 2.6. Variance genome-wide association studies (vGWAS) results for the difference in growing degree days (GDD) to anthesis. Plot of association results from the Brown-Forsythe (BFT) and linkage disequilibrium (r^2) estimates in the genomic region between 152 and 158 Mb on Chromosome 9. The left Y-axis represents the $-\log_{10}(P-values)$ from the vGWAS, while the right Y-axis plots the r^2 values between each SNP and the peak associated marker, which is indicated by an orange triangle. The X-axis is the B73 RefGen_v4 position of this 6-MB region on Chromosome 9. The gray vertical-lines represent the $-\log_{10}(P-values)$ of the SNPs. The blue triangles represent the r^2 values of each SNP relative to the peak associated SNP. The solid purple lines represent the end position of Zm00001d048358 (154,923,673 bp) and the start position of Zm00001d048359 (154,975,225 bp).

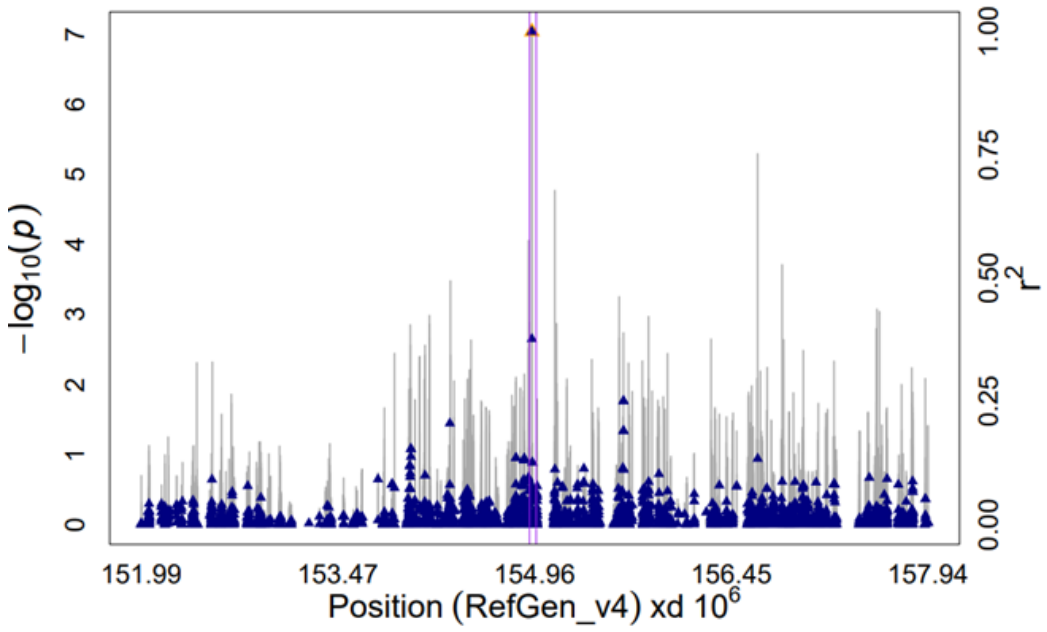

Figure 2.7. Variance genome-wide association studies (vGWAS) results for the difference in growing degree days (GDD) to silking. Plot of association results from the Brown-Forsythe test (BFT) and linkage disequilibrium (r^2) estimates in the genomic region between 152 and 158 Mb on Chromosome 9. The left Y-axis represents the $-\log_{10}(P-values)$ from the vGWAS, while the right Y-axis plots the r^2 values between each SNP and the peak associated marker, which is indicated by an orange triangle. The X-axis is the B73 RefGen_v4 position of this 6-MB region on Chromosome 9. The gray vertical-lines represent the $-\log_{10}(P-values)$ of the SNPs. The blue triangles represent the r^2 values of each SNP relative to the peak associated SNP. The solid purple lines represent the end positions of Zm00001d048358 (154,923,673 bp) and Zm00001d048359 (154,975,225 bp).

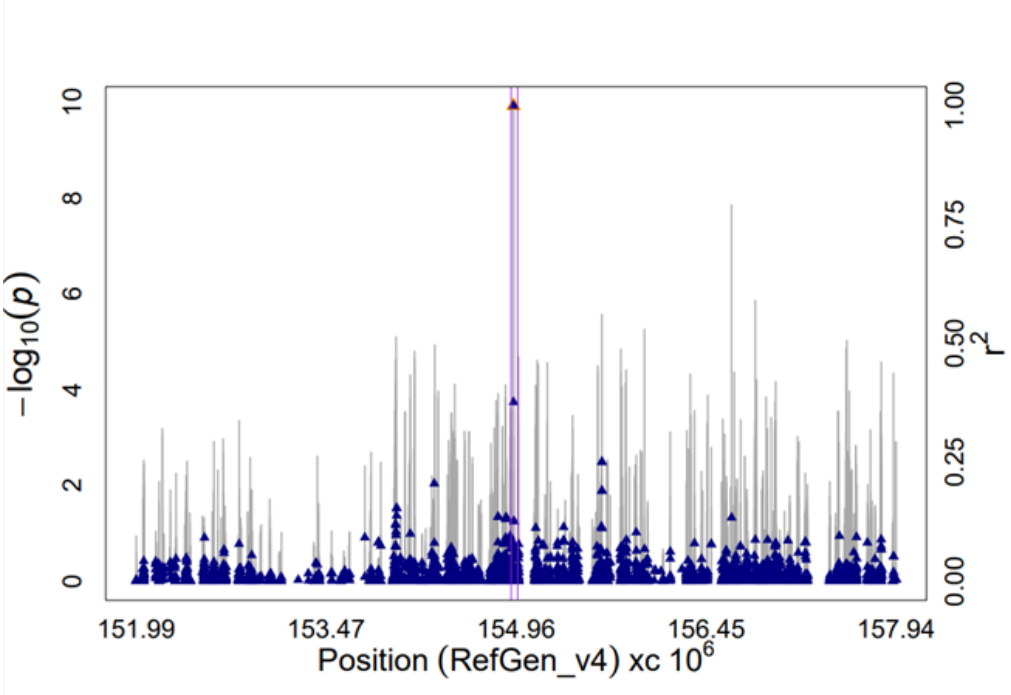

Figure 2.8. Variance genome-wide association studies (vGWAS) results for the difference in growing degree days (GDD) to silking. Scatter plot of association results from the double generalized linear model (DGLM) and linkage disequilibrium (r^2) estimates in the genomic region between 152 and 158 Mb on Chromosome 9. The left Y-axis represents the $-\log_{10}(P-values)$ from the vGWAS, while the right Y-axis plots the r^2 values between each SNP and the peak associated marker, which is indicated by an orange triangle. The X-axis is the B73 RefGen_v4 position of this 6-MB region on Chromosome 9. The gray vertical-lines represent the $-\log_{10}(P-values)$ of the SNPs. The blue triangles represent the r^2 values of each SNP relative to the peak associated SNP. The solid purple lines represent the end positions of Zm00001d048358 (154,923,673 bp) and Zm00001d048359 (154,975,225 bp).

Table 2.1. The number of statistically significant SNPs identified by the variance genome-wide association studies (vGWAS) tests for the difference in growing degree day (GDD) to anthesis and silking. BFT refers to the Brown-Forsythe test, while DGLM refers the double generalized linear model.

Test	Trait	Number of SNPs identified
BFT	Difference in GDD to anthesis	22
DGLM	Difference in GDD to anthesis	102
BFT	Difference in GDD to silking	7
DGLM	Difference in GDD to silking	19

Table 2.2. Variance component estimates from a mixed model fitted to growing degree days to anthesis across the two mega environments. The model is described in detail in the Materials and Methods.

	vcov	vcovPerc
region	384.5954	0.028366
region:trial	892.103	0.065798
genotype	8179.5	0.603287
genotype:region	2578.504	0.19018
residuals	1523.514	0.112368

Table 2.3. Variance component estimates from a mixed model fitted to growing degree days to silking across the two mega environments. The model is described in detail in the Materials and Methods.

	vcov	vcovPerc
region	273.5458	0.021469
region:trial	558.4567	0.043829
genotype	7627.474	0.598622
genotype:region	2721.498	0.213589
residuals	1560.752	0.122491

CHAPTER 3: Exploring the potential for vGWAS to stabilize genomic selection across environments [3]

3.1 Abstract

The ability to breed both for, and against, stability will be paramount for parts of the world most prone to climate change. Plant breeders need to reduce the time and resources needed to identify varieties that still produce sufficient yield in environments with the worst-case scenarios of the resulting abiotic and biotic stresses. Genomic selection (GS) is a promising avenue to get ahead of climate change in plant breeding. Even though GS models that can account for genotype-by-environment (GxE) interactions exist, inaccurate incorporation of these signals could negatively impact their predictive ability. A quantitative genetics phenomenon that may be useful for improving predictive abilities under these circumstances are variance quantitative trait loci (vQTLs). Previous work has shown that GxE is one of the potential genetic sources underlying vQTLs. Therefore, we investigated the potential of using peak-associated markers from a variance genome-wide association study (vGWAS) conducted in a training population to improve the predictive ability for growing degree days to anthesis-silking interval collected across multiple environments in a subset of the US maize (*Zea mays* L.) nested association mapping population. We observed that incorporating peak-associated vGWAS signals tended to lower predictive abilities across all families and environments. This suggests that more development is needed before deploying the strategy of using peak-associated vGWAS signals in a GS framework. We provide specific suggestions for further research investigating how to better utilize GxE manifested as vQTLs into breeding for and against stability.

[3] Matthew Murphy and Alexander Lipka

3.2 Introduction

A rapidly expanding human population and an increase in westernized diets will greatly tax global food production for the rest of the 21st century (Brown & Funk, 2008; Tai et al., 2014). The global food supply will be put under further stress by climate change. Not only will climate change increase the unpredictability and intensity of abiotic stressors relating to precipitation and temperature, but it will also lead to the proliferation of biotic stressors like insects, pathogens, and weeds (Ceccarelli & Grando, 2020; Shahzad et al., 2021). With more unpredictable stressors becoming the new norm, there is a call in plant breeding to prevent massive fluctuations in crop variety performance in not only across years, but even within the same growing season (Leng & Huang, 2017; Raza et al., 2019). A solution in countering these fluctuations and creating yield stable varieties in crops is to select against alleles that display high phenotypic plasticity across different environments, i.e., alleles with genotype-by-environment (GxE) interactions (Bernardo, 2010; Kusmec et al., 2018). In contrast, there are scenarios where breeding for phenotypic plasticity in crops is ideal. A few examples of where it would be ideal to select for phenotypic plasticity in crops are introduction into novel environments, exposure to new management practices in favorable environments, changes in consumerism preference like organic foods, or when new disease pressures arise (Desclaux et al., 2007). Thus, the ability to utilize GxE and plastic alleles in plant breeding have serious implications for food security (de Oliveira Silva et al., 2020; Kusmec et al., 2018; Mahmood et al., 2022).

Genomic selection (GS) has potential to accelerate the development of varieties with optimal breeding values in environments under unprecedented biotic and abiotic stresses arising from climate change (Xiong et al., 2022). The use of GS in plant breeding programs has become widespread due to its proven potential to increase genetic gain while simultaneously decreasing

the resources needed for evaluation (Meuwissen et al., 2001; Heffner et al., 2010; Roorkiwal et al., 2018). Through fitting a linear model in a training population that equates observed phenotypes to genome-wide marker sets plus an error term, GS provides genomic estimated breeding values (GEBVs) for breeding germplasm that has not been phenotyped (Meuwissen et al., 2001; Goddard and Hayes, 2007). Because these GEBVs can be obtained without the need to grow out and phenotype such breeding germplasm, GS typically reduces the time needed for evaluating breeding materials, and thus could expedite the deployment of yield-stable varieties (Goddard and Hayes, 2007; Xiong et al., 2022). However, GS prediction accuracies can be negatively affected if GxE is not accurately quantified, especially across highly diverse growing conditions and environments. Consequently, decreases in prediction accuracies may affect the prospects of selecting optimal individuals based on GEBVs (Spindel and McCouch, 2016). Such decreases could be mitigated due to extensive research in adapting GS models to accurately quantify GxE through multi-kernel frameworks (Burgueño et al., 2012; Heslot et al., 2014; Jarquín et al., 2014; Cuevas et al., 2017).

