

## ORIGINAL ARTICLE

# Near-infrared reflectance spectroscopy phenomic prediction can perform similarly to genomic prediction of maize agronomic traits across environments

Aaron J. DeSalvio<sup>1</sup>  | Alper Adak<sup>2</sup>  | Seth C. Murray<sup>2</sup>  | Diego Jarquín<sup>3</sup>  |  
Noah D. Winans<sup>2</sup>  | Daniel Crozier<sup>2</sup>  | William L. Rooney<sup>2</sup> 

<sup>1</sup>Interdisciplinary Graduate Program in Genetics and Genomics (Department of Biochemistry and Biophysics), Texas A&M University, College Station, Texas, USA

<sup>2</sup>Department of Soil and Crop Sciences, Texas A&M University, College Station, Texas, USA

<sup>3</sup>Department of Agronomy, University of Florida, Gainesville, Florida, USA

## Correspondence

Seth C. Murray, Department of Soil and Crop Sciences, Texas A&M University, College Station, TX 77843-2474, USA.  
Email: [sethmurray@tamu.edu](mailto:sethmurray@tamu.edu).

Assigned to Associate Editor Feng Zhang.

## Funding information

Texas A&M University/Association of Former Students (TAMU/AFS) Graduate Merit Fellowship; Texas Corn Producers Board; Texas A&M AgriLife Research; Texas A&M University Eugene Butler Endowed Chair; National Institute of Food and Agriculture, Grant/Award Numbers: 2020-68013-32371, 2021-67013-33915, Hatch Multistate Research Fund; National Science Foundation, Grant/Award Number: GRFP; Republic of Turkey—Ministry of National Education and Ministry of Agriculture and Forestry

## Abstract

For nearly two decades, genomic prediction and selection have supported efforts to increase genetic gains in plant and animal improvement programs. However, novel phenomic strategies for predicting complex traits in maize have recently proven beneficial when integrated into across-environment sparse genomic prediction models. One phenomic data modality is whole grain near-infrared spectroscopy (NIRS), which records reflectance values of biological samples (e.g., maize kernels) based on chemical composition. Predictions of hybrid maize grain yield (GY) and 500-kernel weight (KW) across 2 years (2011–2012) and two management conditions (water-stressed and well-watered) were conducted using combinations of reflectance data obtained from high-throughput, F<sub>2</sub> whole-kernel scans and genomic data obtained from genotyping-by-sequencing within four different cross-validation (CV) schemes (CV2, CV1, CV0, and CV00). When predicting the performance of untested genotypes in characterized (CV1) environments, genomic data were better than phenomic data for GY ( $0.689 \pm 0.024$ —genomic vs.  $0.612 \pm 0.045$ —phenomic), but phenomic data were better than genomic data for KW ( $0.535 \pm 0.034$ —genomic vs.  $0.617 \pm 0.145$ —phenomic). Multi-kernel models (combinations of phenomic and genomic relationship matrices) did not surpass single-kernel models for GY prediction in CV1 or CV00 (prediction of untested genotypes in uncharacterized environments); however, these models did outperform the single-kernel models for prediction of KW in these same CVs. Lasso regression applied to the NIRS data set selected a subset of 216 NIRS bands that achieved comparable prediction abilities to the full phenomic data set of 3112 bands predicting GY and KW under CV1 and CV00.

**Abbreviations:** CV, cross-validation; GBS, genotyping-by-sequencing; GY, grain yield; KW, 500-kernel weight; NIRS, near-infrared spectroscopy; PP, phenomic prediction; UAS, unoccupied aerial system; WS, water-stressed; WW, well-watered.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *The Plant Genome* published by Wiley Periodicals LLC on behalf of Crop Science Society of America.

### Plain Language Summary

As a common method of identifying high-performing crop varieties, genomic prediction has been successful in research and commercial crop improvement programs for nearly two decades. However, to determine if a nondestructive method of acquiring spectral data could perform comparably with genomic prediction, near-infrared spectroscopy (NIRS) of intact maize kernels was conducted in this study using plants grown in multiple environments. Genomic and phenomic prediction were compared for predicting agronomic traits using combinations of genomic-only, phenomic-only, and combined prediction models. Though in many cases genomic prediction outperformed NIRS-based phenomic prediction, these findings indicate that NIRS can still enable comparable prediction ability, highlighting its potential role as a high-throughput method for predicting economically relevant crop traits. These findings are relevant for breeding programs seeking to screen varieties rapidly and nondestructively.

## 1 | INTRODUCTION

Near-infrared reflectance (NIR) spectroscopy (NIRS) is an active remote sensing approach that leverages techniques in spectroscopy, analytical chemistry, and mathematics (specifically chemometrics) (Pasquini, 2018). NIRS functions by barraging a sample with electromagnetic radiation at wavelengths between 750 and 2500 nm, producing absorption or reflectance values of any compound containing C–H, N–H, S–H, or O–H chemical bonds (Pasquini, 2003, 2018). NIRS measurements are generally recorded in laboratory settings using dried tissue or grain (Ferrio et al., 2004), although portable spectrometers can provide active or passive, non-destructive measurements of plant tissue in the field. NIR spectrometers assess each sample as a single measurement at each discrete band and differ from hyperspectral NIR imaging, which creates a “hyperspectral data cube” with spatial information but generally measures fewer discrete bands. Hernandez et al. (2015) demonstrated the ability to predict grain yield (GY) using canopy NIRS measurements collected at anthesis and grain fill as a proxy trait for GY in a diverse wheat panel. Recently, some small plot combines and forage harvesters have been equipped with NIRS to minimize labor and bias of subsampling grain or biomass for composition (Polytec GmbH, 2022). Given the existing prevalence of NIRS in plant science and breeding programs for kernel composition and quality testing, many researchers are already equipped to incorporate NIRS as a method for phenomic prediction (PP) using reflectance values as phenomic markers instead of genomic markers, in turn facilitating PP and potentially phenomic selection (PS). In the first major NIRS-based PP study, Rincet et al. (2018) demonstrated the ability of NIR spectra to enable prediction accuracy of complex traits on par with marker-based prediction, despite subjecting the

models to environments distinct from the one in which NIRS data were recorded. The study employed covariance matrices based on Yamada et al. (1988) to predict wheat yield in a separate environment, highlighting the role NIRS-based relationship matrices can play in demarcating genetic relationships between lines. NIR spectra have been used previously to estimate complex plant compositional traits and then to predict GY based on the relationships that exist between GY and composition. Spectra can also be implemented in combination with molecular markers to bolster genomic prediction accuracy or coupled with a more commonly recognized genomic relationship (kinship) matrix (Robert, Brault et al., 2022). Phenomic prediction was proposed on the basis that phenomic markers, such as those afforded by NIRS, capture genetic relatedness, enabling across-environment predictions from spectra obtained in different environments. This concept, demonstrated in Rincet et al. (2018), produced predictions with higher accuracy at times than genomic prediction (Robert, Brault et al., 2022). Weiß et al. (2022) reported more stable prediction ability of NIRS phenomic versus genomic prediction when predicting dry grain matter content, GY, and phosphorus concentration across diverse breeding material. Weiß et al. (2022) also found genomic prediction to be less tolerant to genetic divergence between training and test data sets. Robert, Auzanneau et al. (2022) demonstrated the combined predictive ability of genotypic and NIRS data and emphasized the impact of environment on predictive ability as captured by NIRS using 2 years of winter bread wheat candidate breeding lines. In a large panel of soybean recombinant inbred lines, Zhu et al. (2021a) found NIRS-based PP ability to be comparable, and in some cases greater than, genomic prediction ability within and between segregating families. Zhu et al. (2021a) also found that NIRS data sets from distinct environments can be substituted for data sets from other environments

and applied successfully in PP. This is relevant for predicting how genotypes will perform in untested environments, paving the way for translatable applications of PP and eventual selection.

In this study, building on NIRS data from Lane et al. (2020) and genomic data from Farfan et al. (2015), phenomic (NIRS) and genomic data of maize lines evaluated as hybrids on a common tester were used as predictors either separately or together to predict GY and 500-kernel weight (KW) using a novel analysis approach. Notably, we accessed genomic data not presented in Lane et al. (2020) to directly compare genomic and PP in addition to further dissecting the phenomic data set using lasso regression to uncover autocorrelated NIRS bands. To assess prediction accuracy under relevant plant breeding scenarios, the performance of tested and untested maize hybrids was predicted in observed and unobserved environments across water-stressed (WS) and well-watered growth conditions over 2 years (four environments total) under four distinct cross-validation (CV) schemes. In summary, this research sought to (i) develop combined genomic and NIRS models capable of predicting traits such as GY and KW across environments; (ii) assess the utility of including NIR spectra in predictive models alone and alongside genomic data; (iii) explore the predictive ability of a subset of NIRS bands nominated/selected by lasso variable importance scores for predicting GY and KW; and (iv) examine the performance of multi-kernel genomic and PP models versus single-kernel models.

## 2 | MATERIALS AND METHODS

### 2.1 | Experimental design

Field trials were conducted at the Texas A&M AgriLife Experiment Station in Burleson County, TX, during the summers of 2011 and 2012, as described by Farfan et al. (2015). Briefly, a subset of 346 maize lines from the 302-line USDA Goodman maize association panel (Flint-Garcia et al., 2005) and the 300-line diverse southern subtropical Williams/Warburton panel (Warburton et al., 2013) were testcrossed to two variants of Tx714 (Betrán et al., 2004). These two variants of Tx714, isogenic lines with lipoxigenase 4 and 5 knockouts, were not significantly different for any measured trait and were therefore combined throughout analysis. Testcross hybrids were grown in randomized complete block designs with two replications in each year/management combination, consisting of well-watered (WW) and WS management conditions in 2011 and 2012. These will hereafter be denoted as CS11\_WS, CS11\_WW, CS12\_WS, and CS12\_WW, indicating four different growing environments that contained 196, 263, 310, and 213 hybrids, respectively (Figure S1). A common set of 146 hybrids formed the basis of the train/test split that enabled sparse genomic and PP.

### Core Ideas

- Near-infrared spectroscopy (NIRS) phenomic and genomic prediction had similar accuracies for grain yield and 500-kernel weight.
- Models with only genomic/phenomic data often outperformed multi-kernel models.
- Lasso regression removed correlated NIRS bands with limited reduction in prediction ability.

