The effects of artificial selection and planting density on performance stability across environments and yield component traits in maize (*Zea mays* L.*)*

by

Bridget A. McFarland

A dissertation submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
(Plant Breeding and Plant Genetics)

at the
University of Wisconsin - Madison
2021

Date of final oral examination: 01/08/2021

This dissertation is approved by the following members of the Final Oral Committee:

Natalia de Leon, Professor in Agronomy
Shawn M. Kaeppler, Professor in Agronomy
Edgar P. Spalding, Professor in Botany
Jeffrey Endelman, Associate Professor in Horticulture

THE EFFECTS OF ARTIFICIAL SELECTION AND PLANTING DENSITY ON PERFORMANCE STABILITY ACROSS ENVIRONMENTS AND YIELD COMPONENT TRAITS IN MAIZE (*Zea mays* L.*)*

Under the supervision of Professor Natalia de Leon

ABSTRACT

Plant breeders selectively breed plants to maximize productivity within the context of the target environment(s). These environments can be viewed as entire fields or regions with common features, such as weather or soil characteristics, or specific growing conditions unique to a single plant within a field. The objectives of this dissertation are: (1) assess the effects of selection and environment cues on plant performance and stability using maize hybrids derived from a common genetic background and (2) evaluate the effect of planting density on yield component traits in maize. Both of these studies utilize resources and datasets that are part of the Genomes To Fields (G2F) Initiative.

Chapter One provides background on the history of maize and its importance, plant development and various abiotic influences on grain yield, and an overview of genotype-by-environment interaction (G × E) and stability. Chapter Two examines how breeding for productivity has influenced trait stability and which environmental variables are most influential in hybrid performance. Across a range of environments, we observed increased stability and improved performance in lines that had undergone multiple cycles of selection relative to unselected lines across most productivity traits (such as, stand count, flowering time, and grain yield), except stalk lodging. The environmental variables that were most influential on plant

performance were those related to soil classification and day length. When comparing the environmental variables estimates across models, using genotype (G) and G × E variance in place of the raw phenotypic trait values generated environmental that were significantly correlated to the traditional stability environmental rankings. This suggests that environmental variance is not a good indicator of environment ranking, while G+ G×E better explains hybrid performance.

In Chapter Three, an ever-increasing density (EID) plot design was used to evaluate the response of hybrids to increased planting densities using image-based phenotyping of grain yield components. This study used a set of three biparental populations sharing one parent in common, the others representing a highly selected, an almost complete unselected, and an intermediately selected parent. Kernel size traits were the most sensitive to increases in planting density and decreased significantly, while ear and cob width were the least sensitive and did not significantly change. The lines derived from the least selected parent produced the heaviest cobs and kernels, and largest kernel size, while the lines derived from the commercially relevant and highly selected parent produced the lightest cobs and smallest kernels. When connecting density traits data with production-level G2F data, ear height in the production-level environments was significantly correlated with ear height at two of the EID treatments. The known correlation between these two formats supports the continued use of the EID design to evaluate varying planting density effects.

Overall, this work emphasizes the utility of dissecting environments at multiple levels to better understand the driving forces of plant performance and stability, and an alternative planting density scheme to understand the effects of variable planting density on yield component traits, and genetically dissect grain yield components for continued improvement.

Dedication

I dedicate this dissertation to the strong women whose support, love, and memory made it possible.

Mia Terese Guion
November 13, 1994 – June 8, 2018

Melanie Arlene Thwing-Eastman
March 15, 1989 – June 5, 2020

Acknowledgments

To all who helped me get here (mentioned below and not), thank you. This journey has challenged me in so many ways to help me become a better scientist, friend, and person. And for that I'm forever grateful.

I would first like to thank my advisor and mentor, Dr. Natalia de Leon. Thank you for taking a chance on me after only having me visit during field season in 2016. You always encouraged me to pursue whatever interested me and reminded me that life is an experiment without replication. I'm so thankful for your mentorship and support of my growth personally and professionally. And to the other leader of the Field Maize group, Dr. Shawn Kaeppler. Thank you for your thoughtful comments and suggestions throughout my research projects. I would also like to acknowledge the other members of my committee, Drs. Edgar Spalding and Jeff Endelman. Thank you for helping me grow as a scientist during my graduate journey.

A big thank you to the Field Corn lab group, both past and present members, for their willingness to help with planting, data collection, and harvesting. Dustin Eilert, Marina Borsecnik, Rachel Perry, Ryan Alpers, Ben Fisher, Julia Butler, and Melissa Draves for their technical assistance in the field and lab. Fellow graduate students and scientists for their input on data analysis, discussion, and comradery, specifically Kathryn Michel, Jose Varela, Jonas Rodriguez, Nathaniel Schleif, Calli Anibas, Dylan Schoemaker, Martin Costa, Alden Perkins, Dr. Dayane Lima, Dr. Mike White, Dr. Joseph Gage, and Dr. Alejandro Castro Aviles.

The data compilation and statistical analyses presented here would not have been possible without the guidance and expertise of Drs. Nathan Miller, David Guedex-Cross, and Karl Broman.

The Genomes To Fields Initiative collaborators, lab technicians, postdocs, and graduate students for the opportunity to be part of inspiring research and large-scale project management. I'd also like to thank the following sources for funding my research and ability to present at conferences throughout my graduate career: USDA NIFA Award #2016-67024-2219 and Iowa Corn Growers Association.

The Agronomy and PBPG faculty, staff, and students – especially, Amy Cottom, Jillene Fisch, Caitlin Collies, Sally Muller, and Joanna Schuth.

The ELPA community for encouraging me to always ask questions. Specifically, my minor advisor, Dr. Cliford Conrad, for always reminding me to pursue my burning questions and to learn with, from, and for others.

The CALS Equity and Diversity Committee for the opportunity to serve the greater CALS community with incredibly passionate members, especially Thomas Browne and Erika Anna.

The Dog Mom squad for the daily encouragement, coffee/Candy Cart/lunch breaks, weddings, trips to the Library, virtual Prelim celebrations due to COVID-19, and much needed dog playdates. Lindsay Chamberlain, Dr. Maddy Oravec, and Dr. Marian Bolton, my graduate school

journey and life would not be the same without you all. I'm very grateful that Madison brought us together.

To previous graduate students for friendship and guidance: Dr. Adam Bolton, Derek Potratz, and Sarah Striegel.

My lifelong friends from Helser: Kaylee Hahn, Kristen Eiles, and Emily Cowles. From undergrad through graduate school, we've been through the highest highs and the lowest lows together. I'm very lucky to have had you all by my side.

My mother, Sheila McFarland, for introducing me to the cornfields at age 14. You will always get 100% of the credit for helping me find my passion in agriculture. Whether it was pollinating in the field or shelling ears during the COVID-19 pandemic, I could always count on you and Jack McFarland to be there for me without hesitation. Mike Ruffolo and Ann (Dolly) and Jack (G-Paw) Thole for your support and love throughout this process. I don't know where I would be without any of you. Thank you for always being the wind beneath my wings.

One of my biggest sources of joy, Murphy. You have undoubtedly changed my life for the better with your unconditional love, willingness to nap, and constant yearning to be outdoors.

And finally, thank you to my partner and best friend, Christian Jones. Through every high and each low, you have been on every step of this journey with me. Your unwavering love and daily assurance mean more than you'll ever know. I look forward to many more adventures together.

Table of Contents

List of Tables

List of Figures

CHAPTER 1. LITERATURE REVIEW

1.1 History of Maize (*Zea mays*, L.) and its Importance

Maize is one of the most important cereal crops in the world. Over 6,000 years ago, maize was domesticated in Mexico (Wilkes, 1967; Doebley et al, 1990) leading to its evolution from teosinte (*Zea mays* ssp. *parviglumis*) to the modern maize plant structure (*Zea mays* L.). This domestication led to reductions on the activation of axillary meristem and subsequent reduction in branchiness, elimination of kernel shattering, and several ear and kernel modifications including changes in kernel shape, size, conformation and color and the formation of a cob. With adaptations to multiple temperate and tropical environments, maize is one of the most economically valuable crops worldwide (USDA FAS, 2019). Within the United States, the largest number of hectares harvested of all crops in 2019 were of maize, resulting in an estimated $52.71 billion dollar value (USDA WASDE, 2020).

Substantial and systematic increases in maize yield have coincided with the efforts of plant breeders, which has been recorded and published by the USDA since 1866 (USDA NASS, 2018). Plant breeders manipulate plants to improve appearance and performance for various markets (feed, food, fuel, and fiber). This requires evaluating genetic variation observed in collections of genotypes and creating changes to meet specific objectives or goals. Until the mid-1930's, open-pollinated maize varieties were used, and the average grain yield remained relatively constant at approximately 1.75 t/ha (Figure 1). While these inbred lines individually showed low yield and vigor, George Shull first demonstrated that crossing two inbred lines resulted in progeny superior to the two parents, a tendency of plants termed heterosis (Shull, 1908). Since inbred lines had poor seed set, limiting potential hybrid production, D. F. Jones proposed a double cross hybrid breeding scheme. This involved crossing two inbred lines and

crossing the hybrid progeny with the hybrid progeny of two other inbred lines, resulting in higher seed set (D.F. Jones, 1917). The implementation of double cross hybrids was very popular and resulted in yield improvements of 0.05 t/ha/year, from 1937 through 1955 (Figure 1).

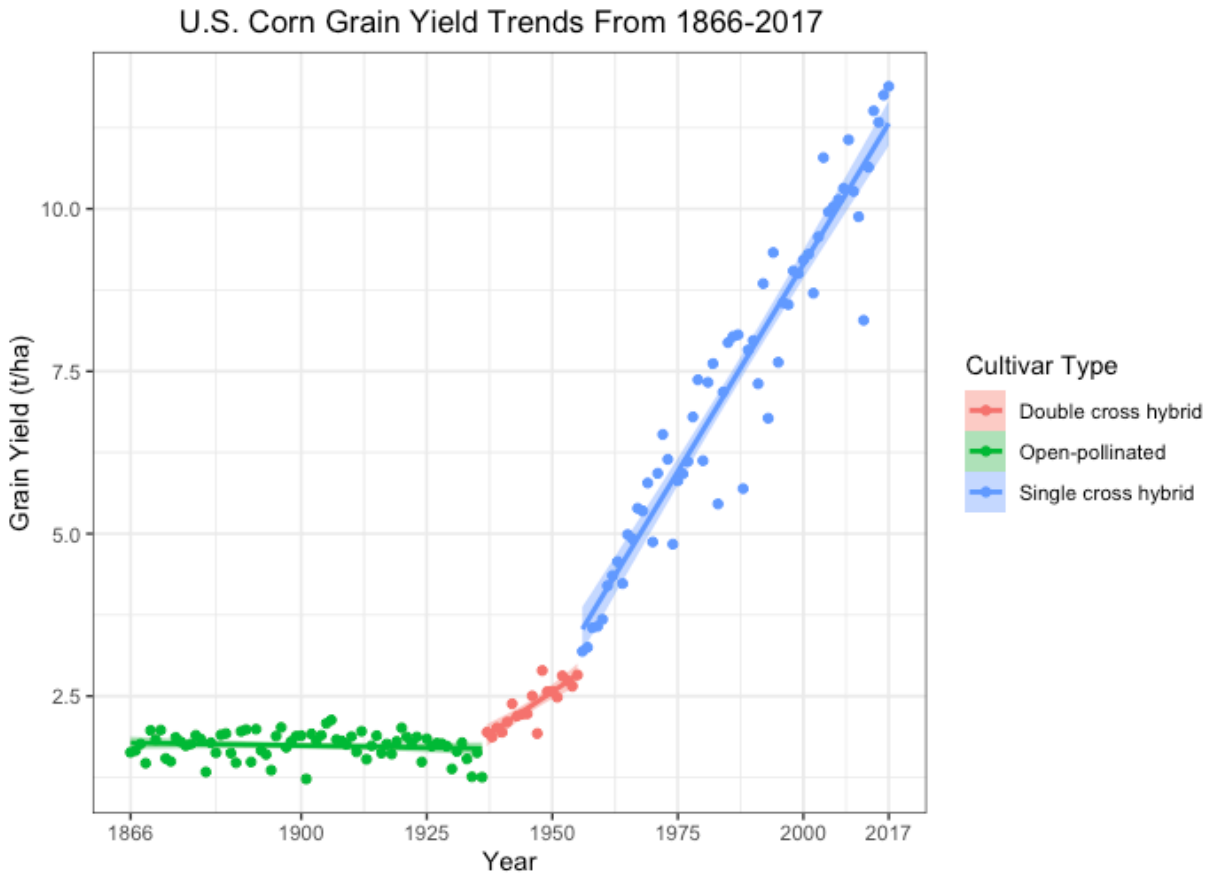


Figure 1.1. Historical maize grain yield trends in the United States from 1866 to 2017 (Data from USDA-NASS, 2020).

The use of single cross hybrids replaced the previous breeding scheme in the 1960's (Figure 1) and are the predominant cultivar type in the U.S. and other parts of the world today. Crop genetics, management practices, and production technologies contribute to the high productivity of maize hybrids (Edgerton, 2009). For example, in North America, recommended planting densities have increased from an average of 30,000 plants ha^{-1} in the 1930's to over 81,000 plants ha^{-1} today (Duvick, 2005a; Butzen and Burnison, 2014). The most recent maize

improvement phase that started in the mid-1950's has resulted in approximately 0.13 t/ha/year increases.

The improvement in productivity of maize as mostly a hybrid crop is coupled with the improvement in inbred performance. Therefore, heterosis is expected to be reduced over time for the productivity trait under selection. Most breeders do not routinely measure inbred performance, therefore heterosis cannot specifically be measured. However, breeders do exploit heterosis by crossing two inbreds from differing heterotic groups to form hybrids. Within North America, maize germplasm encompasses a few primary heterotic groups that when crossed together can optimize hybrid performance. These hybrids are then evaluated to determine superior productivity (Mikel and Dudley, 2006). While the nomenclature of the groups is subjective (Tracy and Chandler, 2006), the most widely recognized heterotic groups are: Stiff Stalk, non-Stiff Stalk and Iodent (Lu and Bernardo, 2001; Gethi et al., 2002). The Stiff Stalk heterotic group refers to lines that are derived from the Iowa Stiff Stalk Synthetic (BSSS) population. The BSSS population was synthesized in the 1930's by Dr. George Sprague by intermating 16 genotypes with above-average stalk quality (Sprague, 1946; Lamkey, 1992). This population has since undergone 13 complete cycles of selection (Edwards, 2016) and is an important part of the U.S. commercial germplasm.

Previous studies using BSSS populations and lines derived from them have looked at genetic diversity (Helms et al., 1989; Messmer et al., 1991; Lu and Bernardo, 2001), grain yield and heterosis (Mungoma and Pollak, 1988), and different aspects of plant morphology (Brekke et al., 2011; Edwards, 2011; Gage et al., 2018). Founding inbred lines derived from the BSSS, which include B14, B37, B73, and B84, have been central components of the commercial germplasm for maize breeding in North America (Darrah and Zuber, 1986; Messmer et al.,

1991). Inbred B73, a BSSS cycle 5 product, was the first publicly available maize reference genome (Schnable et al., 2009) and Mikel and Dudley (2006) identified B73 as the most prevalent publicly available maize line protected by the U.S. Patent and U.S. Plant Variety Protection (PVP) Act with known pedigree. The BSSS and its derivatives provide a common genetic background and ideal framework for evaluating the effect of artificial selection which will be further explored in Chapter Two.

Beyond the heterotic groups, the creation of maize populations specifically for genetic mapping has been a vital tool in uncovering the genetic architecture of complex traits. Different types of populations can be constructed for genetic mapping (such as recombinant inbred lines (RILs), F_2, association panels, among others), each possessing specific advantages and drawbacks regarding the sampling of allelic variation, power to detect shared quantitative trait loci (QTL), power to detect rare QTL within families, among other characteristics (Verhoeven et al., 2006; Holland et al., 2007; Jamann et al., 2015). A QTL is a statistical association between a phenotype and genomic polymorphism identified using genetic markers. In order for genetic mapping to be successful, sufficient genetic diversity and population-wide recombination events are essential. The maize Nested Association Mapping (US NAM) population was specifically constructed for high-resolution QTL mapping (Buckler et al., 2009). In the US NAM population, reference line B73 was crossed to twenty-five diverse inbred lines, which collectively represented approximately 80% of the variability recognized in maize at the time. The F_1 progeny for each of the 25 crosses were then self-pollinated for six generations to generate 200 homozygous RILs from each family. The resulting 5,000 RILs maintained a portion of the genetic diversity from the founder lines, while also representing approximately 135,000 recombination events (Kump et al., 2011). The large allele diversity and population size provide

substantial power for the detection and resolution of QTLs, allowing for the exploitation of components of both, linkage and association mapping approaches (Kump et al., 2011; Cook et al., 2012; Wallace et al., 2014). There have been numerous NAM- inspired populations developed for maize since 2009, such as the Chinese NAM (Li et al., 2015), the European NAM (Bauer et al., 2013), TeoNAM (Chen et al., 2019), as well as a number of others in different crops (rice, Fragoso et al., 2012; soybean, Song et al., 2017; sorghum, Bouchet et al., 2017; Gage et al., 2020). The PHW65 NAM population was designed as a reduced version of the maize NAM, consisting of a single parental reference line, PHW65, crossed to three diverse founder, inbred lines, PHN11, Mo44, and MoG (Gage et al., 2018). Each of the three families consists of 200 double-haploid (DH) derived lines from the F_2 progeny, resulting in 600 total inbred lines. The PHW65 NAM has been used for genetic mapping for various phenotypic and yield-component traits (Haase, 2015; Gage et al., 2018). It should be noted that because only three parents were used in the PHW65 NAM, the possible number of alleles at each locus is four which is more than the two possible alleles from a single biparental cross, but less than 26 potential alleles in the US NAM. The PHW65 NAM population will be used in Chapter Three to assess the phenotypic diversity associated with yield component traits and their response to increasing planting density treatments.

1.2 Maize Development and Influences on Grain Yield

Maize is a diclinous, monoecious plant, meaning that the male and female flowers are on the same plant but in separate locations. This allows for cross-pollination between the staminate inflorescence in the tassel and the pistillate on the ears of neighboring plants as well as self-pollination when pollen lands on the silks of the same plant. The main growth stages of maize

can be described as vegetative (V) and reproductive (R). Vegetative stage can be determined by counting the number of collared leaves sequentially (V1 to VN) followed by the stage in which the tassel emerges (VT). Reproduction then takes place followed by kernel development (R1 to R6) and concludes with the plant senescing and dying (Purdue University Extension, 2009). Availability of sunlight, water, nutrients, and the presence of heat (measured as growing degree units, GDUs) can impact a plant's growth stages.

From V6 toV8, the number of kernel rows in the maize ear is being determined, as the reproductive meristem takes shape (Texas A&M AgriLife Extension, 2017). From V12- V15, the maximum kernels per row is defined based on the number of ovules that fully develop prior to pollination. Nutrient deficiencies, or any type of stress throughout V10- V15 growth stages can significantly reduce kernel numbers and ear size, therefore reducing overall grain yield. The silks begin to elongate from V15- V18 until VT when the tassel emerges. At this point, the plant transitions into reproduction, the first stage of which is described as R1 (Iowa State University Extension, 2017; Nielsen, 2019). Once the silks have been successfully pollinated by pollen from the tassel, grain development occurs from R2 (kernels begin to look like blisters following fertilization) and R3 (kernels have a "milky" interior as starch accumulation is increasing). Available nutrients reallocate from the leaves and other photosynthetic tissues to the cob to contribute to grain filling until R4 (kernel interior has a "dough"-like consistency and kernels begin to form an indent at the top due to increasing starch deposition and moisture loss). Stages R5- R6 are when the plant has reached kernel physiological maturity, leaves begin to naturally senesce and plants are typically ready to be harvested (Kansas State University Extension Agronomy, 2016). At R5, kernels have fully formed "dents" at the top due to declining moisture and increasing starch. The final stage, R6, results in kernels that have formed an abscission layer

at the kernel base, eliminating further dry matter accumulation, also known as “black layer.” Two extremely important abiotic factors that influence a plant’s progression through these V and R stages are growing degree/heat units, or GDUs, and day length, or photoperiod.

During the first half of 18th century, researchers began to propose the idea that each vegetable species has a certain amount of heat that is necessary in order to flower or fructify and below which neither of these two physiological functions could take place (Réaumur, 1735). The relationship between heat units and plant development is commonly used in maize, barley, wheat, canola, potato, sugar beet, and alfalfa among others (NDSU Extension, 2020). While different methods exist for calculating heat units for each crop (because the base temperature is crop-specific), the most commonly accepted method for maize in the U.S. uses the average daily temperature (in degrees Fahrenheit) minus 50. This method also limits the daily maximum and minimum temperatures allowed in the calculation as 86°F and 50°F, respectively (Gilmore and Rogers, 1958; Barger, 1969). These GDUs are added together to represent cumulative GDUs accrued at a given physiological state. The cumulative GDUs needed for maize plants to reach developmental stages, such as reproduction, are variety specific.

Certain plants consistently open and close their flowers at particular times of the day (Linneaus, 1751), but maize flowers do not open and close. Instead, maize flowering is heavily controlled by temperature and photoperiod, or daylength (Allison and Daynard, 1979). Photoperiod plays an especially large role in cultivar adaptation in maize. First reported by Garner and Allard in 1923, maize populations originating from Peru were influenced by the duration of daily illumination. They observed that these populations reached reproduction (flowering stages) earlier when grown under short day lengths. Teosinte, maize’s progenitor, is photoperiod sensitive and therefore reproduction is typically delayed under long day conditions,

compared to temperate maize that has evolved to have a reduced sensitivity to photoperiod (Holland et al., 2012). Understanding photoperiodic responses is essential for adaptation to temperate and tropical environments (Major, 1980) as well as to determine proper planting date to maximize productivity. With longer photoperiod conditions, there is an increase in the amount of vegetative growth and development prior to tassel initiation at reproduction, which is expected to increase productivity potential up to a certain point (i.e., the period is not extended too far and the plant never reaches reproductive stages). This has been observed in maize through an increase in number of emerged leaves, longer stem length, and greater plant dry weight at reproduction (Kannenberg et. al., 1974).

Though GDUs and photoperiod are essential for plant development, other abiotic stresses, such as plant population density and associated reduction on sunlight incidence, can lead to an imbalance in plant productivity and survival. It was suggested at the end of the 20th century that to continue making future yield gains in maize, tolerance to high plant densities will be essential (Duvick, 1997; Duvick and Cassman, 1999) and numerous plant traits were identified as being affected by density changes, both directly and indirectly. For example, commercial hybrids that have been directly selected for high yield performance in these planting density conditions have also greater correlation with the following morphological traits: more upright leaves, shorter plant height, and smaller tassels (Duvick, 2005b; Gage et al., 2018). It has also been observed that modern hybrids are less responsive to density-induced environmental conditions and have less stalk lodging at increased densities compared to older, less selected varieties (Sangoi et al, 2002; Fellner et al., 2006). Increasing the number of plants within a plot means that there is more interplant competition for light and nutrients, both of which are necessary for photosynthesis and growth (Tetio-Kagho and Gardner, 1988). As maize plants

grow taller and develop greater leaf areas, a canopy structure is formed. This canopy structure is the aboveground portion of the plants, primarily composed of the leaves, that provide shade from the ground below when they begin to overlap with neighboring plants. This means that while the canopy structure helps suppress weed growth under the canopy, it can affect the photosynthetic and light interception capabilities for nearby plants (Andrade et al., 1993). For example, Edmeades and Daynard (1979) found that increasing planting densities from 50,000 to 100,000 plants ha^{-1} resulted in decreased per plant yields and reduced kernel numbers per maize ear. Plant prolificacy, or the number of ears, was also reduced with increased planting density (Jacobs and Pearson, 1991).

Plant responses to canopy shade involve the perception of low red to far-red light ratios (R:FR) by phytochromes. Phytochromes are types of photoreceptor proteins that act as a biological light switch for the plant, monitoring the level, intensity, and duration of light in the environment (Franklin and Quail, 2010). Within the light spectrum, phytochromes allow the absorption of red (R) and far-red (FR) light between 600 and 730 nm wavelengths, respectively (Quail, 1997). Absorption of light by the photoreceptor sets in motion a cascade of events that ultimately results in a developmental response (Hopkins and Huner, 2009).

Radiation within and below a canopy is deficient on red and blue light because these wavelengths are absorbed by the chlorophyll of the overlying leaves. In comparison, chlorophyll is transparent to FR light and the reduction of it is limited solely to reflection. For example, in response to anticipated shading (reduced R:FR ratios from the leaves of non-shading neighbors) or actual shading due to high planting densities, plants trigger a variety of responses as a result of phytochromes being converted into an active form that activates stem elongation or alters resource partitioning (Martínez-García et al., 2014). In maize, it has been documented that this

expression of light-responsive genes contributes to plant height, ear node height, culm width, leaf sheath, leaf angle, and internode length (Satter and Whetherell, 1968; Sheehan et al., 2007). Plants may grow taller than neighbors to compete for light (resulting in more assimilates being allocated to plant tissue development instead of grain fill development), accompanied by a reduction in stem diameter and an increase in ear barrenness. It is also important to note the interplant competition that is taking place below ground. Neighboring roots are able to detect each other by the alteration in resource availability (Goldberg, 1990), and are forced to compete for nutrients or water (Schenk, 2006). Previous research has focused on decreasing space between plants, to evaluate planting density tolerance among varieties and advance those that are the most efficient (i.e., most productive in terms of yield) to continue to provide steady increases in grain yield overtime.

The majority of the published research on high plant population density effects in maize utilize one of two experimental designs; entire fields with two- or four- row plots planted at a single planting density or blocks with different planting densities within a field. This specific design is then replicated across several fields, which requires a significant number of resources such as number of seeds and overall field space availability. White et al. (unpublished), developed a novel planting scheme called the Ever-Increasing Density (EID), in which plant spacing within a single plot was systematically increased to represent several planting densities within a row. The EID example described in Chapter Three includes six different densities and requires twenty-seven kernels per plot. A single plant from each planting density within each plot was selected for phenotypic evaluation. The space between plants was always in the same order, decreasing from 30 inches down to two between plants within a row, representing 17,209 plants ha^{-1} and 258,148 plants ha^{-1} as extreme densities. Results from White et al. aligned with

previous findings of the phenotypic response in maize plant morphology to increased planting density using this EID scheme (Stinson and Moss, 1960; Tetio-Kagho and Gardner, 1988; Cox, 1996). Further implementation of the EID planting scheme will be addressed in Chapter Three.

Because yield is a complex, quantitative trait, it often presents low heritability (Hallauer and Miranda, 1988) and it is highly affected by the environment. For this reason, it can be useful to break down yield into component traits that can be analyzed separately by plant breeders interested in dissecting what specific components are predominately contributing to overall yield under different conditions. Some of these components include kernel number, kernel weight, ear size, and number of ears per hectare (Lauer, 2012). Other studies have dissected specific components further such as kernel number into kernel row number times number of kernels per row (Austin and Lee, 1996, 1998; Messmer et al., 2009; Guo et al., 2011; Liu et al., 2012; Stange et al., 2013; Li et al., 2015; Iowa State University Extension and Outreach, 2017).

To assess plant tolerance and response to stresses, such as planting density, it is useful to evaluate yield component traits as opposed to overall yield to reveal the genetic components of these responses. Studies have elucidated aspects of the genetic architecture of stress responses to water stress (Austin and Lee, 1998), planting density (Guo et al., 2011), and nitrogen (N) use efficiency (Gallais and Hirel, 2003) among others. Austin and Lee (1998) detected 59 QTL associated with kernel weight, ear length, and ear diameter, under water stress or non-stress environments with 10 QTL being in common to both growing conditions. Using two planting densities (52,500 and 90,000 plants/hm^2), Guo et al. (2011) identified QTL clusters on chromosomes 1, 3, 4, 5, and 8, for the yield-related traits of ear length, kernel thickness, kernel weight, row number, cob width, and ear width. Gallais and Hirel (2004) evaluated yield, kernel weight, and kernel number under a standard (175 kg N/ha) N application compared to no N

application levels. A reduction in yield was observed in the no N treatments and kernel number was the most affected, while kernel weight was the least affected. All of the previously mentioned studies were able to deduce the various stages of maize growth and development that the stresses had the most impact based on the yield components that were affected. Another benefit of using yield components is that these traits tend to have higher heritability and the ability to detect trait QTL is affected by the heritability of the trait of interest (Bernardo, 2004). In breeding, heritability is used to estimate the proportion of phenotypic variation in a population that is due to genetic variation (Falconer and Mackay, 1996). The accuracy of trait measurement will affect the experimental error, directly influencing the estimation of heritability. Some yield components are difficult to measure accurately by hand so for this reason, these traits can benefit greatly from precision phenotyping.

Advances in high-throughput phenotyping using digital imaging of yield component traits have facilitated more accurate and efficient measurements of these traits compared to hand measurements (Miller et al., 2017; Makanza, Zaman-Allah, Cairns, et al., 2018). The automated image analysis pipeline developed by Miller et al., 2017 allows imaging of maize ears, cobs, and kernels using flat-bed scanners. Custom software measures: ear maximum width, ear average width, ear length, ear weight, cob maximum width, cob average width, cob length, cob color, cob weight, kernel row number, kernels per row, kernel weight, kernel area, kernel width, kernel length, and kernel depth. Using this image pipeline, Falcon et al. (2019) identified more QTL for kernel row number compared to other kernel specific traits (i.e., width, depth, area), supporting previous findings (Bommert et al., 2013; Li et al., 2014; Tian et al., 2014). Evaluating yield component traits across a range of planting densities could identify genomic regions to be explored to improve grain yield while building tolerance to increased planting density. This topic

will be explored in Chapter Three combining the PHW65 NAM population, the EID scheme, and an automated image analysis pipeline (Miller et al., 2017).

1.3 Overview of Genotype-by-Environment Interaction (G × E) and Stability

The phenotypic traits that breeders evaluate are the result of the plant's genetics, or genotype (G), and growing environment (E). The environment is the collection of nongenetic factors that can include soil characteristics, climatic variables, management practices and combinations of different biotic and abiotic stressors, among others (Comstock and Moll, 1963). Genotype-by-environment interaction, or G × E, occurs when environmental sensitivity is not consistent for all genotypes (Falconer, 1952). This differential response has been studied using various mathematical models to understand what portion of the observed variation in plant performance is a product of G, E, or G × E. The relative performance of genotypes in different environments is of high importance to plant breeders, as this influences their recommendations of best-performing varieties for specific growing regions. The performance of traits across locations can be modelled using four general patterns (Ouyang et al, 1995). These patterns will be further described using two genotypes and two environments, but they apply to any number of genotypes and environments greater than two.

Pattern 1 is when one genotype is superior to the other genotype across the two environments and their performance is parallel to one another across these environments. This response does not involve any G × E. Pattern 2 is when one of the genotypes is superior to the other genotype but the performance across the two environments is not parallel, as in Pattern 1. There is G × E in this pattern, but it is considered a non-crossover interaction because the rank of the genotype performance does not change between environments. Pattern 3 involves a crossover interaction in which the best performing genotype differs between the two environments, but the

magnitude of the performance change (i.e., correlation between the environments) is identical. Pattern 4 also represents a crossover interaction, but the magnitude of the correlation differs between environments.

G × E can be a significant source of variation in breeding programs. Breeders have dealt with it using three general approaches: ignore, avoid, or exploit it (Eisemann et al, 1990; Bernardo, 2002). By ignoring G × E, breeders test varieties in a range of environments and select the best performers across all locations. This means that varieties are not necessarily bred for specific environments but rather to have reasonable performance across all environments. To reduce or avoid G × E, the number of environments is clustered into smaller subgroups that represent similar areas based on different environmental and management characteristics. Finally, breeders may opt to exploit G × E by selecting varieties that perform well in specific environments at the expense of poorer performance in other environments to maximize productivity in target locations. This "exploitation" approach can also include taking into account stability analyses (Yates and Cochran, 1938) by looking at the performance of genotypes relative to one another across environments. When "exploiting G × E" and incorporating stability analyses, there are several different ways to calculate stability, but they all serve the same purpose – to identify genotypes that have reliable performance across a range of environments (Becker and Leon, 1988). Stability analyses enable breeders to compare a specific genotype's performance relative to other genotypes across these environments. In order to differentiate environments from one another, a measurement, such as average genotype performance for a given trait, can be used (Bernardo, 2010). Based on the analysis that is chosen, the results correspond to genotype differential response across environments, or G × E. These analyses enable breeders to maximize productivity and select or recommend varieties that will perform

well in target environment(s) given the specific breeding goals. Overall, while G × E can be viewed as a major complication for breeders, it can also be positively interpreted. Ceccarelli (1996) argues that smaller, resource-poor farmers are often overlooked in cultivar development because of the generally low yields. However, it is possible to exploit the diversity of variety performance to continue developing new and improved cultivars for these growing conditions by breeding for specific adaptation.

From a conceptual standpoint, G × E can be studied as a biological or statistical phenomenon. Evolutionary biologists view G × E in the biological sense, focusing on a plant's ability to adapt to the environment, as an individual plant or on a population bases. For example, varieties that are considered stable in regard to environmental stresses are valued for their predictable yields. Because these varieties are bred to predictably respond to specific stresses such as drought and intense temperatures, they tend to have more resilient performance in these situations. When these environmental stressors aren't present however, these varieties' stability is less valuable. Instead, it would be more advantageous for these varieties to take advantage of more favorable conditions (i.e., more available water, fewer pests, etc.) and produce higher yields (Kusmec et al., 2018). Tollenaar and Lee (2002) argued that most of the increase in maize yields in the US is associated with stress tolerance accruing from improvement in genotype-by-management interaction and concluded that high yield potential and yield stability may not be mutually exclusive. While stable, predictable yields are beneficial, it is important to note that a loss of allelic diversity might be associated with reduced plasticity at a population level (Gage et al., 2017). If loci for yield and yield stability are genetically linked, selection for yield would be expected to affect allele frequencies at the loci for yield stability (Kusmec et al., 2018). This could constrain adaptation to future yield in environments that present new challenges or

peculiarities (Gage et al., 2017). A recent study in maize found that water captured in roots differed between planting density treatments under drought conditions but did not differ between double or single cross hybrids (Messina et al., 2020). The absence of a differential genetic gain under well-watered conditions suggests that artificial selection did not necessarily improve water capture but has improved the ability of individual plants to support reproductive structures under stress conditions.

Alternatively, G × E can be interpreted and dissected statistically. Falconer (1952) suggested that a single trait evaluated in two environments could be studied as two different traits and therefore their correlation could be measured. Low genetic correlation would suggest that different genes control the trait in different environments, and low phenotypic correlation means that the performance of this trait in these two environments is not consistent, suggesting the presence of significant G × E. In a situation of no G × E, that is no change in performance rank of the trait between environments, a single genotype would prevail in any environment and it would be very simple to correctly select the best performing genotype(s) (Gauch and Zobel, 1996; Hongyu et al., 2014). Both approaches to studying G × E bring important concepts and questions. To complement one another and bridge the gap of G × E being viewed as being solely biological versus statistical, the use of environmental information in prediction models allows for the statistical models, such as regressions, to be enhanced with biological knowledge (van Eeuwijk et al., 2005). The fusion of whole-crop physiological models and quantitative genetics in the development of these tools and cultivars will be essential.

One effort currently focused on G × E research is the Genomes To Fields (G2F) Initiative. The G2F is a multi-institutional, public collaborative that works to develop tools to connect genomic, climatic, and phenotypic information in maize to better understand

performance at the genetic level across diverse environments (McFarland et al., 2020). Since 2014, G2F has been working on the Maize G × E Project, which evaluates common sets of maize hybrids in a range of locations with numerous plot-level and field-level measurements. By creating publicly available genotypic, phenotypic, and climatic datasets, scientists are able to utilize data generated across many locations and years to improve predictability of response. The work for the following studies leverages germplasm, data, and infrastructure from G2F. Chapter Two includes data that was part of the 2016- 2017 G2F experiments and Chapter Three utilizes a subset of the 2018- 2019 G2F germplasm.

Two studies using G2F data have explored G × E and stability (Gage et al., 2017; Falcon et al., 2019) and their relationship with various traits in maize. One of the methods used to quantify stability mentioned in both is the Finlay-Wilkinson regression (FWR) method (Finlay and Wilkinson, 1963) which uses a linear regression approach, to assess the performance of varieties as a linear function of productivity across a gradient of environments. In FWR, the x-axis is the environmental index, most commonly the average trait performance of common genotypes for a given environment, ordering all environments from least to best performing. The y-axis represents the trait performance and each genotype's performance is plotted for each environment. From there, a linear regression is performed for each genotype. Since the same units are used for the x- and y-axis, the interpretations of the linear regression's slope and deviation can be further described by three types of stability (Lin et al, 1986). Type I describes varieties with little variance across environments. These varieties have a regression slope close or equal to 0, reflecting perfect stability, and are considered biologically stable. By being biologically stable, varieties don't respond favorably to improved environments but also have some capacity to buffer poorer environments (Baker, 1987). Type II describes varieties with

performance similar or proportional to the environmental index. Typically, the environmental index is a gradient of the mean of all varieties, or a subset of common varieties tested in the trial. The slope used in Types I and II is indicative of how a variety's performance compares to what we would expect in an environment, given the selected set of common varieties across all locations. The larger the range in slopes for a trait for a set of individuals in a population, the larger G × E it is expected to be observed for that trait.

Type III is measured based on the relative variability of individuals' performance in relationship to the regression line. Low mean squared error (MSE) after modeling genotypic responses across environments indicates a more stable genotype. An assumption for this type of stability is an identifiable pattern response across environments, many times assumed to be linear. Using these parameters for stability analyses, breeders would focus on whichever type of stability is most important for their goals and objectives. If a localized, target region has been identified, the MSE-based stability would be of interest. On the other hand, the slope-based stability would be more important if breeders focus on a wider range of environments. These types of stability do not serve as specific numerical "cutoff" decisions when breeding, but simply to provide general guidance. Each trait should be assessed independently rather than assuming that a genotype that is stable for one trait will necessarily be stable for other traits (Zaidi et al., 2008).

Environmental indices for stability analyses have traditionally been estimated using the mean yield performance across all genotypes, or a subset of common genotypes, in each location to create an overall gradient of environments ranging from worst to best performing. This index approach has been used in stability analyses for a broad range of crops over several decades. For example, one study evaluated 146 barley cultivars across 50 locations and used the mean yield in

each environment to create the environmental index which all cultivars were regressed on (Kraakman, et al., 2004). Yield ranged from 4,810.0 to 6,377.1 kg/ha, representing a wide range of values for the environmental index of the stability analysis. Conversely, a study in maize utilized a subset of only four hybrids to determine the environmental index that was then used for commercial hybrids in hundreds of multi-location yield trials. The goal of this study was to identify and produce widely adapted maize hybrids, increasing the desire for stable yield performance across a larger number of production environments (Bradley et al., 1988). Another study, focused on short-season maize, found it was more practical to characterize genotypes on a group basis due to the large number of genotypes and environments typically evaluated by breeders (Francis and Kannenberg, 1978). These studies support the value of having a range of environments to develop the environmental index and the use of group-based stability evaluations.

Despite its widespread use, it is important to acknowledge certain limitations and critiques of the FWR stability approach. As previously mentioned, an assumption that is typically made is that there is a linear relationship between the variety's trait performance and the range of environments these varieties are tested in. If this is not true, the linear regression can lead to misleading results depending on the genotypes evaluated and the environments that they are evaluated in. As with all G × E and stability studies, a large number of environments is advantageous because too many extremes (high or low) can skew results (Walsh and Lynch, 2014).

