Received: 9 December 2022 | Accepted: 15 June 2023 | Published online: 11 July 2023

DOI: 10.1002/csc2.21051

Crop Science

ORIGINAL ARTICLE

Crop Breeding & Genetics

An application of vGWAS to differences in flowering time in maize across mega-environments

Matthew D. Murphy | **Alexander E. Lipka**

Department of Crop Sciences, University of Illinois Urbana-Champaign, Urbana, Illinois, USA

Correspondence
Alexander E. Lipka, Department of Crop Sciences, University of Illinois Urbana-Champaign, Turner Hall, 1102 S Goodwin Ave, Urbana, IL 61801, USA.
Email: alipka@illinois.edu

Assigned to Associate Editor Paulino Pérez-Rodríguez.

Funding information
NSF, Grant/Award Numbers: 1355406, 1733606

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 gene-by-environment (G × E) loci is vital, identifying significantly associated loci is challenging due to the multiple testing correction. Consequently, G × E loci of small- to moderate effect sizes may never be identified via traditional genome-wide association studies (GWAS). Variance GWAS (vGWAS) have been previously shown to identify G × E 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 G × E effects. We used publicly available genotypic and phenotypic data in maize (*Zea mays* L.) to test the ability of two vGWAS approaches to identify G × E loci controlling two flowering traits. We observed high inflation of $-\log_{10}(p-\text{values})$ from both approaches. This suggests that these two vGWAS approaches are not suitable to the task of identifying G × E loci. We advocate that similar future applications of vGWAS use more sophisticated models that can adequately control the inflation of $-\log_{10}(p-\text{values})$. Otherwise, the application of vGWAS to search for G × E effects that are critical for combating the effects of climate change will not reach its full potential.

1 | INTRODUCTION

Climate change is already undermining all 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 & 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 & Grando, 2020). In addition, climate change will make environments more favorable to biotic stressors such as increased weed, insect, and pathogen pressures (Ceccarelli & Grando, 2020; Shahzad et al., 2021). With these stressors becoming more prevalent due to climate change, the needs to breed both against and for more yield-stable crops are of utmost importance (Langridge et al., 2021). One approach for achieving these breeding objectives is to exploit

Abbreviations: BFT, Brown–Forsythe test; BLUPs, best unbiased linear predictions; DGLM, double generalized linear model; GDD, growing degree days; GWAS, genome-wide association studies; G × E, gene-by-environment; LD, linkage disequilibrium; MLM, mixed linear model; NAM, nested association mapping; PCs, principal components; RILs, recombinant inbred lines; vGWAS, variance genome-wide association studies; vQTL, variance quantitative trait loci.

gene-by-environment (G × E) interactions (Bernardo, 2010). Under this approach, breeders would select against G × E loci to reduce phenotypic variance across environments and years (Langridge et al., 2021; Reckling et al., 2021). Contrastingly, breeders could also select for G × E 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 G × E 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 G × E (Des Marais et al., 2013; Li et al., 2021; van Eeuwijk et al., 2010) underscores the attention that has been given to better understand this critical component of trait variability. Harnessing G × E interactions is vital for plant breeding, and there are several important challenges that need to be addressed. First, the identification of putative G × E 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 G × E 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 G × E interactions increases when not all environmental covariates are accounted for (Westerman et al., 2022). Finally, G × E loci that behave non-additively (e.g., having different phenotypic variances 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 G × E loci underlying simulated (Murphy et al., 2022) and real plant traits (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 (Forsberg et al., 2015; Hussain et al., 2020; Li et al., 2020; Shen et al., 2012), as well as plant architectural and phenology traits (Zhang & 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 G × E in plants.

There are two commonly-used vGWAS statistical approaches whose potential for identifying G × E loci have been previously shown in Murphy et al. (2022). The first approach is to conduct the Brown–Forsythe test (BFT) at each marker (Brown & 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 (Córdova-Palomera et al., 2020; Hong et al., 2017). The other approach is the double generalized linear model (DGLM; Smyth, 1989; Rönnegård & Valdar, 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 G × E loci at the largest possible tested sample size of $N = 2815$.

Core Ideas

- We ran two commonly used vGWAS models to search for G × E interactions for two flowering-time traits in maize.
- We identified SNPs associated with the flowering-time traits.
- We observed severe inflation of false positives for all models and traits.