### 2.2 | Agronomic traits

GY and 500-kernel weight (KW) were recorded on a plot basis, and values are reported in t/ha and mass in grams of 500 kernels, respectively.

### 2.3 | Phenomic data: Near-infrared spectroscopy

High-throughput, whole-kernel NIRS scans were obtained with a Thermo Scientific Antaris II Fourier transformed interferometer (Thermo Fisher Scientific) using calibrations reported in Meng et al. (2015) and Christman (2017). NIRS data from Lane et al. (2020) from grain harvested from the testcross hybrids developed by Farfan et al. (2015) were used. In brief, whole grain samples of 196 (CS11\_WS), 269 (CS11\_WW), 310 (CS12\_WS), and 213 (CS12\_WW) hybrids were scanned in a rotating cup using the interferometer. Scanning and spectral processing methodologies are described in Meng et al. (2015), and initial analyses and quality control parameters are discussed in Lane et al. (2020). As a result, reflectance values corresponding to 3112 NIRS bands were extracted from each sample. A brief overview of the scanning methodology is as follows: (a) whole grain samples in ~175 g batches were analyzed in triplicate with 128 scans per sample; (b) with the scanner set to record diffuse reflectance, scans were performed over the instrument's integrating sphere module; (c) all spectra were recorded at ambient temperature and were calculated at 4 cm<sup>-1</sup> resolution, thus encompassing 4000–10,000 cm<sup>-1</sup>; (d) preprocessing techniques were employed such as multiplicative signal correction, standard normal variate, data treatment with first and second derivative transformations, and Savitzky-Golay filtering was performed using Thermo Scientific's TQ Analyst Pro Edition Software (Meng et al., 2015).

Using the *caret* package in R (Kuhn, 2008), the lasso algorithm was run with fivefold three-replication CV using the maximum available hybrids in each environment to determine the most important NIRS bands (cm<sup>-1</sup>) associated with GY and kernel composition traits. Model performance was

assessed using  $R^2$  (DeSalvio, 2023). The lasso algorithm nominated important NIRS bands linked to GY and KW in each environment. GY and KW were set as response variables within the *lasso* R package *train* function using fivefold CV, centering and scaling, and “method = glmnet.” Hyperparameter grid search settings were left at default, with alpha and lambda set to one and zero, respectively. Variable importance scores were extracted using the *varImp* function. By selecting bands with variable importance scores exceeding zero, lasso nominated 216 NIRS bands for GY and 131 bands for KW. No significant advantages in predictive ability were observed when constructing the matrix using the combined bands from GY and KW (347 total bands, 216 from GY and 131 from KW). As a result, phenomic, genomic, and combined prediction methods were evaluated after constructing the phenomic relationship matrix with the bands from GY only.

## 2.4 | Genotypic data

Using the genotyping-by-sequencing (GBS) method outlined in Elshire et al. (2011), genome-wide single-nucleotide polymorphism (SNP) genotypic data were obtained for 213 lines from the Williams/Warburton panel from the USDA-ARS Corn Host Plant Resistance Research Unit (Mississippi State University). Genotypic information was obtained from 133 of the 282 inbred lines comprising the USDA Goodman maize panel from Panzea (Zhao et al., 2006). These GBS data were later used for genomic and combined genomic/PP models. In this study, 346 total unique hybrids were planted, which consisted of 213 and 133 lines originating from the Williams/Warburton and USDA Goodman panels, respectively. The other 149 lines from the USDA Goodman maize panel are poorly adapted to Texas and did not generate sufficient testcross seed for yield trials and phenotyping.

Quality control of genomic data from an initial array of 61,410 polymorphic loci were performed as follows: (i) heterozygote calls were set as missing data; (ii) markers with >20% missing values were removed; and (iii) markers with <0.05 minor allele frequency were removed. After quality control, 41,456 markers (according to maize genome version AGPV2) were used to construct the genomic relationship matrix. A resulting set of 329 hybrids had genomic data of sufficient quality to be used for genomic and genomic/PP analyses (Figure S1).

## 2.5 | Analysis of variance (ANOVA)

To estimate genetic and other experimental variance components, a multi-environment ANOVA, Equation (1), was applied to the raw data for GY, KW, and each NIRS band in series (3112 bands total) to obtain the best linear unbiased

predictor (BLUP) pedigree effects via the *lme4* package in R (Bates et al., 2015).

$$Y_{jklmn} = \mu + \text{Pedigree}_j + \text{Env}_k + \text{Env} \times \text{Pedigree}_{jk} + \text{Range}_l + \text{Row}_m + \text{Env} \times \text{Rep}_{nk} + \varepsilon_{jklmn} \quad (1)$$

$Y$  is a vector of length 982 (the sum of the number of hybrids in all environments) denoting each trait value for each hybrid in one of the four environments for GY, KW, and the reflectance values for each of the 3112 NIRS bands for observations of the  $j$ th maize hybrid in the  $k$ th environment,  $l$ th range,  $m$ th row, and  $n$ th replication. Here,  $\mu$  denotes the grand mean; Pedigree denotes the overall effect of the  $j$ th maize hybrid, where  $\text{Pedigree}_j \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_j^2)$ ; Env denotes the effect of the  $k$ th environment where  $\text{Env}_k \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_k^2)$ ; Env  $\times$  Pedigree denotes the effect of the  $j$ th hybrid within the  $k$ th environment where  $\text{Env} \times \text{Pedigree}_{jk} \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_{jk}^2)$ ; Range denotes the effect of  $l$ th range where  $\text{Range}_l \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_l^2)$ ; Row denotes the effect of the  $m$ th row where  $\text{Row}_m \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_m^2)$ ; Env  $\times$  Rep denotes the effect of the  $n$ th replication within the  $k$ th environment where  $\text{Env} \times \text{Rep}_{nk} \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_{nk}^2)$ ; and  $\varepsilon$  denotes the error term addressing the non-explained variability with  $\sigma_{\text{error}}^2$  as the associated variance component such that  $\varepsilon_{jklmn} \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_{\varepsilon}^2)$ .

Repeatability was calculated using Equation (2).

$$\text{Repeatability} = \frac{\sigma_{\text{Pedigree}_j}^2}{\sigma_j^2 + \frac{\sigma_{jk}^2}{k} + \frac{\sigma_{\varepsilon_{jklmn}}^2}{nk}} \quad (2)$$

Here,  $\sigma_j^2$ ,  $\sigma_{jk}^2$ , and  $\sigma_{\varepsilon_{jklmn}}^2$  are the variance components of the overall effects of pedigree, the interaction between pedigree and environment, and the error term, respectively,  $n$  is the number of replicates, and  $k$  is the number of environments. Pedigree effects (BLUPs) for GY, KW, and each NIRS band are shown in the [Supporting Information](#) (DeSalvio, 2023).

## 2.6 | Sparse prediction: ACROSS-environment genomic and PPs

For sparse prediction across the four environments, the 146 maize hybrids common to all four environments were split approximately equally into fivefold, with fourfold used for training each model. The *BGLR* package was used in R to perform predictions of both GY and KW (Pérez & de Los Campos, 2014), considering genomic and/or phenomic covariates. Models were fit under a hierarchical Bayesian framework with 5000 Gibbs sampler iterations, 20% burn-



**TABLE 1** Summary of prediction models for across-environment predictions. M1 indicates genomic prediction; M2 indicates phenomic prediction; M3 indicates phenomic prediction in which the phenomic relationship matrix was constructed with a subset of near-infrared spectroscopy (NIRS) bands nominated by lasso as important for predicting the trait of interest; M4 combines aspects of M1 and M2; and M5 combines aspects of M1 and M3.

Model	Formula	Description
M1	$Y_{jpk} = \mu + g_j + E_k + gE_{jk} + \varepsilon_{jk}$	Genomic
M2	$Y_{jpk} = \mu + q_j + E_k + qE_{jk} + \varepsilon_{jk}$	Phenomic
M3	$Y_{jpk} = \mu + q_j + E_k + qsE_{jk} + \varepsilon_{jk}$	Subset phenomic
M4	$Y_{jpk} = \mu + g_j + q_j + E_k + gE_{jk} + qE_{jk} + \varepsilon_{jk}$	Genomic + phenomic
M5	$Y_{jpk} = \mu + g_j + q_j + E_k + gE_{jk} + qsE_{jk} + \varepsilon_{jk}$	Genomic + subset phenomic

Note: The subset phenomic relationship matrix in M3 and M5 was constructed with 216 NIRS bands nominated as having high variable importance by lasso regression.

in, and the thinning parameter was set at 10 for a total of 400 iterations used to compute the posterior mode of the different parameters involved in the linear predictors. Environment effects were calculated using Bayesian ridge regression, and all other kernels were specified as Reproducing kernel Hilbert spaces (specified as “BRR” and “RKHS,” respectively, in the *BGLR* package documentation), which under this implementation corresponds to the conventional GBLUP approach.

Five genomic and phenomic best linear unbiased prediction (GBLUP/PBLUP) models, including main and interaction effects between genomic/phenomic markers and environments modeled with different covariance structures, were compared in this study (Table 1). The data preparation and matrix multiplication to obtain each kernel are described in the [Supporting Information](#) repository (DeSalvio, 2023) and in the following section. In the series of equations in Table 1,  $Y$  represents a  $982 \times 1$  vector of BLUPs from Equation (1) for GY or KW across all four environments. Each element of  $Y_{jkq}$  corresponding to the BLUP value for the  $j$ th hybrid in the  $k$ th environment ( $k = 1, \dots, 4$ ).  $\mu$  represents the grand mean, and  $g_j$ ,  $q_j$ , and  $qs_j$  denote the multi-environment genomic, phenomic, and subset phenomic relationship matrices, respectively.  $E_k$  represents the  $982 \times 4$  environment incidence matrix, and  $gE_{jk}$ ,  $pE_{jk}$ , and  $psE_{jk}$  represent the interaction effect of environment with genomic, phenomic, and subset phenomic effects, respectively;  $\varepsilon_{jkq}$  denotes the residual error addressing the non-explained variability.

