Received: 6 May 2019 | Accepted: 15 October 2019 | Published online: 29 January 2020

DOI: 10.1002/csc2.20035

Crop Science

ORIGINAL RESEARCH ARTICLE

Crop Breeding & Genetics

Relative utility of agronomic, phenological, and morphological traits for assessing genotype-by-environment interaction in maize inbreds

Celeste M. Falcon[1] | Shawn M. Kaeppler[1] | Edgar P. Spalding[2] | Nathan D. Miller[2] | Nicholas Haase[1] | Naser AlKhalifah[1,3] | Martin Bohn[4] | Edward S. Buckler[5,6] | Darwin A. Campbell[3] | Ignacio Ciampitti[7] | Lisa Coffey[8] | Jode Edwards[9] | David Ertl[10] | Sherry Flint-Garcia[11] | Michael A. Gore[12] | Christopher Graham[13] | Candice N. Hirsch[14] | James B. Holland[15] | Diego Jarquín[16] | Joseph Knoll[17] | Nick Lauter[9] | Carolyn J. Lawrence-Dill[3,18] | Elizabeth C. Lee[19] | Aaron Lorenz[14] | Jonathan P. Lynch[20] | Seth C. Murray[21] | Rebecca Nelson[22] | M. Cinta Romay[6] | Torbert Rocheford[23] | Patrick S. Schnable[8,18] | Brian Scully[24] | Margaret Smith[11] | Nathan Springer[25] | Mitchell R. Tuinstra[23] | Renee Walton[3] | Teclemariam Weldekidan[26] | Randall J. Wisser[26] | Wenwei Xu[27] | Natalia de Leon[1]

[1]Dep. of Agronomy, Univ. of Wisconsin, 1575 Linden Drive, Madison, WI 53706

[2]Dep. of Botany, Univ. of Wisconsin, 430 Lincoln Dr, Madison, WI 53706

[3]Dep. of Genetics, Development and Cell Biology, Iowa State Univ., Ames, IA 50011

[4]Dep. of Crop Sciences, Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801

[5]USDA-ARS Plant, Soil, and Nutrition Research Unit, Cornell Univ., Ithaca, NY 14853

[6]Institute for Genomic Diversity, Cornell Univ., Ithaca, NY 14853

[7]Dep. of Agronomy, Kansas State Univ., Manhattan, KS 66506

[8]Plant Sciences Institute, Iowa State Univ., Ames, IA 50011

[9]USDA-ARS Corn, Insects, and Crop Genetics Research Unit, Iowa State Univ., Ames, IA 50011

[10]Iowa Corn Promotion Board, 5505 NW 88th Street, Johnston, IA 50131

[11]USDA-ARS Plant Genetics Research Unit, Univ. of Missouri, Columbia, MO 65211

[12]Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell Univ., Ithaca, NY 14853

[13]Plant Science Dep., West River Agricultural Center, South Dakota State Univ., Rapid City, SD 57769

[14]Dep. of Agronomy and Plant Genetics, Univ. of Minnesota, St. Paul, MN 55108

[15]USDA-ARS Plant Science Research Unit, North Carolina State Univ., Raleigh, NC 27695

Abbreviations: AEC, average environment coordination; AMMI, additive main effects and multiplicative interactions; ASOS, Automated Surface Observing Systems; BLUP, best linear unbiased predictor; ex-PVP, expired Plant Variety Protection; G × E, genotype × environment interaction; GDU, growing degree units; GGE, genotypic main effects and G × E; MSE, mean squared error; NSS, non-stiff stalk; NWS, National Weather Service; PC, principal component; SSS, stiff-stalk synthetic.

[16]Dep. of Agronomy and Horticulture, Univ. of Nebraska, Lincoln, NE 68583
[17]USDA-ARS Crop Genetics and Breeding Research Unit, Tifton, GA 31793
[18]Dep. of Agronomy, Iowa State Univ., Ames, IA 50011
[19]Dep. of Plant Agriculture, Univ. of Guelph, Guelph, ON N1G 2W1, Canada
[20]Dep. of Plant Science, Penn State Univ., University Park, PA 16802
[21]Dep. of Soil and Crop Sciences, Texas A&M Univ., College Station, TX 77843
[22]Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell Univ., Ithaca, NY 14853
[23]Dep. of Agronomy, Purdue Univ., West Lafayette, IN 47907
[24]USDA-ARS, US Horticulture Research Laboratory, Fort Pierce, FL
[25]Dep. of Plant and Microbial Biology Univ. of Minnesota, St. Paul, MN 55108
[26]Dep. of Plant and Soil Sciences, Univ. of Delaware, Newark, DE 19716
[27]Texas A&M AgriLife Research, Texas A&M University, Lubbock, TX 79403

Correspondence
Natalia de Leon, Dep. of Agronomy, Univ. of Wisconsin, 1575 Linden Drive, Madison, WI 53706.
Email: ndeleongatti@wisc.edu

Funding information
Wisconsin Corn Promotion Board; U.S. Department of Agriculture, Grant/Award Number: Hatch for several co-authors; Texas Corn Producers Board and Texas A&M AgriLife; Nebraska Corn Board; National Science Foundation, Grant/Award Numbers: IOS, 1444456; National Institute of Food and Agriculture; National Corn Growers Association; Minnesota Corn Research and Promotion Council; Iowa Corn Promotion Board; Illinois Corn Marketing Board; Agricultural Research Service, Grant/Award Numbers: provided for ESB, JE, SF-G, JBH, JK, NL, BS

Assigned to Associate Editor Weikai Yan.

Abstract

Plant breeders face the challenge of genotype × environment interaction (G × E) in comprehensively breeding for expanded geographic regions. An improved understanding of G × E sensitivity of traits and the environmental features that effectively discriminate among genotypes will enable more efficient breeding efforts. In this study of 31 maize (*Zea mays* L.) inbreds grown in 36 environments that are part of the Genomes to Fields Initiative, we measured 14 traits, including flowering date, height, and yield components (i.e., ear and kernel dimensions) to (i) identify traits that are the most sensitive indicators of G × E; (ii) determine how geographic location and weather factors influence environments' discriminability of inbreds; and (iii) detect patterns of stability in better and worse discriminating environments. Genotype × environment interaction explained between 9.0–20.4% of the phenotypic variance with greater effects in the yield-component traits. Discriminability of environments varied by trait. Midwest locations (where 26 of the 31 inbreds were developed) were among the most discriminating environments for more traits, while environments in the West and East tended to be less discriminating. Weather factors during silking were significantly different between the most and least discriminating environments more often than average weather across the season or during the period from planting to silking. Stability of genotypes varied by trait, and performance was usually not correlated with stability. The dissection of complex traits, such as yield into component traits, appears to be a useful approach to understand how environmental factors differentially affect phenotype.

1 | INTRODUCTION

The variation in plants' observed phenotypes can be partitioned into three main factors of interest—genotype, environment, and genotype × environment interaction (G × E)—in addition to other minor factors and measurement error. In plant breeding, G × E plays an important role, as the relative performance of different genotypes in different environments influences plant breeders' recommendations of best-performing varieties for specific regions. Typically, plant breeders have minimized G × E by producing cultivars that are appropriate for regions that share common environmental characteristics (Bernardo, 2010). With improved understanding of specific components of genotype, environment, and G × E, breeders may be able to more deliberately use data-based approaches to enhance their ability to position a larger number of genotypes in environments to maximize productivity. Grain yield is of chief importance in breeding maize (*Zea mays* L.), and is commonly considered alongside several traits that affect it indirectly or directly such as flowering, height, and yield-component traits. Due to their differences in heritability and sensitivity to environmental factors, these

different types of traits may show different levels of G × E. Flowering traits, such as days to anthesis and days to silking, provide information on the genotype's degree of adaptation to the environment, and therefore its potential performance and typically shows less G × E than productivity traits. In addition to serving as an indicator of overall plant health or stress, plant and ear height affect several factors that, in turn, influence yield, such as potential for lodging and suitability for machine harvest. Plant and ear height are highly heritable traits (Hung et al., 2012; Peiffer et al., 2014), suggesting that they may show less G × E. While grain yield has low heritability and is sensitive to environmental influences, yield-component traits tend to have higher heritability (Austin & Lee, 1996, 1998; Messmer et al., 2009).

The magnitude of the G × E variance (in other words, the proportion of phenotypic variation due to the G × E component) varies across traits and is related to how genetic and environmental factors affect that trait throughout the plant's development. While plant breeders and geneticists may refer to these differences as G × E "sensitivity" in general terms, few field studies of crop species have investigated relative G × E variance across traits. Here, we introduce only a few papers that focused on the magnitude of G × E across different traits, recognizing that an extensive body of literature exists for G × E in more general terms. Two studies of protein quality in maize examined the effect of G × E on several traits, and one of them found different levels of G × E sensitivity across traits. In one study, 62 maize hybrids or open-pollinated cultivars were tested at 13 tropical locations across four continents. Analysis of variance (ANOVA) indicated that G × E had a significant effect ($P < .05$) on all the traits assessed, namely grain yield, endosperm modification, protein content of grain, tryptophan content of grain, and tryptophan content of protein (Pixley & Bjarnason, 2002). The other study assessed 30 experimental hybrids in six trials, including two with low-nitrogen stress and two with drought stress. Genotype × environment interaction had a significant effect for anthesis-silking interval, senescence score, ears per plant, and grain yield ($P < .01$) as well as protein in grain and lysine in protein ($P < .05$) but not for ear rot, tryptophan in grain, lysine in grain, or tryptophan in protein (Zaidi, Vasal, Maniselvan, Jha, & Singh, 2008).

To be useful for breeding, a test environment should be both representative of a growing region and discriminating of genotypes, meaning that it provides information about differences in their performance (Yan, Kang, Ma, Woods, & Cornelius, 2007). An environment's discriminating power can be quantified in terms of the within-environment heritability or the standard deviation of genotype means in the environment, which corresponds to the length of the environment vector in the "discriminating power vs. representativeness" view of the genotypic main effects and genotype × environment interactions (GGE) biplot (Yan et al., 2007). In a study of 13 maize hybrids grown in 10 locations over 2 yr, a significant G × E effect explaining 15.9% of the variance in grain yield was detected (Fan et al., 2007). Using GGE analysis, the authors identified two least-discriminating locations and one least-representative location, ultimately classifying four environments as the least ideal (three sites with one location classified as non-ideal in both years). They concluded that GGE analysis would be useful for making better use of limited resources available for the testing program by choosing appropriate test environments. Only a few other studies have used the GGE analysis to evaluate test environments for various traits in other crops, including yield in soybean [*Glycine max* (L.) Merr.] (Yan & Rajcan, 2002), lint yield and fiber length in cotton (*Gossypium hirsutum* L.; Blanche & Myers, 2006), yield in common bean (*Phaseolus vulgaris* L.; Kang, Aggarwal, & Chirwa, 2006), and cane yield, sucrose yield, and theoretical-recoverable sucrose in sugarcane (*Saccharum* spp.) cultivars (Glaz & Kang, 2008). These studies did not investigate which weather factors may be associated with discriminability.

Numerous studies in maize have assessed the stability of inbreds or hybrids, using linear regression, additive main effects and multiplicative interactions (AMMI), and GGE analyses. For example, 101 maize hybrids were grown at three Illinois locations over 2 or 3 yr (depending on location) and under varying agronomic treatments (low versus high nitrogen fertilizer and standard versus high plant density; Mastrodomenico, Haegele, Seebauer, & Below, 2018). The authors concluded that yield stability, in terms of slope from linear regression analysis, depended more on a hybrid's response to nitrogen level than on its tolerance of high plant density. Using AMMI analysis, Zaidi et al. (2008) identified quality protein maize hybrids that were least or most stable for several traits. They concluded that many hybrids were stable for some traits, like tryptophan in grain, while relatively few hybrids were stable for other traits like grain protein content. The "mean vs. stability" view of the GGE biplot offers another method for quantifying stability. In this view, genotype stability is approximated based on the average environment coordination (AEC) ordinate axis, which corresponds to genotypes' contributions to the G × E interaction. Genotypes lying near the AEC abscissa are more stable than those farther from it and are ranked consistently across environments (Yan et al., 2007). Meseka et al. (2016) used GGE analysis in a study of nine three-way cross yellow maize hybrids grown in 17 environments in the savannas of West Africa. They identified two hybrids with particularly good stability across all environments, including three with drought stress conditions.