The purpose of this study was to apply the findings of Murphy et al. (2022) on the potential for vGWAS approaches to identify simulated G × E 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. 2012) to Midwestern and southern mega-environments and then assessed the ability of two different vGWAS approaches to detect G × E 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 G × E making a small but non-zero contribution to total phenotypic variance (Buckler et al., 2009). The latter implies that the effect sizes of G × E 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 G × E loci responsible for differences in flowering time across environments.

2 | MATERIALS AND METHODS

2.1 | 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 to 2009. This diversity panel was grown alongside the US maize nested association mapping panel (NAM; Buckler et al., 2009; McMullen et al., 2009; Yu et al., 2008), and the field design is described in Section 2 of Hung et al. (2012). To study G × E, 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; Chen et al., 2015; Xu et al., 2012). With the expectation of Ponce, PR, all locations had 2 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 (MLM) fitted at each of these two mega-environments using the lme4 R package (Bates et al., 2015). This MLM is written as follows:

$$Y_{ij} = \mu + G_i + E_j + \varepsilon_{ij}, \quad (1)$$

where Y_{ij} represents the observed trait value of the ith genotype grown in the jth environment, μ is the grand mean, G_i represents the random effect of the ith genotype, E_j is the random effect of the jth environment, and ε_{ij} represents the error term for the ith genotype grown in the jth environment. Because the 5702 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; Rice et al., 2020). The resulting BLUPs from each mega-environment are provided in their entirety in Table S1 and visualized in Figure S1.

The process of configuring the BLUPs for GDD to anthesis and GDD to silking from each mega-environment for vGWAS is visualized in Figure 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 was fitted to a unified MLM similar to Yu et al. (2006) in TASSEL 5.0 (Bradbury et al., 2007). The Bayesian information criterion (Schwarz, 1978) 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 fitted 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 G × E traits. All of these steps, along with each vGWAS, are outlined Figure 1, which was created using the "DiagrammeR" R package (Iannone & Iannone, 2022).

2.2 | Genotypic data

We used fully sequenced genotypic data from the Goodman-Buckler diversity panel for the G × E traits (Bukowski et al., 2018; Flint-Garcia et al., 2005). 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.

2.3 | Competing vGWAS models

We used two common vGWAS approaches to search for G × E, namely, the BFT and the DGLM. Both of these approaches are described in detail in Murphy et al. (2022). Essentially, both approaches test for unequal population variances of the differences between trait values across mega-environments (after they were filtered through the pipeline previously described in the "Phenotypic data" section and illustrated in Figure 1) at each tested marker. In general, the BFT tests for equality of population variances across different groups by running a standard analysis of variance on a median-derived transformation of the response variable (Brown & 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 (Murphy et al., 2022; Shen et al., 2012). The CAR R package was used to fit the BFT using the Levene.Test function with the central argument set to "median" (Fox & Weisberg, 2011).

