CellPress
Partner Journal

Molecular Plant
Resource Article

An integrated framework reinstating the environmental dimension for GWAS and genomic selection in crops

Xianran Li[1,5], Tingting Guo[1,5], Jinyu Wang[1], Wubishet A. Bekele[2], Sivakumar Sukumaran[3], Adam E. Vanous[1], James P. McNellie[1], Laura Tibbs Cortes[1], Marta S. Lopes[3], Kendall R. Lamkey[1], Mark E. Westgate[1], John K. McKay[4], Sotirios V. Archontoulis[1], Matthew P. Reynolds[3], Nicholas A. Tinker[2], Patrick S. Schnable[1] and Jianming Yu[1,*]

[1]Department of Agronomy, Iowa State University, Ames, IA 50011, USA

[2]Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Ottawa, ON, Canada

[3]International Maize and Wheat Improvement Center (CIMMYT), Mexico City, Mexico

[4]Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO 80523, USA

[5]These authors contributed equally to this article.

***Correspondence: Jianming Yu (jmyu@iastate.edu)**

https://doi.org/10.1016/j.molp.2021.03.010

ABSTRACT

Identifying mechanisms and pathways involved in gene–environment interplay and phenotypic plasticity is a long-standing challenge. It is highly desirable to establish an integrated framework with an environmental dimension for complex trait dissection and prediction. A critical step is to identify an environmental index that is both biologically relevant and estimable for new environments. With extensive field-observed complex traits, environmental profiles, and genome-wide single nucleotide polymorphisms for three major crops (maize, wheat, and oat), we demonstrated that identifying such an environmental index (i.e., a combination of environmental parameter and growth window) enables genome-wide association studies and genomic selection of complex traits to be conducted with an explicit environmental dimension. Interestingly, genes identified for two reaction-norm parameters (i.e., intercept and slope) derived from flowering time values along the environmental index were less colocalized for a diverse maize panel than for wheat and oat breeding panels, agreeing with the different diversity levels and genetic constitutions of the panels. In addition, we showcased the usefulness of this framework for systematically forecasting the performance of diverse germplasm panels in new environments. This general framework and the companion CERIS-JGRA analytical package should facilitate biologically informed dissection of complex traits, enhanced performance prediction in breeding for future climates, and coordinated efforts to enrich our understanding of mechanisms underlying phenotypic variation.

Key words: phenotypic plasticity, genotype by environment interaction, gene–environment interplay, genome-wide association studies, genomic selection, flowering time, reaction norm

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., and Yu J. (2021). An integrated framework reinstating the environmental dimension for GWAS and genomic selection in crops. Mol. Plant. **14**, 874–887.

INTRODUCTION

Climate change and climate variability present a mounting challenge to sustainable food production (Wheeler and von Braun, 2013). Developing climate-resilient crops can be greatly enhanced through a comprehensive understanding of the global crop production system and by developing an actionable blueprint to build such an understanding (Lobell et al., 2008; McCouch et al., 2013). Most traits relevant to adaptation and productivity are affected by the complex interplay between

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

genes and the environment, but in-depth functional analysis of genetic pathways and physiological effects under varied genetic backgrounds and different environmental conditions is not always feasible. Fortunately, advances in genomic technologies have enabled many innovative gene discovery approaches in crops and promoted a concerted effort to tackle long-standing challenges in biology and agriculture (Hickey et al., 2019).

How to explain and predict phenotypes is a prominent question in biology and evolution (Mackay et al., 2009; Boyle et al., 2017; Exposito-Alonso et al., 2019). Answering this question requires the holistic examination of genomes, environments, and their interaction throughout the spatial and temporal dimensions of an organism's life cycle. Gene–environment-wide association studies were proposed to incorporate environmental information into genomic and pathway databases to inform analyses of gene–environment interaction (Thomas, 2010). Recent progress in a few key research areas has set the stage for tackling this challenge through an integrated approach. First is the comprehensive genome profiling of a large number of diverse genotypes in major crops and coordinated phenotyping across varied environments (Bevan et al., 2017). Second is the better appreciation and modeling of the intertwined relationship between genotype and environment (Wilczek et al., 2009; Lobell et al., 2014; Exposito-Alonso et al., 2019; Peng et al., 2020). Quantitative measurements of environmental factors in natural field conditions are being incorporated into performance modeling and forecasting. Finally, research into phenotypic plasticity and genotype by environment interaction (G × E) has received revived attention (Des Marais et al., 2013; Malosetti et al., 2013; Fan et al., 2016; Gage et al., 2017; Kusmec et al., 2017).

Genome-wide association studies (GWAS) have been conducted to understand the genetic architecture of complex traits (Lipka et al., 2015; Tibbs Cortes et al., 2021). In addition to the single trait mixed model approach (Yu et al., 2006; Kang et al., 2010; Zhang et al., 2010), the GWAS method has been advanced to identify pleiotropic loci (Grotzinger et al., 2018), gene–environment interactions (Moore et al., 2019), and gene expression effects across varying conditions and tissues (Urbut et al., 2019). Genomic selection (GS) has been conducted to leverage predictions of untested individuals based on a genotype-phenotype model established with training individuals (Crossa et al., 2017; Xu et al., 2020). However, it is still a common practice to first summarize phenotypic data collected over multiple environments before investigating association with genotyping data to either identify genomic regions associated with the aggregated trait values or generate prediction models. Different approaches have been proposed to incorporate environmental information in GS so that performance prediction can be context-specific and achieved with better accuracy (Heslot et al., 2014; Jarquin et al., 2014; Cooper et al., 2016; Millet et al., 2019; Costa-Neto et al., 2021). Nevertheless, some of these approaches require extensive environmental data, phenotyping data of component traits, or crop modeling, which may not be readily available or conducted. Connections between genetic effects and major environmental factors are still challenging to establish.

Here, we demonstrate that, by taking a simple but critical step, we can establish a novel, integrated framework to add the environmental dimension to both GWAS and GS of complex traits. Although the concept of a performance-free environmental index was suggested decades ago (Finlay and Wilkinson, 1963; Eberhart and Russell, 1966), limited research has been reported about taking this step until recently (Li et al., 2018; Millet et al., 2019; Guo et al., 2020), and the environmental dimension has not been explicitly addressed in typical GWAS with diverse populations. It is critical to demonstrate that complex traits measured for diverse genetic materials across environments can be dissected through GWAS with an explicit environmental dimension. Identifying the major environmental pattern also facilitates the prediction and forecasting of complex traits through GS. Moreover, focusing on extracting the main environmental factors to differentiate field environments would help translate findings from targeted studies under controlled environmental conditions (Blackman, 2017; Scheres and van der Putten, 2017) and with well-equipped field sensors (Millet et al., 2019) to reveal the major patterns across environments. With data for three major crops, we showed that the changing effects of genes along an environmental index can be revealed for quantitative traits, and that patterns uncovered from apparently complex phenotype dynamics can be exploited for genomics-enabled, cross-environment performance prediction. Profiling gene effect continua along an environmental gradient may serve as an umbrella framework for dissecting and validating genomic determinants underlying complex trait variation and their interactions with the environment.

RESULTS

Phenotypic plasticity exhibited by crops

Diverse inbred lines from three major crops were grown at multiple year–location combinations (Figure 1). The maize population had 282 inbred lines (Flint-Garcia et al., 2005), the wheat population had 288 inbred lines (Sukumaran et al., 2017), and the oat population had 433 inbred lines (Esvelt Klos et al., 2016). While all three populations were assembled to cover diverse genetic backgrounds, the wheat and oat populations comprised elite breeding materials. In total, there were 51 year–location combinations (i.e., environments) (Supplemental Table 1) covering a range of latitudes and longitudes that are representative of important production areas for each crop (Figure 1). A total of 35 212 phenotypic values for three traits (flowering time, plant height, and grain yield) were recorded for these inbred lines (Supplemental Figure 1). Given that the geographical regions involved were much wider than those in earlier studies, environmental effect accounted for a large proportion of the phenotypic variance, followed by genotype effect in most cases (Figure 1). These phenotypes appeared to represent complex quantitative patterns that were not easily explained by latitude or time individually (Supplemental Figure 1). However, after ordering these environments based on the trait means (i.e., environmental means) and fitting a regression on environmental mean for each inbred line, a clear pattern emerged in the reaction norms of these genotypes (Supplemental Figure 2). This widely adopted process is termed joint regression analysis (Finlay and Wilkinson, 1963; Eberhart and Russell, 1966). While we sought to expand this concept to