The purpose of this experiment was to determine the most useful approach to study G × E, especially in terms of which traits are most informative. An in-depth analysis of the interplay of G × E sensitivity among several traits, discriminability

among environments, and stability among genotypes will improve our understanding of the relationships among these factors in plant breeding. This knowledge will enable more comprehensive breeding to exploit G × E, instead of focusing on breeding for a defined region to reduce G × E. Accordingly, our objectives for this research were to (i) determine whether G × E interaction is present for several agronomic, phenological, and morphological traits in 31 diverse maize inbred lines across more than 30 environments and to establish which traits are most sensitive to G × E; (ii) determine which environments are best at discriminating among genotypes for various traits and what geographical and environmental factors contribute to discriminability for those traits; and (iii) ascertain the level of stability for the set of maize inbred lines and detect differences in stability between sub-groups of inbreds.

2 | MATERIALS AND METHODS

2.1 | Germplasm, environments, phenotypic, and weather data collection

A set of 31 maize inbreds, chosen to represent a range of maturities and release dates, was selected for this study (Table 1). We chose to use inbreds for this study due to their higher G × E sensitivity compared to hybrids. The inbreds were grown in 36 environments (15 trials in 12 US states in 2014 and 21 trials in 14 states in 2015; Table 2; Supplemental Figure S1) that are part of the Genomes to Fields Initiative (Gage et al., 2017). The locations of these trials were geographically diverse, ranging from 30.55–45.00° latitude and from −75.75 to −102.93° longitude. We defined environment as the combination of site and year because the physical locations of specific trials were often different each year, even when they were nearby. For example, IA_1 in 2014 is not the same field location as IA_1 in 2015, as indicated by the latitude and longitude for these locations. The trials were grown in a randomized complete block design using two replications per environment (with the exceptions of GA2_2014, which had three replications and WI1_14, which had four replications). Planting density ranged from about 39,500–123,000 plants per hectare.

Phenotypic data were collected for 14 phenological, morphological, and yield-component traits. Days to anthesis referred to the number of days between planting and when 50% of the plants in a plot exhibited anther exertion on more than half of the tassel main spike. Days to silking referred to the days between planting and when 50% of the plants in a plot showed silk emergence. These flowering traits were converted to growing degree units (GDU) accumulated from planting to the flowering date using air temperature (°C) data collected in the field at each location. First, temperatures above 30°C were set to 30°C, and temperatures below 10°C were set to 10°C (Gilmore & Rogers, 1958). Temperature was converted from °C to °F, and then the equation GDU = $[(T_{min} + T_{max})/2] - 50$ was used to calculate the GDU for each day from planting to the flowering date. Plant height (cm) referred to the distance from the base of the plant to the ligule of the uppermost leaf at reproductive maturity. Ear height (cm) was measured as the distance from the ground to the uppermost ear bearing node at reproductive maturity.

To measure yield-component and ear and kernel morphology traits, the primary ear was collected from three representative plants per plot at grain physiological maturity. The ears from each trial location were shipped to the University of Wisconsin for imaging where they were dried to approximately 10–15% kernel moisture. Images of the dried ears and were collected using EPSON Perfection V700 PHOTO flatbed scanners. The ears were shelled, and the weight of the shelled kernels from the three ears was recorded as plot grain weight before taking a sample of kernels for imaging. Measurements of ear and kernel traits were automatically computed from these images using a previously described analysis pipeline (Miller et al., 2017) and as implemented in workflows publicly available on CyVerse.

Image-based measurements were obtained for ear length (mm), ear width (mm), kernel depth (mm), kernel thickness (mm), kernel width (mm), and kernel area (mm^2). The analysis pipeline also counted the number of kernels in the image, and these kernels were weighed manually. These measurements were combined to calculate mean kernel weight (g). Kernels per row was also calculated based on the ear length and kernel thickness. Kernel row number was counted manually as the number of kernel rows around the middle to lower third of an ear.

Weather data were collected in each environment using WatchDog Model 2700 (Spectrum Technologies, East-Plainfield, IL) weather stations. At 30-min intervals throughout the growing season, data were collected for air temperature (°C) and rainfall (mm). Several environments were irrigated (Table 2); at some of these, irrigation was tracked along with any precipitation received (DE1_14, GA2_14, DE1_15, GA1_15, and NC1_15) while in other environments, irrigation was not tracked (NC1_14, TX1_14, NY2_15, and TX3_15). To verify calibration and identify erroneous data points, weather station observations were compared to nearby National Weather Service (NWS) Automated Surface Observing Systems (ASOS) data. The calibrated dataset includes observations from the NWS ASOS as well as a "calibrated" column for most elements. Where available, these calibrated data were used for our analyses. The phenotypic and weather data sets are described in more detail (AlKhalifah et al., 2018) and are available to view and download (2014 agronomic and weather data: https://doi.org/10.7946/P2V888; 2015 agronomic and

TABLE 1 Thirty-one inbred genotypes evaluated across 36 environments and their pedigree, heterotic group, Plant Variety Protection status, and year of release

Genotype	Pedigree	Group[a]	Ex-PVP[b]	Origin	Year of Release
740	Mexican Deep Kernel X Mo17(4)	NSS	Ex-PVP	MN	1988
2369	2702H X B73(2)	SSS	Ex-PVP	CO	1989
A619	A171 X Oh43(2)	NSS	Public	MN	1961
A632	Mt42 X B14(4)	SSS	Public	MN	1963
A634	Mt42 X B14(4)	SSS	Public	MN	1966
B14	BSSS-C0	SSS	Public	IA	1953
B37	BSSS-C0	SSS	Public	IA	1958
B73	Selected from advanced recurrent selection population (C5) of Iowa Stiff Stalk Synthetic (BSSS)	SSS	Public	IA	1972
C103	Lancaster Surecrop (from Noah Hershey)	NSS	Public	CT	1949
CM105	CMV3 X B14(2)	SSS	Public	MB	Circa 1980
LH123HT	Pioneer hybrid 3535	NSS	Ex-PVP	IA	1985
LH145	A632Ht X CM105	SSS	Ex-PVP	IA	1984
LH162	ND246 x Mo17	NSS	Ex-PVP	IA	1991
LH195	LH117 X LH132	SSS	Ex-PVP	IA	1991
LH198	LH132(2) X B84	SSS	Ex-PVP	IA	1993
LH74	A632 x B73	SSS	Ex-PVP	IA	1983
LH82	Holden line 610 (610 = W153R type) X LH7	NSS	Ex-PVP	IA	1985
Mo17	C.I.187–2 X C103	NSS	Public	MO	1964
PB80	(10670–1 X B73) X B73Ht(BC6)	SSS	Ex-PVP	IL	1988
PH207	PHG3BD2 X PHG3RZ1	NSS	Ex-PVP	IA	1984
PHB47	SD105 X B37(3)	SSS	Ex-PVP	MN	1984
PHG35	G3BD2 X H7FS6	NSS	Ex-PVP	IA	1983
PHG39	A33GB4 X A34CB4	SSS	Ex-PVP	IN	1983
PHG47	PH041 X MKSDTE	NSS	Ex-PVP	MN	1987
PHJ40	PHB09 X PHB36	SSS	Ex-PVP	ON	1987
PHN82	PHG29 X HD38	NSS	Ex-PVP	IA	1990
PHV63	PH555 X Zap < 4CB (Zapalote chico and Corn Belt inbreds put together to develop insect tolerance)	NSS	Ex-PVP	TN	1988
PHW52	B73 X G39	SSS	Ex-PVP	IA	1989
PHZ51	PH814 X PH848	NSS	Ex-PVP	IA	1987
W117	643 X Minnesota #13	NSS	Public	WI	1963
Wf9	Reid Yellow Dent (Indiana station strain)	NSS	Public	IN	1936

[a]Heterotic group abbreviations: NSS-non-stiff stalk synthetic, SSS-stiff stalk synthetic
[b]Ex-PVP, expired Plant Variety Protection; Public, a publicly released inbred.

weather data: https://doi.org/10.7946/P24S31; 2014 and 2015 yield-component imaging data: https://doi.org/10.7946/P2C34P).

2.2 | Data cleaning and statistical analysis

Extreme outliers beyond five standard deviations from the mean of each trait were removed from the data set. This data cleaning step allowed us to remove erroneous data points, especially those that may have been generated as part of our automated, high-throughput imaging system (Miller et al., 2017). The phenotypic data were analyzed using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) for R (R Core Team, 2019) based on the model $Y_{ijk} = \mu + E_i + r(E)_j + G_k + GE_{jk} + \varepsilon_{ijk}$, where Y_{ijk} was the response variable of the kth genotype (G) in the jth replication (r) nested in the ith environment (E). The residual error ε_{ijk} was assumed to be independent and following a normal distribution [$\sim$ iidN$(0, \sigma^2_\varepsilon)$]. Except for the grand mean (μ), all factors

TABLE 2 Thirty-six environments where the set of 31 inbreds were evaluated, their location, latitude (Lat), longitude (Long), plot area, planting density, planting and harvest dates, and irrigation status

ID	Location	Lat	Long	Plant density Plants ha^{-1}	Planting date[a]	Harvest date	Irrigation
DE1_14	Newark, DE	39.67	−75.75	86112	2014/05/13	2014/10/07	yes
GA2_14	Tifton, GA	31.43	−83.58	47075	2014/04/17	2014/08/25	yes
IA1_14	Ames, IA	41.99	−93.69	69445	2014/05/24	2014/09/30	no
IA2_14	Ames, IA	42.02	−93.77	69445	2014/05/09	2014/09/30	no
IA3_14	Ames, IA	42.01	−93.67	39541	2014/05/20	2014/11/10	no
IL1_14	Urbana, IL	40.06	−88.23	73810	2014/05/06	2014/09/30 (rep 1)	no
						2014/10/07 (rep 2)	
IN1_14	West Lafayette, IN	40.49	−87.01	123017	2014/05/30	2014/10/28	no
MN2_14	St. Paul, MN	45.00	−93.18	78284	2014/05/06	2014/10/10	no
NC1_14	Clayton, NC	35.67	−78.50	106224	2014/04/14	2014/08/29	yes
NE1_14	Lincoln, NE	40.83	−96.66	129167	2014/05/24	2014/10/15	no
NY1_14	Aurora, NY	42.73	−76.65	68890	2014/05/29	2014/12/08	no
PA1_14	Pennsylvania Furnace, PA	40.71	−77.96	71760	2014/05/31	2014/08/19	no
TX1_14	College Station, TX	30.55	−96.43	102514	2014/03/14	2014/07/25	yes
TX2_14	College Station, TX	30.55	−96.43	102514	2014/03/14	2014/07/29	no
WI1_14	West Madison, WI	43.06	−89.53	90644	2014/05/09	2014/10/29	no
DE1_15	Newark, DE	39.67	−75.75	86112	2015/06/16 and 2015/06/17	2015/10/21	yes
GA1_15	Tifton, GA	31.51	−83.56	64071	2015/04/22	2015/08/07	yes
IA1_15	Ames, IA	42.02	−93.77	69445	2015/05/13	2015/09/30	no
IA2_15	Ames, IA	42.01	−93.79	69445	2015/06/09	2015/10/07	no
IA3_15	Ames, IA	42.00	−93.66	53820	2015/05/19	2015/10/22	no
IA4_15	Ames, IA	42.04	−93.71	45264	2015/06/02	2015/11/20	no
IL1_15	Urbana, IL	40.06	−88.23	73810	2015/04/30	2015/10/01	no
IN1_15	West Lafayette, IN	40.48	−87.00	123017	2015/05/14	2015/11/03	no
KS1_15	Manhattan, KS	39.21	−96.60	71760	2015/04/23	2015/09/21	no
MN1_15	St. Paul, MN	44.99	−93.18	78284	2015/05/05	2015/10/26	no
MO1_15	Columbia, MO	38.90	−92.21	74750	2015/05/05	2015/09/23	no
MO2_15	Columbia, MO	38.90	−92.21	74750	2015/05/05	2015/09/23	no
NC1_15	Clayton, NC	35.67	−78.51	106224	2015/04/29	2015/09/16	yes
NY1_15	Aurora, NY	42.79	−76.65	68890	2015/06/05	approx. 2015/11/23	no
NY2_15	Aurora, NY	42.72	−76.66	68890	2015/05/24	2015/10/09	yes
SD1_15	New Underwood, SD	44.21	−102.93	51667	2015/05/22	2015/10/28	no
TX1_15	College Station, TX	30.55	−96.43	102514	2015/03/07	2015/07/27	no
TX2_15	College Station, TX	30.55	−96.43	102514	2015/04/02	2015/07/28	no
TX3_15	Halfway, TX	34.19	−101.95	80730	2015/04/25	2015/09/15	yes
WI1_15	West Madison, WI	43.06	−89.53	90644	2015/06/09	2015/10/09	no
WI2_15	Arlington, WI	43.33	−89.33	90644	2015/05/29	2015/10/09	no