The DGLM is a 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),

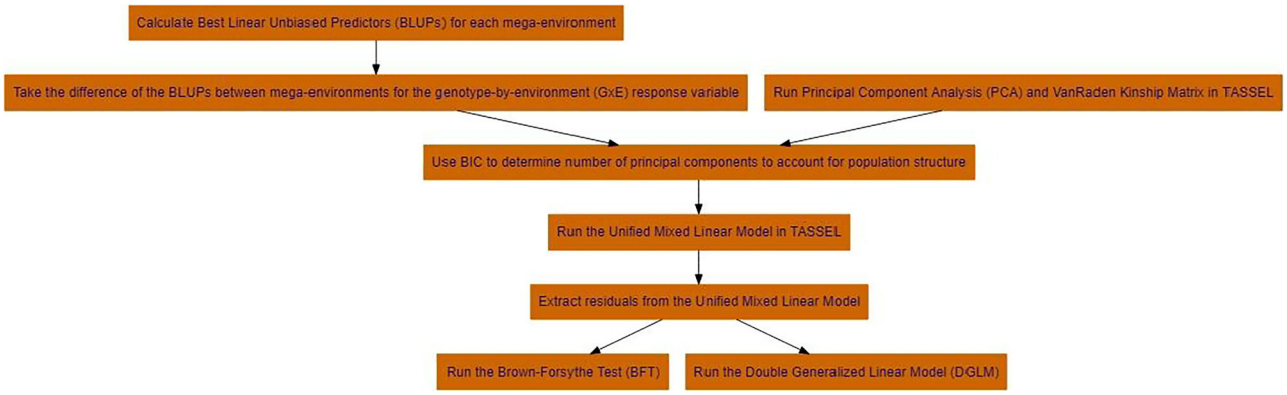

FIGURE 1 Flowchart summarizing the steps taken to process the difference in best linear unbiased predictions (BLUPs) for growing degree dates to anthesis and silking calculated between the Midwestern and southern mega-environments, factor out sources of genetic variability attributed to population structure and familial relatedness, and then run a variance genome-wide association study using the Brown–Forsythe test and double generalized linear model. 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.

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 et al., 2009).

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 QQ-plots and Manhattan plots for all the tests (Yin, 2020). We used the genomic control (Devlin & Roeder, 1999) to quantify the degree of inflation of the test statistics from both vGWAS approaches.

2.4 | 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 $-\log_{10}(p-\text{values})$ 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 $-\log_{10}(p-\text{values})$ with respect to base pairs in the B73 reference genome (version 4) using an R script previously used in Lipka et al. (2013).

2.5 | Mixed model fitted across mega-environments

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

$$Y_{ijk} = \mu + ME_i + Trial_{k(i)} + G_j + (ME \times G)_{ij} + \varepsilon_{ijk}, \quad (2)$$

where Y_{ijk} represents the observed trait value of the jth genotype grown in the kth trial (i.e., particular year at a particular location) nested within the ith mega-environment; μ represents the grand mean; ME_i represents the random effect of the ith mega-environment; $Trial_{k(i)}$ is the random effect of the kth trial nested within the ith mega-environment; G_j is the random effect of the jth genotype; $(ME \times G)_{ij}$ is the random two-way interaction effect between the ith mega-environment and the jth genotype; and ε_{ijk} represents the error term for the jth genotype grown in the kth trial (i.e., particular year at a particular location) nested within the ith mega-environment. After fitting this model in the "statgenGxE" R package (van Rossum et al., 2021), we used the variance component estimates to assess the contribution of $(ME \times G)_{ij}$ variance component (which quantifies G × E) to the overall phenotypic variability of both traits.

3 | RESULTS

3.1 | Midwestern and southern mega-environments capture G × E 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 G × E term (i.e., the variance component estimate of the $(ME \times G)_{ij}$ random effect) accounted for approximately 20% of the total variability of both traits (Tables S2 and S3). This

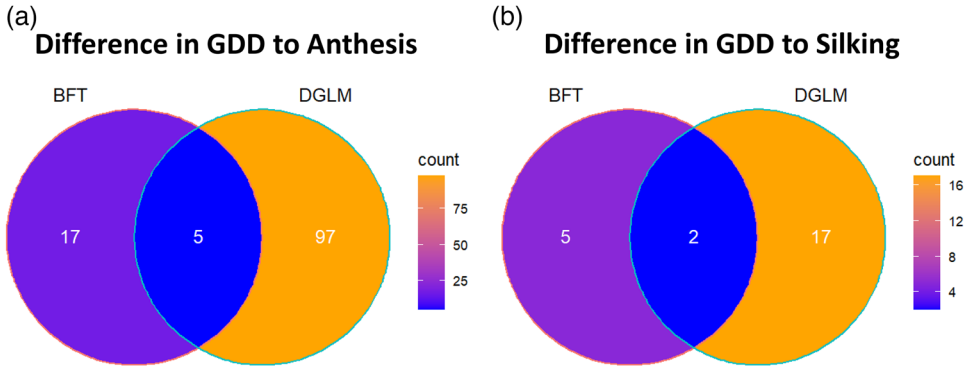

FIGURE 2 Venn diagrams showing the number of statistically significant SNPs for each trait identified by two variance genome-wide association studies (vGWAS) approaches, the Brown–Forsythe test (BFT) as the left circle, and the double generalized linear model (DGLM) as the right circle. The overlap between 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.

suggests that the subdivision of the publicly available field trials into Midwestern and southern mega-environments captures a sufficient amount of G × E for both of the studied traits to justify conducting the ensuing vGWAS.