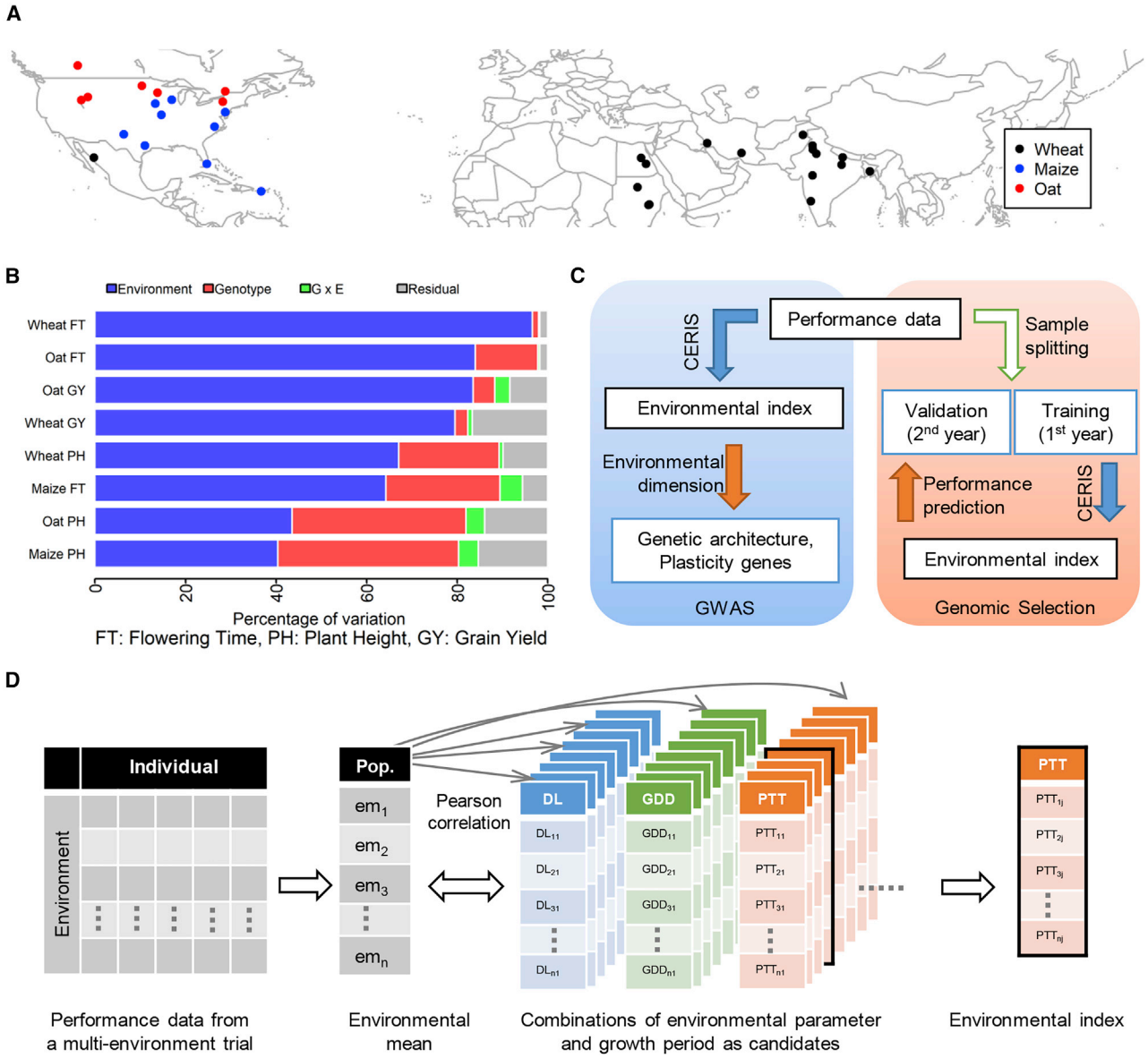

Figure 1. Phenotypic variation of diverse crop genotypes across varied environments.

(A) Geographic distribution of experimental locations for maize, wheat, and oat.

(B) Variances due to environment, genotype, G × E, and residual error for eight crop–trait combinations.

(C) Schematic design of the whole study, which aimed to explicitly incorporate environmental index in genome-wide association study (GWAS) and genomic selection.

(D) The Critical Environmental Regressor through Informed Search (CERIS) procedure. Environmental index is identified as the combination of environmental parameter and growth period that has the strongest correlation with environmental mean (em) of the population (pop.) and a reasonable physiological interpretation. DL, day length; GDD, growing degree days; PTT, photothermal time.

quantify the contribution of individual genes, we also recognized a fundamental shortcoming in this analysis: that the mean performance in a new environment is not known in advance. Relying only on environmental means, we could not quantify growing conditions with specific environmental parameters.

Identifying the environmental index: A critical, enabling step

Temperature and photoperiod have been widely regarded as the two most critical environmental factors affecting phenotypes in many crop species and used in crop development modeling (Robertson, 1968; Angus et al., 1981; Hammer et al., 1982; Masle et al., 1989; Brachi et al., 2010; Des Marais et al., 2013; Blackman, 2017; Scheres and van der Putten, 2017). They varied substantially among the examined environments (Supplemental Figure 3). We recognized that this concept may provide a solution to addrss the shortcoming of the established joint regression analysis. Replacing the environmental means with a quantitative environmental index is highly desirable because once such an index is obtained, we cannot only model the observed phenotype with this environmental index, but also

predict phenotypic performance in new environments by using historical weather averages, in-season weather, or forecasted weather data (Figure 1).

Building on the findings from a single trait (flowering time) in narrow genetic backgrounds (biparental mapping populations) (Li et al., 2018; Guo et al., 2020), we expanded the search for an environmental index to a list of quantitative traits obtained for diverse genetic materials to identify environmental indices with sufficient explanatory power to approximate the observed environmental means for these traits (Supplemental Table 2). We termed this process Critical Environmental Regressor through Informed Search (CERIS). If the whole population is viewed as a single mega-genotype, its performance in different environments (i.e., environmental means) indicates the difference in crop growth and development among environments. Replacing environmental means would then allow us to quantitatively connect all tested environments with a common index, which subsequently bridges the tested environments with new environments. We focused on three complex traits: flowering time, plant height, and grain yield. In brief, for a trait measured in multi-environment trials, CERIS searches the sets of average values obtained for an environmental parameter from different periods of time (windows) and calculates the correlation of these values to environmental means. The parameter–window combination that yields the strongest correlation to the observed environmental means is chosen as the environmental index for a particular trait. Four environmental parameters were examined: photoperiod (day length [DL]), temperature (expressed as growing degree days [GDD]), photothermal time (PTT = GDD × DL), and photothermal ratio (PTR = GDD/DL).

Using the CERIS algorithm, we found that environmental indices based on PTT or PTR were consistently more strongly correlated with environmental means than indices from individual environmental factors (Figure 2 and Supplemental Figures 4–6). Biological relevance was inspected to ensure that the identified environmental index agreed with the general understanding of physiological processes and was quantifiable before the stage when the target trait was observed. This avoided the search for irrelevant growth periods and enabled the use of in-season weather data for forecasting. While maize flowering time was best modeled by PTR, flowering time in wheat and oat was best modeled by PTT. For plant height in maize and wheat, PTT was the best choice, while PTR was the best choice for modeling plant height in oat. For grain yield, PTR was the best choice for both wheat and oat. Given the chosen environmental indices for flowering time, we also examined correlations for other environmental parameters within the fixed window. For maize flowering time, PTR from 22 to 37 days-after-planting (DAP) had the strongest correlation with environmental mean ($r = -0.97$, Figure 2). Temperature (in GDD) was negatively correlated with environmental mean, as higher temperature promotes early flowering ($r = -0.95$, Figure 2), but photoperiod (DL) had a positive correlation ($r = 0.87$) with environmental mean, as short DL promotes early flowering (Figure 2). The ratio of GDD to DL, PTR, captured these antagonistic effects and yielded a stronger correlation to flowering time. On the other hand, for wheat and oat, PTT correlated with flowering time better than the primary factors (DL and GDD) with synergistic effects (Supplemental Figure 7).

One requisite for the environmental index is that it should be obtainable from a growth period before the trait establishment. For flowering time, the identified window was at least 1 week earlier than 97% of the 13 088 flowering time observations recorded across the three crops (Supplemental Figure 8). For plant height and grain yield, the identified windows were also generally earlier than the dates when traits were fully developed (Figure 2 and Supplemental Figure 8). For instance, maize plant height was correlated with PTT from 52 to 59 DAP, while the average flowering time across nine environments ranged from 56 to 83 DAP. Therefore, this 52–59-DAP window overlaps with the exponential growth stage.

To further demonstrate that the CERIS algorithm can identify environmental indices, we conducted searches for other agronomic and physiological traits collected for the maize panel. Indeed, replacing environmental means with environmental indices could be achieved for different traits (Supplemental Table 2).

With the identified environmental indices, phenotypic performance can be systematically modeled with two reaction-norm parameters, intercept and slope, obtained by regressing observed phenotypic values of each inbred line on this index (Figure 2). While intercept measures the expected performance of a genotype at the average point (center) of the environmental range, slope measures the degree of plasticity of an individual genotype along the environmental gradient. Deviations of the observed values from the regression-fitted values represent effects due to other local environmental factors and random errors. Across three diverse panels of genotypes, we observed a range of values for intercept and slope (Supplemental Figure 9). Using the environmental index, traditional joint regression analysis was transformed into a framework, where not only the performance of a genotype but also genetic effects at different loci can now be quantified with the environmental gradient. We termed this overall framework CERIS-Joint Genomic Regression Analysis (CERIS-JGRA).

Dissecting genetic architecture with an environmental dimension

To reveal gene–environment interplay underlying phenotypic plasticity, we conducted GWAS using reaction-norm parameters (slope and intercept) obtained for all trait-crop combinations (Supplemental Figures 10–12). Here, we focused on flowering time because this trait has been better characterized and previous studies have uncovered a list of confirmed or candidate genes. Because both parameters were derived from regressing the flowering time phenotype on the environmental index, we expected GWAS to identify genomic regions harboring genes responsible for the environmental factors involved in trait expression. Indeed, using 12 million single-nucleotide polymorphisms (SNPs), we detected many genomic regions significantly associated with the slope for flowering time and one for the intercept in maize (Figure 3). The genomic regions with the strongest signals associated with slope variation contained flowering time genes contributing to maize adaptation: *ZmCCT* on chromosome 10 and *Vgt1* on chromosome 8 (Salvi et al., 2007; Hung et al., 2012;Bouchet et al., 2013; Yang et al., 2013; Romero Navarro et al., 2017).