[a]Dates in YYYY/MM/DD format.

were considered to be random effects. Each factor's significance was tested using the ranova function of the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) for R (R Core Team, 2019). Best linear unbiased predictions (BLUPs) of genotype performance were calculated from this model. Inbred line mean heritability across all environments was calculated as $h^2 = \frac{\sigma^2_G}{\sigma^2_G+(\sigma^2_{GE}/e)+(\sigma^2_\varepsilon/re)}$, where σ^2_G was the genotypic variance, σ^2_{GE} was the G × E variance, σ^2_ε was the error variance, for e number of environments and r number of replications per environment. The variance component values were also used in calculating the percent of phenotypic

variance explained by each factor of the model. Pearson correlations were calculated between all traits measured using BLUPs. Pearson correlation coefficients were also calculated between trait performance (in terms of BLUP values) and three stability parameters (Type II stability, mean square error, and stability from the GGE analysis described below). In both cases, correlations were considered significant if $P < .05$. To compare different concepts of stability, the Spearman rank correlation was calculated between the three types of stability (defined as the median stability value across all traits for each genotype).

To compare stability among groups of genotypes, we used two categorizations of the inbreds. First, we divided the group of inbreds based on the most typical heterotic split known in North American maize germplasm: the non-stiff stalk (NSS) versus stiff-stalk synthetic (SSS) heterotic pools. Inbreds from the SSS pool, with their history as the female parents of choice to provide enough seed to make hybrid production economical (Reif, Hallauer, & Melchinger, 2005) and ongoing improvement for relevant traits (Duvick, Smith, & Cooper, 2004) are expected to have better kernel production. Non-stiff stalks are generally any lines that did not fall into the SSS pool. Second, we looked at the recently expired Plant Variety Protection (ex-PVP) inbreds versus the public inbreds. With one exception (CM105), the public inbreds were released earlier (in or before 1972) than the ex-PVP inbreds and are considered to have a lower degree of selectness. To test for significant differences in the mean regression slope, Type II stability, MSE, and GGE stability values between the respective categories, *t*-tests were conducted in R (R Core Team, 2019). The stability of inbred lines in the Midwest environments (locations with longitude between −85° and −95°: Iowa, Illinois, Indiana, Minnesota, Montana, and Wisconsin in both years) was compared with their stability in all other environments (West, longitude west of −95°: locations in Kansas, Nebraska, South Dakota, and Texas in both years; and East, longitude east of −85°: locations in Delaware, Georgia, North Carolina, New York, and Pennsylvania in both years) using *t*-tests.

Several weather data parameters were calculated for each environment during three time periods (i) over the entire season, (ii) from planting date to the start of silking, and (iii) the period bracketing silking (14 d before and after silking). Total precipitation was calculated by summing the daily precipitation for each time period. Environments where irrigation was not tracked were excluded from analyses of the effect of precipitation on discriminability. Minimum, median, and maximum temperature during each time period was calculated. Additionally, the number of hours when temperature was greater than 30, 25, and 22°C or less than 22°C was calculated for each time period.

2.3 | Quantification of G × E interaction

2.3.1 | Regression stability analysis

Prior to performing the regression analysis, trait values were standardized using the following equation to allow comparison of MSE values across traits, which have different scales of measurement: $Y_{ijk_{\text{standardized}}} = [Y_{ijk} - mean(Y)]/sd(Y)$, where Y_{ijk} was the response variable of the kth genotype (G) in the jth replication (r) nested in the ith environment (E) and Y was the response of all genotypes in all replications in all environments. This standardization of each trait ensured that the value of MSE reflected variability and not the absolute scale of a given trait (e.g., anthesis with values in the 1000s of GDU versus kernel thickness which has values in the 1s of mm). Then, for each trait, values for each genotype were regressed on the environmental means of all genotypes (Finlay & Wilkinson, 1963) following the model $Y_{ij} = \mu_i + \beta_i I_j + \delta_{ij}$, where Y_{ij} is the mean phenotypic performance of genotype i in environment j; μ_i is the mean of genotype i across environments; β_i is the linear regression coefficient of Y_{ij} on I_j; I_j is the effect of environment j (i.e., the environmental index, or in this study specifically, the environmental mean); and δ_{ij} is the deviation of sY_{ij} from the regression fitted value of genotype i in environment j. The linear regression coefficient, β_i (i.e., slope of the regression line), was used to quantify Type II stability ($\beta_i = 1$) as the absolute value of slope minus one ($|\beta_i - 1|$), and the mean squared error (MSE) was used to quantify Type III stability. Type I ($\beta_i = 0$) was not examined since it is not useful for breeding efforts. To quantify trait sensitivity to G × E, the variance of slope and of MSE in the linear regression model were compared among traits.

2.3.2 | AMMI analysis and biplot

We used AMMI analysis and biplots to examine G × E sensitivity since it represents the genotype and environment factors. To conduct AMMI analysis and construct AMMI biplots, the *agricolae* package (de Mendiburu, 2019) for R (R Core Team, 2019) was used. Since the AMMI2 biplot, which depicts Principal Components 1 and 2 (PC1 vs. PC2), tends to represent much of the G × E pattern (Gauch, 1988). The total Euclidean distance between all pairs of genotypes and between all pairs of environments in this biplot was calculated as another quantification of trait sensitivity to G × E. As described above for the linear regression, the phenotypic data were standardized before running the AMMI analysis so that the Euclidean distance could be compared across traits.

2.3.3 | GGE analysis and biplot

We used GGE analysis and biplots, which examine genotype and G × E, to look at genotype stability and environment discriminability. Prior to performing the GGE analysis, trait values were standardized (as described for the linear regression analysis) to allow comparison of discriminability and stability values across traits. To prepare the data for GGE analysis, a mean value across replications for each genotype in each environment was calculated. Environments and genotypes with more than 30% missing data were removed from the data set. Subsequently, remaining missing data points were estimated based on the mean of the available data for the corresponding environment and genotype. The GGE models were created using the *GGEBiplots* package (Dumble, 2017) for R (R Core Team, 2019).

Each environment's discriminating power was quantified based on the "discriminating power vs. representativeness" view of the GGE biplot. In this type of plot, the length of the environment vector (from environment to origin) corresponds to the standard deviation of genotype means in the environment, which is a measurement of discriminability (Yan et al., 2007). We did not investigate representativeness since inbred lines, which were used in this study, may not give meaningful results on representativeness of environments for testing hybrids. For each trait, the five most discriminating and five least discriminating environments were compared based on their latitude, longitude, planting density, and weather data. These comparisons were tested for statistical significance using a *t*-test. Genotype stability was quantified based on the "mean vs. stability" view of the GGE biplot with genotypes lying near the AEC abscissa being considered as more stable than those farther from it. Custom R scripts were used to extract the length of the environment vector and the genotype distance from the AEC abscissa. Scripts for all the analyses described are available at https://github.com/cmfalcon/G2F_GxE_inbreds.

3 | RESULTS

3.1 | Variation in agronomic and yield-component traits

Across environments and genotypes, a wide range of phenotypic values were observed (Supplemental Figure S2, S3). Between 1.20–2.17-fold differences in mean values were observed among the agronomic and yield-component traits (Supplemental Table S1). In terms of BLUPs, on average, LH123HT was the last inbred to reach anthesis (1547.66 GDU), and Mo17 was the last to reach silking (1555.13 GDU). Across all environments, LH123HT was the tallest inbred (191.35 cm), and inbred 2369 had the greatest ear height (87.21 cm). Inbred 2369 had the greatest plot grain weight (263.20 g), and PHG39 had the greatest kernel weight (0.31 g) on average. C103 had the longest ears (167.25 mm), while PB80 had the widest ears (41.60 mm) across environments. Inbred 2369 had the greatest kernels per row (34.78), and LH82 had the greatest kernel row number (19.54). Inbred 740 had the greatest kernel area (69.40 mm^2), the longest kernels (10.59 mm), and the widest kernels (8.65 mm). PHG39 had the thickest kernels (5.98 mm). The strongest positive correlation among traits was observed between silking and anthesis (Pearson $r = .96$). The strongest negative correlation was observed between ear length and kernel weight (Pearson $r = -.70$; Supplemental Figure S4).

Inbred line mean heritability estimates ranged from .95 for plot grain weight to .99 for kernel row number (Supplemental Table S1). The environment term was significant ($P < .001$) for each trait and explained between 4.46–51.24% of the phenotypic variance. The genotype term was significant ($P < .001$) for each trait as well and explained between 15.58–62.76% of the phenotypic variance. For anthesis, silking, plant height, ear height, plot grain weight, and kernel length, environment explained the majority of the observed variation. For ear length, ear width, kernel row number, kernel weight, kernel area, and kernel width, genotype explained the majority of the observed variation. The G × E interaction term was significant ($P < .001$) for each trait and explained between 8.99 and 20.36% of phenotypic variance. It never explained the majority of variation. The factor replication nested within environment explained between 0.00–7.36% of the phenotypic variance, which was the smallest portion of phenotypic variance among all sources of variation. This term was significant ($P < .01$) for all traits except kernels per row, kernel row number, and kernel thickness. Residual error accounted for between 8.84–40.51% of the phenotypic variance and, in the case of kernel thickness and kernels per row, it represented the majority of observed phenotypic variance (Figure 1).

3.2 | G × E interaction: sensitivity of traits

Trait sensitivity was quantified four ways: (i) the variance of the slopes, (ii) the variance of the mean square errors (MSEs), and (iii) MSE per se from the linear regression model, as well as (iv) the Euclidean distances among points in the AMMI biplot. The median r^2 value (fit of the linear model) across all inbreds for each trait was between .14 for kernel width and .76 for anthesis (Supplemental Table S2). Based on linear regression, anthesis had the lowest median MSE of all traits (MSE = 0.202 and the smallest MSE variance, var = 0.009; Figure 2b; Supplemental Table S3), while silking had the smallest slope variance (var = 0.021; Figure 2a; Supplemental Table S4). Kernel weight had the greatest median MSE of all traits (MSE = 0.470; Figure 2b; Supplemental Table S3).