3.2 | 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 $\alpha = 0.05$, each of the two tested vGWAS approaches identified markers that were statistically significantly associated with both G × E 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 were significantly associated with the difference in GDD to anthesis than the difference in GDD to silking. 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 is summarized in Figure 2 and presented in detail in Table S4.

3.3 | All two vGWAS approaches did not sufficiently control for false positives

For each of the tested traits, we noted that both vGWAS approaches yielded $-\log_{10}(p-\text{values})$ that were highly inflated relative to what would be expected under the corresponding null hypotheses tested at each marker (Figure 3). The greatest amount of such inflation was observed for the DGLM. Nevertheless, the observed increases in $-\log_{10}(p-\text{values})$ for all the analyses performed suggest that these two vGWAS approaches inadequately control for false positive associations for these data.

3.4 | 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 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.

3.5 | Chromosome 9 region containing plausible candidate genes consistently identified across both vGWAS approaches for both traits

To illustrate the potential of these vGWAS approaches to highlight putative candidate genes, we present detailed results for GDD to anthesis using the DGLM in the genomic region on Chromosome 9 in Figure 5. We also created similar figures for the other vGWAS scans, which can be found in Figures S2–S4. Across all vGWAS approaches and

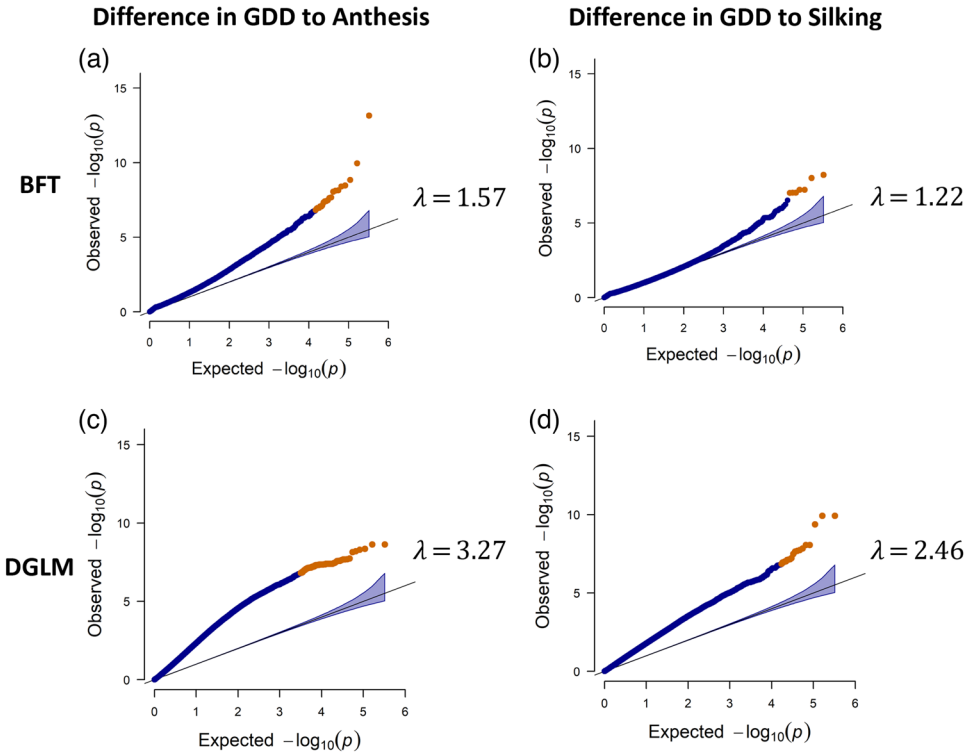

FIGURE 3 Quantile–quantile (QQ)-plots of all the tested variance genome-wide association studies (vGWAS) approaches (rows) tested in all traits (columns). For each plot, the observed $-\log_{10}(p-\text{values})$ from testing each marker is presented on the *Y*-axis, whereas the expected $-\log_{10}(p-\text{values})$ 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.