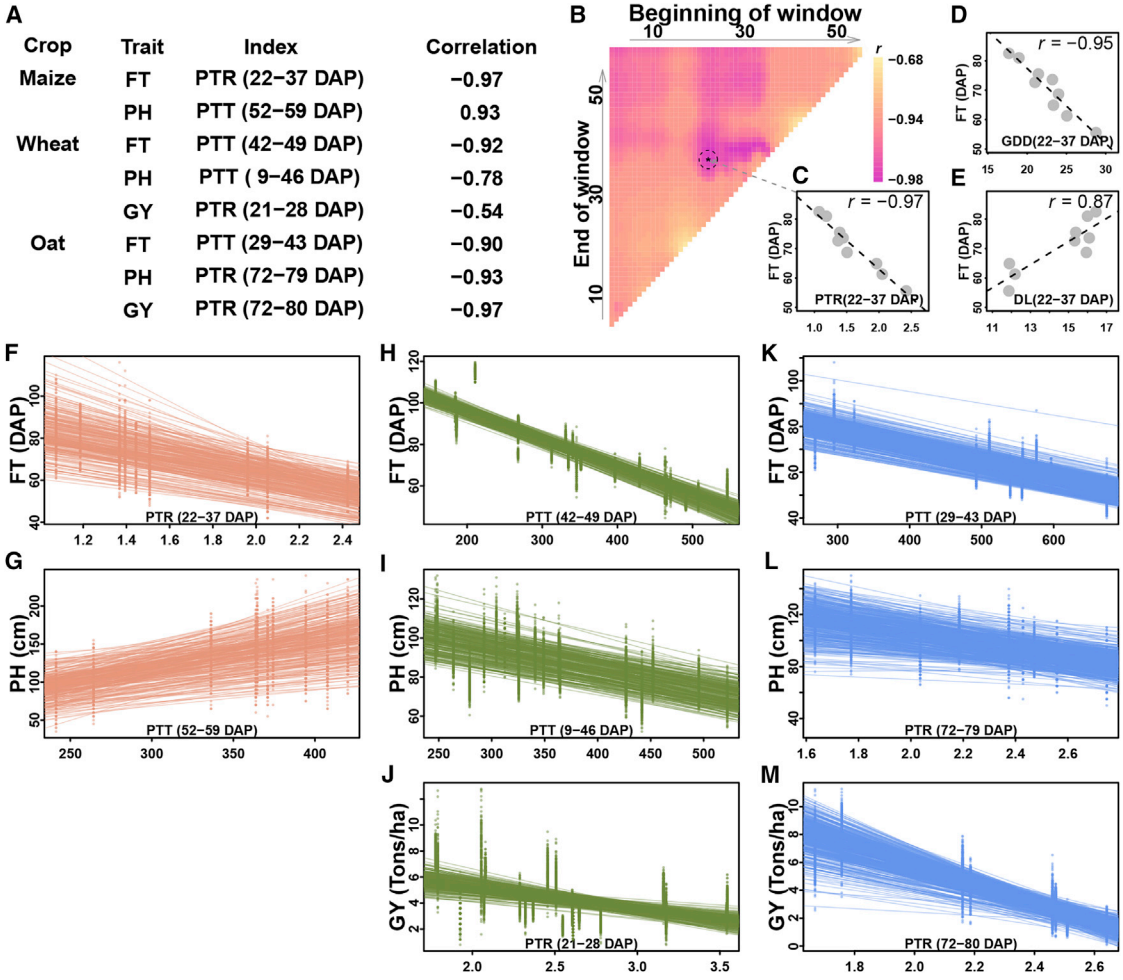

Figure 2. Environmental indices underlying phenotypic plasticity.

(A) Environmental indices identified by CERIS for each trait-crop combination. The index is identified by searching through combinations of environmental parameters and growth periods for the combination with the strongest correlation with the observed environmental means of the population.

(B) Photothermal ratio (PTR) from the 22–37 days-after-planting (DAP) window has the strongest correlation with the environmental means for maize flowering time (FT).

(C–E) PTR captures the antagonistic effects on maize FT by temperature **(D)** and photoperiod **(E)**.

(**F–M**) Performance dynamics modeled by environmental indices. **(F)** Maize FT. **(G)** Maize plant height (PH). **(H)** Wheat FT. **(I)** Wheat PH. **(J)** Wheat grain yield (GY). **(K)** Oat FT. (**L**) Oat PH. (**M**) Oat GY. In **(F)**–**(M)**, each dot denotes the observed phenotypic value for each genotype in an environment, and each line is the regression-fitted values for each genotype.

We further investigated four detected genes (*ZMM22*/*ZmMADS69*, *ZMM5*, *Conz1*, and *ZCN8*) by comparing sequences from eight publicly available whole-genome assemblies (Supplemental Figure 13). The strongest association for intercept variation implicated *ZMM22* on chromosome 3, encoding a MADS-box gene, and its expression level was reported to be associated with flowering time variation (Hirsch et al., 2014; Lin et al., 2017; Liang et al., 2019). Transposon insertions in the introns caused three haplotypes among these eight maize inbreds. *ZMM22* in B73, B104, and CML247 contained an extra 10.5 kb of sequence due to insertions of a Copia retrotransposon (8.9 kb) and a hAT transposon (1.6 kb), while in W22 contained an 8.6 kb Gypsy retrotransposon (XILON1) insertion in the last intron. We observed transposons within the intron of another MADS-box gene (*ZMM5*) on chromosome 9 and in the upstream promoter regions of *Conz1* on chromosome 9. In the promoter of *ZCN8* on chromosome 8, one extra segment with a protein coding sequence resulted in two haplotypes. In previous studies, *ZCN8* was designated as the candidate gene underlying QTL *Vgt2* because it encodes florigen and its gene expression is associated with flowering time (Meng et al., 2011; Bouchet et al., 2013; Guo et al., 2018). *Conz1* is a homolog of *Arabidopsis CO* (*CONSTANS*) and rice *Hd1* (*Heading date 1*) (Miller et al., 2008), both being validated flowering time genes. Our observation of transposon variants in these four genes provides relevant evidence for follow-up studies of functional polymorphisms.

Unlike maize, where many loci with known or candidate genes were detected, only a few major loci were detected for flowering-time phenotypic plasticity in wheat and oat (Supplemental Figures 11 and 12). This may be due in part to the use of

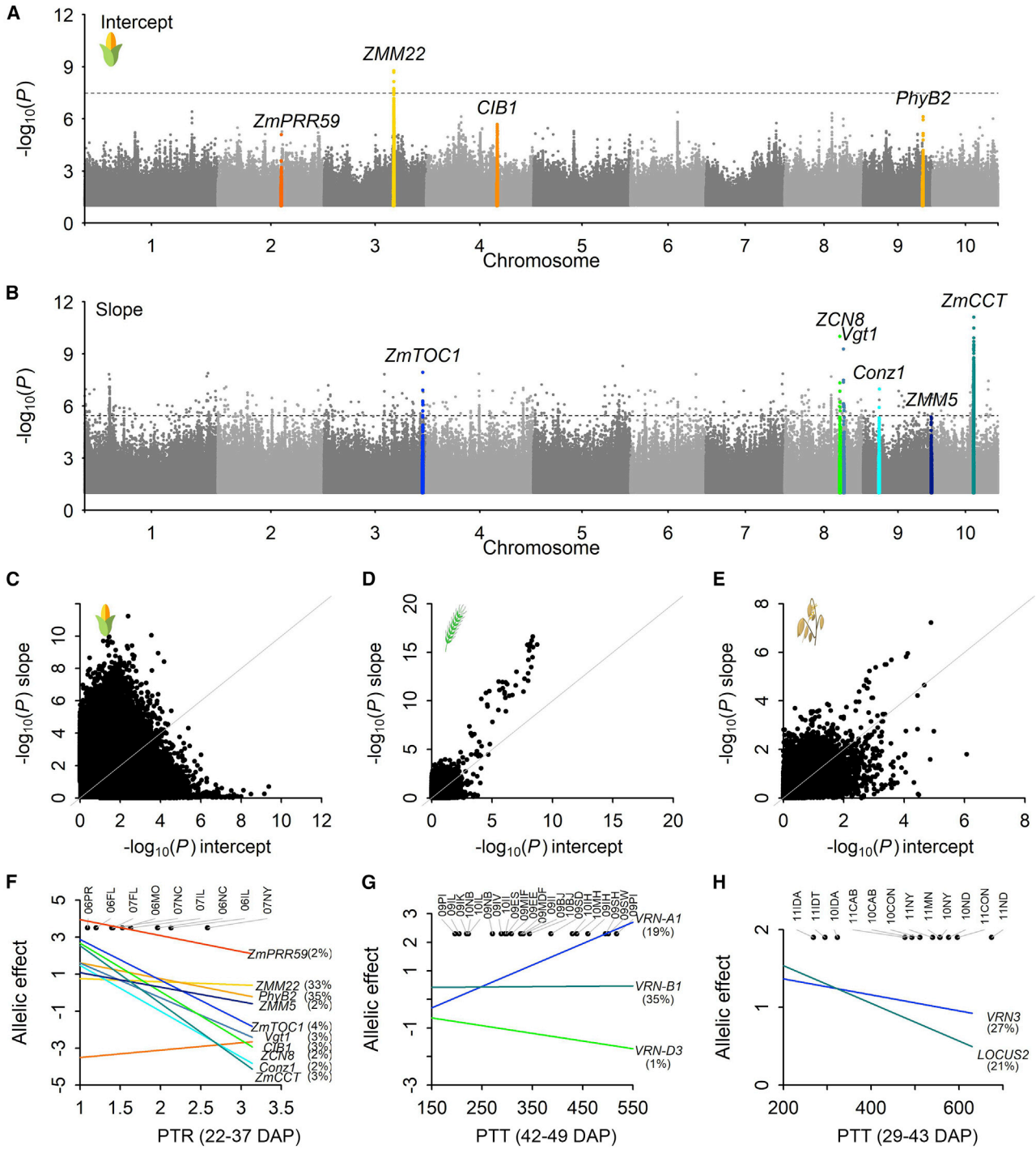

Figure 3. Genetic architecture of FT phenotypic plasticity and gene–environment interplay.