**Genomic data preparation**—Here,  $g$  symbolizes the multi-environment additive genomic effect matrix, modeled via a linear combination between  $n$  additive markers and their corresponding effects based on the first method outlined by VanRaden (2008) using the *AGHmatrix* package in R (Amadeu et al., 2016). Because only one pollinator (Tx714) was used in the topcross, only an additive genomic relationship matrix was calculated. Preparation of the  $g$  matrix was as follows: (i) the complete  $329 \times 41,456$  SNP matrix of all maize hybrids was used to create separate matrices of 196, 263, 310, and 213 rows, each with 41,456 columns, corresponding to SNPs for each hybrid grown within CS11\_WS, CS11\_WW, CS12\_WS,

and CS12\_WW, respectively; (ii) each matrix was scaled and centered column-wise; (iii) these matrices were then stacked to obtain a  $982 \times 41,456$  SNP matrix denoted as  $Z_g$ ; (iv) the multi-environment  $g$  matrix was produced by matrix multiplication of  $Z_g$  with its transpose and then divided by the number of SNPs (41,456), given by:  $G = \frac{Z_g Z_g'}{41,456}$ , and  $g$  is  $982 \times 982$ . The genomic effects are considered outcomes from the normal distribution according to  $g \sim N(0, G\sigma_g^2)$ ,  $\sigma_g^2$  as the additive genetic variance. The interaction between the genotypes and the environments was modeled based on Jarquín et al. (2014) as the Hadamard product ( $\odot$ ) between their corresponding covariance structures such that  $Z_E$  represents the incidence matrix that connects phenotypic records with environments. Hence, the covariance structure of the interaction components is given by  $[Z_E Z_E'] \odot [\frac{Z_g Z_g'}{41,456}]$ , therefore  $gE \stackrel{i.i.d.}{\sim} N(0, [Z_E Z_E'] \odot [\frac{Z_g Z_g'}{41,456}] \sigma_{gE}^2)$ , where the genotype  $\times$  environment variance components are given by  $\sigma_{gE}^2$ .

**Phenomic data preparation: Entire data set**—Here,  $q$  symbolizes the multi-environment phenomic effect matrix, modeled as a linear combination of 3112 NIRS bands and their effects within each environment. A similar data preparation procedure was implemented with NIRS data: (i) NIRS data corresponding to BLUPs of reflectance values for each hybrid in each environment were prepared as separate matrices of 196, 263, 310, and 213 rows with 3112 columns; (ii) each matrix was scaled and centered column-wise; (iii) the matrices were stacked to obtain a  $982 \times 3112$  matrix denoted as  $N$ ; (iv) the multi-environment phenomic relationship matrix was obtained by multiplying  $N$  by its transpose and dividing by the number of NIRS bands, given by:  $q = \frac{N N'}{3112}$ , where  $q$  is  $982 \times 982$ . Each NIRS effect within  $q$  is assumed to be sampled from normal distribution with  $N_q \stackrel{i.i.d.}{\sim} N(0, \sigma_q^2)$ , where the NIRS variance components are given by  $\sigma_q^2$ . The interaction between the environment kernel and the NIRS effect kernel was modeled with the covariance structure given by the Hadamard product of  $E$  and  $q$ :  $EE \odot q$ , therefore  $qE \stackrel{i.i.d.}{\sim} N(0, [Z_E Z_E'] \odot [\frac{N N'}{3112}] \sigma_{qE}^2)$ .

The NIRS  $\times$  environment variance components are given by  $\sigma_{qEE}^2$ .

**Phenomic data preparation: Subset of NIRS bands determined by lasso**—Here,  $qs$  symbolizes the phenomic effects computed with the subset multi-environment phenomic matrix. Identical data processing procedures were implemented above, except that in each instance, out of the 3112 initial bands, only 216 bands were used. These correspond to those with variable importance scores greater than zero as determined by the lasso function implemented via the *caret* package in R (Kuhn, 2008). To summarize, (i) NIRS data corresponding to BLUPs of reflectance values for each hybrid in each environment were again prepared as separate matrices of 196, 263, 310, and 213 rows, but now with only the 216 NIRS bands nominated by lasso; (ii) each matrix was scaled and centered column-wise; (iii) the matrices were stacked to obtain a  $982 \times 216$  matrix denoted as  $NS$ ; (iv) the multi-environment subset phenomic relationship matrix  $qs$  was obtained by multiplying  $NS$  by its transpose and dividing by the number of NIRS bands (now 216 vs. 3112 above), given by:  $Qs = \frac{NSNS'}{216}$ , where  $Qs$  is  $982 \times 982$ . NIRS effects within  $qs$  are again assumed to be sampled from normal distributions with  $qs \sim N(0, Qs\sigma_{qs}^2)$ , and the variance components of the subset of NIRS bands are given by  $\sigma_{qs}^2$ . The interaction between the environment kernel and the subsetted NIRS effect kernel was modeled with the covariance structure given by the Hadamard product of  $E$  and  $ps$ : therefore,  $psE \sim N(0, [Z_E Z_E'] \odot [\frac{NSNS'}{216} \sigma_{psE}^2])$  or simply  $psE \stackrel{i.i.d.}{\sim} N(0, [Z_E Z_E'] \odot Qs \sigma_{psE}^2)$ . The subset NIRS  $\times$  environment variance components are given by  $\sigma_{qsE}^2$ .

## 2.7 | Cross-validation performance evaluation

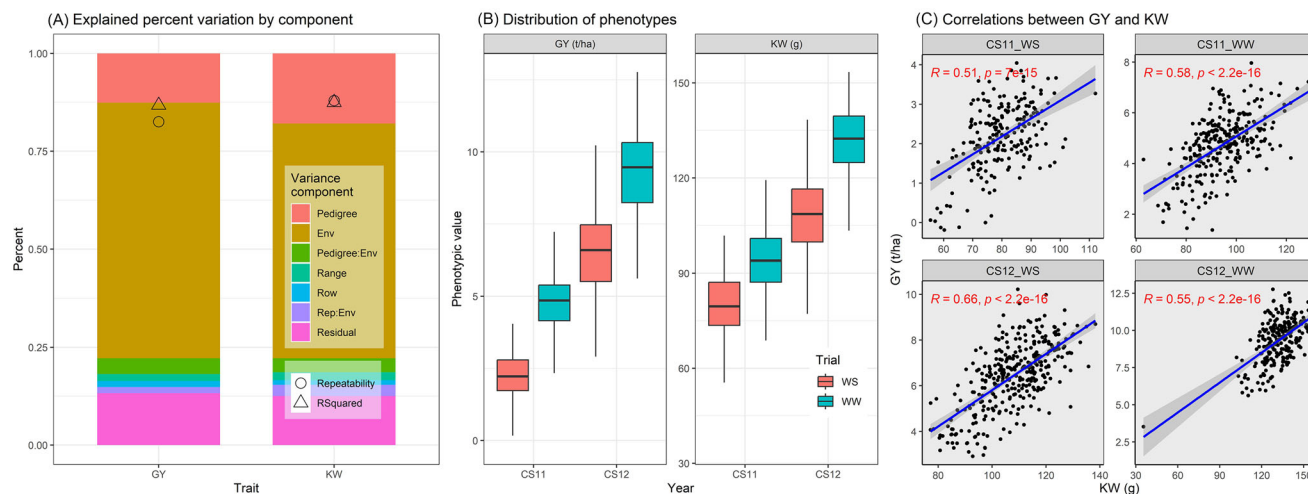
The prediction performance of each model was assessed using a fivefold 20-replication scheme. Four prediction scenarios were implemented based on the naming convention proposed by Jarquín et al. (2017), where CV2 represents prediction of tested genotypes in characterized environments, CV1 represents prediction of untested genotypes in characterized environments, CV0 represents prediction of tested genotypes in uncharacterized environments, and lastly, CV00 represents prediction of untested genotypes in uncharacterized environments. In CV1 and CV2, training data consisted of genotypes belonging to all four environments, so all environments were considered tested or characterized, whereas in CV0 and CV00, training data consisted of only three out of four environments, leaving one environment as uncharacterized or untested. CV schemes are visually described in Figure S2.

The 146 common hybrids grown in each of the four environments (Figure S1) were partitioned into fivefold,

four of which were used to train each model. In CV1 and CV2, phenotypic observations belonging to hybrids from fourfold (116 out of 146, approximately 80%) in each of the four environments were kept, while all other values were set as missing values to be predicted by the model. In this way, the training data set consisted of 464 phenotypic records ( $116 \times 4$ ) while the evaluation set consisted of 518 records. The evaluation set was composed of 30 of 146 common hybrids not allotted to the training set in all four environments plus the remaining records from each of the four environments ( $30 \times 4 + 398 = 518$ ). The same 116 hybrids in all four environments were kept as training data, meaning that they were partitioned by genotype, not by pooling records from all four environments and randomly partitioning after pooling. The prediction accuracy of CV2 was evaluated as the correlation between actual GY or KW values of the 464 training records and predicted values of the same 464 training records. This simulated predicting tested hybrids in characterized environments since the training data set had phenotypic observations from all environments. The prediction accuracy of CV1 was evaluated as the correlation between actual values of hybrids not found in the training set (518 records) and predicted values of those records, which simulated prediction of untested hybrids in characterized environments since the training data set again had phenotypic observations from all environments (Figure S2).

In CV0 and CV00, the 146 hybrids common to all environments were again split into fivefold, with GY or KW values from fourfold (116 hybrids) being kept in only three of the four environments, meaning trait values from one of the environments (CS12\_WW for the example in Figure S2) were set as missing values. In this way, the training data set is comprised strictly of records from four of fivefold (116 hybrids) from the 146 common hybrids in CS11\_WS, CS11\_WW, and CS12\_WS (only three environments instead of four previously), and predictions for CS12\_WW had to borrow information from the remaining three environments. In total, the training data size consisted of 348 phenotypic records ( $116 \times 3$ ) while the evaluation data set consisted of the remaining 634 records, amounting to 30 of 146 common hybrids in all four environments, the 146 records from the withheld environment, and the remaining 368 hybrids distributed throughout all four environments ( $116 + (30 \times 4) + 368 = 634$ ). All possible combinations of setting values from one of the four environments as missing were considered and are reported below.

