Mega-Environment Analysis To Assess Adaptability, Stability, And Genomic Predictions In Grain Sorghum Hybrids

Jales Mendes Oliveira Fonseca (mailto:jalesmof@hotmail.com)
Texas A&M University  https://orcid.org/0000-0002-1195-6193

Ramasamy Perumal
Kansas State University

Patricia E. Klein
Texas A&M University

Robert R. Klein
USDA-ARS Southern Plains Agricultural Research Center

William L. Rooney
Texas A&M University

Research Article

Keywords: Sorghum breeding, Elite germplasm, Multi-environment trials, Mega-environment, genomic prediction, Stability

Posted Date: December 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1157295/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Mega-environment analysis to assess adaptability, stability, and genomic predictions in grain sorghum hybrids

Jales M. O. Fonseca¹, Ramasamy Perumal² Patricia E. Klein³, Robert R. Klein⁴, William L. Rooney¹

¹Department of Soil & Crop Sciences, Texas A&M University, College Station, TX; 77843, USA
²Kansas State University, Agriculture Research Center, Hays, KS, USA; 67601
³Department of Horticultural Sciences, Texas A&M University, College Station, TX; 77843, USA
⁴USDA-ARS Southern Plains Agricultural Research Center, College Station, TX 77845, USA

Received:_________

Correspondence: Jales Mendes Oliveira Fonseca and William L. Rooney, Department of Soil & Crop Sciences, Texas A&M University, College Station, TX; 77843, USA.

Email: jalesmof@hotmail.com, jfonseca@tamu.edu, wlr@tamu.edu

https://orcid.org/0000-0002-1195-6193 - Jales M. O. Fonseca
https://orcid.org/0000-0002-0649-8853 - Ramasamy Perumal
https://orcid.org/0000-0002-5704-1052 - Patricia E. Klein
https://orcid.org/0000-0001-6421-5770 – Robert R. Klein
https://orcid.org/0000-0001-7953-1856 - William L. Rooney
Abstract

Multi-environment trials (MET) are fundamental for assessing genotype-by-environment interaction (GxE) effects, adaptability and stability of genotypes and provide valuable information about target regions. As such, a MET involving grain sorghum hybrid combinations derived from elite inbred lines adapted to diverse sorghum production regions was developed to assess agronomic performance, stability, and genomic-enabled prediction accuracies within mega-environments (ME). Ten females and ten males from the Texas A&M and Kansas State sorghum breeding programs were crossed following a factorial mating scheme to generate 100 hybrids. Grain yield, plant height, and days to anthesis were assessed in a MET consisting of ten environments across Texas and Kansas over two years. Genotype plus Genotype-by-block-of-environment biplot (GGB) assessed ME, while the "mean-vs-stability" view of the biplot and the Bayesian Finlay-Wilkinson regression evaluated hybrid adaptability and stability. A genomic prediction model including the GxE effect was applied within ME to assess prediction accuracy. Results suggest that grain sorghum hybrid combinations involving lines adapted to different target regions can produce superior hybrids. GGB confirmed distinct regions of sorghum adaption in the U.S. Further, genomic predictions within ME reported inconsistent results, suggesting that additional effects rather than the correlations between environments are influencing genomic prediction of grain sorghum hybrids.

Keywords: Sorghum breeding, Elite germplasm, Multi-environment trials, Mega-environment, genomic prediction, Stability
1. Introduction

Sorghum \( \textit{Sorghum bicolor} \) (L.) Moench is the fifth most important cereal crop in the world, following corn \( \textit{Zea mays} \) L., wheat \( \textit{Triticum aestivum} \) L., rice \( \textit{Oryza sativa} \) L., and barley \( \textit{Hordeum vulgare} \) L. with 57.8 million tons of global annual production in 2019 (FAOSTAT 2021). While Africa is responsible for almost 50% of the total grain sorghum production, grain yields are significantly higher in developed countries. For example, the average sorghum yield in 2019 was 4.58- vs. 1.23-tons ha\(^{-1}\) in the U.S. and Nigeria, respectively (FAOSTAT 2021). Within the U.S. states, Kansas and Texas account for 82% of total U.S. production, with average yields of 5.72- and 5.38-tons ha\(^{-1}\), respectively (NASS 2020).

Grain sorghum production is predominantly cultivated in agricultural areas where water is limited (Monk et al. 2014). Sorghum is favored in these regions because the species possesses various drought-tolerance mechanisms, including stay-green and pre-flowering drought-tolerance (Rooney 2016). Drought tolerance in sorghum provides greater yield stability compared to other crops such as maize \( \textit{Zea mays} \) that are more susceptible to water-stress conditions. Because sorghum production targets stress-prone environments, developing improved grain sorghum genotypes requires special considerations. For instance, sorghum hybrids that are highly responsive to farm inputs are best grown in non-limiting environments, but these do not reflect the typical sorghum production conditions. Hence, assessing the stability of grain sorghum hybrids across target environments becomes particularly important for the effective deployment of sorghum hybrids.

Multi-environment trials (MET) are commonly used to assess genotype-by-environment interaction (GxE) effects. When MET detect GxE that do not involve crossover interaction (COI), GxE may be ignored since the best genotypes are consistent across environments. However, GxE is often associated with COI, and judicious hybrid deployment requires breeders to investigate the adaptability and stability of specific genotypes (Crossa et
al. 2004; Burgueño et al. 2008). Many stability tests contemplate univariate measures of
stability (Wricke 1962; Finlay and Wilkinson 1963; Shukla 1972), while others have
proposed the use of multivariate measures based on biplot technique (Gabriel 1971)
i.e., additive main effects and multiplicative interaction, AMMI, Gauch and Zobel, 1997;
genotype + GxE, GGE, Yan et al., 2000), and nonparametric measures (Nassar and Huhn
1987; Huehn 1990). Depending on the test, genotypic stability can be classified as static or
dynamic (Becker 1981; Becker and Leon 1988).

As MET consist of a continuous series of testing genotypes across target
environments, not all genotypes are evaluated in each environment. Consequently, MET are
usually unbalanced, which challenges the ability to assess the stability of a genotype. A
common practice to circumvent unbalanced data and assess stability involves removing
genotypes missing in some environments and performing stability tests via classical models,
assuming genotype and environment effects as fixed (Finlay and Wilkinson 1963; Gauch
1988; Yan et al. 2000). A suitable alternative to manage unbalanced data involves applying
linear mixed models that accommodate random effects and include (co)variance structure to
model genetic and environmental effects.

Among the models that assume effects as random, factor analytics models are
commonly applied to MET (Piepho 1994; Smith et al. 2001, 2005; Piepho et al. 2008). Linear
mixed models permit better prediction of those missing data by incorporating the genetic
relationship matrix of tested and untested genotypes (Jarquin et al. 2020). Many studies have
reported on the benefits of using linear mixed models to assess MET (Piepho 1997, 1998;
Smith et al. 2001, 2019; Resende and Thompson 2004; Piepho and Möhring 2005; Burgueño
et al. 2008; Balestre et al. 2010; Hu 2015; Jarquin et al. 2020). Bayesian models have also
demonstrated their advantages for analyzing genotypic data (Sorensen and Gianola, 2007;
Crossa et al., 2010; Alves et al., 2019; Fonseca et al., 2021a). For instance, Lian and de los
Campos (2016) reported the stability of wheat varieties using a novel approach of the Finlay and Wilkinson stability test under a Bayesian framework. The stability of grain sorghum hybrids based on the Bayesian framework is lacking.

In addition to the unbalanced genotypic effect in the MET, the number of genotypes and environments vary depending on the stage (i.e., early vs. advance cycles) during the breeding process. Conventionally, many genotypes are tested in limited environments during initial stages of selection, whereas advanced stages evaluate a few genotypes across many environments. Although this approach is still generally applicable, breeding pipelines are being modified to accommodate predictive breeding technologies (Crossa et al. 2021). For instance, genomic-enabled prediction is most effective if it replaces the first phase of hybrid trials in grain sorghum breeding schemes (Fonseca et al., 2021a). Further, MET permit the development of genomic-enabled prediction models to include GxE, and results indicate better prediction accuracy (Burgueño et al. 2012; Lopez-Cruz et al. 2015; Lado et al. 2016; Cuevas et al. 2016; Acosta-Pech et al. 2017; Basnet et al. 2019; Costa-Neto et al. 2020).