Other methods to study stability include multiplicative statistical models such as Additive Main-effects and Multiplicative Interaction (AMMI) and Genotype and Genotype × Environment (GGE) biplots. The AMMI model uses an analysis of variance and principal component analysis

(PCA) to examine G × E sensitivity, separating the G from the G × E effects (Dyer, 1994). Principal component analysis (PCA) is a technique to reduce dimensionality of datasets and using the magnitude and direction of variables to explain patterns and the results are presented in a biplot. PCA is used for the visualization and interpretation of AMMI results. An AMMI biplot can help identify genotypes which have the highest productivity for a trait and are well adapted to specific environments and specifically, G × E that the FWR linear regression is unable to explain (in non-linear reactions). GGE biplots also utilize PCA and biplots for visual interpretation, but they differ from AMMI in that the G and G × E effects are jointly analyzed, and environment discriminability and genotype performance can be more clearly illustrated (Miranda et al., 2009). GGE biplots visually present two PCAs and the fit of the biplot can be interpreted using patterns described in Yan, Kang, Ma, Woods, and Cornelius (2007) such as which genotype "won" in which environment. A study on complex traits in wheat used FWR, AMMI, and GGE biplot approaches to identify the most stable breeding lines (Sjoberg et al., 2020). The most stable varieties according to the AMMI stability value were not always consistent with FWR, however, the varieties with the worst stability value ranking were also the least stable according to FWR. The rankings of stable performing varieties from the GGE analysis were significantly correlated with those from FWR. GGE and AMMI biplots are limited by the fact that they are a visual tool, not a statistical test. The PCA results, or principal components (PC), are only useful if the first few PCs account for a large percentage of the variation (Bernardo, 2010). It has been found that the first few PCs don't always capture a large portion of the G × E (the first PC is due to repeatable G × E while the second PC is the "noise" or nonrepeatable G × E), so when visualizing AMMI and GGE biplots with these PCs, the graph could be misleading and lose usefulness in partitioning environments into subgroups (Bernardo,

2010; Sjoberg et al., 2020). Conversely, FWR is easier to interpret in regard to the three types of stability that can be estimated from the analysis. When comparing AMMI and FWR to assess alfalfa performance, Annicchiarico (1992) found that the two analyses provided similar results when environmental stress was minimal. A FWR is expected to most effectively describe G × E when the average site yield and the occurrence of interactions are affected by the same environmental factor(s), such as cold or drought stress.

Previous work expressed the desire of having the environmental index of stability evaluations be based on environmental factors that influence the performance of these genotypes rather than the average of common set of genotypes evaluated across the environments under study, yet the lack of reliable technology and understanding of how to incorporate these additional sources of data have been the cited barriers (Eberhart and Russell, 1966; Lin et al., 1986). Work in genomic selection and crop modelling has incorporated environmental conditions into models (Jarquin et al., 2014; Cooper et al., 2016) and the inclusion of environmental covariates has improved prediction ability even in challenging situations such as prediction of previously unobserved environments (Malosetti et al., 2016). With a growing interest in understanding what environmental and management factors affect the determination of yield potential of different cultivars, new statistical tools have been implemented by plant breeders (González-Barrios et al., 2017). In rice, Monteverde et al. (2019), used 54 environmental covariates in a partial least square regression to identify the environmental variables that have the most influence on milling quality in rice. Incorporating the most influential environmental variables (humidity and solar radiation) into prediction models also allowed for improved prediction of rice yield in untested environments. In maize, a recent study (Li et al., 2018) was able to redefine the environmental index using direct climatic measurements of the testing

locations, independent from plant performance data. A similar approach can be applied to the FWR stability analysis to establish a biologically meaningful and environmentally descriptive index. This environmentally-based index could then be parameterized by a set of hybrids. If this index captures variability and performance influenced by ECs, then slope and MSE estimates would be comparable across experiments and not limited to the specific composition of a given experiment. Currently, it is unknown how the results from the environmentally-based index compare to those generated from the traditional performance-based index. And while breeders have evaluated plants across environments to make selections throughout history, it is unclear how stability responses have been influenced, directly and indirectly, through this selection based on plant performance. Chapter Two will focus on assessing stability in the BSSS population and the implementation of environmental data into the FWR stability index.

Maize is a staple crop worldwide, used for food, feed, fiber, and fuel. Over 43 billion bushels of maize were produced globally in 2018- 2019, with over one-third coming from the U.S. (USDA FAS, 2019). With the anticipated population growth worldwide, the need for greater and more consistent crop production is essential (Ray et al., 2013; United Nations, 2015; Bailey-Serres et al., 2019). Plant breeders have contributed greatly to increases in crop productivity by selecting plant varieties that have both high productivity and consistent performance across environments. The objectives of this work focus on: 1) assessing the effects of selection and environment cues on plant stability and performance of maize hybrids derived from a common genetic background (Chapter Two); and 2) evaluating the effect of varying planting density on yield component traits in maize (Chapter Three).

LITERATURE CITED

Allison, J.C.S., and T.B. Daynard. 1979. Effect of Change in Time of Flowering, Induced by Altering Photoperiod or Temperature, on Attributes Related to Yield in Maize. *Crop Science.* 19(1): 1-4. https://doi.org/10.2135.cropsci1979.0011183X001900010001x.

Andrade, F.H., S.A. Uhart, and A. Cirilo. 1993. Temperature affects radiation use efficiency in maize. *Field Crops Research.* 32(1): 17-25. https://doi.org/10.1016/0378-4290(93)90018-1.

Annicchiarico, P., and P. Annicchiarico. 1992. Cultivar adaptation and recommendation from alfalfa trials in Northern Italy. *Journal of Genetics and Breeding.* 46: 269-269.

Austin, D.F., and M. Lee. 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., and M. Lee. 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.

Bailey-Serres, J., J.E. Parker, E.A. Ainsworth, G.E.D. Oldroyd, and J.I. Schroeder. 2019. Genetic strategies for improving crop yields. *Nature* 575(7781), 109-118. https://doi.org/10.1038/s41586-019-1679-0.

Baker, K.F. 1987. Evolving Concepts of Biological Control of Plant Pathogens. *Annual Review of Phytopathology.* 25:67-85.

Barger, G.L. 1969. Total Growing Degree Days. Weekly Weather & Crop Bulletin 56:18. U.S. Dept. of Commerce and USDA, Washington, D.C.

Bauer, E., M. Falque, H. Walter, C. Bauland, C. Camisan, L. Campo, N. Meyer, N. Ranc, R. Rincent, W. Schipprack, T. Altmann, P. Flament, A.E. Melchinger, M. Menz, J. Moreno-González, M. Ouzunova, P. Revilla, A. Charcosset, O.C. Martin, and C.C. Schön. 2013. Intraspecific variation of recombination rate in maize. *Genome Biology.* 14(9): R103. https://doi.org/10.1186/gb-2013-14-9-r103.

Becker, H.C., and J. Léon. 1988. Stability Analysis in Plant Breeding. *Plant Breeding.* 101(1): 1–23.

Bernardo, R. 2002. Breeding for quantitative traits in plants. Stemma, Woodbury, MN.

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

Bommert, P., N.S. Nagasawa, and D. Jackson. 2013. Quantitative variation in maize kernel row

number is controlled by the FASCIATED EAR2 locus. *Nature Genetics.* 45(3): 334-337. https://doi.org/10.1038/ng.2534.

Bouchet, S., M.O. Olatoye, S.R. Marla, R. Perumal, T. Tesso, J. Yu, M. Tuinstra, G.P. Morris. 2017. Increased power to dissect adaptive traits in global sorghum diversity using a nested association mapping population. Genetics 206(2): 573-585. https://doi.org/10.1534/genetics.116.198499.

Bradley, J.P., K.H. Knittle, and A.F. Troyer. 2013. Statistical Methods in Seed Maize Product Selection. *Journal of Production Agriculture* 1(1): 34-38. https://doi.org/10.2134/jpa1988.0034.

Brekke, B., J. Edwards, and A. Knapp. 2011. Selection and adaptation to high plant density in the Iowa Stiff Stalk Synthetic maize (Zea mays L.) population: II. plant morphology. *Crop Science* 51(6): 2344–2351.

Buckler, E.S., J.B. Holland, P.J. Bradbury, C.B. Acharya, P.J. Brown, C. Browne, E. Ersoz, S. Flint-Garcia, A. Garcia, J.C. Glaubitz, M.M. Goodman, C. Harjes, K. Guill, D.E. Kroon, S. Larsson, N.K. Lepak, H. Li, S.E. Mitchell, G. Pressoir, J.A. Peiffer, M.O. Rosas, T.R. Rocheford, M.C. Romay, S. Romero, S. Salvo, H.S. Villeda, H. Sofia da Silva, Q. Sun, F. Tian, N. Upadyayula, D. Ware, H. Yates, J. Yu, Z. Zhang, S. Kresovich, and M.D. McMullen. 2009. The Genetic Architecture of Maize Flowering Time. *Science*, 325(5941): 714-719. https://doi.org/10.1126/science.1174276.

Butzen, S., and M. Burnison. 2014. Maize seeding rate considerations for 2014. Crop Insights 24(3). DuPont Pioneer.

Ceccarelli, S. 1996. Positive interpretation of genotype by environment interaction in relation to sustainability and biodiversity. In Plant Adaptation and Crop Improvement. p. 467-487.

Chen, Q., C.J. Yang, A.M. York, W. Xue, L.L. Daskalska, C.A. DeValk, K.W. Krueger, S.B. Lawton, B.G. Spiegelberg, J.M. Schnell, M.A. Neumeyer, J.S. Perry, A.C. Peterson, B. Kim, L. Bergstrom, L. Yang, I.C. Barber, F. Tian, and J.F. Doebley. 2019. TeoNAM: A nested association mapping population for domestication and agronomic trait analysis in maize. *Genetics* 213(3): 1065-1078. https://doi.org/10.1534/genetics.119.302594.

Comstock, R.E., and R.H. Moll. 1963. Genotype-by-environment interactions. *Statistical genetics and plant breeding*. Natl. Acad. Sci., Natl. Res. Counc., Washington, D.C. p. 164-194.

Cook, J.P., M.D. McMullen, J.B. Holland, F. Tian, P. Bradbury, J. Ross-Ibarra, E.S. Buckler, and S.A. Flint-Garcia. 2012. Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiology* 158(2). https://doi.org/10.1104/pp.111.185033.

Cooper, M., F. Technow, C. Messina, C. Gho, and L. Radu Totir. 2016. Use of crop growth

models with whole-genome prediction: Application to a maize multienvironment trial. *Crop Science* 56(5): 2141-2156. https://doi.org/10.2135/cropsci2015.08.0512.

Cox, W.J. 1996. Whole-plant physiological and yield responses of maize to plant density. *Agronomy Journal* 88(3): 489-496. https://doi.org/10.2134/agronj1996.00021962008800030022x.

Darrah, L.L., and M.S. Zuber. 1986. 1985 United States Farm Maize Germplasm Base and Commercial Breeding Strategies. *Crop Science.* 26(6): 1109-1113. https://doi.org/10.2135/cropsci1986.0011183X002600060004x.

Doebley, J., A. Stec, J. Wendel, and M. Edwards. 1990. Genetic and morphological analysis of a maize-teosinte F2 population: Implications for the origin of maize. *Proceedings of the National Academy of Sciences of the United States of America.* 87(24): 9888-9892. https://doi.org/10.1073/pnas.87.24.9888.

Duvick, D.N. 2005a. Genetic progress in yield of United States maize (Zea mays L.). *Maydica.* 50(3): 193-202.

Duvick, D.N. 2005b. The Contribution of Breeding to Yield Advances in maize (Zea mays L.). *Advances in Agronomy.* 86: 83-145. https://doi.org/10.1016/S0065-2113(05)86002-X.

Duvick, D.N., and K.G. Cassman. 1999. Post–Green Revolution Trends in Yield Potential of Temperate Maize in the North-Central United States Breeding Methods and Investment. *Crop Science* 39(6). https://doi.org/10.2135/cropsci1999.3961622x.

Dyer, C.J. 1994. Statistical analysis of regional yield trials: AMMI analysis of factorial designs. *Field Crops Research* 37(2):146-147.

Eberhart, S.A., and W.A. Russell. 1966. Stability parameters for comparing varieties. *Crop Science* 6: 36-40.

Edgerton, M.D. 2009. Increasing crop productivity to meet global needs for feed, food, and fuel. *Plant Physiology* 149(1):7-13. https://doi.org/10.1104/pp.108.130195.

Edmeadest, G.O., and T.B. Daynard. THE RELATIONSHIP BETWEEN FINAL YIELD AND PHOTOSYNTHESIS AT FLOWERING IN INDIVIDUAL IMAIZE PLANTS. *Canadian Journal of Plant Science*. 59(3): 585-601. https://doi.org/10.4141.cjps79-097.

Edwards, J.W. 2016. Genotype × environment interaction for plant density response in maize (Zea mays L.). *Crop Science* 56(4): 1493-1505. https://doi.org/10.2135/cropsci2015.07.0408.

Edwards, J. 2011. Changes in plant morphology in response to recurrent selection in the Iowa StiffStalk Synthetic maize population. *Crop Science* 51(6): 2352-2361. https://doi.org/10.2135/cropsci2010.09.0564.

Eisemann, R.L., M. Cooper, and D.R. Woodruff. 1990. Beyond the analytical methodology. Better interpretation and exploitation of genotype-by-environment interaction in breeding? *In* Genotype by environment interaction and plant breeding.

Falcon, C.M., S.M. Kaeppler, E.P. Spalding, N.D. Miller, N. Haase, N. AlKhalifah, M. Bohn, E.S. Buckler, D.A. Campbell, I. Ciampitti, L. Coffey, J. Edwards, D. Ertl, S. Flint-Garcia, M.A. Gore, C. Graham, C.N. Hirsch, J.B. Holland, D. Jarquín, J. Knoll, N. Lauter, C.J. Lawrence-Dill, E.C. Lee, A. Lorenz, J.P. Lynch, S.C. Murray, R. Nelson, M.C. Romay, T. Rocheford, P.S. Schnable, B. Scully, M. Smith, N. Springer, M.R. Tuinstra, R. Walton, T. Weldekidan, R.J. Wisser, W. Xu, and N. de Leon. 2020. Relative utility of agronomic, phenological, and morphological traits for assessing genotype-by-environment interaction in maize inbreds. *Crop Science* 60(1): 62–81.

Falconer, D.S. 1952. The Problem of Environment and Selection. *American Naturalist* 86(830): 293-298.

Falconer, D.S., and Mackay, T.F.C. 1996. Introduction to Quantitative Genetics. 4th Edition.

Fellner, M., E. David Ford, and E. Van Volkenburgh. 2006. Development of erect leaves in a modern maize hybrid is associated with reduced responsiveness to auxin and light of young seedlings in vitro.*Plant Signaling and Behavior* 1(4): 201-211. https://doi.org/10.4161/psb.1.4.3106.

Finlay, K.W., and G.N. Wilkinson. 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.

Fragoso, C.A., M. Moreno, Z. Wang, C. Heffelfinger, L.J. Arbelaez, J.A. Aguirre, N. Franco, L.E. Romero, K. Labadie, H. Zhao, S.L. Dellaporta, M. Loriaux. 2017, Genetic architecture of a rice nested association mapping population. *G3* 7(6):1913-1926. https://doi.org/10.1534/g3.117.041608.

FRANCIS, T.R., and L.W. KANNENBERG. 1978. YIELD STABILITY STUDIES IN SHORT-SEASON MAIZE. I. A DESCRIPTIVE METHOD FOR GROUPING GENOTYPES. *Canadian Journal of Plant Science* 58:1029-1034.

Franklin, K.A., and P.H. Quail. 2010. Phytochrome functions in Arabidopsis development. *Journal of Experimental Biology* 61(1):11-24. https://doi.org/10.1093/jxb/erp304.

Gage, J.L., B. Monier, A. Giri, and E.S. Buckler. 2020. Ten Years of the Maize Nested Association Mapping Population: Impact, Limitations, and Future Directions. *Plant Cell* 32:2083-2093. https://doi.org/10.1105/tpc.19.00951.

Gage, J.L. Jarquin, D., Romay, C. et al. 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

Gage, J.L., M.R. White, J.W. Edwards, S. Kaeppler, and N. de Leon. 2018. Selection Signatures Underlying Dramatic Male Inflorescence Transformation During Modern Hybrid Maize Breeding. Genetics 210(3): 1125 LP – 1138Available at http://www.genetics.org/content/210/3/1125.abstract.

Gallais, A., and B. Hirel. 2004. An approach to the genetics of nitrogen use efficiency in maize. *Journal of Experimental Botany* 55(1):295-306. https://doi.org/10.1093/jxb/erh006.

GARNER, W.W., and H.A. ALLARD. 1920. EFFECT OF THE RELATIVE LENGTH OF DAY AND NIGHT AND OTHER FACTORS OF THE ENVIRONMENT ON GROWTH AND REPRODUCTION IN PLANTS. *Journal of Agricultural Research* 18:553-606.

Gauch R., H., and W. Zobel. 1996. AMMI Analysis of Yield Trials. *In* Genotype-by-Environment Interaction.

Gethi, J.G., J.A. Labate, K.R. Lamkey, M.E. Smith, and S. Kresovich. 2002. SSR variation in important U.S. maize inbred lines. *Crop Science* 42(3): 951-957. https://doi.org/10.2135/cropsci2002.9510.

Gilmore, E.C., and J.S. Rogers. 1958. Heat Units as a Method of Measuring Maturity in Maize 1. *Agronomy Journal*, 50(10), 611-615. https://doi.org/10.2134/agronj1958.0002196200.

Goldberg, D.E. 1990. Components of Resource Competition in Plant Communities. *In* Perspectives on Plant Competition.

Gollob, H.F. 1968. A statistical model which combines features of factor analytic and analysis of variance techniques. *Psychometrika* 33:73-115. https://doi.org/10.1007/BF02289676.

González-Barrios, P., M. Castro, O. Pérez, D. Vilaró, and L. Gutiérrez. 2017. Genotype by environment interaction in sunflower (Helianthus annus L.) to optimize trial network efficiency. *Spanish Journal of Agricultural Research* 15(4). https://doi.org/10.5424/sjar/2017154-11016.

Guo, J., Z. Chen, Z. Liu, B. Wang, W. Song, W. Li, J. Chen, J. Dai, and J. Lai. 2011. Identification of genetic factors affecting plant density response through QTL mapping of yield component traits in maize (Zea mays L.). *Euphytica* 182(409). https://doi.org/10.1007/s10681-011-0517-8.

Haase, N. (2015). *Phenotypic Analysis and Genetic Dissection of Yield Component Traits in Maize (Zea mays L.)* (Doctoral dissertation, University of Wisconsin-Madison). Available from ProQuest Dissertations and Thesis database. (UMI No. 10187943).

Hallauer, A.R., and D. Malithano. 1976. Evaluation of maize varieties for their potential as breeding populations. *Euphytica* 25:117-27.

Hallauer, A. R., and J. Miranda, 1988 Quantitative genetic in maize breeding, 2nd Edition. Iowa State University Press, Ames, Iowa.

Helms, T.C., A.R. Hallauer, O.S. Smith. 1989. Genetic drift and selection evaluated from recurrent selection programs in maize. *Crop Science* 29(2):602-607. https://doi.org/10.2135/cropsci1989.0011183X002900030009x.

Holland, J.B. 2007. Genetic architecture of complex traits in plants.*Current Opinion in Plant Biology* 10(2), 156-161. https://doi.org/10.1016/j.pbi.2007.01.003.

Hongyu, K., M. García-Peña, L.B. de Araújo, and C.T. dos Santos Dias. 2014. Statistical analysis of yield trials by AMMI analysis of genotype × environment interaction. Biometrical Letters 51(2), 89-102. https://doi.org/10.2478/bile-2014-0007.

Hopkins, W., and N. Huner. 2009. Resonses of plant to envirnomental stress. *In* An Introduction to Plant Physiology.

Hunter, R.B., L.A. Hunt, and W. Kannenberg. 1974. Photoperiod and temperature effects on maize. *Canadian Journal of Plant Science* 57: 1127-1133.

Iowa State University Extension and Outreach. 2017. Estimating Maize Yields Using Yield Components. Available at https://crops.extension.iastate.edu/cropnews/2017/08/estimating-corn-yields-using-yield-components.

Jacobs, B.C., and C.J. Pearson. 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.

Jamann, T.M., P.J. Balint-Kurti, and J.B. Holland. 2015. QTL mapping using high-throughput sequencing. *In* Plant Functional Genomics: Methods and Protocols: Second Edition.

Jarquín, D., J. Crossa, X. Lacaze, P. Du Cheyron, J. Daucourt, J. Lorgeou, F. Piraux, L. Guerreiro, P. Pérez, M. Calus, J. Burgueño, and G. de los Campos. 2014. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretial and Applied Genetics* 127(3), 595-607. https:doi.org/10.1007/s00122-013-2243-1.

Jones, D.F. 1917. Dominance of Linked Factors as a Means of Accounting for Heterosis. *Genetics* 2(5), 466-479.

Kansas State University Extension Agronomy. 2016. Maize Growth and Development. Available at https://www.agronomy.k-state.edu/extension/crop-production/corn/corn-growth-development/.

Kraakman, A.T.W., R.E. Niks, P.M.M.M. Van Den Berg, P. Stam, and F.A. Van Eeuwijk. 2004. Linkage disequilibrium mapping of yield and yield stability in modern spring barley

cultivars. *Genetics* 168(1): 435-446. https://doi.org/10.1534/genetics.104.026831.

Kump, K.L., P.J. Bradbury, R.J. Wisser, E.S. Buckler, A.R. Belcher, M.A. Oropeza-Rosas, J.C. Zwonitzer, S. Kresovich, M.D. McMullen, D. Ware, P.J. Balint-Kurti, and J.B. Holland. 2011. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nature Genetics* 43(2): 163–168.

Kusmec, A., N. de Leon, and P.S. Schnable. 2018. Harnessing phenotypic plasticity to improve maize yields. *Fronteis in Plant Science* 9:1377.

Lamkey, K.R. 1992. Fifty Years of Recurrent Selection in the Iowa Stiff Stalk Synthetic Maize Population. *Maydica* 37(1), 19-28.

Lauer, J., 2012 Maximizing Maize Grain and Forage Yield, pp. in Field Crops. University of Wisconsin - Agronomy Department, http://maize.agronomy.wisc.edu/AA/pdfs/A088.pdf.

Li, C., Y. Li, P.J. Bradbury, X. Wu, Y. Shi, Y. Song, D. Zhang, E. Rodgers-Melnick, E.S. Buckler, Z. Zhang, Y. Li, and T. Wang. 2015. Construction of high-quality recombination maps with low-coverage genomic sequencing for joint linkage analysis in maize. *BMC Biology* 13-78. https://doi.org/10.1186/s12915-015-0187-4.

Li, F., H.T. Jia, L. Liu, C.X. Zhang, Z.J. Liu, and Z.X. Zhang. 2014. Quantitative trait loci mapping for kernel row number using chromosome segment substitution lines in maize. *Genetics and Molecular Research* 13(1), 1707-1716. https://dx.doi.org/10.4238/2014.January.17.1.

Li, L., Y. Du, X. Shen, M. Li, W. Sun. 2015. KRN4 Controls Quantitative Variation in Maize Kernel Row Number. *PLoS Genetics* 11(11) e1005670. https://doi.org/10.1371/journal.pgen.1005670.

Li, X., T. Guo, Q. Mu, X. Li, and J. Yu. 2018. Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proceedings of the National Academy of Sciences* 115(26), 6679-6684. http://www.pnas.org/content/115/26/6679.abstract.

Lin, C. S., Binns, M. R. & Lefkovitch, L. P. 1986. Stability analysis: where do we stand? *Crop Science* 26(5), 894-900. https://doi.org/10.2135/cropsci1986.0011183X002600050012x.

Linneaeus, C. 1751. Philosophia Botanica.

Liu, R., H. Jia, X. Cao, J. Huang, F. Li, Y. Tao, F. Qiu, Y. Zheng, and Z. Zhang. 2012. Fine Mapping and Candidate Gene Prediction of a Pleiotropic Quantitative Trait Locus for Yield-Related Trait in Zea mays. *PLoS ONE* 7(11) e49836. https://doi.org/10.1371/journal.pone.0049836.

Lu, H., and R. Bernardo. 2001. Molecular marker diversity among current and historical maize inbreds. T*Theoretical and Applied Genetics* 103, 613-617. https://doi.org/10.1007.PL00002917.

Lynch, M. and Walsh, B. 1998. Genetics and Analysis of Quantitative Traits, 1st edn. Sinauer, Sunderland, MA.

Major, D.J. 1980. Environmental Effects on Flowering. *In* Hybridization of Crop Plants.

Makanza, R., M. Zaman-Allah, J.E. Cairns, J. Eyre, J. Burgueño, Á. Pacheco, C. Diepenbrock, C. Magorokosho, A. Tarekegne, M. Olsen, and B.M. Prasanna. 2018. High-throughput method for ear phenotyping and kernel weight estimation in maize using ear digital imaging. *Plant Methods* 14(49). https://doi.org/10.1186/s13007-018-0317-4.

Malosetti, M., D. Bustos-Korts, M.P. Boer, and F.A. Van Eeuwijk. 2016. Predicting responses in multiple environments: Issues in relation to genotype × Environment interactions. *Crop Science* 56(5), 2210-2222. https://doi.org/10.2135/cropsci2015.05.0311.

Mansfield, B.D., and R.H. Mumm. 2014. Survey of plant density tolerance in U.S. maize germplasm. *Crop Science* 54(1), 157-173. https://doi.org/10.2135/cropsci2013.04.0252.

Martínez-García, J.F., M. Gallemí, M.J. Molina-Contreras, B. Llorente, M.R.R. Bevilaqua, and P.H. Quail. 2014. The shade avoidance syndrome in Arabidopsis: The antagonistic role of phytochrome A and B differentiates vegetation proximity and canopy shade. *PLoS ONE* 9(10) e109275. https://doi.org/10.1371/journal.pone.0109275.

McFarland, B.A., N. Alkhalifah, M. Bohn, J. Bubert, E.S. Buckler, I. Ciampitti, J. Edwards, D. Ertl, J.L. Gage, C.M. Falcon, S. Flint-Garcia, M.A. Gore, C. Graham, C.N. Hirsch, J.B. Holland, E. Hood, D. Hooker, D. Jarquin, S.M. Kaeppler, J. Knoll, G. Kruger, N. Lauter, E.C. Lee, D.C. Lima, A. Lorenz, J.P. Lynch, J. McKay, N.D. Miller, S.P. Moose, S.C. Murray, R. Nelson, C. Poudyal, T. Rocheford, O. Rodriguez, M.C. Romay, J.C. Schnable, P.S. Schnable, B. Scully, R. Sekhon, K. Silverstein, M. Singh, M. Smith, E.P. Spalding, N. Springer, K. Thelen, P. Thomison, M. Tuinstra, J. Wallace, R. Walls, D. Wills, R.J. Wisser, W. Xu, C.T. Yeh, and N. De Leon. 2020. Maize genomes to fields (G2F): 2014-2017 field seasons: Genotype, phenotype, climatic, soil, and inbred ear image datasets. *BMC Research Notes* 13(1): 4–9. https://doi.org/10.1186/s13104-020-4922-8.

Messina, C.D., M. Cooper, D. McDonald, H. Poffenbarger, R. Clark, A. Salinas, Y. Fang, C. Gho, T. Tang, G. Graham. Reproductive resilience but not root architecture underpin yield improvement in maize (Zea mays L.). bioRxiv.

Messmer, R., Y. Fracheboud, M. Bänziger, M. Vargas, P. Stamp, and J.M. Ribaut. 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.

Messmer, M.M., A.E. Melchinger, M. Lee, W.L. Woodman, E.A. Lee, and K.R. Lamkey. 1991. Genetic diversity among progenitors and elite lines from the Iowa Stiff Stalk Synthetic (BSSS) maize population: comparison of allozyme and RFLP data. *Theoretical and Applied Genetics*, 83(1), 97-107. https://doi.org/10.1007/BF00229231.

Mikel, M.A., and J.W. Dudley. 2006. Evolution of North American dent maize from public to proprietary germplasm. *Crop Science,* 46(3), 1193-1205. https://doi.org/10.2135/cropsci2005.10-0371.

Miller, N.D., N.J. Haase, J. Lee, S.M. Kaeppler, N. de Leon, and E.P. Spalding. 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.

Mungoma, C., and L.M. Pollak. 1988. Heterotic Patterns among Ten Maize Belt and Exotic Maize Populations. *Crop Science,* 28(3), 500-504. https://doi.org/10.2135/cropsci1988.0011183X002800030015x.

NDSU Extension. 2020. North Dakota Agricultural Weather Network.

Nielsen, R.L. 2017. Historical Maize Grain Yields for the U.S, available at https://www.agry.purdue.edu/ext/maize/news/timeless/YieldTrends.html.

Nielsen, R.L. 2019. Predict leaf stage development in corn using thermal time. Purdue University Extension. https://www.agry.purdue.edu/ext/corn/news/timeless/VStagePrediction.html.

Ouyang, Z., Mowers, R.P., Jensen A., Wang, S. and Zheng, S.1995. Cluster analysis for genotype x environment interaction with unbalanced data. *Crop Science*. 35, 1300-1305. https://doi.org/10.2135/cropsci1995.0011183X003500050008x.

Purdue University Extension. 2009. Field Crops IPM.

Quail, P.H. 1997. The phytochromes: A biochemical mechanism of signaling in sight? *Bioessays* 19(7), 571-579. https://doi.org/10.1002/bies.950190708.

Ray, D.K., N.D. Mueller, P.C. West, and J.A. Foley. 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS One* 8(6) e66426. https://doi.org/10.1371/journal.pone.0066428.

Réaumur, R. 1735. Report of the Commissioner of Agriculture: For the Year of 1869.

Sangoi, L., M.A. Gracietti, C. Rampazzo, and P. Bianchetti. 2002. Response of Brazilian maize hybrids from different eras to changes in plant density. *Field Crops Research* 79(1), 39-51. https://doi.org/10.1016/S0378-4290(02)00124-7.

Satter, R.L., and D.F. Wetherell. 1968. Photomorphogenesis in Sinningia speciosa , cv. Queen Victoria II. Stem Elongation: Interaction of a Phytochrome Controlled Process and a Red-requiring, Energy Dependent Reaction. *Plant Physiology* 43(6), 953-960. https://doi.org/10.1104/pp.43.6.953.

Schnable, P.S., D. Ware, R.S. Fulton, J.C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L.

Fulton, T.A. Graves, P. Minx, A.D. Reily, L. Courtney, S.S. Kruchowski, C. Tomlinson, C. Strong, K. Delehaunty, C. Fronick, B. Courtney, S.M. Rock, E. Belter, F. Du, K. Kim, R.M. Abbott, M. Cotton, A. Levy, P. Marchetto, K. Ochoa, S.M. Jackson, B. Gillam, W. Chen, L. Yan, J. Higginbotham, M. Cardenas, J. Waligorski, E. Applebaum, L. Phelps, J. Falcone, K. Kanchi, T. Thane, A. Scimone, N. Thane, J. Henke, T. Wang, J. Ruppert, N. Shah, K. Rotter, J. Hodges, E. Ingenthron, M. Cordes, S. Kohlberg, J. Sgro, B. Delgado, K. Mead, A. Chinwalla, S. Leonard, K. Crouse, K. Collura, D. Kudrna, J. Currie, R. He, A. Angelova, S. Rajasekar, T. Mueller, R. Lomeli, G. Scara, A. Ko, K. Delaney, M. Wissotski, G. Lopez, D. Campos, M. Braidotti, E. Ashley, W. Golser, H. Kim, S. Lee, J. Lin, Z. Dujmic, W. Kim, J. Talag, A. Zuccolo, C. Fan, A. Sebastian, M. Kramer, L. Spiegel, L. Nascimento, T. Zutavern, B. Miller, C. Ambroise, S. Muller, W. Spooner, A. Narechania, L. Ren, S. Wei, S. Kumari, B. Faga, M.J. Levy, L. McMahan, P. Van Buren, M.W. Vaughn, K. Ying, C.-T. Yeh, S.J. Emrich, Y. Jia, A. Kalyanaraman, A.-P. Hsia, W.B. Barbazuk, R.S. Baucom, T.P. Brutnell, N.C. Carpita, C. Chaparro, J.-M. Chia, J.-M. Deragon, J.C. Estill, Y. Fu, J.A. Jeddeloh, Y. Han, H. Lee, P. Li, D.R. Lisch, S. Liu, Z. Liu, D.H. Nagel, M.C. McCann, P. SanMiguel, A.M. Myers, D. Nettleton, J. Nguyen, B.W. Penning, L. Ponnala, K.L. Schneider, D.C. Schwartz, A. Sharma, C. Soderlund, N.M. Springer, Q. Sun, H. Wang, M. Waterman, R. Westerman, T.K. Wolfgruber, L. Yang, Y. Yu, L. Zhang, S. Zhou, Q. Zhu, J.L. Bennetzen, R.K. Dawe, J. Jiang, N. Jiang, G.G. Presting, S.R. Wessler, S. Aluru, R.A. Martienssen, S.W. Clifton, W.R. McCombie, R.A. Wing, and R.K. Wilson. 2009. The B73 Maize Genome: Complexity, Diversity, and Dynamics. Science (80-.). 326(5956): 1112 LP – 1115. http://science.sciencemag.org/content/326/5956/1112.abstract.

Schenk, M.K. 2006. Nutrient efficiency of vegetable crops. *American Horticulture* 700, 21-33.

Sheehan, M.J., L.M. Kennedy, D.E. Costich, and T.P. Brutnell. 2007. Subfunctionalization of PhyB1 and PhyB2 in the control of seedling and mature plant traits in maize. *The Plant Journal* 49(2), 338-353. https://doi.org/10.1111/j.1365-313X.2006.02962.x.

Shull, G.H. 1908. The composition of a field of maize.*Journal of Heredity* 4(1), 296-301. https://doi.org/10.1093/jhered/os-4.1.296.

Sjoberg, S.M., A.H. Carter, C.M. Steber, and K.A. Garland-Campbell. 2020. Unraveling complex traits in wheat: Approaches for analyzing genotype × environment interactions in a multienvironment study of falling numbers. *Crop Science* 60(6), 3013-3026. https://doi.org/10.1002/csc2.20133.

Song, Q., et al. 2017. Genetic characterization of the soybean nested association mapping population. *The Plant Genome* 10(2). https://doi.org/10.3835/plantgenome2016.10.0109.

Sprague, G.F. 1946. Early Testing of Inbred Lines of Maize. *Journal American Society of Agronomy* 38, 108-117.

Stange, M., T.A. Schrag, H.F. Utz, C. Riedelsheimer, E. Bauer, and A.E. Melchinger. 2013. High-density linkage mapping of yield components and epistatic interactions in maize with doubled haploid lines from four crosses. *Molecular Breeding* 32(3), 533-546.

Stinson, H.T., and D.N. Moss. 1960. Some Effects of Shade upon Maize Hybrids Tolerant and Intolerant of Dense Planting. *Agronomy Journal* 52(8), 482-484.https://doi.org/10.2134/agronj1960.00021962005200080019x.

Tetio-Kagho, F., and F.P. Gardner. 1988. Responses of Maize to Plant Population Density. I. Canopy Development, Light Relationships, and Vegetative Growth. *Agronomy Journal*, 80(6), 930-935. https://doi.org/10.2134/agronj1988.00021962008000060018x.

Tian, B., J. Wang, and G. Wang. 2014. Confirmation of a major QTL on chromosome 10 for maize kernel row number in different environments. *Plant Breeding*, 133(2), 184-188. https://doi.org/10.1111/pbr.12132

Tollenaar, M., E.A. Lee. 2002. Yield potential, yield stability and stress tolerance in maize. *Field Crops Research*, 75(2), 161-169. https://doi.org.10.1016/S0378-4290(02)00024-2.

Tracy, W.F., and M.A. Chandler. 2008. The Historical and Biological Basis of the Concept of Heterotic Patterns in Maize Belt Dent Maize. *In* Plant Breeding: The Arnel R. Hallauer International Symposium.

United Nations, Department of Economic and Social Affairs, Population Division, 2015. World Population Prospects: The 2015 Revision, Key Findings and Advance Tables. Working Paper No. ESA/P/WP.241.

USDA-FAS, United States Department of Agriculture – Foreign Agricultural Service, 2019. Grain: World Markets and Trade.

USDA-NASS, United States Department of Agriculture – National Agricultural Statistics Service, 2020. Quick Stats, available at https://www.quickstats.nass.usda.gov.

USDA-NASS, United States Department of Agriculture – National Agricultural Statistics Service, 2018. Crop Production Historical Track Records, available at https://www.nass.usda.gov/Publications/Todays_Reports/reports/croptr18.pdf.

USDA-WASDE, United States Department of Agriculture – World Agricultural Supply and Demand Estimates, 2020. Maize Crop Value 2017.

Van Eeuwijk, F.A., M. Malosetti, X. Yin, P.C. Struik, and P. Stam. 2005. Statistical models for genotype by environment data: From conventional ANOVA models to eco-physiological QTL models. *Australian Journal of Agricultural Research*, 56(9), 883-894. https://doi.org/10.1071/AR05153.

Verhoeven, K.J.F., J.L. Jannink, and L.M. McIntyre. 2006. Using mating designs to uncover QTL and the genetic architecture of complex traits. *Heredity*, 96(2),139-149. https://doi.org/10.1038/sj.hdy.6800763.

Wallace, J.G., S.J. Larsson, and E.S. Buckler. 2014. Entering the second century of maize quantitative genetics. *Heredity,* 112(1), 30-38.

Washburn, J.D., M.B. Burch, and J.A.V. Franco. 2020. Predictive breeding for maize: Making use of molecular phenotypes, machine learning, and physiological crop models. *Crop Science*, 60(2), 622-638.

White, M.R. (submitted). An Evaluation of Allelic Diversity and Heterosis of Expired Plant Variety Protection Germplasm and a Novel Scheme to Evaluate Hybrid Response to Density in Maize (*Zea mays* L.) (Doctoral dissertation, University of Wisconsin-Madison). Available from ProQuest Dissertations and Thesis database. (UMI No. TBD).

Wilkes, H. Garrison. 1967. Teosinte: The Closest Relative of Maize. Bussey Inst., Harvard Univ.: Cambridge.

Yan, W., M.S. Kang, B. Ma, S. Woods, and P.L. Cornelius. 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Science*, 47(2), 643-655. https://doi.org/10.2135/cropsci2006.06.3074.

Yates, F., and W.G. Cochran. 1938. The analysis of groups of experiments. *The Journal of Agricultural Science*, 28(4), 556-580. https://doi.org/10.1017/S0021859600050978.

Zaidi, P.H., S.K. Vasal, P. Maniselvan, G.C. Jha, Mehrajjudin, and R.P. Singh. 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.

CHAPTER 2. AN INVESTIGATION ON HOW BREEDING HAS INFLUENCED TRAIT STABILITY IN MAIZE

Abbreviations list: G×E: genotype-by-environment interaction; BSSS: Iowa Stiff Stalk Synthetic population; FWR: Finlay-Wilkinson regression; PLSR: Partial Least Squares Regression; EC: environmental covariate

ABSTRACT

Plant breeders utilize selection to maximize productivity of crops for specific environments while considering the reliability of that performance across environments that vary spatially and temporally. The goals of this research are to assess how breeding has influenced trait stability and to determine which environmental variables are most influential in hybrid performance. A set of 102 hybrids was generated by crossing inbreds originating from the Iowa Stiff Stalk Synthetic (BSSS) population with varying levels of selection to tester DK3IIH6. Hybrids were evaluated across 31 environments as part of the Genomes to Fields Initiative using a randomized complete block design with two field replications per location in 2016 and 2017. Finlay-Wilkinson linear regression (FWR) was used to assess trait stability for productivity and related agronomic and phenological traits. Partial least squares regression (PLSR) was then used to identify the environmental covariates (EC) related to day length, field management, precipitation, soil, temperature, and wind, that were most influential on each of the phenotypic traits and across the groupings of lines representing different selection levels. PLSR was run using phenotypic trait models composed of E+G+ G×E and G+ G×E variance separately. We observed increased stability and improved performance in newer, highly selected inbreds relative to unselected inbreds for all traits, except stalk lodging. While soil classifications, such as

phosphorous and organic matter, were of chief importance across all hybrids, we did not find that groupings having undergone more selection responded significantly more to agronomic inputs than other groupings. When comparing ECs across PLSR models, the G+ G×E model generated environmental rankings based on predicted hybrid performance that were significantly correlated to the FWR environmental rankings, suggesting that environmental variance (E+G+ G×E) is not a good indicator of environment ranking, while G+ G×E better explains hybrid performance. Our results illustrate that selection in maize has improved performance and stability across environments and that different ECs were relevant depending on the trait considered.