(A and B) Manhattan plot from GWAS of intercept **(A)** and slope **(B)** derived for maize FT phenotypic plasticity.

(C) Scatter plot of $-\log_{10}(P)$ from the GWAS of slope and intercept for maize FT.

(D) Scatter plot of $-\log_{10}(P)$ from the GWAS of slope and intercept for wheat FT.

(E) Scatter plot of $-\log_{10}(P)$ from the GWAS of slope and intercept for oat FT.

(F) Genetic effect continua of phenotypic-plasticity genes for maize FT.

(G) Genetic effect continua of phenotypic-plasticity genes for wheat FT.

(H) Genetic effect continua of phenotypic-plasticity loci for oat FT. The false discovery rate significance threshold of 0.05 is indicated with a dotted line in **(A–B)**. Minor allele frequency is shown in parentheses after the locus name in **(F–H)**.

diversity panels that were biased toward breeding materials adapted to a set of target environments, in contrast to the maize panel, which was considerably more diverse. Moreover, while different sets of genes were associated with intercept and slope in maize, similar sets of genes were found to be associated with intercept and slope of flowering-time phenotypic plasticity in wheat and oat (Figure 3). In wheat, the region with the strongest association with both slope and intercept was located near *VRN-A1*, which encodes a MADS-box transcription factor involved in a flowering time regulatory network (Yan et al., 2003). The recessive *VRN-A1* allele is one of the main targets for breeding of spring wheat. Multiple *VRN-A1* alleles have been discovered with indel polymorphisms in the promoter. Similarly, *VRN3* on oat chromosome 1 (Esvelt Klos et al., 2016) and one unknown locus on chromosome 2 were associated with both intercept and slope. Although known candidate genes were not found in several other genomic regions identified from GWAS of slope and intercept in wheat and oat, these detected regions provide critical information for future targeted studies of individual genes that may become possible as mature reference genomes and molecular tools are developed.

To further understand the gene–environment interplay, we examined genetic effect dynamics along the identified environmental index. We first obtained genetic effects within individual environments for SNPs associated with slope and intercept of flowering time. Then we regressed these genetic effects onto the environmental index to obtain fitted lines describing the genetic effect continua for these loci (Figure 3). Several known genes (*PhyB2*, *CIB1*, and *ZmPRR59* for maize; *VRN-B1* and *VRN-D3* for wheat) were also plotted to show the potential reasons why they did not pass the significance threshold in GWAS (Supplemental Figures 10 and 11): low minor allele frequency, relatively small effect size at the mean value of the environmental index, or relatively small change in effect size across the values of the environmental index. We followed the nomenclature from earlier publications to interpret these findings (Des Marais et al., 2013; Li et al., 2018). Differential sensitivity, i.e., genetic effect varying in degree but not in direction, was the primary mode of gene action across all three crops. Conditional neutrality, i.e., genetic effect detected in some but not in all environments, was found for a single locus (*VRN-A1*) in wheat but for a few other loci in maize, including *ZMM5*. Antagonistic pleiotropy, i.e., genetic effect varying in direction, was also detected for several loci in maize. These findings revealed different patterns of gene–environment interplay underlying the observed phenotypic plasticity. Since we know that genes do not function in isolation and that gene networks are involved, our interpretation is that the marginal effects of individual genes involved in the network can vary according to different environmental conditions.

Genomics-enabled, on-target performance forecasting across environments

We tested the CERIS-JGRA framework for performance forecasting. Each crop population was planted in two consecutive years; we established the prediction model with data from only the first year. CERIS generally identified the same environmental indices from similar periods (Supplemental Table 3 and Figure 2). The specific value of the environmental index for trials conducted in the second year was then used by the model to obtain predictions for all genotypes. In this relatively simple case, no sampling of genotypes (G) was made (thus, the case was termed CERIS-JRA). Predictions made with the CERIS-JRA contained the environmental effect and G × E interaction (Supplemental Figure 14). This approach resulted in higher prediction accuracy across all environments than obtained using the first-year averages computed for all locations as predictions (Supplemental Figure 15). This was also clear when the distributions of the observed-predicted ratios were compared between CERIS-JRA and the approach of using the previous-season averages. Moreover, although prediction accuracy within environments was comparable between the two approaches, which was expected due to the small amount of G × E interaction (Figure 1), having differential performance prediction through CERIS-JRA is more desirable than simply relying on the fixed sets of previous year averages.

Next, we considered a more challenging forecasting scenario where genome-wide (G) SNP information was used and sampling of genotypes (G) was involved, i.e., predicting performance of untested genotypes in new seasons. CERIS was implemented to identify the environmental index using environmental data and the performance data from half of the panel in the first year. Genomic prediction models were established for reaction-norm parameters, which were then fed into JGRA to obtain predicted performance values for the other half of the panel in the second year by factoring in both SNP information and environmental-index values (Figure 4). It was encouraging to see that CERIS-JGRA had generally high accuracy for most trait–species combinations. Notice that, in this case, prediction was only possible with the CERIS-JGRA approach, because of the connection through genomics between the two halves of the panel, i.e., the "G" in "JGRA." In other words, genomics connected the different sets of genetic materials included in the training and validation sets. The relatively modest prediction accuracy for individual trials of some trait–species combinations was partially due to the low number of replications and may be improved with additional environmental indices. We also investigated one less-challenging forecasting scenario, where performance prediction of untested genotypes in untested environments was conducted by leaving one environment and 50% of the genotypes out. CERIS-JGRA generated consistent results (Supplemental Figure 16). In both scenarios (Figure 4 and Supplemental Figure 16), it was not realistic to expect the prediction accuracy to be high for all environments, as location–season-specific environmental factors can alter the growth and development of genotypes. The combination of generally high overall prediction accuracy, mostly modest within-environment prediction accuracy, and sometimes low within-environment prediction accuracy not only demonstrated the capacity of this approach but also revealed the complex biological reality.

DISCUSSION

Many complementary approaches have been developed to incorporate environmental data, physiological insights, and crop models into genetic analysis (Heslot et al., 2014; Jarquin et al., 2014; Cooper et al., 2016). These earlier studies stimulated research aimed at bringing environmental context

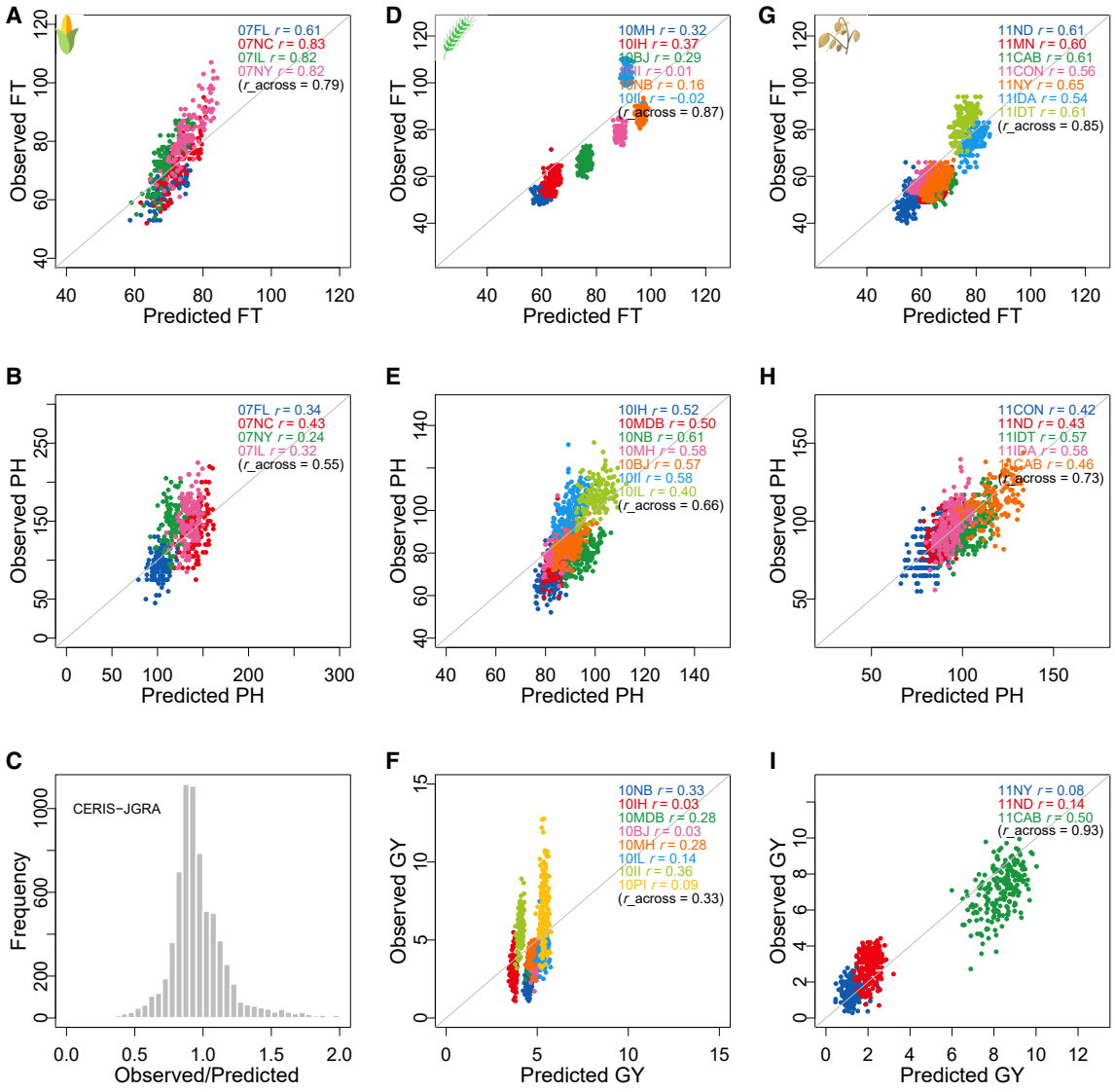

Figure 4. Performance forecasting with the CERIS-JGRA framework.