Although selecting for or against GxE alleles both have merit for plant breeding, there are barriers in identifying these types of loci. These include both an extensive multiple testing correction and that GxE effects may not always behave additively (Yang, 2014; Wang et al., 2021; Waters et al., 2023). Ultimately, such barriers could result in the underutilization of moderate-sized, but biologically important GxE loci in breeding (Struchalin et al., 2012; Malosetti et al., 2013). Alternate quantitative genetics approaches that seek to identify loci associated with the population variance of a trait (i.e., variance quantitative trait loci or vQTLs; Rönnegård & Valdar, 2012) have potential to bypass these barriers in detecting GxE interactions. Analogous to detecting mean QTLs (hereafter called mQTLs), variance genome-wide

association studies (vGWAS) and vQTL linkage mapping approaches have been developed to detect and quantify vQTLs (Corty & Valdar, 2018). While many statistical approaches have been deployed for vGWAS, two commonly-used approaches are the Brown-Forsythe Test (BFT) and the double generalized linear model (DGLM). While useful because of its simplicity, the BFT cannot directly control for population structure and familial relatedness (Hong et al., 2017). In contrast, the DGLM can control for population structure as fixed effects but cannot account for familial relatedness (Y. Lee & Nelder, 1996; Rönnegård & Valdar, 2012). Our previous simulation study has demonstrated that it is possible for both of these vGWAS approaches to identify GxE effects simulated using marker data (Murphy et al. 2022). Thus, GxE loci identified by vGWAS could theoretically be incorporated into a GS model as fixed-effect covariates or as a separate kernel; the inclusion of such loci could facilitate breeding efforts seeking to select for germplasm with stable phenotypes across environmental fluctuations. While the potential for vGWAS and related analyses to identify GxE loci have been evaluated for real traits in animals (Mouresan et al., 2019; Braz et al., 2021), these approaches have been underexplored for real traits in plants (Song et al., 2022; Murphy and Lipka, in press).

To the best of our knowledge, the performance of GS models that include GxE loci identified through a vGWAS as fixed-effect covariates has never been evaluated. However, a study conducted by Mouresan et al. (2019) could shed light on the potential of vQTL to improve GS prediction accuracy. Using both simulated and real trait data, Mouresan et al. (2019) observed that GS models including peak-associated markers from a vGWAS as a fixed-effect covariate tended predict GEBVs more accurately than traditional GS models for traits that were controlled by vQTLs. Even though this model shows promise, this model employed by Mouresan et al. (2019) did not explicitly allow for a GxE term. Nevertheless, this study suggests

that it is merited to evaluate the impact of incorporating vGWAS results into GS models used to breed for phenotypic stability.

The purpose of this study was to investigate if incorporating vGWAS signals into a GxE genomic selection model can generate consistent GEBVs across environments to implement this for breeding for stability. We used publicly available genotypic and multi-environment flowering-time phenotypic data from the US nested association mapping (NAM) populations in maize (*Zea mays* L.) (Yu et al., 2008; Buckler et al., 2009; McMullen et al., 2009). We hypothesized that incorporating peak vGWAS associations into the GS model can increase prediction accuracies under certain circumstances, as seen in both Mouresan et al. (2019) and in parallel with studies that investigated accounting for mQTLs in genomic prediction models (Bernardo, 2014; Rice & Lipka, 2019).

3.3 Materials and Methods

Genotypic and Phenotypic Data

We used genotypic and phenotypic data from the US maize nested association mapping (NAM) panel, which has been previously described (Yu et al., 2008; Buckler et al., 2009; McMullen et al., 2009). Briefly, this panel was created by crossing 25 genetically diverse maize inbred lines to B73, the common parent, to form 25 recombinant inbred line (RIL) families. The resulting panel comprises of approximately 5,000 recombinant inbred lines (RILs), with the typical RIL family size being approximately 200 RILs. To perform our analyses, we used 14,772 markers measured in 4,635 of the NAM RILs that have been used and described in previous analyses (Diepenbrock et al., 2017, Diepenbrock et al. 2021). In summary, these marker data were obtained from genotyping by sequencing (Elshire et al., 2011; Glaubitz et al., 2014) using a procedure where markers were imputed every 0.1 cM, as described in (Ogut et al., 2015).

For the phenotypic dataset, we analyzed growing degree days (GDD) to anthesis-silking interval (GDDASI) from the traitMatrix_maize282NAM_v15-130212.txt from Panzea (http://www.panzea.org/#!phenotypes/c1m50 accessed March 2020). These data consist of the Goodman-Buckler Diversity Panel (Flint-Garcia et al., 2005) and the US maize NAM panel grown at seven locations. The data collection procedures and experimental design details are described in (Hung et al., 2012a).

Due to the computational complexities of the ensuing analyses, we used a subset of the NAM panel and available locations. Specifically, we used six NAM families, each with a diverse founder representing at least one of the heterotic maize groups summarized in Liu et al. (2003) and that segregated for photoperiod sensitivity (Coles et al., 2010; Hung et al., 2012b). To emulate the full NAM panel's genetic composition, most of the families had a tropical/subtropical line as a diverse founder that still displayed variation for flowering time. Accordingly, five of the families we selected were Z002 (B73 x CML103; tropical), Z003 (B73 x CML228; tropical), Z010 (B73 x Hp301; popcorn), Z011 (B73 x IL14H; sweet), and Z014 (B73 x Ky21; non-stiff stalk) (Flint-Garcia et al., 2005; Gage et al., 2020). We also selected NAM family Z020 (B73 x NC350) due to the broad adaptability of NC350, a subtropical/tropical line adapted to temperate environments (Nelson et al., 2016; Woodhouse et al., 2021). The subsets of environments we considered were at Aurora, NY, Urbana, IL, and Clayton, NC, due to the lower rates of missing data at these locations. Accordingly, we considered the available GDDASI data from both the 2006 and 2007 field seasons at each of these environments. Inclusion of these phenotypic data from the 2007 field season data sets are of particular interest because the United States experienced a severe drought that year, which is ideal for studying abiotic stress (Price et al., 2011).

Stability Analysis

Stability estimates were obtained from fitting the Finlay-Wilkinson Regression (FWR) model (Finlay and Wilkinson, 1963) to every RIL in the six NAM families we selected, using an approach analogous to Gage et al. (2017). Thus, for each RIL we fitted the following simple linear regression model:

$$Y_{ij} = \beta_{0j} + \beta_{1j} x_i + \varepsilon_{ij} \quad , (3.1)$$

where Y_{ij} is the observed GDDASI value of the j^{th} NAM RIL in the i^{th} environment, β_{0j} is the intercept parameter for the j^{th} NAM RIL, β_{1j} is the slope parameter of the j^{th} NAM RIL, x_i is the GDDASI value of B73 checks measured in the i^{th} environment, and ε_{ij} is the random error term for the j^{th} NAM RIL in the i^{th} environment ~NID(0, $\sigma^2_{\varepsilon_j}$). We fitted this FWR model using the lm() function in R and we extracted the least squares estimate of the slope. The resulting slope estimates all of the NAM RILs were used as the response variables in parts of the ensuing analyses.

Cross-Validation Scheme employed for ensuing analysis

We used the cross-validation 00 (CV00) scheme to assess how accurate the GS models described below can predict GEBVs for unobserved RILs in unobserved environments (Guo et al., 2020; Jarquin et al., 2020, Jarquin et al., 2021). Under this scheme, the validation sets consisted of the RILs of a given NAM family in a given environment. The corresponding training set consisted of the GDDASI values from the RILs of the remaining NAM families observed in the remaining environments. This scheme resulted in a total of 36 different combinations of training and validation sets.

vGWAS Analysis

To assess the potential for variance quantitative trait loci (vQTLs) to produce stable GEBVs across environments, we undertook a two-step process using the Brown-Forsythe Test (BFT) and double generalized linear model (DGLM). First, we conducted a vGWAS on each training set to find vQTLs associated with the slope estimates obtained from FWR. We previously described the BFT in the simulation study conducted in Murphy et al. (2022) to identify GxE loci that manifest themselves as vQTLs. The marker from this vGWAS with the strongest association (i.e., with the lowest *P*-value from testing H_0: Population variances of the slope estimates from FWR are equal across all genotypes at the tested marker) was then included as a fixed-effect covariate in the dispersion part of the following DGLM:

$$Y_{ijk} = \mu_m + \varepsilon_{ijk} \qquad (3.2)$$

where Y_{ijk} corresponds to the GDDASI of the k^{th} RIL in the j^{th} environment and the i^{th} family; μ_m is the intercept; and ε_{ijk} is the error term corresponding to the k^{th} RIL in the j^{th} environment and the i^{th} family $\sim N(0, \sigma^2_{ijk})$. This DGLM links the error population variance σ^2_{ijk} to the following linear combination of explanatory variables:

$$\log(\sigma^2_{ijk}) = \mu_v + a_v s_k \qquad (3.3)$$

where $\sigma^2_{\varepsilon_{ijk}}$ is the residual variance to the k^{th} RIL in the j^{th} environment and the i^{th} family; variance intercept; a_v is the effect size of the SNP with the lowest *P*-value from the vGWAS scan ran using the BFT; and s_k is the observed genotype of the SNP with the lowest *P*-value

from the vGWAS scan at the k^{th} RIL, encoded as 0, 1, or 2. This DGLM was fitted using the dglm R package (Symth, 1989; Dunn and Symth, 2020), with code derived from Hussian et al (2020) and Murphy et al (2022). The resulting residuals from this DGLM are considered as one of the two response variables for the GS model described in the next section.

Marker by Environment Genomic Selection Model

We considered the Marker by Environment (MxE) genomic selection model that has been proposed in Lopez-Cruz et al. (2015). Briefly, this is a multi-trait model, where the phenotypic values across each of the six environments are set to be an individual trait. The MxE model considers both common marker effects across all environments and environment-specific marker effects; the latter is potentially useful for quantifying the contribution of GxE to the overall genetic architecture of the trait of interest (Lopez-Cruz et al., 2015; Crossa et al., 2017). The MxE model is written as follows in matrix notation:

$$\begin{pmatrix} \boldsymbol{Y_1} \\ \boldsymbol{Y_2} \\ \vdots \\ . \\ \boldsymbol{Y_6} \end{pmatrix} = \begin{pmatrix} \mathbf{1}\mu_1 \\ \mathbf{1}\mu_2 \\ \vdots \\ . \\ \mathbf{1}\mu_6 \end{pmatrix} + \begin{pmatrix} \boldsymbol{u_{01}} \\ \boldsymbol{u_{02}} \\ \vdots \\ . \\ \boldsymbol{u_{06}} \end{pmatrix} + \begin{pmatrix} \boldsymbol{u_{11}} \\ \boldsymbol{u_{12}} \\ \vdots \\ . \\ \boldsymbol{u_{16}} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon_1} \\ \boldsymbol{\varepsilon_2} \\ \vdots \\ . \\ \boldsymbol{\varepsilon_6} \end{pmatrix}, \textbf{(}3.4\textbf{)}$$

where $\boldsymbol{Y_i}$ is an *N*-dimensional vector of phenotypic observations of the i^{th} environment; $\mathbf{1}$ is an *N*-dimensional vector of 1's; μ_i is the population mean of $\boldsymbol{Y_i}$; $\boldsymbol{u_0} = (\boldsymbol{u'_{01}}, \boldsymbol{u'_{02}}, \dots, \boldsymbol{u'_{06}})' \sim MVN(\mathbf{0}, \sigma^2_{G_0}\boldsymbol{G_0})$ is an *Nx6*-dimensional vector of random genetic effects of the RILs that are common to all environments; $\boldsymbol{u'_1} = (\boldsymbol{u'_{11}}, \boldsymbol{u'_{12}}, \dots, \boldsymbol{u'_{1j}}) \sim MVN(\mathbf{0}, \boldsymbol{G_1})$ is an *Nx6*-dimensional vector of random genetics effects that are specific to each environment; $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon'_1}, \boldsymbol{\varepsilon'_2}, \dots, \boldsymbol{\varepsilon'_6})' \sim MVN(\mathbf{0}, \sigma^2_\varepsilon \boldsymbol{I})$; $\sigma^2_{G_0}$ is the population variance of each element of $\boldsymbol{u_0}$; σ^2_ε is the population variance of each element of $\boldsymbol{\varepsilon}$; and $\boldsymbol{I}$ is an $nk\ x\ nk$ identity matrix. $\boldsymbol{G_0}$ is an additive genetic relatedness calculated as follows:

$$\boldsymbol{G_0} = \begin{pmatrix} \boldsymbol{X_1 X'_1} & \cdots & \boldsymbol{X_1 X'_j} \\ \vdots & \ddots & \vdots \\ \boldsymbol{X_1 X'_j} & \cdots & \boldsymbol{X_6 X'_j} \end{pmatrix} /p, \text{ (3.5)}$$

where $\boldsymbol{X_j}$ is an Nxp matrix of genotype values for p markers across the N RILs grown in the j^{th} environment. Each marker is encoded as 0,1, and 2. $\boldsymbol{G_1}$ is analogous to $\boldsymbol{G_0}$, but is a block-diagonal matrix that i.) does not consider relatedness between RILs across environments and ii.) has separate variance components at each environment. $\boldsymbol{G_1}$ is calculated as follows:

$$\boldsymbol{G_1} = \begin{pmatrix} \sigma^2_{u11}\boldsymbol{X_1 X'_1} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \sigma^2_{u61}\boldsymbol{X_6 X'_6} \end{pmatrix} /p, \text{ (3.6)}$$

where $\boldsymbol{0}$ is an n x n matrix of zeros, σ^2_{u1i} is the population variance of the i^{th} element of $\boldsymbol{u_1}$, and all other terms are as previously described.