2.1 Introduction

For centuries, humans have bred plants for superior productivity, more recently using statistical tools and experimental design approaches to separate effects due to genotype, environment, and genotype x environment interactions. Breeding strategies depend on priority trait(s), target environment(s), and related resource allocation considerations, but it can be argued that plant breeders typically select for consistent, high performance across relevant environments. In this context, stability can be defined as genotype performance that is consistent with changes in the environment. This "dynamic" view also suggests that genotypes are not penalized for responding to favorable environments, so long as it is in a predictable fashion (Becker, 1981). An alternative way to think about stability is that a genotype should have minimal variance over environments (Röemer, 1917). But as environments continue to improve through the use of relevant management and mitigation strategies, this "static" characterization can be thought of as too rigid and unsustainable for most economically important traits of interest.

Instances when plant performance stability is not achieved can be attributed to genotype-by-environment (G×E) interaction or error in measurement of the phenotype of interest. G×E occurs when environmental sensitivity is not consistent across all genotypes (Falconer, 1952). Because the environment includes all nongenetic factors affecting plant growth, G×E is influenced by factors such as soil characteristics, climatic variables, and management practices in combination with the variable genetics of cultivars tested (Comstock and Moll, 1963). While there are some predictable environmental differences (such as soil types), there are many unpredictable aspects of the environment (such as year to year variation in weather). Due to the effects of G×E, superior performance in one environment does not necessarily correspond to superior cultivar performance in other environmental conditions. To address this, plant breeders

have implemented evaluation schemes that encompass locations and years, commonly referred to as multi-environment trials (MET; Bornhofen et al., 2018; Burgueño et al., 2008; Dias et al., 2018; Oliveira et al., 2020). The diversity of these METs allows plant breeders to assess the genotypic, phenotypic and environmental variability. Both, G×E and stability of plant performance, can be evaluated using METs, however they are analyzed in different ways.

Examining the performance of a plant variety across environments and relative to other varieties is accomplished through the use of stability analyses (Bernardo, 2002). One of the first analyses, described by Finlay and Wilkinson (1963), calculates coefficients (slope and deviation from the slope) from the regression of cultivar performance against an environmental index based on the average performance at each environment. Using Finlay-Wilkinson Regression (FWR), stability can be described in three ways. Type I stability describes genotypes that perform the same in all environments, illustrated by a regression slope of approximately 0 (Becker and Leon, 1988). Type II describes genotypes that change performance as environments improve, with an average cultivar having a slope of 1. Cultivars with a slope greater than 1 respond positively to more inputs (inputs being agronomic management strategies to increase performance, such as additional fertilizer or irrigation, for example). Type III is a measure of the deviations from the linear response calculated as mean squared error (MSE). Larger deviations of a variety to the expected performance across environments leads to an increase in Type III stability. Cultivars with large MSE could, for example, could be ones that are particularly sensitive to biotic stress at one or a few environments, therefore dramatically reducing performance. Using these parameters for stability analyses, there is a general consensus among breeders to aim to develop cultivars that perform near the top in all environments. For example, a breeder may be selecting for disease resistance. In this case, it would be advantageous to have

low disease damage across all locations to minimize the probability of a disaster, opting for Type I stability (Eskridge et al., 1991). Alternatively, the goal could be to identify genotypes that respond linearly with increased inputs. One practice of this is selecting and developing cultivars that consistently outperform check cultivars across a broad range of environments, using reliability of performance as a decision tool. This reliability is functionally related to Type II stability (Eskridge and Mumm, 1992).

Other types of models that can be used for these analyses include additive main effects, multiplicative interactions (AMMI) models, factorial regressions, and biadditive models. In an AMMI model, variance components for the main effects of genotype, environment, and G×E are partitioned, and principal component analysis is subsequently applied to the G×E portion with the goals of delineating mega-environments and selecting genotypes to maximize productivity and increase accuracy to improve recommendation and selections (Gauch 1988, 2013). For example, an advantage of FWR over AMMI, is that the values of slope and MSE are a direct dissection of stability measurements. This allows for more straightforward values of stability on an individual genotypes (varieties) or group basis, resulting in meaningful comparisons and conclusions as opposed to variances and principal components. However, AMMI is able to reduce the multi-dimensionality of G×E, while still capturing the variation across genotypes and environments, into fewer dimensions (Walsh and Lynch, 2014). Historically, FWR environmental indices have been estimated using the mean performance across all genotypes, or a subset of common genotypes, in each location to create an overall gradient of environments ranging from worst to best performing. Earlier researchers expressed the desire for an environmental index based on environmental factors that influence the performance of these genotypes, citing the lack of reliable technology and understanding of how to implement these

additional data as the reason for using average performance across available genotypes (Eberhart and Russell, 1966; Lin et al., 1986). Technologies now exist to measure plants and environments in efficient ways allowing for the potential to determine environmental potential independent of the composition of experiments. Work in genomic selection, prediction, and crop modelling have incorporated environmental variables into models to improve prediction accuracies, facilitate the prediction of genotypes performance in untested environments, and advance our understanding of the most influential climatic measurements on plant performance (Crossa et al., 1999; Heslot et al., 2014; Jarquin et al., 2014; Malosetti et al., 2016; Vargas et al., 1998; Vargas et al., 1999). Li et al., (2018) redefined the environmental index using direct climatic measurements of the testing locations, independent from plant performance data in maize. The climatic measurements can then be used in FWR to establish a biologically meaningful and environmentally descriptive index. Further research remains to specifically define the environmental variables most influential on plant performance traits and how these variables might differ across germplasm that have undergone differing amounts of selection.

Historically, plant breeders have sought to minimize and harness G×E by selecting broadly adapted cultivars and strategically placing varieties in well-suited environments. This study sought to determine if selection for productivity aligned with increased stability by comparing unselected materials from the Iowa Stiff Stalk Synthetic Population (BSSS) population to inbreds developed from one or more cycles of breeding founded in this originator population. The BSSS population was created in the 1930s by intermating 16 lines of maize with above average stalk strength (Lamkey, 1992) and is one of the most relevant maize populations contributing germplasm, primarily in the temperate regions of the world. Using BSSS-derived

lines provides a closed genetic system that shares a common genetic background and allows for assessments to be made on the effect of selection on plant performance stability.

With this unique collection of germplasm combined with the infrastructure of the Genomes To Fields (G2F) Initiative, a multi-environment evaluation was conducted in 2016 and 2017. The objectives of this study were to: (i) determine the role of varying levels of selection in phenotypic trait performance and trait stability and (ii) identify the most influential environmental variables on these phenotypic traits. Results from these analyses provide insight on how phenotypic performance and stability have been shaped by selection and the value of incorporating specific environmental measurements in future analyses, such as crop growth models and genomic prediction.

2.2 Materials and Methods

Germplasm

A total of 102 unique hybrids were planted in 31 environments ranging from 34.73 to 43.32 latitude and -75.43 to -94.72 longitude (Table 2.1) across 11 states in the United States in two years for a total of 3,042 plots in 2016 and 2,991 plots in 2017 as part of the G2F Initiative. Hybrids were developed by crossing the following inbreds founded in, or derived from, the BSSS maize population, to the non-Stiff Stalk expired-PVP (exPVP) inbred DK3IIH6. The 102 inbreds used reflect varying levels of selection: 10 of the 16 original BSSS founder lines, 19 lines randomly derived from BSSS cycle 0, 18 lines selected directly from a BSSS inbred, 35 lines with one or more BSSS-derived inbred(s) as a generational relative, and 20 double haploid (DH) inbred lines derived from a synthetic population that combines six core BSSS lines (B73, B84, LH145, PHB47, NKH8431 and PHJ40). Together, all 102 of these inbreds comprise four main groupings that will be evaluated in this study: Unselected, Selected, Advanced, and Synthetic (Supplementary Table 2.1; visual schematic in Supplementary Figure 2.1). Because the founder and C0 lines represent the original, unselected population, they were combined into a single group, Unselected, for analysis purposes. The Selected group contains lines that were selected from BSSS populations, while the Advanced group contains lines that derived from populations generated by lines that have been previously selected. The Advanced group contains lines that were both publicly and privately released for superior performance. Unlike the Unselected, Selected, and Advanced groups, the Synthetic is comprised of lines that vary in selection but were recombined to form this group. Five of the six Synthetic parents were grown in this study and are part of the Selected (B73 and B84) and Advanced (LH145, PHB47, and NKH8431) groupings. The Synthetic group will provide unique insight into the effect of

recombination in the maintenance of performance and stability compared to the Selected and Advanced groupings.

Along with using known pedigree information (Mazaheri et al., 2018), the four groupings were corroborated using genotypic data. The genotypic data was generated using genotyping-by-sequencing (GBS) single nucleotide polymorphisms (SNP; Elshire et al., 2011) and is available for 98 of the 102 inbred lines. After filtering for minor allele frequencies (MAF) less than 0.05, a total of 995,690 markers were used to assess population structure and confirm groupings. A distance matrix was calculated in TASSEL as 1 – identity by state (IBS; Bradbury et al., 2007). The resulting principal components 1 (PC1) and 2 (PC2) were used to create a multi-dimensional scaling (MDS) plot. The MDS plot was used to assess population structure across the 98 genotyped inbreds as the points that are closer together have a more similar genetic composition, while points further apart are more genetically different. Available GBS data can be found at doi.org/10.25739/w560-2114.

Experimental Design

All experiments were planted using a randomized complete block design (RCBD) with two replications per location per year, resulting in 31 location-year combinations to represent each unique environment. Two-row plot length ranged from 5.30 to 7.62 m long. Plot area ranged from 7 to 11.77 m^2 and planting density ranged from 75,753 to 102,790 seeds/hectare (Table 2.1; McFarland et al., 2020).

Data Collection

Phenotypic data were collected and/or estimated for the following traits (McFarland et al., 2020): flowering- anthesis date (MM/DD/YY), flowering- silking date (MM/DD/YY), ear height (cm), plant height (cm), root lodging (% of plot), stalk lodging (% of plot), stand count (% of seeds planted in plot), grain moisture (%), test weight (kg/hL), and grain yield (ton/ha). Flowering dates were taken after 50% of the plants in a plot were extruding anthers on more than half of the main tassel spike (anthesis) and after 50% of the plants in a plot exhibited silk emergence (silking). Ear height was measured as the distance from the root crown to the primary ear bearing node and plant height was measured as the distance from the base of the plant to the ligule of the flag leaf. Root lodging was measured as the percent of plants that show stems leaning substantially ($\geq 15^{\circ}$ from vertical) and stalk lodging was measured as the percent of plants broken between the ground and the primary ear node at harvest. Stand count reflects the percent of seeds planted in a plot that germinated and survived the growing season. Plot weight was the shelled grain weight per plot, grain moisture was the water content in grain at harvest, test weight was the shelled grain weight per hectoliter and grain yield was estimated in tons per hectare using specific parameters depending on the plot size and planting density for each environment. Test weight, grain moisture, and yield were all collected by grain combines at harvest.

The flowering traits were converted from days after planting to growing degree units (GDU) using the temperature collected leading up to that point. GDU was calculated using the equation: GDU = $\frac{minimum\ temperature\ (°F)+maximum\ temperature\ (°F)}{2} - 50$. If the daily temperature was greater than 86°F, the maximum temperature was set to 86 and if the daily temperature was less than 50°C then the minimum temperature was set to 50 in the equation

(Barger, 1969). The daily GDUs were added together depending on the date that each hybrid flowered for cumulative GDUs.

Environmental data contains weather/climatic data, soil data, and agronomic inputs. Weather data were collected at 22 of the 31 environments using Spectrum Technologies, Inc. WatchDog 2700 Weather Stations. Data were collected every 30 minutes for the entire season from planting through harvest and then compared to the nearest airport weather observation station to verify accuracy and fill in missing elements using the Iowa Environmental Mesonet archive (https://mesonet.agron.iastate.edu/request/download.phtml?network=OH_ASOS) as described in the 2016 doi.org/10.7946/P3MW27 and 2017 doi.org/10.25739/w560-2114 DOI. Soil data were collected from 25 environments once during the growing season using bulk soil samples of at least 10 soil cores per field at a depth of 30 cm. Soil samples were processed by Ward Laboratories, Inc. to estimate a variety of soil characteristics. Finally, agronomic inputs were the various management practices reported at 18 of the 31 environments throughout the growing season. These inputs occurred pre-planting and post-planting and were unique to each environment. Environmental variables were aggregated to reflect critical timepoints, or developmental stages, such as planting, flowering, and the accumulation from planting to harvest. The environmental data can be separated into six “areas” based on what they measure: day length, management, precipitation, soil, temperature, and wind. All of the aforementioned environmental data are listed and described in detail in Supplementary Table 2.2.

Statistical Analysis

Phenotypic Data

The phenotypic data used in this study only included hybrids that were represented in at least 12 of the 31 locations (Supplementary Table 2.1). While all hybrids were grown across all environments, poor germination (mainly for Unselected lines) and missing phenotypic notes in specific locations resulted in imbalances in the data set. The number of environments in which phenotypic traits were measured ranged from 21 (flowering- anthesis and flowering- silking) to all 31 (stand count, grain moisture, and grain yield).

Data were analyzed using the model $Y_{ijk} = \mu + E_i + r(E_j) + G_k + G*E_{ik} + \varepsilon_{ijk}$. 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 G×E ($G*E_{ik}$) was the interaction between the genotype and each environment. The i^{th} environment (E) represents the 31 unique environments, combination of location and evaluation year. This analysis was implemented using the *lme4* package (Bates et al., 2015) in R (R Core Team, 2018). All factors were analyzed as random effects and the average value of traits in each environment were used as the phenotype modeled in the regression analysis. To estimate the contribution of location versus year in the G×E, or $G*E_{ik}$ estimate, a second model was used: $Y_{ijk} = \mu + E_i + r(e)_j + G_k + G*L + G*Y + G*L*Y + \varepsilon_{ijk}$. All of the factors are the same as previously described, but the i^{th} environment of E_{ik} was broken down into *L,* the location tested in, and *Y,* the year tested in. To compute significance of the random-effect terms used in each model, likelihood ratio tests were used.

Summary phenotypic trait statistics (minimum, mean, maximum, and standard deviation) were calculated for the whole dataset across both years and each year, separately. The mean and standard deviation for each phenotypic trait were then calculated for each grouping separately.

Tukey's honest significant difference test was used to compare the means of each grouping to every other grouping across phenotypic traits to determine significant differences between them. This analysis used the *agricolae* package in R (Mendiburu, 2020) and a confidence level of 0.95. Pearson correlations between phenotypic traits along with significance were also calculated.

Finlay- Wilkinson Regression Stability Analysis

For the stability analysis, the average of each hybrid's trait performance was linearly regressed on a performance-based index, following the FWR. This performance-based index was based on the average performance of all 102 hybrids for a given trait in each environment according to the methods described in Gage et al., 2017 (https://github.com/joegage/GxE_scripts). Slope and MSE were calculated for each hybrid following a linear regression. To evaluate changes in stability with selection, the slope and MSE statistics were summarized on a group-level for each trait.

Partial Least Squares Regression Analysis

A partial least squares regression (PLSR; Wold et al, 2001) was used to identify the environmental covariates (EC) that have the most influence on each of the measured traits. PLSR is a dimension reduction approach that accommodates a large number of correlated variables simultaneously and was conducted using the *pls* package in R (Mevik and Wehrens, 2007). The goal of this regression analysis was to identify the ECs that are most correlated with the measured phenotypic traits across environments. The phenotypic measurements for nine traits (stand count, flowering-anthesis, flowering-silking, plant height, ear height, root lodging, stalk lodging, grain moisture, and grain yield) at each of the 6,033 plots was used as the response

variables, while the 56 ECs were the predictor variables. All environmental data (ECs) were centered and scaled and then a nonlinear iterative partial least squares (NIPALS) algorithm was run for each individual phenotypic trait using the 56 ECs (Supplementary Table 2.2). Any EC that had greater than 50% missing values for the phenotypic trait being evaluated was removed from that analysis. The PLSR resulted in a model comprised of EC components that best explain the variance between the response phenotypic trait and other predictors. The coefficients of the multiple response variables for each of the traits can then be used to assess how well the identified ECs model the phenotypic trait being evaluated. Model selection was based on the number of components, or ECs, that minimized the random mean squared error of prediction (RMSEP) and the number of variables matching the optimal component number with the largest magnitude of coefficients were selected.

To assess the accuracy of these ECs in modeling the phenotypic traits, a training set was constructed by randomly sampling 80% of the original 6,033 plots and using leave-one-out as a means for cross-validation. Using the 20% of plots previously withheld from the training model, to create the testing set, a multiple linear regression used the selected ECs for a given trait and multiplied the coefficient by the values for the specific environment that each hybrid was grown in. To account for random sampling bias, the training and testing subsetting and cross- validation were performed 100 times. The prediction accuracy was then estimated using the Spearman correlation between the predicted and actual phenotypic values on the 20% testing set. A Spearman rank test was used to assess the correlation between the predicted and actual phenotypic traits.

Three PLSR analyses were conducted. First, PLSR was used to identify the ECs most influential on each of the nine phenotypic traits (stand count, flowering-anthesis, flowering-

silking, plant height, ear height, root lodging, stalk lodging, grain moisture, and grain yield). Next, PLSR was run on each of the groupings (Unselected, Selected, Advanced, and Synthetic) for each phenotypic trait. Finally, best linear unbiased estimates (BLUEs) were calculated for each phenotypic trait with the same equation previously used for assessing trait variance. $Y_{ijk} = \mu + E_i + r(E_j) + G_k + G*E_{ik} + \varepsilon_{ijk}$. 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 G×E ($G*E_{ik}$) was the interaction between the genotype and each environment. The genotype (G) and G×E components were evaluated as fixed variables with everything else being random variables. The PLSR was then run using the trait BLUEs as the response variables, assessing which ECs are most influential on the G and G×E variances, compared to the first PLSR analysis which is assessing all variances contributing to the phenotype (mainly the environment (E), G, and G×E).

From the E+G+ G×E and G+ G×E PLSR analyses, the resulting response EC coefficients were used to calculate the mean phenotypic trait in each environment. Because stability analyses utilize an environmental index ordered from poorest to best performance, we wanted to see how the inclusion of ECs would influence an environment's "ranking." To evaluate how the ordering of these environments compared between the FWR, PLSR E+G+ G×E, and PLSR G+ G×E, a Spearman rank correlation was calculated.

.

2.3 Results

Grouping Genetic Analysis

The MDS plot indicates patterns of genetic similarity and dissimilarity between the 98 genotyped inbred lines (Supplementary Figure 2.2). Together, PC1 and PC2 accounted for 50.30% of the variation and individual lines. The Unselected and Synthetic groupings produce much tighter and distinct clusters, while the Selected and Advanced groupings are larger and overlap with multiple groupings. The Selected group overlaps with the Unselected group, which aligns with pedigree information since the Selected individuals are the result of selections made on the initial BSSS populations. As expected, the Selected and Advanced groupings also show overlap, given that Advanced lines are more closely related generationally to the Selected lines through subsequent selections, than they are the original Unselected lines.

The five parents of the Synthetic grouping with available genotypic information are highlighted on the MDS plot. The parents' location on the plot is surrounding the collection of Synthetic lines that resulted from their recombination. Since all of the Synthetic lines share the same parents, it is expected for these points to be close in terms of their genetic similarity. The location of B73, a Selected line, is more distant from the other Selected lines. This could be due to B73's pervasiveness in Stiff Stalk germplasm and, therefore, high genetic relation to Advanced lines (Mikel and Dudley, 2006), or be an artifact since B73 is the reference genome used to call SNPs.

Phenotypic Data Variation and G×E

Across the 31 environments, mean grain yield was 10.61 (ton/ha) with unique environment averages ranging from 7.30 ton/ha (MO1_16) to 12.73 ton/ha (WI1_17). The average stand count ranged from 63% in IL1_17 to 86.03% in WI2_16, with the average across

all environments being 75.8%. Flowering time for all hybrids across environments ranged from 854.51 to 1658.81 GDUs for anthesis and 906.11 to 1854.1 for silking. Both stalk and root lodging displayed wide ranges of percentages, from 0% to 90% and 0% to 100%, respectively, with plot averages of 9.53% for stalk lodging and 5.03% for root lodging. When comparing phenotypic traits in 2016 to 2017, the average grain yield was higher in 2017 (11.2 ton/ha) than in 2016 (9.98 ton/ha) and the range of flowering times in 2017 (435.46 GDUs for anthesis and 448.1 GDUs for silking) was narrower than in 2016 (831.55 GDUs for anthesis and 999.6 GDUs for silking). All traits were significantly different ($P<0.05$) between years, indicating that there are nongenetic features across the environments that are inconsistent from year to year. This, along with the fact that not all locations were grown in multiple years, supports the decision to analyze each combination of location and evaluation year separately instead of combining locations across years.

Strong positive Pearson correlations ($r>0.75$; Figure 2.1) were observed between flowering- anthesis and flowering- silking, and ear height and plant height. Moderately weak Pearson correlations ($r<0.25$) were observed between stalk lodging and grain yield, and flowering- silking and grain yield.

Comparing mean trait performance across the selection groupings demonstrates that selection has played an important role. The Unselected grouping consistently had the lowest average performance across all groupings (stand count, 75.5%; stalk lodging, 13.7%; and grain yield, 9.33 ton/ha) and the latest flowering (anthesis, 1325 GDUs; silking, 1357 GDUs). Conversely, there are significant differences ($P<0.05$) and positive performance improvements from the Unselected to the Selected grouping for stand count (77.8%), stalk lodging (8.74%), test weight (70.1kg/hL), and grain yield (11.4 ton/ha). As expected, the Advanced group was the

earliest of all groupings (anthesis, 1280 GDUs; silking, 1300 GDUs), had the lowest plant and ear heights (235 and 114 cm, respectively), and the lowest stalk and root lodging (7.77 and 5.55%, respectively). These findings are consistent with known breeding goals over the years for better stand establishment and early seedling vigor, shorter plant stature, and higher overall productivity, demonstrating the influence of selection on this set of germplasm.

Recombination of selected inbreds with good performance results in transgressive segregation, including population members and overall mean with reduced agronomic performance and yield. For example, the Synthetic grouping's stand count is significantly improved compared to Unselected (76.4% compared to 75.5%), but less than Selected (77.8%); flowering (anthesis and silking) is significantly earlier (anthesis, 1290 GDU; silking, 1300 GDU) than Unselected and Selected but still later than Advanced; and grain yield is significantly higher (10.9 ton/ha) than Unselected, but lower than Selected and Advanced.

Genotypic, environmental, and G×E variances were significant ($P<0.05$) for all traits. Across all phenotypic traits, the percent of G×E variance was least for flowering (0.71% for silking and 0.73% for anthesis) and greatest for grain yield (7.33%) and stand count (7.74%; Table 2.2). For stand count, root lodging, and stalk lodging, the G×E variance was greater than the genotype variance. This means that the latter are likely be the most unstable traits evaluated in this study. When dissecting G×E and evaluating genotype × location, genotype × year, and genotype × location × year effects, the majority of variance can be attributed to the genotype × location × year and genotype × location components (Table 2.3). While climatic variables can alter plant performance and G×E from year to year in a given location, the overall geographic locations that these hybrids were tested in resulted in more phenotypic variation than year-to-year.

FWR Stability

The FWR was used to evaluate the stability of 102 hybrids across the test environments. Using the previously mentioned groupings, stability was described on a per-group basis using the extracted slope and MSE values from the linear regressions. The average MSE did not significantly change across groupings for stand count, flowering- anthesis, plant height, ear height, test weight, or grain moisture (Figure 2.2). The most significant differences ($P<0.05$) were in the MSE reductions from Unselected to Advanced for flowering- silking and grain yield and the MSE increases for stalk lodging. The ranges of MSE were the largest for root and stalk lodging. The general trend from Unselected to Selected to Advanced groupings was a reduction in average MSE for grain yield, stand count, flowering- anthesis, flowering- silking, plant height, ear height, and grain moisture. When comparing the Advanced and Synthetic groups, the average MSE was higher for the Synthetic group for all traits, except root lodging.

The overall range in slope values was from 0.09 (root lodging) to 1.51 (grain moisture; Figure 2.3). Since no genotypes performed exactly the same in all environments (which would be indicated by a slope of 0), it can be concluded that Type I stability was not observed in this experiment. The average slopes showed no significant changes across the groupings for stand count, flowering- anthesis, flowering- silking, and grain yield. However, there were significant differences ($P<0.05$) from Unselected to Advanced for plant height, ear height, stalk lodging, root lodging, and test weight. The slopes for these traits were closer to 1 for the Advanced grouping, except for stalk lodging, which had an average slope of 0.79. Stalk lodging was the only trait that had significant differences in the mean slope across all groupings. Overall, the stability of Advanced across most trait performances reflected Types II and III stability (slope of approximately 1 and low MSE, respectively) compared to the Unselected grouping, supporting

the idea that performance stability has improved with selection, with the exception of stalk lodging.

Environmental Covariates and PLSR

The 56 ECs that were used as potential predictor variables in the PLSR showed some multicollinearity (Supplementary Figure 2.3), and ultimately 20 were deemed influential based on the number of components and regression coefficients across the three analyses. These ECs related to the environmental "areas" of day length (n=2), management (n=6), precipitation (n=1), soil (n=12), temperature (n=5), and wind (n=4).

When using PLSR to identify ECs across phenotypic traits, the number of ECs ranged from two variables (flowering- anthesis and flowering- silking) to six (stand count and stalk lodging). While no ECs were influential on all phenotypic traits, the photoperiod at flowering-anthesis, soil phosphorous, soil soluble salts, and soil organic matter were the most commonly identified in four of the traits. It was also observed that the highest yielding environments had greater clay content and more organic matter in the soil, while the lowest yielding environments were classified as sandier soil textures.

Using the regression coefficients of selected ECs, trait predictability can be assessed to see how much variability of the performance is captured by the PLSR EC model. The mean prediction accuracy was greater than 0.66 across all phenotypic traits, demonstrating relatively high predictive ability and illustrating the utility of ECs in these prediction models. Prediction accuracies for the 100 permutations of each trait ranged from 0.45±0.04 for root lodging and 0.84±0.01 for flowering- anthesis (Figure 2.4). The traits with the highest prediction accuracies exhibited the highest environmental (E) variance and the lowest G×E variance (flowering-

anthesis and silking). Alternatively, the traits with the lowest prediction accuracies exhibited the lowest E variance and the highest G×E variance (stalk and root lodging). This suggests that when using all variances contributing to the phenotype in the PLSR analysis, the selected ECs are mainly predicting the environment. An interesting note was that stalk lodging and root lodging had G×E variance that was greater than the estimated genotype variance. This small genotypic variance indicates that there weren't differences among genotypes, but the ranking of genotype performance was not consistent across environments. From the FWR analyses, root lodging and stalk lodging were two of the least stable traits; both having the largest MSE values and stalk lodging slope deviating more greatly from 1 after selection.

To assess whether selection has affected which ECs were the most influential across phenotypic traits, PLSR was subsequently conducted on a per-group basis for each of the traits (Supplementary Table 2.3). While no ECs were influential on all phenotypic-grouping combinations, the most commonly identified ECs were soil phosphorus (n=17), photoperiod at anthesis (n=13), soil nitrogen (n=12), soil organic matter (n=12), and soil pH (n=11). When looking at the 145 identified phenotypic-grouping ECs, the majority are related to soil measurements (n=79) and the least are related to management (n=14) and precipitation (n=2). No significant differences in ECs or EC coefficient values were identified across groupings ($P<0.05$). This suggests that genotypes having undergone more selection don't respond significantly more to agronomic inputs than the other groupings based on the environmental measurements in this experiment. The substantial role of soil characteristics on plant performance and productivity is also evident.

Finally, BLUEs of the phenotypic traits composed of the genetic and G×E variances were used as the response variables. The ECs and corresponding coefficients were different from the

first PLSR results of phenotypic traits composed of E+G+ G×E effects, such that the ECs for flowering traits changed from soil characteristics with coefficients near 0 (E+G+ G×E) to daylength and temperature with coefficients greater than 10 (G+ G×E model; Figure 2.5). The EC coefficients from the G+ G×E and E+G+ G×E models for grain yield were used estimate mean grain yield performance of each environment and then the ranking of environments (lowest to highest grain yield) was compared using a Spearman rank correlation. Between the G+ G×E and E+G+ G×E models, there was no correlation between the rankings (ρ=0.1, P=0.77). When comparing the ranking of environments from the previously mentioned FWR index, there was no significant correlation with the E+G+ G×E model (ρ=0.43, P=0.19), but a significant correlation with the G+ G×E environment ranking (ρ=0.56, P=0.08).

While these two PLSR models are measuring different components of the observed phenotypic variation, as expected, there are substantial differences in the ECs identified between the two models and also in relationship to the FWR results. These results suggest that using BLUEs of the G+ G×E effect of a trait generate similar environmental rankings as the traditionally used FWR linear regression. While the E+G+ G×E model produces environmental rankings different from the G+ G×E and FWR approaches, this model does explain the environmental variance, which is a dominant contributor to several phenotypic traits.

2.4 Discussion

Our analyses using germplasm originating from the BSSS population reveals how performance and stability have been shaped by selection. When assessing stability measurements across the groupings, the Advanced grouping showed improved slope near 1 compared to the Unselected grouping for all traits except stalk lodging. Previous literature has argued that selection for high-yielding environments not only selects for lines with higher mean performance, but also for lines that are more responsive to "good" environments, reflected by a slope greater than 1 (Simmonds, 1991). From the FWR, we found an increase in stability (slope near 1 and reduction in deviation from the linear regression) from Unselected to Advanced groupings for stand count, flowering, plant height, ear height, and grain yield. These traits have been known breeding goals over the years for better germination, earlier maturity, shorter plants, and higher overall productivity. Previous work has noted that the genotypes exhibiting the highest yield differences above location averages were cultivars advanced by private companies and widely used for commercial production (Boyles et al., 2019; Campos et al., 2004), and a number of the well-established lines from the Advanced grouping are exPVPs. Our findings suggest that selecting for higher performance did not mean sacrificing stability across most traits.

The trends from the Unselected to Advanced groupings suggest that while selection for productivity also increased stability, stability itself is heritable. The trends from the Advanced to the Synthetic group show a decrease in stability. By recombining lines with differing levels of stability from the Unselected and Selected groups to form the Synthetic group, variation is reintroduced. As previously, there is evidence for a genetic basis for stability (Kraakman, et al., 2004; Pigliucci, 2005; Rae et al., 2007; Mickelbart et al., 2015). Acknowledging that stability estimates are relative to the range of environments being tested in, and that the deviation from the linear regression line is influenced by model fit (Lin et al., 1986), the measurements of slope

across groupings still demonstrate a shift toward a slope of 1. This reinforces the idea that similar to other traits, performance stability or instability can be introduced, maintained, or lost through breeding and is transmissible to progeny.

The traits measured in this study showed significant variation due to G×E. The traits that expressed the least G×E variance and most stability were those that directly control the plant's ability to reproduce (i.e., flowering traits). The growth and maturation of the tassels and silks are controlled primarily by daylength and temperature of individual environments (Eisele, 1938), which explains why one of the most commonly influential ECs across all phenotypic-grouping combinations was related to day length. Using the identified EC coefficients, the flowering traits, which exhibited the least G×E and are dominated by environmental variance had the highest prediction accuracy. Alternatively, traits that expressed the greatest G×E variance and comparatively less environmental variance were those that represent agronomic strength and productivity- stalk and root lodging, stand count, and grain yield. Wind ECs were only influential on lodging and productivity-related traits.

Deriving stability measurements from linear regression coefficients has been practiced for decades using an index based on plant performance and uses the extracted values of slope and MSE to describe stability. However, these aforementioned indices are only useful for previously assessed environments and only represent a single measurement of the environment. By incorporating ECs into this linear regression, we can estimate changes in single or multiple ECs and predict performance in new environments with relatively high accuracy (Oliveira et al., 2020). Two of the ECs that were most influential across the most traits were photoperiod at anthesis and soil phosphorus. Photoperiod at anthesis, or the hours of daylight at the time of flowering-anthesis, has been heavily studied as it is closely related to grain yield. With longer

photoperiods, there is an increase in vegetative growth and development prior to tassel initiation (Kannenberg et al., 1974; Kanemasu and Warrington, 1983). Phosphorous is considered an essential macronutrient in plant growth and development. It has a critical role in photosynthesis and respiration reactions so without it, plant growth is diminished, maturity delayed, and yield reduced (Mollier and Pellerin, 1999; Iowa State University Extension and Outreach, 2000). An interesting finding was the level of importance of soil characteristics in trait performance compared to the minimal effect of management practices, especially on highly selected genotypes. Even though management practices (i.e., irrigation, herbicide application, fertilizer application, etc.) and the addition of inputs onto fields can positively impact maize yield improvements, we did not find that management was more influential for any single grouping.

Future work should focus on identifying ECs at different growth stages instead of simply planting, flowering, and end of season stages/traits, which was a limitation of the available data. The results of this work have demonstrated that plant performance stability has improved (in regard to slope near 1 and deviation from the linear regression line) with selection for all traits except stalk lodging, and that the most influential ECs on plant performance were related to soil and day length characteristics. Our work also suggests that PLSR can be used to identify ECs that best explain the environmental variance attributed to a given trait (E+G+ G×E model) or ECs that more explain the genetic and G×E variance. The G+ G×E model environmental rankings are significantly correlated to those of FWR linear regression, while the E+G+ G×E model rankings are not correlated with either of the previously mentioned rankings. This suggests that environmental variance estimated across environments is not necessarily indicative of hybrid performance rankings and if the goal is to rank environment based on hybrid performance, the G+ G×E model is better able to explain performance. Being able to connect

plant performance with G×E and stability has illustrated the role of selection in maize and the utility of dissecting environments into separate components (i.e., soil characteristics, precipitation, day length, management practices, etc.).

Acknowledgments

The authors would like to thank the Genomes to Fields consortium for providing the data and materials used in this study. This consortium involves more than 30 researchers representing more than 20 research institutions. Details about the initiative and publicly available resources can be found at www.Genomes2Fields.org. Funding has been provided for this research and publication from the USDA National Institute of Food and Agriculture (NIFA) project award #2016-67024-2219.

LITERATURE CITED

Barger, G.L. 1969. Total Growing Degree Days. Weekly Weather & Crop Bulletin 56:18. U.S. Dept. of Commerce and USDA, Washington, D.C.

Bates, D., M. Mächler, B.M. Bolker, and S.C. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1-48.

Becker, H.C., and J. Léon. 1988. Stability Analysis in Plant Breeding. *Plant Breeding.* 101(1): 1–23.

Bernardo, R. 2002. Breeding for quantitative traits in plants. Stemma, Woodbury, MN.

Bornhofen, E., G. Benin, L. Storck, L.G. Woyann, T. Duarte, M.G. Stoco, and S.V. Marchioro. 2017. Statistical methods to study adaptability and stability of wheat genotypes. *Bragantia* 76(1). https://dx.doi.org/10.1590/1678-4499.557.

Boyles, R.E., Z.W. Brenton, and S. Kresovich. 2019. Genetic and genomic resources of sorghum to connect genotype with phenotype in contrasting environments. *The Plant Journal* 97(1): 19-39. https://doi.org/10.111/tpj.14113.

Burgueño, J., J. Crossa, P.L. Maizeelius, and R.C. Yang. 2008. Using factor analytic models for joining environments and genotypes without crossover genotype x environment interaction. *Crop Science* 48(4): 1291-1305. https://doi.org/10.2135/cropsci2007.11.0632.

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23(19): 2633-2635. https://doi.org/10.1093/bioinformatics/btm308.

Campos, H., M. Cooper, J.E. Habben, G.O. Edmeades, and J.R. Schussler. 2004. Improving drought tolerance in maize: A view from industry. *Field Crops Research* 90(1):19-34. https://doi.org/10.1016/j.fcr.2004.07.003.

Comstock, R.E., and R.H. Moll. 1963. Genotype-by-environment interactions. *Statistical genetics and plant breeding*. Natl. Acad. Sci., Natl. Res. Counc., Washington, D.C. p. 164-194.

Cornelius, P.L., and J. Crossa. 1999. Prediction assessment of shrinkage estimators of multiplicative models for multi-environment cultivar trials. *Crop Science* 39(4):998-1009. https://doi.org/10.2135/cropsci1999.0011183X003900040007x.

Darrah, L.L., and M.S. Zuber. 1986. 1985 United States Farm Maize Germplasm Base and Commercial Breeding Strategies. *Crop Science.* 26(6): 1109-1113. https://doi.org/10.2135/cropsci1986.0011183X002600060004x.

De, F., M. Maintainer, and F. De Mendiburu. 2017. Package "agricolae". *Statistical Procedures*

for Agricultural Research. https://myaseen208.github.io/agricolae/https://cran.r-project.org/package=agricolae.

Dias, K.O.D.G., S.A. Gezan, C.T. Guimarães, A. Nazarian, L. Da Costa E Silva, S.N. Parentoni, P.E. De Oliveira Guimarães, C. De Oliveira Anoni, J.M.V. Pádua, M. De Oliveira Pinto, R.W. Noda, C.A.G. Ribeiro, J.V. De Magalhães, A.A.F. Garcia, J.C. De Souza, L.J.M. Guimarães, and M.M. Pastina. 2018. Improving accuracies of genomic predictions for drought tolerance in maize by joint modeling of additive and dominance effects in multi-environment trials. *Heredity* 121(1):24-37. https://doi.org/10.1038/s41437-018-0053-6.

Eberhart, S.A., and W.A. Russell. 1966. Stability parameters for comparing varieties. *Crop Science* 6: 36-40.

Edwards, J.W. 2016. Genotype × environment interaction for plant density response in maize (Zea mays L.). *Crop Science* 56(4): 1493-1505. https://doi.org/10.2135/cropsci2015.07.0408.

Eisele, H. F. 1938. Influence of environmental factors on the growth of the maize plant under field conditions. Iowa Agriculture and Home Economics Experiment Station Research Bulletin.

Elshire, R.J, J.C. Glaubitz, S. Qi, J.A. Poland, K. Kawamoto, E.S. Buckler, S.E. Mitchell. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* 6(5): e19379. https://doi.org/10.1371/journal.pone.0019379.

Eskridge, K.M., P.F. Byrne, J. Crossa. 1991. Selection of stable cultivars by minimizing the probability of disaster. *Field Crops Research* 27(1-2): 169-181. https://doi.org/10.1016/0378-4290(91)90029-U.

Eskridge, K.M., and R.F. Mumm. 1992. Choosing plant cultivars based on the probability of outperforming a check. *Theoretical and Applied Genetics* 84:494-500. https://doi.org/10.1007/BF00229512.