Environment profile from the first year, and phenotype and genotype data of half of each crop panel from the first year were used to identify the environmental indices and build the forecasting models. Predicted values for the other half of the panel were obtained using CERIS-JGRA for the second year and compared with the observed values.

(A) Maize FT.

(B) Maize PH.

(C) Distribution of the ratio of observed versus predicted values across all species–trait combinations.

(D) Wheat FT.

(E) Wheat PH.

(F) Wheat GY.

(G) Oat FT.

(H) Oat PH.

(I) Oat GY. Within each crop–trait combination, colors represent different trials. The gray diagonal line marks the desired 1:1 relationship between observed and predicted values.

into genetic mapping and genomic prediction. Besides improving modeling and prediction capacity, connecting findings of genetic and environmental determinants to show the relationship between a gene, organism, and environment is also critical (Li et al., 2018; Guo et al., 2020). By reinstating the environmental dimension to GWAS and GS with an environmental index, our proposed framework can be used for both complex trait dissection and prediction. By recognizing different biological responses to environmental cues, such as temperature and photoperiod, which may be either synergistic or antagonistic, the CERIS algorithm successfully identified an effective environmental index to enable the quantitative analysis of

phenotypic plasticity. Using the index built from these environmental cues, phenotypic plasticity observed in a multi-environment trial can be not only visualized and modeled by reaction norms of genotypes along the environmental gradient but also further dissected into genetic effect continua along the identified environmental gradient. With extensive characterization of environments, "envirotyping" and "enviromics" (Xu, 2016; Resende et al., 2021), we expect to see increased efforts to understand the environmental context of the observed phenotype complexity (Guo et al., 2020).

It is desirable to systematically show that individual genes underlie the plastic responses of diverse genotypes along an explicit environmental dimension for diverse natural field conditions. Because it attributes the environmental response to quantifiable factors, the CERIS-JGRA framework can facilitate comparison of genes detected under varied conditions in the same GWAS study with multiple environments or in different GWAS studies. Furthermore, it can document the genetic effect continua and gene network dynamics in a way that promotes the elucidation of molecular mechanisms (Des Marais et al., 2013; Blackman, 2017; Scheres and van der Putten, 2017). This framework can also be used to evaluate engineered alleles and modified networks from genome editing (Eshed and Lippman, 2019; Lin et al., 2020). Interestingly, we may ask how many genomic regions would be detected at different input values across the range of a single environmental index, or multiple different environmental indices, i.e., providing evidence for the recently proposed "omnigenic" nature of complex traits (Boyle et al., 2017).

On the prediction side, we believe that introducing the environmental dimension helps to establish optimally designed multi-environment trials for forecasting crop performance at regional or global scales. Two major advantages of the current approach are that photoperiod and temperature profiles are generally readily available for field experiments and that they account for a large proportion of environmental variation. On the other hand, with advances in high-throughput phenotyping, additional effort in full-scale crop modeling with data collected from field sensors would further improve the capacity for general performance forecasting and in-season forecasting, particularly for complex traits, such as grain and biomass yield, under the influence of a lot more environmental factors (regional and local) (Millet et al., 2019). Such forecasting may impact the process of price determination in commodity markets of the agricultural sector through better modeling of expected supply (Trostle, 2010). Widely used benchmarks estimated by government agencies, such as the United States Department of Agriculture, have relied on self-reported performance from contracted growers. Because the varieties released into the markets, which generally have 3–4 years of product life (Magnier et al., 2010), have been extensively tested in multiple years and locations by developers, forecasting models based on environmental indices that are tuned to specific varieties could be used. The obvious extension of this concept is to improve our ability to anticipate long-term effects of climate change, and even to anticipate, deploy, and modify specific genes and gene networks that will enable mitigation or leverage of these changes.

In terms of genetic architecture, differences were found for flowering time in maize versus wheat and oat. Although all species respond to two primary environmental factors (temperature and photoperiod), the responses among maize, wheat, and oat varied, being predicted best by either PTT or PTR. This, as well as the difference in detection of either distinct (maize) or overlapping genes (wheat and oat) for slope and intercept, may be explained partly by the specific materials involved (diverse maize panel versus breeding panels in wheat and oat) and different ranges of environments examined, although we cannot rule out the possiblity that different underlying regulatory mechanisms may have evolved in these species (Imaizumi and Kay, 2006; Andres and Coupland, 2012). While two sets of mostly non-overlapping loci were detected for slope and intercept in maize, supporting the structural gene theory (Scheiner and Lyman, 1989; Scheiner, 1993; Kusmec et al., 2017), similar sets of loci were detected for slope and intercept in wheat and oat as well as biparental mapping populations of sorghum (Li et al., 2018) and rice (Guo et al., 2020), supporting the allelic sensitivity theory (Via and Lande, 1985; Via, 1993). It is reasonable to speculate that the degree of overlap between genes associated with slope and intercept is an emergent property of the underlying genetic mechanisms and environmental conditions examined. Further analyses of diversity panels with a diversity level comparable to that of the maize panel over a comparable geographical and climatic range would be informative. Additional bivariate GWAS analysis may also be conducted to exploit the varied relationship between intercept and slope. Nevertheless, our current and previous research (Li et al., 2018; Guo et al., 2020) provided initial population-level support to show that patterns may be uncovered from the outcome of the plant perceptron interacting with environmental and developmental cues (Scheres and van der Putten, 2017) and they can be profiled for core genes in the omnigenic model (Boyle et al., 2017). It is possible that the effects of some peripheral genes are non-detectable under certain environmental conditions. For a specific environment, the effect estimates for core genes and peripheral genes can also change if strong local environmental factors drive the gene–environment interplay in different directions than the major environmental forces.

In summary, our development and demonstration of the CERIS-JGRA framework using three major crops have offered mechanistic insights into gene–environment interplay, enabling accurate, whole-genome, cross-environment performance prediction. This framework can be extended to further incorporate gene expression, regulatory network, and other functional characterization data to shed light on questions in genetics, ecology, and evolution.

METHODS

Germplasm and phenotype

The maize diversity panel was assembled to capture the genetic diversity in maize germplasm (Flint-Garcia et al., 2005). This population was evaluated in 10 environments in 2006 and 2007 and phenotype data were deposited in Panzea (https://www.panzea.org/phenotypes). The wheat diversity panel was assembled from advanced spring wheat lines selected from globally distributed nursery fields by the International Maize and Wheat Improvement Center. Phenotypes of these advanced lines were evaluated in 2009 and 2010, and data were deposited in the International Maize and Wheat Improvement Center Dataverse (https://data.cimmyt.org/) (Sukumaran et al., 2017). The oat diversity panel population was a set of spring oat genotypes nominated by breeders

and evaluated in 2010 and 2011 (Esvelt Klos et al., 2016). Field information and phenotypes were deposited in T3/oat (https://triticeaetoolbox.org/oat/). GPS coordinates and planting dates are listed in Supplemental Table 1. Not all traits were scored in all environments. For some traits, environments with a high missing observation rate or with a low correlation between replications were removed from further analysis. For example, one maize environment (06NY) was removed from analysis because 25% of genotypes did not have flowering time records.

The maize diversity panel with 282 inbred lines was planted in Puerto Rico, Florida, North Carolina, Missouri, Illinois, and New York. Each location was arranged in an augmented lattice design with a single replicate. Each genotype was planted in a single-row plot of 8–15 plants. Flowering time (days to anthesis) was measured as the days from planting to median anthesis in a plot. Plant height was measured as the distance in centimeters from the soil line of the plant to the base of the flag leaf at reproductive maturity. The wheat diversity panel of 287 elite lines was phenotyped in major wheat-growing areas of India, Pakistan, Nepal, Bangladesh, Iran, Egypt, Sudan, and Mexico. In each location, plants were grown in plots with an α-lattice design with two replications. Flowering time (days to heading) was measured from planting to when 50% of the plants in the plot showed spikes coming out of the boot leaves. Plant height was measured as the length of individual culms from the soil surface to the tip of the spike. Grain yield was the grain weight per square meter and converted to tons per hectare. The oat diversity panel with 434 genotypes was planted by sowing 50 g of seed in unreplicated four to five row plots of 1.2–2.2 m length in Idaho, North Dakota, and New York, United States and in 3.25–4.6 m^2 plots in Lacombe and Ottawa, Canada. In addition, 15–20 seeds were planted in hill plots in Minnesota in the United States. Flowering time (days to heading) was measured from planting to when half of the plants in a plot had emerged inflorescences. Plant height was measured from the soil surface to the tip of panicle. Grain yield was recorded as grams per square meter and converted to tons per hectare.

Genomic data

Genotyping data of the maize population were a subset of maize HapMap3 (Bukowski et al., 2018), and 12 million maize SNPs were obtained after filtering out SNPs with a missing rate higher than 20% and minor allele frequency less than 1%. The wheat population was genotyped with an Illumina iSelect 90K SNP assay providing data for 26 814 SNPs (Sukumaran et al., 2015). The oat population was genotyped at 35 176 genotyping-by-sequencing-derived SNPs, which were anchored to a consensus linkage map (Bekele et al., 2018).