Although MET primarily focus on identifying superior genotypes for target environments, established breeding programs also apply MET to generate relevant information about tested environments (Gauch and Zobel, 1997; Yan et al., 2000; Yan and Kang, 2002; Laffont et al., 2013; Gauch, 2013; Yan, 2015, 2016). This analysis can be useful to identify mega-environments (ME) across a wider production area (Rakshit et al. 2012; Nielsen and Vigil 2018; Dalló et al. 2019; Sharma et al. 2020; Ansarifard et al. 2020). By definition, ME reduce GxE and allow breeders to extract performance information about genotypes more effectively within a ME (González-Barrios et al. 2019). Lado et al. (2016) indicated that wheat line performance was more predictable when ME information is considered. Alves et al. (2021) showed that the prediction accuracy of hybrids in weakly-
correlated environments can be increased when including additional phenotypic information from other correlated trials.

Lopez-Cruz et al. (2015) presented a marker-by-environment genomic best linear unbiased prediction (MxE GBLUP) model to account for GxE effects in genomic-enabled prediction models. The MxE model borrows genetic information across environments and thereby increases prediction accuracies for untested genotypes in tested environments.

Acosta-Pech et al. (2017) further expanded the MxE GBLUP model to include combining abilities, environment, and combining abilities-by-environment interactions and reported higher prediction accuracies in maize. Fonseca et al. (2021a) recently applied the Acosta-Pech approach to sorghum and prediction accuracies were promising. Most genomic-enabled prediction models that include GxE effects assume the GxE (co)variance matrix to be positive semi-definite (Cuevas et al. 2016; Crossa et al. 2019). Thus, higher correlations between environments are expected to yield higher prediction accuracies. The combination of the Acosta-Pech model under ME may generate good prediction accuracies as ME are expected to increase the correlation between environments, causing a reduction in the residual variance, and potentially capturing greater genetic variation.

A common breeding approach for multinational commercial breeding programs is to cross elite germplasm adapted to diverse target environments (Cooper et al., 2014). Studies have demonstrated the potential advantages of utilizing this strategy in hybrid crops (Podlich and Cooper, 1998; Technow et al., 2020). This strategy is less common in public institutions (Fonseca et al. 2021a), because these programs have more limited target area with fewer genetic resources. Nevertheless, the exchange of germplasm between distinct breeding programs could foster a collaborative system between public institutions and mitigate the decrease in number and funding of public breeding programs (Shelton and Tracy, 2017; Coe et al., 2020; Fonseca et al., 2021b).
The present research aimed to develop ME for grain sorghum production regions and assess the adaptability and stability of grain sorghum hybrids derived from geographically distinct U.S. public breeding programs within ME. Further, genomic predictions within ME were investigated. Finally, a comparison between genomic predictions based on ME and without ME is presented.

2. Material and Methods

2.1. Genetic material

The grain sorghum hybrids utilized in this study are as previously described by Fonseca et al. (2021a). In brief, 100 hybrids were generated from crossing ten female with ten male lines in a factorial mating scheme (Comstock and Robinson 1952). The elite inbreds originated from the sorghum breeding programs at Texas A&M AgriLife Research (College Station, Texas) and Kansas State Research Center (Hays, Kansas). Each program provided five females (A-lines) and five males (R-lines) developed for their respective target environments and that produced agronomically acceptable hybrids (Table 1). The 100 hybrids were subdivided into four groups to represent hybrid combinations derived from elite lines within each breeding programs and between breeding programs (Figure 1).

2.2. Experimental data

Hybrids were planted in a randomized complete block design (RCBD) with sets in replicate adjustment; each set was composed of one of the four groups of 25 hybrids (Figure 1), with randomization occurring within and between sets. In 2018 trials, each location had three replicates, while in 2019 each location had two replicates. A plot consisted of two adjacent rows, approximately 5.3 m in length, with row spacing that ranged from 0.76 to 1.0 m, depending on the production practices in each environment. In 2018, trials were grown in Monte Alto, Victoria and College Station in Texas, and Garden City and Colby in Kansas. In 2019, the trials were in Taft, Victoria and College Station in Texas and Hays and Colby in Kansas. These Texas and Kansas locations represent distinct adaptation zones and, in each
test, agronomic practices standard to the location were followed (Table 2) (Fonseca et al., 2021a). GPS coordinates of each environment were used to collect weather data from the NASA POWER database (NASA 2021).

Three agronomic traits were measured including days to anthesis (DA), plant height (PHT), and grain yield (GY). DA were recorded as the number of days from planting to when 50% of the plants in the plot were at mid-anthesis. In Kansas locations, DA was collected on one replicate only. Just prior to harvest, PHT was measured in cm using a representative plant in each plot at the length from the soil to the tip of the panicle. Plots were harvested using plot combines fitted with a Harvest Master GrainGage System (Juniper Systems), to measure grain weight and moisture content. After adjusting yield to a 14% moisture, GY was converted to tons ha⁻¹ (Fonseca et al., 2021a).

2.3. Statistical analysis
For each environment, data were analyzed following a linear mixed model (Henderson 1984):

\[ y = \mathbf{1}\mu + \mathbf{Z}_1\mathbf{h} + \mathbf{Z}_2\mathbf{r} + \mathbf{e}, \] (1)

where \( y \) is the vector of phenotypes; \( \mu \) is the intercept; \( \mathbf{h} \) is the random effect of hybrid, \( \mathbf{h} \sim N(0, \sigma_h^2 \mathbf{I}) \), \( \mathbf{r} \) is the random effect of replicates, \( \mathbf{r} \sim N(0, \sigma_r^2 \mathbf{I}) \); \( \mathbf{e} \) is the vector of residuals, \( \mathbf{e} \sim N(0, \sigma_e^2 \mathbf{I}) \); \( \mathbf{1} \) is a vector of ones; \( \mathbf{Z}_1 \) and \( \mathbf{Z}_2 \) are incidence matrixes; \( \sigma_h^2 \), \( \sigma_r^2 \) and \( \sigma_e^2 \) are variance components for hybrids, replicates and residuals, respectively. Model (1) was extended to incorporate environment and GxE effects for performing a combined analysis as followed:

\[ y = \mathbf{1}\mu + \mathbf{Z}_1\mathbf{h} + \mathbf{Z}_2\mathbf{s} + \mathbf{Z}_3\mathbf{hs} + \mathbf{Z}_4\mathbf{r(s)} + \mathbf{e}, \] (2)
where \( \mathbf{s} \) is the vector of environmental effects, \( \mathbf{s} \sim N(\mathbf{0}, \sigma_s^2 \mathbf{I}) \); \( \mathbf{h} \) is the vector of the GxE effect \( \mathbf{h} \sim N(\mathbf{0}, \sigma_{hs}^2 \mathbf{I}) \), \( \mathbf{r} \) is the vector of replicate effect nested within environment, \( \mathbf{r} \sim N(\mathbf{0}, \sigma_{r(s)}^2 \mathbf{I}) \); \( \mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3, \mathbf{Z}_4 \), are incidence matrixes for the corresponding effect.

Variance components were estimated via restricted maximum likelihood (REML) method (Patterson and Thompson 1971) using the lmer function of the lme4 R package (Bates et al. 2015), and its significance assessed by the likelihood ratio test (LRT) using the ranova function of the lmerTest R package (Kuznetsova et al. 2017). All analysis were conducted in R software (R Core Team 2020). From the variance component estimates, broad-sense heritability at each environment was calculated as \( R = \frac{\sigma_h^2}{\sigma_h^2 + \sigma_s^2} \), whereas, for combined environments, \( H^2 = \frac{\sigma_h^2}{\sigma_h^2 + \sigma_{hs}^2 + \sigma_{re}^2} \) where \( t \) is the number of environments and \( r \) is the number of replicates. Best linear unbiased estimates (BLUEs) of each hybrid at each location were obtained using the R package emmeans (Lenth 2021) from model (1) assuming hybrids as fixed. BLUEs were used to develop the training set in genomic prediction models.