The prediction accuracy of CV0 was again calculated as the correlation between actual and predicted values of the  $116 \times 4 = 464$  common hybrids in all four environments, but the model was generated using records from only three environments. In this way, the model predicted tested hybrids in uncharacterized environments. For CV00, the prediction accuracy was calculated as the correlation between actual



**FIGURE 1** (A) Results are displayed for analysis of variance (ANOVA) of grain yield (GY) and 500-kernel weight (KW) according to Equation (1), where  $Y_{jklmn} = \mu + \text{Pedigree}_j + \text{Env}_k + \text{Env} \times \text{Pedigree}_{jk} + \text{Range}_l + \text{Row}_m + \text{Env} \times \text{Rep}_{nk} + \varepsilon_{jklmn}$ . Black circles denote repeatability, and white triangles indicate  $R^2$  values. (B) Phenotypic data distributions are shown for each trait with red denoting water stress and blue denoting well-watered conditions. Dashed lines indicate the means of each distribution. (C) Correlations between GY and KW are displayed across each environment. Statistically significant positive correlations were observed in all cases. CS, College Station; WS, water-stressed; WW, well-watered.

and predicted values of the remaining 518 hybrids, but in this scheme, none of the hybrids being predicted were in the training set. Here, the model predicted untested hybrids in uncharacterized environments (Figure S2).

## 2.8 | Prediction ability

To quantify prediction ability of each CV/model/trait combination, a weighted average correlation was implemented according to Tiezzi et al. (2017). This calculation accounted for the unequal numbers of hybrids grown in each environment and ensured proper weighting of the results from environments with differing numbers of hybrids:

$$r_\varphi = \frac{\sum_{i=1}^I \frac{r_i}{V(r_i)}}{\sum_{i=1}^I \frac{1}{V(r_i)}} \quad (3)$$

where  $r_i$  is the Pearson's correlation between actual and predicted values belonging to the  $i$ th environment,  $V(r_i) = \frac{1-r_i^2}{n_i-2}$  is the sampling variance, and  $n_i$  is the number of hybrids in each environment.

## 3 | RESULTS

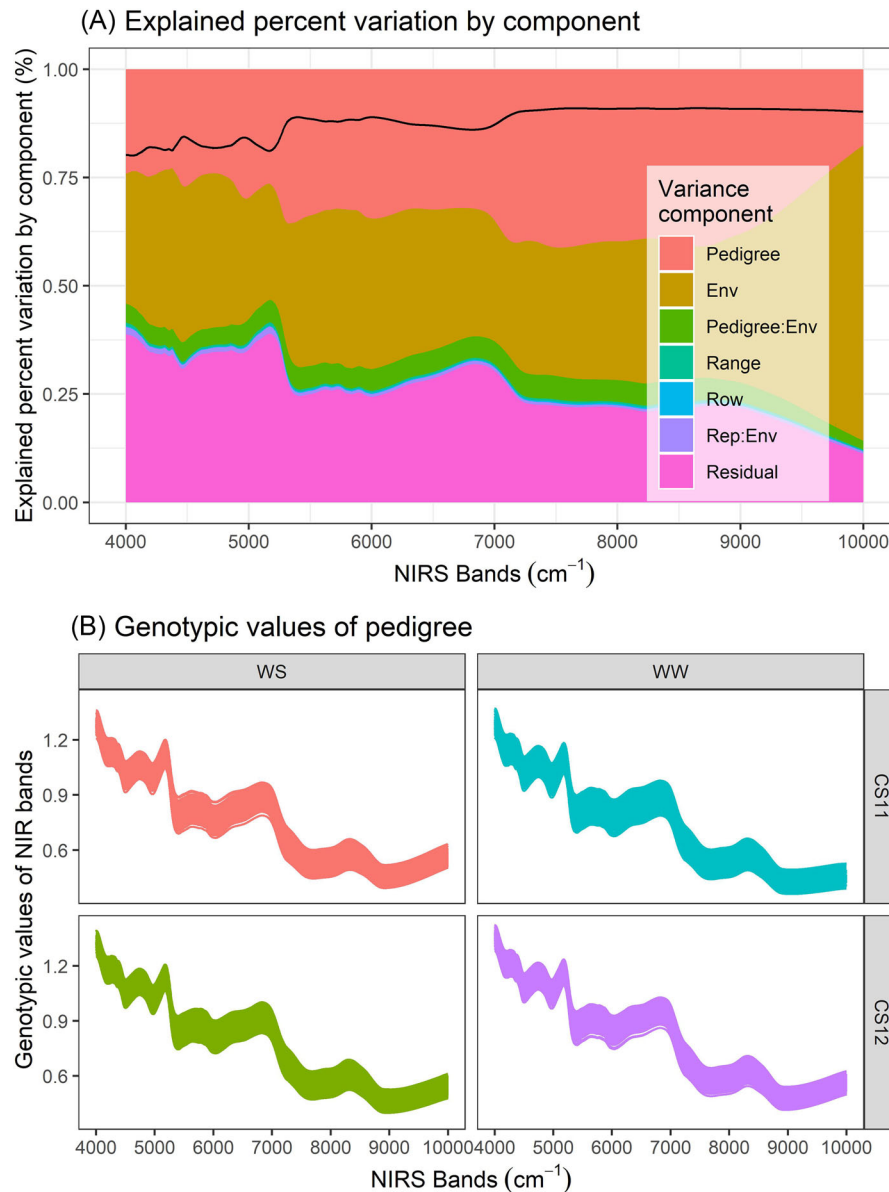
### 3.1 | Variance component analysis from multi-environment ANOVA

Results of ANOVA from Equation (1) indicated that for GY and KW, the environment component explained 65.2% and

59.9% of each trait, respectively (Figure 1A). Spatial variation (range, row, and the interaction between replication and environment) each individually explained little of the experimental variation, with the interaction between replication and environment displaying the largest effect at 7.00% for GY.  $R^2$  and repeatability reached 0.868 and 0.826 for GY, respectively, and both metrics exceeded 0.874 for KW. Pedigree explained 19.7% and 12.3% of variation for GY and KW, while the interaction between pedigree and environment accounted for 4.04% and 3.57%, respectively, indicating relatively modest genotype  $\times$  environment interaction. Distributions of pedigree BLUPs for GY and KW were stratified by environment, with the atypical drought conditions in CS11\_WS visible in the low GY and KW values for this environment, whereas CS12\_WW demonstrated the highest values for both traits (Figure 1B).

### 3.2 | Variance component analysis for NIRS bands

The results of ANOVA using Equation (1) across all 3112 NIRS wavelength bands revealed variability in percentages of explained variation in different regions of the spectrum, with pedigree accounting for as much as 41.2% at  $\sim 7532 \text{ cm}^{-1}$  and as little as 17.6% at  $\sim 9997 \text{ cm}^{-1}$  (Figure 2A). On average, pedigree variance explained a comparable percentage of experimental variation across the spectrum for GY and KW (33.0% vs. 35.1%, respectively), while the third-largest component was environmental variation (25.6%). The interaction between pedigree and environment explained <5% of



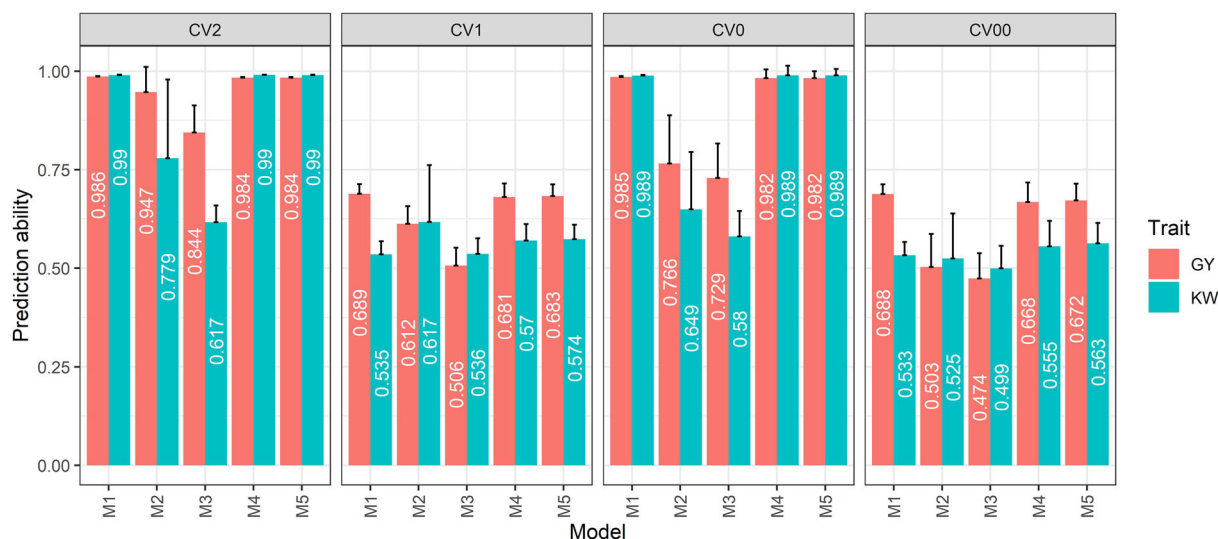
**FIGURE 2** (A) Explained percent variation across all 3112 near-infrared spectroscopy (NIRS) bands in each of the four environments as computed by Equation (1), where  $Y_{jklmn} = \mu + \text{Pedigree}_j + \text{Env}_k + \text{Env} \times \text{Pedigree}_{jk} + \text{Range}_l + \text{Row}_m + \text{Env} \times \text{Rep}_{nk} + \varepsilon_{jklmn}$ . (A) is displayed as a proportional stacked area chart. The solid black line denotes repeatability values as they change across the bands (Equation 2). (B) Genotypic values of all NIRS bands are displayed for each environment as calculated in Equation (1). CS, College Station; WS, water-stressed; WW, well-watered.

variation. Spatial effects such as range and row accounted for <5% and 3%, respectively, while the interaction between replication and environment on average accounted for 7.70%. Repeatability, as computed by Equation (2), ranged between 0.800 and 0.910 (given by the solid black line in Figure 2A). The highest overall values for repeatability were observed between ~7100 and 10,000 cm⁻¹. Phenotypic values of NIRS bands over the entire spectrum followed similar trends in each environment (Figure 2B).