CERIS

Daily temperature and DL data (civil twilight, hours) from the entire growing season were retrieved from National Centers for Environmental Information (NOAA, http://www.ncdc.noaa.gov) and The Astronomical Applications Department of the U.S. Naval Observatory (http://aa.usno.navy.mil/data/index.php), respectively. For the i^{th} DAP in the e^{th} environment, daily GDD was calculated as follows:$GDD_{ei} = (T_{max\_ei} + T_{min\_ei}) \div 2 - T_{base}$; where $T_{max_ei}$is the maximum temperature (°F, with the highest observation of 86°F for maize, and 100°F for wheat and oat), $T_{\min\_ei}$ is the minimum temperature, and $T_{base}$is the species-specific base temperature (50°F for maize and 32°F for wheat and oat) (Bauer et al., 1984; McMaster and Wilhelm, 1997; Cousens et al., 2003). For $T_{\min\_ei} < T_{base}$, $T_{\min\_ei}$ was set to T_{base}. Two composite environmental parameters, PTT and PTR, were derived from temperature and photoperiod as follows: $PTT_{ei} = GDD_{ei} \times DL_{ei}$; $PTR_{ei} = GDD_{ei} \div DL_{ei}$.

We implemented the CERIS algorithm with R to identify the environmental indices (https://github.com/jmyu/CERIS_JGRA). In brief, for any target trait evaluated in n (n>3) environments, the algorithm calculates the average performance of the evaluated population at each environment (environmental mean of the trait). The pipeline then calculates the average value of the environment parameters (DL, GDD, PTT, or PTR) for the window from DAP_i to DAP_j, where DAP_i is the i^{th} day after planting, DAP_j is the j^{th} day after planting, and $i<j-5$. The correlation between the environmental mean vector and each parameter–window value vector is calculated and stored separately. The parameter–window combination with the strongest correlation (either positive or negative) is then chosen as the environmental index and used for further analyses. We restricted the search for the window to the times before the trait is fully developed, following the sequence of physiological growth. Because of the overlapping nature of windows, the window with the strongest correlation for an environmental parameter is generally supported by the surrounding windows. To minimize spurious search results, we restricted the window to a span of at least 5 days, and conducted leave-one-environment-out cross-validation (i.e., the average of n correlations was reported for each window).

The high correlation between environmental index and environmental means is interpreted as follows. For each environment, the environmental mean of the measured trait was obtained by averaging across all genotypes: 282 lines in maize, 287 in wheat, and 434 in oat. With a large sample size, the average value (environmental mean) is an accurate reflection of the environment effect on the whole population. In other words, if the whole population is regarded as a single entity, its performance (represented by environmental means) is explained adequately by the environmental effect captured by the environmental index. When environmental parameters (DL, GDD, PTT, or PTR) within different windows of time are searched, we expect to observe a high level of correlation when the environmental parameter with the period of the growing season is a major differentiating factor underlying the differences among the environmental means dervied from the performance data. Moreover, the parameter–window combination selected as the environmental index should be surrounded by other windows with slightly decreased correlation values.

Performance prediction

CERIS-JGRA performance prediction includes three scenarios: (1) predicting the performance of tested genotypes in untested environments, (2) predicting the performance of untested genotypes in tested environments, and (3) predicting the performance of untested genotypes in untested environments (Li et al., 2018; Guo et al., 2020). Unlike the traditional joint regression analysis, CERIS-JGRA involves performance prediction for genotypes without performance data by exploiting the genomic relationship between tested and untested genotypes, and performance prediction for individuals without performance data in future environments by exploiting both genomic relationship and environmental index.

For the first scenario, the leave-one-environment-out cross-validation was conducted. (1) Let the j^{th} (j = 1, 2, 3…, m) environment be the untested environment, and the remaining the training (tested) environments. (2) Search for the environmental index with CERIS by using the environmental means from the tested genotypes in the tested environments. (3) For the i^{th} (i = 1, 2, 3…, n) genotype, regress the observed phenotypes from the tested environments on the identified environmental index to obtain intercept and slope estimates. (4) Predict the phenotype in the j^{th} untested environment by supplying the fitted linear models (regression models from step 3) with the value of the corresponding environmental index value from the j^{th} untested (to be predicted) environment. (5) Repeat steps 1 to 4 until each environment is predicted.

For the second scenario, the leave-one-half-of-genotypes-out cross-validation was conducted. (1) Equally split n genotypes into tested genotypes and untested genotypes. (2) Search for the environmental index with CERIS by using the environmental means from the tested genotypes. (3) Regress the observed phenotypes on the identified environmental index to obtain intercept and slope estimates for each tested genotype. (4) Treating intercept and slope as new "traits", run genomic prediction through "rrBLUP" (Endelman, 2011) to predict the intercept and slope

for each untested genotype. (5) Predict the phenotypes of the untested genotypes with the predicted intercept and slope and the environmental-index value of each environment.

For the third scenario, the joint leave-one-environment-out and one-half-of-genotypes-out cross-validation was conducted. (1) Let the j^{th} (j = 1, 2, 3…, m) environment be the untested environment, and the remaining the training (tested) environments. (2) Equally split n genotypes into tested genotypes and untested genotypes. (3) Search the environmental index with CERIS by using the environmental means from the tested genotypes in the tested environments. (4) Regress the observed phenotypes on the identified environmental index to obtain intercept and slope estimates for each tested genotype. (5) Treating intercept and slope as new "traits", run genomic prediction through rrBLUP to predict the intercept and slope for each untested genotype. (6) Predict phenotypes of the untested genotypes in the untested environment with the predicted intercept and slope and the environmental-index value from the untested environment. (7) Repeat steps 1 to 6 until each environment is processed.

The joint regression analysis model is as follows: $P_{ij} = g_i + b_i t_j + e_{ij}$, where P_{ij} is the phenotypic value of genotype i in environment j; g_i is the genotypic effect (intercept); b_i is the regression coefficient (slope); t_j is the environmental-index value of environment j centered by the average environmental-index values across the environments; and e_{ij} is the residual. Average values from two replications were used for wheat traits, while single-replication observations were used for maize and oat traits. The genomic prediction model in JGRA uses the following mixed model: $y = 1\mu + Zg + e$, where y is the phenotypic values (i.e., intercept or slope derived for the original phenotype), μ is the overall mean (fixed effect), g is the vector of random genetic effects, and e is the vector of residuals. 1 and Z are incidence matrices. The variance of the random effects u is $var(u) = A\sigma_g^2$, where A is the genomic relationship matrix and σ_g^2 is the additive genetic variance. The variance of the residuals e is $var(e) = I\sigma_e^2$, where I is the identity matrix.

JGRA is a generic framework that is applicable to input datasets with sizes ranging from small to large. For genomic prediction with the genome-wide relationship or genome-wide marker effect estimation, rrBLUP is used as the default setting, which can be customized to accommodate other methods. In addition, splitting environments into training and testing environments is user-defined, so is splitting genotypes.

When the multi-environment trial involves 2 years with adequate sites within each year, the generic term "prediction" can be replaced with "forecasting". The purpose of CERIS-JRA forecasting is to predict the performance of tested genotypes in the second year. We split the multi-environment trial into training environments and testing environments based on the planting year. Trials planted in the first year were used as training environments and those planted in the second year as testing environments. Because the prediction scenario predicts tested genotypes in untested environments (i.e., genotypes are all tested), genomics is not involved in this forecasting. The purpose of CERIS-JGRA forecasting is to predict performances of untested genotypes in the second year. In this case, we equally split genotypes into two halves, one as a training set and the other as a testing set. By applying CERIS-JGRA, performance forecasting can be done for the second half of genotypes in the second-year environments.

Prediction accuracy was calculated as the correlation between observed and predicted values. The predicted values were generated from the model developed from the training set of tested genotypes under tested environments. Prediction accuracy was assessed across all environments and at the individual environment level. When sampling of the genotypes was conducted, the average value of prediction accuracy across 50 runs was calculated. A representative run is shown in the figure for illustration.

Identification of loci associated with phenotypic plasticity through GWAS

Reaction norm relates the environments to which a particular genotype is exposed to the phenotypes produced by the genotype. With the identified environmental index, two reaction norm-norm parameters (intercept and slope) were obtained for each trait using the joint regression analysis of the observed trait values across environments. Treating the estimates of intercept and slope as two derived traits, the established mixed model GWAS procedure (Yu et al., 2006; Zhang et al., 2010) implemented in GAPIT (Tang et al., 2016) was used, separately for each trait, to identify genomic loci underlying the variation observed for slope and intercept across different genotypes.

Both principal components and relative kinship were used to control genetic relatedness through the default model selection process before testing individual SNPs. Genome-wide significance thresholds were first determined using the false discovery rate at the 0.05 level (Benjamini and Hochberg, 1995). For maize GWAS, when a false discovery rate at the 0.05 level was not obtainable, a commonly used suggestive threshold value ($-\log_{10}(P) = 6$) was plotted as a visual aid. For wheat and oat GWAS with less SNP coverage than maize, we used the SimpleM method to obtain the significance threshold (Gao et al., 2008, 2010). To model the genetic effect dynamics along the environmental gradient, SNPs significantly associated with intercept and slope were tested for association with flowering time within each environment using the mixed model method in GAPIT. The same procedure was done for five additional known genes: *PhyB2*, *CIB1*, and *ZmPRR59* for maize and *VRN-B1* and *VRN-D3* for wheat. These separate genetic effects were then regressed on the environmental index to generate the fitted lines as the genetic effect continua, i.e., varied genetic effects with different environmental inputs.