Mega-environments (ME) were developed using the gge function of the gge R package (Wright and Laffont 2020). Genotype plus genotype × environment (GGE) biplot and which-won-where polygon were used to assess grain sorghum hybrid performance, while genotype plus genotype × block of environments (GGB) biplots were used to define ME. Biplots were scaled and environment-centered. For GGE biplot, stability was assessed using loading projections on the average environment coordinate (AEC) at each ME. For Bayesian Finlay-Wilkinson regression (BFW - Lian and De Los Campos, 2016), static and dynamic stability was assessed based on the slope of genotypes across environment. Regression lines presenting slopes close to 0 classified static stability while slopes close to 1 defined dynamic stability. For BFW, genomic information of hybrids was included in the model as (co)variance matrix of the genotypic effect. Posterior distribution was calculated
based on 10,000 interactions, burn-in of 5,000 and a thin of 5. Genomic relationship of
hybrids were developed \textit{in silico} by the Kronecker product of the parental marker matrix. For
details see Fonseca et al. (2021a).

2.4. Genotyping, Genomic prediction model, and cross validation scheme

DNA extraction, genotyping-by-sequencing (GBS), and development of molecular
markers were described in Fonseca et al. (2021a). The genomic-enabled prediction model
used in this study was first introduced by Acosta-Pech et al. (2017) and first applied in grain
sorghum hybrids by Fonseca et al. (2021a). Herein, the model was applied for each ME
separately and included combining abilities, environment and GxE effects to predict the
performance of sorghum hybrids as follows:

\[
y = Z_s \beta + Z_1 f + Z_2 m + Z_3 h + u_f + u_m + u_h + e, 
\]

where \( y = [y_1, \ldots, y_j, \ldots, y_s]' \) is the vector of observations of hybrids (\( i = 1, \ldots, n_j \)) in the
\( j \)th environment (\( j = 1, \ldots, s \)); \( \beta \) is a vector of the environmental effects, \( \beta \sim N(0, \sigma_\beta^2 I) \), \( f \) is
the vector of GCA effects for females, \( f \sim N(0, \sigma_f^2 G_f) \), where \( j \) represents environment; \( m \)
is a vector of GCA effects for males, \( m \sim N(0, \sigma_m^2 G_m) \); \( h \) is a vector of SCA effects for
hybrid combinations, \( h \sim N(0, \sigma_h^2 H) \); \( u_f \) is a random vector to include the interaction
between GCA of females and environment, \( u_f \sim N(0, \sigma_{f_s}^2 V_f) \); \( u_m \) is a random vector that
models the interaction between males and environment, \( u_m \sim N(0, \sigma_{m_s}^2 V_m) \); \( u_h \) is a vector
that takes into account the interaction between hybrids and environment, \( u_h \sim N(0, \sigma_{h_s}^2 V_h) \);
\( e \) is a vector of residuals, \( e \sim N(0, \sigma_e^2 I) \); \( Z_s, Z_1, Z_2, \) and \( Z_3 \) are incidence matrixes for
environment, female, male, and hybrid, respectively. \( G_f, G_m, \) and \( H \) are the genomic
relationship of matrix for females, males, and hybrids, respectively. The elements of matrix
\( H \) were calculated \textit{in silico} by the Kronecker product of \( G_f \) and \( G_m \) as \( H = G_f \otimes G_m \)
(Technow et al., 2014; Acosta-Pech et al., 2017). \( V_f, V_m, \) and \( V_H \) are variance-covariance
matrices that associate each respective genomic-main-effect-by-environment variance component, i.e., $\sigma_{fS}^2$, $\sigma_{mS}^2$, and $\sigma_{hS}^2$. The variance-covariance matrix was calculated as $V_f = Z_1G_fZ'_1#Z_sZ'_s$, $V_m = Z_mG_mZ'_2#Z_sZ'_s$, and $V_{f_i} = Z_3HZ'_3#Z_sZ'_s$ where # stands for the Hadamard product [see Acosta-Pech et al. (2017) for further details about the derivations].

To validate models and estimate the predictability of parents, a simple leave-one-out cross-validation (CV) scheme was applied following Fonseca et al. (2021a). Briefly, seed parents in hybrid combinations were removed, and models trained with the remaining records where that seed parent was absent. The same process was applied for males. Thus, the CV scheme led to prediction of hybrids when a female or a male parent were removed. The prediction accuracy ($r$) at each ME was estimated as the Pearson’s correlation between observed and predicted values of hybrids.

Genomic prediction model was fit using Bayesian methods implemented in the BGLR statistical package (Pérez and de los Campos 2014) available for R. Inferences were based on 10,000 Gibbs sampler iterations with a burn-in of 5,000, and a thin of 5.

3. Results and discussion

3.1. Analysis by location

Genetic variation was significant for all traits in all environments (Table 3). Broad-sense heritability estimates were generally high, indicating that most of the variation within a location was associated with genetic effects. GY ranged from 4.31 to 8.34 ton ha$^{-1}$, PHT ranged from 121 to 145 cm, and DA ranged from 59 to 80 days across all environments. The lowest environmental mean GY occurred in 2018 Monte Alto (18RF), TX, while the greatest was in 2019 Colby (19COL), KS (Table 3). In 2020, the National Agricultural Statistics Service reported grain sorghum average yields of 4.23 and 5.71 ton ha$^{-1}$ for Texas (NASS 2021a) and Kansas (NASS 2021b), respectively. Such grain sorghum averages indicate that
hybrids presented in this study generate yields similar to those of commercial hybrids, thus meeting the expectation of high-yielding germplasm.

3.2. Combined analysis and development of mega-environments

The likelihood ratio test (LRT) for the combined analysis detected highly significant effects due to genetics, environments, and GxE for all traits (Table 4). Environment effects explained most of the variation for GY (48.9%) and DA (76.7%), while genetic components were the largest for PHT (40.5%). For GY, the GxE variance component accounted for almost 12% of the total variation, exceeding the proportion of the genetic variance (9.8%). Yan et al. (2000) suggested the existence of mega environments when GxE effects explain proportionally more than genotype effects; thus, the likelihood of finding a single best sorghum hybrid for all test environments is unlikely. Hence, mega-environment 1 (ME1) and mega-environment 2 (ME2) were developed (Table 4).

The GGE biplot analysis for GY explained more than 50% of the variation with the first two principal components (PC) (Figure 2). The which-won-where polygon suggested that three ME exist. Nonetheless, the relative lack of representativeness of the environment 19HAY combined with empirical knowledge about sorghum production regions led to the establishment of only two ME (Figure 3 and Table 4). These two ME are consistent with previous grain sorghum studies that define production regions as subtropical and temperate environments (Monk et al., 2014). Mega-environment 1 included Texas locations (i.e., Monte Alto, Taft, Victoria, and College Station), while mega-environment 2 described Kansas locations (i.e., Hays, Colby, and Garden City) (Figure 3; Table 5).

3.3. Adaptability and stability of grain sorghum hybrids derived from public breeding programs

Several grain sorghum hybrids from both programs were highest yielding and highly stable (Figures 4, 5; Table 5). For instance, within the subtropical environment (i.e., ME1), the highest yielding hybrid was a Kansas/Texas combination (KF1/TM3) (Figure 4). More
important, the KF1/TM3 hybrid exhibited good stability across environments (Figure 4). According to Finlay-Wilkinson (1963), linear regressions of genotype yield parallel to the population mean yield regression categorize ‘dynamic stability’. Further, an “ideal genotype” outperforms other genotypes in both low- and high-yielding environments. Based on the BFW regression method applied herein, the KF1/TM3 hybrid exhibits dynamic stability and represents a potential “ideal genotype” (Figure 4). For instance, the KF1/TM3 hybrid was among the highest yielding in the 18RF and 19CS environments (Figure 4). Therefore, KF1/TM3 exemplifies a crucial characteristic that determines high-performing and stable grain sorghum hybrids: hybrids that are adapted to marginal areas and responsive to more favorable environments (Monk et al., 2014).