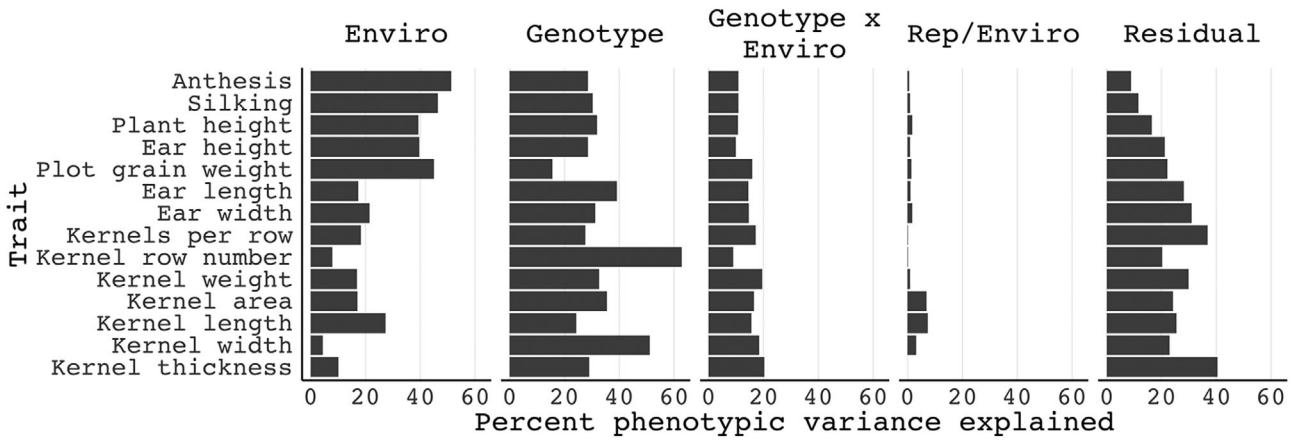


FIGURE 1 Percent of phenotypic variance explained by each analysis of variance model term fitted as random effects for 14 quantitative traits—two flowering traits (anthesis and silking), two height traits (plant and ear height), plot grain weight, four ear traits (ear length and width, kernels per row, and kernel row number), and five kernel traits (kernel weight, area, length, width, and thickness)—measured on 31 inbred lines in 36 environments. Enviro, environment, Rep, replications

Kernel thickness had the largest MSE variance (var = 0.291; Figure 2b; Supplemental Table S3) as well as the largest slope variance (var = 0.318; Figure 2a; Supplemental Table S4).

For the AMMI analysis, PC1 and PC2 together represented between 32.5–52.4% of the variation observed for each trait (Supplemental Table S5). In the AMMI biplots, the median Euclidean distance between pairs of genotypes was higher than the mean Euclidean distance between pairs of environments for all traits except kernels per row and kernel thickness, even in cases where the environment explained more of the observed phenotypic variance. The yield-component traits tended to show higher sensitivity to G × E compared to morphological and flowering time traits except for kernel row number, which had the lowest total Euclidean distance (1.07). The flowering date traits showed lower sensitivity to G × E measured as the total Euclidean distance (1.11 for anthesis and 1.15 for silking) as they did for the MSE and slope variance estimates of G × E sensitivity. Kernel length had the greatest total Euclidean distance (1.57) of any trait (Figure 3; Supplemental Table S6). Traits with a larger percent variance explained by G × E tended to have a larger total Euclidean distance (among genotypes plus among environments; Pearson $r = .83$).

3.3 | G × E interaction: discriminability of environments

Depending on the trait, the GGE biplot explained 57.2–84.6% of the total variation due to genotype and G × E (Supplemental Table S7). For each trait, the discriminability of each environment was quantified based on the length of the environment vector in GGE biplots, which corresponds to the standard deviation of genotype means in the environment. By this measure, the most discriminating environment was MN2_14 with a median discriminability value of 4.36 across traits. The least discriminating environment was NY1_14 with a median discriminability value of 1.05 across traits (Figure 4a). For the most and least discriminating environments for each trait, we report the rank of that environment based on the mean value across genotypes to contextualize the relationship of these environments to one another (Supplemental Table S8). Environments with high discriminability values for the yield-component traits tended to have low discriminability for flowering date traits, although this correlation was not significant (Figure 4b). Examining the five most and five least discriminating environments for each trait suggested that environments located in the Midwest showed high discriminability for more traits than environments to the West or East (Figure 5).

Comparing weather factors in the five least and five most discriminating environments suggested that the weather during silking was influential for more traits than the average weather over the whole season or from planting to silking. Over the whole season, no weather factors were significantly different between the least and most discriminating environments for any trait. From planting to silking, seven comparisons among the eight weather factors and 14 traits showed significant differences. Discriminability for anthesis and silking was influenced by the number of hours when temperature was greater than 25°C and the number of hours when temperature was greater than 22°C. Discriminability for kernel weight was influenced by median temperature, the number of hours when temperature was less than 22°C, and precipitation. During silking, 15 comparisons were significantly different between the least and most discriminating environments. Discriminability for flowering traits was influenced by median and maximum temperature, the number of hours

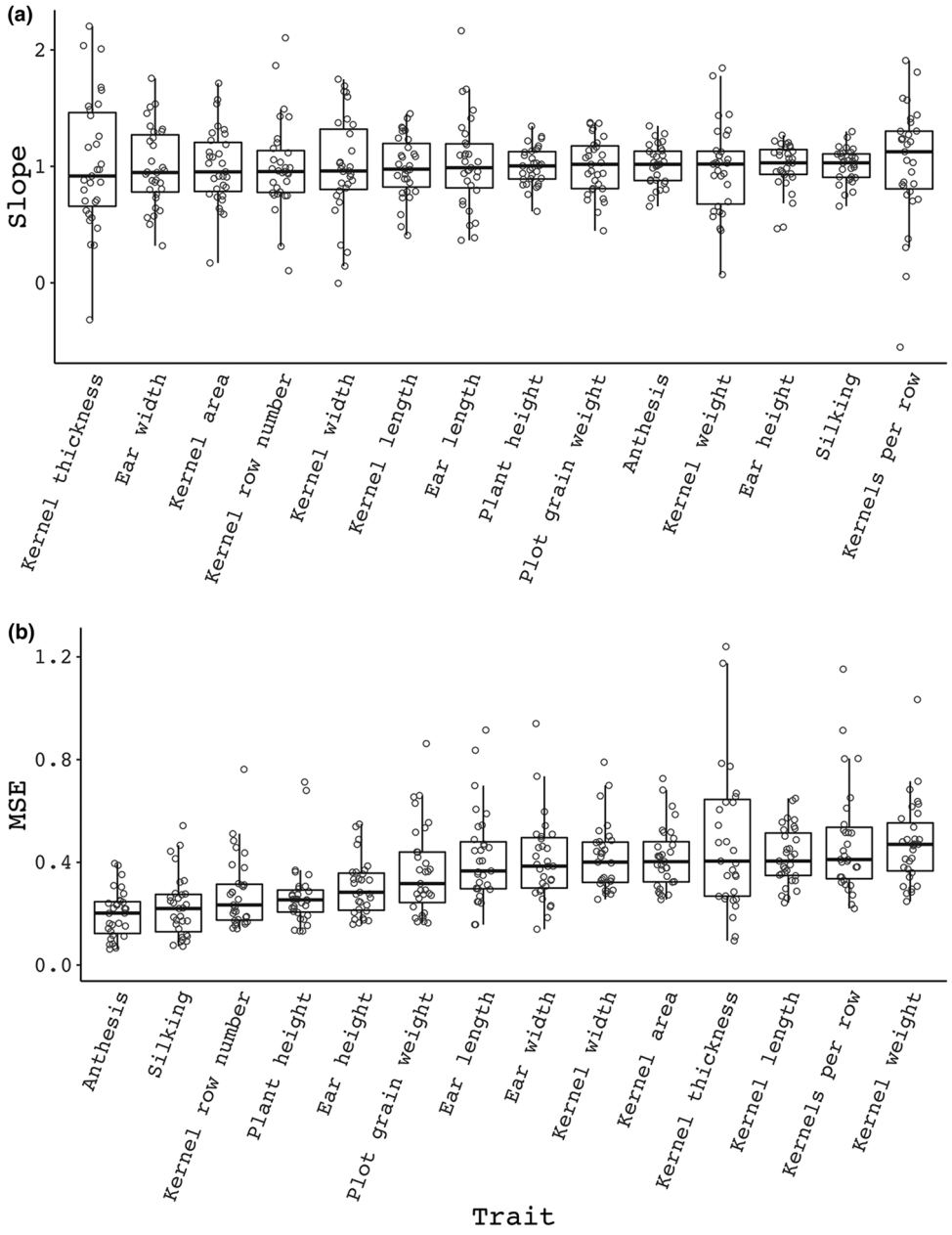


FIGURE 2 Boxplots of (a) slope and (b) mean square error (MSE) values from linear regression analysis for the evaluation of 31 inbreds across 36 environments for each of 14 traits, which were standardized by using the equation $Y_{ijk_{\text{standardized}}} = [Y_{ijk} - mean(Y)]/sd(Y)$ where Y_{ijk} was the response variable of the *k*th genotype (G) in the *j*th replication (*r*) nested in the *i*th environment (E) and *Y* was the response of all genotypes in all replications in all environments. The set of 31 inbreds was used to determine the environmental index used in the regression. The *y*-axis was truncated to display a less compressed view of the boxplots. As a result, several outlier data points are not included: kernel thickness, MSE = 1.96 and MSE = 2.66; ear width, MSE = 1.56; kernels per row, MSE = 1.45

when temperature was greater than 30, 25, and 22°C and less than 22°C, and precipitation during silking. Discriminability for ear width was influenced by minimum temperature and the number of hours when temperature was less than 22°C. Discriminability for kernel row number, kernels per row, kernel area, and kernel width was influenced by the number of hours when temperature was less than 22°C (Figure 6).

3.4 | G × E interaction: stability of genotypes

The stability of each genotype was quantified in terms of Type II stability and MSE from the linear regression model, and as the distance from the AEC abscissa in the GGE analysis. The values for these different stability parameters showed a wide range of variation. Slope values ranged from −0.55–2.20 across genotypes and traits. Likewise, MSE ranged from

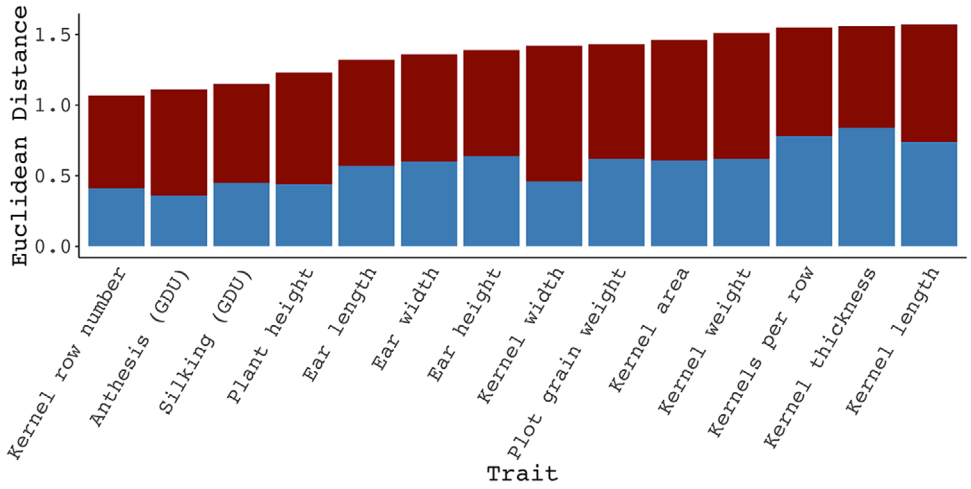


FIGURE 3 Stacked barplot of median Euclidean distance between all pairs of genotypes (red) and between all pairs of environments (blue) from additive main effects and multiplicative interactions Principle Component 1 vs. Principle Component 2 biplots for each of 14 traits, which were standardized by using the equation $Y_{ijk}{}_{\text{standardized}} = [Y_{ijk} - mean(Y)]/sd(Y)$ where Y_{ijk} was the response variable of the *k*th genotype (G) in the *j*th replication (*r*) nested in the *i*th environment (E) and *Y* was the response of all genotypes in all replications in all environments. The set of genotypes includes 31 maize inbred lines evaluated across 36 environments. Along the *x*-axis, traits are sorted by the total Euclidean distance (i.e., among genotypes plus among environments)

0.06–2.66, and the GGE stability value ranged from 0–5.86. As the GGE stability is most appropriately interpreted in relation to the genotypes' mean performance values, we have reported the rank (based on mean performance value) for the most and least stable (based on GGE stability) genotypes to contextualize this measure of stability (Supplemental Table S9). While more patterns of stability versus mean performance show up among the traits in this study than we can concisely identify here, we recognize that GGE stability is more meaningful in situations where the genotypes exhibit similar mean performance. Each of these quantifications represent different concepts of stability. In fact, the Spearman rank correlations between these three types of stability (defined as median stability across all traits for each genotype) were all low (Type II and MSE: $\rho = .19$; Type II and GGE: $\rho = .31$; MSE and GGE: $\rho = .41$) and nonsignificant at the $\alpha = .05$ level. Type II and GGE stability are similar in that they indicate a genotype that has ranked consistently across environments, but they were not strongly correlated in our study.