Falcon, C.M., S.M. Kaeppler, E.P. Spalding, N.D. Miller, N. Haase, N. AlKhalifah, M. Bohn, E.S. Buckler, D.A. Campbell, I. Ciampitti, L. Coffey, J. Edwards, D. Ertl, S. Flint-Garcia, M.A. Gore, C. Graham, C.N. Hirsch, J.B. Holland, D. Jarquín, J. Knoll, N. Lauter, C.J. Lawrence-Dill, E.C. Lee, A. Lorenz, J.P. Lynch, S.C. Murray, R. Nelson, M.C. Romay, T. Rocheford, P.S. Schnable, B. Scully, M. Smith, N. Springer, M.R. Tuinstra, R. Walton, T. Weldekidan, R.J. Wisser, W. Xu, and N. de Leon. 2020. Relative utility of agronomic, phenological, and morphological traits for assessing genotype-by-environment interaction in maize inbreds. *Crop Science*. 60(1): 62–81.

Falconer, D.S. 1952. The Problem of Environment and Selection. *American Naturalist* 86(830): 293-298.

Gauch R., H., and W. Zobel. 1996. AMMI Analysis of Yield Trials. *In* Genotype-by-Environment Interaction.

Gauch G. H. 2013. A Simple Protocol for AMMI Analysis of Yield Trials. *Crop Science* 53(5):1860-1869. https://doi.org/10.2135/cropsci2013.04.0241.

Helms, T.C., A.R. Hallauer, O.S. Smith. 1989. Genetic drift and selection evaluated from recurrent selection programs in maize. *Crop Science* 29(3):602-607. https://doi.org/10.2135/cropsci1989.0011183X002900030009x.

Heslot, N., D. Akdemir, M.E. Sorrells, and J.L. Jannink. 2014. Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theoretical Applied Genetics* 127(2):463-480. https://doi.org/10.1007/s00122-013-2231-5.

Hunter, R.B., L.A. Hunt, and W. Kannenberg. 1974. Photoperiod and temperature effects on maize. *Canadian Journal of Plant Science* 57: 1127-1133.

Iowa State University Extension and Outreach. 2000. Phosphorus Basics. Available at https://crops.extension.iastate.edu/encyclopedia/phosphrus-basics.

Jarquín, D., J. Crossa, X. Lacaze, P. Du Cheyron, J. Daucourt, J. Lorgeou, F. Piraux, L. Guerreiro, P. Pérez, M. Calus, J. Burgueño, and G. de los Campos. 2014. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretial and Applied Genetics* 127(3), 595-607. https:doi.org/10.1007/s00122-013-2243-1.

Kraakman, A.T.W., R.E. Niks, P.M.M.M. Van Den Berg, P. Stam, and F.A. Van Eeuwijk. 2004. Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics* 168(1):435-446. https://doi.org/10.1534/genetics.104.026831.

Lamkey, K.R. 1992. Fifty Years of Recurrent Selection in the Iowa Stiff Stalk Synthetic Maize Population. *Maydica* 37(1), 19-28.

Li, X., T. Guo, Q. Mu, X. Li, and J. Yu. 2018. Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proceedings of the National Academy of Sciences* 115(26), 6679-6684. http://www.pnas.org/content/115/26/6679.abstract.

Lin, C. S., Binns, M. R. & Lefkovitch, L. P. 1986. Stability analysis: where do we stand? *Crop Science* 26(5), 894-900. https://doi.org/10.2135/cropsci1986.0011183X002600050012x.

Lynch, M. and Walsh, B. 1998. Genetics and Analysis of Quantitative Traits, 1st edn. Sinauer, Sunderland, MA.

Malosetti, M., D. Bustos-Korts, M.P. Boer, and F.A. Van Eeuwijk. 2016. Predicting responses in multiple environments: Issues in relation to genotype × Environment interactions. *Crop Science* 56(5), 2210-2222. https://doi.org/10.2135/cropsci2015.05.0311.

Mazaheri, M., M. Heckwolf, B. Vaillancourt, J.L. Gage, B. Burdo, S. Heckwolf, K. Barry, A. Lipzen, C.B. Ribeiro, T.J.Y. Kono, H.F. Kaeppler, E.P. Spalding, C.N. Hirsch, C. Robin Buell, N. de Leon, and S.M. Kaeppler. 2019. Genome-wide association analysis of stalk biomass and anatomical traits in maize. *BMC Plant Biology* 19(1). https://doi.org/10.1186/s12870-019-1653-x.

McFarland, B.A., N. Alkhalifah, M. Bohn, J. Bubert, E.S. Buckler, I. Ciampitti, J. Edwards, D. Ertl, J.L. Gage, C.M. Falcon, S. Flint-Garcia, M.A. Gore, C. Graham, C.N. Hirsch, J.B. Holland, E. Hood, D. Hooker, D. Jarquin, S.M. Kaeppler, J. Knoll, G. Kruger, N. Lauter, E.C. Lee, D.C. Lima, A. Lorenz, J.P. Lynch, J. McKay, N.D. Miller, S.P. Moose, S.C. Murray, R. Nelson, C. Poudyal, T. Rocheford, O. Rodriguez, M.C. Romay, J.C. Schnable, P.S. Schnable, B. Scully, R. Sekhon, K. Silverstein, M. Singh, M. Smith, E.P. Spalding, N. Springer, K. Thelen, P. Thomison, M. Tuinstra, J. Wallace, R. Walls, D. Wills, R.J. Wisser, W. Xu, C.T. Yeh, and N. De Leon. 2020. Maize genomes to fields (G2F): 2014-2017 field seasons: Genotype, phenotype, climatic, soil, and inbred ear image datasets. *BMC Research Notes* 13(1): 4–9. https://doi.org/10.1186/s13104-020-4922-8.

Mevik, B.H., and R. Wehrens. 2007. The pls package: Principal component and partial least squares regression in R. *Journal of Statistical Software* 18(2). https://doi.org/10.18637/jss.v018.i02.

Mickelbart, M. V, P.M. Hasegawa, and J. Bailey-Serres. 2015. Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. *Nature Review Genetics* 16: 237 https://doi.org/10.1038/nrg3901.

Mikel, M.A., and J.W. Dudley. 2006. Evolution of North American dent maize from public to proprietary germplasm. *Crop Science* 46(3): 1193-1205. https://doi.org/10.2135/cropsci2005.10-0371.

Mollier, A., and S. Pellerin. 1999. Maize root system growth and development as influenced by phosphorus deficiency. *Journal of Experimental Botany* 50:487-497.

Oliveira, I.C.M., J.H.S. Guilhen, P.C. de O. Ribeiro, S.A. Gezan, R.E. Schaffert, M.L.F. Simeone, C.M.B. Damasceno, J.E. de S. Carneiro, P.C.S. Carneiro, R.A. da C. Parrella, and M.M. Pastina. 2020. Genotype-by-environment interaction and yield stability analysis of biomass sorghum hybrids using factor analytic models and environmental covariates. *Field Crops Research* 257.

Pigliucci, M. 2005. Evolution of phenotypic plasticity: where are we going now? *Trends in Ecology & Evolution* 20(9):481-486. https://doi.org/10.1016/j.tree.2005.06.001

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Röemer J. 1917. Sinde die ertagsreichen sorten ertagissicherer? Mitteilungen DLG.

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

Vargas, M., J. Crossa, F.A. Van Eeuwijk, M.E. Ramírez, and K. Sayre. 1999. Using partial least squares regression, factorial regression, and AMMI models for interpreting genotype x environment interaction. *Crop Science* 39(4):955-967. https://doi.org/10.2135/cropsci1999.0011183X003900040002x.

Vargas, M., J. Crossa, K. Sayre, M. Reynolds, M.E. Ramírez, and M. Talbot. 1998. Interpreting genotype x environment interaction in wheat by partial least squares regression. *Crop Science* 38(3):679-689. https://doi.org/10.2135/cropsci1998.0011183X003800030010x.

Warrington, I.J., and E.T. Kanemasu. 1983. Maize Growth Response to Temperature and Photoperiod II. Leaf-Initiation and Leaf-Appearance Rates 1. *Agronomy Journal* 75(5):755-761. https://doi.org/10.2134/agronj1983.00021962007500050009x.

Wold, S., M. Sjöström, and L. Eriksson. 2001. PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* 58(2):109-130. https://doi.org/10.1015/S0169-7439(01)00155-1.

Table 2.1 Testing locations' latitude, longitude, planting density, planting and harvest dates for each year for 31 locations that are part of the Genomes to Fields project field evaluation in 2016 and 2017.

Location	Location Code	Year	Latitude (°)	Longitude (°)	Plant density (plants hectacre^{-1})	Planting date(s)	Harvest date(s)
Jonesboro, AR	AR1_16	2016	34.73	-90.76	75,753	4/7/16	8/30/16
	AR1_17	2017	34.73	-90.76	75,753	4/17/17	9/11/17
Marianna, AR	AR2_17	2017	35.67	-90.08	75,753	4/25/17	9/16/17
Georgetown, DE	DE_16	2016	38.65	-75.45	101,141	4/25/16	9/14/16
	DE_17	2017	38.67	-75.43	101,141	4/28/17	9/7/17, 9/8/17
Crawfordsville, IA	IA1_16	2016	41.20	-91.49	99,168	4/22/16	10/1/16
	IA1_17	2017	41.20	-91.50	99,168	4/25/17	10/26/17
Carroll, IA	IA2_16	2016	42.07	-94.73	99,168	4/25/16	10/11/16
	IA2_17	2017	42.06	-94.72	99,168	5/6/17	10/31/17
Keystone, IA	IA3_16	2016	41.99	-92.26	99,168	4/24/16	10/6/16
	IA3_17	2017	41.98	-92.24	99,168	4/25/17	10/20/17
Ames, IA	IA4_16	2016	41.99	-92.26	99,168	4/26/16	10/17/16
	IA4_17	2017	41.99	-92.24	99,168	5/7/17	10/16/17, 10/17/17
Champaign, IL	IL_16	2016	40.06	-88.23	98,964	4/26/16, 5/6/16	10/9/16
	IL_17	2017	40.06	-88.23	98,964	4/26/17	10/9/17, 10/16/17
West Lafayette, IN	IN_16	2016	40.48	-86.99	97,117	5/19/16	10/5/16, 10/6/16
	IN_17	2017	40.47	-86.99	97,117	5/16/17	10/19/17
East Lansing, MI	MI_16	2016	42.41	-84.30	97,117	5/24/16	11/16/16
	MI_17	2017	42.68	-84.49	97,117	5/22/17	10/20/17
Columbia, MO	MO_16	2016	38.90	-92.21	89,014	5/23/16	10/3/16, 10/4/16, 10/5/16, 10/7/16
	MO_17	2017	38.89	-92.21	89,014	5/15/17	10/19/17
Aurora, NY	NY1_16	2016	42.73	-76.66	94,281	5/19/16	11/14/16
	NY1_17	2017	42.73	-76.65	94,281	5/20/17	11/22/17
Aurora, NY	NY2_16	2016	42.73	-76.66	94,281	5/10/16	11/15/16, 11/18/16, 12/8/16
	NY2_17	2017	42.73	-76.65	94,281	5/18/17	11/24/17
South Charleston, OH	OH_16	2016	39.86	-83.68	102,274	5/27/16	10/15/16
	OH_17	2017	39.86	-83.67	102,274	5/16/17	10/20/17
Madison, WI	WI1_16	2016	43.06	-89.53	85,773	5/9/16	10/14/16
	WI1_17	2017	43.06	-89.53	85,773	5/5/17	10/19/17
Arlington, WI	WI2_16	2016	43.33	-89.34	102,790	5/24/16	10/25/16
	WI2_17	2017	43.32	-89.34	85,773	5/11/17	11/6/17

Table 2.2. Percent of phenotypic variance explained by each component of the random effects model for various phenotypic traits. These are based on the measurements of 102 genotypes in 31 environments of the Genomes to Fields project evaluated in 2016 and 2017.

		Variance of Phenotypic Traits (in %)									
		Stand Count	Flowering-Anthesis	Flowering-Silk	Plant Height	Ear Height	Root Lodging	Stalk Lodging	Grain Moisture	Test Weight	Grain Yield
Source of variation	Environment	41.60	89.37	86.74	86.47	67.35	68.81	43.59	45.25	78.69	46.61
	Genotype	4.19	4.16	6.27	3.83	10.16	1.48	6.70	14.1	1.3	13.77
	Genotype x Environment	7.85	0.82	0.84	1.35	3.31	9.34	15.92	10.94	0.95	8.76
	Rep Within Environment	4.91	2.7	2.26	0.73	1.28	0.99	5.04	3.23	0.59	2.52
	Residual	41.44	2.96	3.89	7.45	17.64	19.38	28.75	25.14	18.47	28.34

Table 2.3. Percent of phenotypic variance using the random effects model, comparing G*Y, G*L and G*L*Y across phenotypic traits. These are based on the measurements of 102 genotypes in 31 environments of the Genomes to Fields project evaluated in 2016 and 2017.

		Variance of Phenotypic Traits (in %)									
		Stand Count	Flowering-Anthesis	Flowering-Silk	Plant Height	Ear Height	Root Lodging	Stalk Lodging	Grain Moisture	Test Weight	Grain Yield
Source of variation	Genotype x Year (G*Y)	5.04	0.19	0.07	0.11	0.58	0	1.01	0.76	0.31	0.96
	Genotype x Location (G*L)	0	0.23	0.33	0.37	0.71	1.1	0	5.54	0.21	2.75
	Genotype x Location x Year (G*L*Y)	5.51	0.47	0.43	0.93	2.25	8.25	15.54	5.5	0.33	5.22

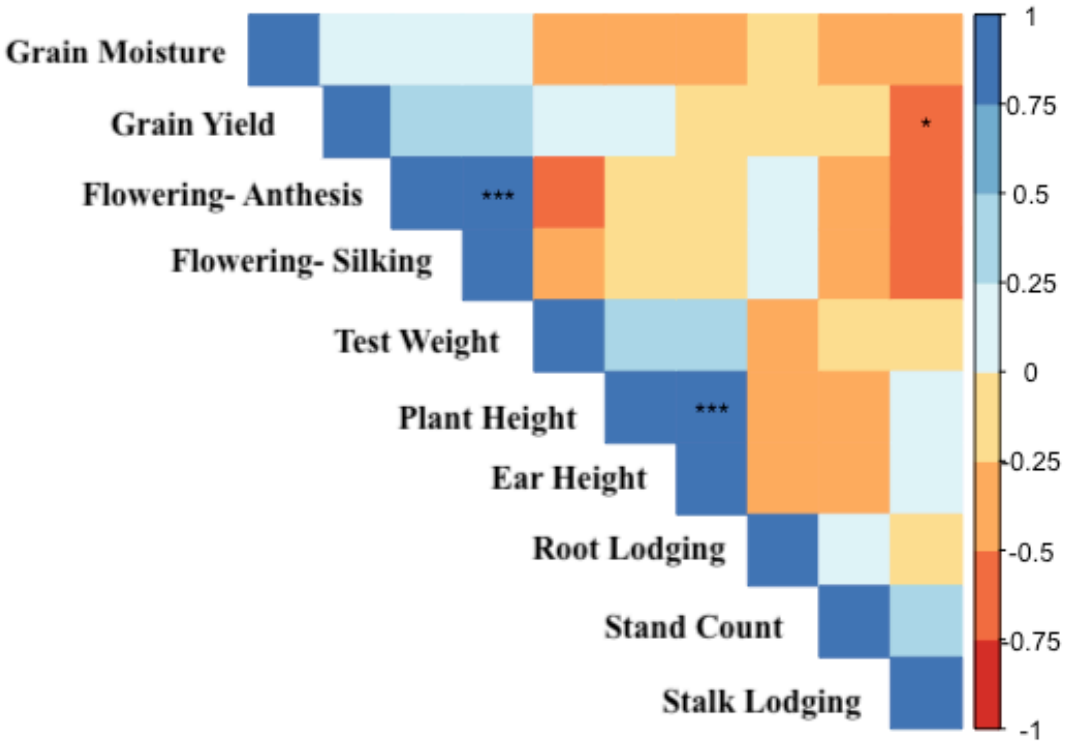


Figure 2.1. Pearson correlations between 10 phenotypic traits measured across 102 maize hybrids in 21 locations of the Genomes to Fields experiment in 2016 and 2017. Magnitude ranges from -1 (red) to 1 (dark blue) and significance of relationship is indicated inside box ($P<0.05$ = *, $P<0.001$ = ***).

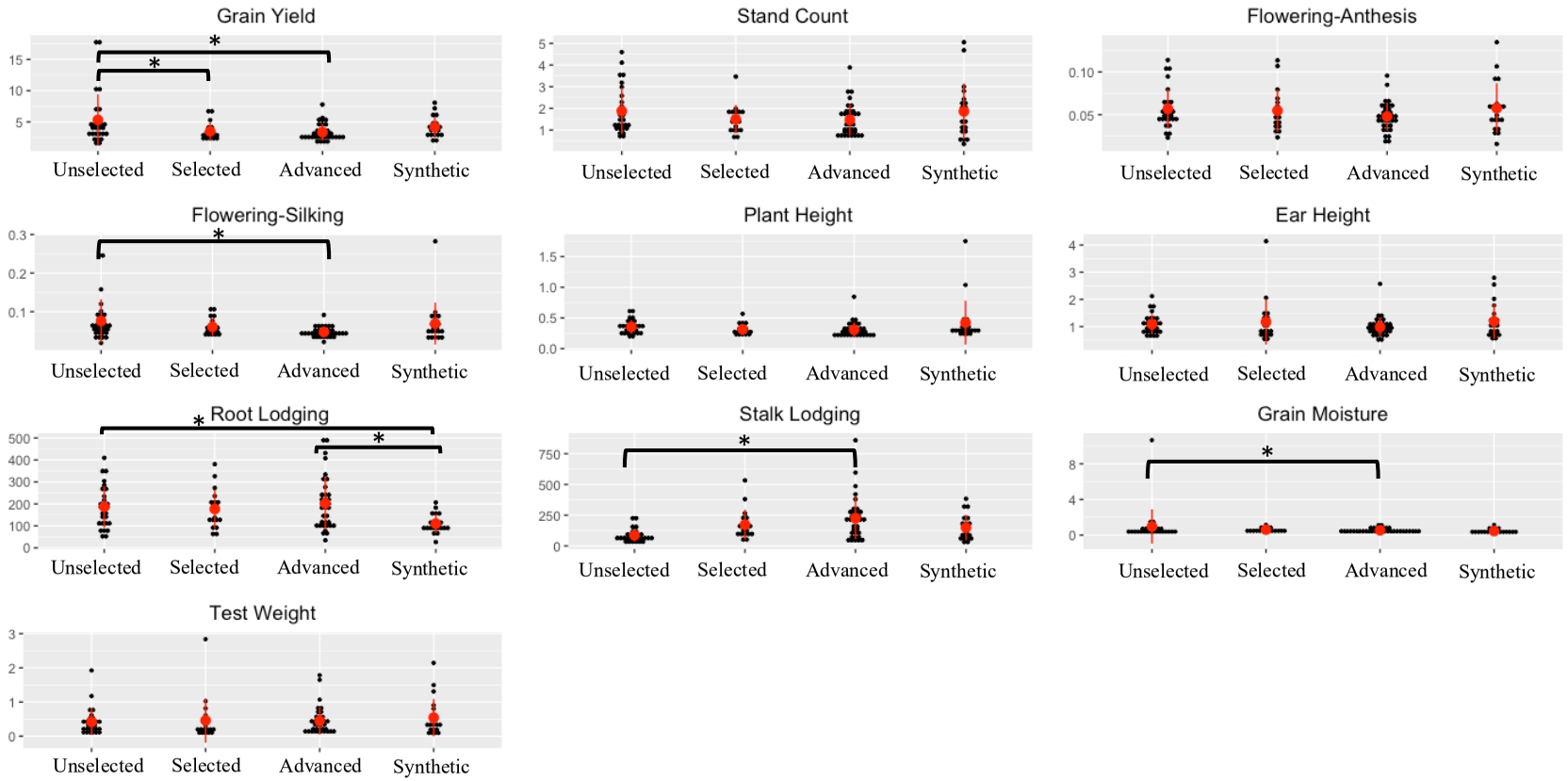


Figure 2.2. Dot plots of mean squared error (MSE) extracted from the performance-based stability index for various phenotypic traits and organized by grouping: Unselected (n=29), Selected (n=16), Advanced (n=37), and Synthetic (n=20). Black dots indicate individual genotype MSE, red dots represent the mean, and the red line is the mean plus or minus the standard deviation. Tukey HSD significance (P<0.05 = *) is indicated only between significant groupings.

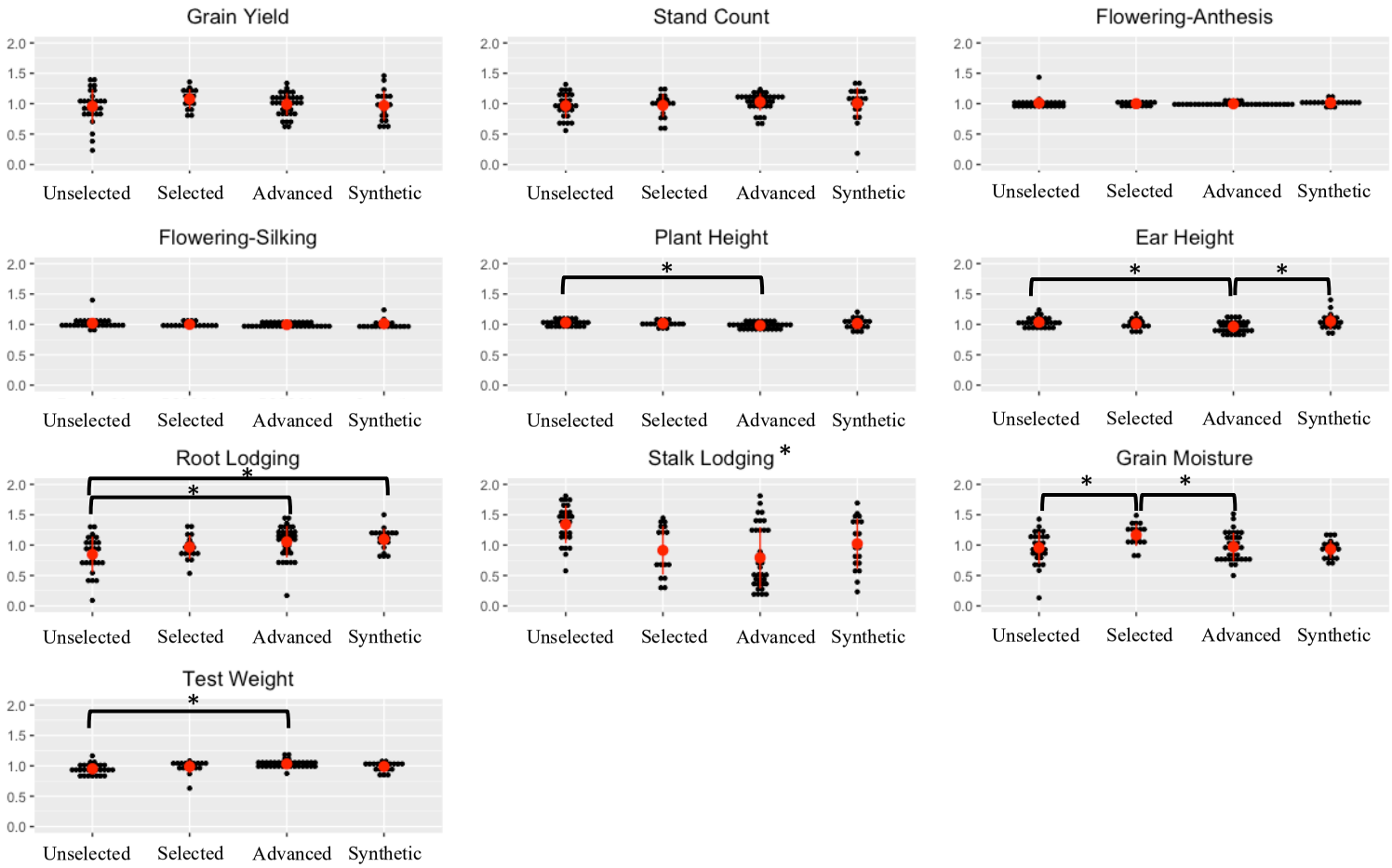


Figure 2.3. Dot plots of slope extracted from the performance-based stability index for various phenotypic traits and organized by grouping: Unselected (n=29), Selected (n=16), Advanced (n=37), and Synthetic (n=20). Black dots indicate individual genotype slope, red dots represent the mean, and the red line is the mean plus or minus the standard deviation. Tukey HSD significance ($P<0.05$ = *) is indicated only between significant groupings, except for Stalk Lodging in which all groupings were significantly different.

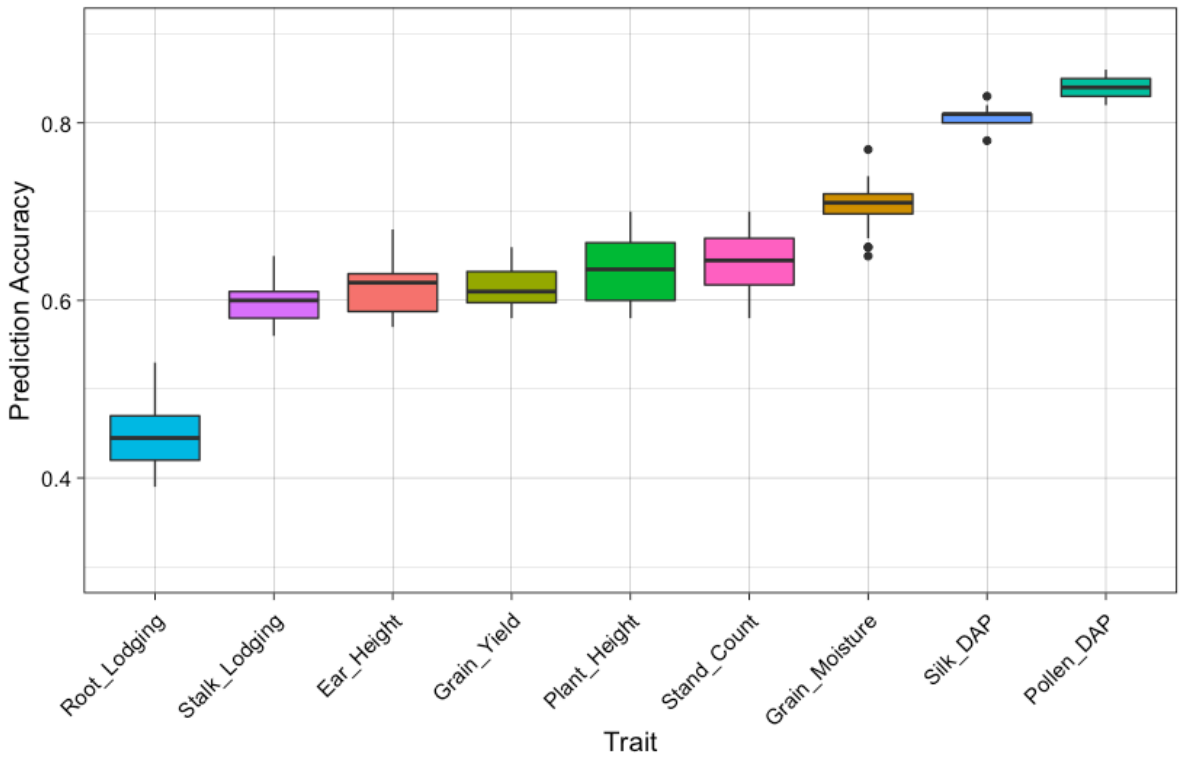


Figure 2.4. Prediction accuracies estimated across nine phenotypic traits using the Spearman correlation between the PLSR EC-predicted and actual phenotypic performance values. Traits are ordered from lowest to highest average prediction accuracy.

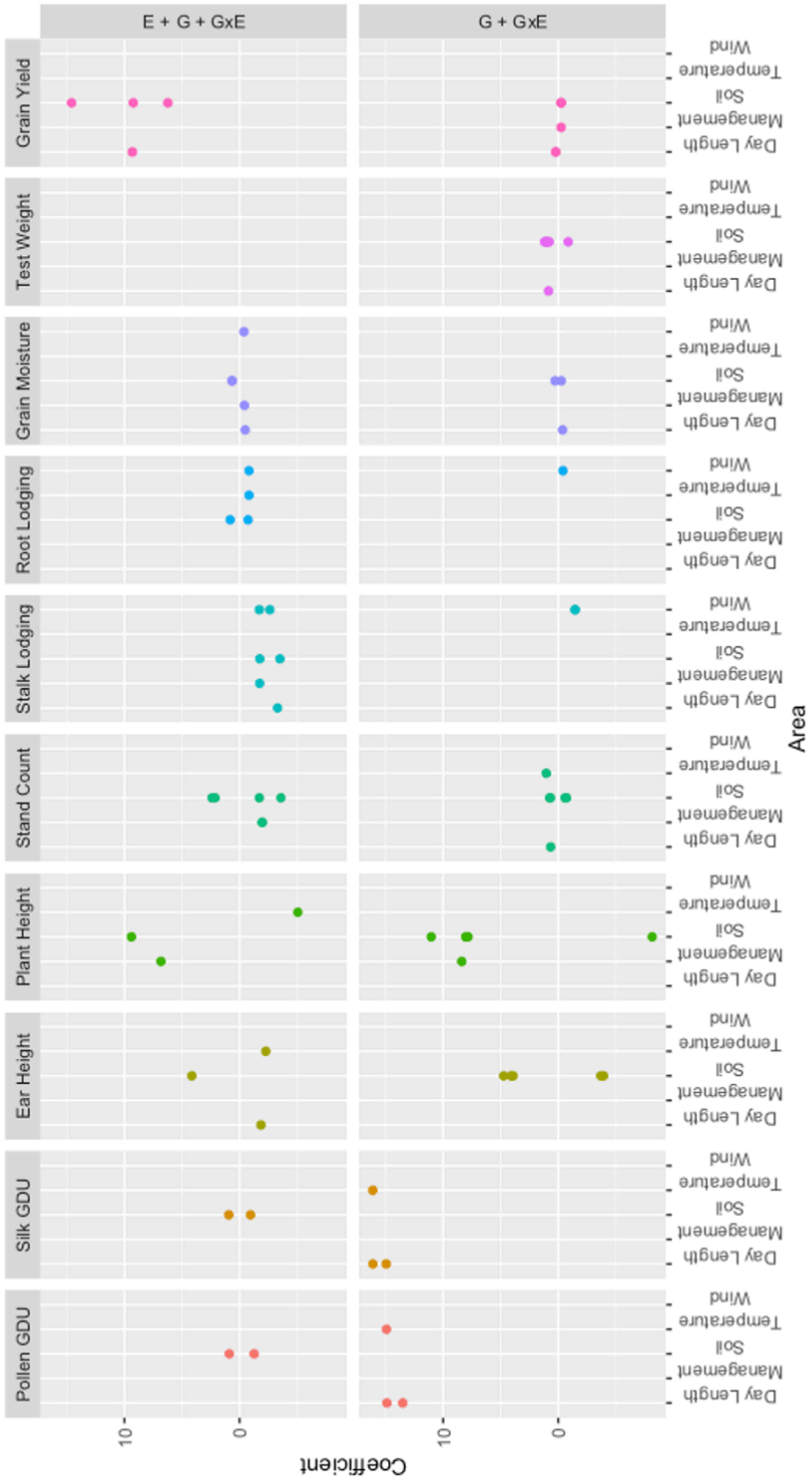


Figure 2.5. PLSR EC coefficients that are most influential across phenotypic traits for each of the environmental areas (day length, management, day length, soil, temperature, and wind). The identified ECs are grouped by the PLSR models of the raw phenotypic traits (E+G+ G×E) and the BLUEs of genotypic and G×E variance of the phenotypic traits (G+ G×E).

Supplemental Information

Supplementary Table 2.1. Table of 102 hybrid genotypes, the number of environments grown in, and assigned groupings used in Chapter Two study.

Genotype	**No. of Environments Present**	**Grouping**
2FACC/3IIH6	31	Advanced
4N506/3IIH6	31	Advanced
6F629/3IIH6	31	Advanced
A3G-3-3-1-313/3IIH6	12	Unselected
A632/3IIH6	31	Advanced
A634/3IIH6	31	Advanced
A635/3IIH6	31	Advanced
A679/3IIH6	31	Advanced
A680/3IIH6	31	Advanced
AH83/3IIH6	16	Unselected
B104/3IIH6	31	Selected
B105/3IIH6	31	Selected
B109/3IIH6	30	Advanced
B110/3IIH6	31	Selected
B111/3IIH6	31	Selected
B119/3IIH6	31	Selected
B14/3IIH6	30	Selected
B2/3IIH6	15	Unselected
B73/3IIH6	31	Selected
B84/3IIH6	31	Selected
BSSSC0_001/3IIH6	31	Unselected
BSSSC0_002/3IIH6	31	Unselected
BSSSC0_003/3IIH6	31	Unselected
BSSSC0_005/3IIH6	30	Unselected
BSSSC0_008/3IIH6	31	Unselected
BSSSC0_009/3IIH6	29	Unselected
BSSSC0_012/3IIH6	31	Unselected
BSSSC0_015/3IIH6	30	Unselected
BSSSC0_017/3IIH6	31	Unselected
BSSSC0_020/3IIH6	31	Unselected
BSSSC0_024/3IIH6	31	Unselected
BSSSC0_026/3IIH6	31	Unselected
BSSSC0_029/3IIH6	30	Unselected

BSSSC0_033/3IIH6	15	Unselected
BSSSC0_038/3IIH6	31	Unselected
BSSSC0_048/3IIH6	23	Unselected
BSSSC0_054/3IIH6	31	Unselected
BSSSC0_057/3IIH6	23	Unselected
BSSSC0_060/3IIH6	31	Unselected
CG102/3IIH6	31	Selected
CG106/3IIH6	26	Advanced
CI540/3IIH6	31	Unselected
CO258/3IIH6	31	Selected
78010/3IIH6	31	Advanced
DKFBHJ/3IIH6	31	Advanced
FBLA/3IIH6	31	Advanced
FE/3IIH6	17	Unselected
FR19/3IIH6	31	Advanced
H122W/3IIH6	30	Advanced
H84/3IIH6	31	Advanced
H91/3IIH6	30	Advanced
I159/3IIH6	31	Unselected
Ill.Hy/3IIH6	31	Unselected
LE23/3IIH6	10	Unselected
LH145/3IIH6	31	Advanced
LH149/3IIH6	31	Advanced
LH195/3IIH6	31	Advanced
LH198/3IIH6	31	Advanced
LH205/3IIH6	31	Advanced
LH209/3IIH6	31	Advanced
LH220Ht/3IIH6	30	Advanced
LH222/3IIH6	31	Advanced
MS222/3IIH6	30	Selected
N192/3IIH6	31	Advanced
N209/3IIH6	31	Selected
N217/3IIH6	31	Selected
N218/3IIH6	31	Selected
N28Ht/3IIH6	30	Selected
N501/3IIH6	31	Selected
NKH8431/3IIH6	30	Advanced
Os420/3IIH6	31	Unselected
PHB09/3IIH6	31	Advanced

PHB47/3IIH6	31	Advanced
PHBB3/3IIH6	28	Advanced
PHBW8/3IIH6	31	Advanced
PHG80/3IIH6	31	Advanced
PHP85/3IIH6	31	Advanced
PHT69/3IIH6	31	Advanced
PHV37/3IIH6	31	Advanced
R229/3IIH6	29	Advanced
SD101/3IIH6	31	Advanced
TR 9-1-1-6/3IIH6	30	Unselected
W10004_0007/3IIH6	31	Synthetic
W10004_0008/3IIH6	29	Synthetic
W10004_0011/3IIH6	29	Synthetic
W10004_0026/3IIH6	30	Synthetic
W10004_0041/3IIH6	31	Synthetic
W10004_0045/3IIH6	31	Synthetic
W10004_0062/3IIH6	31	Synthetic
W10004_0072/3IIH6	31	Synthetic
W10004_0080/3IIH6	28	Synthetic
W10004_0095/3IIH6	31	Synthetic
W10004_0119/3IIH6	31	Synthetic
W10004_0121/3IIH6	31	Synthetic
W10004_0140/3IIH6	27	Synthetic
W10004_0148/3IIH6	31	Synthetic
W10004_0178/3IIH6	15	Synthetic
W10004_0208/3IIH6	30	Synthetic
W10004_0212/3IIH6	31	Synthetic
W10004_0225/3IIH6	31	Synthetic
W10004_0258/3IIH6	30	Synthetic
W10004_0292/3IIH6	31	Synthetic

Supplementary Table 2.2. Table of environmental covariates, specifically their abbreviations, units, and descriptions used in Chapter Two study.

Area	Environmental Covariate	Abbreviation	Units	Description
Day Length	Photoperiod at Planting	Photoperiod_p	hours	Number of hours the plant is exposed to light in a 24-hour period at planting for each plot.
	Photoperiod at Anthesis	Photoperiod_an	hours	Number of hours the plant is exposed to light in a 24-hour period during flowering-anthesis for each plot.
	Photoperiod at Silking	Photoperiod_si	hours	Number of hours the plant is exposed to light in a 24-hour period during flowering-silking for each plot.
	Photothermal at Anthesis	Photothermal_an	PTU	Product of photoperiod and temperature (in GDUs) at flowering-anthesis for each plot.
	Photothermal at Silking	Photothermal_si	PTU	Product of photoperiod and temperature (in GDUs) at flowering-silking for each plot.
Management	Total Irrigation	Times_Irrigated_Total	number of times	Total number of times irrigated from planting to harvest. The specific irrigation quantity varied across locations, so the total number of irrigation application events was used to represent this EC.
	Pre-Silking Irrigation	Times_Irrigated_PreSilk	number of times	Total number of times irrigated from planting to flowering- silking. The specific irrigation quantity varied across locations, so the total number of irrigation application events before the first flowering-silking plot was used to represent this EC.

	Pre-Planting Herbicide Application	PrePlant_Herb_Total.oz.A0	L/ha	Total liters per hectare of herbicide applied before planting. The specific pre-planting herbicide used (i.e., Callisto, Princep, Lexar, etc.) varied across locations. For this reason, all applications labeled as pre-plant herbicide were combined to represent this EC.
	Post-Planting Herbicide Application	PostPlant_Herb_Total.oz.A0	L/ha	Total liters per hectare of herbicide applied after planting. The specific post-planting herbicide used (i.e., Atrazine, Callisto, Permit, etc.) used varied across locations. For this reason, all applications labeled as post-plant herbicide were combined to represent this EC.
	Insecticide Application	Insecticide_Total.oz.A	L/ha	Total liters per hectare of insecticide applied from planting to harvest. The specific insecticide used (i.e., Force 3G, Sniper) varied across locations. For this reason, all applications labeled as insecticide were combined to represent this EC.
	Fertilizer Application	Fertilizer_Total.oz.A	L/ha	Total liters per hectare of fertilizer applied from planting to harvest. The specific fertilizer used (i.e., Nitrogen, potash, urea, N-P-K mixture, etc.) varied across locations. For this reason, all applications labeled as fertilizer were combined to represent this EC.
Precipitation	Total Rainfall	TotalRainfall_s	mm	Total rainfall from planting to harvest in millimeters.
	Total Rainfall to Anthesis	TotalRainfall_an	mm	Total rainfall from planting to flowering-anthesis in millimeters.
	Total Rainfall to Silking	TotalRainfall_si	mm	Total rainfall from planting to flowering-silking in millimeters.