For each of the training sets obtained from the CV00 cross-validation scheme, we fitted this model twice. The response variable for the first model was the observed GDDASI value, while the response variable for the second model was the residuals from the DGLM models described in (2) and (3). Thus, the former response variable does not account for peak vGWAS associations, while the latter response variable does. We measured predictive ability from the resulting GEBVs in each corresponding validation set by estimating the Pearson correlation between the observed GDDASI values and the GEBVs. When using the residuals from the second model, we estimated the Pearson correlation between the residuals and GEBVs. To assess stability across environments, we calculated Pearson correlations between the median GDDASI values of each RIL across all environments and the median GEBVs across all environments. We used the "DiagremmeR" R package to illustrate all of these steps in our proposed pipeline (Iannone and Iannone, 2022) (Figure 3.1).

3.4 Results

Overall, predictive abilities of GEBVs for GDDASI were consistently lower and tended to be negative when accounting for peak vGWAS associations from a training set (Figure 3.2). When the peak-associated vGWAS signals considered, Families Z003 and Z020 had the highest and lowest overall predictive abilities, respectively. Irrespective of integrating peak-associated vGWAS signals, Family Z010 tended to yield higher predictive abilities, while Family Z014 tended to yield lower overall predictive abilities.

Did including peak-associated markers from vGWAS improve prediction accuracy within environments?

In general, predictive abilities were observed to be lower at each environment when peak vGWAS associations were included (Figure 3.3). However, a few exceptions were noted, especially for Aurora 2006. Incorporating vGWAS signals into the analysis also tended to yield negative predictive abilities. These results refute our hypothesis that incorporating peak vGWAS associations can improve predictive abilities within environments.

Did including peak-associated markers from vGWAS improve prediction accuracy across environments?

The Pearson correlation between the median GEBVs and the median GDDASI values across environments tended to be higher when peak vGWAS associations were not considered. (Figure 3.4). Family Z002 had the highest predictive ability when peak vGWAS associations were considered, while Family Z020 had the lowest predictive ability. When peak vGWAS signals where not considered, Family Z010 had the highest predictive ability, while Family Z014 had the lowest.

3.5 Discussion

We incorporated peak-associated vGWAS signals into cross-environment predictions of GEBVs for GDDASI in a subset of the US maize NAM population grown in multiple environments. We observed that such an incorporation yielded a substantial reduction in predictive abilities. These results suggest that incorporating peak-associated vGWAS signals into genomic prediction is not a viable approach for assisting plant breeding efforts to select for or against stability across environments.

Prospects of using vQTLs to assist in genomic selection and breeding for stability

Because of the theory behind a GxE locus manifesting itself as a vQTL (Rönnegård & Valdar, 2011), this GS study focused on incorporating peak-associated vGWAS signals into a GxE framework via the GS model proposed by Lopez-Cruz et al. (2015). The research presented here suggests that incorporating peak-associated vGWAS signals into GS models reduces predictive abilities. Nevertheless, we observed low predictive abilities regardless of whether or not peak vGWAS associations were considered. A possible explanation for this result is the cross-validation scheme itself. We sought to predict GEBVs of untested RILs in untested environments (i.e., in a CV00 cross-validation scheme), and therefore we expected lower overall predictive abilities, as reported in previous studies (Ankamah-Yeboah et al., 2020; Jarquin et al., 2020). It is worth investigating how different cross-validation schemes influence the predictive abilities when incorporating peak-associated vGWAS signals.

A few trends stood out when observing the predictive abilities across families. Family Z020 has surprisingly low prediction accuracies. The diverse founder of this family, NC350, is a

tropical inbred line adapted to temperate climates. We initially predicted that Z020 would have the highest predictive abilities because NC350 can be grown in temperate, subtropical, and tropical environments (Nelson et al., 2016; Woodhouse et al., 2021). However, the tropical families (Z002 and Z003) had the highest raw (Figure 3.2) and median-based (Figure 3.4) predictive abilities when peak vGWAS associations were accounted for.

The Aurora 2006 environment had the highest predictive abilities among the validation environments when peak-associated vGWAS signals were accounted for (Figure 3.3). In contrast, Aurora 2006 yielded the lowest predictive abilities when no peak-associated vGWAS signals were not accounted for. This may suggest that the vGWAS conducted on the training sets may have detected a genomic locus that is highly predictive of breeding values for this particular environment.

We were initially going to investigate this approach with all possible locations grown during the 2006-2009 field seasons that are publicly available. However, we only focused on these locations due to their low rate of missing phenotypic data. Investigating the predictive abilities in more tropical locations, such as Homestead, Florida, or Ponce, Puerto Rico, would be interesting to investigate to see how a tropical environment may impact predictive abilities.

Considerations in using peak-associated vGWAS signals for breeding for stability

Despite the overall low predictive abilities observed when peak-associated vGWAS signals were incorporated into GS, our findings suggest that there are potential areas that warrant further investigation. First, we explored only one trait (GDDASI) in our study. Investigating a more comprehensive range of traits with contrasting genetic architectures should shed light on the generalizability of our findings. Second, we chose to analyze the US NAM population to better control for the confounding effects of population substructure and familial relatedness often dealt

with diversity panels (Flint-Garcia et al., 2005; McMullen et al., 2009). This was important because we previously noticed *P*-value inflation when using the BFT and DGLM to identify GxE loci in the Goodman-Buckler maize diversity panel (Murphy and Lipka, in press). Thus, the choice of germplasm should also be considered when using peak-associated vGWAS signals. If one were to use this approach with a diversity panel, then population structure and familial relatedness could be addressed implicitly by extracting residuals from a linear mixed model as described in the vGWAS literature (Shen et al., 2012; Forsberg et al., 2015; Murphy et al., 2022). Nevertheless, we demonstrated in Murphy and Lipka (in press) that even such an approach does not always guarantee adequate control of false positives. Future research should consider studying other multi-parent populations like multi-parent advanced generation inter-cross (MAGIC) and random-open-parent association mapping (ROAM) populations (Dell'Acqua et al., 2015; Xiao et al., 2016).

For calculating stability estimates, we used the widely utilized the FWR model to obtain slope estimates. Finlay-Wilkinson regression is often used for assessing stability in plant breeding due to its parsimonious interpretation (Walsh and Lynch 2014). However, FWR does have drawbacks, including increased bias from missing data across multiple environments and potentially inducing large sampling variance from assuming fixed effects (Lian & De Los Campos, 2016). It may be worthwhile to investigate additional stability statistical models such as the Additive Main Effects and Multiplicative Interaction Model to reduce any extraneous sources of variation when calculating stability estimates (Cornelius, 1993; Piepho, 1995).

We ran the BFT on the slope estimates from FWR to identify peak-associated vGWAS signals. The rationale for using BFT instead of the DGLM is that we demonstrated in Murphy et al. (2022) that the BFT and DGLM performed similarly when searching for GxE loci. However,

one caveat Murphy et al. (2022) identified in using vGWAS for GxE genomic prioritization is that a large sample size was often needed in maize to use vGWAS for this matter successfully. The NAM subset we used may not have had a large enough sample size to identify GxE loci that manifested as a peak-associated vGWAS signal. It may be worthwhile to also investigate the performance of DGLM, but we hypothesize that DGLM will perform on par with BFT.

Finally, our study used the MxE GS model from Lopez-Cruz et al. (2015). The rationale for using this model was that it allows for marker effects to change across environments, which is theoretically optimal for quantifying GxE. However, a drawback of the MxE model is that it assumes that the contributions of GxE are additive. In addition, Lopez-Cruz et al. (2015) did note that MxE performed best when the environments had a positive correlation with one another. It is possible that our selected environments did not have all positive correlations. An alternate model that theoretically addresses these drawbacks was proposed by Cuevas et al. (2016), in which a non-Gaussian kernel is used to account for GxE in the model.

How incorporating peak-associated vGWAS signals parallels incorporating peak-associated mGWAS markers in a GS model

The rationale for incorporating peak-associated vGWAS signals is analogous to what has been demonstrated with incorporating peak-associated mean GWAS (mGWAS) signals in genomic selection (Bernardo, 2014; Rice & Lipka, 2019). The theory is that predictive abilities can be improved when single or multiple peak-associated mGWAS signals accounting for at least 10% of the phenotypic variance are incorporated into GS models better to quantify GEBVs. While Rice and Lipka (2019) demonstrated that predictive abilities improved in some simulations, most of the simulations yielded lower predictive abilities when including peak-associated mGWAS signals into the GS model. Considering what has been observed with such incorporation of peak-

associated mGWAS signals, there may be some explanations for why incorporating peak-associated vGWAS signals lowered predictive abilities. One is that we used this approach in an outcrossing species (maize) (Wallace et al., 2014). Rice and Lipka (2019) did note that predictive abilities in simulation settings in sorghum, an inbreeding species, tended to be higher than that of maize when incorporating peak-associated mGWAS signals. Incorporating peak-associated vGWAS signals into genomic selection models should also be explored in inbreeding species as well.

Another plausible explanation of why incorporating peak-associated vGWAS signals lowered predictive abilities is that the peak-associated vGWAS signals may have been due to either a pure vQTL, an epistatic QTL, or additional confounding in the error term (Rönnegård & Valdar, 2011). In addition, Rice and Lipka (2019) also mentioned that the population size of the training and validation populations might also factor into the success of incorporating peak-associated GWAS signals. This is especially important because i.) our validation families tended to be less than 200 RILs and ii.) larger sample sizes are often needed when searching for vQTLs compared to mQTLs (Lee and Nelder, 2006; Rönnegård & Valdar, 2012).

3.6 Conclusions

Our findings clearly suggest that further research needs to be invested into incorporating GxE loci into GS to breed for stability. We recommend that researchers investigate how to best incorporate GxE with a broader range of germplasm, traits with differing genetic architectures, different GS models, stability measurements, and genotyping platforms. If the results of vGWAS were to be considered in a GS model, we call for developing such models that can directly incorporate peak-associated vGWAS signals while simultaneously allowing a GxE interaction term, in contrast to our approach that used residuals from a DGLM fitted prior to running the GS

model. We also recommend that simulation studies be used to see how different quantitative genetics parameters can influence the effectiveness of this approach in breeding for stability. These measures should be taken in future research to fully explore the extent to which using the results from a vGWAS can bolster GS-based approaches to breed for stability, if at all.