Type II stability and MSE were each significantly correlated with the BLUP values of genotype performance for four traits. Pearson *r*-values ranged from −0.49–0.52 for the correlation of the BLUP values and Type II stability and from −0.41–0.77 for the correlation of the BLUP values and MSE. The correlation of Type II stability and the BLUP values is explained by the fact that Type II stability is the directional responsiveness of the genotype to the environmental index. The GGE stability showed significant correlation with the BLUP values for only one trait, kernel thickness (Pearson $r = .42$; Table 3). These results suggest that performance and sensitivity to differential environmental influences are not correlated for most traits.

Based on the slope of each genotype in the linear regression model (Supplemental Table S10), PHN82 exhibited the highest Type II stability (mean slope value across traits = 1.00), and PHG47 exhibited the lowest (mean slope value across traits = 0.59). Based on the MSE from the linear regression model (Supplemental Table S11), PHG39 was the least stable genotype (mean MSE value across traits = 0.61), and LH198 was the most stable (mean MSE value across traits = 0.23) across all traits. Using the GGE biplot analysis to quantify stability, C103 was the most stable genotype (median stability value across traits = 0.43), and B37 was the least stable genotype (median stability value across traits = 1.46) across all traits (Figure 7a). Stability ranking differed for each trait for all of these methods. Even within types of traits (i.e., flowering, height, or yield-component traits), correlations of stability rankings were nonsignificant or weak (Figure 7b; Supplemental Figure S5). When considering kernel thickness, the trait that showed the greatest variance for both slope and MSE (Supplemental Tables S4 and S5), LH74 line showed the highest (slope = 1.02) and W117 showed the lowest (slope = −0.32) Type II stability (Supplemental Table S10), whereas PHB47 was the most stable (MSE = 0.09) and B37 was the least stable (MSE = 2.66) based on MSE values (Supplemental Table S11). For GGE analysis, ear width showed the most variance in stability value (Supplemental Table S12), PHZ51 line was the most stable (stability value = 0.04), and B37 was the least stable (stability value = 5.86) in this case (Supplemental Table S13).

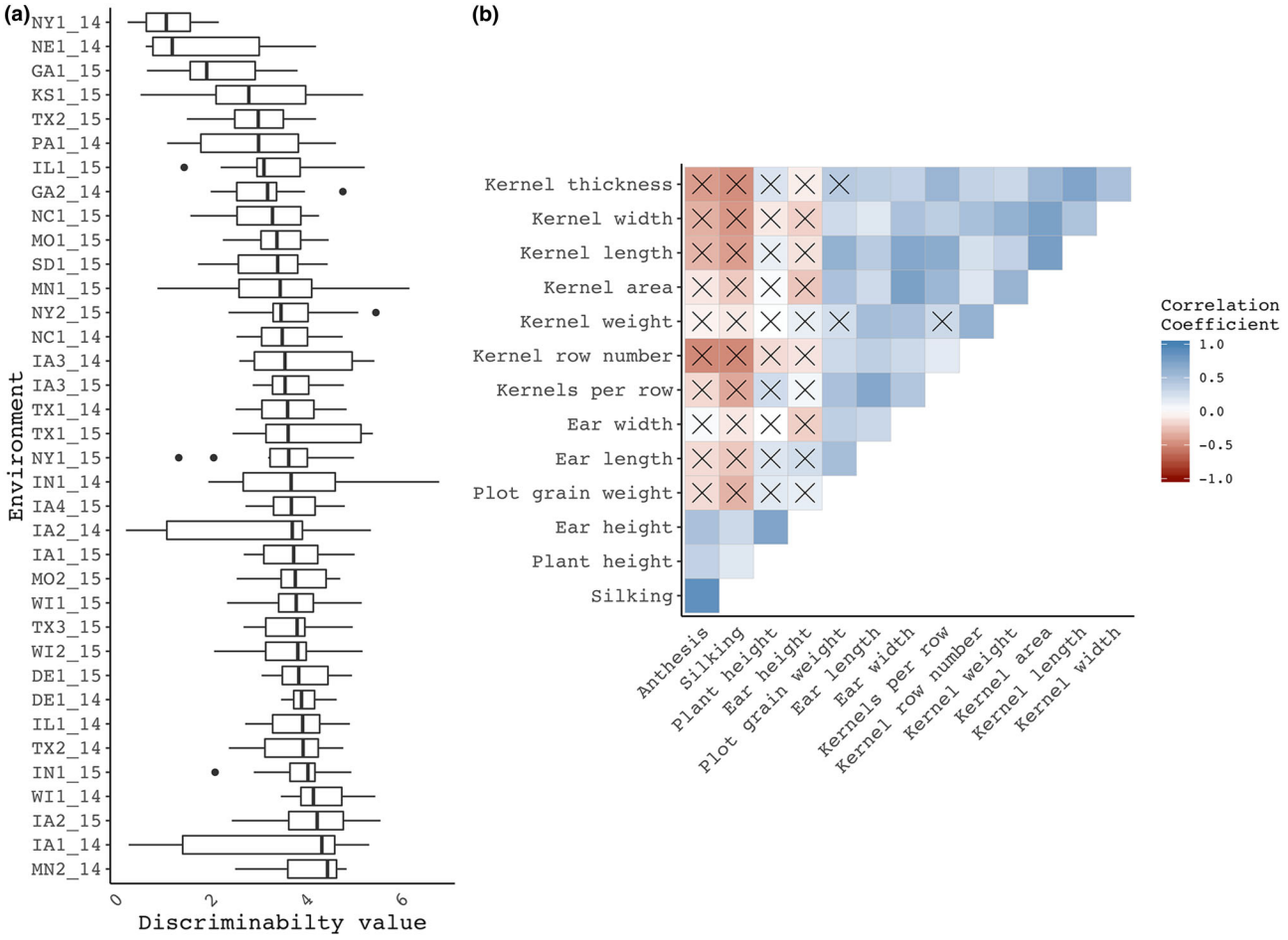
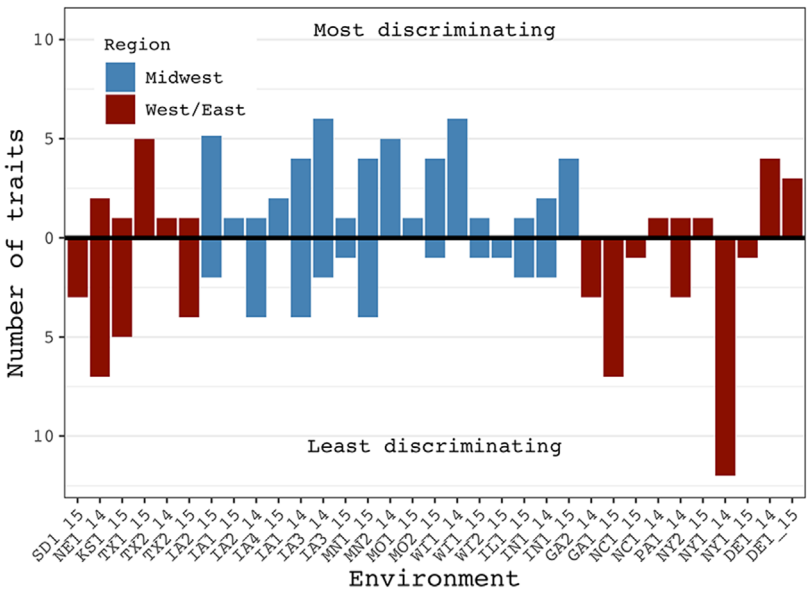


FIGURE 4 Each environment's discriminating power was quantified based on the "discriminating power vs. representativeness" view of the genotype main effects and genotype × environment interaction (GGE) biplot. (a) GGE discriminability of 14 traits (which were standardized using the equation where Y_{ijk} was the response variable of the kth genotype (G) in the jth replication (r) nested in the ith environment (E) and Y was the response of all genotypes in all replications in all environments) evaluated for 31 inbreds for across the 36 environments used in this study ordered by the median discriminability. Lower discriminability values (x-axis) indicate lower discriminability while higher values indicate greater discriminability. (b) Rank correlations among discriminability values for each trait. The symbol " × " marks nonsignificant ($\alpha = .05$) correlations

FIGURE 5 Number of traits out of 14 evaluated for which an environment was among the five least (below the x-axis) or five most (above the x-axis) discriminating environments. This study included 31 maize inbred lines evaluated across 36 environments in the United States. Environments are arranged along the x-axis by longitude from west to east with Midwest environments (locations with longitude between −85° and −95°: Iowa, Illinois, Indiana, Minnesota, Montana, and Wisconsin) indicated in blue and other environments (West, longitude west of −95°, and East, longitude east of −85°:locations in Delaware, Georgia, Kansas, North Carolina, Nebraska, New York, Pennsylvania, South Dakota, and Texas) indicated in red. Environments not shown were not among the five least or most discriminating environments for any trait



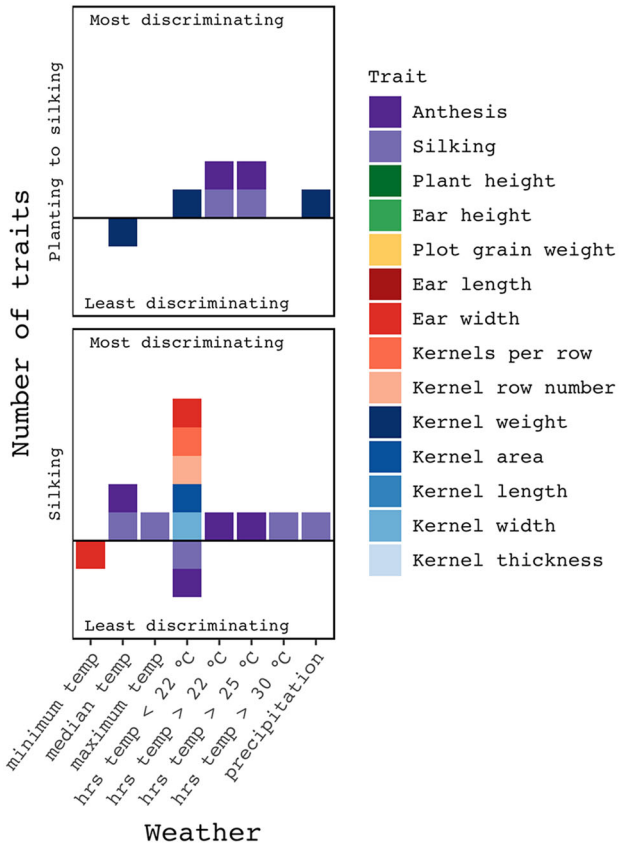


FIGURE 6 Summary of significant differences in weather factors (minimum, median, and maximum temperature; number of hours when temperature was less than 22°C or was greater than 22, 25, or 30°C; and precipitation) between the five least and five most discriminating environments for each of 14 traits evaluated. The evaluation included 36 different environments in the United States. Statistical significance was determined by a *t*-test ($P < .05$). Bars are placed above the *x*-axis where the weather factor was greater in the most discriminating environment or below the *x*-axis where the weather factor was greater in the least discriminating environments. Height of the bars illustrates the number of traits significantly greater in the most or least discriminating environment for each weather factor

Having observed that the Midwest locations (as defined by longitude) tended to show higher discriminability, we examined the stability of inbred performance in the Midwest versus other regions (West and East). Type II stability was significantly lower in the Midwest for plant and ear height as well as plot grain weight and ear length and was significantly higher in the Midwest for kernel row number. With significantly lower MSE, the inbreds were more stable in the Midwest for ear length, kernel row number, kernel weight, kernel area, kernel length, and kernel width. For GGE stability, the inbreds were significantly more stable in the Midwest for plant height, kernel row number, and kernel width. However, for anthesis, silking, and ear length, the inbreds were less stable in the Midwest (Figure 8). The effect of latitude on stability was also examined, but no pattern was identified.