Soil	Soil Sulfate M-3	S_ppm	ppm	Available Sulfur in parts per million.
	Soil Boron	B_ppm	ppm	Available Boron in parts per million.
	Soil Sodium NH4OAc	Na_ppm	ppm	Available Sodium in parts per million.
	Sodium Saturation	%Na_sat	%	Percent saturation of sodium.
	Soil Phosphorous M-3	P_ppm	ppm	Available Phosphorous in parts per million.
	Soil Nitrate-N KCl	N_ppm	ppm	Available Nitrogen in parts per million.
	Soil Magnesium NH4OAc	Mg_ppm	ppm	Available Magnesium in parts per million.
	Magnesium Saturation	%Mg_sat	%	Percent saturation of magnesium.
	Soil Potassium NH4OAc	K_ppm	ppm	Available Potassium in parts per million.
	Potassium Saturation	%K_sat	%	Percent saturation of potassium.
	Soil Calcium NH4OAc	Ca_ppm	ppm	Available Calcium in parts per million.
	Calcium Saturation	%Ca_sat	%	Percent saturation of calcium.
	Soil Organic Matter	OM_LOI	%	Percentage of organic matter in soil sample measured by loss on ignition.
	Cation Exchange Capacity	CEC	me/100g	Capacity of a soil to hold exchangeable cations.
	Hydrogen Saturation	%H_sat	%	Percent saturation of Hydrogen.
	Soil pH	WDRF_pH	-	Woodruff method for measuring total soil acidity.
	Soil pH	X1.1_soilpH	-	Soil: water (1:1) solution for measuring total soil acidity.
	Soluble salts 1:1	X1.1_salts	mmho/cm	Soluble salts concentration in soil sample.
	Soil Texture	soil_text	-	Category of soil texture based on the percentage of sand, silt, and clay.
	Soil Clay	%clay	%	Percentage of clay composition in soil sample.
	Soil Sand	%sand	%	Percentage of sand composition in soil sample.

	Soil Silt	%silt	%	Percentage of silt composition in soil sample.
Temperature	Minimum Temperature at Anthesis	MinTemp_an	°C	Minimum air temperature at flowering-anthesis in Celsius.
	Maximum Temperature at Anthesis	MaxTemp_an	°C	Maximum air temperature at flowering-anthesis in Celsius.
	Minimum Temperature at Silking	MinTemp_si	°C	Minimum air temperature at flowering-silking in Celsius.
	Maximum Temperature at Silking	MaxTemp_si	°C	Maximum air temperature at flowering-silking in Celsius.
	Growing Degree Units at Anthesis	GDU_Anthesis	GDU	Growing degree units at flowering-anthesis.
	Minimum Temperature at Planting	MinTemp_p	°C	Minimum air temperature at planting in Celsius.
	Average Temperature at Planting	AvTemp_p	°C	Average air temperature at planting in Celsius.
	Maximum Temperature at Planting	MaxTemp_p	°C	Maximum air temperature at planting in Celsius.
	Minimum Relative Humidity at Planting	MinRH_p	%	Minimum air relative humidity at planting as a percentage.
	Average Relative Humidity at Planting	AvRH_p	%	Average air relative humidity at planting as a percentage.
	Maximum Relative Humidity at Planting	MaxRH_p	%	Maximum air relative humidity at planting as a percentage.
	Minimum Dew Point at Planting	MinDew_p	°C	Minimum air dew point at planting in Celsius.
	Average Dew Point at Planting	AvDew_p	°C	Average air dew point at planting in Celsius.
	Maximum Dew Point at Planting	MaxDew_p	°C	Maximum air dew point at planting in Celsius.

Wind	Minimum Wind Speed at Planting	MinWSpeed_p	m/s	Minimum wind speed in meters per second at planting
	Average Wind Speed at Planting	AvWSpeed_p	m/s	Average wind speed in meters per second at planting.
	Maximum Wind Speed at Planting	MaxWSpeed_p	m/s	Maximum wind speed in meters per second at planting.
	Minimum Wind Gust at Planting	MinWGust_p	m/s	Minimum wind speed over 30-minute period in meters per second at planting.
	Average Wind Gust at Planting	AvWGust_p	m/s	Average wind speed over 30-minute period in meters per second at planting.
	Maximum Wind Gust at Planting	MaxWGust_p	m/s	Maximum wind speed over 30-minute period in meters per second at planting.

Supplementary Table 2.3. The 145 ECs identified by Partial Least Squares Regression (PLSR) on each trait and BSSS grouping used in Chapter Two study. The EC, BSSS grouping, PLSR model coefficient corresponding to the EC, trait, and Area of climatic data that the EC belongs to are listed.

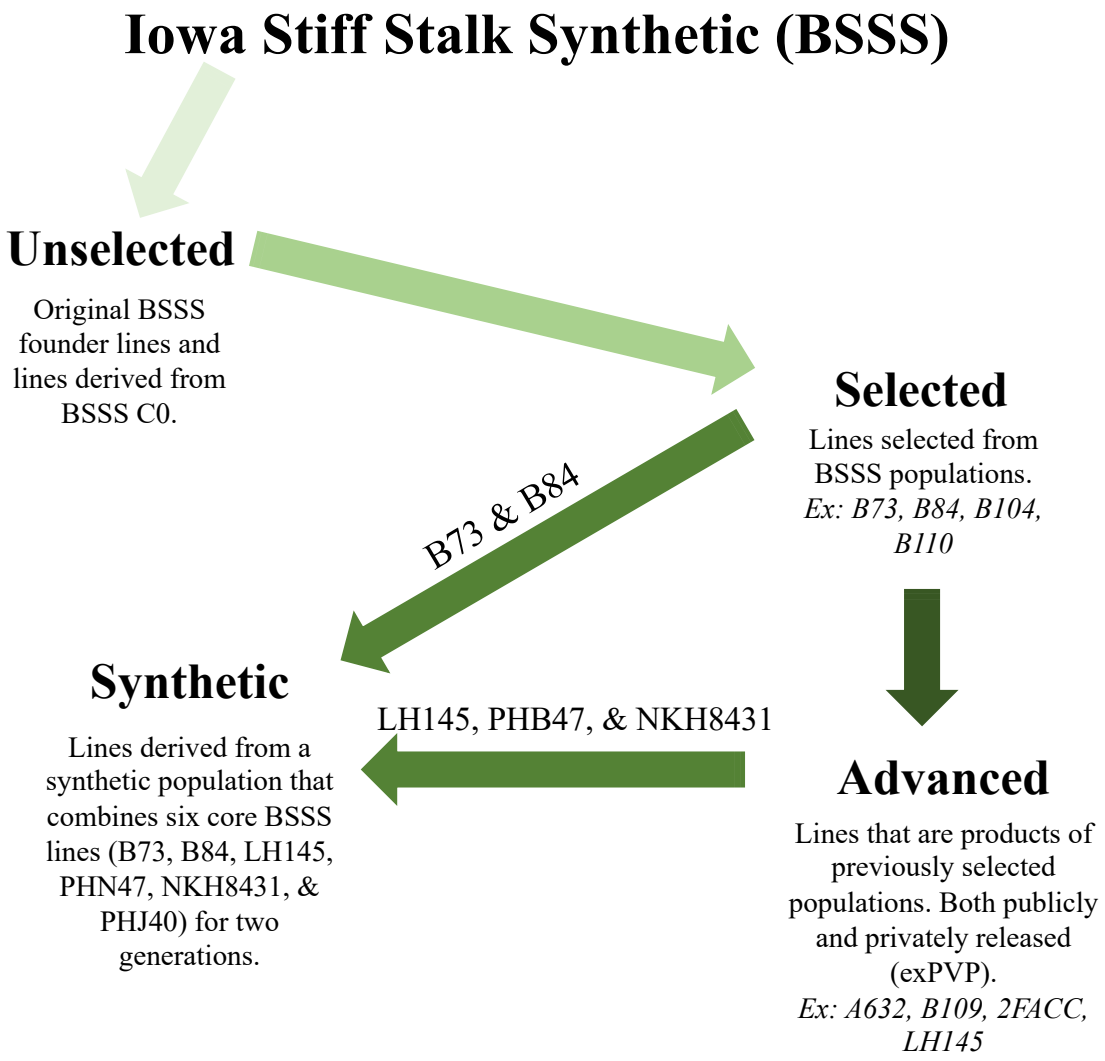
EC	Grouping	Coefficient	Trait	Area
AvDew_p	Synthetic	-0.181292	GrainMoisture	Temperature
AvRH_p	Synthetic	-7.2926817	GrainYield	Temperature
AvRH_p	Advanced	-0.3648044	SilkGDU	Temperature
AvRH_p	Unselected	0.18866895	GrainMoisture	Temperature
AvRH_p	Synthetic	0.94455961	RootLodging	Temperature
AvTemp_p	Selected	-1.205242	RootLodging	Temperature
AvWGust_p	Unselected	-1.8729701	StalkLodging	Wind
AvWGust_p	Selected	-1.7235656	StalkLodging	Wind
AvWGust_p	Advanced	-1.6966771	StalkLodging	Wind
AvWGust_p	Synthetic	-1.5129477	StalkLodging	Wind
AvWGust_p	Selected	-0.9382421	RootLodging	Wind
AvWGust_p	Synthetic	-0.2461561	GrainMoisture	Wind
AvWGust_p	Advanced	2.3178215	GrainYield	Wind
AvWSpeed_p	Advanced	-0.3060488	GrainYield	Wind
AvWSpeed_p	Synthetic	0.32032305	StandCount	Wind
Ca_ppm	Advanced	0.48692127	RootLodging	Soil
Ca_ppm	Synthetic	0.65203707	RootLodging	Soil
Fertilizer_Total.oz.A.	Synthetic	-5.8497499	GrainYield	Management
GDU_Anthesis	Selected	-0.9638094	StalkLodging	Temperature
GDU_Anthesis	Advanced	-0.7041825	StalkLodging	Temperature
GDU_Anthesis	Unselected	-0.5576445	StandCount	Temperature
GDU_Anthesis	Advanced	-0.2978409	PollenGDU	Temperature
GDU_Anthesis	Unselected	0.42412364	RootLodging	Temperature
GDU_Anthesis	Advanced	0.44441752	RootLodging	Temperature
GDU_Anthesis	Synthetic	1.20152117	PlantHeight	Temperature
Insecticide_Total.oz.A.	Advanced	-3.0876687	GrainYield	Management
Insecticide_Total.oz.A.	Selected	-2.334918	StandCount	Management
Insecticide_Total.oz.A.	Unselected	0.29022209	GrainMoisture	Management
Insecticide_Total.oz.A.	Advanced	0.3547393	GrainMoisture	Management
K_ppm	Synthetic	13.4128318	GrainYield	Soil
K_ppm	Selected	13.9391445	GrainYield	Soil
K_ppm	Unselected	15.2624986	GrainYield	Soil
Mg_ppm	Synthetic	0.19991843	GrainMoisture	Soil
Mg_ppm	Advanced	0.32523659	SilkGDU	Soil

Mg_ppm	Advanced	0.7323248	RootLodging	Soil
Mg_ppm	Selected	0.9746229	RootLodging	Soil
Mg_ppm	Synthetic	1.00482385	RootLodging	Soil
Mg_ppm	Synthetic	1.95225882	StalkLodging	Soil
MinTemp_an	Unselected	0.14883223	GrainMoisture	Temperature
MinTemp_an	Unselected	6.07567331	EarHeight	Temperature
MinTemp_an	Unselected	10.7298892	PlantHeight	Temperature
MinWGust_p	Unselected	-2.796657	StalkLodging	Wind
MinWGust_p	Advanced	-2.5259595	StalkLodging	Wind
MinWGust_p	Selected	-2.0263143	StalkLodging	Wind
MinWGust_p	Synthetic	-1.5055556	StalkLodging	Wind
MinWGust_p	Selected	-0.4526538	GrainMoisture	Wind
MinWSpeed_p	Unselected	-1.5148347	StalkLodging	Wind
N_ppm	Selected	-9.0819695	PlantHeight	Soil
N_ppm	Advanced	-7.8399542	PlantHeight	Soil
N_ppm	Advanced	-6.6650168	EarHeight	Soil
N_ppm	Selected	-6.3112608	EarHeight	Soil
N_ppm	Synthetic	-6.2213976	EarHeight	Soil
N_ppm	Unselected	-5.0198399	EarHeight	Soil
N_ppm	Synthetic	-1.4002364	StalkLodging	Soil
N_ppm	Synthetic	1.96181415	StandCount	Soil
N_ppm	Advanced	2.5123393	StandCount	Soil
N_ppm	Selected	2.57380777	StandCount	Soil
N_ppm	Unselected	2.72453465	StandCount	Soil
N_ppm	Unselected	7.34831638	GrainYield	Soil
OM_LOI	Unselected	-1.5587327	StalkLodging	Soil
OM_LOI	Selected	-1.5455998	StalkLodging	Soil
OM_LOI	Advanced	-0.3617332	GrainMoisture	Soil
OM_LOI	Unselected	0.92689489	PollenGDU	Soil
OM_LOI	Synthetic	1.00669297	SilkGDU	Soil
OM_LOI	Synthetic	1.06049175	PollenGDU	Soil
OM_LOI	Synthetic	1.70153528	StandCount	Soil
OM_LOI	Selected	1.86971381	StandCount	Soil
OM_LOI	Advanced	2.11241283	StandCount	Soil
OM_LOI	Unselected	8.787724	GrainYield	Soil
OM_LOI	Advanced	9.12232959	GrainYield	Soil
OM_LOI	Selected	9.26751739	GrainYield	Soil
P_ppm	Unselected	-4.8972055	StalkLodging	Soil
P_ppm	Advanced	-3.4338574	StalkLodging	Soil

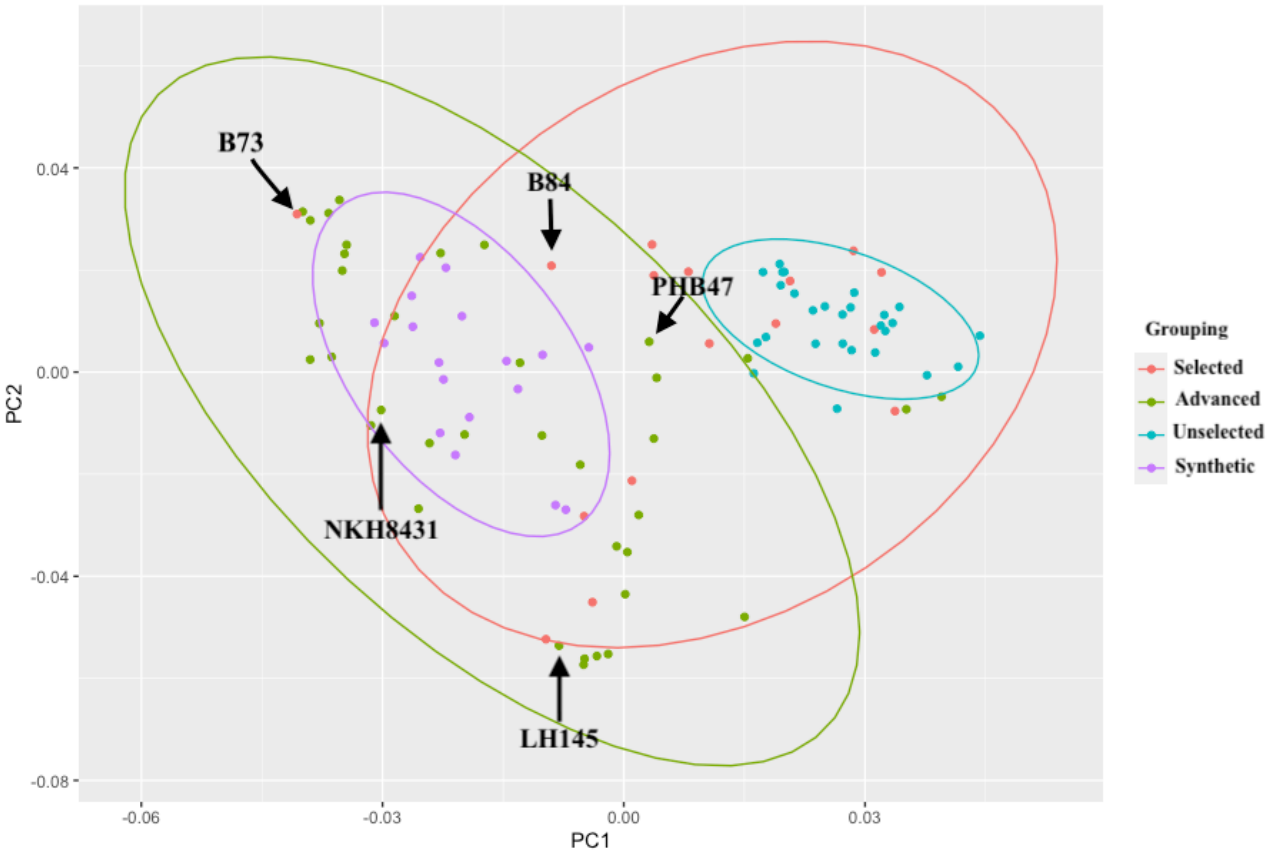
P_ppm	Advanced	-3.3981006	StandCount	Soil
P_ppm	Unselected	-3.3137636	StandCount	Soil
P_ppm	Selected	-3.309419	StandCount	Soil
P_ppm	Unselected	0.70062946	SilkGDU	Soil
P_ppm	Synthetic	0.98282513	PollenGDU	Soil
P_ppm	Selected	1.08665144	PollenGDU	Soil
P_ppm	Synthetic	4.29682965	EarHeight	Soil
P_ppm	Advanced	4.8324268	EarHeight	Soil
P_ppm	Selected	5.78133332	EarHeight	Soil
P_ppm	Advanced	6.19807794	GrainYield	Soil
P_ppm	Advanced	8.99905861	PlantHeight	Soil
P_ppm	Selected	9.31635929	GrainYield	Soil
P_ppm	Unselected	11.3924476	GrainYield	Soil
P_ppm	Selected	11.6505469	PlantHeight	Soil
P_ppm	Unselected	12.0937544	PlantHeight	Soil
Photoperiod_Anthesis	Advanced	-3.3362744	StalkLodging	Day Length
Photoperiod_Anthesis	Selected	-3.120489	StalkLodging	Day Length
Photoperiod_Anthesis	Synthetic	-1.6796905	StalkLodging	Day Length
Photoperiod_Anthesis	Synthetic	-1.3263156	PlantHeight	Day Length
Photoperiod_Anthesis	Unselected	-1.1183748	RootLodging	Day Length
Photoperiod_Anthesis	Advanced	-0.9765424	RootLodging	Day Length
Photoperiod_Anthesis	Selected	-0.5955229	GrainMoisture	Day Length
Photoperiod_Anthesis	Advanced	-0.475728	GrainMoisture	Day Length
Photoperiod_Anthesis	Selected	1.01043952	SilkGDU	Day Length
Photoperiod_Anthesis	Synthetic	1.05945268	SilkGDU	Day Length
Photoperiod_Anthesis	Advanced	5.58571508	GrainYield	Day Length
Photoperiod_Anthesis	Selected	6.87251663	GrainYield	Day Length
Photoperiod_Anthesis	Synthetic	7.66081939	GrainYield	Day Length
Photothermal_Anthesis	Selected	-1.2986349	StalkLodging	Day Length
Photothermal_Anthesis	Advanced	-1.0448355	StalkLodging	Day Length
Photothermal_Anthesis	Unselected	-0.6656757	StandCount	Day Length
Photothermal_Anthesis	Advanced	-0.2420621	PollenGDU	Day Length
Photothermal_Anthesis	Unselected	0.32185258	RootLodging	Day Length
Photothermal_Anthesis	Synthetic	1.06699515	PlantHeight	Day Length
PostPlant_Herb_Total.oz.A.	Selected	0.316365	StandCount	Management
PostPlant_Herb_Total.oz.A.	Synthetic	0.56035697	StandCount	Management
PostPlant_Herb_Total.oz.A.	Unselected	0.83001943	StandCount	Management
PrePlant_Herb_Total.oz.A0	Unselected	-2.3301803	StandCount	Management
PrePlant_Herb_Total.oz.A0	Selected	-1.5896253	StandCount	Management

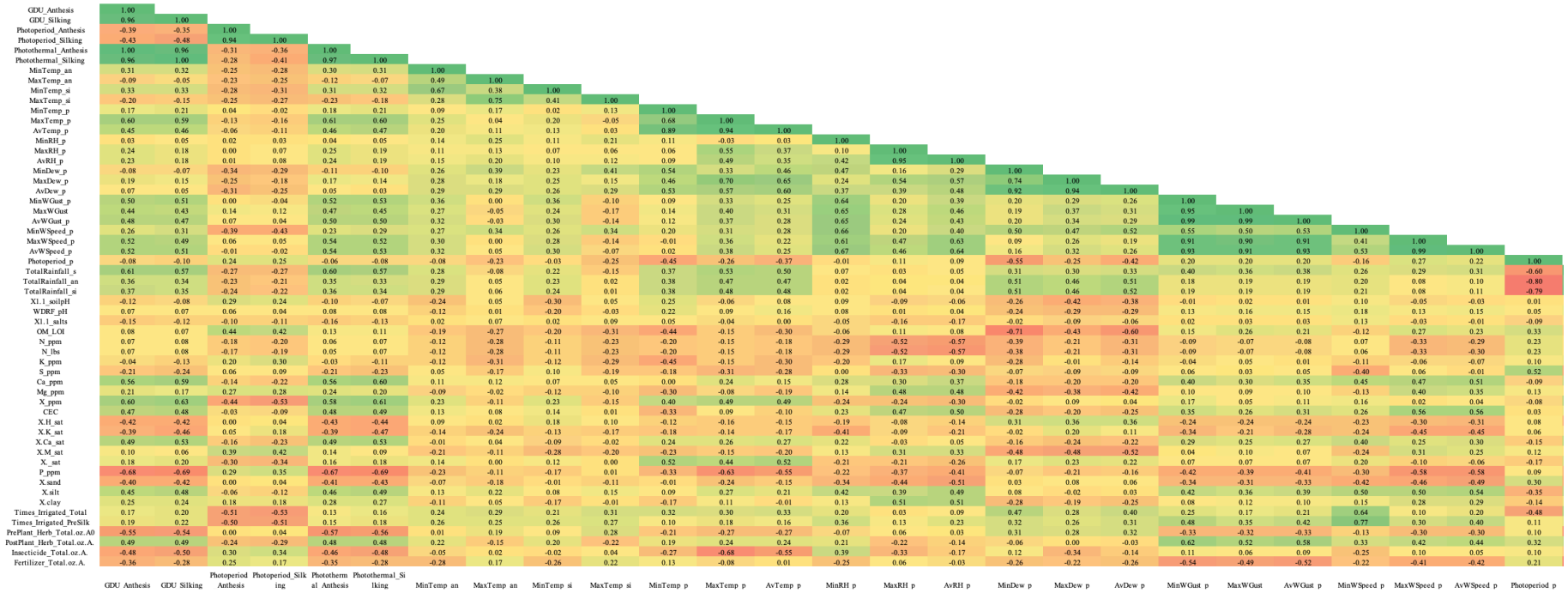
PrePlant_Herb_Total.oz.A0	Selected	-0.5015745	GrainMoisture	Management
PrePlant_Herb_Total.oz.A0	Advanced	4.26670178	EarHeight	Management
S_ppm	Synthetic	1.2242313	StandCount	Soil
Times_Irrigated_PreSilk	Unselected	-2.229755	StalkLodging	Management
Times_Irrigated_PreSilk	Advanced	-2.0230932	StandCount	Management
TotalRainfall_s	Unselected	-0.307529	RootLodging	Precipitation
TotalRainfall_s	Synthetic	0.21245965	StalkLodging	Precipitation
WDRF_pH	Unselected	-10.815164	PlantHeight	Soil
WDRF_pH	Selected	-10.015734	PlantHeight	Soil
WDRF_pH	Advanced	-8.9955081	PlantHeight	Soil
WDRF_pH	Unselected	-5.4257617	EarHeight	Soil
WDRF_pH	Selected	-4.7998481	EarHeight	Soil
WDRF_pH	Synthetic	-4.7472675	EarHeight	Soil
WDRF_pH	Synthetic	0.48159611	GrainMoisture	Soil
WDRF_pH	Advanced	0.54633198	GrainMoisture	Soil
WDRF_pH	Unselected	0.80059706	GrainMoisture	Soil
WDRF_pH	Selected	0.97597156	GrainMoisture	Soil
WDRF_pH	Advanced	1.4917584	GrainYield	Soil
X_ppm	Selected	-1.3716021	PollenGDU	Soil
X_ppm	Unselected	-1.0929769	PollenGDU	Soil
X_ppm	Selected	-1.0183715	SilkGDU	Soil
X_ppm	Unselected	-0.9195994	SilkGDU	Soil
X_ppm	Advanced	-0.5238682	StandCount	Soil
X.clay	Synthetic	0.3761729	StandCount	Soil
X1.1_salts	Synthetic	-9.2406476	GrainYield	Soil
X1.1_salts	Synthetic	-2.1992491	StalkLodging	Soil
X1.1_salts	Synthetic	-0.72618	RootLodging	Soil
X1.1_salts	Advanced	0.64690113	StandCount	Soil
X1.1_soilpH	Selected	-0.940732	RootLodging	Soil
X1.1_soilpH	Synthetic	0.35588615	GrainMoisture	Soil
X1.1_soilpH	Unselected	0.55571176	GrainMoisture	Soil
X1.1_soilpH	Selected	0.56324432	GrainMoisture	Soil
X1.1_soilpH	Advanced	0.64732319	GrainMoisture	Soil

Supplementary Figure 2.1. Schematic of the four groupings used in the Chapter One study: Unselected, Selected, Advanced, and Synthetic. Brief descriptions of each grouping as well as examples of some inbreds.

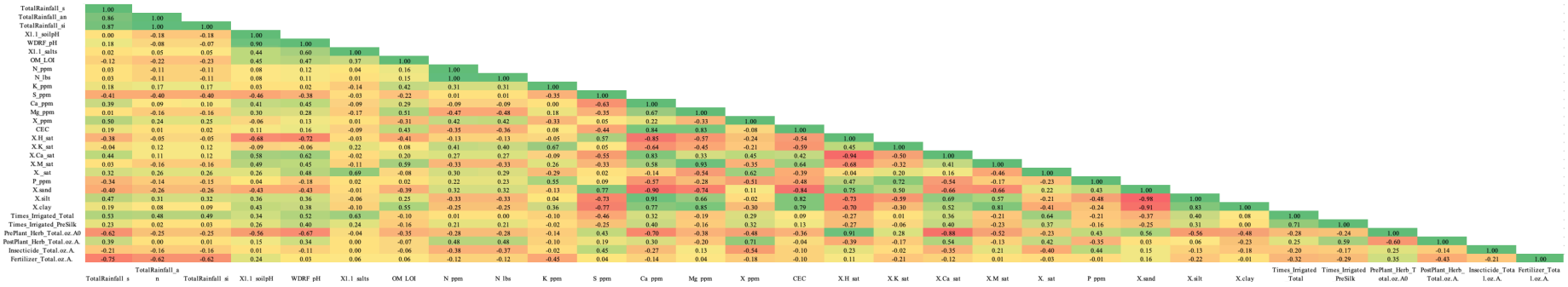


Supplementary Figure 2.2. Multi-dimensional scaling (MDS) plot of genetic relationships between the 98 inbreds, each represented by a point and colored by grouping. The five inbreds indicated with arrows are the parents of the Synthetic population.





Supplementary Figure 2.3. Pearson correlation across 56 environmental covariates (ECs) used in the partial least squares regression (PLSR).


TotalRainfall_s
TotalRainfall_an
TotalRainfall_si
Xl.1_soilpH
WDRF_pH
Xl.1_salts
OM_LOI
N_ppm
N_lbs
K_ppm
S_ppm
Ca_ppm
Mg_ppm
X_ppm
CEC
X.H_sat
X.K_sat
X.Ca_sat
X.M_sat
X._sat
P_ppm
X.sand
X.silt
X.clay
Times_Irrigated_Total
Times_Irrigated_PreSilk
PrePlant_Herb_Total.oz.A0
PostPlant_Herb_Total.oz.A.
Insecticide_Total.oz.A.
Fertilizer_Total.oz.A.

CHAPTER 3. THE EFFECTS OF INCREASED PLANTING DENSITY ON YIELD COMPONENT TRAITS IN MAIZE

Abbreviations list: EID: Ever-increasing density; NAM: Nested association mapping; SNP: Single nucleotide polymorphism; QTL: quantitative trait locus; BLUE: Best Linear Unbiased Estimator; LOD: Logarithm of the odds

ABSTRACT

Grain yield in maize has been steadily increasing since the 1930's in North America and increased number of plants per hectare, or planting density, has been one of the most influential aspects associated with these increases. Dissecting specific yield components, including ear, cob, and kernel characteristics, could further elucidate which component traits have been primarily modified while breeding for higher productivity in the context of increasing planting density. The goals of this work are to understand the phenotypic variation of yield component traits for three biparental recombinant inbred lines families with inbred PHW65 as the common parent, which reflect older, minimally selected and commercially relevant germplasm, respectively, across a range of planting densities in connection with production-level yield data from the Genomes to Field (G2F) Initiative; and to genetically dissect the architecture of kernel yield component responses across the varying planting densities. To assess the effect of different planting densities in plant development and components of yield, an "Ever-Increasing Density" (EID) plot format was used. The EID scheme deploys altered plant spacing within each plot to systematically simulate planting densities from 17,205 to 258,148 plants ha^{-1}. The evaluation was conducted at a Madison, Wisconsin location in 2018 and 2019. Ear, cob, and kernel traits (yield component traits) were evaluated from the uppermost ear on one representative plant per

plot using an image-based phenotyping tool. The components of ear height, ear length and weight, cob length and weight, cob width, kernel depth, width and area, kernel row number, and kernel weight were significantly influenced by the density treatment and they varied across families; however, ear width did not significantly change across treatments or families. The less selected family (PHW65 x MoG) produced the heaviest cobs and kernels, and largest kernel size, while the commercially relevant and highly selected family (PHW65 x PHN11) produced the lightest cobs and smallest kernels. The less selected family consistently had the widest range of values for numerous traits (ear height, maximum ear width, total cob weight, and kernel depth), while the highly selected family displayed less variability and had the smallest range of values for kernel weight and kernel width. Across all families, increased planting density resulted in decreases in ear, cob, and kernel weight, as well as kernel size (width, depth, and area). Joint linkage mapping identified seven genomic regions on chromosomes 6, 7, 8 and 9 associated with kernel size traits and ear height at the three highest planting densities of 51,604, 85,925, and 258,385 plants ha^{-1}. When connecting EID data with production-level G2F data at comparable planting densities, the yield component traits that were most positively correlated with grain yield were number of ears, while kernel size was negatively correlated. The significant correlation between ear height in the production-level environments with ear height at two of the EID treatments, the same phenotypic trait with an expected correlation to one another, supports the use of the EID design to evaluate varying planting density effects. Our findings demonstrate the utility of alternative planting density schemes to understand the effects of variable planting density on yield component traits and genetically dissect grain yield.

3.1 Introduction

While various factors contribute to grain yield, such as management and weather, increasing the number of plants per hectare, or planting density, has been one of the most influential aspects contributing to higher grain yield in maize (Duvick, 2005). Within North America, maize planting densities have increased substantially from 30,000 plants ha^{-1} in the 1930's to over 81,000 plants ha^{-1} today (Duvick, 2005; Butzen and Burnison, 2014; Mansfield and Mumm, 2014). Studies have found that older hybrids (released in 1960) perform best at low planting densities (grain yield of 2,500 kg ha^{-1} at a density of 40,000 plants ha^{-1}), compared to hybrids released after 1990 that perform the best at higher planting densities (grain yield of 7,500 kg ha^{-1} at a density of 65,000 plants ha^{-1}; Duvick, 2005; Smith et al., 2014). This suggests that US maize breeders have selected for plants with greater stress resistances imposed by higher planting densities.

Grain yield, as one of the most important traits considered by maize breeders, is determined by several yield components and in some instances evaluated as such due to the higher heritability of the components compared to overall yield (Austin and Lee, 1996; Gupta et al, 2006; Messmer et al., 2009). Generally, grain yield in maize can be broken down into the ear density, or ears per area, kernel number per ear (a product of row number and kernels per row), and kernel weight (Lauer, 2008). Duvick (2005) evaluated which yield components and plant characteristics that have changed from the 1930's through the early 2000's, coinciding with the increase in hybrid grain yield in maize during that time frame. The total number of ears per 100 plants, leaf angle score, and staygreen, or delayed leaf senescence, increased in newer hybrids. While the number of leaves per plant did not change, there were changes in tassel size. The tassel branch number and overall weight decreased consistently when evaluating commercial hybrids that were introduced from 1934 to 2001 (Meghji et al., 1984), with a reduction of 2.5 fewer

branches per decade. Commercial hybrids that have been selected for high yield performance in these planting density conditions have also greater correlation with the following morphological traits: more upright leaves, shorter plant height, and smaller tassels (Duvick, 2005; Gage et al., 2018). Studies have often focused on kernel number and weight, but there is a tradeoff between the two traits, increasing the difficulty in genetic improvement for yield (Henery and Westoby, 2001; Peng et al., 2011). Notably, it was observed that newer hybrids exhibited an increase in kernel weight over the years rather than an increase in the kernel number per plant, highlighting the idea that kernel weight might be a greater source of potential continued grain yield improvements (Crosbie, 1982; Russell, 1985; Duvick, 1997; Edmeades et al., 2003; Barker et al., 2005; Duvick, 2005). There are numerous reasons to focus on kernel size-related traits. Kernel size is highly correlated to kernel weight, which is an important component in calculating grain yield, as indicated previously (Duvick, 2005; Messmer et al., 2009), and unlike grain yield, the genetic architecture associated with kernel traits are more stable across diverse environments (Peng et al., 2011). Kernel size can also be broken down further into different shape traits, such as kernel length, width, and thickness (Li et al., 2009). Focusing on kernel-related traits to identify major quantitative trait loci (QTL) would further the understanding and potential breeding areas for improved maize yield, as suggested by Chen et al. (2016) in their investigation of the genetic basis of natural variation in maize seed size and yield related components.

Planting density can be used as a discriminator of multiple stress response. An increased number of plants per area generates increased competition for water, nutrients, and light (Troyer, 1996) and crowding stress in plants is a persistent stress, continuously affecting plants throughout the entire crop cycle (Hütsch et al., 2018). In maize, for example, environmental stress posed during pollination or post-pollination, such as drought or extreme heat temperatures,

can result in lack of fertilization and ultimately reducing the number of kernels at the end of the season greatly increases the likelihood for kernel abortion, resulting in yield loss (Iowa State University Extension, 2017). This stress, among others, could also be a result of more plant competition that occurs as spacing between plants decreases.

To assess the effects of the changes in planting density through the years of breeding programs, an "Ever-Increasing Density" (EID) plot format can be used. This plot format allows for planting at varying plant spacing at an increasing density throughout the plot. The EID concept is that spacing between plants varies systematically within each row of a plot rather than between plots or across whole sections of plots representing a density treatment. EID also enables more genotypes to be evaluated at a variety of density treatments in the same area. By improving resource allocation (i.e., reducing the amounts of seed and field space necessary), the EID plot format reduces the confounding variable that exists between the planting density treatment and field conditions. When evaluating a single planting density in one field and comparing it to another planting density in a neighboring field, differences in the soil fertility, water drainage, planting dates, and other variables could lead to observed differences in plant performance that are not associated directly with planting density (Farnham, 2001). However, since it is not possible to separate the plant density treatment effect from the field effect, this could impact conclusions. The ability to analyze a large number of genotypes, across multiple planting densities within a given field improves the utility of specific experiments looking to assess genetic variability across genotypes, for example.

Quantitative trait locus (QTL) analyses identify statistical associations between DNA polymorphisms (for example, single nucleotide polymorphisms, or SNPs) and phenotypic variation (Doerge, 2002). The ability to detect QTL depends highly on the structure of the

population that is being used, as well as other factors. One way to evaluate phenotype and genotype associations is to use designed populations. One possible population structure includes the cross of two parental lines, commonly referred to as biparental crosses. Biparental populations aren't able to detect any genetic loci that don't vary between the two parents and that the number of recombination events is limited. Other types of populations, such as multi-parent populations (MPP), have been proposed to improve the likelihood of QTL detection (Cavanagh et al., 2008). One of the most well-known MPP reference designs are Nested Association Mapping (NAM) populations (McMullen et al., 2009). In these types of populations, mapping resolution is enhanced due to the combination of historic recombination unique to each founder parent and the more recent recombination during population development. Developed in 2009, the maize NAM population is comprised of 25 diverse maize lines that were crossed to a common reference parent, B73, and the progeny from each cross were subsequently self-pollinated for six generations to create 200 homozygous lines per family, totaling 5000 lines for the entire NAM population.

Collections of diverse genotypes can be used to associate genotypes and phenotypes using an association mapping approach (Zhu et al., 2008). Association mapping is highly dependent on the amount of linkage disequilibrium (LD), or non-random association of alleles at two or more different loci, across the genome (Flint-Garcia et al., 2003). If LD decays in a short distance, mapping resolution is expected to be high as long as there are a large number of markers (Rafalski, 2002). Alternatively, if LD extends for a longer distance, the mapping resolution will be lower but fewer markers are necessary. A potential drawback of this approach is that these methods generally use single SNPs in the analysis, or bi-allelic marker models at a given QTL position, which has a possibility of failing to reflect the allele diversity potentially

present at within diverse populations (Garin et al., 2017). Taking into account the level of genetic relatedness (kinship) and the effect of population structure are two ways to help distinguish true associations from false positives generated by the population structure (Yu et al., 2006). The R/qtl2 software environment (Broman et al., 2019) allows for the inclusion of kinship matrices, population structure, and can handle multiple founder alleles, which allows for further flexibility to deploy association mapping populations with diverse population structures.

Introduced in Chapter 1, the PHW65 NAM population was designed as a reduced version of the maize NAM, consisting of a single parental reference line, PHW65, crossed to three diverse founder, inbred lines, PHN11, Mo44, and MoG (Gage et al., 2018). The PHW65 NAM has been used for genetic mapping in various plant phenotypic and yield-component traits (Haase, 2015; Gage et al., 2018). A yield-component analysis was conducted on the PHW65 NAM population in 2013 and 2014 (Haase, 2015) with plants grown at a single density of 65,000 plants ha^{-1}. This linkage mapping analysis identified 58 QTL associated with kernel yield-components across all 10 chromosomes. Of these QTL, 22 were related to kernel width or depth measurements and 19 to kernel row number, highlighting the genomic regions associated with kernel-related yield components. To get a better understanding of how the PHW65 NAM yield component traits relate to production- level grain yield, data from the Genomes To Fields (G2F) Initiative was used. As described in Chapter 1, G2F is a public collaboration that evaluates maize hybrids across a range of environments and results in plot-level measurement. Lines from the PHW65 NAM were planted in two environments in Wisconsin and 26 additional environments across the U.S. during 2018 and the resulting, publicly available data will be compared to yield components measured from the same lines.

Using the EID plot format to simulate six planting densities and the PHW65 NAM population, the objectives of this study were to: (i) assess the phenotypic diversity for yield component traits under a wide range of planting density treatments, (ii) dissect the genetic architecture for kernel yield component traits across planting density treatments, and (iii) leverage production-level yield data through the Genomes to Field (G2F) Initiative with EID yield component traits.

3.2 Materials and Methods

Germplasm

Biparental families were generated from crosses between the common parent PHW65 and inbred lines MoG, Mo44, and PHN11. The crosses PHW65 x PHN11, PHW65 x Mo44, and MoG x PHW65 formed the three biparental populations (referred to as families moving forward) and individuals within the families (hereafter referred to as lines were derived as doubled haploid (DH) from the F_2 of the original parents. The DHs were crossed to expired plant variety protection (ex-PVP) inbred PHT69 and grown as part of the Genomes to Fields Initiative (G2F) in 2018 across 12 unique environments (Genomes to Fields, 2020). While these three families share PHW65, the other parents are genetically diverse. PHN11 is an Iodent ex-PVP that is related to PH207 (commercially relevant germplasm) and highly selected, Mo44 was derived from inbred line Mo22 (which was developed from the open-pollinated variety, Laguna) and Pioneer Mexican Synthetic 17 (Flint-Garcia et al., 2005) and has received moderate selection for productivity and MoG is derived from a population named Mastodon and has undergone little selection for grain yield.