3.7 Figures

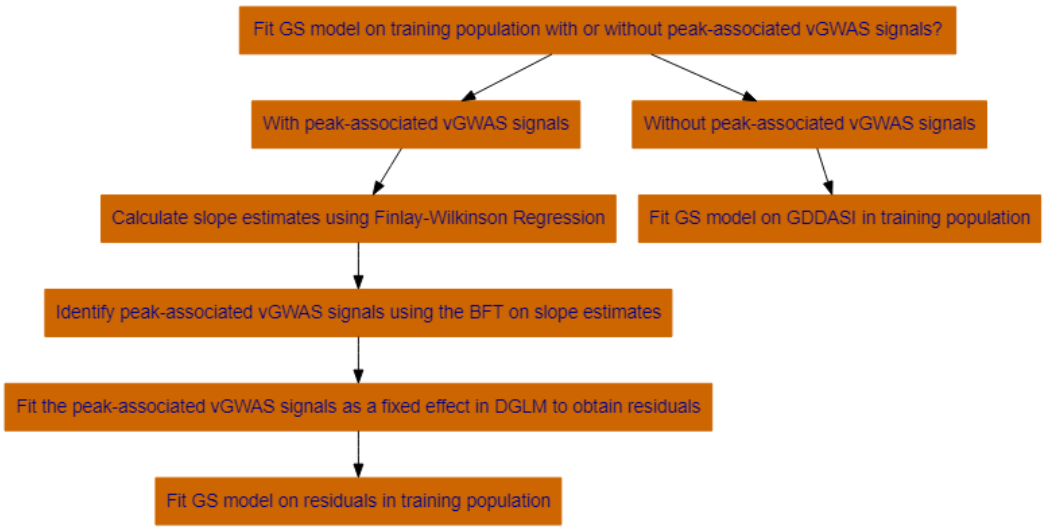

Figure 3.1. Flow chart. Summarizing the steps to perform genomic selection with or without incorporating a peak-associated vGWAS signal. If peak-associated vGWAS signals were incorporated, slope estimates were calculated using Finlay-Wilkinson Regression, run a variance genome-wide association study using the Brown-Forsythe Test (BFT), fit the peak-associated vGWAS signal as a fixed effect in DGLM to obtain residuals, and then run the genomic selection model on the residuals for each training population. If peak-associated vGWAS signals were not incorporated, the genomic selection model was fitted on the raw growing degree days to anthesis silking interval (GDDASI) measurements. Orange rectangles represent each step in this pipeline, and the black arrows represent how these steps are related

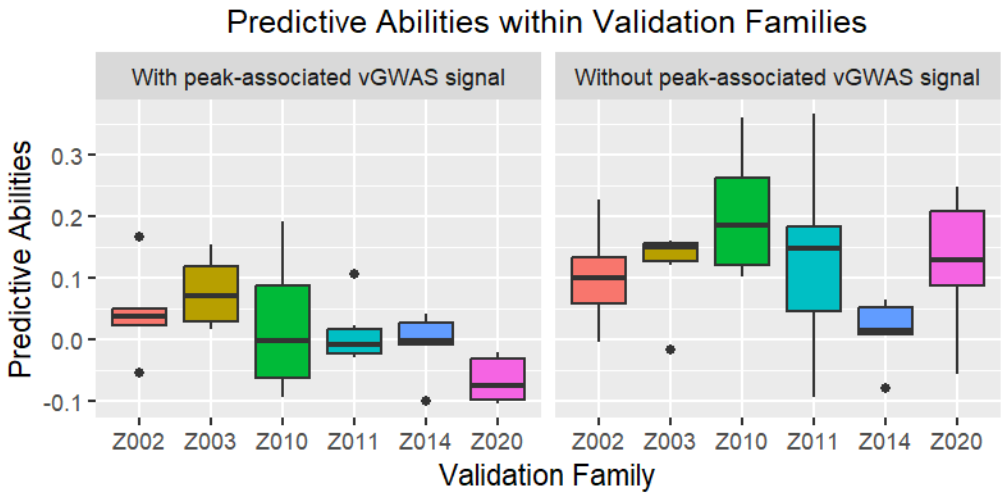

Figure 3.2. Distribution of predictive ability for each family across environments. The x-axis represents each of the tested nested association mapping families. Briefly, the NAM families used in this study include B73 x CML103 (Z002), B73 x CML228 (Z003), B73 x Hp301 (Z010), B73 x IL14H (Z011), B73 x Ky21 (Z014), and B73 x NC350 (Z020). The y-axis represents the Pearson correlation between Growing Degree Days Anthesis Silking Interval (GDDASI) and genomic estimated breeding values (GEBVs) (Predictive Ability). The left panel represents the scenario where the MxE model was ran with peak-associated vGWAS signals incorporated the scenario where the Marker by Environment (MxE) genomic selection model was ran without peak-associated variance genome-wide association study (vGWAS) signals incorporated. The right panel represents the scenario where the Marker by Environment (MxE) genomic selection model was ran without peak-associated variance genome-wide association study (vGWAS) signals incorporated

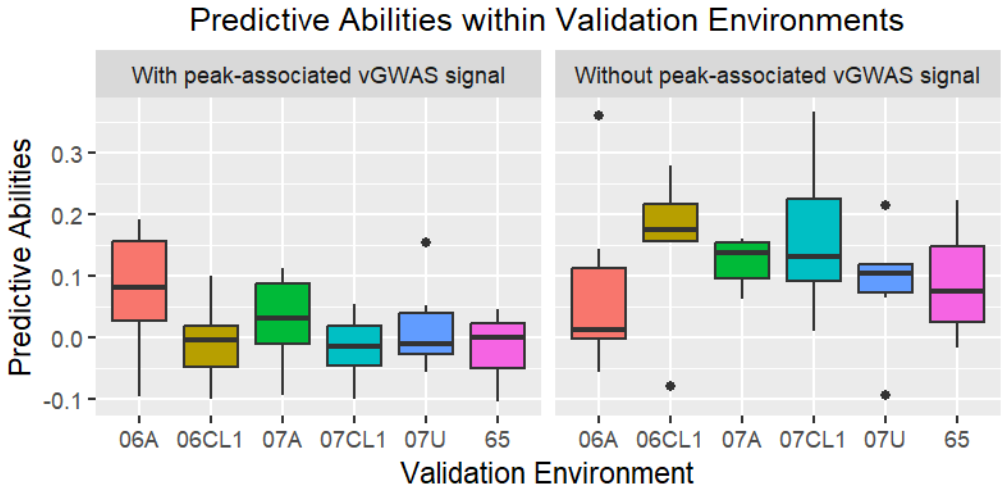

Figure 3.3. Distribution of predictive ability for each environment across families. The x-axis represents each of the tested environments. Briefly, the environments used in this study where Aurora 2006 (06A), Clayton 2006 (06CL1), Aurora 2007 (07A), Clayton 2007 (07CL1), Urbana 2007 (07U), and Urbana 2006 (65). The y-axis represents the Pearson correlation between Growing Degree Days Anthesis Silking Interval (GDDASI) and genomic estimated breeding values (GEBVs) (Predictive Ability). The left panel represents the scenario where the MxE model was ran with peak-associated vGWAS signals incorporated. The right panel represents the scenario where the Marker by Environment (MxE) genomic selection model was ran without peak-associated variance genome-wide association study (vGWAS) signals incorporated.

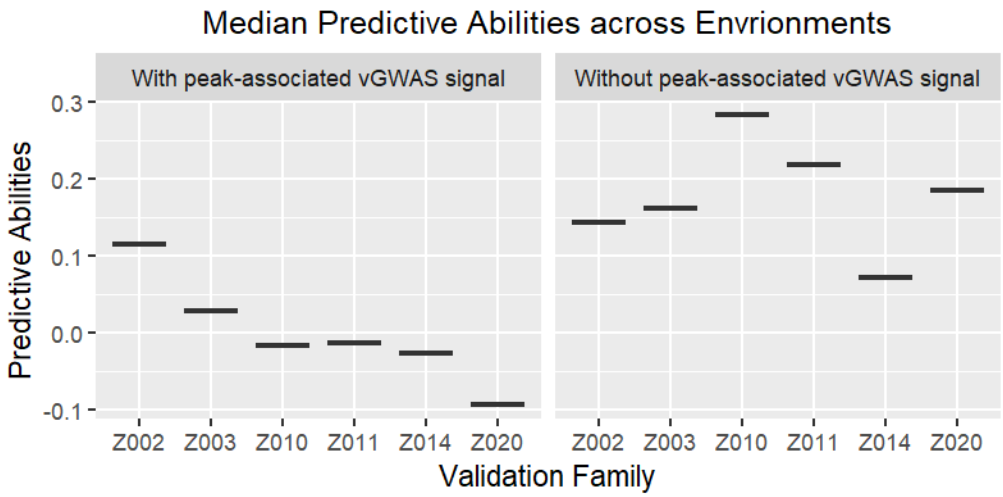

Figure 3.4. Distribution of the median predictive ability for each family across the tested environments. The x-axis represents each of the tested nested association mapping families. Briefly, the NAM families used in this study include B73 x CML103 (Z002), B73 x CML228 (Z003), B73 x Hp301 (Z010), B73 x IL14H (Z011), B73 x Ky21 (Z014), and B73 x NC350 (Z020). The y-axis represents the Predictive Ability from the median growing degree days anthesis silking interval (GDDASI) and the median genomic estimated breeding values (GEBVs) for each RIL across all environments. The left panel represents the scenario where the MxE model was ran with peak-associated vGWAS signals incorporated represents the scenario where the Marker by Environment (MxE) genomic selection model was ran without peak-associated variance genome-wide association study (vGWAS) signals incorporated. The right panel represents the scenario where the Marker by Environment (MxE) genomic selection model was ran without peak-associated variance genome-wide association study (vGWAS) signals incorporated.

CHAPTER 4: The parallel universe of vGWAS to traditional GWAS

While vQTLs can be a valuable tool in plant breeding and quantitative genetics, especially for GxE, further research must be conducted before realizing this full potential. While being first noted and described in the earlier days of when mean QTL mapping approaches were being developed, vQTLs have only recently received attention within the plant science community (Weller et al., 1988; Shen et al., 2012; Forsberg et al., 2015). While most plant science vGWAS publications have focused on using vGWAS to identify putatively epistatic loci, the application of vGWAS to elucidate regions likely to harbor GxE interactions is still in its infancy (Hussain et al., 2020; Li et al., 2020; Murphy et al., 2022; Song et al., 2022). Depending on the breeding goal, identifying GxE loci through vGWAS can help breeders identify such loci of small to moderate effect that may be missed from traditional GxE analyses due to the need to test every marker and environment interaction (Dempfle et al., 2008; Bustos-Korts et al., 2017; Gauderman et al., 2017). Breeders can either select against a vQTL allele to breed for stability and unpredictable growing environments or for a vQTL when introducing a crop to a new environment in its strictest sense (Desclaux et al., 2008; Kusmec et al., 2018; Langridge et al., 2021; Reckling et al., 2021).

The biggest hurdle in using vGWAS on a more widespread scale for GxE and other studies is the issue of false positives. Currently, the most widely used vGWAS statistical models, the BFT and DGLM, cannot explicitly control for familial relatedness. This problem can theoretically be indirectly addressed by extracting residuals from a linear mixed model like the unified linear mixed model in Yu et al. (2006). However, such an approach does not always guarantee that false positives will be controlled, as I observed from my work when using vGWAS to detect GxE loci in flowering-time traits in maize (Murphy and Lipka, in press).

Some alternative models can directly address the confounding factors from population substructure and familial relatedness, namely the HGLM and DHGLM (Lee and Nelder, 1996; Lee and Nelder, 2006; Rönnegård and Valdar, 2012). However, the bottleneck preventing the wide-scale deployment of these models is the computational burden required to fit them to the scale of modern plant genotypic and phenotypic data. For vGWAS to be deployed routinely in plant breeding, we need to further develop the computational infrastructure needed to fit these models.

Despite these bottlenecks in vGWAS, many parallels with traditional GWAS may provide exciting opportunities to expand the scope of vGWAS. So far, within plant science and breeding, vGWAS has been performed in *Arabidopsis*, maize, wheat, and mung bean. While we focused on flowering-time traits in maize, which did have vQTLs identified from independent germplasm, there are other traits where vQTLs have been identified such as plant height in maize. It is warranted to look at other germplasm to see if the same vQTLs are identified (Shen et al., 2012; Forsberg et al., 2015; Li et al., 2020; Ageev et al., 2021; Zhang and Qi, 2021; Song et al., 2022). Many crop and plant communities may benefit from using vGWAS to help aid in breeding for uniformity and stability, responsiveness, and, of course, as a way to prioritize genomic regions with epistasis and or GxE signals. There are some plant communities, especially, that may benefit from running vGWAS more often. For example, an important breeding objective in potato breeding is to breed for uniformity in potato size for ease of processing (Bradshaw et al., 2006). Another crop that may also benefit from breeding for uniformity is maize. Even though vQTLs have been described in maize, one maize breeding community that may benefit the most from vQTLs is sweetcorn, where uniformity in ear size,

nutrient profile, and healthy-looking ears is desired (Tracy, 1996). These are just a few examples where vQTLs may be valuable in plant breeding by selecting against vQTLs (Ordas et al., 2008).