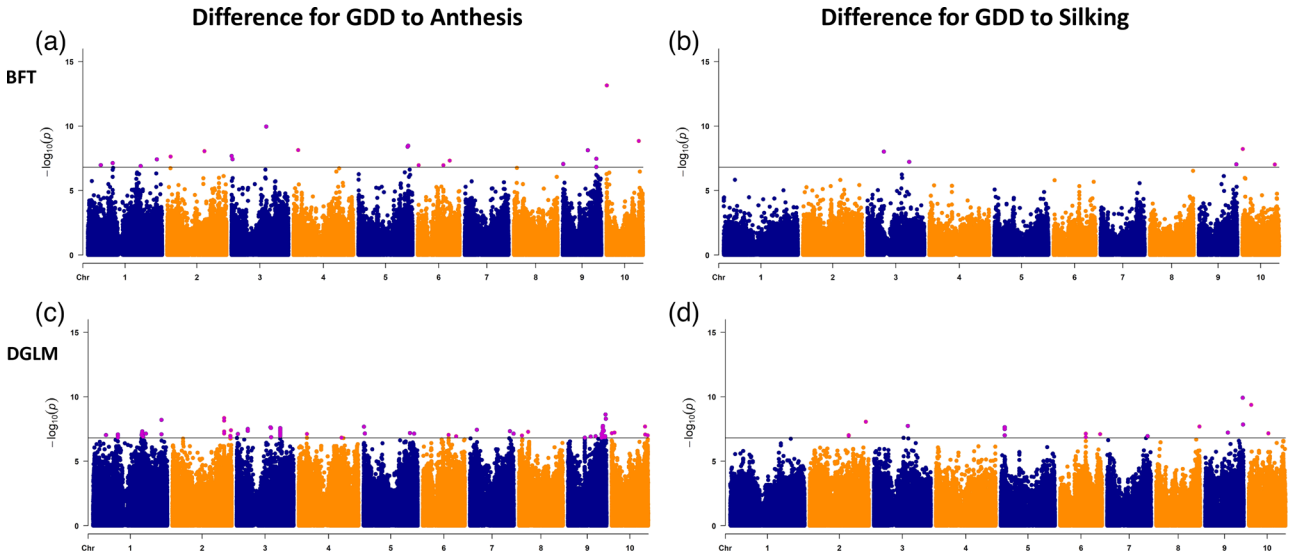

FIGURE 4 Summary of two variance genome-wide association studies (vGWAS) results. (a–d) Manhattan plots of the association results from the two vGWAS approaches across the ten maize chromosomes. The *Y*-axis represents the $-\log_{10}(p-\text{values})$ plotted with respect to B73 RefGen_v4 genome position (*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.

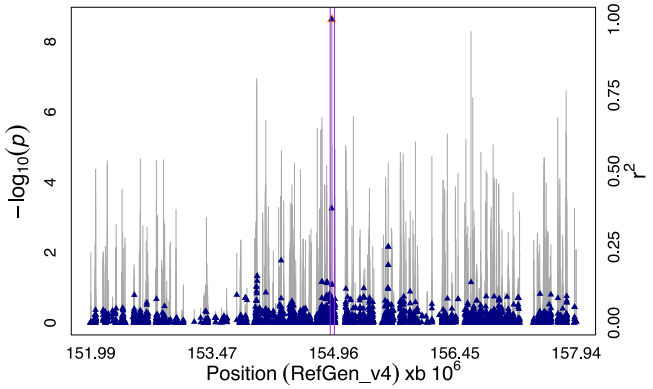

FIGURE 5 Variance genome-wide association studies (vGWAS) results for the difference in growing degree days (GDD) to anthesis. Plot of association results from the double generalized linear model (DGLM) and linkage disequilibrium (r^2) estimates in the genomic region between 152 and 158 Mb on Chromosome 9. The left *Y*-axis presents $-\log_{10}(p-\text{values})$ from the vGWAS, whereas 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 $-\log_{10}(p-\text{values})$ 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).

traits, the strongest peak associations (S9_151791144 and S9_151791148; 154,942,763 and 154,942,767 bp, respectively, *p*-value 1.19×10^{-10}) were 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.

4 | 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 G × E effects to simulated traits. Therefore, we evaluated the potential for two different vGWAS approaches to identify genomic regions likely to contain loci with G × E 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 $-\log_{10}(p-\text{values})$. Thus, this analysis did not provide any further insights into the contribution of G × E 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.