TABLE 3 Significant ($P < .05$) Pearson correlations between best linear unbiased prediction (BLUP) values for 31 inbreds evaluated across 36 environments and stability parameters for each trait evaluated. Nonsignificant correlations indicated by "ns."

Trait	Type II[a]	MSE[b]	GGE
Anthesis	ns	ns	ns
Silking	ns	ns	ns
Plant height	ns	ns	ns
Ear height	−0.44	ns	ns
Plot grain weight	−0.49	ns	ns
Ear length	ns	ns	ns
Ear width	ns	ns	ns
Kernels per row	ns	−0.41	ns
Kernel row number	0.52	0.48	ns
Kernel weight	ns	0.54	ns
Kernel area	ns	ns	ns
Kernel length	−0.37	ns	ns
Kernel width	ns	ns	ns
Kernel thickness	ns	0.77	0.42

[a]Type II stability calculated as the absolute value of slope – 1
[b]MSE, mean square error; GGE, genotype main effect and genotype × environment interaction.

3.5 | Comparing stability among categories of germplasm

Comparing the NSS to SSS inbred types, silking showed a significant difference for Type II stability ($P < 0.05$) with the SSS inbreds being more stable (slope closer to 1). No trait showed a significant difference for MSE or for the GGE quantification of stability. For the comparison of public and ex-PVP genotypes, kernel width showed a significant difference in Type II stability ($P < 0.05$), with the public inbreds being more stable (slope closer to 1). Ear length showed a significant difference for MSE with the ex-PVP inbreds having lower MSE (greater Type III stability). For the GGE quantification of stability, no trait showed a significant difference (Supplemental Table S14).

4 | DISCUSSION

4.1 | Presence of G × E interaction

In this study, we measured 14 traits for 31 genotypes in 36 environments. Inbred line mean heritability estimates in this study were very high, likely due to the fact that data were collected from many environments. By studying inbreds, we were able to maximize the diversity that we considered and more easily interpret the G × E patterns since we were not looking at the average performance of two inbred parents via hybrid genotypes. Genotype × environment interaction

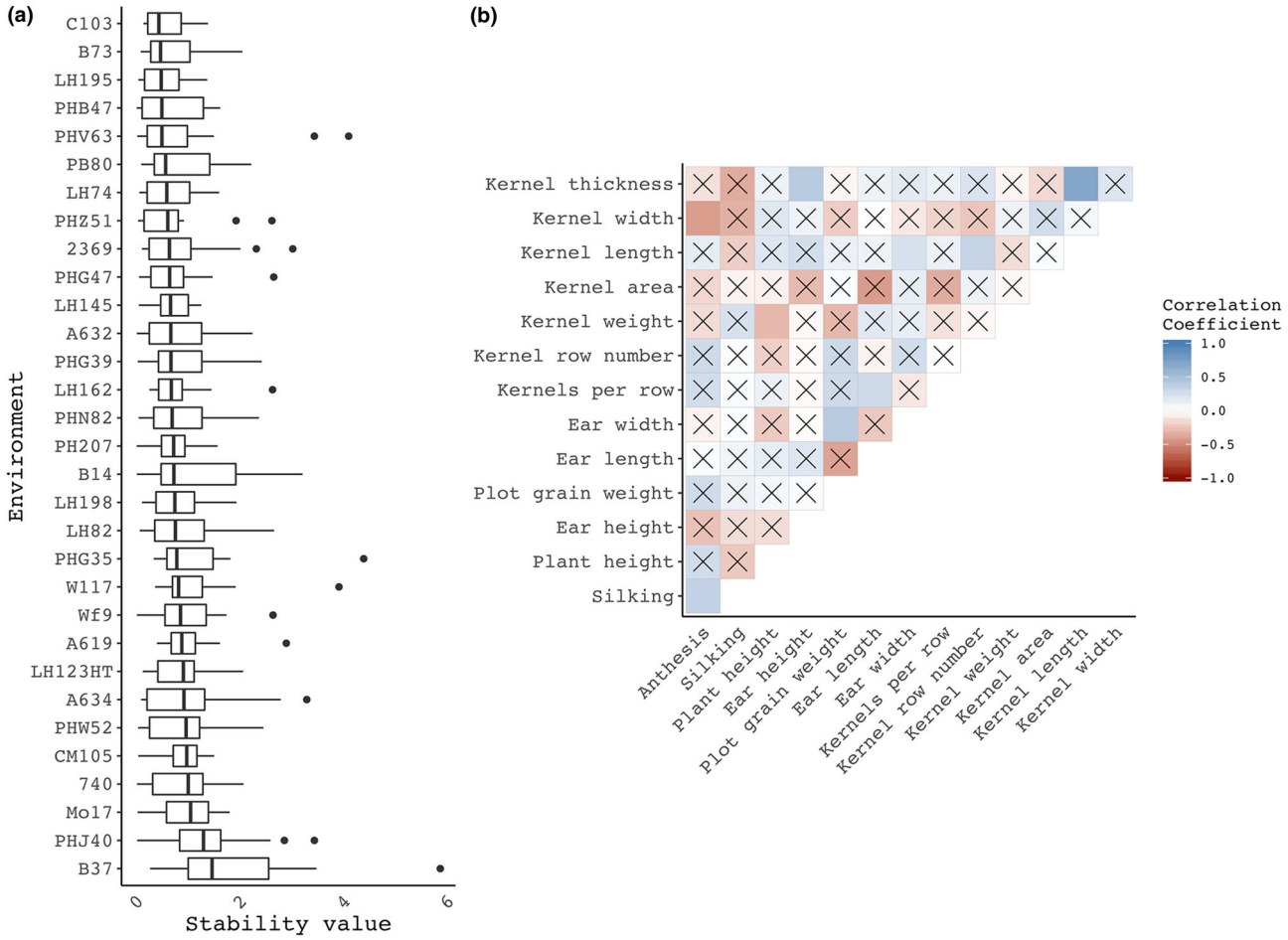


FIGURE 7 Each genotype's stability was quantified based on the "mean vs. stability" view of the genotype main effects and genotype × environment interaction (GGE) biplot. (a) GGE stability of the 14 traits used in this study, which were standardized using the equation $Y_{ijk_{\text{standardized}}} = [Y_{ijk} - mean(Y)]/sd(Y)$ where Y_{ijk} was the response variable of the kth genotype (G) in the jth replication (r) nested in the ith environment (E) and Y was the response of all genotypes in all replication in all environments, for each of the 31 genotypes in this study ordered by median stability. Lower stability values (x-axis) indicate greater stability while larger stability values indicate lower stability. (b) Rank correlation among stability values for each trait

was present at some level for all traits, with the percent variance explained by this factor ranging between 8.99–20.36%. The flowering date traits, anthesis and silking, exhibited a smaller effect of G × E, likely because of the more important main effect of environment for these traits. Plant height and ear height exhibited a low-level G × E effect with a balance of environment, genotype, and G × E effects on these traits. Kernel row number also exhibited a low level of G × E effect with a more prevalent effect of genotype. The remaining yield-component traits showed higher G × E effects with varying levels of genotype and environment effects. The pronounced effect of environment for many of these traits may be due to the large number and geographic spread of environments in this study. Similarly, an experiment of 835 maize hybrids grown at 21 locations found that the wide range of climatic conditions across the environments resulted in a large proportion of variance, between 42–74%, being explained by the environment term (Gage et al., 2017). The small effect of replications is favorable as it indicates consistent performance across the two replications in each environment. The residual error ranged widely with low values for the flowering date traits and higher values for the height and yield-component traits. This may be due to the variable goodness of fit of the linear model for different traits as well as the accuracy and precision of measurement for each trait. Residual error represented the majority of observed phenotypic variance for kernel thickness and kernels per row. Such large residual errors for these traits may be due to low accuracy in the method of measuring kernel thickness (and therefore kernels per row since it is derived from kernel thickness) as well as the tendency of inbreds to exhibit more variability due to poor seed fill. Importantly, however, the relative effects of each model factor (environment, genotype, G × E, and replicates nested within environment) depends on the environments and

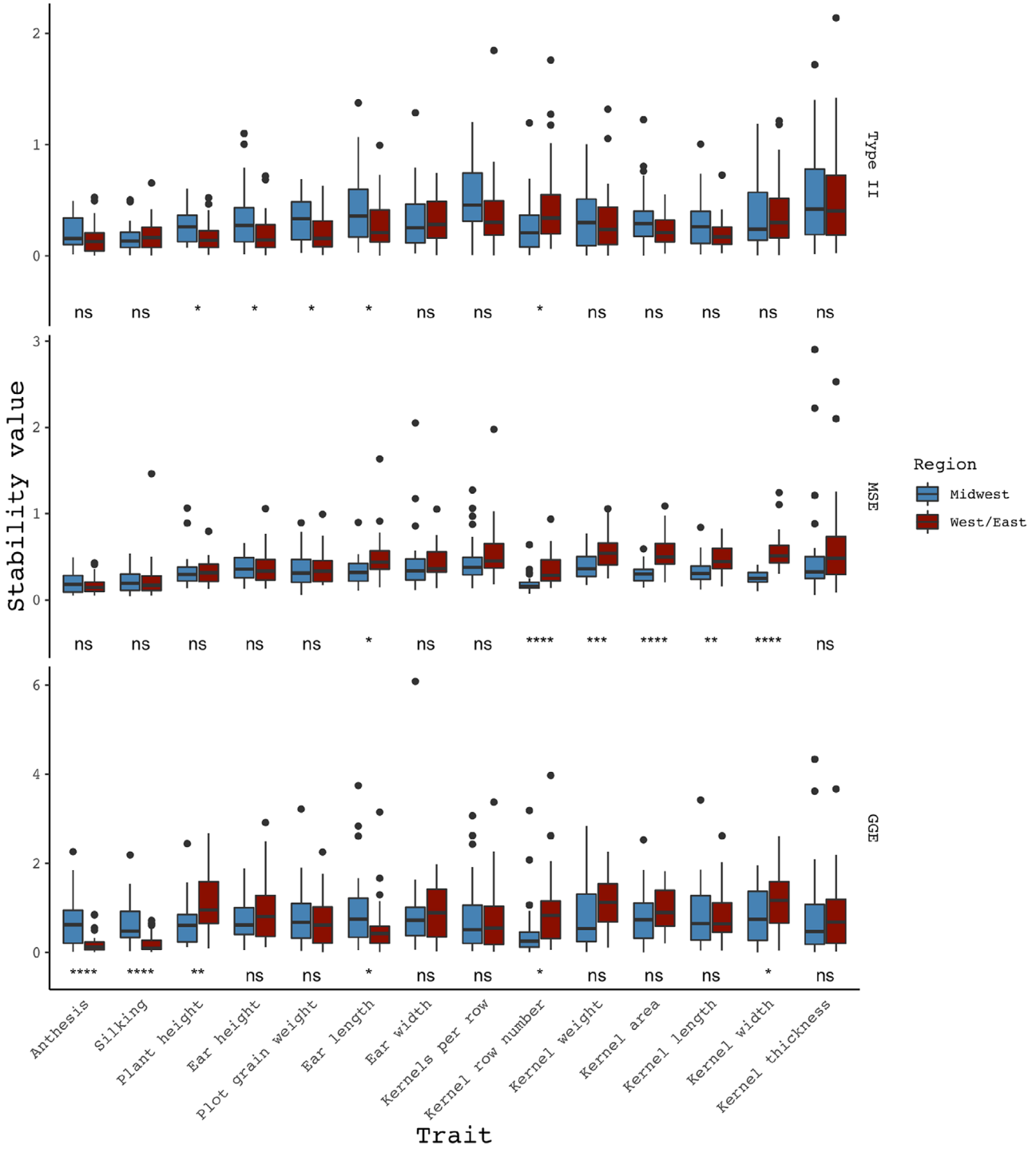


FIGURE 8 Boxplots of Type II stability (absolute value of slope – 1), mean square error (MSE), and the stability value calculated from the genotype main effects and genotype × environment interaction (GGE) biplot for 14 traits evaluated in 31 maize inbred lines across 36 environments in the United States. *t*-tests were used to identify significant differences in the stability values in the Midwest environments (locations with longitude between −85° and −95°: Iowa, Illinois, Indiana, Minnesota, Montana, and Wisconsin), which are indicated in blue, versus the West and East environments (West, longitude west of −95°; East, longitude east of −85°:locations in Delaware, Georgia, Kansas, North Carolina, Nebraska, New York, Pennsylvania, South Dakota, and Texas), which are indicated in red. Significance level for the *t*-test is shown below each pair of boxes: ns, nonsignificant, where $P > 0.05$; *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$

genotypes examined and will vary in other experiments. Overall, by studying several different types of traits in a large number of environments, we were able to capture the variable effect of G × E.