Another way that vQTLs can be identified is in traditional single-marker linkage mapping approaches. The DGLM can be used for this exact purpose, as demonstrated in Corty and Valdar (2018). Using vQTL biparental mapping populations has yet to be explored in the plant breeding and science realm. Using a vQTL identified from biparental mapping will provide a complementary approach to vGWAS in a similar vein to traditional GWAS and QTL mapping (Tibbs Cortes et al., 2021). While QTL mapping in biparental crosses do not suffer from false positives arising from population structure issue and familial relatedness, biparental populations can severally limit the genetic diversity present in these populations (Keurentjes et al., 2011). Multi-parent populations and joint-linkage mapping can address this shortcoming (Scott et al., 2020); while such populations and approaches would be valuable to vQTL mapping, there has (to my current knowledge) yet to be an approach developed specifically for vQTL mapping.

Currently, vGWAS and vQTL mapping only can model a single trait at a time. While considering only a single trait may answer the biological question being asked, multiple traits in one quantitative genetics model may answer additional biological questions like pleiotropy (Auge *et al.*, 2019; Kim *et al.*, 2009; Rice *et al.*, 2020). There are theoretical benefits to looking at vQTLs through a multi-trait framework. Multi-trait statistical approaches can increase statistical power in detecting a vQTL by utilizing genetic and trait correlations. This would address the shortcoming of identifying a vQTL (and potentially a GxE or epistatic QTL) with improved precision, especially if one of the traits has either low heritability or small overall genetic effects (Rönnegård & Valdar, 2012). From the biological perspective, it has been noted that epistasis and GxE can be multi-trait in nature (Majumdar *et al.*, 2020; Konigorski and

Glicksberg, 2021). In theory, a multi-trait vGWAS approach may detect multi-trait GxE and epistasis. Interest has been pushed in breeding for stability for multiple traits at once (Olivoto et al., 2019). Selecting against a multi-trait vQTL allele may help realize this objective.

My work with genomic selection and peak-associated vGWAS signals has the greatest need for further development. While Mouresan et al. (2019) provided a framework for including peak-associated vGWAS signals for genomic prediction, further research is needed to develop and deploy this model into a plant breeding program geared towards selecting for or against stability. While my approach of incorporated peak vGWAS signals lowered the overall prediction accuracies, my work establishes a baseline for future studies to improve this approach.

All the statistical approaches employed in this this dissertation were frequentist. So far, only one Bayesian model called the Bayesian test for heteroskedasticity (BTH) has been proposed for vGWAS in Dumitrascu et al. (2019). A rationale that Dumitrascu et al. (2019) gave for using their Bayesian approach is that priors can give more reliable estimates for the mean and variance effects and can distinguish between a true vQTL and non-genetic sources or error. That said, this model performed on par with the DGLM in some instances, but the main limitation of deploying the BTH is the computational demands. However, the BTH would be worth investigating within plant breeding applications. In addition to BTH, Bayesian forms of DGLM and HGLM exist that can be modelled within R (Bonner et al., 2021). This may be worthwhile to investigate not just for vGWAS but also for GS.

The incorporation of time series into vGWAS approaches have great potential. To date, time series applications to traditional GWAS has been proposed as a way to identify more candidate loci involved with the developmental trajectory of a trait throughout the plant's life cycle (Moore et al., 2013; Miao et al., 2020). One intriguing observation noted by Li et al. (2020)

is that epistasis tended to be more enriched the earlier development stages of grain moisture, whereas GxE tended to be enriched for later developmental trajectories of grain moisture content in maize. Epistatic interactions have also been reported in other studies investigating time-series GWAS (Muraya et al., 2017; Knoch et al., 2020). These observations warrant the investigation of using a time-series vGWAS to aid in identifying additional epistatic and GxE loci in this framework.

In conclusion, vQTLs have the potential to aid in detecting GxE loci, but more development is needed to realize this potential. As vGWAS and accompanying analyses gain momentum in plant breeding for studying GxE, many things to consider come to mind. Inferences based off vGWAS could provide insight into not only GxE loci but for many different biological phenomena. At the same time, the current limitations of vGWAS identified in this dissertation, including false positives with current methods, computational demands of running more sophisticated vGWAS and vQTL models, and how to best incorporate vGWAS results into GS, need to be addressed in future research. Ultimately, mitigating these limitations could result in realizing the full potential of vGWAS to assist plant breeding efforts to study and better understand GxE in a changing global environment.

REFERENCES

Ageev, A., Lee, C. R., Ting, C. T., Schafleitner, R., Bishop-Von Wettberg, E., Nuzhdin, S. V., Samsonova, M., & Kozlov, K. (2021). Modeling of flowering time in vigna radiata with approximate bayesian computation. *Agronomy*, *11*(11), 1–16. https://doi.org/10.3390/agronomy11112317

Agresti, A. (2003). *Categorical data analysis* (Vol. 482). John Wiley & Sons.

Alonso-Blanco, C., Andrade, J., Becker, C., Bemm, F., Bergelson, J., Borgwardt, K. M. M., Cao, J., Chae, E., Dezwaan, T. M. M., Ding, W., Ecker, J. R. R., Exposito-Alonso, M., Farlow, A., Fitz, J., Gan, X., Grimm, D. G. G., Hancock, A. M. M., Henz, S. R. R., Holm, S., … Zhou, X. (2016). 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. *Cell*, *166*(2), 481–491. https://doi.org/10.1016/j.cell.2016.05.063

Ankamah-Yeboah, T., Janss, L. L., Jensen, J. D., Hjortshøj, R. L., & Rasmussen, S. K. (2020). Genomic Selection Using Pedigree and Marker-by-Environment Interaction for Barley Seed Quality Traits From Two Commercial Breeding Programs. *Frontiers in Plant Science*, *11*(May). https://doi.org/10.3389/fpls.2020.00539

Ansarifar, J., Akhavizadegan, F., & Wang, L. (2020). Performance prediction of crosses in plant breeding through genotype by environment interactions. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-68343-1

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

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 (Methodological)*, *57*(1), 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bernardo, R. (2010). *Breeding for quantitative traits in plants* (Vol. 1, p. 369). Woodbury: Stemma press.

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

Bonat, W. H., & Jørgensen, B. (2016). Multivariate covariance generalized linear models. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, *65*(5), 649–675.

https://doi.org/10.1111/rssc.12145

Bonhomme, R., Derieux, M., & Edmeades, G. O. (1994). Flowering of diverse maize cultivars in relation to temperature and photoperiod in multilocation field trials. *Crop Science*, *34*(1), 156–164. https://doi.org/10.2135/cropsci1994.0011183X003400010028x

Bonner, S., Mutzel, A., Kim, H. N., Westneat, D., Wright, J., & Schofield, M. (2021). dalmatian: A Package for Fitting Double Hierarchical Linear Models in R via JAGS and nimble. *Journal of Statistical Software*, *100*(10). https://doi.org/10.18637/JSS.V100.I10

Bouchet, S., Servin, B., Bertin, P., Madur, D., Combes, V., Dumas, F., Brunel, D., Laborde, J., Charcosset, A., & Nicolas, S. (2013). Adaptation of Maize to Temperate Climates: Mid-Density Genome-Wide Association Genetics and Diversity Patterns Reveal Key Genomic Regions, with a Major Contribution of the Vgt2 (ZCN8) Locus. *PLoS ONE*, *8*(8). https://doi.org/10.1371/journal.pone.0071377

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. https://doi.org/10.1093/bioinformatics/btm308

Bradshaw, J. E., Bryan, G. J., & Ramsay, G. (2006). Genetic resources (including wild and cultivated Solanum species) and progress in their utilisation in potato breeding. *Potato Research*, *49*(1), 49–65. https://doi.org/10.1007/s11540-006-9002-5

Braz, C. U., Rowan, T. N., Schnabel, R. D., & Decker, J. E. (2021). Genome-wide association analyses identify genotype-by-environment interactions of growth traits in Simmental cattle. *Scientific Reports*, *11*(1), 13335. https://doi.org/10.1038/s41598-021-92455-x

Brown, M. B., & Forsythe, A. B. (1974). The Small Sample Behavior of Some Statistics Which Test the Equality of Several. *Technometrics*, *16*(1), 129–132. https://doi.org/10.1080/00401706.1974.10489158

Brown, M. E., & Funk, C. C. (2008). Food security under climate change. *Science*, *319*(5863), 580-581. https://doi.org/10.1126/science.1154102

Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., Ersoz, E., Flint-Garcia, S., Garcia, A., Glaubitz, J. C., Goodman, M. M., Harjes, C., Guill, K., Kroon, D. E., Larsson, S., Lepak, N. K., Li, H., Mitchell, S. E., Pressoir, G., … McMullen, M. D. (2009). The genetic architecture of maize flowering time. *Science*, *325*(5941), 714–718. https://doi.org/10.1126/science.117426

Bukowski, R., Guo, X., Lu, Y., Zou, C., He, B., Rong, Z., Wang, B., Xu, D., Yang, B., Xie, C., Fan, L., Gao, S., Xu, X., Zhang, G., Li, Y., Jiao, Y., Doebley, J. F., Ross-Ibarra, J., Lorant, A., … Xu, Y. (2018). Construction of the third-generation Zea mays haplotype map. *GigaScience*, *7*(4), 1–12. https://doi.org/10.1093/gigascience/gix134

Burgueño, J., de los Campos, G., Weigel, K., & Crossa, J. (2012). Genomic prediction of breeding values when modeling genotype× environment interaction using pedigree and dense molecular markers. *Crop Science*, *52*(2), 707-719. https://doi.org/10.2135/cropsci2011.06.0299

Bustos-Korts, D., Malosetti, M., Chapman, S., & Eeuwijk, F. V. (2016). Modelling of genotype by environment interaction and prediction of complex traits across multiple environments as a synthesis of crop growth modelling, genetics and statistics. In *Crop systems biology* (pp. 55-82). Springer, Cham.

Cai, X., Ballif, J., Endo, S., Davis, E., Liang, M., Chen, D., Dewald, D., Kreps, J., Zhu, T., & Wu, Y. (2007). A putative CCAAT-binding transcription factor is a regulator of flowering timing in arabidopsis. *Plant Physiology*, *145*(1), 98–105. https://doi.org/10.1104/pp.107.102079

Ceccarelli, S., & Grando, S. (2020). Evolutionary Plant Breeding as a Response to the Complexity of Climate Change. *IScience*, *23*(12), 101815. https://doi.org/10.1016/j.isci.2020.101815

Chen, Q., Zhong, H., Fan, X. W., & Li, Y. Z. (2015). An insight into the sensitivity of maize to photoperiod changes under controlled conditions. *Plant Cell and Environment*, *38*(8), 1479–1489. https://doi.org/10.1111/pce.12361

Chen, M., Ji, M., Wen, B., Liu, L., Li, S., Chen, X., Gao, D., & Li, L. (2016). GOLDEN 2-LIKE transcription factors of plants. *Frontiers in Plant Science*, *7*(OCTOBER2016), 1–5. https://doi.org/10.3389/fpls.2016.01509

Cheng, M., McCarl, B., & Fei, C. (2022). Climate change and livestock production: A literature review. *Atmosphere*, *13*(1). https://doi.org/10.3390/atmos13010140

Clopper, C. J., & Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, *26*(4), 404-413. https://doi.org/10.2307/2331986

Coles, N. D., McMullen, M. D., Balint-Kurti, P. J., Pratt, R. C., & Holland, J. B. (2010). Genetic

control of photoperiod sensitivity in maize revealed by joint multiple population analysis. *Genetics*, *184*(3), 799-812. https://doi.org/10.1534/genetics.109.110304

Cook, J. P., McMullen, M. D., Holland, J. B., Tian, F., Bradbury, P., Ross-Ibarra, J., ... & Flint-Garcia, S. A. (2012). Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant physiology*, *158*(2), 824-834. https://doi.org/10.1104/pp.111.185033