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

Identifying specific loci with G × E effects is vital because they can be selected for (or against) to achieve specific breeding goals. Selecting for such G × E 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 G × E 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 G × E loci once the issues of inflated $-\log_{10}(p-\text{values})$ are addressed.

4.2 | Unable to make inferences on contribution of G × E loci to genetic architecture of flowering time in maize due to inflated high false positives

The genetic architecture of flowering time in maize is consistent with the Fisher–Orr model (Fisher, 1919; Orr, 1998) in that it consists of many small-effect genes, as well as a small number of large-effect genes (Bouchet et al., 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., Bouchet et al., 2013; Romay 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 G × E loci contribute 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 G × E loci (by reducing the severity of the multiple testing correction) could highlight which particular loci contribute to the small G × E 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 G × E loci for flowering time in maize. This could lead to similar, complementary advances that were already made for understanding the role of G × E for flowering time in *Arabidopsis* (Sasaki et al., 2015), sorghum (Li et al., 2018), and rice (Guo et al., 2020).

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

The most important finding from this study was the severe inflation of $-\log_{10}(p-\text{values})$ from both vGWAS approaches, especially those from the DGLM. This degree of inflation was surprising because both G × E flowering time traits were first fitted to a unified MLM that accounted for population structure and familial relatedness (please see Section 2 for details). The residuals from this fitted model were used in our subsequent vGWAS evaluations. Similar approaches have been implemented in previous studies (Forsberg et al., 2015; Li et al., 2020; Shen et al., 2012), 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 and the double hierarchical generalized linear model (Lee & Nelder, 1996, 2006; Rönnegård & Valdar, 2012) needs to be reduced so that researchers can feasibly use them in a vGWAS.

4.4 | Follow-up studies need to address the inflation of $-\log_{10}(p-\text{values})$ before investigating the putative G × E genomic regions identified in our study

Even though the tested vGWAS statistical approaches were prone to severe inflation of $-\log_{10}(p-\text{values})$, 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 G × E genomic regions. However, follow-up studies using more sophisticated vGWAS models 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 (Korte & Farlow, 2013; Nordborg & Weigel, 2008).

5 | CONCLUSION

Although our previous work, in Murphy et al. (2022), highlighted the potential of the BFT and DGLM to identify G × E loci in simulated traits, the analysis conducted here clearly shows these approaches yielded highly inflated $-\log_{10}(p-\text{values})$ when applied to searching for G × E 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 G × E loci real trait data. 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 G × E loci for agronomically important traits like flowering time in maize will not be realized.

AUTHOR CONTRIBUTIONS

Matthew D. Murphy: Conceptualization; formal analysis; methodology; software; visualization; writing-original draft; writing-review and editing. **Alexander E. Lipka**: Conceptualization; formal analysis; investigation; methodology; project administration; supervision; visualization; writing-original draft; writing-review and editing.

ACKNOWLEDGMENTS

The research presented in this manuscript is funded by National Science Foundation project accession numbers 1355406 (A.E.L and M.D.M) and 1733606 (A.E.L. and M.D.M.), as well as University of Illinois Urbana-Champaign Department of Crop Science's Illinois Corn Marketing Growers Fellowship (M.D.M). Additionally, we would like to thank Sarah J. Widener for critically evaluating this manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

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.

ORCID

Alexander E. Lipka https://orcid.org/0000-0003-1571-8528

REFERENCES

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), 1–48. https://doi.org/10.18637/jss.v067.i01

Bernardo, R. (2010). *Breeding for quantitative traits in plants* (Vol. 1, pp. 369). Stemma Press.

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

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), e71377. 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

Brown, M. B., & Forsythe, A. B. (1974). American society for quality the small sample behavior of some statistics which test the equality of several means linked references are available on JSTOR for this article : 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

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.1174276

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

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.