Within the temperate ME (ME2), the top two performers resulted from crosses involving Texas inbred lines (TF2/TM2 and TF2/TM1) (Figure 6). That implies that elite lines developed outside a particular target environment can produce good grain sorghum hybrids in other sub-tropical ME. Within U.S. sorghum improvement programs it is generally easier to move germplasm from lower to higher latitudes (i.e., from South to North in the U.S.) than in the opposite direction. Results presented herein support such a statement as Texas elite sorghum lines performed well in hybrid combination in temperate Kansas environments.

It is interesting to note that the top performers within ME1 exhibited above-average PHT and below-average DA across environments (Figure 5). This emphasizes the positive correlation between PHT and GY, and further highlights the relationship between earlier hybrids and their avoidance of biotic and abiotic stresses (i.e., midge damage and post-anthesis drought tolerance). However, for ME2, the top hybrids were shorter than the average high across environments (Figure 7). In ME2, taller plants might be adversely affected by abiotic stress associated with this ME, such as wind and/or drought that can lead to lodging.
For those environments, shorter plants are likely to suffer less from lodging and stalk breakage problems, which could explain the yield advantage and justify their selection for these environments.

3.4. **Genomic predictions modeling GxE within ME**

Genomic predictions of grain sorghum hybrid performance within ME produced inconsistent results (Table 6). When compared to the prediction without ME, the leave-one-out female CV scheme had higher prediction accuracies within ME1, while the leave-one-out male CV scheme had better predictions in ME2 for GY and PHT. Overall, there was a slight increase in predicting GY and DA within ME1. These results indicate that closely related environments can but do not necessarily increase prediction accuracies (Cuevas et al. 2016; Crossa et al. 2019).

The accuracy of genomic prediction models is influenced by many effects, including but not limited to the quality of the phenotypic data, the heritability of the trait, and the size and structure of the training population. For example, predictions across all environments consider the entire data set, predictions within a ME only include information from a specific ME. While the ME division can increase the correlations between environments located within the ME, the advantages for predicting untested genotypes maybe limited due to the concomitant reduction in the number of hybrids present in the training set. As such, developing prediction models for specific ME is recommended when the target environment is well-defined and sufficient evaluated to build robust prediction models (Table 6). It is also possible to increase the prediction accuracy by including phenotypic records derived from addition sets of correlated environments (Alves et al., 2021). This is observed when prediction accuracies are compared between ME2 and without ME (Table 6). In situations where predictions of untested hybrids involve new environments, reaction norm principles and envirotypic-assisted genomic prediction models indicate promising results as both
genomic and environmental information are included in the model (Jarquín et al. 2014; Li et al. 2018; Costa-Neto et al. 2020; Buntaran et al. 2021).

4. Conclusion

The identification of ME reduces GxE effects, provides better information about MET, and continues to be a common practice among plant breeders interested in both genotype and environment assessment. A MET for grain sorghum hybrids grown across Texas and Kansas confirmed the existence of two distinct ME that coincide with historically different regions of adaptation generally designated as subtropical and temperate environments. Hybrid combinations between elite inbred lines from Texas and Kansas public sorghum breeding programs indicate a potential benefit of exchanging geographically-distinct germplasm to generate high-performing and stable grain sorghum hybrids across temperate and subtropical environments. Based on the present results, additional public (and commercial) sorghum breeding programs are encouraged to explore elite germplasm exchanges to increase genetic diversity within and across programs with an ultimate goal of enhancing genetic gain and hybrid performance. This is especially critical in crops in which investment in public breeding programs is anticipated to continue to decline and thereby exploit the shared resources of multiple programs. Genomic enabled prediction models were inconsistent in their ability to predict hybrids performance within ME, yet highly correlated environments can generate greater prediction accuracies with smaller training sets. Further research is warranted to identify additional factors affecting the accuracy of genomic prediction models within ME grain sorghum hybrids trials. Finally, the inclusion of non-genetic effects to model the correlation among environments should be considered in future studies to assess their merit to increase prediction accuracies in MET.

5. Acknowledgements

Funds from the Borlaug-Monsanto Chair in Plant Breeding were used to support this work.
6. Author contributions

Jales M O Fonseca: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Visualization; Writing-original draft; Writing-review & editing.

Perumal Ramasamy: Conceptualization; Investigation; Resources; Writing-review & editing.

Patricia E Klein: Investigation; Writing-review & editing. Robert Klein: Investigation; Writing-review & editing. William L Rooney: Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Writing-review & editing.

7. Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

8. Compliance with ethical standards

9. Conflict of interest

The authors declare that there is no conflict of interest.

10. References

Acosta-Pech R, Crossa J, de los Campos G, et al (2017) Genomic models with genotype × environment interaction for predicting hybrid performance: an application in maize hybrids. Theor Appl Genet 130:1431–1440. https://doi.org/10.1007/s00122-017-2898-0

Alves FC, Galli G, Matias FI, et al (2021) Impact of the complexity of genotype by environment and dominance modeling on the predictive accuracy of maize hybrids in multi-environment prediction models. Euphytica 217:. https://doi.org/10.1007/s10681-021-02779-y

Alves FC, Granato ÍSC, Galli G, et al (2019) Bayesian analysis and prediction of hybrid performance. Plant Methods 15:14. https://doi.org/10.1186/s13007-019-0388-x

Ansarifard I, Mostafavi K, Khosroshahli M, et al (2020) A study on genotype–environment interaction based on GGE biplot graphical method in sunflower genotypes (Helianthus annuus L.). Food Sci Nutr 8:3327–3334. https://doi.org/10.1002/fsn3.1610

Balestre M, Santos VB dos, Soares AA, Reis MS (2010) Stability and adaptability of upland rice genotypes. Crop Breed Appl Biotechnol 10:357–363. https://doi.org/10.1590/S1984-70332010000400011

Basnet BR, Crossa J, Dreisigacker S, et al (2019) Hybrid wheat prediction using genomic, pedigree, and environmental covariables interaction models. Plant Genome 12:1–13. https://doi.org/10.3835/plantgenome2018.07.0051

Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:. https://doi.org/10.18637/jss.v067.i01

Becker HC (1981) Correlations among some statistical measures of phenotypic stability. Euphytica 30:835–840. https://doi.org/10.1007/BF00038812
17

Becker HC, Leon J (1988) Stability Analysis in Plant Breeding. Plant Breed 101:1–23. https://doi.org/10.1111/j.1439-0523.1988.tb00261.x

Buntaran H, Forkman J, Piepho H-P (2021) Projecting results of zoned multi-environment trials to new locations using environmental covariates with random coefficient models: accuracy and precision. Theor Appl Genet 1:3. https://doi.org/10.1007/s00122-021-03786-2

Burgueño J, Crossa J, Cornelius PL, Yang RC (2008) Using factor analytic models for joining environments and genotypes without crossover genotype × environment interaction. Crop Sci 48:1291–1305. https://doi.org/10.2135/cropsci2007.11.0632

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

Coe MT, Evans KM, Gasic K, Main D (2020) Plant breeding capacity in U.S. public institutions. Crop Sci 60:2373–2385. https://doi.org/10.1002/csc2.20227

Comstock RE, Robinson HF (1952) Estimation of average dominance of genes. Heterosis 2:494–516

Cooper M, Messina CD, Podlich D, et al (2014) Predicting the future of plant breeding: Complementing empirical evaluation with genetic prediction. Crop Pasture Sci 65:311–336. https://doi.org/10.1071/CP14007

Costa-Neto G, Fritsche-Neto R, Crossa J (2020) Nonlinear kernels, dominance, and envirotyping data increase the accuracy of genome-based prediction in multi-environment trials. Heredity (Edinb). https://doi.org/10.1038/s41437-020-00353-1

Crossa J, Campos G de los, Pérez P, et al (2010) Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics 186:713–724. https://doi.org/10.1534/genetics.110.118521

Crossa J, Fritsche-Neto R, Montesinos-Lopez OA, et al (2021) The Modern Plant Breeding Triangle: Optimizing the Use of Genomics, Phenomics, and Enviromics Data. Front Plant Sci 12:1–6. https://doi.org/10.3389/fpls.2021.651480

Crossa J, Martini JWR, Gianola D, et al (2019) Deep Kernel and Deep Learning for Genome-Based Prediction of Single Traits in Multi-environment Breeding Trials. Front Genet 10:1–13. https://doi.org/10.3389/fgen.2019.01168

Crossa J, Yang R-C, Cornelius PL (2004) Studying crossover genotype × environment interaction using linear-bilinear models and mixed models. J Agric Biol Environ Stat 9:362–380. https://doi.org/10.1198/108571104X4423

Cuevas J, Crossa J, Soberanis V, et al (2016) Genomic Prediction of Genotype × Environment Interaction Kernel Regression Models. Plant Genome 9:1–20. https://doi.org/10.3835/plantgenome2016.03.0024

Dalló SC, Zdziarski AD, Woyann LG, et al (2019) Across year and year-by-year GGE biplot analysis to evaluate soybean performance and stability in multi-environment trials.