EID Evaluation

A subset of the lines was selected and evaluated as hybrids in the EID experiment in 2018 and 2019 at the West Madison Agricultural Research Station in Verona, WI. Across both years, a total of 385 lines were evaluated (122 from MoG, 132 from Mo44, 131 from PHN11). Hybrids were planted in a modified split-plot design with the treatment of density nested within each of the individual genotypes and with two field replications. The lines were completely randomized throughout both replications. Each main plot was one row, 4.26 meters in length, and contained a single genotype. Each line was planted at six density treatments and the treatments were always

in the same order, from lowest to highest density. The spacing between each plant increased from 0.051 to 0.76 meters. Within a plot, these spacings represented planting densities of 17,209; 23,456; 36,790; 51,604; 85,925; and 258,385 plants ha^{-1}, respectively. To maximize precision and spacing between plants, a planter adapted with a graduated planting disk was used. The planting schematic and density treatments are labeled in Figure 3.1.

Within each plot, a single competitive and representative plant from each of the planting density treatments was identified and tagged once plants reached kernel physiological maturity at the end of the growing season. On the tagged plants, ear height was measured as the distance from the base of the crown root to the primary ear bearing node (cm). All ears from each of the tagged plants were then harvested and dried to approximately 10-15% kernel moisture. To assess yield-component traits, ears were imaged using an EPSON Perfection V700 flatbed scanner and processed using custom MATLAB algorithms as described in Miller et al. (2017) and executed in CyVerse. One ear per plot per plant was used and the kernel row number (number of kernel rows around the middle to lower third of the ear) were manually counted. After the number of kernel rows were manually counted, each ear per plot treatment was weighed separately on a scale. Ears from each planting density within a plot were imaged separately on the flatbed scanner. The ears were imaged at 1200 digital pixels per inch (dpi). Since seed fill can be irregular around an ear, the ears were rotated 90 degrees and imaged a second time before shelling them to separate the kernels from the cob. The measurements between the two images for each ear were averaged for the ear length (distance from the base of the ear to the tip in cm), ear width (distance across the ear in cm), and maximum ear width (median of the top 200 ear width measurements to account for imperfections and variability across the ear; in cm).

Once shelled, cobs were weighed on a scale to measure the total cob weight (g) and a 20 mL volume of kernels were weighed for the total kernel weight (g). The cobs were also imaged at a resolution of 1200 dpi to measure the distance in pixels from the base to the tip of cob (referred to as the cob length, in cm) and the distance across the cob (cob width in cm). The maximum cob width (cm) was estimated by taking the greatest 200 width measurements to correct for variability within the image. For kernel measurements, the 20 mL volume of kernels were spread out on the flatbed scanner and imaged once. A maximum likelihood estimation was used to determine the major axis through a kernel, from the tip of the kernel to the base, or kernel depth (mm). Kernel width (mm) measures the distance across the kernel, or minor axis. Finally, kernel area was the total count of pixels inside of the outlined kernel space (mm^2), averaged across all kernels in the sample. The total number of kernels in the image were counted and the weight per kernel was estimated (g). Sample images of these measurements can be found in Supplementary Figures 3.1 and 3.2.

All previously mentioned yield components were measured on the hybrids grown in 2018. Due to facility and personnel limitations surrounding the COVID-19 pandemic, only a portion of these measurements were collected (ear height, kernel row number, total kernel weight, number of kernels, average weight per kernel, kernel width, kernel depth, and kernel area) on the 2019 field season materials and are indicated as such in Table 3.1.

G2F Evaluation

For G2F in 2018, there were a total of 287 unique DHs grown in Verona, WI (WI1) and Arlington, WI (WI2) across both environments (180 from MoG, 98 from Mo44, 167 from PHN11). These hybrids were planted in a randomized complete block design at a single density

of 78,938 plants ha^{-1}. There were 10 other environments across the U.S. that ranged in planting densities of 78,938 to 106,155 plants ha^{-1}. A variety of phenotypic measurements were taken, but only grain yield (ton ha^{-1}) and ear height (cm) will be used in this study to connect back to the yield component traits to see the correlations between EID yield components and production-level data. The 12 G2F environment locations used in this study, planting densities, and average grain yields are listed in Table 3.2.

Statistical Analysis

Phenotypic Data

EID Evaluation

Phenotypic variation was evaluated across all traits for the 385 lines (Line) in each of the six planting density treatments (Density) for 2018 and 2019 as well as the three Families across each of the traits. Data cleaning included removing outliers beyond three standard deviations from the mean, which included removal of total ear, cob, and kernel weights of 5 g or less; removal of number of kernel measurements less than 20 and observations beyond four standard deviations from the mean for kernel width, depth, and area. This large range of standard deviation was chosen due to the observed variability, and sensitivity of these kernel traits, across traits in the planting density treatments. Across both years, a total of 6,730 Line by Density observations remained. Pearson correlations between all phenotypic traits were calculated.

To assess the variability of all Lines across each of the Treatments, a linear mixed model was used. While the experiment was planted as a split-plot, the Treatments were not randomized throughout the main plot, limiting the calculation of the subplot error term. For this reason, the plots were all statistically analyzed as modified randomized complete block designs. The model in Equation 1 traits accounted for the fixed effects of Year, Density, Family, Year x Density,

Year x Line nested in Family, Density x Line nested in Family, and Year x Density x Line nested in Family. The random effects were Replication nested within Year and Line nested within Family. This model was only used for the following traits that were evaluated across both years: ear height, kernel row number, total kernel weight, number of kernels, weight per kernel, kernel width, kernel depth, and kernel area.

Equation 1:

$$Y_{ijklm} = \mu + Y_i + D_j + F_k + L(F)_l + YD_{ij} + YL(F)_{il} + DL(F)_{jl} + YDL(F)_{ijl} + R(Y)_m + \varepsilon_{ijklm}$$

μ = population mean

Y_i = Year effect (i= 1-2)

D_j = Density effect (j= 1-6)

F_k = Family effect (k= 1-3)

$L(F)_l$ = main Line effect (l= 1-385) nested within Family (F; 1-3)

YD_{ij} = Year by Density interaction

$YL(F)_{il}$ = Year by Line nested in Family interaction

$DL(F)_{jl}$ = Density by Line nested in Family interaction

$YDL(F)_{ijl}$ = Year by Density by Line nested in Family interaction

$R(Y)_m$ = Replication (m= 1-2) nested in Year (Y; 1-2)

ε_{ijklm} = Error

To assess the variability of all Lines across each of the Densities for the traits only evaluated in 2018, a linear mixed model was used (Equation 2). accounted for the fixed effects of

Density, Family, and Density x Line nested in Family. The random effects were Replication nested within Year and Line nested within Family.

Equation 2:

$$Y_{jklm} = \mu + D_j + F_k + L(F)_l + DL(F)_{jl} + R_m + \varepsilon_{jklm}$$

$$\mu = \text{population mean}$$

$$D_j = \text{Density effect (j= 1-6)}$$

$$F_k = \text{Family effect (k= 1-3)}$$

$$L(F)_l = \text{Line effect (l= 1-385) nested within Family (F; 1-3)}$$

$$DL(F)_{jl} = \text{Density by Line nested in Family interaction}$$

$$R_m = \text{Replication effect (m= 1-2)}$$

$$\varepsilon_{ijklm} = \text{Error}$$

Heritability was estimated across traits on an entry-mean basis (Equation 3; Fehr, 1987) for all lines. For the traits that were only measured in 2018, this is more appropriately called repeatability.

Equation 3:

$$h^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{GE}^2}{e} + \frac{\sigma_\epsilon^2}{re}}$$

where σ_G^2 was the Line variance, σ_{GE}^2 was the Line by Year variance, σ_ϵ^2 was the error variance, for *e* number of Years and *r* number of Replications per Year.

When modelling the Line by Density interaction across traits, there were issues with model convergence failure. The majority of the genetic variance was often explained by the Family effect, which resulted in minimal variance attributable to the individual Lines. In order to understand the variability of Family across all Densities, the following model was used (Equation 4). This model for traits accounted for the fixed effects of Year, Density, Family, Year x Density, Year x Family, Density x Family, and Year x Density x Family. The random effect was Replication nested within Year. This model was used for the traits of: ear height, kernel row number, total kernel weight, number of kernels, weight per kernel, kernel width, kernel depth, and kernel area.

Equation 4:

$$Y_{ijkm} = \mu + Y_i + D_j + F_k + YD_{ij} + YF_{ik} + DF_{jk} + YDF_{ijk} + R(Y)_m + \varepsilon_{ijkm}$$

μ = population mean

Y_i = Year effect (i= 1-2)

D_j = Density effect (j= 1-6)

F_k = Family effect (k= 1-3)

YD_{ij} = Year by Density interaction

YF_{ik} = Year by Family interaction

DF_{jk} = Density by Family interaction

YDF_{ijk} = Year by Density by Family interaction

$R(Y)_m$ = Replication (m= 1-2) nested in Year (Y; 1-2)

ε_{ijkm} = Error

To evaluate the Family variance for the traits only evaluated in 2018 across all Densities, a linear mixed model was used (Equation 5). This model accounted for the fixed effects of Density, Family, Density x Family. The random effect was Replication.

Equation 5:

$$Y_{jkm} = \mu + D_j + F_k + DF_{jk} + R_m + \varepsilon_{jkm}$$

$$\mu = \text{population mean}$$

$$D_j = \text{Density effect (j= 1-6)}$$

$$F_k = \text{main Family effect (k= 1-3)}$$

$$DF_{jk} = \text{Density by Family interaction}$$

$$R_m = \text{Replication (m= 1-2)}$$

$$\varepsilon_{jkm} = \text{Error}$$

G2F Evaluation

For evaluating the production-level grain yield and ear heights reported by G2F, a mixed-effects linear model was used (Equation 6). The Line nested in Family and Family terms were fixed, while Environment, Replication nested in Environment, and Line nested in Family x Environment were all considered random effects. Best linear unbiased estimators (BLUEs) were estimated for each line using this model for grain yield and ear height within the two G2F Wisconsin environments and then across all 12 G2F environments. The calculated BLUEs were then used in Pearson correlations with BLUEs estimating the EID yield component traits.

Equation 6:

$$Y_{jklm} = \mu + E_j + F_k + L(F)_l + EL(F)_{jl} + R(E)_m + \varepsilon_{jklm}$$

μ = population mean

E_j = Environment effect (j= 1-2 for WI environments, j=1-12 for all G2F environments)

F_k = Family effect (k= 1-3)

$L(F)_l$ = Line effect (l= 1-385) nested within Family (F; 1-3)

$EL(F)_{jl}$ = Environment by Line nested in Family interaction

$R(E)_m$ = Replication (m= 1-2) nested in Environment

ε_{ijklm} = Error

All analyses for Equations 1 through 6 were done with the *lme4* package (Bates et al., 2015) in R (R Core Team, 2020) and ANOVAs were conducted to assess sources of variation. To make comparisons across families for the traits with significant Family x Density interactions from Equations 4 and 5, least square means were estimated using the *lsmeans* package (Lenth, 2016) in R. For comparisons between Family and Density factors, the P-values were adjusted using the Tukey HSD test with a significance level of $\alpha = 0.05$.

Genetic Mapping

Genetic information was used only in combination with the EID evaluation for genetic mapping. The lines were sequenced at ~5x coverage and read against a database populated with various genome assemblies' sequences that have been aligned to the B73 v4 genome (Jiao et al., 2017) to be inferred using the Practical Haplotype Graph (PHG, https://bitbucket.org/bucklerlab/practicalhaplotypegraph/wiki/Home). As described by Jensen et

al. (2020), with PHG, all parental genotypes are sequenced at high coverage, and loaded as parental haplotypes in a relational database. Progeny are then sequenced at lower coverage and a Hidden Markov Model (HMM) is used to infer which parental haplotypes/genotypes from the database are most likely present in a given progeny. Genic or gene-proximal regions are then called as SNPs. Initial filtering removed monomorphic markers, markers that were missing calls in all samples (Genomes to Fields, 2020), and markers with minor allele frequency (MAF) less than 0.05. There was genotypic information for 214 of the original 385 DH lines that were planted in EID (39 from MoG, 82 from Mo44, 93 from PHN11, as well as the three Founder parents and the reference parent, PHW65; Supplementary Table 3.1). A distance matrix was calculated in TASSEL (Bradbury et al., 2007) as 1 – identity by state (IBS) and used to show relationships between all 218 genotyped lines.

For genetic mapping, best linear unbiased estimates (BLUEs) were calculated for each of the lines as Line nested in Family plus Family effects for each Treatment using *lme4* and Equation 1. BLUEs were calculated only for the traits that were grown in both years – ear height, kernel row number, total kernel weight, number of kernels, weight per kernel, kernel width, kernel depth, and kernel area. A matrix of the 218 lines genotyped at 1,027,702 SNPs was analyzed using the R/qtl2 (Broman et al., 2019) package in R for joint linkage mapping. The cross type used was for doubled haploids and to account for the three families, family was used as a covariate in the model (*intcovar*). Genotype probabilities were calculated using a HMM (*calc_genoprob*) and kinship was also calculated using the leave one chromosome out method (*calc_kinship*). A mixed linear model was fit using the genotype probabilities, kinship matrix, and family covariate (*scan1*) for all nine phenotypes (ear height, kernel row number, number of ears, total kernel weight, number of kernels, weight per kernel, kernel width, kernel depth, and

kernel area) with two years of data. To establish statistical significance, 100 permutations for each phenotype and a 5% significance threshold was used to determine the logarithm of the odds (LOD) and identify a locus as significant (*scan1perm*). This linkage analysis was conducted for each of the six density treatments across the nine kernel measurements.

3.3 Results

EID Evaluation

The six planting density treatments used in the EID format are reflective of previous, current, and potential planting densities. Planting density in the United States in the 1930's was around 30,000 plants ha^{-1} (treatments 1 and 2 are below this value at 17,205 and 23,456, respectively) and has increased steadily to 81,000 plants ha^{-1} today (treatments 3 and 4 are below this value with densities of 36,790 and 51,604 plants ha^{-1}, respectively). Treatment 5 is slightly above the value at 85,925 plants ha^{-1} and treatment 6 is drastically larger at 258,385 plants ha^{-1}, representing potential planting densities if increases continue for treatment 5 and a substantially higher density for treatment 6.

The phenotypic variability across all lines for yield component traits was summarized for each of the treatments (Figures 3.3, 3.4, and 3.5). There were significant differences between years for ear height, number of ears, total kernel weight, weight per kernel, kernel depth, and kernel area. When evaluating the mean performance across all lines and density treatments, 2018 resulted in higher average ear heights (123.2 compared to 117.14 cm), kernel depth (510.53 versus 550.04 mm), and larger kernel area (157,202.7 compared to 139,775.4 mm^2). Overall, 2019 produced higher average kernel weights (31.84 versus 18.77 g) and weight per kernel (0.48 compared to 0.34 g). All of the previously mentioned traits exhibited significant Family x Year interaction. While the performance itself differed, the ranking was not significantly different (Spearman rank correlation of ρ=0.37) across Years, supporting the decision to combine years for evaluation.

For the yield components that were evaluated in both 2018 and 2019 (Equation 1; ear height, number of ears, kernel row number, total kernel weight, number of kernels, weight per kernel, kernel width, kernel depth, and kernel area), all were significantly affected by the varying

planting density ($P<0.05$). The number of kernels per 20 mL increased from treatments 1 through 6, with averages ranging from 60.01 kernels at treatment 1 to an average of 71.92 kernels at treatment 6. The average area of each kernel decreased from treatments 1 through 6 and showed the largest range of variability of all kernel components, ranging from 83,851.43 to 226,688.70 mm^2. Kernel weight of 20 mL samples ranged from 15 g (treatment 6) to 26.4 g (treatment 3) and large variability was observed across all treatments. The number of ears had the smallest range across lines and planting densities with a minimum of 1 and a maximum of 2 and because 90% of plots did not have more than one ear, all secondary ears were removed from further analysis. In density treatments 1 to 3, plants produced an average of 1.25 ears. On average, planting density treatments 1 and 2 resulted in the maximum yield component performance, except for ear height (minimum of 50 cm at treatment 1 and maximum of 185 cm at treatment 6) and the number of kernels, which increased with planting density treatments.

There was less variability observed across treatments for ear and cob traits measured only in 2018 (Equation 2) compared to the kernel traits previously described. The treatments had no significant effect ($P<0.05$) on maximum ear width (mean of 93.19 $\pm$ 50 cm), average ear width (mean of 26.31 $\pm$ 1.14 cm), ear length (mean of 20.34 $\pm$ 6.77 cm), or maximum cob width (mean of 76.78 $\pm$ 15.68 cm). Cob and ear weight were both significantly affected by treatment ($P<0.05$) and were the heaviest at treatment 2 (mean of 30.22 and 274.37 g, respectively) and the lightest at treatment 6 (mean of 13.75 and 80.89 g, respectively).

When considering differences across families for yield component traits measured in 2018 and 2019 using Equation 4, ear height, total kernel weight and kernel row number were significantly different ($P<0.05$) across both planting density treatments and families. Variability of trait performance was largest in the MoG family and smallest in the PHN11 family for a

subset of traits (Table 3.3). For example, MoG produced maximum ear height (185 cm at treatment 6), kernel width (treatment 3), kernel depth and area (both at treatment 2). The PHN11 family had the smallest range for total kernel weight and kernel width. Mo44 lines tended to perform midway in trait values between PHN11 and MoG for yield components. There were no significant differences between families for several of the ear and cob traits (Equation 5). While maximum ear width, average ear width, ear length, maximum cob width, average cob width, and cob length were not significantly ($P<0.05$) across families, ear and cob weights were ($P<0.05$). PHN11 produced the lightest ears (mean of 160.42 g) and cobs (mean of 24.28 g) while MoG produced the heaviest cobs (mean of 28.1) and Mo44 the heaviest ears (mean of 187.8).

There were significant interactions between treatment and family for the number of ears ($P<0.001$), kernel depth ($P<0.01$), kernel area ($P<0.01$), and total ear weight ($P<0.01$). Planting density treatments 1 to 3 averaged 1.25 ears per plant, while treatments 4 to 6 consistently produced only one ear per plant across all three families. Mo44 overcame its steady middle rank and produced the highest number of ears per plant at treatments 2 and 3 (Figure 3.6A). The total weight of the ears produced increased from treatments 1 to 2 but declined steadily in treatments 3 to 6. For treatment 2, MoG did produce significantly heavier ears than PHN11 and Mo44 and it was also the largest weight across all treatments (Figure 3.6B). Both the kernel area and depth increased slightly from treatments 1 to 2 across families but followed a similar trend and declined in value from treatments 3 to 6. An outlier to this trend was Mo44, which at treatment 4 produced significantly larger kernels (in terms of area and depth) than the other two families (Figures 3.6C and 3.6D).

Across all treatments, there were numerous correlations between ear, cob, and kernel traits (Figure 3.7). The Pearson correlations show significantly negative correlations ($P<0.05$)

between the number of kernels and total ear weight, total cob weight, weight per kernel, kernel depth, kernel area and kernel width. Significant negative correlations ($P<0.05$) were also observed between the number of ears and cob length, and total kernel weight. Kernel row number and kernel width, number of kernels and weight per kernel also show significant negative correlations ($P<0.05$). There were positive correlations (>0.60; $P<0.05$) between total kernel weight and cob length, kernel depth and kernel area, and total ear weight and ear weight per ear. Using Equation 3, heritabilities were calculated for all traits across all treatments (Figure 3.2). The heritabilities across traits ranged from 0.1 for cob length to 0.72 for ear height. Ear height and kernel related traits had relatively higher heritabilities, while the ear and cob traits tended to be lower.

G2F Evaluation

The average grain yield across all families was significantly different ($P<0.05$) between WIH1 (12.55 ± 2.3 ton ha^{-1}) and WIH2 (8.73 ± 2.61 ton ha^{-1}). Across all 12 G2F environments, average grain yield of lines ranged from 7.05 to 12.53 ton $ha^{-1}$with an average of 9.43 ± 2.42 ton ha^{-1}. When comparing the BLUPs across families using Tukey HSD, the highest average grain yield was the PHN11 family at 9.76 ton ha^{-1} which was not significantly different ($P<0.05$) from the Mo44 family at 9.54 ton ha^{-1}. The MoG family averaged the lowest yield (9.0 ton ha^{-1}) and was significantly less than both PHN11 and Mo44.

The average ear height across all 12 G2F environments ranged from 61 to 167 cm with an average of 111.8 ± 17.21 cm and the average ear height of each family was significantly different. MoG produced the tallest ear heights at 116.81 cm as well as the largest range of values (61 to 160 cm), the Mo44 family averaged ear heights of 112.23 cm, and PHN11

produced the shortest ear heights at 106.95 cm as well as the smallest range of values (70 to 167 cm).

The planting density used at the two WI G2F environments (78,939 plants ha^{-1}) was most similar to the EID planting density treatment 5 (85,925 plants ha^{-1}). Using the BLUEs calculated from Equation 6, Pearson correlations were calculated between grain yield and ear height and the yield component traits measured in the EID experiment at planting density treatments 2, 3, 4, and 5 (Tables 3.7 and 3.8). Significant Pearson correlations of yield component traits with grain yield ($P<0.05$) ranged from r=-0.22 to 0.21 within the WI environments with kernel size traits being negatively correlated and the number of ears and cob size being positively correlated. Across the 12 G2F environments, significant Pearson correlations ($P<0.05$) ranged from r=-0.30 to 0.26 with kernel traits being the most negatively correlated and the cob size being positively correlated. As grain yield increases, the number of ears per plant and cob size also increase. Alternatively, increases in grain yield are correlated with fewer kernels in a 20 mL volume. The significant Pearson correlations with ear height ($P<0.05$) ranged from -0.21 to 0.25 in WI environments and -0.22 to 0.48 in the G2F environments. Across both sets of correlations, the most positively correlated yield component traits were ear height (at treatments 3 and 5) and cob size traits were the most negatively correlated.

Joint Linkage Genetic Mapping

The MDS plot with principal components 1 (PC1) and 2 (PC2) shows genetic relationships among the 218 lines in PHW65 NAM. The reference line, PHW65, which all DHs share, is located in the origin and the three families are spread out into three distinct areas outside from the PHW65 point. Further out from the family clusters are the respective founder parents of

MoG, Mo44, and PHN11, showing their genetic dissimilarity from one another, but relatedness to their progeny lines.

From the joint linkage analysis, we would expect to identify QTL for kernel-related yield components across multiple planting densities. Using kernel yield components BLUEs (Equation 1), seven QTL were identified on chromosomes 6, 7, 8, and 9 in the PHW65 NAM (Table 3.8). Five of the QTL were identified in density treatment 4, one in treatment 5, and one in treatment 6. There are two "clusters" of QTL peaks: chromosome 7 at 135.48 Mb (overlap of kernel area and kernel depth) and chromosome 8 at 57.44 Mb (range of 125 bp and an overlap of kernel depth, width, and area). The percent of phenotypic variation explained by the QTL ranged from 10.79% for ear height at planting density 6 up to 22.38% for kernel depth at planting density 4.

The peaks on chromosomes 6, 8, and 9 did not fall within regions of any documented genes. The LOD interval on chromosome 7 for kernel depth and area is located very near the gene *GRMZM2G436981*, or *krp11*. The gene *krp11* is 6,449 bp in length and relates to kernel-size variation. Mutants with *krp11* exhibit a varied-kernel-size phenotype and often have segregated small seeds of varying sizes (Huang et al., 2019). . When estimating QTL effects of the kernel area peak on chromosome 7, it can be parsed out by each of the four alleles (MoG, Mo44, PHN11, and PHW65 as the reference). The Mo44 alleles attributed most to this peak, which aligns with previously mentioned findings that Mo44 had significant Family x Treatment interaction at treatment 4.

3.4 Discussion

Dissection of grain yield into yield component traits relating to ear, cob, and kernel traits revealed a range of responses to increasing planting density. The ear length, ear width, and cob width did not significantly change with an increase in density from 17,209 to 258,148 plants ha^{-1}. Our study found that kernel row number was only reduced at density treatment 6 (258,148 plants ha^{-1}), whereas kernel size traits (depth, width, and area) were highly susceptible to changes in plant density. This was reflected in the slight increases in size at treatments 1 through 3 and then decreases in size from 4 through 6.

The materials used in this study represent known differing levels of selection: MoG being the less selected, Mo44 having undergone moderate levels of selection, and PHN11 being an exPVP and the most commercially relevant grouping. Across all EID planting densities, PHN11 consistently produced the lightest cobs and shortest ear heights. MoG, on the other hand, produced the heaviest cobs and kernels, largest kernels, and greatest ear heights. It has been observed that grain yield has been steadily increasing since the 1930's in North America and this trend was corroborated in the production-level G2F setting with PHN11 yielding the greatest and MoG yielding the least, on average, across the 12 environments. As previous studies have indicated, plant breeders have been breeding for shorter plant stature (Duvick, 2005) and this observation was validated with significant decreases in ear height from MoG to Mo44 to PHN11 derived lines. When comparing the EID data to production-level BLUEs, ear height had positive correlations with the EID ear height at treatments 3 and 5 as well as kernel traits at treatment 4. There were negative correlations between G2F ear height and cob length across treatments 2, 3, 4, and 5. It was observed that the grain yield BLUEs across all G2F environments were positively correlated with the EID number of ears at treatment 2 and cob size traits. Alternatively, grain yield was negatively correlated with kernel size traits and ear width.

Previous studies have noted a decrease in kernel size negatively impacts yield (Coretz-Mendoza and Hallauer, 1979; Salazar and Hallauer, 1986; Lopez-Reynoso and Hallauer, 1998). Knowing that there are correlations for the same phenotypic measurements between these two formats (ear height at the production-level to ear height at EID treatments 3 and 5), supports the use of EID as a surrogate model to evaluate varying planting density effects. To evaluate planting density response, conventional yield trials typically require large amounts of field space and the number of lines being evaluated can be limited by the space and necessary seed requirements. While the EID plot design was only used to evaluate the pressures of varying planting densities on yield components in this study, future studies could use this schematic to understand other abiotic stresses that are observed with increased planting density.

The yield components that were significantly affected by both family and treatment (separately) were related to the kernels. It has been argued that kernel number is responsible for most yield variation (Early et al., 1967; Otegui, 1995; Chapman and Edmeades, 1999). Kernel traits are a function of the physiological conditions of the plant at the period just before and after silking (Andrade et al., 1993). The number of kernels and their size depend on the accumulation and allocation of nutrients from the plant and the efficiency of using this biomass specifically for kernel set (Borrás et al., 2018). Our study found that as planting density increased, the size and weight of the kernels on the primary ears declined. Planting density treatments 1 to 3 averaged 1.25 ears per plant, while treatments 4 to 6 consistently produced only one ear per plant across all three families, and this reduction in ear production with increased plant competition has been observed in previous studies (Tollenaar et al., 1992; Otegui, 1995).

The relationship for kernel components was rarely linear across the six planting density treatments. While the increase in planting density was not even across treatment, the observed

relationships align with previous findings (Nienhuis and Singh, 1985; Tetio-Kagho and Gardner, 1988; Hashemi-Dezfouli and Herbert, 1992). The two types of relationships between plant performance and planting density that are usually found are: asymptotic and parabolic (Villalobos et al., 2016). An asymptotic relationship describes when yield increases with density, but reaches a plateau eventually, as observed with kernel row number. A parabolic relationship describes one when yield reaches a maximum at a given density but decreases for any densities above or below that, which was observed across many yield components at a planting density of 23,456 plants ha^{-1} (treatment 2 of EID) including number of ears, total cob weight, total ear weight, total kernel weight, and kernel width, depth, and area.

Of the kernel measurements that had significant Family x Treatment interactions, the Mo44 family produced significantly larger kernels (in terms of kernel depth and area) at density treatment 4. This stark increase in kernel size could be indicative of optimum planting density for productivity of specific lines within Mo44. When evaluating kernel size traits at the genetic level through genetic mapping, these same phenotypes at density treatment 4 resulted in a significant association peak position. Given the relatedness of some of the kernel measurements (i.e., kernel width and depth to kernel area), colocalization of QTL was not surprising. This also suggests that because genetic architecture of kernel traits is related, it might not be necessary to measure all traits to conduct initial genetic mapping studies if materials or resources are limited. A previous study using the EID plot design combined with the image-based phenotyping with yield components identified 171 total QTL with 24 being related to kernel yield components. These QTL were identified in density treatments 1, 3, 4, and 5 and on all chromosomes, except 2. There was also co-localization of two traits (kernel area and width) at a peak on chromosome 9 (White et al., unpublished). Another study on kernel trait components in maize identified 28 QTL related

to various measurements and noted that 10 of the QTLs involved in three traits were found in common genomic regions. This suggests that traits that are closely linked might lead to the co-localization for different traits (Jiang et al., 2015). An area for further exploration would be to re-sequence more of the lines for joint linkage analysis. Only 55% of the lines that were phenotyped were successfully genotyped (214/385). The absence of most MoG genotypic information (only 31% of lines were genotyped) could have resulted in fewer MoG alleles being detected that are contributing to phenotypic variation that was observed in the yield components. The relatively small sample size for genetic mapping is likely to hinder QTL detection of smaller effects and overestimate the effect size of any QTL that were identified (Beavis et al., 1991).

There are some known limitations of the EID study due to its design. Since the same planting disk was used for each plot and the disk always planted from low to high density treatments, it was not possible to randomize the planting density treatments. It was also observed that there weren't consistently 0.61 m of space between treatment 6 at the end of one plot and treatment 1 of the following plot. This lapse in precision could confound results since there was only a single plant in Treatment 1 to use for phenotypic and image analyses and this plant underwent more competition than what was intended.

The EID plot design is able to measure the genotype's response to various forms of abiotic stress but requires fewer resources compared to traditionally used multi-environment trials. Future research could focus on parsing out the specific abiotic stresses at each planting density to better understand a genotype's tolerance, or lack thereof, to weed pressure, limited moisture, sunlight, and soil nutrients, to name a few. Through the use of image-based phenotyping, the phenotypic variation of kernel yield components with increased planting densities has been documented as well as the variability that exists across different genetic

backgrounds. Kernel size and weight have been associated with seedling vigor (Peng et al., 2011; Sekhon et al., 2014) so further investigation in how these multiple stressors impact plant emergence and establishment could be useful in identifying lines that will be most resilient to increased planting density.

Acknowledgments

The author would like to thank Dr. Nathan Miller for support with yield component data compilation and Drs. Mike White, Karl Broman, and Lucia Gutierrez for assistance with data analysis. The author would also like to thank the Genomes to Fields consortium for providing the data and materials used in this study. This consortium involves more than 30 researchers representing more than 20 research institutions. Details about the initiative and publicly available resources can be found at www.Genomes2Fields.org. Funding has been provided for this research from the USDA National Institute of Food and Agriculture (NIFA) project award #2016-67024-2219.

LITERATURE CITED

Andrade, F.H., S.A. Uhart, and A. Cirilo. 1993. Temperature affects radiation use efficiency in maize. *Field Crops Research.* 32(1): 17-25. https://doi.org/10.1016/0378-4290(93)90018-1.

Austin, D.F., and M. Lee. 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.

Bates, D., M. Mächler, B.M. Bolker, and S.C. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1-48.

Beavis, W.D., Grant, D., Albertsen, M., Fincher, R. 1991. Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. *Theoretical and Applied Genetics.* 83(2):141-145. https://doi.org/10.1007/BF00226242.

Borrás, L., and L.N. Vitantonio-Mazzini. 2018. Maize reproductive development and kernel set under limited plant growth environments. *Journal of Experimental Botany* 69(13):3235-3242. https://doi.org/10.1093/jxb/erx452.

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23(19): 2633-2635. https://doi.org/10.1093/bioinformatics/btm308.

Broman, K.W., D.M. Gatti, P. Simecek, N.A. Furlotte, P. Prins, Ś. Sen, B.S. Yandell, and G.A. Churchill. 2019. R/qtl2: Software for mapping quantitative trait loci with high-dimensional data and multiparent populations. *Genetics* 211(2):495-502. https://doi.org/10.1534/genetics.118.301595.

Butzen, S., and M. Burnison. 2014. Maize seeding rate considerations for 2014. Crop Insights 24(3). DuPont Pioneer.

Cavanagh, C., M. Morell, I. Mackay, W. Powell. 2008. From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Current Opinion in Plant Biology* 11(2):215-221. https://doi.org/10.1016/j.pbi.2008.01.002.

Chen, J., L. Zhang, S. Liu, Z. Li, R. Huang, Y. Li, H. Cheng, X. Li, B. Zhou, S. Wu, W. Chen, J. Wu, J. Ding. 2016. The Genetic Basis of Natural Variation in Kernel Size and Related Traits Using a Four-Way Cross Population in Maize. *PLoS One*, 11(4): e0153428. https://doi.org/10.1371/journal.pone.0153428.

Cortez-Mendoza, H., and A.R. Hallauer. 1979. Divergent Mass Selection for Ear Length in Maize. *Crop Science* 19(2):175-178. https://doi.org/cropsci1979.0011183X001900020001x.

Doerge, R.W. 2002. Mapping and analysis of quantitative trait loci in experimental populations.

Nature Review Genetics 3(1):43-52. https://doi.org/10.1038/nrg703.

Duvick, D.N. 2005. Genetic progress in yield of United States maize (Zea mays L.). *Maydica* 50(3):193-202.

Edmeades, G.O., and S.C. Chapman. 1999. Selection Improves Drought Tolerance in Tropical Maize Populations: II. Direct and Correlated Responses among Secondary Traits. *Crop Science* 39(5):1315-1324. https://doi.org/10.2135/cropsci1999.3951315x.

Falconer, D.S., and T.F.C. Mackay. 1996. Introduction to Quantitative Genetics (Fourth Edition).

Farnham, D.E. 2001. Row Spacing, Plant Density, and Hybrid Effects on Maize Grain. Yield and Moisture. *Agronomy Journal* 93(5):1049-1053.

Fehr, W. 1987. Principles of cultivar development: Theory and techniques.

Fehr, W. 1991. Principles of cultivar development: Theory and techniques.

Flint-Garcia, S.A., A.C. Thuillet, J. Yu, G. Pressoir, S.M. Romero, S.E. Mitchell, J. Doebley, S. Kresovich, M.M. Goodman, and E.S. Buckler. 2005. Maize association population: A high-resolution platform for quantitative trait locus dissection. *The Plant Journal* 44(6):1054-1064. https://doi.org/10.1111/j.1365-313X.2005.02591.x.

Gage, J.L., M.R. White, J.W. Edwards, S. Kaeppler, and N. de Leon. 2018. Selection Signatures Underlying Dramatic Male Inflorescence Transformation During Modern Hybrid Maize Breeding. *Genetics* 210(3): 1125 LP – 1138 Available at http://www.genetics.org/content/210/3/1125.abstract.

Garin, V., V. Wimmer, S. Mezmouk, M. Malosetti, and F. van Eeuwijk. 2017. How do the type of QTL effect and the form of the residual term influence QTL detection in multi-parent populations? A case study in the maize EU-NAM population. *Theoretical and Applied Genetics* 130(8):1753-1764. https://doi.org/10.1007/s00122-017-2923-3.

Genomes to Fields. 2020. Genomes to Fields 2018 Data Set (Version 1). CyVerse Data Commons. Available at https://datacommons.cyverse.org/browse/iplant/home/shared/commons_repo/curated/GenomesToFields_G2F_Data_2018

Haase, N. (2015). *Phenotypic Analysis and Genetic Dissection of Yield Component Traits in Maize (Zea mays L.)* (Doctoral dissertation, University of Wisconsin-Madison). Available from ProQuest Dissertations and Thesis database. (UMI No. 10187943).

Hashemi-Dezfouli, A., Herbert, S.J. 1992. Intensifying plant density response of maize with artificial shade. *Agronomy Journal*. 84: 547-551. https://doi.org/10.2134/agronj1992.00021962008400040001x.

Henry, M.L., and Westoby, M. 2001. Seed mass and seed nutrient content as predictors of seed output variation between species. *Oikos* 92(3): 479-490. https://doi.org/10.1034/j.1600-0706.2001.920309.x.

Hütsch, B.W., D. Jahn, S. Schubert. 2018. Grain yield of wheat (Triticum aestivum L.) under long-term heat stress is sink-limited with stronger inhibition of kernel setting than grain filling. *Journal of Agronomy and Crop Science* 205(1):22-32. https://doi.org/10.1111/jac.12298.

Jensen, S.E., J. Rigaud Charles, K. Muleta, P.J. Bradbury, T. Casstevens, S.P. Deshpande, M.A. Gore, R. Gupta, D.C. Ilut, L. Johnson, R. Lozano, Z. Miller, P. Ramu, A. Rathore, M.C. Romay, H.D. Upadhyaya, R.K. Varshney, G.P. Morris, G. Pressoir, E.S. Buckler, G.P. Ramstein. 2020. A sorghum practical haplotype graph facilitates genome-wide imputation and cost-effective genomic prediction. *The Plant Genome* 13(1):e20009. https://doi.org/10.1002/tpg2.20009.

Jiang, L., Ge, M., Zhao, H., Zhang, T. 2015. Analysis of Heterosis and Quantiative Trait Loci for Kernel Shape Related Traits Using Triple Testcross Population. In Maize. *PLoS ONE* 10(4): e0124779.

Jiao, Y., P. Peluso, J. Shi, T. Liang, M.C. Stitzer, B. Wang, M.S. Campbell, J.C. Stein, X. Wei, C.S. Chin, K. Guill, M. Regulski, S. Kumari, A. Olson, J. Gent, K.L. Schneider, T.K. Wolfgruber, M.R. May, N.M. Springer, E. Antoniou, W.R. McCombie, G.G. Presting, M. McMullen, J. Ross-Ibarra, R.K. Dawe, A. Hastie, D.R. Rank, and D. Ware. 2017. Improved maize reference genome with single-molecule technologies. *Nature* 546(7659):524-527. https://doi.org/10.1038/nature22971.

Lauer, J. 2008. Mid- and Late- Season Yield Enhancing Management Practices for Maize. *Field Crops*. Available at http://corn.agronomy.wisc.edu/AA/pdfs/A058.pdf.

Lenth, R. V. 2016. Least-squares means: The R package lsmeans. *The Journal of Statistical Software* 69(1). https://doi.org/10.18637/jss.v069.i01.

Li, Y., Wang, Y., Shi, Y., Song, Y., Wang, T., Li, Y. 2009. Correlation analysis and QTL mapping for traits of kernel structure and yield components in maize. *Scientia Agricultura Sinica* 42(2):408-418.

Lopez-Reynoso, J.J., and A.R. Hallauer. 1998. Twenty-seven cycles of divergent mass selection for ear length in maize. *Crop Science* 38(4):1099-1107. https://doi.org/10.2135/cropsci1998.0011183X003800040035X.

Mansfield, B.D., and R.H. Mumm. 2014. Survey of plant density tolerance in U.S. maize germplasm. *Crop Science* 54(1), 157-173. https://doi.org/10.2135/cropsci2013.04.0252.

Meghji, M.R., Dudley, J.W., Lambert, R.J., and Sprague, G.F. 1984. Inbreeding depression, inbred and hybrid grain yields, and other traits of maize genotypes representings three eras.

Crop Science 24(3):545-549. https://doi.org/10.2135/cropsci1984.0011183X002400030028x.

Messmer, R., Y. Fracheboud, M. Bänziger, M. Vargas, P. Stamp, and J.M. Ribaut. 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.