### 3.3 | Assessment of prediction ability

Average prediction abilities for all models grouped by CV and trait (in order of GY and KW) were  $0.949 \pm 0.028$  and  $0.873 \pm 0.049$  for CV2,  $0.634 \pm 0.036$  and  $0.566 \pm 0.059$  for CV1,  $0.889 \pm 0.050$  and  $0.839 \pm 0.051$  for CV0, and lastly,  $0.601 \pm 0.053$  and  $0.535 \pm 0.064$  for CV00 (Figure 3). Within CV2 and CV0, in which tested hybrids are predicted in characterized and uncharacterized environments,





**FIGURE 3** Prediction abilities of each across-environment prediction model for grain yield (GY) and 500-kernel weight (KW) are displayed within each cross-validation scheme. Red indicates GY, and blue indicates KW. Model notation is described in detail in Table 1. Error bars denote standard deviations (only the top bracket is shown). M1: genomic ( $g$ ), environment ( $E$ ), plus their interaction ( $gE$ ); M2: phenomic ( $q$ ), environment ( $E$ ), plus their interaction ( $qE$ ); M3: phenomic with only near-infrared spectroscopy (NIRS) bands nominated by lasso (216 bands) ( $qs$ ), environment ( $E$ ), plus their interaction ( $qsE$ ); M4: combination of M1 and M2; M5: combination of M1 and M3. CV, cross-validation.

respectively, genomic (M1), combined genomic plus phenomic (M4), and genomic plus subset phenomic (M5) models performed nearly identically, each above  $\sim 0.98$ .

For GY prediction of untested genotypes in characterized environments (CV1), genomic prediction (M1) had the highest accuracy ( $0.689 \pm 0.024$ ), with PP trailing by 12.5% at  $0.612 \pm 0.045$ . The model incorporating the phenomic relationship matrix derived from a subset of 216 lasso nominated NIRS bands achieved a prediction ability of  $0.506 \pm 0.045$ . Multi-kernel methods (M4 and M5) did not outperform M1. Prediction of GY in CV00, where phenotypic values of untested genotypes are predicted in uncharacterized environments, was again best performed by M1 ( $0.688 \pm 0.024$ ).

Prediction of KW in CV1 was best performed by M2, though this model had a higher standard deviation of prediction than M1 ( $0.617 \pm 0.150$  vs.  $0.535 \pm 0.034$ , respectively). Neither multi-kernel model was able to surpass the performance of M2, though M4 and M5 had standard deviations of prediction that were 71.1% and 74.7% lower, respectively. In CV00, KW was predicted best by the multi-kernel genomic plus subset phenomic model (M5,  $0.563 \pm 0.052$ ).

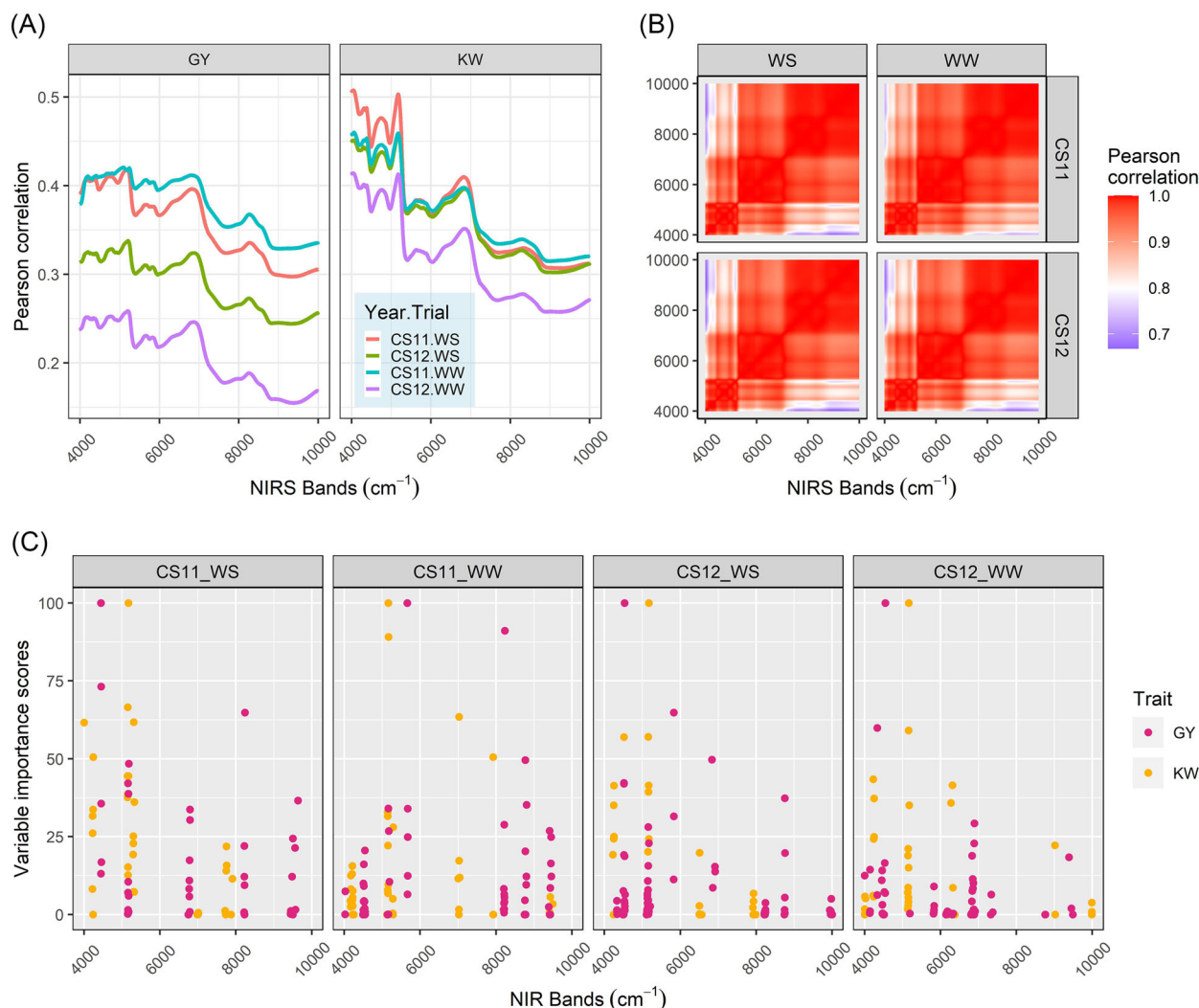
### 3.4 | Correlations of NIRS bands with GY, KW, and all pairwise bands

Similar patterns of correlation between GY and KW with the NIRS bands were observed in each of the four environments, however, spectra from CS11\_WS had the highest average cor-

relation for GY (averaging 0.378 across the entire spectrum), while CS12\_WW was markedly lower at 0.207 (Figure 4A). Correlations between NIRS bands and KW across the four environments were more tightly grouped (Figure 4A). For both GY and KW, correlation trends decreased as the wavelengths increased (Figure 4A). Correlations of the 3112 NIRS bands with each other revealed that every pairwise correlation was above 0.66 (Figure 4B), prompting dissection of the NIRS data set using lasso (Figure 4C). Three main clusters of NIRS bands were visible and consistent across four environments (e.g., cluster 1:  $\sim 4000$ – $5000$   $\text{cm}^{-1}$ ; cluster 2:  $\sim 5000$ – $7000$   $\text{cm}^{-1}$ ; and cluster 3:  $\sim 7000$ – $10,000$   $\text{cm}^{-1}$ ) (Figure 4B). Certain NIRS bands displayed consistency across environments and both traits, such as in the range of  $\sim 5000$ – $5200$   $\text{cm}^{-1}$  (Figure 4C).

## 4 | DISCUSSION

Improvements in historical crop productivity are largely attributable to traditional plant breeding methods (Evenson & Gollin, 2003) and improved agronomic practices (Rizzo et al., 2022). Recent improvements of classical techniques are due in part to the integration of genomic selection into breeding paradigms. Phenomic prediction and selection approaches may further improve this paradigm, both through increased measurement quantity leading to more accurate genetic value estimations (Lane & Murray, 2021) and through better understanding and modeling of genetic relationships based on how genotypes interact with their environments. Given the novelty



**FIGURE 4** (A) The correlations of grain yield (GY) and 500-kernel weight (KW) with each near-infrared spectroscopy (NIRS) band. Colors denote each of the four environments. (B) The correlations of the 3112 NIRS bands between each other separated by environment. (C) Variable importance scores calculated by the lasso regression are reported for GY and KW in each environment. Bands with high variable importance are displayed as points with values closer to 100. Lasso nominated 216 and 133 NIRS bands for GY and KW, respectively. Subset phenomic prediction was performed with bands nominated for GY, as these were advantages over subset prediction using KW alone or GY and KW combined. CS, College Station; NIR, Near-infrared reflectance; NIRS, near-infrared spectroscopy; WS, water-stressed; WW, well-watered.

of NIRS-based PP approaches, this technology has yet to be integrated into breeding programs at scale. The present study builds on the body of knowledge demonstrated by other studies regarding the potential value of NIRS to predict agronomic field performance (Brault et al., 2022; Rincet et al., 2018; Weiß et al., 2022), however, it highlights limitations of the technique when applied in a high-throughput manner to scans of whole unground maize grain. Variance component estimation obtained from Equation (1) along all 3112 NIRS bands demonstrated that pedigree variance accounted for between 33.0% and 35.1% of variation with moderate to high repeatability (Figure 2A), indicating that reflectance values obtained by NIRS have the potential to capture effects related

to the unique genetic backgrounds of the germplasm that were scanned.

Here, we used whole grain intact samples as opposed to finely ground, homogenized samples. This method affords much greater throughput and practicality but likely lowers prediction ability due to the shallow depth of light penetration into the sample, sample heterogeneity, and inconsistent light scattering. This may have been a critical limitation in the present study and offers a likely explanation why genomic prediction (M1) was better in some scenarios (GY prediction in CV00 and CV1) than PP with all NIRS bands (M2) (Figure 3). Spielbauer et al. (2009) hypothesized that the larger and asymmetric nature of grain can produce differing NIRS reflectance

values for whole grain depending on whether the germinal or abgerminal side is facing the instrument (Heffner et al., 2011; Orman & Schumann, 1992; Weinstock et al., 2006). A potentially insightful follow-up to this study might be to scan finely ground kernel samples with NIRS to see if the scanning methodology improves the predictive ability of the PP models (M2 and/or M3). Alternatively, to avoid the laborious process of grinding samples, hyperspectral imaging via novel methods developed by Varela et al. (2022) may allow researchers to obtain accurate estimations of kernel composition traits non-destructively, in addition to obtaining morphometric features such as width and length of the kernels. An additional limitation to the present study's findings on the use of PP includes the limited effects of  $G \times E$  for GY and KW, as the interaction effect between pedigree and environment accounted for only 4.04% and 3.57% for GY and KW, respectively (Figure 1A). This phenomenon was also observable in the ANOVA results of each NIRS band, with the average percentage explained by the interaction between pedigree and environment only reaching 4.71%. Additional  $G \times E$  may have increased the value of PP when compared with genomic prediction.