Cordell, H. J. (2002). Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Human molecular genetics*, *11*(20), 2463-2468. https://doi.org/10.1093/hmg/11.20.2463

Córdova-Palomera, A., van der Meer, D., Kaufmann, T., Bettella, F., Wang, Y., Alnæs, D., ... & Westlye, L. T. (2021). Genetic control of variability in subcortical and intracranial volumes. *Molecular psychiatry*, *26*(8), 3876-3883. https://doi.org/10.1038/s41380-020-0664-1

Cornelius, P. L. (1993). Statistical tests and retention of terms in the additive main effects and multiplicative interaction model for cultivar trials. *Crop Science*, *33*(6), 1186–1193. https://doi.org/10.2135/cropsci1993.0011183X003300060016x

Corty, R. W., & Valdar, W. (2018). QTL mapping on a background of variance heterogeneity. *G3: Genes, Genomes, Genetics*, *8*(12), 3767–3782. https://doi.org/10.1534/g3.118.200790

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., Burgueño, J., González-Camacho, J. M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., & Varshney, R. K. (2017). Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends in Plant Science*, *22*(11), 961–975. https://doi.org/10.1016/j.tplants.2017.08.011

Cuevas, J., Crossa, J., Montesinos-López, O. A., Burgueño, J., Pérez-Rodríguez, P., & de Los Campos, G. (2017). Bayesian genomic prediction with genotype× environment interaction kernel models. *G3: Genes, Genomes, Genetics*, *7*(1), 41-53. https://doi.org/10.1534/g3.116.035584

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

de Oliveira Silva, A., Slafer, G. A., Fritz, A. K., & Lollato, R. P. (2020). Physiological Basis of Genotypic Response to Management in Dryland Wheat. *Frontiers in Plant Science*, *10*(January), 1–19. https://doi.org/10.3389/fpls.2019.01644

Debat, V., & David, P. (2001). Mapping phenotypes: Canalization, plasticity and developmental stability. *Trends in Ecology and Evolution*, *16*(10), 555–561. https://doi.org/10.1016/S0169-5347(01)02266-2

Dell'Acqua, M., Gatti, D. M., Pea, G., Cattonaro, F., Coppens, F., Magris, G., Hlaing, A. L., Aung, H. H., Nelissen, H., Baute, J., Frascaroli, E., Churchill, G. A., Inzé, D., Morgante, M., & Pè, M. E. (2015). Genetic properties of the MAGIC maize population: A new platform for high definition QTL mapping in Zea mays. *Genome Biology*, *16*(1), 1–23. https://doi.org/10.1186/s13059-015-0716-z

Dempfle, A., Scherag, A., Hein, R., Beckmann, L., Chang-Claude, J., & Schäfer, H. (2008). Gene-environment interactions for complex traits: Definitions, methodological requirements and challenges. *European Journal of Human Genetics*, *16*(10), 1164–1172. https://doi.org/10.1038/ejhg.2008.106

Des Marais, D. L., Hernandez, K. M., & Juenger, T. E. (2013). Genotype-by-environment interaction and plasticity: Exploring genomic responses of plants to the abiotic environment. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 5–29. https://doi.org/10.1146/annurev-ecolsys-110512-135806

Desclaux, D., Nolot, J. M., Chiffoleau, Y., Gozé, E., & Leclerc, C. (2008). Changes in the concept of genotype × environment interactions to fit agriculture diversification and decentralized participatory plant breeding: Pluridisciplinary point of view. *Euphytica*, *163*(3), 533–546. https://doi.org/10.1007/s10681-008-9717-2

Devlin, B., & Roeder, K. (1999). Genomic control for association studies. *Biometrics*, *55*(4),997-1004. https://doi.org/10.1111/j.0006-341X.1999.00997.x

Diepenbrock, C. H., Ilut, D. C., Magallanes-Lundback, M., Kandianis, C. B., Lipka, A. E., Bradbury, P. J., Holland, J. B., Hamilton, J. P., Wooldridge, E., Vaillancourt, B., Góngora-Castillo, E., Wallace, J. G., Cepela, J., Mateos-Hernandez, M., Owens, B. F., Tiede, T., Buckler, E. S., Rocheford, T., Buell, C. R., … DellaPenna, D. (2021). Eleven biosynthetic genes explain the majority of natural variation in carotenoid levels in maize grain. *Plant Cell*, *33*(4), 882–900. https://doi.org/10.1093/plcell/koab032

Diepenbrock, C. H., Kandianis, C. B., Lipka, A. E., Magallanes-Lundback, M., Vaillancourt, B., Góngora-Castillo, E., Wallace, J. G., Cepela, J., Mesberg, A., Bradbury, P. J., Ilut, D. C., Mateos-Hernandez, M., Hamilton, J., Owens, B. F., Tiede, T., Buckler, E. S., Rocheford, T., Buell, C. R., Gore, M. A., & DellaPenna, D. (2017). Novel loci underlie natural variation in vitamin E levels in maize grain. *Plant Cell*, *29*(10), 2374–2392. https://doi.org/10.1105/tpc.17.00475

Dumitrascu, B., Darnell, G., Ayroles, J., & Engelhardt, B. E. (2019). Statistical tests for detecting variance effects in quantitative trait studies. *Bioinformatics*, *35*(2), 200–210. https://doi.org/10.1093/bioinformatics/bty565

Dunn P.K., Smyth G.K., Dunn M.P.K. (2020) Package 'dglm'

van Eeuwijk, F. A., Bink, M. C., Chenu, K., & Chapman, S. C. (2010). Detection and use of QTL for complex traits in multiple environments. *Current opinion in plant biology*, *13*(2), 193-205. https://dx.doi.org/10.1016/j.pbi.2010.01.001

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), 1–10. https://doi.org/10.1371/journal.pone.0019379

Fernandes, S. B., & Lipka, A. E. (2020). simplePHENOTYPES: SIMulation of pleiotropic, linked and epistatic phenotypes. *BMC Bioinformatics*, *21*(1), 1–10. https://doi.org/10.1186/s12859-020-03804-y

Finlay, K. W., & Wilkinson, G. N. (1963). The analysis of adaptation in a plant-breeding programme. *Australian journal of agricultural research*, *14*(6), 742-754. https://doi.org/10.1071/AR9630742

Fisher, R. A. (1918). Reproduced by permission of the Royal Society of Edinburgh from Transactions of the Society, vol. 52: 399-433 (1918). *Society*, *52*, 399–433.

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

Forsberg, S. K. G., Andreatta, M. E., Huang, X. Y., Danku, J., Salt, D. E., & Carlborg, Ö. (2015). The Multi-allelic Genetic Architecture of a Variance-Heterogeneity Locus for

Molybdenum Concentration in Leaves Acts as a Source of Unexplained Additive Genetic Variance. In *PLoS Genetics* (Vol. 11, Issue 11). https://doi.org/10.1371/journal.pgen.1005648

Forsberg, S. K., & Carlborg, Ö. (2017). On the relationship between epistasis and genetic variance heterogeneity. *Journal of experimental botany*, *68*(20), 5431-5438. https://doi.org/10.1093/jxb/erx283

Fox, J., & Weisberg, S. (2019). Using car functions in other functions.

Gage, J. L., Jarquin, D., Romay, C., Lorenz, A., Buckler, E. S., Kaeppler, S., ... & De Leon, N. (2017). The effect of artificial selection on phenotypic plasticity in maize. *Nature communications*, *8*(1), 1348. https://doi.org/10.1038/s41467-017-01450-2

Gage, J. L., de Leon, N., & Clayton, M. K. (2018). Comparing genome-wide association study results from different measurements of an underlying phenotype. *G3: Genes, Genomes, Genetics*, *8*(11), 3715-3722. https://doi.org/10.1534/g3.118.200700

Gage, J. L., Monier, B., Giri, A., & Buckler, E. S. (2020). Ten years of the maize nested association mapping population: Impact, limitations, and future directions. *Plant Cell*, *32*(7), 2083–2093. https://doi.org/10.1105/tpc.19.00951

Gauderman, W. J., Mukherjee, B., Aschard, H., Hsu, L., Lewinger, J. P., Patel, C. J., ... & Chatterjee, N. (2017). Update on the state of the science for analytical methods for gene-environment interactions. *American journal of epidemiology*, *186*(7), 762-770. https://doi.org/10.1093/aje/kwx228

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0090346

Goddard, M. E., & Hayes, B. J. (2007). Genomic selection. *Journal of Animal breeding and Genetics*, *124*(6), 323-330. https://doi.org/10.1111/j.1439-0388.2007.00702.x

Guo, T., Mu, Q., Wang, J., Vanous, A. E., Onogi, A., Iwata, H., Li, X., & Yu, J. (2020). Dynamic effects of interacting genes underlying rice flowering-time phenotypic plasticity and global adaptation. *Genome Research*, *30*(5), 673–683. https://doi.org/10.1101/gr.255703.119

Heffner, E. L., Lorenz, A. J., Jannink, J. L., & Sorrells, M. E. (2010). Plant breeding with Genomic selection: Gain per unit time and cost. *Crop Science*, *50*(5), 1681–1690.

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

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*, 463-480. https://doi.org/10.1007/s00122-013-2231-5

Hill, W. G., & Zhang, X. S. (2004). Effects on phenotypic variability of directional selection arising through genetic differences in residual variability. *Genetics Research*, *83*(2), 121-132. https://doi:10.1017/S0016672304006640

Hill, W. G., & Mulder, H. A. (2010). Genetic analysis of environmental variation. *Genetics research*, *92*(5-6), 381-395. https://doi:10.1017/S0016672310000546

Hong, C., Ning, Y., Wei, P., Cao, Y., & Chen, Y. (2017). A semiparametric model for vQTL mapping. *Biometrics*, *73*(2), 571–581. https://doi.org/10.1111/biom.12612

Hung, H. Y., Browne, C., Guill, K., Coles, N., Eller, M., Garcia, A., Lepak, N., Melia-Hancock, S., Oropeza-Rosas, M., Salvo, S., Upadyayula, N., Buckler, E. S., Flint-Garcia, S., McMullen, M. D., Rocheford, T. R., & Holland, J. B. (2012a). The relationship between parental genetic or phenotypic divergence and progeny variation in the maize nested association mapping population. *Heredity*, *108*(5), 490–499. https://doi.org/10.1038/hdy.2011.103

Hung, H. Y., Shannon, L. M., Tian, F., Bradbury, P. J., Chen, C., Flint-Garcia, S. A., ... & Holland, J. B. (2012b). ZmCCT and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. *Proceedings of the National Academy of Sciences*, *109*(28), E1913-E1921. https://doi.org/10.1073/pnas.1203189109

Hussain, W., Campbell, M. T., Jarquin, D., Walia, H., & Morota, G. (2020). Variance heterogeneity genome-wide mapping for cadmium in bread wheat reveals novel genomic loci and epistatic interactions. *Plant Genome*, *13*(1), 1–13. https://doi.org/10.1002/tpg2.20011

Iannone, R., & Iannone, M. R. (2022). Package 'DiagrammeR'.