McMullen, M.D., S. Kresovich, H.S. Villeda, P. Bradbury, H. Li, Q. Sun, S. Flint-Garcia, J. Thornsberry, C. Acharya, C. Bottoms, P. Brown, C. Browne, M. Eller, K. Guill, C. Harjes, D. Kroon, N. Lepak, S.E. Mitchell, B. Peterson, G. Pressoir, S. Romero, M.O. Rosas, S. Salvo, H. Yates, M. Hanson, E. Jones, S. Smith, J.C. Glaubitz, M. Goodman, D. Ware, J.B. Holland, and E.S. Buckler. 2009. Genetic Properties of the Maize Nested Association Mapping Population. *Science* 325(5941):737-740. https://doi.org/10.1126/science.1174320.

Miller, N.D., N.J. Haase, J. Lee, S.M. Kaeppler, N. de Leon, and E.P. Spalding. 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.

Nienhuis, J. and Singh, S.P. 1985. Effect of location and plant density on yield and architectural traits in dry beans. *Crop Science* 25(4): 579-584. https://doi.org/10.2135/cropsci1985.0011183X002500040001x.

Otegui, M.E., M.G. Nicolini, R.A. Ruiz, P.A. Dodds. 1995. Sowing Date Effects on Grain Yield Components for Different Maize Genotypes. *Agronomy Journal* 87(1):29-33. https://doi.org/10.2134/agronj1995.00021962008700010006x.

Peng, B., Li, Y., Wang, Y. et al. 2011. QTL analysis for yield components and kernel-related traits in maize across multi-environments. *Theoretical and Applied Genetics* 122:1305-1320. https://doi.org/10.1007/s00122-011-1532-9.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rafalski, J.A. 2002. Novel genetic mapping tools in plants: SNPs and LD-based approaches. *Plant Science* 162(2):329-333. https://doi.org/10.1016/S0168-9452(01)00587-8.

Salazar, M.A., and A.R. Hallauer. 1986. Divergent mass selection for ear length in maize. *Brazilian Journal of Genetics* 9:281-294.

Sekhon, R.S., C.N. Hirsch, K.L. Childs, M.W. Breitzman, P. Kell, S. Duvick, E.P. Spalding, C.R. Buell, N. de Leon, S.M. Kaeppler. 2014. Phenotypic and Transcriptional Analysis of Divergently Selected Maize Populations Reveals the Role of Developmental Timing in Seed Size Determination. *Plant Physiology* 165(2):658-669. https://doi.org/10.1104/pp.114.235424.

Smith, S., Cooper, M., Gogerty, J., Loffler, C., Borcherding, D. and Wright, K. 2014. Maize.

Yield grains in major U.S. field crops: CSSSA Special Publication 33 (Chapter 6, pp. 125-172).

Tollenaar, M., A. Aguilera. 1992. Radiation use efficiency of an old and a new maize hybrid. *Agronomy Journal* 84(3):536-541. https://doi.org/10.2134/agronj1992.00021962008400030033x.

Troyer, A.F. 1996. Breeding widely adapted, popular maize hybrids. *Euphytica* 92:163-174. https://doi.org/10.1007/BF00022842.

Tetio-Kagho, F., Gardner, F.P. 1988. Responses of maize to plant population density. I. Canopy development, light relationship, and vegetative growth. *Agronomy Journal* 80(6):930-935. https://doi.org/10.2134/agronj1988.00021962008000060018x.

Villalobos, F.J., Sandras, V.O., Fereres, E. 2016. Plant Density and Competition. *Principles of Agronomy for Sustainable Agriculture*:159-168.

White. M.R. unpublished. An Evaluation of Allelic Diversity and Heterosis of Expired Plant Variety Protection Germplasm and a Novel Scheme to Evaluate Hybrid Response to Density in Maize. (Zea mays L.). (Doctoral dissertation, University of Wisconsin-Madison).

Yu, J., G. Pressoir, W.H. Briggs, I.V. Bi, M.Yamaski et al. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38(2):203-208. https://doi.org/10.1038/ng1702.

Table 3.1. List of the phenotypic traits measured in EID, respective trait units, and the years measured in. The traits are broken up into the yield components of general development, ear, cob, and kernel.

	Measurement	**Abbreviation**	**Units**	**Measured in 2018 and 2019**
General	Ear Height per treatment	Ear height	cm	X
	Number of Ears per plant per treatment harvested	Number of ears	#	X
Ear	Kernel Row Number	KRN	#	X
	Total Ear Weight for all ears per treatment	Total ear weight	g	
	Average Ear Width for each ear per treatment	Average ear width	mm	
	Maximum Ear Width for each ear per treatment	Maximum ear width	mm	
	Ear Length for each ear per treatment	Ear length	mm	
Cob	Total Cob Weight for all cobs per treatment	Total cob weight	g	
	Average Cob Width for each cob per treatment	Average cob width	mm	
	Maximum Cob Width for each cob per treatment	Maximum cob width	mm	
	Cob Length for each cob per treatment	Cob length	mm	
Kernel	Total Kernel Weight for all kernels in 20 mL sample cup per treatment	Total kernel weight	g	X
	Number of kernels in 20 mL sample cup per treatment	Number of kernels	#	X
	Average Weight per kernel per treatment	Weight per kernel	g	X
	Average Kernel Minor Axis per treatment	Kernel width	mm	X
	Average Kernel Major Axis per treatment	Kernel depth	mm	X
	Average Kernel Area per treatment	Kernel area	mm^2	X

Table 3.2 Testing environments' latitude, longitude, planting density, and average grain yield with standard error for the 12 locations that are part of the Genomes to Fields project field evaluation in 2018.

Environment	Latitude (°)	Longitude (°)	Planting Density (plants/hectare)	Average Grain Yield ± SE (t/ha)
MN1	44.08	-93.51	78,938	8.06 ± 1.92
WI1	43.30	-89.39	78,938	12.46 ± 2.11
WI2	43.05	-89.53	78,938	9.05 ± 2.77
MO1	38.90	-92.21	87,686	8.96 ± 1.63
NC1	35.67	-78.49	88,464	10.58 ± 2.05
NY1	42.73	-76.65	92,548	9.16 ± 1.97
MI1	42.69	-84.50	93,908	10.94 ± 1.50
NY2	42.73	-76.66	94,249	10.2 ± 2.24
OH1	39.86	-83.67	98,116	12.18 ± 1.45
IA4	42.00	-93.70	99,353	12.53 ± 1.37
DE1	38.65	-75.45	104,412	7.05 ± 2.31
ON2	42.45	-81.88	106,155	11.53 ± 1.01

Table 3.3 Variability of general, ear, cob, and kernel yield components across the three families, MoG, Mo44, and PHN11. The minimum, mean, standard deviation, and maximum, are reported for the families across all six planting density treatments.

Phenotypic Trait		MoG			Mo44			PHN11		
		Minimum	Mean ± SD	Maximum	Minimum	Mean ± SD	Maximum	Minimum	Mean ± SD	Maximum
General	**Ear Height (cm)**	85	130.81 ± 16.22	185	75	123.9 ± 16.66	165	65	113.57 ± 15.94	160
	# of ears	1	1.15 ± 0.36	2	1	1.21 ± 0.41	2	1	1.13 ± 0.34	2
	# of kernel rows	12	16.66 ± 2.03	22	8	16.22 ± 1.89	22	10	16.39 ± 1.93	22
Ear	**Total ear weight (g)**	8.2	171.39 ± 125.14	411.7	13.2	187.8 ± 100.33	397.3	6	160.42 ± 124.8	425.7
	Average ear width (mm)	24.33	26.43 ± 0.91	27.89	24.32	26.4 ± 0.89	27.96	24.13	26.56 ± 0.77	27.83
	Maximum ear width (mm)	121.14	122.35 ± 0.71	125.53	121.12	122.29 ± 0.72	124.57	121.14	122.41 ± 0.69	124.79
	Ear length (mm)	5.43	19.93 ± 6.9	38.89	5.45	20.35 ± 7.19	38.89	5.22	20.51 ± 6.93	38.68
Cob	**Total cob weight (g)**	3	28.1 ± 15.1	68.9	3	27.69 ± 12.89	66.5	3	24.28 ± 12.94	68.5
	Average cob width (mm)	37.17	43.31 ± 3.37	56.92	26.54	43.25 ± 3.63	62.95	37.25	43.22 ± 3.68	67.86
	Maximum cob width (mm)	34.66	76.48 ± 16.31	125	21.62	74.31 ± 16.24	121.41	23.13	74.43 ± 14.9	119.75
	Cob length (mm)	119.7	122.46 ± 0.88	124.66	119.76	122.47 ± 0.92	126.69	119.58	122.36 ± 0.87	128.53
Kernel	**Total kernel weight (g)**	15	18.75 ± 2.59	25.8	15	18.75 ± 2.47	26.4	15	18.68 ± 2.19	24.7
	# of kernels	28	55.11 ± 10.99	98	31	56.13 ± 12.39	98	34	59.64 ± 12.54	115
	Kernel weight per kernel (g)	0.18	0.34 ± 0.063	0.57	0.16	0.34 ± 0.084	0.65	0.15	0.32 ± 0.07	0.64
	Kernel width (mm)	278.55	358.18 ± 26.43	454.43	257.6	353.19 ± 28.32	453.97	267.4	346.86 ± 26.05	441.74
	Kernel depth (mm)	399.6	564.04 ± 46.64	700.94	376.68	555.49 ± 50.82	659.09	379.53	541.22 ± 53.12	661.96
	Kernel area (mm2)	99,546.48	164,212.3 ± 21,650.01	226,688.70	86,260.42	160,878.5 ± 21,650.01	211,907.90	83,851.43	153,449.5 ± 23,295.41	214,460.60

Table 3.4. Variance significance output from ANOVA for general and kernel traits from 2018 and 2019. Significance is indicated (Not significant = NS, P<0.1 = . , P<0.05 = *, P<0.01 = **, P<0.001 = ***).

Source of Variance	Ear Height	Number of Ears	KRN	Total Kernel Weight	Number of Kernels	Kernel Weight per Kernel	Kernel Width	Kernel Depth	Kernel Area
Year	.	***	NS	**	*	***	***	***	***
Rep Within Year	***	NS	NS	NS	NS	NS	NS	NS	NS
Treatment	***	***	***	***	***	***	***	***	***
Family	***	***	***	***	NS	NS	NS	***	**
Year*Treatment	NS	***	NS	NS	NS	NS	NS	**	**
Year*Family	***	***	***	***	***	**	***	***	***
Treatment*Family	***	**	NS	NS	**	*	*	***	***
Treatment*Family*Year	NS	NS	NS	NS	NS	.	NS	NS	NS

Table 3.5. Variance significance output from ANOVA for ear traits from 2018. Significance is indicated (Not significant = NS, P<0.1 = . , P<0.01 = **, P<0.001 = ***).

Source of Variance	Total Ear Weight	Maximum Ear Width	Average Ear Width	Ear Length
Treatment	***	NS	NS	NS
Family	**	NS	NS	**
Treatment*Family	**	NS	NS	NS
Rep	NS	***	.	***

Table 3.6. Variance significance output from ANOVA for cob traits from 2018. Significance is indicated (Not significant = NS, P<0.1 = . , P<0.01 = **, P<0.001 = ***).

Source of Variance	Total Cob Weight	Maximum Cob Width	Average Cob Width	Cob Length
Treatment	***	NS	.	.
Family	***	NS	NS	NS
Treatment*Family	NS	NS	NS	NS
Rep	***	NS	NS	NS

Table 3.7. Pearson correlations between production-level BLUEs of G2F grain yield and the EID yield component traits across planting density treatments 2, 3, 4, and 5. BLUEs were calculated for the two Wisconsin G2F environments as well as all 12 G2F environments. Correlations were calculated using α=0.05 significance level and only significant relationships are shown. Magnitude of correlation ranges from -1 (colored red) to 1 (colored green).

EID Trait_Treatment	Grain Yield (WI)	Grain Yield (All G2F)
Ear Height_2	-	-0.04
Ear Height_3	0.21	-
Ear Height_4	0.02	-0.1
Ear Height_5	-0.06	-0.18
Number of Ears_2	0.15	0.13
Number of Ears_3	0.16	-
Number of Ears_4	-0.13	-0.3
Total Ear Weight_2	0.11	-0.06
Total Ear Weight_3	-	-0.13
Total Ear Weight_4	-0.03	-0.04
Total Ear Weight_5	0.07	0.26
Ear_Length_2	-0.22	-0.1
Ear_Length_3	0.02	-0.06
Ear_Length_4	-0.19	-0.28
Ear Max Width_2	-	0.08
Ear Max Width_5	-0.12	-0.2
Ear Avg Width_3	-0.12	-0.21
Ear Avg Width_4	-0.07	-0.07
Total Cob Weight_2	-0.14	-0.18
Total Cob Weight_3	0.17	0.19
Total Cob Weight_4	-0.05	-0.22
Total Cob Weight_5	0.08	0.18
Cob Max Width_3	-0.02	-0.06
Cob Max Width_4	-	0.05
Cob Max Width_5	0.1	0.22
Cob Avg Width_2	-	-0.08
Cob Avg Width_3	-	0.03
Cob Avg Width_4	-	0.12
Cob Avg Width_5	-	0.09
Cob_Length_3	-	0.18
Cob_Length_4	0.12	0.18
Cob_Length_5	0.09	0.13
Total Kernel Weight_3	-	0.12

Total Kernel Weight_5	-0.11	-0.17
Number of kernels_3	-0.07	-0.17
Number of kernels _4	0.06	0.06
Number of kernels _5	-0.15	-0.21
Weight per kernel_2	0.11	-
Weight per kernel_3	0.1	0.1
KRN_2	0.13	0.13
KRN_3	-	0.09
KRN_4	-0.14	-0.16
KRN_5	0.06	0.21
Kernel_Width_2	-	0.12
Kernel_Width_3	-0.13	0
Kernel_Width_4	-	-0.01
Kernel_Depth_2	-0.21	-0.07
Kernel_Depth_3	-0.17	-0.13
Kernel_Depth_4	0	-0.01
Kernel_Depth_5	-	0.07
Kernel_Area_2	-0.16	-
Kernel_Area_3	-0.17	-0.12
Kernel_Area_4	-0.05	-0.09

Table 3.8. Pearson correlations between production-level BLUEs of G2F ear height and the EID yield component traits across planting density treatments 2, 3, 4, and 5. BLUEs were calculated for the two Wisconsin G2F environments as well as all 12 G2F environments. Correlations were calculated using α=0.05 significance level and only significant relationships are shown. Magnitude of correlation ranges from -1 (colored red) to 1 (colored green).

EID Trait_Treatment	Ear Height (WI)	Ear Height (All G2F)
Ear Height_3	0.23	0.48
Ear Height_4	0.19	-
Ear Height_5	0.19	0.46
Number of Ears_3	0.09	0.24
Number of Ears_4	0.12	0.28
Total Ear Weight_2	-0.02	-0.05
Total Ear Weight_3	0.02	0.2
Total Ear Weight_4	0.11	0.07
Total Ear Weight_5	-0.14	-0.12
Ear_Length_2	-0.03	0.09
Ear_Length_4	0.2	0.29
Ear_Length_5	0.04	0.12
Ear Max Width_3	0.18	0.31
Ear Max Width_4	-0.2	-0.19
Ear Max Width_5	0.25	0.33
Ear Avg Width_3	0.09	0.19
Total Cob Weight_2	0.02	0.09
Total Cob Weight_3	0.03	-0.07
Total Cob Weight_4	0.16	0.38
Total Cob Weight_5	-0.21	-0.22
Cob Max Width_2	0.08	0.22
Cob Max Width_3	0.03	0.27
Cob Max Width_4	-0.08	0.01
Cob Max Width_5	-0.15	-0.08
Cob Avg Width_2	0.25	0.31
Cob Avg Width_3	-	-0.06
Cob Avg Width_4	-0.16	-0.22
Cob Avg Width_5	-0.14	-0.06
Cob_Length_2	-0.06	-0.11
Cob_Length_3	-0.09	-0.15
Cob_Length_4	-0.17	-0.2
Cob_Length_5	-0.1	-0.17
Total Kernel Weight_3	-0.11	0.01

Total Kernel Weight_5	-0.04	0.05
Number of Kernels_2	-	-0.1
Number of Kernels_4	-0.09	-0.21
Number of Kernels_5	0.04	0.02
Weight per Kernel _2	-0.14	-0.03
Weight per Kernel _3	-0.12	0.09
Weight per Kernel _4	-	-0.12
Weight per Kernel_5	-	-0.08
KRN_2	-0.09	-0.02
KRN_3	0.03	-0.1
KRN_5	-0.03	-0.08
Kernel_Width_2	-	-0.03
Kernel_Width_4	0.02	0.31
Kernel_Depth_2	0.06	-
Kernel_Depth_3	0.15	0.08
Kernel_Depth_4	0.07	0.36
Kernel_Area_3	0.13	0.07
Kernel_Area_4	0.09	0.39

Table 3.9. Significant QTL detected on kernel yield components using the Ever-increasing Density (EID) experiment. The kernel trait, EID treatment, chromosome, position (Mb), LOD peak using a threshold of α=0.05, and the total percent of phenotypic variation explained by the QTL model are reported.

Trait	EID Treatment	Chromosome	Position (Mb)	LOD	Percent Variation (%)
Kernel Area	4	7	135.476074	6.53	15.09
Kernel Area	4	8	57.438367	6.99	16.23
Kernel Width	4	8	57.438492	7.94	18.26
Kernel Depth	4	7	135.474828	8.03	18.49
Kernel Depth	4	8	57.43369	9.56	22.38
KRN	5	9	110.944207	6.38	14.40
Ear Height	6	6	162.451203	4.85	10.79

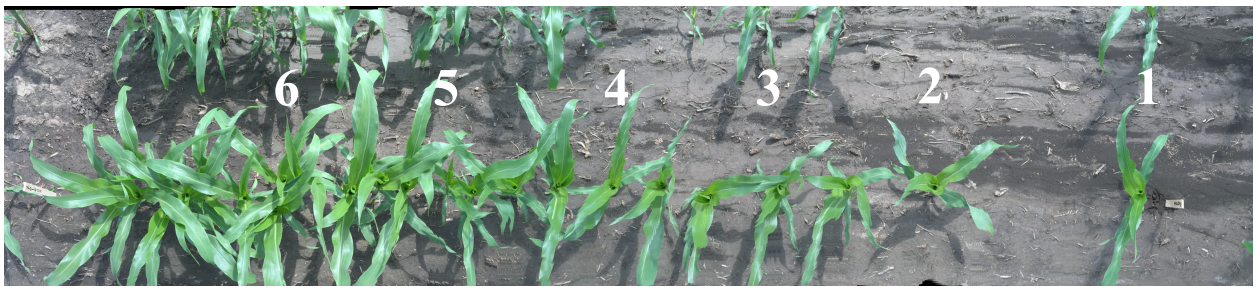


Figure 3.1. Image of an Ever-Increasing Density plot with density treatments labeled 1 through 6. In-row plant spacing ranges from 0.051 m (density treatment 1) to 0.76 m (density treatment 6).

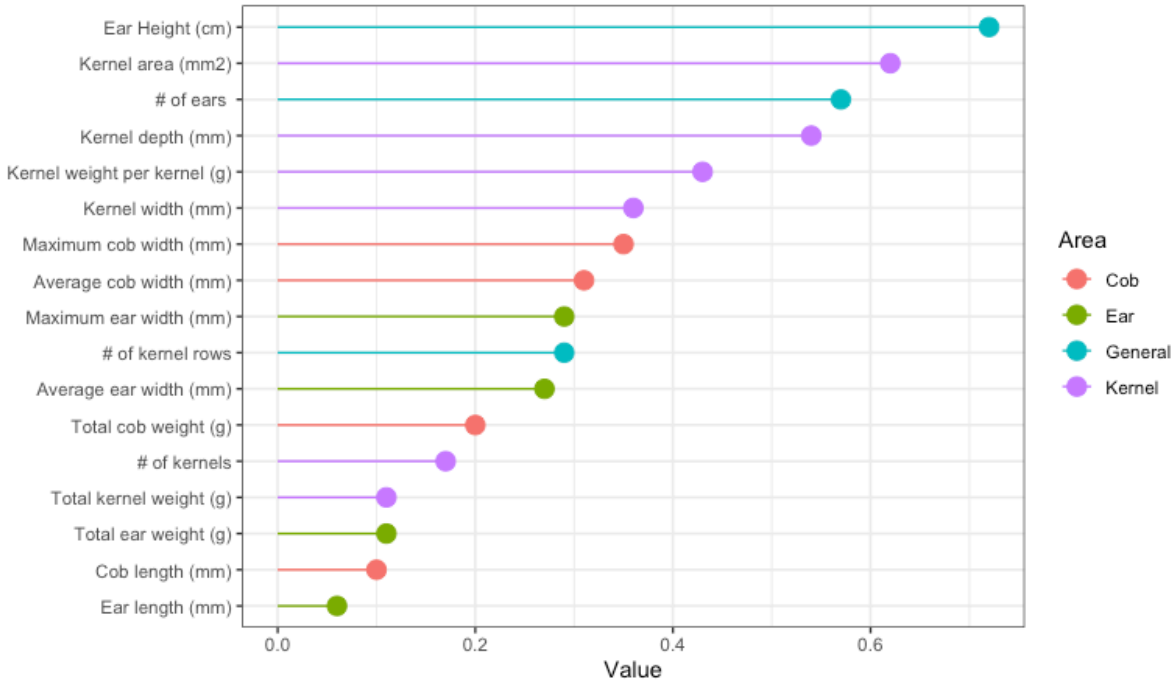


Figure 3.2. Heritability estimates for general, ear, cob, and kernel traits from the 2018 and 2019 growing seasons across all planting density treatments and lines in the EID.

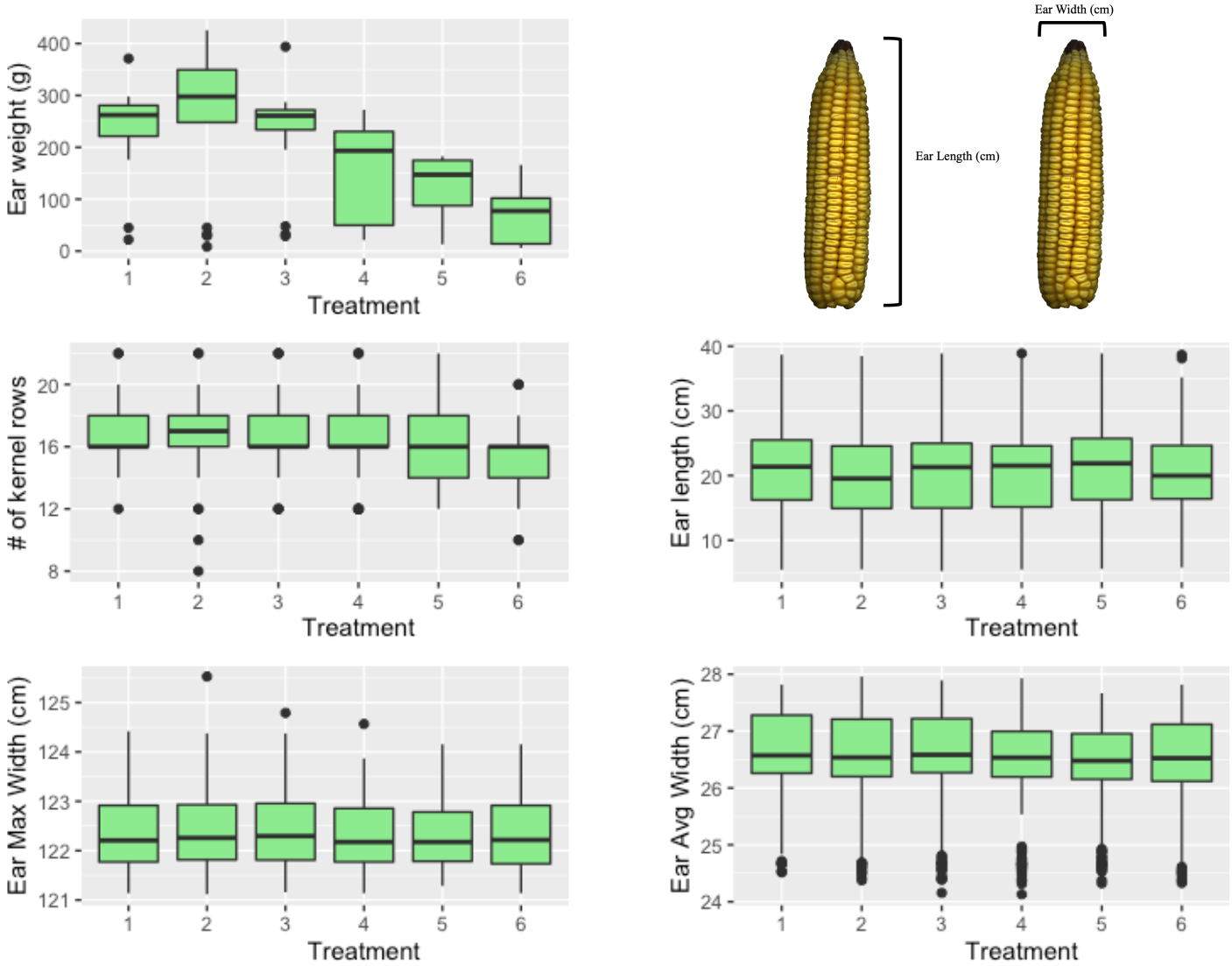


Figure 3.3. Boxplots of ear-related traits (ear weight, number of kernel rows, maximum ear width, ear length, and average ear width) across each of the six planting densities for all lines. A graphic of the ear length and width measurements is also included.

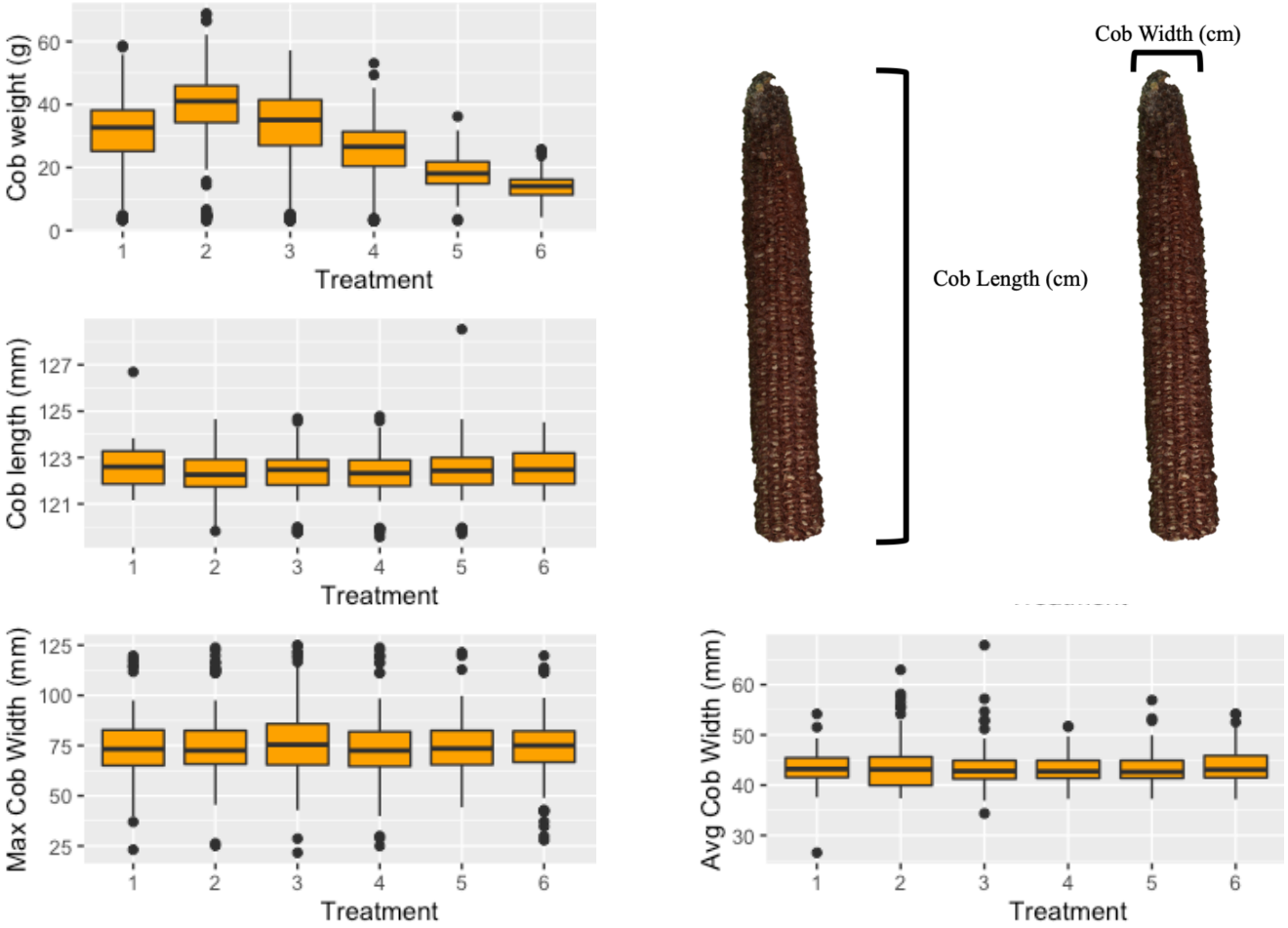


Figure 3.4. Boxplots of cob-related traits (cob weight, cob length, maximum cob width, and average cob width) across each of the six planting densities for all lines. A graphic of the cob length and width measurements is also included.

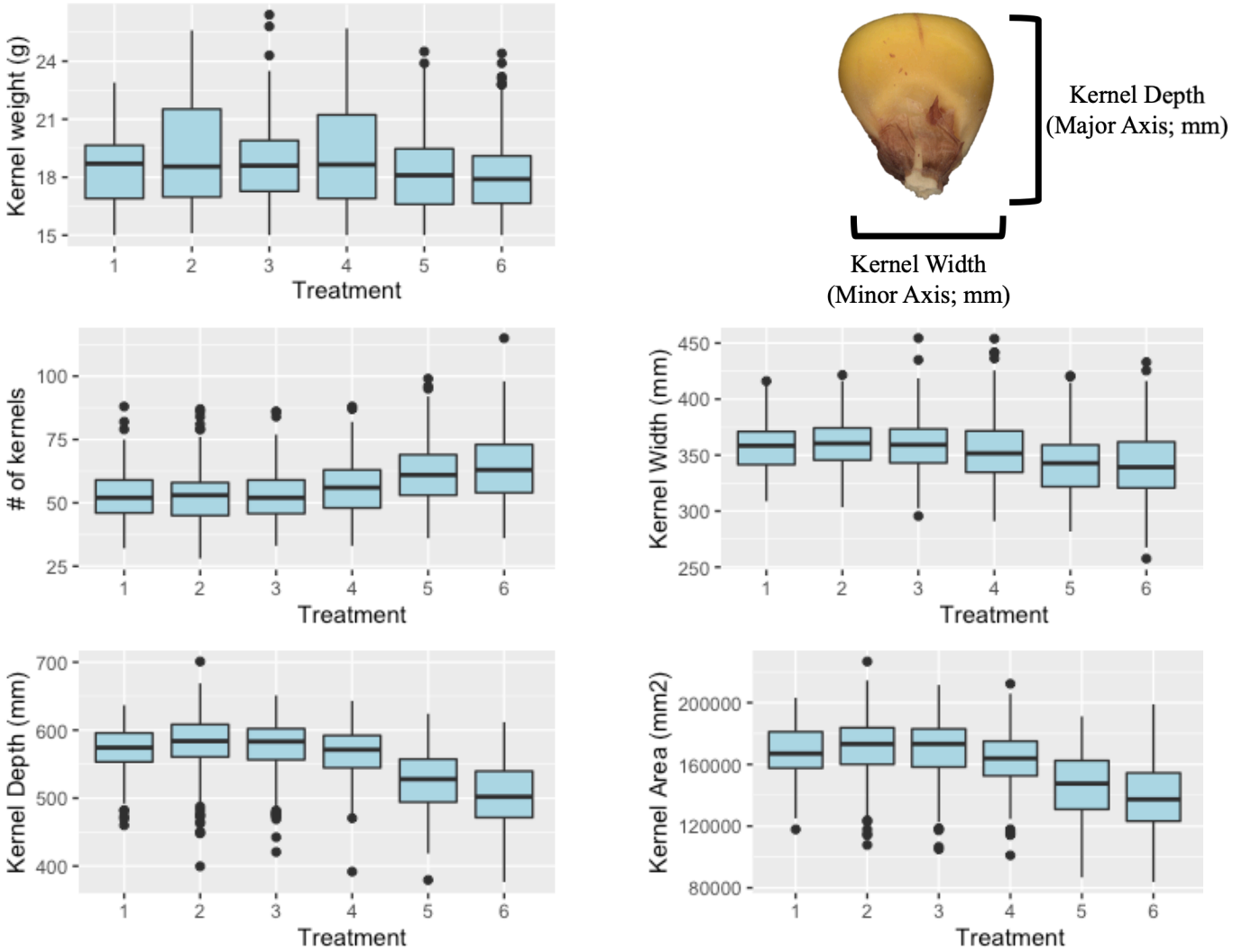


Figure 3.5. Boxplots of kernel-related traits (kernel weight, number of kernels, kernel depth, kernel width, and kernel area) across each of the six planting densities for all lines. A graphic of the kernel width and depth measurements is also included.

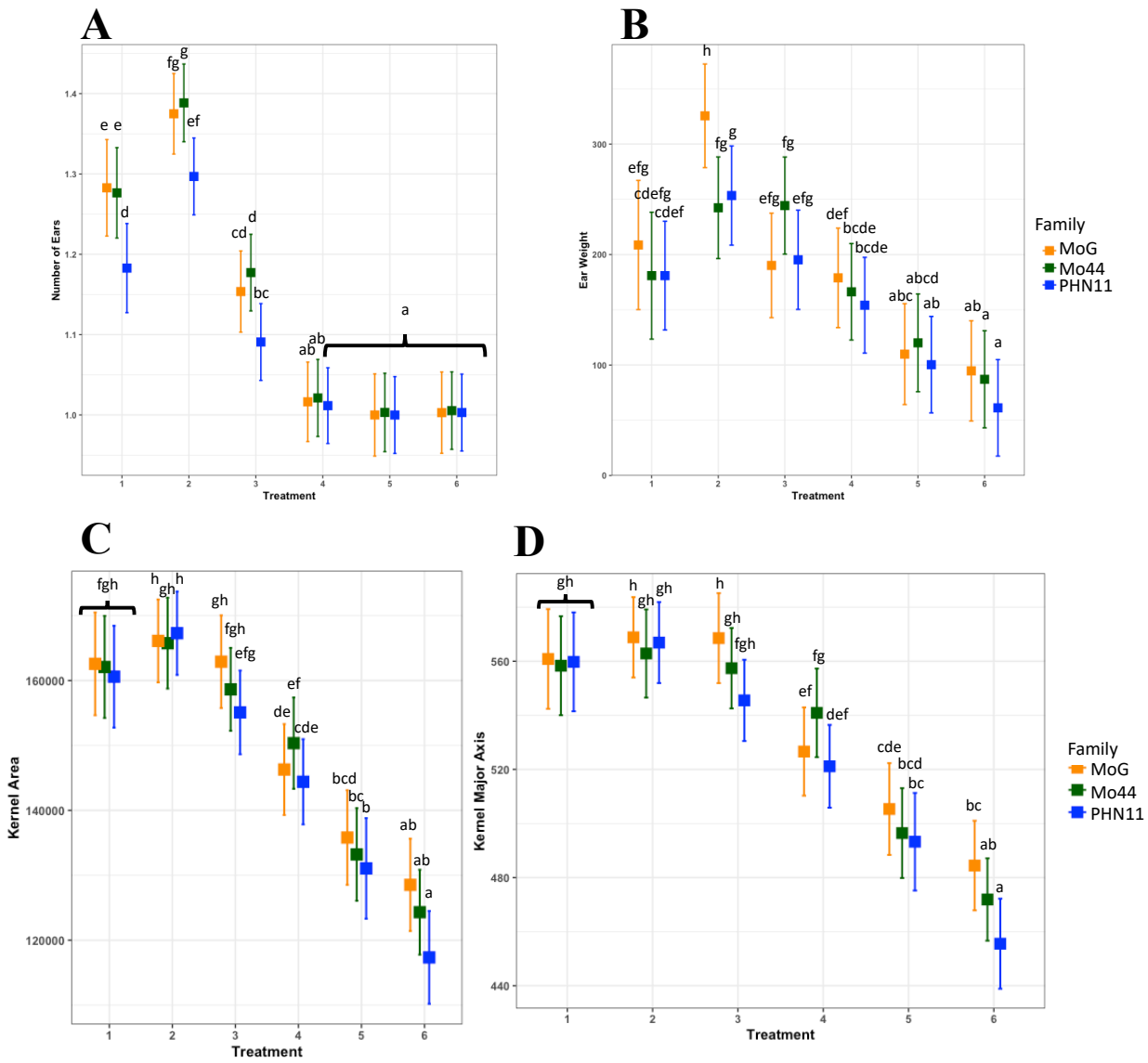


Figure 3.6 A:D. Breakdown of the four traits that have significant Treatment x Family interactions: number of ears (A), ear weight (B), kernel area (C), and kernel major axis (D). Boxes indicate the least square (LS) mean, error bars indicate the 95% confidence interval of the LS mean, and means sharing a letter within the same trait are not significantly different (Tukey-adjusted comparisons). Boxes are colored by Family (MoG is yellow, Mo44 is green, and PHN11 is blue).

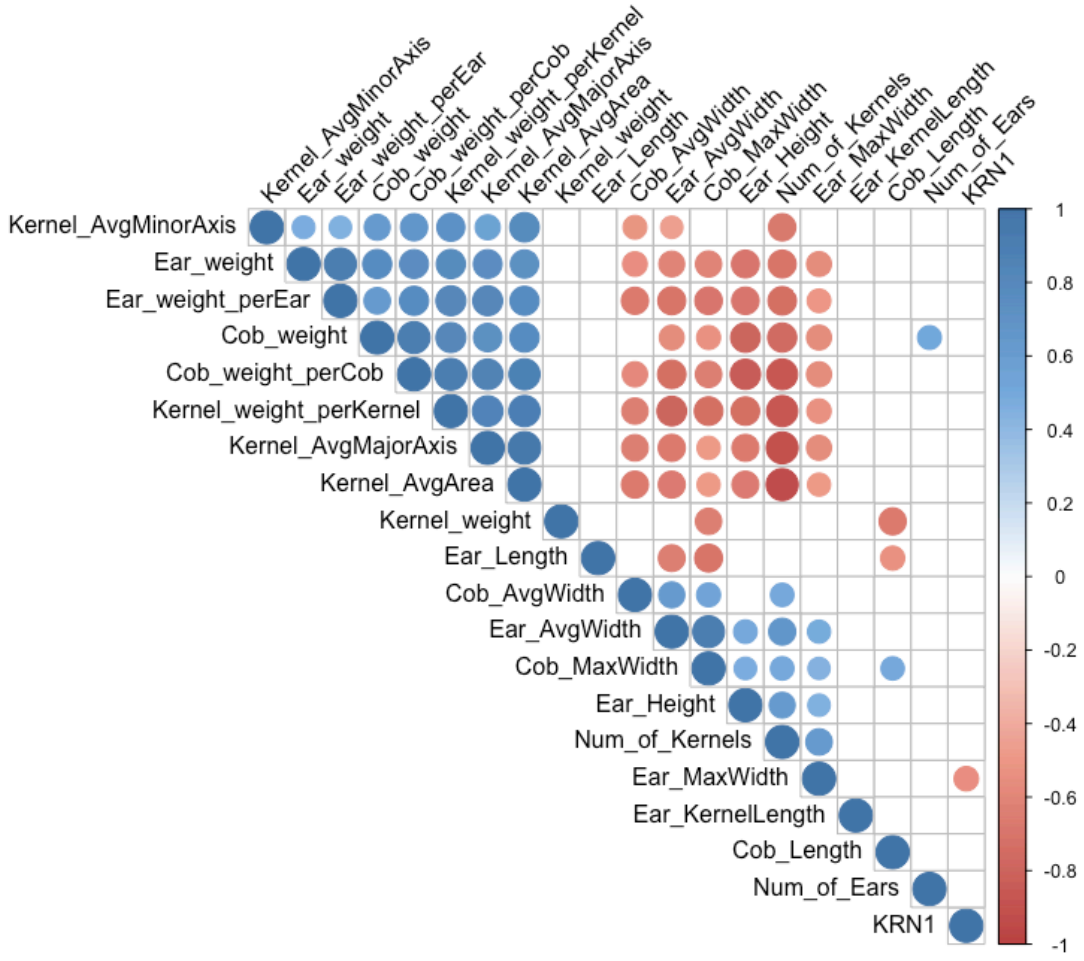


Figure 3.7. Pearson correlations across ear, cob, and kernel measurements. Correlations were calculated using α=0.05 significance level, magnitude ranges from -1 (red) to 1 (blue), and only significant p-values are shown.