17
Euphytica 215:113. https://doi.org/10.1007/s10681-019-2438-x

FAOSTAT 2021. Food and Agriculture Organization of the United Nations. FAOSTAT statistical database. Available online at: http://www.fao.org/faostat/en/#data/QC/visualize

Finlay K, Wilkinson G (1963) The analysis of adaptation in a plant-breeding programme. Aust J Agric Res 14:742. https://doi.org/10.1071/AR9630742

Fonseca J, Perumal R, Klein PE, et al (2021a) Combining abilities and elite germplasm enhancement across US public sorghum breeding programs. Crop Sci scs2.20624. https://doi.org/10.1002/csc2.20624

Fonseca JMO, Klein PE, Crossa J, et al (2021b) Assessing combining abilities, genomic data, and genotype × environment interactions to predict hybrid grain sorghum performance. Plant Genome. https://doi.org/10.1002/pgen.202127

Gabriel KR (1971) The biplot graphic display of matrices with application to principal component analysis. Biometrika 58:453–467. https://doi.org/10.1093/biomet/58.3.453

Gauch HG (1988) Model Selection and Validation for Yield Trials with Interaction. Biometrics 44:705. https://doi.org/10.2307/2531585

Gauch HG, Zobel RW (1997) Identifying Mega-Environments and Targeting Genotypes. Crop Sci 37:311–326. https://doi.org/10.2135/cropsci1997.0011183X003700020002x

Gauch Jr HG (2013) A simple protocol for AMMI analysis of yield trials. Crop Sci 53:1860–1869. https://doi.org/10.2135/cropsci2013.04.0241

González-Barrios P, Díaz-García L, Gutiérrez L (2019) Mega-environmental design: Using genotype × environment interaction to optimize resources for cultivar testing. Crop Sci 59:1899–1915. https://doi.org/10.2135/cropsci2018.11.0692

Henderson CR (1984) Applications of linear models in animal breeding

Hu X (2015) A comprehensive comparison between ANOVA and BLUP to valuate location-specific genotype effects for rape cultivar trials with random locations. F Crop Res 179:144–149. https://doi.org/10.1016/j.fcr.2015.04.023

Huehn M (1990) Nonparametric measures of phenotypic stability. Part 1: Theory. Euphytica 47:189–194. https://doi.org/10.1007/BF00024241

Jarquín D, Crossa J, Lacaze X, et al (2014) A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor Appl Genet 127:595–607. https://doi.org/10.1007/s00122-013-2243-1

Jarquin D, Howard R, Crossa J, et al (2020) Genomic prediction enhanced sparse testing for multi-environment trials. G3 Genes, Genomes, Genet 10:2725–2739. https://doi.org/10.1534/g3.120.401349

Kofoid KD, Harvey TL (2005) Registration of Greenbug Resistant Sorghum Germplasm Lines KS 116 A/B through KS 120 A/B. Crop Sci 45:802.
Kuznetsova A, Brockhoff PB, Christensen RHB (2017) lmerTest package: tests in linear mixed effects models. J Stat Softw 82: https://doi.org/10.18637/jss.v082.i13

Lado B, Barrios PG, Quincke M, et al (2016) Modeling genotype × Environment interaction for genomic selection with unbalanced data from a wheat breeding program. Crop Sci 56:2165–2179. https://doi.org/10.2135/cropsci2015.04.0207

Laffont JL, Wright K, Hanafi M (2013) Genotype plus genotype × block of environments biplots. Crop Sci 53:2332–2341. https://doi.org/10.2135/cropsci2013.03.0178

Lenth R (2021) emmeans: Estimated Marginal Means, aka Least-Squares Means. R Package version 1.6.0

Li X, Guo T, Mu Q, et al (2018) Genomic and environmental determinants and their interplay underlying phenotypic plasticity. Proc Natl Acad Sci U S A 115:6679–6684. https://doi.org/10.1073/pnas.1718326115

Lian L, de los Campos G (2016) FW: An R Package for Finlay–Wilkinson Regression that Incorporates Genomic/Pedigree Information and Covariance Structures Between Environments. G3 Genes|Genomes|Genetics 6:589–597. https://doi.org/10.1534/g3.115.026328

Lopez-Cruz M, Crossa J, Bonnett D, et al (2015) Increased prediction accuracy in wheat breeding trials using a marker × environment interaction genomic selection model. G3 Genes, Genomes, Genet 5:569–582. https://doi.org/10.1534/g3.114.016097

Mbulwe L, Peterson GC, Scott-Armstrong J, Rooney WL (2015) Registration of sorghum germplasm Tx3408 and Tx3409 with tolerance to sugarcane aphid [Melanaphis sacchari (Zehntner)]. J Plant Regist 10:51–56. https://doi.org/10.3198/jpr2015.04.0025cr

Miller FR, Dusek TF, Prihoda KL, Rooney LW (1992) Registration of RTx436 sorghum parental line. Crop Sci 32:1518–1518. https://doi.org/10.2135/cropsci1992.0011183x00320060059x

Monk R, Franks C, Dahlberg J (2014) Sorghum. In: Yield gains in major US field crops. Wiley Online Library, pp 293–310

Nassar R, Huhn M (1987) Studies on Estimation of Phenotypic Stability: Tests of Significance for Nonparametric Measures of Phenotypic Stability. Biometrics 43:45. https://doi.org/10.2307/2531947

NASS (2021a). 2020 state agriculture overview: Texas. USDA National Agricultural Statistics Service. https://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=TEXAS (accessed Mar 25, 2021)

NASS (2021b). 2020 state agriculture overview: Kansas. USDA National Agricultural Statistics Service. https://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=KANSAS (accessed Mar 25, 2021)

NASS (U. S. Department of Agriculture National Agricultural Statistics Service) (2020).
Crop production. Available online at: 
https://www.nass.usda.gov/Publications/Todays_Reports/reports/crop1020.pdf (accessed feb 5 2021).

NASA 2021. Prediction Of Worldwide Energy Resources (POWER). Available online at: https://power.larc.nasa.gov/data-access-viewer/ (accessed April 16, 2021)

Nielsen DC, Vigil MF (2018) Wheat yield and yield stability of eight dryland crop rotations. Agron J 110:594–601. https://doi.org/10.2134/agronj2017.07.0407

Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. Biometrika 58:545–554. https://doi.org/10.1093/biomet/58.3.545

Pérez P, de los Campos G (2014) Genome-wide regression and prediction with the BGLR statistical package. Genetics 198:483–495. https://doi.org/10.1534/genetics.114.164442

Perumal R, Tesso T, Kofoid KD, et al (2019) Registration of six grain sorghum pollinator (R) lines. J Plant Regist 13:113–117. https://doi.org/10.3198/jpr2017.12.0087crp

Piepho H-P (1998) Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. Springer-Verlag

Piepho HP (1994) Best Linear Unbiased Prediction (BLUP) for regional yield trials: a comparison to additive main effects and multiplicative interaction (AMMI) analysis. Theor Appl Genet 89:647–654. https://doi.org/10.1007/BF00222462

Piepho HP (1997) Mixed models with multiplicative terms for cultivar trial data. Adv Biom Genet 215–220

Piepho HP, Möhring J (2005) Best Linear Unbiased Prediction of Cultivar Effects for Subdivided Target Regions. Crop Sci 45:1151–1159. https://doi.org/10.2135/cropsci2004.0398

Piepho HP, Möhring J, Melchinger AE, Büchse A (2008) BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161:209–228

Podlich DW, Cooper M (1998) Modelling plant breeding programs as search strategies on a complex response surface. In: Asia-Pacific Conference on Simulated Evolution and Learning. Springer, pp 171–178

R Core Team (2020) R: A language and environment for statistical computing

Rakshit S, Ganapathy KN, Gomashe SS, et al (2012) GGE biplot analysis to evaluate genotype, environment and their interactions in sorghum multi-location data. Euphytica 185:465–479. https://doi.org/10.1007/s10681-012-0648-6

Resende MDV, Thompson R (2004) Factor analytic multiplicative mixed models in the analysis of multiple experiments. Rev Matemática e Estatística, São Paulo, 22:31–52

Rooney WL (2016) Sorghum: Production and Improvement Practices, 2nd edn. Elsevier Ltd.