Izawa, T. (2007). Adaptation of flowering-time by natural and artificial selection in Arabidopsis and rice. *Journal of Experimental Botany*, *58*(12), 3091–3097. https://doi.org/10.1093/jxb/erm159

Jarquín, 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*, 595-607. https://doi.org/10.1007/s00122-013-2243-1

Jarquin, D., Kajiya-Kanegae, H., Taishen, C., Yabe, S., Persa, R., Yu, J., Nakagawa, H., Yamasaki, M., & Iwata, H. (2020). Coupling day length data and genomic prediction tools for predicting time-related traits under complex scenarios. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-70267-9

Jarquin, D., de Leon, N., Romay, C., Bohn, M., Buckler, E. S., Ciampitti, I., Edwards, J., Ertl, D., Flint-Garcia, S., Gore, M. A., Graham, C., Hirsch, C. N., Holland, J. B., Hooker, D., Kaeppler, S. M., Knoll, J., Lee, E. C., Lawrence-Dill, C. J., Lynch, J. P., … Lorenz, A. (2021). Utility of Climatic Information via Combining Ability Models to Improve Genomic Prediction for Yield Within the Genomes to Fields Maize Project. *Frontiers in Genetics*, *11*(March), 1–11. https://doi.org/10.3389/fgene.2020.592769

Al Kawam, A., Alshawaqfeh, M., Cai, J., Serpedin, E., & Datta, A. (2017, August). Simulating variance heterogeneity in quantitative genome wide association studies. In *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics* (pp. 762-763). https://doi.org/10.1145/3107411.3110407

Keurentjes, J. J. B., Willems, G., Van Eeuwijk, F., Nordborg, M., & Koornneef, M. (2011). A comparison of population types used for QTL mapping in Arabidopsis thaliana. *Plant Genetic Resources: Characterisation and Utilisation*, *9*(2), 185–188. https://doi.org/10.1017/S1479262111000086

Kitano, H. (2004). Biological robustness. *Nature Reviews Genetics*, *5*(11), 826–837. https://doi.org/10.1038/nrg1471

Kliem, L., & Sievers-Glotzbach, S. (2022). Seeds of resilience: the contribution of commons-based plant breeding and seed production to the social-ecological resilience of the agricultural sector. *International Journal of Agricultural Sustainability*, *20*(4), 595–614. https://doi.org/10.1080/14735903.2021.1963598

Knoch, D., Abbadi, A., Grandke, F., Meyer, R. C., Samans, B., Werner, C. R., Snowdon, R. J., & Altmann, T. (2020). Strong temporal dynamics of QTL action on plant growth progression revealed through high-throughput phenotyping in canola. *Plant Biotechnology Journal*, *18*(1), 68–82. https://doi.org/10.1111/pbi.13171

Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods*, *9*(1), 1. https://doi.org/10.1186/1746-4811-9-29

Kusmec, A., de Leon, N., & Schnable, P. S. (2018). Harnessing phenotypic plasticity to improve maize yields. *Frontiers in Plant Science*, *9*(September), 1–4. https://doi.org/10.3389/fpls.2018.01377

Langridge, P., Braun, H., Hulke, B., Ober, E., & Prasanna, B. M. (2021). Breeding crops for climate resilience. *Theoretical and Applied Genetics*, *134*(6), 1607–1611. https://doi.org/10.1007/s00122-021-03854-7

Lee, Y., & Nelder, J. A. (1996). Hierarchical Generalized Linear Models. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(4), 619–656. https://doi.org/10.1111/j.2517-6161.1996.tb02105.x

Lee, Youngjo, & Nelder, J. A. (2006). Double hierarchical generalized linear models. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, *55*(2), 139–185. https://doi.org/10.1111/j.1467-9876.2006.00538.x

Leng, G., & Huang, M. (2017). Crop yield response to climate change varies with crop spatial distribution pattern. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/s41598-017-01599-2

Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., Han, Y., Chai, Y., Guo, T., Yang, N., Liu, J., Warburton, M. L., Cheng, Y., Hao, X., Zhang, P., Zhao, J., Liu, Y., Wang, G., Li, J., & Yan, J. (2013). Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nature Genetics*, *45*(1), 43–50. https://doi.org/10.1038/ng.2484

Li, H., Wang, M., Li, W., He, L., Zhou, Y., Zhu, J., Che, R., Warburton, M. L., Yang, X., & Yan, J. (2020). Genetic variants and underlying mechanisms influencing variance heterogeneity in maize. *Plant Journal*, *103*(3), 1089–1102. https://doi.org/10.1111/tpj.14786

Li, M., Zhang, Y. W., Zhang, Z. C., Xiang, Y., Liu, M. H., Zhou, Y. H., ... & Zhang, Y. M. (2022). A compressed variance component mixed model for detecting QTNs and QTN-by-environment and QTN-by-QTN interactions in genome-wide association studies. *Molecular Plant*, *15*(4), 630-650. https://doi.org/10.1016/j.molp.2022.02.012

Li, X., Guo, T., Mu, Q., Li, X., & Yu, J. (2018). Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(26), 6679–6684. https://doi.org/10.1073/pnas.1718326115

Li, Y. X., Li, C., Bradbury, P. J., Liu, X., Lu, F., Romay, C. M., Glaubitz, J. C., Wu, X., Peng, B., Shi, Y., Song, Y., Zhang, D., Buckler, E. S., Zhang, Z., Li, Y., & Wang, T. (2016). Identification of genetic variants associated with maize flowering time using an extremely large multi-genetic background population. *The Plant Journal : For Cell and Molecular Biology*, *86*(5), 391–402. https://doi.org/10.1111/tpj.13174

Lian, L., & De Los Campos, G. (2016). FW: An R package for Finlay-Wilkinson regression that incorporates genomic/pedigree information and covariance structures between environments. *G3: Genes, Genomes, Genetics*, *6*(3), 589–597. https://doi.org/10.1534/g3.115.026328

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., & Zhang, Z. (2012). GAPIT: Genome association and prediction integrated tool. *Bioinformatics*, *28*(18), 2397–2399. https://doi.org/10.1093/bioinformatics/bts444

Lipka, A. E., Gore, M. A., Magallanes-Lundback, M., Mesberg, A., Lin, H., Tiede, T., ... & DellaPenna, D. (2013). Genome-wide association study and pathway-level analysis of tocochromanol levels in maize grain. *G3: Genes, Genomes, Genetics*, *3*(8), 1287-1299. https://doi.org/10.1534/g3.113.006148

Lipka, A. E., Kandianis, C. B., Hudson, M. E., Yu, J., Drnevich, J., Bradbury, P. J., & Gore, M. A. (2015). From association to prediction: Statistical methods for the dissection and selection of complex traits in plants. *Current Opinion in Plant Biology*, *24*, 110–118. https://doi.org/10.1016/j.pbi.2015.02.010

Liu, K., Goodman, M., Muse, S., Smith, J. S., Buckler, E. D., & Doebley, J. (2003). Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics*, *165*(4), 2117-2128. https://doi.org/10.1093/genetics/165.4.2117

Lopez-Cruz, M., Crossa, J., Bonnett, D., Dreisigacker, S., Poland, J., Jannink, J. L., Singh, R. P., Autrique, E., & de los Campos, G. (2015). Increased prediction accuracy in wheat breeding trials using a marker × environment interaction genomic selection model. *G3: Genes, Genomes, Genetics*, *5*(4), 569–582. https://doi.org/10.1534/g3.114.016097

Mahmood, T., Ahmed, T., & Trethowan, R. (2022). Genotype x Environment x Management (GEM) Reciprocity and Crop Productivity. *Frontiers in Agronomy*, *4*(June), 1–11. https://doi.org/10.3389/fagro.2022.800365

Malosetti, M., Ribaut, J. M., & van Eeuwijk, F. A. (2013). The statistical analysis of multi-environment data: Modeling genotype-by-environment interaction and its genetic basis. *Frontiers in Physiology*, *4 MAR*(March), 1–17. https://doi.org/10.3389/fphys.2013.00044

McMullen, M. D., Kresovich, S., Villeda, H. S., Bradbury, P., Li, H., Sun, Q., Flint-Garcia, S., Thornsberry, J., Acharya, C., Bottoms, C., Brown, P., Browne, C., Eller, M., Guill, K., Harjes, C., Kroon, D., Lepak, N., Mitchell, S. E., Peterson, B., … Buckler, E. S. (2009). Genetic properties of the maize nested association mapping population. *Science*, *325*(5941), 737–740. https://doi.org/10.1126/science.1174320

Metz, C. E. (1978, October). Basic principles of ROC analysis. In *Seminars in nuclear medicine* (Vol. 8, No. 4, pp. 283-298). WB Saunders.

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

Miao, C., Xu, Y., Liu, S., Schnable, P. S., & Schnable, J. C. (2020). Increased power and accuracy of causal locus identification in time series genome-wide association in sorghum1[OPEN]. *Plant Physiology*, *183*(4), 1898–1909. https://doi.org/10.1104/pp.20.00277

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. https://doi.org/10.1534/g3.115.021667

Moore, C. R., Johnson, L. S., Kwak, I. Y., Livny, M., Broman, K. W., & Spalding, E. P. (2013). High-throughput computer vision introduces the time axis to a quantitative trait map of a plant growth response. *Genetics*, *195*(3), 1077–1086. https://doi.org/10.1534/genetics.113.153346

Mouresan, E. F., Selle, M., & Rönnegård, L. (2019). Genomic prediction including SNP-specific variance predictors. *G3: Genes, Genomes, Genetics*, *9*(10), 3333–3343. https://doi.org/10.1534/g3.119.400381

Mulder, H. A., Bijma, P., & Hill, W. G. (2007). Prediction of breeding values and selection

responses with genetic heterogeneity of environmental variance. *Genetics*, *175*(4), 1895-1910. https://doi.org/10.1534/genetics.106.063743

Muraya, M. M., Chu, J., Zhao, Y., Junker, A., Klukas, C., Reif, J. C., & Altmann, T. (2017). Genetic variation of growth dynamics in maize (Zea mays L.) revealed through automated non-invasive phenotyping. *Plant Journal*, *89*(2), 366–380. https://doi.org/10.1111/tpj.13390

Murphy, M. D., Fernandes, S. B., Morota, G., & Lipka, A. E. (2022). Assessment of two statistical approaches for variance genome-wide association studies in plants. *Heredity*, *June 2021*. https://doi.org/10.1038/s41437-022-00541-1

Nelson, P. T., Krakowsky, M. D., Coles, N. D., Holland, J. B., Bubeck, D. M., Smith, J. S. C., & Goodman, M. M. (2016). Genetic characterization of the North Carolina State university maize lines. *Crop Science*, *56*(1), 259–275. https://doi.org/10.2135/cropsci2015.09.0532

Nordborg, M., & Weigel, D. (2008). Next-generation genetics in plants. *Nature*, *456*(7223), 720–723. https://doi.org/10.1038/nature07629

Ogut, F., Bian, Y., Bradbury, P. J., & Holland, J. B. (2015). Joint-multiple family linkage analysis predicts within-family variation better than single-family analysis of the maize nested association mapping population. *Heredity*, *114*(6), 552–563. https://doi.org/10.1038/hdy.2014.123

Olivoto, T., Lúcio, A. D., da Silva, J. A., Sari, B. G., & Diel, M. I. (2019). Mean performance and stability in multi-environment trials II: Selection based on multiple traits. *Agronomy Journal*, *111*(6), 2961-2969. https://doi.org/10.2134/agronj2019.03.0221

Ordas, B., Malvar, R. A., & Hill, W. G. (2008). Genetic variation and quantitative trait loci associated with developmental stability and the environmental correlation between traits in maize. *Genetics Research*, *90*(5), 385–395. https://doi.org/10.1017/S0016672308009762

Orr, H. A. (1998). The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution*, *52*(4), 935–949. https://doi.org/10.1111/j.1558-5646.1998.tb01823.x

Park, C. J., & Seo, Y. S. (2015). Heat shock proteins: a review of the molecular chaperones for plant immunity. *The plant pathology journal*, *31*(4), 323. https://doi.org/10.5423/PPJ.RW.08.2015.0150

Peiffer, J. A., Romay, M. C., Gore, M. A., Flint-Garcia, S. A., Zhang, Z., Millard, M. J., ... & Buckler, E. S. (2014). The genetic architecture of maize height. *Genetics*, *196*(4), 1337-

1356. https://doi.org/10.1534/genetics.113.159152

Pettersson, M. E., & Carlborg, Ö. (2015). Capacitating epistasis—detection and role in the genetic architecture of complex traits. *Epistasis: Methods and Protocols*, 185-196. https://doi.org/10.1007/978-1-4939-2155-3_10

Piepho, H. P. (1995). Robustness of statistical tests for multiplicative terms in the additive main effects and multiplicative interaction model for cultivar trials. *Theoretical and Applied Genetics*, *90*(3–4), 438–443. https://doi.org/10.1007/BF00221987

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. https://doi.org/10.1038/ng1847

Price, K., Jackson, C. R., Parker, A. J., Reitan, T., Dowd, J., & Cyterski, M. (2011). Effects of watershed land use and geomorphology on stream low flows during severe drought conditions in the southern Blue Ridge Mountains, Georgia and North Carolina, United States. *Water Resources Research*, *47*(2). https://doi.org/10.1029/2010WR009340

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., & 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. https://doi.org/10.1086/519795

Raza, A., Razzaq, A., Mehmood, S. S., Zou, X., Zhang, X., Lv, Y., & Xu, J. (2019). Impact of climate change on crops adaptation and strategies to tackle its outcome: A review. *Plants*, *8*(2). https://doi.org/10.3390/plants8020034