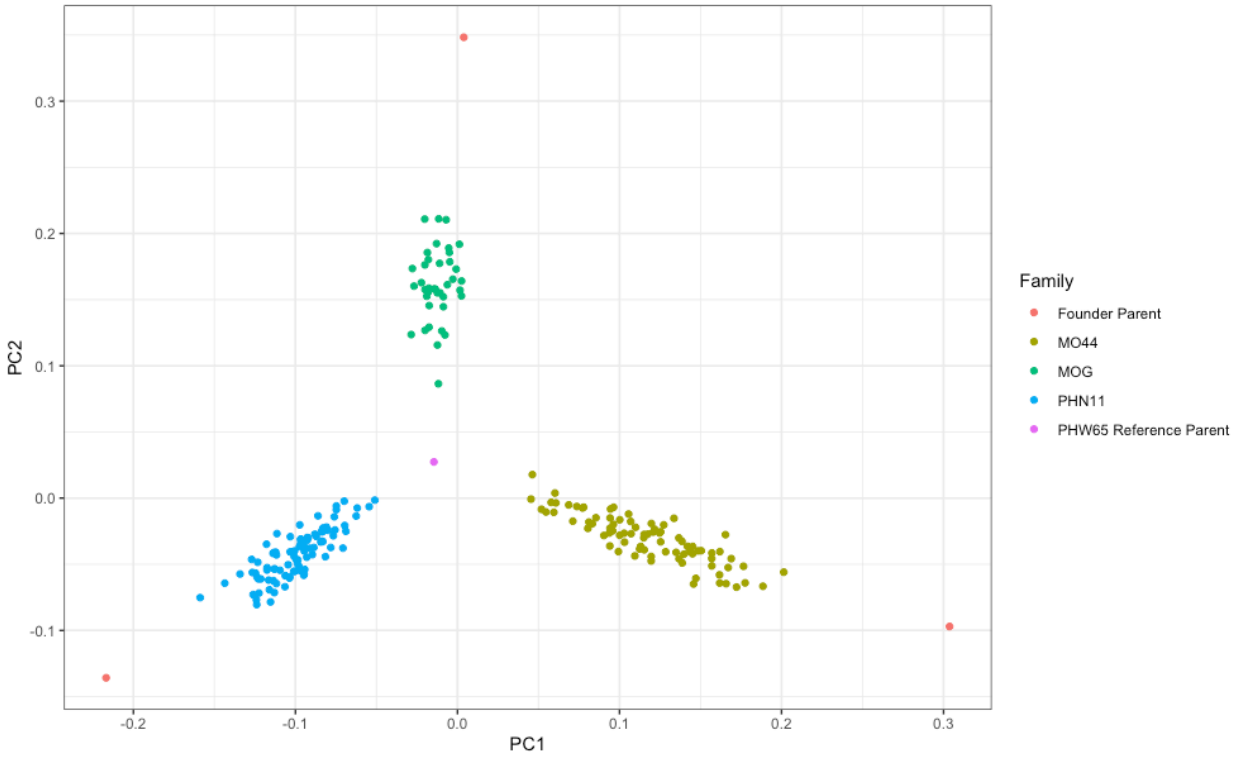


Figure 3.8. A Multi-Dimensional Scaling (MDS) plot of the 218 genotyped individuals using principal components 1 (PC1) and 2 (PC2) from a genome-wide identity by state (IBS) distance matrix.

Supplementary Table 3.1. Breakdown of 385 DH lines' assigned families, the number of total density observations (maximum of 24), whether they were grown in the 2018 and/or 2019 EID experiment (indicated by X), and corresponding genotyping code used in Chapter Three study.

Pedigree	Family	Density Observations	2018	2019	PHG Genotyping Code
MO44	Founder Parent				GE18N001
MO44_PHW65_0001/PHT69	MO44	16	X	X	
MO44_PHW65_0009/PHT69	MO44	21	X	X	GE18N119
MO44_PHW65_0011/PHT69	MO44	21	X	X	GE18N024
MO44_PHW65_0012/PHT69	MO44	17	X	X	
MO44_PHW65_0015/PHT69	MO44	15	X	X	GE18N120
MO44_PHW65_0017/PHT69	MO44	23	X	X	GE18N104
MO44_PHW65_0020/PHT69	MO44	21	X	X	GE18N019
MO44_PHW65_0025/PHT69	MO44	11	X	X	
MO44_PHW65_0027/PHT69	MO44	11	X		
MO44_PHW65_0029/PHT69	MO44	21	X	X	
MO44_PHW65_0031/PHT69	MO44	10	X		
MO44_PHW65_0033/PHT69	MO44	11	X		GE18N121
MO44_PHW65_0034/PHT69	MO44	6	X		
MO44_PHW65_0035/PHT69	MO44	24	X	X	
MO44_PHW65_0037/PHT69	MO44	12	X		GE18N061
MO44_PHW65_0042/PHT69	MO44	23	X	X	GE18N096
MO44_PHW65_0048/PHT69	MO44	24	X	X	GE18N125
MO44_PHW65_0050/PHT69	MO44	13	X	X	GE18N052
MO44_PHW65_0056/PHT69	MO44	10	X		
MO44_PHW65_0060/PHT69	MO44	12		X	
MO44_PHW65_0063/PHT69	MO44	22	X	X	GE18N028
MO44_PHW65_0067/PHT69	MO44	22	X	X	
MO44_PHW65_0068/PHT69	MO44	10	X		
MO44_PHW65_0069/PHT69	MO44	12	X	X	GE18N005
MO44_PHW65_0073/PHT69	MO44	16	X	X	
MO44_PHW65_0074/PHT69	MO44	16	X	X	GE18N037
MO44_PHW65_0083/PHT69	MO44	12	X		GE18N082
MO44_PHW65_0084/PHT69	MO44	23	X	X	GE18N053
MO44_PHW65_0092/PHT69	MO44	16	X	X	GE18N029
MO44_PHW65_0093/PHT69	MO44	20	X	X	GE18N030
MO44_PHW65_0098/PHT69	MO44	23	X	X	GE18N047

MO44_PHW65_0099/PHT69	MO44	12		X	GE18N038
MO44_PHW65_0103/PHT69	MO44	11	X		GE18N039
MO44_PHW65_0104/PHT69	MO44	11	X		
MO44_PHW65_0110/PHT69	MO44	22	X	X	GE18N015
MO44_PHW65_0111/PHT69	MO44	21	X	X	GE18N112
MO44_PHW65_0113/PHT69	MO44	21	X	X	
MO44_PHW65_0115/PHT69	MO44	10	X		GE18N072
MO44_PHW65_0116/PHT69	MO44	8	X		
MO44_PHW65_0118/PHT69	MO44	8	X		
MO44_PHW65_0120/PHT69	MO44	8	X		
MO44_PHW65_0128/PHT69	MO44	22	X	X	
MO44_PHW65_0134/PHT69	MO44	23	X	X	GE18N031
MO44_PHW65_0135/PHT69	MO44	18	X	X	GE18N105
MO44_PHW65_0136/PHT69	MO44	12	X		
MO44_PHW65_0139/PHT69	MO44	23	X	X	GE18N063
MO44_PHW65_0140/PHT69	MO44	23	X	X	GE18N106
MO44_PHW65_0144/PHT69	MO44	22	X	X	
MO44_PHW65_0145/PHT69	MO44	11	X		GE18N048
MO44_PHW65_0147/PHT69	MO44	10	X		GE18N073
MO44_PHW65_0150/PHT69	MO44	21	X	X	
MO44_PHW65_0151/PHT69	MO44	24	X	X	
MO44_PHW65_0154/PHT69	MO44	12		X	GE18N064
MO44_PHW65_0155/PHT69	MO44	22	X	X	GE18N107
MO44_PHW65_0157/PHT69	MO44	9	X		GE18N040
MO44_PHW65_0159/PHT69	MO44	22	X	X	
MO44_PHW65_0160/PHT69	MO44	23	X	X	GE18N011
MO44_PHW65_0163/PHT69	MO44	20	X	X	GE18N006
MO44_PHW65_0164/PHT69	MO44	10	X		
MO44_PHW65_0165/PHT69	MO44	22	X	X	
MO44_PHW65_0166/PHT69	MO44	9	X		GE18N041
MO44_PHW65_0167/PHT69	MO44	11	X		GE18N042
MO44_PHW65_0168/PHT69	MO44	20	X	X	GE18N113
MO44_PHW65_0171/PHT69	MO44	22	X	X	GE18N108
MO44_PHW65_0174/PHT69	MO44	23	X	X	GE18N097
MO44_PHW65_0175/PHT69	MO44	24	X	X	GE18N054
MO44_PHW65_0177/PHT69	MO44	23	X	X	GE18N109
MO44_PHW65_0179/PHT69	MO44	21	X	X	GE18N087
MO44_PHW65_0180/PHT69	MO44	18	X	X	GE18N043
MO44_PHW65_0185/PHT69	MO44	18	X	X	GE18N076

MO44_PHW65_0187/PHT69	MO44	11	X		
MO44_PHW65_0188/PHT69	MO44	20	X	X	GE18N035
MO44_PHW65_0189/PHT69	MO44	15	X	X	GE18N008
MO44_PHW65_0191/PHT69	MO44	20	X	X	GE18N009
MO44_PHW65_0193/PHT69	MO44	11	X		
MO44_PHW65_0195/PHT69	MO44	24	X	X	GE18N088
MO44_PHW65_0200/PHT69	MO44	22	X	X	GE18N099
MO44_PHW65_0207/PHT69	MO44	15	X	X	GE18N044
MO44_PHW65_0210/PHT69	MO44	7	X		GE18N114
MO44_PHW65_0214/PHT69	MO44	23	X	X	GE18N036
MO44_PHW65_0217/PHT69	MO44	15	X	X	GE18N077
MO44_PHW65_0218/PHT69	MO44	22	X	X	
MO44_PHW65_0220/PHT69	MO44	18	X	X	
MO44_PHW65_0223/PHT69	MO44	22	X	X	GE18N089
MO44_PHW65_0224/PHT69	MO44	10	X		GE18N115
MO44_PHW65_0225/PHT69	MO44	19	X	X	GE18N116
MO44_PHW65_0230/PHT69	MO44	11	X		GE18N078
MO44_PHW65_0236/PHT69	MO44	17	X	X	GE18N055
MO44_PHW65_0237/PHT69	MO44	22	X	X	GE18N059
MO44_PHW65_0247/PHT69	MO44	21	X	X	GE18N117
MO44_PHW65_0248/PHT69	MO44	23	X	X	GE18N101
MO44_PHW65_0250/PHT69	MO44	16	X	X	GE18N065
MO44_PHW65_0252/PHT69	MO44	15	X	X	
MO44_PHW65_0253/PHT69	MO44	22	X	X	
MO44_PHW65_0254/PHT69	MO44	21	X	X	GE18N066
MO44_PHW65_0259/PHT69	MO44	8	X		
MO44_PHW65_0260/PHT69	MO44	22	X	X	
MO44_PHW65_0263/PHT69	MO44	17	X	X	
MO44_PHW65_0271/PHT69	MO44	22	X	X	
MO44_PHW65_0288/PHT69	MO44	21	X	X	
MO44_PHW65_0302/PHT69	MO44	23	X	X	GE18N103
MO44_PHW65_0307/PHT69	MO44	18	X	X	GE18N056
MO44_PHW65_0316/PHT69	MO44	19	X	X	GE18N080
MO44_PHW65_0317/PHT69	MO44	21	X	X	GE18N045
MO44_PHW65_0319/PHT69	MO44	22	X	X	GE18N012
MO44_PHW65_0320/PHT69	MO44	18	X	X	GE18N081
MO44_PHW65_0321/PHT69	MO44	23	X	X	
MO44_PHW65_0324/PHT69	MO44	9	X		
MO44_PHW65_0328/PHT69	MO44	22	X	X	GE18N111

MO44_PHW65_0333/PHT69	MO44	22	X	X	
MO44_PHW65_0336/PHT69	MO44	22	X	X	
MO44_PHW65_0339/PHT69	MO44	23	X	X	GE18N007
MO44_PHW65_0343/PHT69	MO44	11	X		
MO44_PHW65_0349/PHT69	MO44	9	X		
MO44_PHW65_0350/PHT69	MO44	10	X		GE18N091
MO44_PHW65_0363/PHT69	MO44	20	X	X	GE18N033
MO44_PHW65_0376/PHT69	MO44	22	X	X	GE18N021
MO44_PHW65_0390/PHT69	MO44	18	X	X	GE18N057
MO44_PHW65_0391/PHT69	MO44	11	X		
MO44_PHW65_0392/PHT69	MO44	10	X		
MO44_PHW65_0396/PHT69	MO44	14	X	X	GE18N092
MO44_PHW65_0397/PHT69	MO44	19	X	X	GE18N018
MO44_PHW65_0404/PHT69	MO44	21	X	X	
MO44_PHW65_0416/PHT69	MO44	16	X	X	
MO44_PHW65_0433/PHT69	MO44	22	X	X	GE18N034
MO44_PHW65_0435/PHT69	MO44	21	X	X	GE18N046
MO44_PHW65_0442/PHT69	MO44	12		X	GE18N068
MO44_PHW65_0446/PHT69	MO44	22	X	X	GE18N022
MO44_PHW65_0455/PHT69	MO44	17	X	X	
MO44_PHW65_0464/PHT69	MO44	22	X	X	GE18N093
MO44_PHW65_0471/PHT69	MO44	22	X	X	
MO44_PHW65_0475/PHT69	MO44	9	X		
MOG	Founder Parent				GE18N437
PHN11	Founder Parent				GE18N433
PHN11_PHW65_0001/PHT69	PHN11	22	X	X	
PHN11_PHW65_0003/PHT69	PHN11	11	X		GE18N160
PHN11_PHW65_0010/PHT69	PHN11	10	X		GE18N132
PHN11_PHW65_0011/PHT69	PHN11	22	X	X	GE18N257
PHN11_PHW65_0015/PHT69	PHN11	22	X	X	
PHN11_PHW65_0016/PHT69	PHN11	11	X		GE18N188
PHN11_PHW65_0017/PHT69	PHN11	12	X		GE18N209
PHN11_PHW65_0021/PHT69	PHN11	6		X	GE18N138
PHN11_PHW65_0023/PHT69	PHN11	22	X	X	GE18N169
PHN11_PHW65_0027/PHT69	PHN11	12	X		
PHN11_PHW65_0028/PHT69	PHN11	10	X		
PHN11_PHW65_0035/PHT69	PHN11	9	X		
PHN11_PHW65_0042/PHT69	PHN11	10	X		GE18N143

PHN11_PHW65_0048/PHT69	PHN11	22	X	X	GE18N162
PHN11_PHW65_0049/PHT69	PHN11	9	X		GE18N298
PHN11_PHW65_0050/PHT69	PHN11	10	X		
PHN11_PHW65_0055/PHT69	PHN11	17	X	X	GE18N280
PHN11_PHW65_0056/PHT69	PHN11	24	X	X	GE18N170
PHN11_PHW65_0057/PHT69	PHN11	19	X	X	GE18N299
PHN11_PHW65_0059/PHT69	PHN11	9	X		GE18N319
PHN11_PHW65_0062/PHT69	PHN11	23	X	X	GE18N210
PHN11_PHW65_0063/PHT69	PHN11	23	X	X	GE18N286
PHN11_PHW65_0064/PHT69	PHN11	15	X	X	GE18N194
PHN11_PHW65_0065/PHT69	PHN11	21	X	X	GE18N295
PHN11_PHW65_0066/PHT69	PHN11	23	X	X	GE18N315
PHN11_PHW65_0069/PHT69	PHN11	20	X	X	GE18N259
PHN11_PHW65_0070/PHT69	PHN11	23	X	X	GE18N189
PHN11_PHW65_0071/PHT69	PHN11	22	X	X	GE18N178
PHN11_PHW65_0072/PHT69	PHN11	9	X		GE18N260
PHN11_PHW65_0073/PHT69	PHN11	23	X	X	GE18N246
PHN11_PHW65_0074/PHT69	PHN11	24	X	X	GE18N261
PHN11_PHW65_0076/PHT69	PHN11	11	X		GE18N310
PHN11_PHW65_0079/PHT69	PHN11	21	X	X	GE18N262
PHN11_PHW65_0080/PHT69	PHN11	19	X	X	GE18N195
PHN11_PHW65_0081/PHT69	PHN11	23	X	X	
PHN11_PHW65_0084/PHT69	PHN11	22	X	X	GE18N247
PHN11_PHW65_0086/PHT69	PHN11	10	X		GE18N263
PHN11_PHW65_0087/PHT69	PHN11	12	X	X	GE18N296
PHN11_PHW65_0089/PHT69	PHN11	17	X	X	
PHN11_PHW65_0092/PHT69	PHN11	22	X	X	GE18N225
PHN11_PHW65_0096/PHT69	PHN11	19	X	X	GE18N191
PHN11_PHW65_0100/PHT69	PHN11	22	X	X	GE18N164
PHN11_PHW65_0101/PHT69	PHN11	20	X	X	GE18N281
PHN11_PHW65_0105/PHT69	PHN11	21	X	X	GE18N264
PHN11_PHW65_0106/PHT69	PHN11	22	X	X	GE18N236
PHN11_PHW65_0107/PHT69	PHN11	17	X	X	GE18N171
PHN11_PHW65_0110/PHT69	PHN11	10	X		GE18N179
PHN11_PHW65_0113/PHT69	PHN11	23	X	X	GE18N311
PHN11_PHW65_0114/PHT69	PHN11	23	X	X	GE18N180
PHN11_PHW65_0117/PHT69	PHN11	21	X	X	GE18N150
PHN11_PHW65_0118/PHT69	PHN11	23	X	X	GE18N249
PHN11_PHW65_0119/PHT69	PHN11	22	X	X	GE18N250

PHN11_PHW65_0125/PHT69	PHN11	21	X	X	GE18N165
PHN11_PHW65_0129/PHT69	PHN11	21	X	X	
PHN11_PHW65_0132/PHT69	PHN11	9	X		GE18N196
PHN11_PHW65_0133/PHT69	PHN11	20	X	X	
PHN11_PHW65_0136/PHT69	PHN11	9	X		GE18N300
PHN11_PHW65_0137/PHT69	PHN11	21	X	X	GE18N322
PHN11_PHW65_0138/PHT69	PHN11	22	X	X	GE18N145
PHN11_PHW65_0139/PHT69	PHN11	18	X	X	
PHN11_PHW65_0140/PHT69	PHN11	19	X	X	GE18N301
PHN11_PHW65_0141/PHT69	PHN11	22	X	X	
PHN11_PHW65_0143/PHT69	PHN11	22	X	X	GE18N157
PHN11_PHW65_0144/PHT69	PHN11	19	X	X	GE18N151
PHN11_PHW65_0146/PHT69	PHN11	17	X	X	
PHN11_PHW65_0149/PHT69	PHN11	21	X	X	GE18N302
PHN11_PHW65_0154/PHT69	PHN11	20	X	X	
PHN11_PHW65_0155/PHT69	PHN11	20	X	X	
PHN11_PHW65_0158/PHT69	PHN11	22	X		GE18N237
PHN11_PHW65_0167/PHT69	PHN11	10	X		GE18N265
PHN11_PHW65_0168/PHT69	PHN11	24	X	X	
PHN11_PHW65_0172/PHT69	PHN11	21	X	X	GE18N182
PHN11_PHW65_0176/PHT69	PHN11	23	X	X	GE18N197
PHN11_PHW65_0178/PHT69	PHN11	17	X	X	GE18N167
PHN11_PHW65_0179/PHT69	PHN11	21	X	X	
PHN11_PHW65_0180/PHT69	PHN11	16	X	X	GE18N267
PHN11_PHW65_0184/PHT69	PHN11	22	X	X	GE18N226
PHN11_PHW65_0189/PHT69	PHN11	10	X		
PHN11_PHW65_0190/PHT69	PHN11	22	X	X	GE18N308
PHN11_PHW65_0194/PHT69	PHN11	14	X	X	
PHN11_PHW65_0205/PHT69	PHN11	18	X	X	
PHN11_PHW65_0206/PHT69	PHN11	17	X	X	
PHN11_PHW65_0209/PHT69	PHN11	15	X	X	GE18N268
PHN11_PHW65_0212/PHT69	PHN11	20	X	X	
PHN11_PHW65_0216/PHT69	PHN11	22	X	X	GE18N238
PHN11_PHW65_0217/PHT69	PHN11	12	X		GE18N303
PHN11_PHW65_0221/PHT69	PHN11	8	X		GE18N251
PHN11_PHW65_0223/PHT69	PHN11	10	X		
PHN11_PHW65_0224/PHT69	PHN11	19	X	X	
PHN11_PHW65_0227/PHT69	PHN11	21	X	X	
PHN11_PHW65_0237/PHT69	PHN11	22	X	X	GE18N313

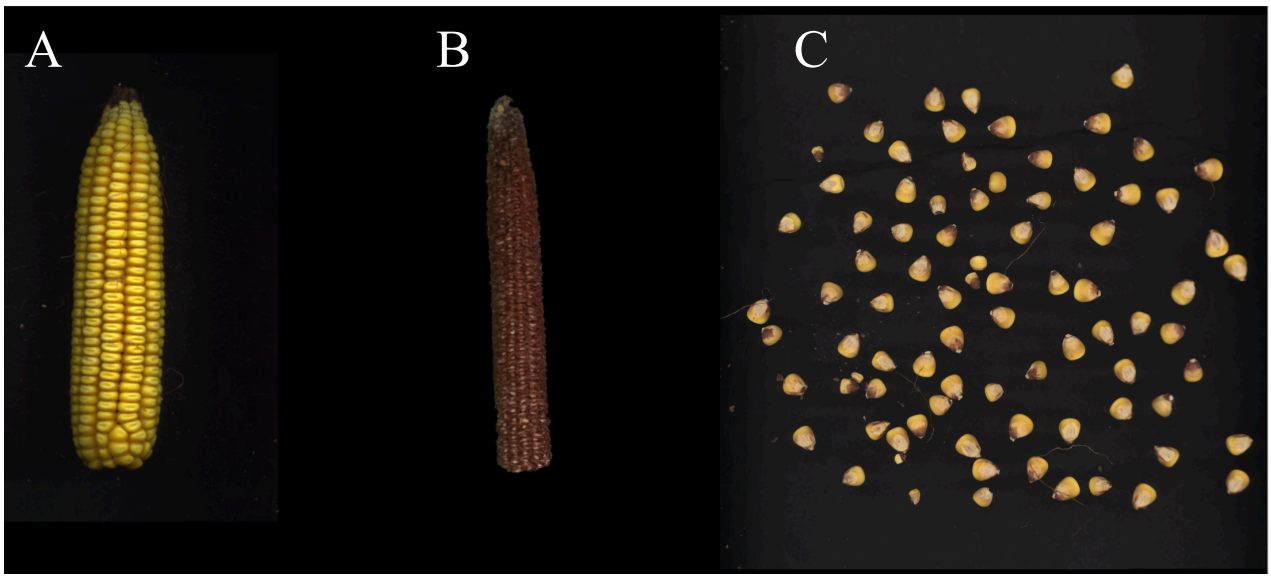
PHN11_PHW65_0241/PHT69	PHN11	20	X	X	
PHN11_PHW65_0242/PHT69	PHN11	12	X		GE18N239
PHN11_PHW65_0245/PHT69	PHN11	21	X	X	GE18N198
PHN11_PHW65_0251/PHT69	PHN11	10	X		GE18N252
PHN11_PHW65_0256/PHT69	PHN11	22	X	X	GE18N316
PHN11_PHW65_0257/PHT69	PHN11	23	X	X	GE18N215
PHN11_PHW65_0258/PHT69	PHN11	23	X	X	GE18N201
PHN11_PHW65_0266/PHT69	PHN11	15	X	X	GE18N217
PHN11_PHW65_0267/PHT69	PHN11	21	X	X	GE18N253
PHN11_PHW65_0274/PHT69	PHN11	15	X	X	GE18N270
PHN11_PHW65_0275/PHT69	PHN11	21	X	X	
PHN11_PHW65_0276/PHT69	PHN11	24	X	X	
PHN11_PHW65_0277/PHT69	PHN11	9	X		GE18N133
PHN11_PHW65_0279/PHT69	PHN11	23	X	X	
PHN11_PHW65_0282/PHT69	PHN11	21	X	X	GE18N287
PHN11_PHW65_0284/PHT69	PHN11	23	X	X	GE18N140
PHN11_PHW65_0285/PHT69	PHN11	16	X	X	GE18N202
PHN11_PHW65_0286/PHT69	PHN11	11	X		
PHN11_PHW65_0288/PHT69	PHN11	22	X	X	GE18N288
PHN11_PHW65_0291/PHT69	PHN11	23	X	X	GE18N174
PHN11_PHW65_0294/PHT69	PHN11	11	X		GE18N204
PHN11_PHW65_0300/PHT69	PHN11	22	X	X	
PHN11_PHW65_0301/PHT69	PHN11	24	X	X	
PHN11_PHW65_0302/PHT69	PHN11	8	X		GE18N240
PHN11_PHW65_0303/PHT69	PHN11	21	X	X	
PHN11_PHW65_0305/PHT69	PHN11	16	X	X	GE18N152
PHN11_PHW65_0336/PHT69	PHN11	22	X	X	GE18N218
PHN11_PHW65_0344/PHT69	PHN11	15	X	X	GE18N317
PHN11_PHW65_0352/PHT69	PHN11	22	X	X	
PHN11_PHW65_0354/PHT69	PHN11	22	X	X	GE18N219
PHN11_PHW65_0359/PHT69	PHN11	23	X	X	GE18N231
PHN11_PHW65_0369/PHT69	PHN11	21	X	X	
PHN11_PHW65_0393/PHT69	PHN11	22	X	X	GE18N309
PHN11_PHW65_0398/PHT69	PHN11	21	X	X	
PHN11_PHW65_0424/PHT69	PHN11	16	X	X	GE18N274
PHN11_PHW65_0432/PHT69	PHN11	8	X		GE18N187
PHN11_PHW65_0464/PHT69	PHN11	8	X		GE18N154
PHN11_PHW65_0496/PHT69	PHN11	23	X	X	GE18N284
PHN11_PHW65_0637/PHT69	PHN11	22	X	X	

PHN11_PHW65_0639/PHT69	PHN11	22	X	X	
PHW65	Reference Parent				GE18N436
PHW65_MOG_0001/PHT69	MOG	24	X	X	GE18N420
PHW65_MOG_0012/PHT69	MOG	22	X	X	GE18N366
PHW65_MOG_0015/PHT69	MOG	20	X	X	
PHW65_MOG_0017/PHT69	MOG	23	X	X	GE18N389
PHW65_MOG_0024/PHT69	MOG	9	X		GE18N331
PHW65_MOG_0036/PHT69	MOG	22	X	X	GE18N367
PHW65_MOG_0038/PHT69	MOG	22	X	X	
PHW65_MOG_0040/PHT69	MOG	10	X		
PHW65_MOG_0041/PHT69	MOG	23	X	X	GE18N390
PHW65_MOG_0048/PHT69	MOG	20	X	X	
PHW65_MOG_0051/PHT69	MOG	19	X	X	
PHW65_MOG_0052/PHT69	MOG	23	X	X	
PHW65_MOG_0054/PHT69	MOG	22	X	X	
PHW65_MOG_0055/PHT69	MOG	24	X	X	
PHW65_MOG_0057/PHT69	MOG	23	X	X	GE18N345
PHW65_MOG_0060/PHT69	MOG	24	X	X	
PHW65_MOG_0061/PHT69	MOG	22	X	X	
PHW65_MOG_0062/PHT69	MOG	14	X	X	
PHW65_MOG_0067/PHT69	MOG	12		X	
PHW65_MOG_0068/PHT69	MOG	12		X	
PHW65_MOG_0071/PHT69	MOG	12		X	
PHW65_MOG_0075/PHT69	MOG	7	X		
PHW65_MOG_0082/PHT69	MOG	12		X	GE18N398
PHW65_MOG_0087/PHT69	MOG	23	X	X	
PHW65_MOG_0089/PHT69	MOG	22	X	X	
PHW65_MOG_0096/PHT69	MOG	12		X	GE18N425
PHW65_MOG_0106/PHT69	MOG	12		X	
PHW65_MOG_0116/PHT69	MOG	22	X	X	
PHW65_MOG_0117/PHT69	MOG	21	X	X	
PHW65_MOG_0118/PHT69	MOG	20	X	X	
PHW65_MOG_0146/PHT69	MOG	10	X		
PHW65_MOG_0160/PHT69	MOG	11	X		
PHW65_MOG_0162/PHT69	MOG	6	X		
PHW65_MOG_0173/PHT69	MOG	19	X	X	
PHW65_MOG_0175/PHT69	MOG	21	X	X	
PHW65_MOG_0179/PHT69	MOG	12		X	GE18N411

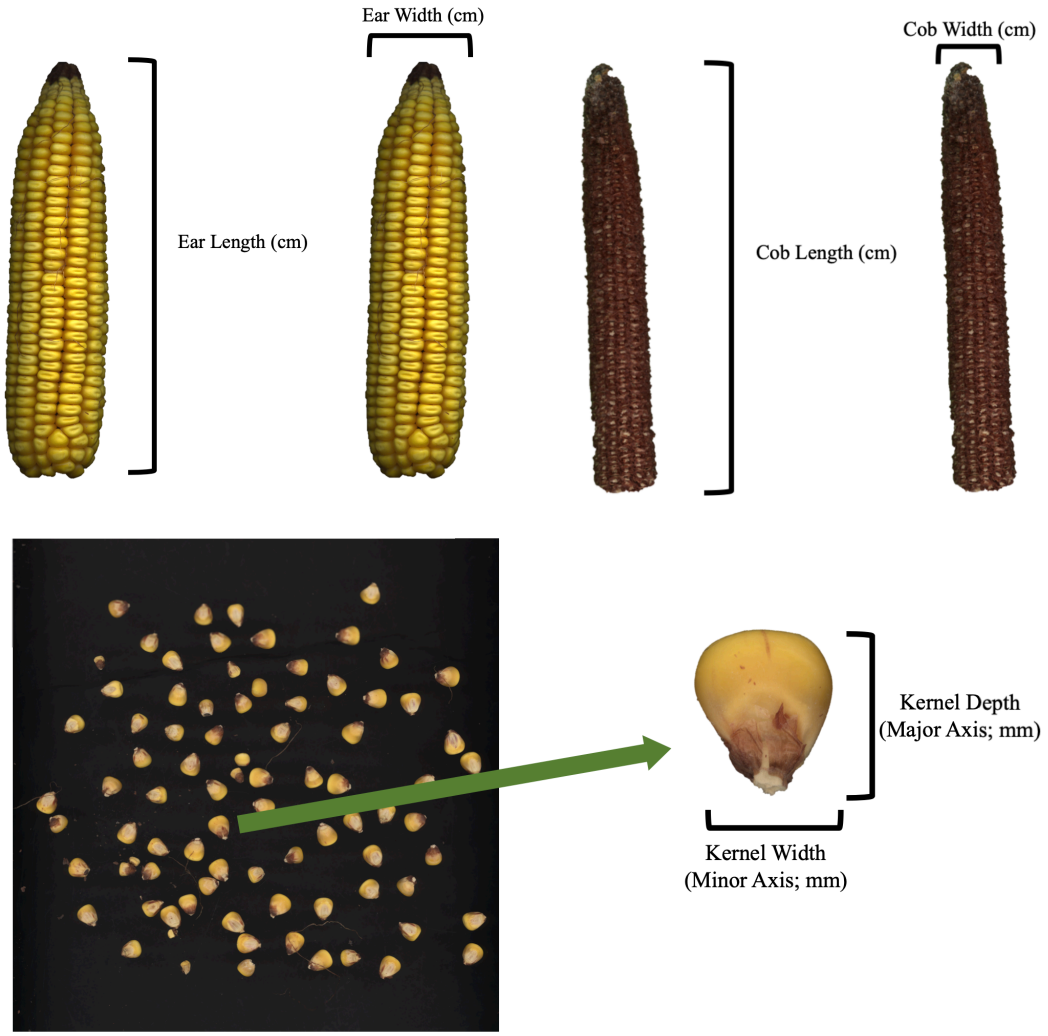
PHW65_MOG_0190/PHT69	MOG	21	X	X	
PHW65_MOG_0196/PHT69	MOG	12		X	GE18N401
PHW65_MOG_0199/PHT69	MOG	12		X	
PHW65_MOG_0200/PHT69	MOG	11	X		
PHW65_MOG_0203/PHT69	MOG	12		X	
PHW65_MOG_0206/PHT69	MOG	23	X	X	
PHW65_MOG_0209/PHT69	MOG	6		X	
PHW65_MOG_0215/PHT69	MOG	22	X	X	GE18N382
PHW65_MOG_0217/PHT69	MOG	22	X	X	GE18N355
PHW65_MOG_0218/PHT69	MOG	12		X	
PHW65_MOG_0221/PHT69	MOG	6		X	
PHW65_MOG_0224/PHT69	MOG	12		X	GE18N393
PHW65_MOG_0232/PHT69	MOG	12		X	
PHW65_MOG_0250/PHT69	MOG	6		X	
PHW65_MOG_0253/PHT69	MOG	12		X	
PHW65_MOG_0262/PHT69	MOG	23		X	GE18N418
PHW65_MOG_0266/PHT69	MOG	12		X	GE18N383
PHW65_MOG_0291/PHT69	MOG	12		X	
PHW65_MOG_0293/PHT69	MOG	23	X	X	
PHW65_MOG_0294/PHT69	MOG	7	X		GE18N336
PHW65_MOG_0308/PHT69	MOG	21	X	X	GE18N338
PHW65_MOG_0310/PHT69	MOG	11	X		
PHW65_MOG_0311/PHT69	MOG	21	X	X	
PHW65_MOG_0314/PHT69	MOG	24	X	X	
PHW65_MOG_0315/PHT69	MOG	20	X	X	
PHW65_MOG_0316/PHT69	MOG	22	X	X	
PHW65_MOG_0326/PHT69	MOG	20	X	X	
PHW65_MOG_0327/PHT69	MOG	23	X	X	GE18N430
PHW65_MOG_0329/PHT69	MOG	16	X	X	GE18N356
PHW65_MOG_0332/PHT69	MOG	23	X	X	
PHW65_MOG_0335/PHT69	MOG	21	X	X	
PHW65_MOG_0340/PHT69	MOG	17	X	X	
PHW65_MOG_0341/PHT69	MOG	22	X	X	
PHW65_MOG_0343/PHT69	MOG	23	X	X	
PHW65_MOG_0352/PHT69	MOG	9	X		
PHW65_MOG_0354/PHT69	MOG	21	X	X	
PHW65_MOG_0375/PHT69	MOG	17	X	X	
PHW65_MOG_0377/PHT69	MOG	17	X	X	GE18N369
PHW65_MOG_0379/PHT69	MOG	10	X		

PHW65_MOG_0381/PHT69	MOG	16	X	X	
PHW65_MOG_0392/PHT69	MOG	23	X	X	
PHW65_MOG_0401/PHT69	MOG	16	X	X	GE18N364
PHW65_MOG_0410/PHT69	MOG	11	X		
PHW65_MOG_0411/PHT69	MOG	23	X	X	GE18N365
PHW65_MOG_0421/PHT69	MOG	24	X	X	GE18N357
PHW65_MOG_0423/PHT69	MOG	24	X	X	
PHW65_MOG_0424/PHT69	MOG	23	X	X	
PHW65_MOG_0426/PHT69	MOG	21	X	X	
PHW65_MOG_0429/PHT69	MOG	20	X	X	GE18N408
PHW65_MOG_0431/PHT69	MOG	22	X	X	GE18N419
PHW65_MOG_0436/PHT69	MOG	20	X	X	GE18N431
PHW65_MOG_0438/PHT69	MOG	9	X		
PHW65_MOG_0439/PHT69	MOG	10	X		
PHW65_MOG_0442/PHT69	MOG	10	X		GE18N405
PHW65_MOG_0444/PHT69	MOG	16	X	X	
PHW65_MOG_0449/PHT69	MOG	15	X	X	
PHW65_MOG_0460/PHT69	MOG	8	X		GE18N384
PHW65_MOG_0461/PHT69	MOG	23	X	X	
PHW65_MOG_0463/PHT69	MOG	6	X		
PHW65_MOG_0467/PHT69	MOG	9	X		
PHW65_MOG_0476/PHT69	MOG	22	X	X	
PHW65_MOG_0487/PHT69	MOG	20	X	X	GE18N385
PHW65_MOG_0499/PHT69	MOG	22	X	X	GE18N424
PHW65_MOG_0511/PHT69	MOG	20	X	X	
PHW65_MOG_0512/PHT69	MOG	21	X	X	
PHW65_MOG_0513/PHT69	MOG	14	X	X	
PHW65_MOG_0522/PHT69	MOG	14	X	X	
PHW65_MOG_0534/PHT69	MOG	10	X		
PHW65_MOG_0539/PHT69	MOG	16	X	X	GE18N386
PHW65_MOG_0542/PHT69	MOG	23	X	X	GE18N427
PHW65_MOG_0545/PHT69	MOG	23	X	X	
PHW65_MOG_0551/PHT69	MOG	24	X	X	
PHW65_MOG_0554/PHT69	MOG	20	X	X	
PHW65_MOG_0567/PHT69	MOG	6	X		GE18N417
PHW65_MOG_0569/PHT69	MOG	20	X	X	GE18N370
PHW65_MOG_0572/PHT69	MOG	21	X	X	GE18N340
PHW65_MOG_0575/PHT69	MOG	10	X		
PHW65_MOG_0588/PHT69	MOG	21	X	X	GE18N395

PHW65_MOG_0591/PHT69	MOG	21	X	X	
PHW65_MOG_0620/PHT69	MOG	10	X		
PHW65_MOG_0621/PHT69	MOG	9	X		GE18N351
PHW65_MOG_0622/PHT69	MOG	23	X	X	GE18N397
PHW65_MOG_0633/PHT69	MOG	21	X	X	
PHW65_MOG_0660/PHT69	MOG	19	X	X	
PHW65_MOG_0675/PHT69	MOG	10	X		
PHW65_MOG_0696/PHT69	MOG	21	X	X	



Supplementary Figure 3.1. The scanned images of ear (A), cob (B), and kernels (C) from the genotype PHN11_PHW65_0286/PHT69 at planting density treatment 6.



Supplementary Figure 3.2. Measurements extracted from processed ear, cob, and kernel images using MATLAB custom algorithms described in Miller et al., 2017. The direct measurements include ear length (cm), maximum and average ear width (cm), cob length (cm), maximum and average cob width (cm), and kernel width (mm) and depth (mm). Sample images are from the genotype PHN11_PHW65_0286/PHT69 at planting density treatment 6.