#### 4.1 | CV1 and CV00 prediction results

For plant breeding programs, CV1 (prediction of untested genotypes in tested environments) and CV00 (prediction of untested genotypes in untested environments) are the most realistic and applicable scenarios, as it is desirable to predict the performance of new genotypes in either familiar or new trial locations. For GY in CV1 and CV00, genomic prediction (M1) performed better than PP alone (M2) and similarly to genomic plus PP (M4), indicating that a trait measured with relatively high repeatability in diverse environments in this study could sufficiently be predicted by genomic data while phenomic data added little value (Figure 3). The latter point contrasts with the findings of Rincet et al. (2018), in which NIR spectra from wheat leaf tissue collected in one environment generated an estimated covariance matrix capable of accurate yield predictions in a weakly correlated separate environment (located ~500 km away). Future studies are warranted to determine whether NIRS-based PP is suitable as environments become more disparate. However, in CV1, KW was best predicted by M2 (which used the entire NIR spectrum), even surpassing combined genomic plus PP (M4). Prediction of KW by M2 in CV00 was on par with M1, differing by only 1% accuracy. The mixed performance of M2 versus M1 indicates that, despite not always being able to surpass genomic prediction, NIRS-based PP can still perform comparably and may add value to research programs as a nondestructive mechanism of spectral measurement.

Lane et al. (2020) demonstrated the utility of using a functional regression model to predict untested hybrids in two out

of four environments (predicting 20% of hybrids untested in two environments but tested in the remaining two), reporting a prediction ability of 0.87 when supplying the model with phenotypic data from WW environments while withholding information from WS. The NIRS BLUP model in the same prediction scenario fared worse, reporting a prediction ability of 0.07. The methods of this study differed from Lane et al. (2020) (namely, from the use of functional and partial least squares [PLS] regression) to test the prediction ability of NIR spectra when treated in a similar manner to genomic marker relationship matrices. The variable selection imposed by PLS in Lane et al. (2020) may have handled the autocorrelation observed when employing the entire spectrum for prediction, and to test this, lasso was used to filter through autocorrelated bands before creating a phenomic relationship matrix based only on important NIRS wavelengths (M3). In CV1 and CV00, M3 performed similarly to M2, indicating that the majority of the NIR spectrum was redundant. Combined genomic and phenomic (M4) and combined genomic and subsetted phenomic (M5) prediction models did not surpass genomic prediction alone for GY, suggesting that at least in this study's germplasm, genomic and phenomic data sets captured similar yield-related features and did not lead to synergistic effects when data modalities were combined. With KW, however, modest improvements in prediction ability were observed with M4 and M5 performance in CV1 and CV00 compared with phenomic or genomic prediction alone (M1/M2/M3), indicating a trait dependency for the use cases of NIRS relationship matrices in predicting agronomic traits.

#### 4.2 | Applicability of genomic and phenomic data for plant breeding

As an alternative or supplement to genomic prediction, PP uses spectral signals instead of molecular markers. As reflectance spectra are representative of tissue biochemical properties, which are genetically controlled, relationship matrices derived from plant spectra can capture genetic signals (Brault et al., 2022) but also likely environmental and genome  $\times$  environmental signals. The advent of affordable remote sensing technologies has ushered phenomics into an era of low-cost and higher throughput data collection, in which spectral data sets (RGB, multispectral, or hyperspectral) enable calculation of vegetation indices that can parameterize predictive models that inform selection decisions (Adak et al., 2021; Galán et al., 2020; Hernandez et al., 2015; Montesinos-López et al., 2017; Rutkoski et al., 2016). Because these approaches and demonstrations are new compared to more extensively investigated genomic approaches, case studies are needed to build knowledge sufficient to improve phenomic methods. Brault et al. (2022) demonstrated that similarities existed between NIRS and genomic

relationship matrices across tissues or years in *Vitis vinifera* (grapevine). These authors, like many others reported to date, found that despite modest genetic variance estimated from NIRS, PP was comparable to that of GP.

In this study, genomic prediction (M1) frequently surpassed the predictive ability of phenomic (M2), subset phenomic (M3), or combined genomic/PP (M4/M5), indicating continued merit in sparse prediction for multi-environment trials. While there were scenarios (KW in CV1 and CV00) in which inclusion of NIR spectra with genomic relationship matrices improved prediction abilities, scans of intact maize kernels may not surpass the performance of genomic prediction due to the nature of asymmetric nutrient deposition during grain filling as discussed previously (Figure 3). However, PP was frequently able to achieve prediction abilities on par with genomic prediction. In this study, samples were scanned nondestructively and there were no laboratory consumables besides electricity for the scanner required to obtain the spectral data. As such, NIRS may (after an initial investment in the spectrometer) serve as an economical and rapid-turnaround alternative to genomic data while potentially providing simultaneous estimates of kernel composition traits such as phenolic antioxidants (Meng et al., 2015).

A further advantage of phenomic measurement technologies such as NIRS is the ability to record environment-specific and genetic by environment-specific metrics with the intent of capturing phenotypic plasticity of the same genotypes grown in disparate environments (Zhu et al., 2021a). However, this advantage is less pronounced with NIRS, as scans must be conducted after the germplasm being studied has been harvested. Given the sensitivity of genomic prediction and selection to training/testing set relatedness (Brauner et al., 2020; Olatoye et al., 2020; Riedelsheimer et al., 2012; Weiß et al., 2022; Zhu et al., 2021b), PP has shown merit in predicting performance of diverse breeding material with varying degrees of relatedness (Weiß et al., 2022). Zhu et al. (2021b) and Brauner et al. (2020) reported decreasing genomic prediction ability as the training set composition moved from full-sib to half-sib to unrelated families in *Glycine max* (L.) Merr. and maize, respectively. The diversity of the hybrid panel in the present study may also have played a role in the ability for models to predict across environments and serves as a key distinction between the elite wheat varieties and association mapping panel of poplar reported by Rincet et al. (2018). Though NIRS-based PP alone underperformed genomic prediction in this study for GY, there may still be advantages to phenomic approaches for plant breeders owing to its rapid turnaround time and fixed costs versus the consumable nature of preparing samples for genotyping. This will be increasingly true where NIRS pipelines are more automated, including NIRS integrated on combine harvesters, though these too will likely suffer from uneven nutrient deposition if applied to maize kernels. As a result, phenomic approaches may enable

increased selection accuracy with existing population sizes or selection intensity if they allow more lines and hybrids to be screened. Furthermore, NIRS phenomic-based approaches could provide additional value in correcting errors in combine yield data or predicting other unmeasurable traits (e.g., flowering time) not directly estimable from either DNA or conventional grain composition measurements.

It is important to note that using NIRS on hybrid F<sub>1</sub> seed before planting would likely be inappropriate in the case of hybrid maize, where the kernel would represent the inbred mother plant and not the genetic potential of the hybrid that will emerge from the seed. In contrast, using the harvested grain, assuming no xenia effect (which is reasonable in this study, Bulant et al. [2000]), will result in estimations of yield from the plants this grain was selected from. In the case of mechanically (e.g., combine) harvested yield trials, this is unlikely to provide any substantial speed advantages regarding mid-growing season selection decisions, but the enhancement in prediction ability afforded by NIRS data could reduce the amount of screening required in subsequent seasons because genomic/PP results can be implemented for either positive or negative selection, especially in untested environments. Though this study examined the predictive ability of spectra obtained from harvested hybrid F<sub>2</sub> grain, future investigation of predicting testcross performance from NIRS data obtained from inbred grain is warranted and would have direct relevance for breeding programs. This concept has been explored for genomic prediction of hybrid maize GY using doubled haploid lines (Albrecht et al., 2011) and hybrid sorghum yield [*Sorghum bicolor* (L.) Moench] using elite inbred lines (Fonseca et al., 2021), but has received little attention from the phenomics research community. Demonstration of comparable prediction accuracies for predicting hybrid performance using phenomic data acquired for inbreds (i.e., NIRS, vegetation indices from unoccupied aerial system [UAS]) could provide breeders with low-cost alternatives to genotyping inbred parents.

Prior studies have demonstrated an array of statistical treatments for NIRS in PS since most models applicable to genomic selection can be adapted to phenomic selection (PS) (Robert, Brault et al., 2022). Functional regression, in which observations are functions outlined on a continuous domain, has also been implemented in PS for yield prediction, leading to greater parsimony among regression models and higher computational efficiency (Lane et al., 2020; Montesinos-López et al., 2018, 2017; Morris, 2015). Variable selection techniques such as lasso and BayesB can reduce multicollinearity and high dimensionality characteristic of NIRS data, where large numbers of wavelengths supersede the number of observations (Lane et al., 2020; Robert, Brault et al., 2022). Variable selection in PS can improve interpretability of findings, as certain predictors (in the case of this research, NIRS bands) could show higher heritability than others



(Ferragina et al., 2015; Galán et al., 2020). However, Aguete et al. (2017) and Galán et al. (2021) both reported losses in prediction accuracies when less than the full wavelength spectrum (hyperspectral data in these cases) was used, and minor decreases in prediction ability in M3 versus M2 were also observed in the present study (Figure 3). Assessment of variable importance in this study among NIRS bands for GY (Figure 4C) via the lasso regression highlighted wavelengths such as  $\sim 5000\text{--}5200\text{ cm}^{-1}$  and indicated a large portion of the spectrum was autocorrelated. As noted by Zhu et al. (2021a), determining a smaller group of predictive bands creates the potential for production of lower-cost spectrometers that omit non-predictive or repetitive bands.