4.2 | Sensitivity of traits to G × E

Yield-component traits tended to show greater G × E sensitivity than flowering or height traits. The lower G × E sensitivity of flowering and height traits may owe to their higher heritability (Buckler et al., 2009; Peiffer et al., 2014; Wang, Yao, Zhang, & Zheng, 2006). Of the yield-component traits, kernel row number was the exception as it exhibited lower values for the various measures of G × E sensitivity. This is likely because of the large genetic effects for this trait, both in our study and in previous work (Hung et al., 2012). Potential kernel row number is determined earlier in the plant's development, but the realized kernel row number is influenced by stress during the silking period, which can lead to kernel abortion (Bänziger, Edmeades, & Lafitte, 2002; Below, Cazetta, & Seebauer, 2000; Gallais & Coque, 2005; Jacobs & Pearson, 1991; Reed, Singletary, Schussler, Williamson, & Christy, 1988; Uhart & Andrade, 1995a; 1995b). The low G × E sensitivity for kernel row number in our study may suggest that the trials had consistent levels of stress (which was likely low considering we did not aim to induce stress). Overall, given these results on trait sensitivity to G × E, including detailed ear and kernel yield component traits in future studies of G × E should improve detection of environmental stress and G × E.

4.3 | Discriminability of environments

Different groups of environments were quantified as more discriminating depending on the specific trait in question. For example, MN2_14 was the most discriminating environment based on the median GGE discriminability value across traits. However, while it was the third most discriminating environment for the yield-component traits, it was ranked 16th for the remaining traits (anthesis, silking, and plant and ear height). In cotton, environment discriminability and representativeness were assessed using GGE analysis and biplots for two traits (Blanche & Myers, 2006). The authors found that different environments were the most desirable depending on the trait. Their approach to reconciling this issue was to determine the most desirable location for an index trait comprising weighted values of the two traits they considered. A study in sugarcane also used GGE analysis and biplots to study environment discriminability and representativeness for three traits (Glaz & Kang, 2008). Here, the most discriminating environments overlapped across traits (cane yield, sucrose yield, and total recoverable sucrose), possibly because of the relatedness of these traits as sucrose yield is calculated from cane yield and total recoverable sucrose.

Thus, as has been noted for other crops, no single environment can be described as *the* most discriminating, as this characteristic differs by trait. Related traits such as those in the sugarcane study or the yield-components in this study show more similarity in discriminability patterns across environments. Since the option of testing at the most desirable environment for each separate trait is usually limited by resources, breeders may wish to choose the most discriminating test environments for the most important trait or based on the results for an index that combines traits of interest. Further research is needed to determine the utility of indices in this context and how to weight traits in those indices. Another way that future research might frame the practical utility of discriminating environments is to consider whether selection in non-target environments with high discriminability translates to gains in a target environment with low discriminability.

4.4 | Differences between least and most discriminating environments

Midwest locations (based on longitude) tended to be in the top five most discriminating environments for more traits than environments in the West or East, while environments in the West and East tended to be among the five least discriminating environments more often. Twenty-six of the 31 inbreds in this study were originally developed in the Midwest. In future research, genotypes from more varied origins should be involved to investigate the patterns of discriminability for germplasm that was not developed in the region of the environments studied.

Similarly, a study of discriminability in sugarcane observed that organic soil locations were more discriminating than the sand soil environment with regard to genotypes that had been developed by testing mostly in organic soil locations (Glaz & Kang, 2008). These results suggest that environments are more discriminating of lines that are well-adapted to them. Stated differently, locations beyond the target set of environments for a set of lines may exhibit lower discriminating power among them. Practically speaking, this observation may factor into the results of response to selection for tropical-to-temperate phenological adaptation in maize (Teixeira et al., 2015). In that study, the authors observed greater gain from selection for flowering time adaptation in the original location of selection (Iowa) than other locations where the multigenerational population was evaluated. The more general conclusion that selecting in a particular environment is the best way to achieve adaptation to that environment has been frequently described in the literature on correlated response to selection (Simmonds, 1991).

Weather during silking was significantly different between the least and most discriminating environments in more cases than weather throughout the season or weather from planting to silking. In particular, weather factors during silking were significantly associated with the discriminability of environments for ear size, kernel number, and kernel size traits. These results are consistent with previous findings that stress during the period bracketing silking intensified ovule, kernel, and ear abortion by influencing the rate of dry matter accumulation and partitioning (Early, McIlrath, Seif, & Hageman, 1967; Echarte & Tollenaar, 2006; Echarte, Andrade, Vega, & Tollenaar, 2004; Edmeades & Daynard, 1979; Grant, Jackson, Kiniry, & Arkin, 1989; Jacobs & Pearson, 1991; Lee & Tollenaar, 2007; Nafziger, 2009; Prine, 1971; Reed et al., 1988; Tollenaar, Dwyer, & Stewart, 1992; Uhart & Andrade, 1995a; 1995b). By affecting each genotype's ability to achieve its potential number of ovules and kernels, greater stress during this influential period likely diminishes the discriminability among genotypes in that environment.

4.5 | Stability of genotypes

Stability of genotypes varied by trait, showing little discernible pattern, even within categories of related traits. Similarly, quality protein maize studies observed variation in slope and mean sum of square of deviations from regression across several traits (Pixley & Bjarnason, 2002; Zaidi et al., 2008). For instance, Pixley and Bjarnason (2002) observed that the open-pollinated cultivars were least responsive (of the cultivar types studied) to environmental potential for grain yield, but were also the most responsive for endosperm modification score. However, these studies suggested more similarity in stability parameter values among related traits. For example, tryptophan and lysine in protein both had low slope values compared to grain yield, protein in grain, tryptophan in grain, and lysine in grain (Zaidi et al., 2008). Taken together, these observations suggest that breeding endeavors for different traits need to assess G × E for each trait independently rather than assuming that a genotype that is stable for one trait will also be stable for others. Investigating the stability for different traits is also useful since the desirability of stability depends on the trait in question. For example, to achieve ideal quality protein maize cultivars, Pixley and Bjarnason (2002) sought lines that would respond favorably to environment and have high yield, protein content, and protein quality in all environments ($\beta_i = 1$, high mean) while also maintaining low endosperm modification score even at unfavorable sites ($\beta_i < 1$, low mean).

In most cases, performance was not correlated with any measure of stability. A previous study of 23 phenotypes in maize showed that candidate genes for mean performance and for linear and nonlinear plasticity (i.e., slope and MSE from the Finley–Wilkinson regression) were structurally and functionally distinct (Kusmec, Srinivasan, Nettleton, & Schnable, 2017). If different genetic regions control trait performance as opposed to trait stability, it may be easier to exploit G × E while also breeding for improved performance. Results for which genotypes were most stable for each trait differed by the definition of stability being considered. In practice, breeders would focus on whichever type of stability was most important for their specific goals. For instance, they would focus on type II stability if breeding for a wide area, but MSE would be more important in breeding for a more localized region.

We found that stability was significantly different in Midwest (more discriminating) versus non-Midwest (less discriminating) locations for several traits. While we would have expected to see greater stability in the Midwest locations where most of the inbreds were originally developed, Type II stability was actually lower for plant height, ear height, plot grain weight, and ear length. The greater deviation from a slope of one for these traits may signal that while inbreds were not stable, they were more responsive to the best environments for these traits. For plant height, ear height, and plot grain weight, the median slope value across genotypes was greater than one, which indicates a better response to favorable environments. Type III stability in terms of MSE performed as we would have expected with greater stability for several kernel number and kernel size traits in the Midwest locations. For GGE stability, several traits—plant height, kernel row number, and kernel width—were more stable in the Midwest, as expected. However, the two flowering traits, anthesis and silking, were less stable in the Midwest. This lower stability may be due to the inbreds' wider range of flowering dates in environments where they are adapted in comparison to a flatter response in less suitable environments. This result reinforces our suggestion that future studies match the germplasm's origins to the target population of environments. Future studies should also consider how hybrids' stability patterns may differ from those found in this study of inbreds since inbreds are developed for their performance in hybrids.

For most traits, stability was not significantly different between NSS and SSS inbreds or between the ex-PVP and public inbreds. We would have expected the SSS inbreds to show greater stability for yield-component traits, owing to their history of being used as female parents. However, the SSS inbreds did not show better stability (Type II, MSE, or GGE) for any yield-component trait. We expected the ex-PVP inbreds to exhibit better stability due to their higher degree of selectness in comparison to the public inbreds. However, only two comparisons of stability were significantly different between these groups of inbreds. This result pointed out that while these inbreds are more highly selected, they were not selected for stability or performance as inbreds but as hybrid parents. As such, future research should investigate the performance and stability of hybrids created from these inbreds.

5 | CONCLUSIONS

In this experiment, we focused on questions of how to improve future studies of G × E, especially in terms of which traits are most informative. Given our results, future studies of G × E should include yield-component traits, which were more sensitive to G × E relative to flowering date and height traits. The discriminating power of environments varied by trait, so the utility of determining discriminability for an index trait should be investigated. Weather factors during the silking period were significantly different between the most and least discriminating environments more often than weather across the season or during the period from planting to silking. These differences may aid in identifying environments that, over seasons, tend to have appropriate weather patterns to achieve maximum discriminability. We found that the most discriminating environments were in the Midwest for more traits, and that stability was higher in the Midwest environments as well. Most of the inbreds we tested originated in the Midwest. Thus, we propose that future work be done to investigate whether these patterns are different when germplasm from more varied origins are involved. We did not observe the expected differences in stability between SSS versus NSS inbreds and in public versus ex-PVP inbreds, highlighting the importance of considering hybrid performance and stability in future G × E studies.