Reckling, M., Ahrends, H., Chen, T. W., Eugster, W., Hadasch, S., Knapp, S., Laidig, F., Linstädter, A., Macholdt, J., Piepho, H. P., Schiffers, K., & Döring, T. F. (2021). Methods of yield stability analysis in long-term field experiments. A review. *Agronomy for Sustainable Development*, *41*(2). https://doi.org/10.1007/s13593-021-00681-4

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. *The Plant Genome*, *12*(1), 180052. https://doi.org/10.3835/plantgenome2018.07.0052

Rice, B. R., Fernandes, S. B., & Lipka, A. E. (2020). Multi-trait genome-wide association studies reveal loci associated with maize inflorescence and leaf architecture. *Plant and Cell Physiology*, *61*(8), 1427–1437. https://doi.org/10.1093/pcp/pcaa039

Romay, M. C., Millard, M. J., Glaubitz, J. C., Peiffer, J. A., Swarts, K. L., Casstevens, T. M., ... & Gardner, C. A. (2013). Comprehensive genotyping of the USA national maize inbred seed bank. *Genome biology*, *14*, 1-18. https://doi.org/10.1186/gb-2013-14-6-r55

Rönnegård, L., Felleki, M., Fikse, F., Mulder, H. A., & Strandberg, E. (2010). Genetic heterogeneity of residual variance-estimation of variance components using double hierarchical generalized linear models. *Genetics Selection Evolution*, *42*, 1-10. https://doi.org/10.1186/1297-9686-42-8

Rönnegård, L., & Valdar, W. (2011). Detecting major genetic loci controlling phenotypic variability in experimental crosses. *Genetics*, *188*(2), 435–447. https://doi.org/10.1534/genetics.111.127068

Rönnegård, L., & Valdar, W. (2012). Recent developments in statistical methods for detecting genetic loci affecting phenotypic variability. *BMC Genetics*, *13*. https://doi.org/10.1186/1471-2156-13-63

Roorkiwal, M., Jarquin, D., Singh, M. K., Gaur, P. M., Bharadwaj, C., Rathore, A., Howard, R., Srinivasan, S., Jain, A., Garg, V., Kale, S., Chitikineni, A., Tripathi, S., Jones, E., Robbins, K. R., Crossa, J., & Varshney, R. K. (2018). Genomic-enabled prediction models using multi-environment trials to estimate the effect of genotype × environment interaction on prediction accuracy in chickpea. *Scientific Reports*, *8*(1), 1–11. https://doi.org/10.1038/s41598-018-30027-2

van Rossum B (2022). *statgenGxE: Genotype by Environment (GxE) Analysis.*

Sasaki, E., Zhang, P., Atwell, S., Meng, D., & Nordborg, M. (2015). "Missing" G x E Variation Controls Flowering Time in Arabidopsis thaliana. *PLoS Genetics*, *11*(10), 1–18. https://doi.org/10.1371/journal.pgen.1005597

Scherer, R., & Scherer, M. R. (2018). Package 'PropCIs'.

Schillaci, M., Gupta, S., Walker, R., & Roessner, U. (2019). The role of plant growth-promoting bacteria in the growth of cereals under abiotic stresses. *Root biology-growth, physiology, and functions*, *28*, 1-21.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 461-464. http://www.jstor.org/stable/2958889

Scott, M. F., Ladejobi, O., Amer, S., Bentley, A. R., Biernaskie, J., Boden, S. A., Clark, M., Dell'Acqua, M., Dixon, L. E., Filippi, C. V., Fradgley, N., Gardner, K. A., Mackay, I. J.,

O'Sullivan, D., Percival-Alwyn, L., Roorkiwal, M., Singh, R. K., Thudi, M., Varshney, R. K., Mott, R. (2020). Multi-parent populations in crops: a toolbox integrating genomics and genetic mapping with breeding. *Heredity*, *125*(6), 396–416. https://doi.org/10.1038/s41437-020-0336-6

Shahzad, A., Ullah, S., Dar, A. A., Sardar, M. F., Mehmood, T., Tufail, M. A., Shakoor, A., & Haris, M. (2021). Nexus on climate change: agriculture and possible solution to cope future climate change stresses. *Environmental Science and Pollution Research*, *28*(12), 14211–14232. https://doi.org/10.1007/s11356-021-12649-8

Shen, X., Pettersson, M., Rönnegård, L., & Carlborg, Ö. (2012). Inheritance Beyond Plain Heritability: Variance-Controlling Genes in Arabidopsis thaliana. *PLoS Genetics*, *8*(8). https://doi.org/10.1371/journal.pgen.1002839

Shen, X., Alam, M., & Rönnegård, L. (2014). Mixed models through the lens of hglm: applications and grand challenges. *JSM*, 1254-1263.

Smyth, G. K. (1989). Generalized linear models with varying dispersion. J. R. Statist. Soc. B, 51, 47–60. https://doi.org/10.1111/j.2517-6161.1989.tb01747.x

Smyth, G., Dunn, P. K., Corty, R. W., & Smyth, M. G. (2022). Package 'dglm'.

Song, H., Wang, X., Guo, Y., & Ding, X. (2022). G × EBLUP: A novel method for exploring genotype by environment interactions and genomic prediction. *Frontiers in Genetics*, *13*(September). https://doi.org/10.3389/fgene.2022.972557

Spindel, J. E., & McCouch, S. R. (2016). When more is better: how data sharing would accelerate genomic selection of crop plants. *New Phytologist*, *212*(4), 814-826. https://doi.org/10.1111/nph.14174

Struchalin, M. V., Amin, N., Eilers, P. H. C., van Duijn, C. M., & Aulchenko, Y. S. (2012). An R package "VariABEL" for genome-wide searching of potentially interacting loci by testing genotypic variance heterogeneity. *BMC Genetics*, *13*. https://doi.org/10.1186/1471-2156-13-4

Tai, A. P., Martin, M. V., & Heald, C. L. (2014). Threat to future global food security from climate change and ozone air pollution. *Nature Climate Change*, *4*(9), 817-821. http://www.nature.com/doifinder/10.1038/nclimate2317

Tian, F., Bradbury, P. J., Brown, P. J., Hung, H., Sun, Q., Flint-Garcia, S., Rocheford, T. R., McMullen, M. D., Holland, J. B., & Buckler, E. S. (2011). Genome-wide association study

of leaf architecture in the maize nested association mapping population. *Nature Genetics*, *43*(2), 159–162. https://doi.org/10.1038/ng.746

Tibbs Cortes, L., Zhang, Z., & Yu, J. (2021). Status and prospects of genome-wide association studies in plants. *Plant Genome*, *14*(1), 1–17. https://doi.org/10.1002/tpg2.20077

van Eeuwijk, F. A., Bink, M. C., Chenu, K., & Chapman, S. C. (2010). Detection and use of QTL for complex traits in multiple environments. *Current Opinion in Plant Biology*, *13*(2), 193–205. https://doi.org/10.1016/j.pbi.2010.01.001

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

Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, *150*(3811), 563-565. https://doi.org/10.1038/150563a0

Wallace, J. G., Larsson, S. J., & Buckler, E. S. (2014). Entering the second century of maize quantitative genetics. *Heredity*, *112*(1), 30–38. https://doi.org/10.1038/hdy.2013.6

Walsh, B., & Lynch, M. (2018). *Evolution and selection of quantitative traits*. Oxford University Press.

Wang, H., Ye, M., Fu, Y., Dong, A., Zhang, M., Feng, L., Zhu, X., Bo, W., Jiang, L., Griffin, C. H., Liang, D., & Wu, R. (2021). Modeling genome-wide by environment interactions through omnigenic interactome networks. *Cell Reports*, *35*(6), 109114. https://doi.org/10.1016/j.celrep.2021.109114

Waters, D. L., van der Werf, J. H. J., Robinson, H., Hickey, L. T., & Clark, S. A. (2023). Partitioning the forms of genotype-by-environment interaction in the reaction norm analysis of stability. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *136*(5), 99. https://doi.org/10.1007/s00122-023-04319-9

Weller, J. I., Soller, M., & Brody, T. (1988). Linkage analysis of quantitative traits in an interspecific cross of tomato (lycopersicon esculentum x lycopersicon pimpinellifolium) by means of genetic markers. *Genetics*, *118*(2), 329–339. https://doi.org/10.1093/genetics/118.2.329

Westerman, K. E., Majarian, T. D., Giulianini, F., Jang, D. K., Miao, J., Florez, J. C., Chen, H., Chasman, D. I., Udler, M. S., Manning, A. K., & Cole, J. B. (2022). Variance-quantitative trait loci enable systematic discovery of gene-environment interactions for cardiometabolic serum biomarkers. *Nature Communications*, *13*(1), 1–11. https://doi.org/10.1038/s41467-

022-31625-5

Woodhouse, M. R., Cannon, E. K., Portwood, J. L., Harper, L. C., Gardiner, J. M., Schaeffer, M. L., & Andorf, C. M. (2021). A pan-genomic approach to genome databases using maize as a model system. *BMC Plant Biology*, *21*(1), 1–10. https://doi.org/10.1186/s12870-021-03173-5

Woodward, A. W., & Bartel, B. (2018). Biology in bloom: A primer on the arabidopsis thaliana model system. *Genetics*, *208*(4), 1337–1349. https://doi.org/10.1534/genetics.118.300755

Xiao, Y., Tong, H., Yang, X., Xu, S., Pan, Q., Qiao, F., Raihan, M. S., Luo, Y., Liu, H., Zhang, X., Yang, N., Wang, X., Deng, M., Jin, M., Zhao, L., Luo, X., Zhou, Y., Li, X., Liu, J., … Yan, J. (2016). Genome-wide dissection of the maize ear genetic architecture using multiple populations. *New Phytologist*, *210*(3), 1095–1106. https://doi.org/10.1111/nph.13814

Xiong, W., Reynolds, M., & Xu, Y. (2022). Climate change challenges plant breeding. *Current Opinion in Plant Biology*, *70*, 102308. https://doi.org/10.1016/j.pbi.2022.102308

Xu, J., Liu, Y., Liu, J., Cao, M., Wang, J., Lan, H., Xu, Y., Lu, Y., Pan, G., & Rong, T. (2012). The genetic architecture of flowering time and photoperiod sensitivity in maize as revealed by QTL review and meta analysis. *Journal of Integrative Plant Biology*, *54*(6), 358–373. https://doi.org/10.1111/j.1744-7909.2012.01128.x

Yadav, A., Dhole, K., & Sinha, H. (2016). Genetic regulation of phenotypic plasticity and canalisation in yeast growth. *PloS one*, *11*(9), e0162326. https://doi.org/10.1371/journal.pone.0162326

Yang, R. C. (2014). Analysis of linear and non-linear genotype × environment interaction. *Frontiers in Genetics*, *5*(JUL), 1–7. https://doi.org/10.3389/fgene.2014.00227

Yang, Q., & Wang, Y. (2012). Methods for analyzing multivariate phenotypes in genetic association studies. *Journal of Probability and Statistics*, *2012*. https://doi.org/10.1155/2012/652569

Yin, L. (2020). CMplot: circle manhattan plot. *R package version*, *3*(2).

Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., ... & Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature genetics*, *38*(2), 203-208. https://doi.org/10.1038/ng1702

Yu, J., Holland, J. B., McMullen, M. D., & Buckler, E. S. (2008). Genetic design and statistical power of nested association mapping in maize. *Genetics*, *178*(1), 539–551.

https://doi.org/10.1534/genetics.107.074245

Ziervogel, G., & Ericksen, P. J. (2010). Adapting to climate change to sustain food security. *Wiley Interdisciplinary Reviews: Climate Change*, *1*(4), 525-540. https://doi.org/10.1002/wcc.56

Zhang, X., & Qi, Y. (2021). Genetic architecture affecting maize agronomic traits identified by variance heterogeneity association mapping. *Genomics*, *113*(4), 1681-1688. https://doi.org/10.1016/j.ygeno.2021.04.009

Zhu, C., Gore, M., Buckler, E. S., & Yu, J. (2008). Status and Prospects of Association Mapping in Plants. *The Plant Genome*, *1*(1), 5–20. https://doi.org/10.3835/plantgenome2008.02.0089