### 4.3 | Future directions

The push to explore the predictive nature of agricultural phenomic data has led to the adoption of UASs (or drones) to capture temporal, also called longitudinal, phenomic data at scale, which enables the elucidation of genotypic variation across a growing season (Adak et al., 2021) and potentially well before harvest. In contrast, genomic information and NIRS of grain samples represent single measurements, which is a primary limitation of the present investigation. Future studies are warranted that take advantage of NIRS, RGB, multispectral, and even hyperspectral relationship matrices in attempting to perform prediction of agronomically relevant traits using novel statistical methods to handle large numbers of predictors throughout growth.

## 5 | CONCLUSIONS

Overall, genomic-only prediction was more suited to across-environment prediction of GY of untested genotypes in characterized and uncharacterized environments, while combined genomic/PP performed best for prediction of 500-kernel weight. NIR spectra were able to delineate between pedigree, environment, spatial, and interactions between pedigree and environment, although these results indicated that high-throughput scans of intact maize kernels may not be able to surpass genomic prediction when used in a relationship matrix prediction context. However, the nondestructive nature of NIRS whole-kernel scans, coupled with its comparable prediction results to genomic prediction, indicates potential value within research programs for selection. Further study of integrating phenomic and genomic data is necessary to enable more accurate and efficient prediction models, reduced need for field-based phenotyping, and an accelerated breeding process. Phenomic relationship matrices using NIRS bands nominated by lasso enabled comparable prediction abilities to relationship matrices based on the entire NIR spectrum, indicating that, as with genomic markers, high degrees of

autocorrelation were present in the phenomic data set. Future studies of combining NIRS and genomic data have the potential to overcome some of the limitations of each individual technique, providing a more comprehensive understanding of the genetic and physiological factors that contribute to both agronomically important and research-specific traits in maize such as GY and KW, which is especially fitting for breeding programs targeting genetic gains and developing locally adapted germplasm.

### AUTHOR CONTRIBUTIONS

**Aaron J. DeSalvio:** Conceptualization; data curation; formal analysis; investigation; methodology; visualization; writing—original draft; writing—review and editing. **Alper Adak:** Conceptualization; data curation; formal analysis; investigation; methodology; visualization; writing—original draft; writing—review and editing. **Seth C. Murray:** Conceptualization; data curation; funding acquisition; project administration; supervision; writing—review and editing. **Diego Jarquin:** Formal analysis; methodology; writing—review and editing. **Noah D. Winans:** Formal analysis; methodology; writing—review and editing. **Daniel Crozier:** Formal analysis; methodology; writing—review and editing. **William L. Rooney:** Conceptualization; methodology; writing—review and editing.

### ACKNOWLEDGMENTS

The authors would like to acknowledge Dr. Ivan Barrero-Farfan and Holly Lane for their efforts in collecting the data upon which this study was based. We would also like to acknowledge the graduate and undergraduate students and high school employees of the Texas A&M Quantitative Genetics and Maize Breeding program for their dedication to maintaining fields and collecting phenotypic data. Financial support for this research has been provided by USDA–NIFA–AFRI Award Nos. 2020-68013-32371 and 2021-67013-33915, USDA–NIFA Hatch funds, Texas A&M AgriLife Research, the Texas Corn Producers Board, and the Eugene Butler Endowed Chair. Aaron J. DeSalvio was supported by the Texas A&M University/Association of Former Students (TAMU/AFS) Graduate Merit Fellowship and the National Science Foundation (NSF) Graduate Research Fellowship (GRFP), and Alper Adak was supported by a fellowship from the Republic of Turkey, Ministry of National Education and Ministry of Agriculture and Forestry.

### CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interests.

### DATA AVAILABILITY STATEMENT

An annotated script of the R code used in this research can be accessed via GitHub (DeSalvio, 2023) [<https://github.com/ajdesalvio/Maize-NIRS-GBS.git>]. All files needed to

reproduce the results are provided in the GitHub repository for users to access.

## ORCID

Aaron J. DeSalvio  <https://orcid.org/0000-0003-1818-4699>

Alper Adak  <https://orcid.org/0000-0002-2737-8041>

Seth C. Murray  <https://orcid.org/0000-0002-2960-8226>

Diego Jarquín  <https://orcid.org/0000-0002-5098-2060>

Noah D. Winans  <https://orcid.org/0000-0003-4185-5665>

Daniel Crozier  <https://orcid.org/0000-0002-3101-235X>

William L. Rooney  <https://orcid.org/0000-0001-7953-1856>

## REFERENCES

- Adak, A., Conrad, C., Chen, Y., Wilde, S. C., Murray, S. C., Anderson, S., & Subramanian, N. K. (2021). Validation of functional polymorphisms affecting maize plant height by unoccupied aerial systems (UAS) discovers novel temporal phenotypes. *G3 Genes|Genomes|Genetics*, 11(6), jkab075. <https://doi.org/10.1093/g3journal/jkab075>
- Aguate, F. M., Trachsel, S., Pérez, L. G., Burgueño, J., Crossa, J., Balzarini, M., Gouache, D., Bogard, M., & de los Campos, G. (2017). Use of hyperspectral image data outperforms vegetation indices in prediction of maize yield. *Crop Science*, 57(5), 2517–2524. <https://doi.org/10.2135/cropsci2017.01.0007>
- Albrecht, T., Wimmer, V., Auinger, H.-J., Erbe, M., Knaak, C., Ouzunova, M., Simianer, H., & Schön, C.-C. (2011). Genome-based prediction of testcross values in maize. *Theoretical and Applied Genetics*, 123, 339–350. <https://doi.org/10.1007/s00122-011-1587-7>
- Amadeu, R. R., Cellon, C., Olmstead, J. W., Garcia, A. A. F., Resende, M. F. R., Jr., & Muñoz, P. R. (2016). AGHmatrix: R package to construct relationship matrices for autotetraploid and diploid species: A blueberry example. *The Plant Genome*, 9(3), plantgenome2016.2001.0009. <https://doi.org/10.3835/plantgenome2016.01.0009>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Betrán, F. J., Bockholt, A., Fojt Iii, F., Mayfield, K., & Pietsch, D. (2004). Registration of Tx714 maize germplasm line. *Crop Science*, 44(3), 1028–1029.
- Brault, C., Lazerges, J., Doligez, A., Thomas, M., Ecartot, M., Roumet, P., Bertrand, Y., Berger, G., Pons, T., & François, P. (2022). Interest of phenomic prediction as an alternative to genomic prediction in grapevine. *Plant Methods*, 18(1), Article 108. <https://doi.org/10.1186/s13007-022-00940-9>
- Brauner, P. C., Müller, D., Molenaar, W. S., & Melchinger, A. E. (2020). Genomic prediction with multiple biparental families. *Theoretical and Applied Genetics*, 133(1), 133–147. <https://doi.org/10.1007/s00122-019-03445-7>
- Bulant, C., Gallais, A., Matthys-Rochon, E., & Prioul, J. (2000). Xenia effects in maize with normal endosperm: II. Kernel growth and enzyme activities during grain filling. *Crop Science*, 40(1), 182–189. <https://doi.org/10.2135/cropsci2000.401182x>
- Christman, J. L. (2017). *Development of near infrared reflectance spectroscopy (NIRS) calibrations of maize kernel phosphorus for the identification of useful breeding material* [Master's thesis, Texas A&M University].
- DeSalvio, A. J. (2023). *Supplementary data—Near infrared reflectance spectroscopy phenomic and genomic prediction of maize agronomic and composition traits across environments*. <https://github.com/ajdesalvio/Maize-NIRS-GBS>
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6(5), e19379. <https://doi.org/10.1371/journal.pone.0019379>
- Evenson, R. E., & Gollin, D. (2003). Assessing the impact of the green revolution, 1960 to 2000. *Science*, 300(5620), 758–762. <https://doi.org/10.1126/science.1078710>
- Farfan, I. D. B., De La Fuente, G. N., Murray, S. C., Isakeit, T., Huang, P.-C., Warburton, M., Williams, P., Windham, G. L., & Kolomiets, M. (2015). Genome wide association study for drought, aflatoxin resistance, and important agronomic traits of maize hybrids in the subtropics. *PLoS One*, 10(2), e0117737. <https://doi.org/10.1371/journal.pone.0117737>
- Ferragina, A., de Los Campos, G., Vazquez, A. I., Cecchinato, A., & Bittante, G. (2015). Bayesian regression models outperform partial least squares methods for predicting milk components and technological properties using infrared spectral data. *Journal of Dairy Science*, 98(11), 8133–8151. <https://doi.org/10.3168/jds.2014-9143>
- Ferrio, J. P., Bertran, E., Nachit, M. M., Catala, J., & Araus, J. L. (2004). Estimation of grain yield by near-infrared reflectance spectroscopy in durum wheat. *Euphytica*, 137(3), 373–380. <https://doi.org/10.1023/B:EUPH.0000040523.52707.1e>
- 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. *The Plant Journal*, 44(6), 1054–1064. <https://doi.org/10.1111/j.1365-3113X.2005.02591.x>
- Fonseca, J. M. O., Klein, P. E., Crossa, J., Pacheco, A., Perez-Rodriguez, P., Ramasamy, P., Klein, R., & Rooney, W. L. (2021). Assessing combining abilities, genomic data, and genotype × environment interactions to predict hybrid grain sorghum performance. *The Plant Genome*, 14(3), e20127. <https://doi.org/10.1002/tpg2.20127>
- Galán, R. J., Bernal-Vasquez, A.-M., Jebsen, C., Piepho, H.-P., Thorwarth, P., Steffan, P., Gordillo, A., & Miedaner, T. (2020). Integration of genotypic, hyperspectral, and phenotypic data to improve biomass yield prediction in hybrid rye. *Theoretical and Applied Genetics*, 133(11), 3001–3015. <https://doi.org/10.1007/s00122-020-03651-8>
- Galán, R. J., Bernal-Vasquez, A.-M., Jebsen, C., Piepho, H.-P., Thorwarth, P., Steffan, P., Gordillo, A., & Miedaner, T. (2021). Early prediction of biomass in hybrid rye based on hyperspectral data surpasses genomic predictability in less-related breeding material. *Theoretical and Applied Genetics*, 134(5), 1409–1422. <https://doi.org/10.1007/s00122-021-03779-1>
- Heffner, E. L., Jannink, J. L., & Sorrells, M. E. (2011). Genomic selection accuracy using multifamily prediction models in a wheat breeding program. *The Plant Genome*, 4(1), 65–75. <https://doi.org/10.3835/plantgenome2010.12.0029>
- Hernandez, J., Lobos, G. A., Matus, I., Del Pozo, A., Silva, P., & Galleguillos, M. (2015). Using ridge regression models to estimate grain yield from field spectral data in bread wheat (*Triticum aestivum* L.) grown under three water regimes. *Remote Sensing*, 7(2), 2109–2126. <https://doi.org/10.3390/rs70202109>

- Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., Piraux, F., Guerreiro, L., Pérez, P., & Calus, M. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical and Applied Genetics*, 127, 595–607. <https://doi.org/10.1007/s00122-013-2243-1>
- Jarquín, D., Lemes da Silva, C., Gaynor, R. C., Poland, J., Fritz, A., Howard, R., Battenfield, S., & Crossa, J. (2017). Increasing genomic-enabled prediction accuracy by modeling genotype  $\times$  environment interactions in Kansas wheat. *The Plant Genome*, 10(2), plantgenome2016–2012. <https://doi.org/10.3835/plantgenome2016.12.0130>
- Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(1), 1–26.
- Lane, H. M., & Murray, S. C. (2021). High throughput can produce better decisions than high accuracy when phenotyping plant populations. *Crop Science*, 61(5), 3301–3313. <https://doi.org/10.1002/csc2.20514>
- Lane, H. M., Murray, S. C., Montesinos-López, O. A., Montesinos-López, A., Crossa, J., Rooney, D. K., Barrero-Farfan, I. D., De La Fuente, G. N., & Morgan, C. L. S. (2020). Phenomic selection and prediction of maize grain yield from near-infrared reflectance spectroscopy of kernels. *The Plant Phenome Journal*, 3(1), e20002. <https://doi.org/10.1002/ppj2.20002>
- Meng, Q., Murray, S. C., Mahan, A., Collison, A., Yang, L., & Awika, J. (2015). Rapid estimation of phenolic content in colored maize by near-infrared reflectance spectroscopy and its use in breeding. *Crop Science*, 55(5), 2234–2243. <https://doi.org/10.2135/cropsci2014.11.0767>
- Montesinos-López, A., Montesinos-López, O. A., Gianola, D., Crossa, J., & Hernández-Suárez, C. M. (2018). Multi-environment genomic prediction of plant traits using deep learners with dense architecture. *G3 Genes| Genomes| Genetics*, 8(12), 3813–3828. <https://doi.org/10.1534/g3.118.200740>
- Montesinos-López, O. A., Montesinos-López, A., Crossa, J., De Los Campos, G., Alvarado, G., Suchismita, M., Rutkoski, J., González-Pérez, L., & Burgueño, J. (2017). Predicting grain yield using canopy hyperspectral reflectance in wheat breeding data. *Plant Methods*, 13(1), 1–23. <https://doi.org/10.1186/s13007-016-0154-2>
- Morris, J. S. (2015). Functional regression. *Annual Review of Statistics and Its Application*, 2, 321–359. <https://doi.org/10.1146/annurev-statistics-010814-020413>
- Olatoye, M. O., Clark, L. V., Labonte, N. R., Dong, H., Dwiyantri, M. S., Anzoua, K. G., Brummer, J. E., Ghimire, B. K., Dzyubenko, E., & Dzyubenko, N. (2020). Training population optimization for genomic selection in miscanthus. *G3 Genes| Genomes| Genetics*, 10(7), 2465–2476. <https://doi.org/10.1534/g3.120.401402>
- Orman, B. A., & Schumann, R. A. (1992). Nondestructive single-kernel oil determination of maize by near-infrared transmission spectroscopy. *Journal of the American Oil Chemists' Society*, 69, 1036–1038. <https://doi.org/10.1007/BF02541073>
- Pasquini, C. (2003). Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, 14, 198–219. <https://doi.org/10.1590/S0103-50532003000200006>
- Pasquini, C. (2018). Near infrared spectroscopy: A mature analytical technique with new perspectives—A review. *Analytica Chimica Acta*, 1026, 8–36. <https://doi.org/10.1016/j.aca.2018.04.004>
- Pérez, P., & de Los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics*, 198(2), 483–495. <https://doi.org/10.1534/genetics.114.164442>
- Polytec GmbH. (2022). *Seed and harvester applications*. <https://www.polytec.com/eu/process-analytics/areas-of-application/seed-harvester-applications>
- Riedelsheimer, C., Czedik-Eysenberg, A., Grieder, C., Lisec, J., Technow, F., Sulpice, R., Altmann, T., Stitt, M., Willmitzer, L., & Melchinger, A. E. (2012). Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nature Genetics*, 44(2), 217–220. <https://doi.org/10.1038/ng.1033>
- Rincint, R., Charpentier, J.-P., Faivre-Rampant, P., Paux, E., Le Gouis, J., Bastien, C., & Segura, V. (2018). Phenomic selection is a low-cost and high-throughput method based on indirect predictions: Proof of concept on wheat and poplar. *G3 Genes| Genomes| Genetics*, 8(12), 3961–3972. <https://doi.org/10.1534/g3.118.200760>
- Rizzo, G., Monzon, J. P., Tenorio, F. A., Howard, R., Cassman, K. G., & Grassini, P. (2022). Climate and agronomy, not genetics, underpin recent maize yield gains in favorable environments. *Proceedings of the National Academy of Sciences*, 119(4), e2113629119. <https://doi.org/10.1073/pnas.2113629119>
- Robert, P., Auzanneau, J., Goudemand, E., Oury, F.-X., Rolland, B., Heumez, E., Bouchet, S., Le Gouis, J., & Rincint, R. (2022). Phenomic selection in wheat breeding: Identification and optimisation of factors influencing prediction accuracy and comparison to genomic selection. *Theoretical and Applied Genetics*, 135(3), 895–914.
- Robert, P., Brault, C., Rincint, R., & Segura, V. (2022). Phenomic selection: A new and efficient alternative to genomic selection. In N. Ahmadi & J. Bartholomé (Eds.), *Genomic prediction of complex traits: Methods and protocols* (pp. 397–420). Springer.
- Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Pérez, L. G., Crossa, J., Reynolds, M., & Singh, R. (2016). Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. *G3 Genes| Genomes| Genetics*, 6(9), 2799–2808. <https://doi.org/10.1534/g3.116.032888>
- Spielbauer, G., Armstrong, P., Baier, J. W., Allen, W. B., Richardson, K., Shen, B., & Settles, A. M. (2009). High-throughput near-infrared reflectance spectroscopy for predicting quantitative and qualitative composition phenotypes of individual maize kernels. *Cereal Chemistry*, 86(5), 556–564. <https://doi.org/10.1094/CCEM-86-5-0556>
- Tiezzi, F., de Los Campos, G., Gaddis, K. L. P., & Maltecca, C. (2017). Genotype by environment (climate) interaction improves genomic prediction for production traits in US Holstein cattle. *Journal of Dairy Science*, 100(3), 2042–2056. <https://doi.org/10.3168/jds.2016-11543>
- 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>
- Varela, J. I., Miller, N. D., Infante, V., Kaeppler, S. M., de Leon, N., & Spalding, E. P. (2022). A novel high-throughput hyperspectral scanner and analytical methods for predicting maize kernel composition and physical traits. *Food Chemistry*, 391, 133264. <https://doi.org/10.1016/j.foodchem.2022.133264>
- Warburton, M. L., Williams, W. P., Windham, G. L., Murray, S. C., Xu, W., Hawkins, L. K., & Duran, J. F. (2013). Phenotypic and genetic characterization of a maize association mapping panel developed for the identification of new sources of resistance to *Aspergillus flavus* and aflatoxin accumulation. *Crop Science*, 53(6), 2374–2383. <https://doi.org/10.2135/cropsci2012.10.0616>

- Weinstock, B. A., Janni, J., Hagen, L., & Wright, S. (2006). Prediction of oil and oleic acid concentrations in individual corn (*Zea mays* L.) kernels using near-infrared reflectance hyperspectral imaging and multivariate analysis. *Applied Spectroscopy*, 60(1), 9–16. <https://doi.org/10.1366/000370206775382631>
- Weiβ, T. M., Zhu, X., Leiser, W. L., Li, D., Liu, W., Schipprack, W., Melchinger, A. E., Hahn, V., & Würschum, T. (2022). Unraveling the potential of phenomic selection within and among diverse breeding material of maize (*Zea mays* L.). *G3 Genes| Genomes| Genetics*, 12(3), jkab445. <https://doi.org/10.1093/g3journal/jkab445>
- Yamada, Y., Itoh, Y., & Sugimoto, I. (1988). Parametric relationships between genotype x environment interaction and genetic correlation when two environments are involved. *Theoretical and Applied Genetics*, 76(6), 850–854. <https://doi.org/10.1007/BF00273671>
- Zhao, W., Canaran, P., Jurkuta, R., Fulton, T., Glaubitz, J., Buckler, E., Doebley, J., Gaut, B., Goodman, M., & Holland, J. (2006). Panzea: A database and resource for molecular and functional diversity in the maize genome. *Nucleic Acids Research*, 34(Suppl\_1), D752–D757. <https://doi.org/10.1093/nar/gkj011>
- Zhu, X., Leiser, W. L., Hahn, V., & Würschum, T. (2021a). Phenomic selection is competitive with genomic selection for breeding of complex traits. *The Plant Phenome Journal*, 4(1), e20027. <https://doi.org/10.1002/ppj2.20027>

- Zhu, X., Leiser, W. L., Hahn, V., & Würschum, T. (2021b). Training set design in genomic prediction with multiple biparental families. *The Plant Genome*, 14(3), e20124. <https://doi.org/10.1002/tpg2.20124>

## 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:** DeSalvio, A. J., Adak, A., Murray, S. C., Jarquín, D., Winans, N. D., Crozier, D., & Rooney, W. L. (2024). Near-infrared reflectance spectroscopy phenomic prediction can perform similarly to genomic prediction of maize agronomic traits across environments. *The Plant Genome*, 17, e20454. <https://doi.org/10.1002/tpg2.20454>