ACKNOWLEDGMENTS

The authors thank the entire G2F Consortium for their help with this study: Emily Rothfusz and Jane Petzoldt from the University of Wisconsin-Madison for assisting with coordination of the project; Dustin Eilert and Marina Borsecnik for assistance organizing and conducting field trials at the University of Wisconsin-Madison; and Miriam Lopez, Grace Kuehne, and Sarah Hennings from the USDA-ARS Corn Insects and Crop Genetics Research Unit at Iowa State University. We also acknowledge contributions from field manager, Nicholas Kaczmar, at Cornell University. This project was supported by the Iowa Corn Promotion Board, the Nebraska Corn Board, the Minnesota Corn Research and Promotion Council, the Illinois Corn Marketing Board, the Wisconsin Corn Promotion Board, and the National Corn Growers Association. Thanks to USDA-NIFA Hatch for supporting a large number of the cooperators. USDA-ARS base funds were also provided for ESB, JE, SF-G, JBH, JK, NL, and BS. Support was also provided through the Corn Promotion Board Endowed Chair in Maize Genetics for PSS. We thank the support of the USDA Hatch under accession numbers 1003971 for MAG and NSF grant IOS-1444456 that supports EPS. We thank the Texas Corn Producers Board and Texas A&M AgriLife Research for support for SCM.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

AlKhalifah, N., Campbell, D. A., Falcon, C. M., Gardiner, J. M., Miller, N. D., Romay, M. C., … Lawrence-Dill, C. J. (2018). Maize Genomes to Fields: 2014 and 2015 field season genotype, phenotype, environment, and inbred ear image datasets. *BMC Research Notes*, *11*(1), 452. https://doi.org/10.1186/s13104-018-3508-1

Austin, D. F., & Lee, M. (1996). Comparative mapping in $F_{2:3}$ and $F_{6:7}$ generations of quantitative trait loci for grain yield and yield components in maize. *Theoretical and Applied Genetics*, *92*(7), 817–826. https://doi.org/10.1007/BF00221893

Austin, D. F., & Lee, M. (1998). Detection of quantitative trait loci for grain yield and yield components in maize across generations in stress and nonstress environments. *Crop Science*, *38*(5), 1296–1308. https://doi.org/10.2135/cropsci1998.0011183×003800050029x

Bänziger, M., Edmeades, G. O., & Lafitte, H. R. (2002). Physiological mechanisms contributing to the increased N stress tolerance of tropical maize selected for drought tolerance. *Field Crops Research*, *75*(2–3), 223–233. https://doi.org/10.1016/S0378-4290(02)00028-X

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Below, F. E., Cazetta, J. O., & Seebauer, J. R. (2000). Carbon/Nitrogen Interactions During Ear and Kernel Development of Maize. In M. Westgate & K. Boote (Eds.), *Physiology and Modeling Kernel Set in Maize* (pp. 15–24). CSSA Spec. Publ. 29 Madison, WI: CSSA and ASA. https://doi.org/10.2135/cssaspecpub29.c2.

Bernardo, R. (2010). *Breeding for Quantitative Traits in Plants* (2nd ed.). Woodbury, Minnesota: Stemma Press.

Blanche, S. B., & Myers, G. O. (2006). Identifying discriminating locations for cultivar selection in Louisiana. *Crop Science*, *46*(2), 946–949. https://doi.org/10.2135/cropsci2005.0279

Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., … McMullen, M. D. (2009). The genetic architecture of maize flowering time, *Science*, *325*(5941), 714–718. https://doi.org/10.1126/science.1174276.

Dumble, S. (2017). GGEBiplots: GGE Biplots with "ggplot2". R package version 0.1.1. Retrieved from https://CRAN.R-project.org/package=GGEBiplots

Duvick, D. N., Smith, J. S. C., & Cooper, M. (2004). Long-term Selection in a Commercial Hybrid Maize Breeding Program. *Plant Breeding Reviews: Part 2: Long-term Selection: Crops, Animals, and Bacteria*, *24*, 109–151. https://doi.org/10.1002/9780470650288.ch4

Early, E. B., McIlrath, W. O., Seif, R. D., & Hageman, R. H. (1967). Effects of shade applied at different stages of plant development on corn (*Zea mays* L.) production. *Crop Science*, *7*(2), 151. https://doi.org/10.2135/cropsci1967.0011183×000700020018x

Echarte, L., Andrade, F. H., Vega, C. R. C., & Tollenaar, M. (2004). Kernel number determination in Argentinean maize hybrids released between 1965 and 1993. *Crop Science*, *44*(5), 1654–1661. https://doi.org/10.2135/cropsci2004.1654

Echarte, L., & Tollenaar, M. (2006). Kernel set in maize hybrids and their inbred lines exposed to stress. *Crop Science*, *46*(2), 870–878. https://doi.org/10.2135/cropsci2005.0204

Edmeades, G. O., & Daynard, T. B. (1979). The relationship between final yield and photosynthesis at flowering in individual maize

plants, *Canadian Journal of Plant Science*, *59*(3), 585–601. https://doi.org/10.4141/cjps79-097

Fan, X. M., Kang, M. S., Chen, H., Zhang, Y., Tan, J., & Xu, C. (2007). Yield stability of maize hybrids evaluated in multi-environment trials in Yunnan, China. *Agronomy Journal*, *99*(1), 220–228. https://doi.org/10.2134/agronj2006.0144

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

Gallais, A., & Coque, M. (2005). Genetic variation and selection for nitrogen use efficiency in maize: A synthesis. *Maydica*, *50*, 531–547. Retrieved from https://journals-crea.4science.it/index.php/maydica

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

Gauch, H. G. (1988). Model Selection and Validation for Yield Trials with Interaction. *Biometrics*, *44*(3), 705–715. https://doi.org/10.2307/2531585

Gilmore, E. C., & Rogers, J. S. (1958). Heat units as a method of measuring maturity in corn. *Agronomy Journal*, *50*(10), 611–615. https://doi.org/10.2134/agronj1958.00021962005000100014x

Glaz, B., & Kang, M. S. (2008). Location contributions determined via GGE biplot analysis of multienvironment sugarcane genotype-performance trials. *Crop Science*, *48*(3), 941–950. https://doi.org/10.2135/cropsci2007.06.0315

Grant, R. F., Jackson, B. S., Kiniry, J. R., & Arkin, G. F. (1989). Water deficit timing effects on yield components in maize. *Agronomy Journal*, *81*(1), 61–65. https://doi.org/10.2134/agronj1989.00021962008100010011x

Hung, H. Y., Browne, C., Guill, K., Coles, N., Eller, M., Garcia, A., … Holland, J. B. (2012). The relationship between parental genetic or phenotypic divergence and progeny variation in the maize nested association mapping population. *Heredity*, *108*(5), 490–499. https://doi.org/10.1038/hdy.2011.103

Jacobs, B. C., & Pearson, C. J. (1991). Potential yield of maize, determined by rates of growth and development of ears. *Field Crops Research*, *27*(3), 281–298. https://doi.org/10.1016/0378-4290(91)90067-6

Kang, M. S., Aggarwal, V. D., & Chirwa, R. M. (2006). Adaptability and stability of bean cultivars as determined via yield-stability statistic and GGE biplot analysis. *Journal of Crop Improvement*, *15*(1), 97–120. https://doi.org/10.1300/J411v15n01_08

Kusmec, A., Srinivasan, S., Nettleton, D., & Schnable, P. S. (2017). Distinct genetic architectures for phenotype means and plasticities in *Zea mays*. *Nature Plants*, *3*(9), 715–723. https://doi.org/10.1038/s41477-017-0007-7

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lee, E. A., & Tollenaar, M. (2007). Physiological basis of successful breeding strategies for maize grain yield. *Crop Science*, *47*(Supplement_3), S–202–S–215. https://doi.org/10.2135/cropsci2007.04.0010IPBS

Mastrodomenico, A. T., Haegele, J. W., Seebauer, J. R., & Below, F. E. (2018). Yield stability differs in commercial maize hybrids in response to changes in plant density, nitrogen fertility, and environment. *Crop Science*, *58*(1), 230–241. https://doi.org/10.2135/cropsci2017.06.0340

de Mendiburu, F. (2019). agricolae: Statistical procedures for agricultural research. *R package version*, *1*, 3–1. Retrieved from https://CRAN.R-project.org/package=agricolae.

Meseka, S., Menkir, A., Olakojo, S., Jalloh, A., Coulibaly, N., & Bossey, O. (2016). Yield stability of yellow maize hybrids in the savannas of West Africa. *Agronomy Journal*, *108*(4), 1313–1320. https://doi.org/10.2134/agronj2015.0366

Messmer, R., Fracheboud, Y., Bänziger, M., Vargas, M., Stamp, P., & Ribaut, J. M. (2009). Drought stress and tropical maize: QTL-by-environment interactions and stability of QTLs across environments for yield components and secondary traits. *Theoretical and Applied Genetics*, *119*(5), 913–930. https://doi.org/10.1007/s00122-009-1099-x

Miller, N. D., Haase, N. J., Lee, J., Kaeppler, S. M., de Leon, N., & Spalding, E. P. (2017). A robust, high-throughput method for computing maize ear, cob, and kernel attributes automatically from images. *The Plant Journal*, *89*(1), 169–178. https://doi.org/10.1111/tpj.13320

Nafziger, E. (2009). Corn. In *Illinois Agronomy Handbook* (pp. 13–26). Urbana IL: University of Illinois Extension.

Peiffer, J. A., Romay, M. C., Gore, M. A., Flint-Garcia, S. A., Zhang, Z., Millard, M. J., … Buckler, E. S. (2014). The genetic architecture of maize height. *Genetics*, *196*(4), 1337–1356. https://doi.org/10.1534/genetics.113.159152

Pixley, K. V., & Bjarnason, M. S. (2002). Stability of grain yield, endosperm modification, and protein quality of hybrid and open-pollinated quality protein maize (QPM) Cultivars. *Crop Science*, *42*(6), 1882–1890. https://doi.org/10.2135/cropsci2002.1882

Prine, G. M. (1971). A critical period for ear development in maize. *Crop Science*, *11*(6), 782–786. https://doi.org/10.2135/cropsci1971.0011183×001100060003x

Core Team, R. (2019). R: A language and environment for statistical computing. Retrieved from https://www.r-project.org/

Reed, A., Singletary, G. W., Schussler, J. R., Williamson, D. R., & Christy, A. L. (1988). Shading effects on dry matter and nitrogen partitioning, kernel number, and yield of maize. *Crop Science*, *28*(5), 819–825. https://doi.org/10.2135/cropsci1988.0011183×002800050020x

Reif, J. C., Hallauer, A. R., & Melchinger, A. E. (2005). Heterosis and heterotic patterns in maize. *Maydica*, *50*(3–4), 215–223. Retrieved from https://journals-crea.4science.it/index.php/maydica

Simmonds, N. W. (1991). Selection for local adaptation in a plant breeding programme. *Theoretical and Applied Genetics*, *82*(3), 363–367. https://doi.org/10.1007/BF02190624

Teixeira, J. E. C., Weldekidan, T., de Leon, N., Flint-Garcia, S., Holland, J. B., Lauter, N., … Wisser, R. J. (2015). Hallauer's Tusón: A decade of selection for tropical-to-temperate phenological adaptation in maize. *Heredity*, *114*(2), 229–240. https://doi.org/10.1038/hdy.2014.90

Tollenaar, M., Dwyer, L. M., & Stewart, D. W. (1992). Ear and kernel formation in maize hybrids representing three decades of grain yield improvement in Ontario. *Crop Science*, *32*(2), 432–438. https://doi.org/10.2135/cropsci1992.0011183×003200020030x

Uhart, S. A., & Andrade, F. H. (1995a). Nitrogen deficiency in maize: I. Effects on crop growth, development, dry matter partitioning, and kernel set. *Crop Science*, *35*(5), 1376–1383. https://doi.org/10.2135/cropsci1995.0011183×003500050020x

Uhart, S. A., & Andrade, F. H. (1995b). Nitrogen deficiency in maize: II. Carbon-nitrogen interaction effects on kernel number and grain-yield. *Crop Science*, *35*(5), 1384–1389. https://doi.org/10.2135/cropsci1995.0011183×003500050021x

Wang, Y., Yao, J., Zhang, Z., & Zheng, Y. (2006). The comparative analysis based on maize integrated QTL map and meta-analysis of plant height QTLs. *Chinese Science Bulletin*, *51*(18), 2219–2230. https://doi.org/10.1007/s11434-006-2119-8

Yan, W., Kang, M. S., Ma, B., Woods, S., & Cornelius, P. L. (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Science*, *47*(2), 643–655. https://doi.org/10.2135/cropsci2006.06.0374

Yan, W., & Rajcan, I. (2002). Biplot analysis of test sites and trait relations of soybean in Ontario. *Crop Science*, *42*(1), 11–20. https://doi.org/10.2135/cropsci2002.1100

Zaidi, P. H., Vasal, S. K., Maniselvan, P., Jha, G. C., & Singh, R. P. (2008). Stability in performance of quality protein maize under abiotic stress. *Maydica*, *53*, 249–260. Retrieved from https://journals-crea.4science.it/index.php/maydica

SUPPORTING INFORMATION

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

How to cite this article: Falcon CM, Kaeppler SM, Spalding EP, et al. Relative utility of agronomic, phenological, and morphological traits for assessing genotype-by-environment interaction in maize inbreds. *Crop Science*. 2020;60:62–81. https://doi.org/10.1002/csc2.20035