Rooney WL, Miller FR, Rooney LW (2003) Registration of RTx437 sorghum parental line.
Rosenow DT, Clark LE, Peterson GC, et al (2021) Registration of Tx3440 through Tx3482 sorghum germplasm. J Plant Regist plr2.20082. https://doi.org/10.1002/plr2.20082

Sharma SP, Leskovar DI, Crosby KM, Ibrahim AMH (2020) GGE biplot analysis of genotype-by-environment interactions for melon fruit yield and quality traits. HortScience 55:533–542

Shelton AC, Tracy WF (2017) Cultivar development in the U.S. public sector. Crop Sci 57:1823–1835. https://doi.org/10.2135/cropsci2016.11.0961

Shukla GK (1972) Some statistical aspects of partitioning genotype environmental components of variability. Heredity (Edinb) 29:237–245

Smith A, Cullis B, Thompson R (2001) Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. Biometrics 57:1138–1147. https://doi.org/10.1111/j.0006-341X.2001.01138.x

Smith AB, Borg LM, Gogel BJ, Cullis BR (2019) Estimation of Factor Analytic Mixed Models for the Analysis of Multi-treatment Multi-environment Trial Data. J Agric Biol Environ Stat 24:573–588. https://doi.org/10.1007/s13253-019-00362-6

Smith AB, Cullis BR, Thompson R (2005) The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. J. Agric. Sci. 143:449–462

Sorensen D, Gianola D (2007) Likelihood, Bayesian, and MCMC methods in quantitative genetics. Springer Science & Business Media

Wricke G (1962) Über eine Methode zur Erfassung der okologischen Streubreite in Feldversuchen. Z pflanzenzuchtg 47:92–96

Yan W (2015) Mega-environment Analysis and Test Location Evaluation Based on Unbalanced Multiyear Data. Crop Sci 55:113–122. https://doi.org/10.2135/cropsci2014.03.0203

Yan W (2016) Analysis and handling of G×E in a practical breeding program. Crop Sci 56:2106–2118

Yan W, Hunt LA, Sheng Q, Szlavnics Z (2000) Cultivar Evaluation and Mega-Environment
Investigation Based on the GGE Biplot. Crop Sci 40:597–605.  
https://doi.org/10.2135/cropsci2000.403597x

Yan W, Kang MS (2002) GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC press
Table 1. List of female (A-) and male (R-) lines from Texas A&M (TAMU) and Kansas State (KSU) breeding programs used in the factorial mating scheme to develop hybrids. The information includes pedigree, seed color, plant color, and publications describing the material if available.

| Parents | Pedigree | Seed color | Plant color | Reference             |
|---------|----------|------------|-------------|-----------------------|
| **KSU A-lines** |           |            |             |                       |
| ARCH11051A | ms3tan/ms3/3/Tx623B/PI 550610B/01714B | White | Tan | Experimental line |
| ARCH11129A | KP8BS2#16/OPB/3/Tx623 B/PI550610B/01714B | White | Tan | Experimental line |
| ARCH11136A | RRSWht-tanB0641701-167B ms3tan/ms3/3/Tx623B/PI 550610B/01714B | White | Tan | Experimental line |
| ARCH11146A |           |            |             |                       |
| KS118A | ms3tan/RTx436/PI550610 | White | Tan | Kofoid and Harvey (2005) |
| **KSU R-lines** |           |            |             |                       |
| ARCH111001R | 87BH8606-6//ms3/PI 550610/3/SC 414-12ER/PI 550610/4/N310B | White | Tan | Perumal et al. (2019) |
| ARCH111002R | 87BH8606-6//ms3/PI 550610/3/SC 414-12E R/PI 550610/4/N310B | White | Tan | Perumal et al. (2019) |
| ARCH111028R | 2006ISO11(Wht-tan RRS)R | White | Tan | Perumal et al. (2019) |
| ARCH111055R | 9614B/5/ms3tan/BHF14- (B1*Bvar)B/4/ms3//ms3/PI 550610 | Red | Tan | Perumal et al. (2019) |
| ARCH111056R | 9614B/5/ms3 tan/BHF14- (B1*Bvar)B/4/ms3//ms3/PI 550610 | Red | Tan | Perumal et al. (2019) |
| **TAMU A-lines** |           |            |             |                       |
| A05071 | BDLO357/BHF88 | White | Pigmented | Experimental line Rosenow et al., Release Notice Mbulwe et al., (2015) |
| ATx645 | BTx3042//BTx625/BTx642 | Red | Pigmented |                       |
| ATx3408 | BTx631/08PR047 | White | Tan |                       |
| ATx3447 | BTx643/BTx635 | White | Tan |                       |
| A08140 | BTx642//B9411/BTx645 | Yellow | Tan | Experimental line |
| **TAMU R-lines** |           |            |             |                       |
| R08304 | Tx430/00CS765 | Red | Pigmented | Experimental line Rooney et al., (2003) |
| RTx437 | 77CS4/RTx430 | White | Tan |                       |
| R07178 | 91BE7414/R01125 | Red | Tan | Experimental line |
| RTx436 | SC120-6-sel/2//Tx7000 | White | Tan | Miller et al., (1992) |
| EON361 | R5646/SC326-6 | White | Tan | Experimental line |
Table 2. Classification of the environments at which hybrid trials were conducted. The information includes growing season, average temperature, precipitation, adaptation zone, and GPS coordinates. Designations for environments are as follows: Monte Alto 2018 (18RF), Taft 2019 (19TA), Victoria 2018 (18VC), Victoria 2019 (19VC), College Station 2018 (18CS), College Station 2019 (19CS), Garden City 2018 (18GC), Garden City 2019 (19GC), Colby 2018 (18COL), Colby 2019 (19COL), and Hays 2019 (19HAY).

| Location | Soil Temperature (°C)† | Precipitation (mm)‡ | Growing season | Adaptation zone | GPS coordinates |
|----------|-------------------------|---------------------|----------------|----------------|-----------------|
| 18RF     | 21                      | 30                  | 306            | Feb - Jun      | Subtropical     |
|          |                         |                     |                |                | 26°21'06.1"N    |
|          |                         |                     |                |                | 97°53'50.3"W.   |
| 19TA     | 21                      | 30                  | 335            | Mar - Jul      | Subtropical     |
|          |                         |                     |                |                | 28°00'05.4"N    |
|          |                         |                     |                |                | 97°15'12.4"W.   |
| 18VC     | 20                      | 29                  | 373            | Mar - Jul      | Subtropical     |
|          |                         |                     |                |                | 28°47'24.4"N    |
|          |                         |                     |                |                | 96°50'22.6"W.   |
| 19VC     | 17                      | 30                  | 371            | Mar - Jul      | Subtropical     |
|          |                         |                     |                |                | 30°32'56.6"N    |
|          |                         |                     |                |                | 96°26'11.5"W.   |
| 18CS     | 18                      | 32                  | 323            | Mar - Jul      | Subtropical     |
|          |                         |                     |                |                | 37°59'21.4"N    |
|          |                         |                     |                |                | 100°48'52.5"W.  |
| 19CS     | 13                      | 30                  | 534            | Mar - Jul      | Subtropical     |
|          |                         |                     |                |                | 39°22'56.6"N    |
|          |                         |                     |                |                | 101°04'45.0"W.  |
| 18GC     | 27                      | 12                  | 555            | Jun - Oct      | Temperate       |
|          |                         |                     |                |                | 38°51'10.8"N    |
|          |                         |                     |                |                | 99°20'24.6"W.   |
| 19GC     | 24                      | 11                  | 241            | Jun - Oct      | Temperate       |
|          |                         |                     |                |                | 39°22'56.6"N    |
|          |                         |                     |                |                | 101°04'45.0"W.  |
| 18COL    | 26                      | 10                  | 359            | Jun - Oct      | Temperate       |
|          |                         |                     |                |                | 39°22'56.6"N    |
|          |                         |                     |                |                | 101°04'45.0"W.  |
| 19COL    | 22                      | 9                   | 345            | Jun - Oct      | Temperate       |
|          |                         |                     |                |                | 39°22'56.6"N    |
|          |                         |                     |                |                | 101°04'45.0"W.  |
| 19HAY    | 22                      | 10                  | 405            | Jun - Oct      | Temperate       |
|          |                         |                     |                |                | 38°51'10.8"N    |
|          |                         |                     |                |                | 99°20'24.6"W.   |