Sequences surrounding the genes (*ZCN8*, *Conz1*, *ZMM5*, and *ZMM22*/*ZmMADS69*) implicated in flowering-time phenotypic plasticity were retrieved from MaizeGDB (https://www.maizegdb.org/genome/assemblies_overview) for B73, B104, CML247, EP1, F7, PH207, W22, and Mo17. Because annotation information for most assemblies was not available, coding sequences from B73 were used as queries in BLAST searches to anchor the corresponding sequences in other assemblies. The collinear sequences among eight inbred lines were BLASTed against each other to identify structural variations. Transposon elements encoded in the targeted sequences were identified with CENSOR (www.girinst.org). Primer 3 (http://bioinfo.ut.ee/primer3/) was used to design primers to amplify sequences near insertion/deletion sites. PCR amplicons were separated in an agarose gel to verify the potential structural variations.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Molecular Plant Online*.

FUNDING

This work is supported by the Agriculture and Food Research Initiative competitive grant (2021-67013-33833) from the USDA National Institute of Food and Agriculture, the Advanced Research Projects Agency-Energy program (DEAR0000826) of the Department of Energy, the National Science Foundation (IOS-1546657), the Iowa State University Raymond F. Baker Center for Plant Breeding, and the Iowa State University Plant Sciences Institute.

AUTHOR CONTRIBUTIONS

Conceptualization and methodology, X.L., T.G., and J.Y.; investigation, X.L., T.G., J.Y., W.A.B., S.S., A.E.V., J.P.McN., L.T.C., and M.S.L.; writing – original draft, X.L., T.G., and J.Y.; writing – review & editing, all authors; resources, K.R.L., M.E.W., J.K.M, S.V.A., M.P.R., N.A.T., P.S.S., and J.Y.

ACKNOWLEDGMENTS

No conflict of interest declared.

Received: September 30, 2020
Revised: February 3, 2021
Accepted: March 9, 2021
Published: March 9, 2021

REFERENCES

Andres, F., and Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. Nat. Rev. Genet. **13**:627–639.

Angus, J.F., Mackenzie, D.H., Morton, R., and Schafer, C.A. (1981). Phasic development in field crops II. Thermal and photoperiodic responses of spring wheat. Field Crops Res. **4**:269–283.

Bauer, A., Frank, A., and Black, A. (1984). Estimation of spring wheat leaf growth rates and anthesis from air temperature. Agron. J. **76**:829–835.

Bekele, W.A., Wight, C.P., Chao, S., Howarth, C.J., and Tinker, N.A. (2018). Haplotype-based genotyping-by-sequencing in oat genome research. Plant Biotechnol. J. **16**:1452–1463.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: lessons from comparative QTL approach to multipletesting. J. R. Stat. Soc. B **57**:289–300.

Bevan, M.W., Uauy, C., Wulff, B.B., Zhou, J., Krasileva, K., and Clark, M.D. (2017). Genomic innovation for crop improvement. Nature **543**:346–354.

Blackman, B.K. (2017). Changing responses to changing seasons: natural variation in the plasticity of flowering time. Plant Physiol. **173**:16–26.

Bouchet, S., Servin, B., Bertin, P., Madur, D., Combes, V., Dumas, F., Brunel, D., Laborde, J., Charcosset, A., and 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**:e71377.

Boyle, E.A., Li, Y.I., and Pritchard, J.K. (2017). An expanded view of complex traits: from polygenic to omnigenic. Cell **169**:1177–1186.

Brachi, B., Faure, N., Horton, M., Flahauw, E., Vazquez, A., Nordborg, M., Bergelson, J., Cuguen, J., and Roux, F. (2010). Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. Plos Genet. **6**:e1000940.

Bukowski, R., Guo, X., Lu, Y., Zou, C., He, B., Rong, Z., Wang, B., Xu, D., Yang, B., Xie, C., et al. (2018). Construction of the third-generation *Zea mays* haplotype map. Gigascience **7**:1–12.

Cooper, M., Technow, F., Messina, C., Gho, C., and Totir, L.R. (2016). Use of crop growth models with whole-genome prediction: application to a maize multienvironment trial. Crop Sci. **56**:2141–2156.

Costa-Neto, G., Fritsche-Neto, R., and Crossa, J. (2021). Nonlinear kernels, dominance, and envirotyping data increase the accuracy of genome-based prediction in multi-environment trials. Heredity **126**:92–106.. https://doi.org/10.1038/s41437-020-00353-1.

Cousens, R.D., Barnett, A.G., and Barry, G.C. (2003). Dynamics of competition between wheat and oat. Agron. J. **95**:1295–1304.

Crossa, J., Perez-Rodriguez, P., Cuevas, J., Montesinos-Lopez, O., Jarquin, D., de Los Campos, G., Burgueno, J., Gonzalez-Camacho, J.M., Perez-Elizalde, S., Beyene, Y., et al. (2017). Genomic selection in plant breeding: methods, models, and perspectives. Trends Plant Sci. **22**:961–975.

Des Marais, L.D., Hernandez, K.M., and Juenger, T.E. (2013). Genotype-by-environment interaction and plasticity: exploring genomic responses of plants to the abiotic environment. Annu. Rev. Ecol. Evol. Syst. **44**:5–29.

Eberhart, S.A., and Russell, W.A. (1966). Stability parameters for comparing varieties. Crop Sci. **6**:36–40.

Endelman, J.B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome **4**:250–255.

Eshed, Y., and Lippman, Z.B. (2019). Revolutions in agriculture chart a course for targeted breeding of old and new crops. Science **366**:eaax0025.

Esvelt Klos, K., Huang, Y.F., Bekele, W.A., Obert, D.E., Babiker, E., Beattie, A.D., Bjornstad, A., Bonman, J.M., Carson, M.L., Chao, S., et al. (2016). Population genomics related to adaptation in elite oat germplasm. Plant Genome **9**. https://doi.org/10.3835/plantgenome2015.10.0103.

Exposito-Alonso, M., Genomes Field Experiment Team, Burbano, H.A., Bossdorf, O., Nielsen, R., and Weigel, D. (2019). Natural selection on the *Arabidopsis thaliana* genome in present and future climates. Nature **573**:126–129.

Fan, S., Hansen, M.E., Lo, Y., and Tishkoff, S.A. (2016). Going global by adapting local: a review of recent human adaptation. Science **354**:54–59.

Finlay, K., and Wilkinson, G. (1963). The analysis of adaptation in a plant-breeding programme. Aust. J. Agric. Res. **14**:742–754.

Flint-Garcia, S.A., Thuillet, A.C., Yu, J., Pressoir, G., Romero, S.M., Mitchell, S.E., Doebley, J., Kresovich, S., Goodman, M.M., and Buckler, E.S. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. Plant J. **44**:1054–1064.

Gage, J.L., Jarquin, D., Romay, C., Lorenz, A., Buckler, E.S., Kaeppler, S., Alkhalifah, N., Bohn, M., Campbell, D.A., Edwards, J., et al. (2017). The effect of artificial selection on phenotypic plasticity in maize. Nat. Commun. **8**:1348.

Gao, X., Becker, L.C., Becker, D.M., Starmer, J.D., and Province, M.A. (2010). Avoiding the high Bonferroni penalty in genome-wide association studies. Genet. Epidemiol. **34**:100–105.

Gao, X., Starmer, J., and Martin, E.R. (2008). A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet. Epidemiol. **32**:361–369.

Grotzinger, A.D., Rhemtulla, M., de Vlaming, R., Ritchie, S.J., Mallard, T.T., Hill, W.D., Ip, H.F., McIntosh, A.M., Deary, I.J., and Koellinger, P.D. (2018). Genomic SEM provides insights into the multivariate genetic architecture of complex traits. BioRxiv, 305029.

Guo, L., Wang, X., Zhao, M., Huang, C., Li, C., Li, D., Yang, C.J., York, A.M., Xue, W., Xu, G., et al. (2018). Stepwise *cis*-regulatory changes in ZCN8 contribute to maize flowering-time adaptation. Curr. Biol. **28**:3005–3015.e4.

Guo, T., Mu, Q., Wang, J., Vanous, A.E., Onogi, A., Iwata, H., Li, X., and Yu, J. (2020). Dynamic effects of interacting genes underlying rice flowering-time phenotypic plasticity and global adaptation. Genome Res. **30**:673–683.

Hammer, G.L., Goyne, P.J., and Woodruff, D.R. (1982). Phenology of sunflower cultivars. 3. Models for prediction in field environments. Aust. J. Agric. Res. **33**:263–274.

Heslot, N., Akdemir, D., Sorrells, M.E., and Jannink, J.L. (2014). Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor. Appl. Genet. **127**:463–480.

Hickey, L.T., Ambar, N.H., Robinson, H., Jackson, S.A., Leal-Bertioli, S.C.M., Tester, M., Gao, C., Godwin, I.D., Hayes, B.J., and Wulff, B.B.H. (2019). Breeding crops to feed 10 billion. Nat. Biotechnol. **37**:744–754.

Hirsch, C.N., Foerster, J.M., Johnson, J.M., Sekhon, R.S., Muttoni, G., Vaillancourt, B., Penagaricano, F., Lindquist, E., Pedraza, M.A.,

Barry, K., et al. (2014). Insights into the maize pan-genome and pan-transcriptome. Plant Cell **26**:121–135.

Hung, H.Y., Shannon, L.M., Tian, F., Bradbury, P.J., Chen, C., Flint-Garcia, S.A., McMullen, M.D., Ware, D., Buckler, E.S., Doebley, J.F., et al. (2012). *ZmCCT* and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. Proc. Natl. Acad. Sci. U S A **109**:E1913–E1921.