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, 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*, 1–5. https://doi.org/10.3389/fpls.2016.01509

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

Cheng, M., McCarl, B., & Fei, C. (2022). Climate change and livestock production: A literature review. *Atmosphere*, *13*(1), 140. https://doi.org/10.3390/atmos13010140

Córdova-Palomera, A., van der Meer, D., Kaufmann, T., Bettella, F., Wang, Y., Alnæs, D., Doan, N. T., Agartz, I., Bertolino, A., Buitelaar, J. K., Coynel, D., Djurovic, S., Dørum, E. S., Espeseth, T., Fazio, L., Franke, B., Frei, O., Håberg, A., Le Hellard, S., … Westlye, L. T. (2020). Genetic control of variability in subcortical and intracranial volumes. *Molecular Psychiatry*, *26*, 3876–3883. https://doi.org/10.1038/s41380-020-0664-1

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

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

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

Dunn, A. P. K., Smyth, G. K., & Smyth, M. G. (2009). Package 'dglm.' 1–14.

Fisher, R. A. (1919). The correlation between relatives on the supposition of Mendelian inheritance. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, *52*(2), 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. *PLoS Genetics*, *11*(11), e1005648. https://doi.org/10.1371/journal.pgen.1005648

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (2nd ed.). Sage. http://z.umn.edu/carbook

Gauderman, W. J., Mukherjee, B., Aschard, H., Hsu, L., Lewinger, J. P., Patel, C. J., Witte, J. S., Amos, C., Tai, C. G., Conti, D., Torgerson, D. G., Lee, S., & 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

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

Hong, C., Ning, Y., Wei, P., Cao, Y., & Chen, Y. (2017). A semi-parametric 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. (2012). 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

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'*.

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

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, Y., & 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

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, 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, X., Guo, T., Wang, J., Bekele, W. A., Sukumaran, S., Vanous, A. E., McNellie, J. P., Cortes, L. T., Lopes, M. S., Lamkey, K. R., Westgate, M. E., McKay, J. K., Archontoulis, S. V., Reynolds, M. P., Tinker, N. A., Schnable, P. S., & Yu, J. (2021). An integrated framework reinstating the environmental dimension for GWAS and genomic selection in crops. *Molecular Plant*, *14*(6), 874–887. https://doi.org/10.1016/j.molp.2021.03.010

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

Lipka, A. E., Gore, M. A., Magallanes-Lundback, M., Mesberg, A., Lin, H., Tiede, T., Chen, C., Buell, C. R., Buckler, E. S., Rocheford, 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., 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

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

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

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*, *129*(2), 93–102. https://doi.org/10.1038/s41437-022-00541-1

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

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

Peiffer, J. A., Romay, M. C., Gore, M. A., Flint-Garcia, S. A., Zhang, Z., Millard, M. J., Gardner, C. A. C., McMullen, M. D., Holland, J. B., Bradbury, P. J., & Buckler, E. S. (2013). The genetic architecture of maize height. *Genetics*, *196*(4), 1337–1356. https://doi.org/10.1534/genetics.113.159152

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), 27. https://doi.org/10.1007/s13593-021-00681-4

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., Elshire, R. J., Acharya, C. B., Mitchell, S. E., Flint-Garcia, S. A., McMullen, M. D., Holland, J. B., Buckler, E. S., & Gardner, C. A. (2013). Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biology*, *14*(6), R55. https://doi.org/10.1186/gb-2013-14-6-r55

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

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

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464.

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), e1002839. https://doi.org/10.1371/journal.pgen.1002839

Smyth, G. K. (1989). Generalized linear models with varying dispersion. *Journal of the RoyalStatistical Society: Series B (Methodological)*, *51*(1), 47–60.

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), 972557. https://doi.org/10.3389/fgene.2022.972557

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

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

van Rossum, B.-J., van Eeuwijk, F., Boer, M., & Malosetti, M. (2021). Package 'statgenGxE.' CRAN. https://cran.r-project.org/web/packages/statgenGxE/index.html

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

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

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

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

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

Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., & 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

Zeng, L., Liu, X., Zhou, Z., Li, D., Zhao, X., Zhu, L., Luo, Y., & Hu, S. (2018). Identification of a G2-like transcription factor, OsPHL3, functions as a negative regulator of flowering in rice by co-expression and reverse genetic analysis. *BMC Plant Biology*, *18*(1), 1–12. https://doi.org/10.1186/s12870-018-1382-6

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

SUPPORTING INFORMATION

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

How to cite this article: Murphy, M. D., & Lipka, A. E. (2023). An application of vGWAS to differences in flowering time in maize across mega-environments. *Crop Science*, *63,* 2807–2817. https://doi.org/10.1002/csc2.21051