†Average soil surface temperature at beginning, end of the growing season
‡Cumulative precipitation across growing season
Table 3. Phenotypic performance of hybrids developed from Texas A&M and Kansas State elite lines under a factorial II mating design across environments. Estimates include genetic variance ($\sigma^2_{hyb}$), residual variance ($\sigma^2_e$), coefficient of variation (CV), and broad-sense heritability ($H^2$). Designations for environments are as follows: Monte Alto 2018 (18RF), Victoria 2018 (18VC), College Station 2018 (18CS), Garden City 2018 (18GC), Colby 2018 (18COL), Taft 2019 (19TA), Victoria 2019 (19VC), College Station 2019 (19CS), Hays 2019 (19HAY), and Colby 2019 (19COL).

| Trait | Parameter | Environments |
|-------|-----------|--------------|
|       |           | 18RF | 18VC | 18CS | 18GC | 18COL | 19TA | 19VC | 19CS | 19HAY | 19COL |
| GY    | Mean (t/ha) | 4.31 | 5.18 | 6.26 | 5.58 | 6.65 | 6.46 | 7.78 | 6.51 | 8.34 |
|       | $\sigma^2_{hyb}$ | 0.70* | 0.28* | 0.66* | 0.67* | 0.57* | 0.47* | 0.49* | 0.77* | 0.47* | 1.36* |
|       | $\sigma^2_e$ | 0.86 | 0.15 | 0.66 | 0.95 | 1.24 | 0.18 | 0.29 | 1.36 | 1.55 | 0.84 |
|       | CV (%) | 21.49 | 7.35 | 12.95 | 17.47 | 21.51 | 6.37 | 8.29 | 14.97 | 19.14 | 10.96 |
|       | $H^2$ | 0.71 | 0.85 | 0.75 | 0.68 | 0.58 | 0.84 | 0.77 | 0.53 | 0.38 | 0.76 |
| PHT   | Mean (cm) | 121 | 127 | 126 | 145 | 137 | 138 | 140 | 140 | 122 | 129 |
|       | $\sigma^2_{hyb}$ | 64.6* | 91.0* | 155.7* | 111.7* | 94.8* | 80.3* | 71.2* | 137.4* | 89.3* | 78.1* |
|       | $\sigma^2_e$ | 26.31 | 19.33 | 48.76 | 27.15 | 25.32 | 30.14 | 14.32 | 53.02 | 28.05 | 67.18 |
|       | CV (%) | 4.22 | 3.45 | 5.5 | 3.58 | 3.67 | 3.96 | 2.7 | 5.2 | 4.33 | 6.32 |
|       | $H^2$ | 0.88 | 0.93 | 0.91 | 0.93 | 0.92 | 0.84 | 0.91 | 0.84 | 0.86 | 0.76 |
| DA    | Mean (day) | 67.9 | 67.2 | 78.7 | 70.6 | 72.4 | 69.6 | 59.5 | 80.4 | 64.2 | 69.1 |
|       | $\sigma^2_{hyb}$ | 11.1* | 4.7* | 10.9* | - | - | 1.6* | 3.5* | 4.2* | - | - |
|       | $\sigma^2_e$ | 4.77 | 0.99 | 4.36 | - | - | 0.54 | 0.78 | 6.88 | - | - |
|       | CV (%) | 3.22 | 1.48 | 2.65 | - | - | 1.06 | 1.48 | 3.26 | - | - |
|       | $H^2$ | 0.87 | 0.93 | 0.88 | - | - | 0.86 | 0.90 | 0.55 | - | - |

* Significant at 0.05 probability level
Table 4. Combined analysis for grain yield (GY), plant height (PHT), and days to anthesis (DA) across 10 environments (Combined), Mega-environment 1 (ME1), and Mega-environment 2 (ME2). Estimates include variance components of model effects, mean of trait, coefficient of variation (CVe), and broad-sense heritability ($H^2$) for each dataset.

| Trait | Variance Components | Combined | ME1 | ME2 |
|-------|---------------------|----------|-----|-----|
|       | Estimate            | %        | Estimate | % | Estimate | % |
| GY    | Hybrid              | 0.29**   | 9.75   | 0.28** | 10.68 | 0.54** | 13.92 |
|       | Environment         | 1.45**   | 48.92  | 1.40** | 52.73 | 1.96** | 50.90 |
|       | Hybrid × Env        | 0.35**   | 11.71  | 0.28** | 10.45 | 0.19*  | 4.93  |
|       | Rep(Env)            | 0.08**   | 2.72   | 0.13** | 4.76  | 0.01   | 0.36  |
|       | Residual            | 0.80     | 26.91  | 0.57   | 21.37 | 1.15   | 29.89 |
|       | Mean (t/ha)         | 6.22     | 6.10   | 6.40   |       |       |      |
|       | CVe (%)             | 14.35    | 12.36  | 16.77  |       |       |      |
|       | $H^2$               | 0.83     | 0.78   | 0.79   |       |       |      |
| PHT   | Hybrid              | 82.39**  | 40.49  | 86.81**| 42.94 | 84.60**| 37.14 |
|       | Environment         | 68.50**  | 33.66  | 64.30**| 31.80 | 97.06**| 42.62 |
|       | Hybrid × Env        | 16.62**  | 8.17   | 14.15**| 7.00  | 12.04**| 5.28  |
|       | Rep(Env)            | 3.34**   | 1.64   | 5.27** | 2.61  | 0.43   | 0.19  |
|       | Residual            | 32.65    | 16.04  | 31.64  | 15.65 | 33.63  | 14.77 |
|       | Mean (cm)           | 132.9    | 132.4  | 133.6  |       |       |      |
|       | CVe (%)             | 4.30     | 4.25   | 4.34   |       |       |      |
|       | $H^2$               | 0.97     | 0.95   | 0.94   |       |       |      |
| DA    | Hybrid              | 5.33**   | 10.72  | 4.27** | 6.07  | -      | -     |
|       | Environment         | 38.11**  | 76.70  | 60.41**| 85.87 | -      | -     |
|       | Hybrid × Env        | 2.46**   | 4.96   | 2.01** | 2.86  | -      | -     |
|       | Rep(Env)            | 0.58**   | 1.16   | 0.58** | 0.82  | -      | -     |
|       | Residual            | 3.22     | 6.47   | 3.08   | 4.38  | -      | -     |
|       | Mean (days)         | 69.9     | 70.5   | 69.07  |       |       |      |
|       | CVe (%)             | 2.56     | 2.49   | 2.49   |       |       |      |
|       | $H^2$               | 0.94     | 0.89   | 0.89   |       |       |      |

*, ** Significant at 0.01 and 0.001, respectively
Table 5 – Names, codes, and origin of grain sorghum elite lines used to generate hybrid combinations following a factorial II mating scheme. Information includes environment names and codes involved in the multi environment trial to assess grain yield, plant height, and days to anthesis of grain sorghum hybrids derived from U.S. public breeding programs.