Imaizumi, T., and Kay, S.A. (2006). Photoperiodic control of flowering: not only by coincidence. Trends Plant Sci. **11**:550–558.

Jarquin, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., Piraux, F., Guerreiro, L., Perez, P., Calus, M., et al. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor. Appl. Genet. **127**:595–607.

Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.Y., Freimer, N.B., Sabatti, C., and Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. **42**:348–354.

Kusmec, A., Srinivasan, S., Nettleton, D., and Schnable, P.S. (2017). Distinct genetic architectures for phenotype means and plasticities in *Zea mays*. Nat. Plants **3**:715–723.

Li, X., Guo, T.T., Mu, Q., Li, X.R., and Yu, J.M. (2018). Genomic and environmental determinants and their interplay underlying phenotypic plasticity. Proc. Natl. Acad. Sci. U S A **115**:6679–6684.

Liang, Y., Liu, Q., Wang, X., Huang, C., Xu, G., Hey, S., Lin, H.Y., Li, C., Xu, D., Wu, L., et al. (2019). ZmMADS69 functions as a flowering activator through the ZmRap2.7-ZCN8 regulatory module and contributes to maize flowering time adaptation. New Phytol. **221**:2335–2347.

Lin, H.Y., Liu, Q., Li, X., Yang, J., Liu, S., Huang, Y., Scanlon, M.J., Nettleton, D., and Schnable, P.S. (2017). Substantial contribution of genetic variation in the expression of transcription factors to phenotypic variation revealed by eRD-GWAS. Genome Biol. **18**:192.

Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., Wang, Y., Anzalone, A.V., Raguram, A., Doman, J.L., et al. (2020). Prime genome editing in rice and wheat. Nat. Biotechnol. **38**:582–585.

Lipka, A.E., Kandianis, C.B., Hudson, M.E., Yu, J., Drnevich, J., Bradbury, P.J., and Gore, M.A. (2015). From association to prediction: statistical methods for the dissection and selection of complex traits in plants. Curr. Opin. Plant Biol. **24**:110–118.

Lobell, D.B., Burke, M.B., Tebaldi, C., Mastrandrea, M.D., Falcon, W.P., and Naylor, R.L. (2008). Prioritizing climate change adaptation needs for food security in 2030. Science **319**:607–610.

Lobell, D.B., Roberts, M.J., Schlenker, W., Braun, N., Little, B.B., Rejesus, R.M., and Hammer, G.L. (2014). Greater sensitivity to drought accompanies maize yield increase in the U.S. Midwest. Science **344**:516–519.

Mackay, T.F., Stone, E.A., and Ayroles, J.F. (2009). The genetics of quantitative traits: challenges and prospects. Nat. Rev. Genet. **10**:565–577.

Magnier, A., Kalaitzandonakes, N., and Miller, D.J. (2010). Product life cycles and innovation in the US seed corn industry. Int. Food Agribus. Man **13**:17–36.

Malosetti, M., Ribaut, J.M., and van Eeuwijk, F.A. (2013). The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. Front Physiol. **4**:44.

Masle, J., Doussinault, G., Farquhar, G.D., and Sun, B. (1989). Foliar stage in wheat correlates better to photothermal time than to thermal time. Plant Cell Environ. **12**:235–247.

McCouch, S., Baute, G.J., Bradeen, J., Bramel, P., Bretting, P.K., Buckler, E., Burke, J.M., Charest, D., Cloutier, S., Cole, G., et al. (2013). Agriculture: feeding the future. Nature **499**:23–24.

McMaster, G.S., and Wilhelm, W.W. (1997). Growing degree-days: one equation, two interpretations. Agric. For. Meteorol. **87**:291–300.

Meng, X., Muszynski, M.G., and Danilevskaya, O.N. (2011). The *FT*-like *ZCN8* gene functions as a floral activator and is involved in photoperiod sensitivity in maize. Plant Cell **23**:942–960.

Miller, T.A., Muslin, E.H., and Dorweiler, J.E. (2008). A maize *CONSTANS*-like gene, *conz1*, exhibits distinct diurnal expression patterns in varied photoperiods. Planta **227**:1377–1388.

Millet, E.J., Kruijer, W., Coupel-Ledru, A., Alvarez Prado, S., Cabrera-Bosquet, L., Lacube, S., Charcosset, A., Welcker, C., van Eeuwijk, F., and Tardieu, F. (2019). Genomic prediction of maize yield across European environmental conditions. Nat. Genet. **51**:952–956.

Moore, R., Casale, F.P., Bonder, M.J., Horta, D., Franke, L., Barroso, I., Stegle, O., and Consortium, B. (2019). A linear mixed-model approach to study multivariate gene-environment interactions. Nat. Genet. **51**:180.

Peng, B., Guan, K., Tang, J., Ainsworth, E.A., Asseng, S., Bernacchi, C.J., Cooper, M., Delucia, E.H., Elliott, J.W., Ewert, F., et al. (2020). Towards a multiscale crop modelling framework for climate change adaptation assessment. Nat. Plants **6**:338–348.

Resende, R.T., Piepho, H.P., Rosa, G.J.M., Silva-Junior, O.B., FF, E.S., de Resende, M.D.V., and Grattapaglia, D. (2021). Enviromics in breeding: applications and perspectives on envirotypic-assisted selection. Theor. Appl. Genet. **134**:95–112.

Robertson, G.W. (1968). A biometeorological time scale for a cereal crop involving day and night temperatures and photoperiod. Int. J. Biometeorol. **12**:191–223.

Romero Navarro, J.A., Willcox, M., Burgueno, J., Romay, C., Swarts, K., Trachsel, S., Preciado, E., Terron, A., Delgado, H.V., Vidal, V., et al. (2017). A study of allelic diversity underlying flowering-time adaptation in maize landraces. Nat. Genet. **49**:476–480.

Salvi, S., Sponza, G., Morgante, M., Tomes, D., Niu, X., Fengler, K.A., Meeley, R., Ananiev, E.V., Svitashev, S., Bruggemann, E., et al. (2007). Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. Proc. Natl. Acad. Sci. U S A **104**:11376–11381.

Scheiner, S.M. (1993). Genetics and evolution of phenotypic plasticity. Annu. Rev. Ecol. Syst. **24**:35–68.

Scheiner, S.M., and Lyman, R.F. (1989). The genetics of phenotypic plasticity I. Heritability. J. Evol. Biol. **2**:95–107.

Scheres, B., and van der Putten, W.H. (2017). The plant perceptron connects environment to development. Nature **543**:337–345.

Sukumaran, S., Crossa, J., Jarquin, D., Lopes, M., and Reynolds, M.P. (2017). Genomic prediction with pedigree and genotype x environment interaction in spring wheat grown in south and West Asia, North Africa, and Mexico. G3 (Bethesda) **7**:481–495.

Sukumaran, S., Dreisigacker, S., Lopes, M., Chavez, P., and Reynolds, M.P. (2015). Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. Theor. Appl. Genet. **128**:353–363.

Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., Tian, F., Su, Z., Pan, Y., Liu, D., Lipka, A.E., et al. (2016). GAPIT version 2: an enhanced integrated tool for genomic association and prediction. Plant Genome **9**. https://doi.org/10.3835/plantgenome2015.11.0120.

Thomas, D. (2010). Gene–environment-wide association studies: emerging approaches. Nat. Rev. Genet. **11**:259–272.

Tibbs Cortes, L., Zhang, Z., and Yu, J. (2021). Status and prospects of genome-wide association studies in plants. Plant Genome **14**:e20077.. https://doi.org/10.1002/tpg2.20077.

Trostle, R. (2010). Global Agricultural Supply and Demand: Factors Contributing to the Recent Increase in Food Commodity Prices (BiblioGov).

Urbut, S.M., Wang, G., Carbonetto, P., and Stephens, M. (2019). Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions. Nat. Genet. **51**:187.

Via, S. (1993). Adaptive phenotypic plasticity: target or by-product of selection in a variable environment? Am. Nat. **142**:352–365.

Via, S., and Lande, R. (1985). Genotype-environment interaction and the evolution of phenotypic plasticity. Evolution **39**:505–522.

Wheeler, T., and von Braun, J. (2013). Climate change impacts on global food security. Science **341**:508–513.

Wilczek, A.M., Roe, J.L., Knapp, M.C., Cooper, M.D., Lopez-Gallego, C., Martin, L.J., Muir, C.D., Sim, S., Walker, A., Anderson, J., et al. (2009). Effects of genetic perturbation on seasonal life history plasticity. Science **323**:930–934.

Xu, Y. (2016). Envirotyping for deciphering environmental impacts on crop plants. Theor. Appl. Genet. **129**:653–673.

Xu, Y., Liu, X., Fu, J., Wang, H., Wang, J., Huang, C., Prasanna, B.M., Olsen, M.S., Wang, G., and Zhang, A. (2020). Enhancing genetic gain through genomic selection: from livestock to plants. Plant Commun. **1**:100005.

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., and Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene *VRN1*. Proc. Natl. Acad. Sci. U S A **100**:6263–6268.

Yang, Q., Li, Z., Li, W., Ku, L., Wang, C., Ye, J., Li, K., Yang, N., Li, Y., Zhong, T., et al. (2013). CACTA-like transposable element in *ZmCCT* attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. Proc. Natl. Acad. Sci. U S A **110**:16969–16974.

Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. **38**:203–208.

Zhang, Z., Ersoz, E., Lai, C.Q., Todhunter, R.J., Tiwari, H.K., Gore, M.A., Bradbury, P.J., Yu, J., Arnett, D.K., Ordovas, J.M., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. Nat. Genet. **42**:355–360.