| Origin | Female Codes | Male Codes | Environment Codes |
|--------|--------------|------------|-------------------|
| Texas  | A05071 TF1   | EON361 TM1 | Monte Alto 2018 18RF |
|        | A08140 TF2   | R07178 TM2 | Victoria 2018 18VC |
|        | ATx3447 TF3  | R08304 TM3 | College Station 2018 18CS |
|        | ATx3408 TF4  | RTx436 TM4 | Garden City 2018 18GC |
|        | ATx645 TF5   | RTx437 TM5 | Colby 2018 18COL |
|        | AARCH11051 KF1 | RARCH11001 KM1 | Taft 2019 19TA |
| Kansas | AARCH11129 KF2 | RARCH11002 KM2 | Victoria 2019 19VC |
|        | AARCH11136 KF3 | RARCH11028 KM3 | College Station 2019 19CS |
|        | AARCH11146 KF4 | RARCH11055 KM4 | Hays 2019 19HAY |
|        | AKS118 KF5   | RARCH11056 KM5 | Colby 2019 19COL |
Table 6. Prediction accuracies of genomic GxE effects model for grain yield (GY), plant height (PHT), and days to anthesis (DA). Predictions are presented for genomic GxE effects model with and without partitioning data into mega-environments (ME). The environments from Texas composed ME1, while ME2 accounted for Kansas year-location combination as determined by Genotype plus genotype × environment (GGE) biplot. Numbers in bold show the overall prediction accuracy across females and males.

| Parents  | Prediction accuracy without ME | Prediction accuracy in ME1 | Prediction accuracy in ME2 |
|----------|-------------------------------|----------------------------|---------------------------|
|          | GY   | PHT | DA  | GY   | PHT | DA  | GY   | PHT | DA  |
| Female   |      |     |     |      |     |     |      |     |     |
| A05071   | 0.65 | 0.85 | 0.76 | 0.86 | 0.89 | 0.84 | 0.35 | 0.78 | 0.65 |
| A08140   | 0.59 | 0.85 | 0.69 | 0.75 | 0.89 | 0.81 | 0.36 | 0.80 | 0.56 |
| AARCH11051 | 0.67 | 0.91 | 0.76 | 0.79 | 0.92 | 0.79 | 0.51 | 0.90 | 0.72 |
| AARCH11129 | 0.37 | 0.87 | 0.69 | 0.51 | 0.91 | 0.77 | 0.17 | 0.79 | 0.56 |
| AARCH11136 | 0.69 | 0.82 | 0.86 | 0.71 | 0.80 | 0.90 | 0.64 | 0.83 | 0.79 |
| AARCH11146 | 0.63 | 0.79 | 0.81 | 0.75 | 0.81 | 0.79 | 0.45 | 0.77 | 0.83 |
| AHF14    | 0.60 | 0.80 | 0.83 | 0.66 | 0.83 | 0.88 | 0.46 | 0.75 | 0.80 |
| AKS118   | 0.71 | 0.82 | 0.84 | 0.78 | 0.83 | 0.85 | 0.54 | 0.78 | 0.83 |
| ATx3408  | 0.68 | 0.90 | 0.85 | 0.73 | 0.89 | 0.84 | 0.59 | 0.88 | 0.90 |
| ATx645   | 0.67 | 0.83 | 0.62 | 0.74 | 0.81 | 0.81 | 0.58 | 0.87 | 0.35 |
| Mean     | 0.63 | 0.84 | 0.77 | 0.73 | 0.86 | 0.83 | 0.47 | 0.81 | 0.70 |
| Male     |      |     |     |      |     |     |      |     |     |
| EON361   | 0.55 | 0.68 | 0.74 | 0.40 | 0.65 | 0.73 | 0.76 | 0.72 | 0.78 |
| R07178   | 0.64 | 0.84 | 0.83 | 0.68 | 0.83 | 0.81 | 0.59 | 0.87 | 0.85 |
| R08304   | 0.51 | 0.83 | 0.80 | 0.51 | 0.81 | 0.77 | 0.45 | 0.86 | 0.83 |
| RARCH11001 | 0.62 | 0.83 | 0.75 | 0.52 | 0.76 | 0.81 | 0.82 | 0.89 | 0.68 |
| RARCH11002 | 0.71 | 0.86 | 0.75 | 0.74 | 0.82 | 0.76 | 0.63 | 0.90 | 0.71 |
| RARCH11028 | 0.47 | 0.69 | 0.74 | 0.44 | 0.68 | 0.76 | 0.56 | 0.72 | 0.70 |
| RARCH11055 | 0.58 | 0.81 | 0.63 | 0.44 | 0.80 | 0.65 | 0.78 | 0.82 | 0.57 |
| RARCH11056 | 0.56 | 0.77 | 0.83 | 0.48 | 0.76 | 0.86 | 0.71 | 0.79 | 0.80 |
| RTx436   | 0.49 | 0.74 | 0.67 | 0.52 | 0.76 | 0.77 | 0.45 | 0.69 | 0.54 |
| RTx437   | 0.58 | 0.81 | 0.73 | 0.61 | 0.83 | 0.84 | 0.53 | 0.78 | 0.59 |
| Mean     | 0.57 | 0.79 | 0.75 | 0.53 | 0.77 | 0.77 | 0.63 | 0.80 | 0.70 |
| Overall mean | 0.60 | 0.82 | 0.76 | 0.63 | 0.82 | 0.80 | 0.55 | 0.81 | 0.70 |

‡Prediction accuracies are as presented in Fonseca et al. (2021a)
Figure 1. A- and R-lines from Texas (maroon) and Kansas (purple) sorghum breeding programs used in a factorial mating scheme to generate four sets of hybrid combinations. Maroon and purple squares represent hybrid combinations derived from Texas and Kansas programs, respectively. White squares represent hybrids generated across programs.
Figure 2. Genotype plus genotype × environment (GGE) biplot based on grain yield of sorghum hybrids evaluated in ten environments. GGE biplot was environment-centered and singular value scaling via nipals algorithm. Information includes the first and second principal components (PC) and the which-won-where polygon. Hybrids are identified in gray, while environments in dark orange.
Figure 3. Mosaic plot and genotype plus genotype × block of environments (GGB) biplot of the multi-environment trial visualizing the two-way partitioning of the total sums of squares (TSS) into genotype (G), genotype × blocks of environments (GB), and residuals (R) along each GGB biplot axis for yield. The two mega-environments are denoted 1 (ME1) and 2 (ME2). ME1 included Texas environments as defined in Table 2. Hybrids are labeled according to each female x male combination. TF and TM codes for female and males derived from Texas A&M sorghum breeding programs, respectively, whereas females and males derived from Kansas breeding programs are identified as KF and KM respectively. Number following the inbred line origin identifies different parents.
Figure 4. The “mean vs stability” representation of the genotype plus genotype × environment (GGE) biplot based on mega-environment 1 (left), and Bayesian Finlay-Wilkinson (BFW) regression (right) for grain yield. For GGE, data were environment-centered and not scaled; the biplot was focused on singular value partitioning using nipals algorithm and grain yield scores projections on the average environment coordinate displayed the stability of each hybrid. Underlined hybrids in the GGE biplot indicated the top and bottom two and an intermediate less stable hybrid. The underlined hybrids were assessed using BFW regression to further investigate static and dynamic stability.
Figure 5. Bayesian Finlay-Wilkinson (BFW) regression of hybrids based on the top and bottom two hybrids and an intermediate less stable hybrid to assess plant height (left) and days to anthesis (right) stability of grain sorghums within mega-environment 1.
Figure 6. The “mean vs stability” representation of the genotype plus genotype × environment (GGE) biplot based on mega-environment 2 (left), and Bayesian Finlay-Wilkinson (BFW) regression (right) for grain yield. For GGE, data were environment-centered and not scaled. The biplot was genotype-focused singular value partitioning using nipals algorithm. Grain yield score projections on the average environment coordinate displayed the stability of each hybrid. Underlined hybrids in the GGE biplot indicated the top and bottom two hybrids and an intermediate less stable hybrid. The underlined hybrids were assessed using BFW regression to further investigate static and dynamic stability.
Figure 7. Bayesian Finlay-Wilkinson (BFW) regression of hybrids based on the top and bottom two hybrids and an intermediate less stable hybrid to assess plant height (left) and days to anthesis (right) stability of grain sorghums within mega-environment 2.