GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN IN PUBLICA COMMODA SEIT 1737

Evaluating a Collection of Agronomic Traits to Better Understand the Importance of Epistasis and Epistasis by Environment Interaction in Maize.

Dissertation

to attain the doctoral degree Dr. sc. agr.

of the Faculty of Agriculture Sciences,

Georg-August-Universität Göttingen

Submitted by

Barış Alaca

born on the 01.01.1992 in Çanakkale Türkiye

Göttingen, (16.01.2024)

First Referee: Prof. Dr. Stefan Scholten

Second Referee: Prof. Dr. Timothy Mathes Beissinger

Third Referee: Prof. Dr. Reimund P. Rötter

Date of oral examination: 08.03.2024

Table of Contents

Summary **6**

Chapter 1: General Introduction **9**

Background information 9

Infinitesimal model and linkage disequilibrium (LD) 11

Genetic markers and the importance of single nucleotide polymorphisms (SNPs) 12

Importance of G2F 12

Importance of environment and QTL-by-environment interactions 14

Epistasis 15

Epistasis-by-environment interactions 18

Epigenetics and epigenetic inheritance 19

Importance of gene expression and origin-of-seed effect 23

Aim of thesis 25

Summary of material and method 25

Chapter 2: Identification of Significant Quantitative Trait Loci (QTL) in Diverse Mega Environments (MEs) and Interactions Between Epistasis by Environment in Maize **26**

Abstract **26**

Material & Methods **27**

Population 27

Mega environments (MEs) 28

Genotyping 29

Principle Component Analysis (PCA) 30

Filtering, marker trimming, and linkage mapping 30

Best linear unbiased estimation (BLUEs) 32

Broad sense heritability 41

QTL mapping 41
Epistasis QTL mapping: 42
Epistasis by environment interactions: 43
Results 46
Linkage mapping 46
Phenotype 53
Correlations 54
Heritability 55
QTL mapping 56
Epistasis QTL mapping 67
Epistasis by environment interactions 70
Discussion 80
Correlations & heritability 80
Linkage and QTL mapping 81
QTL environment interactions 84
Epistasis 84
Epistasis by environment interactions 85
Chapter 3: Gene Expression Analysis to Detect the Effect of the Origin of Seeds for an Epistasis Mapping Population Experiment in Maize 87
Abstract 87
Material & Method 88
Population and growth condition 88
RNA isolation, gene expression analysis 89
Quality control, RNA alignment, data collection 90
Outlier, principle component analysis (PCA), and statistical analysis 92
Results 93
Cook's distances and PCA 93
DEG analysis 94

Discussion 99

Epistasis mapping population (EMP) 99

Origin of seed 99

Chapter 4: General Discussion 102

Importance of epistasis and previous findings 102

QTL by environment and epistasis, and epistasis by environment interactions 103

Epistasis mapping population (EMP) and differential expression genes (DEG) 104

General conclusion 107

References 110

Appendix 129

Acknowledgments 129

Curriculum Vitae 130

Summary

Maize, a key cereal crop, stands alongside wheat and rice in global importance. Its domestication from teosinte in southern Mexico 10,000 years ago marked a turning point in human agriculture. Over the centuries, maize has spread globally, becoming a staple crop vital for food, feed, and biofuel. Modern plant breeding techniques and improved genetic backgrounds have propelled maize production, making it adaptable to diverse environments and significantly contributing to global food security.

The increase in agronomic traits owes to multiple factors, including advancements in plant genetics, agronomic practices, and technological innovations. Agronomic practices, precision agriculture, and remote sensing contribute to higher yields, underscoring the multidimensional approach to crop improvement. Molecular markers and genomic selection enable breeders to identify genes controlling crucial agronomic traits like yield, flowering time, and plant height. Genetic markers, especially single nucleotide polymorphisms (SNPs), have been pivotal in developing improved maize varieties.

The integration of sequencing technology has revolutionized plant breeding, making genetic information more accessible. The decreasing cost of sequencing has facilitated genome exploration, enabling the identification of genes controlling agronomic traits.

Furthermore, researchers have identified that additive genetic variation predominantly contributes to the variation within a population. However, it is acknowledged that additive variation does not fully account for heritability, leaving a portion known as missing heritability. Consequently, an increasing number of studies are directing their focus toward understanding and incorporating this missing heritability to enhance the prediction of quantitative agronomic traits. Epistasis, representing complex interactions between genes, is one approach employed to calculate this missing heritability. However, detecting epistatic interactions poses challenges, notably the multiple testing problem. Additionally, environmental factors can modify gene effects, complicating their detection. To address these complexities and identify epistatic interactions while considering the impact of the environment on genotypes, a unique population must be tested across various environments.

The initial study aimed to identify epistasis by environment interactions using genomes-to-field (G2F) data tested in multiple environments. In the second study, we aimed to investigate origin-of-seed effects utilizing an Epistasis Mapping Population (EMP).

Epistasis by environment (EEI) study delves into the complicated relationships between various agronomic traits in maize, exploring correlations, heritability, and quantitative trait loci (QTL) mapping across multiple environments to find the interactions between QTLs, QTLs by environment, and epistasis by environment interactions. The research investigated the following quantitative traits: pollination and silking days, plant and ear height, stand percentage, and yield. The initial analysis reveals strong correlations between pollen and silk days after pollination (DAP), as well as between plant height (PH) and ear height (EH). Moderate correlations exist among PH, EH, pollen DAP, and silk DAP, indicative of the intricate network of influences among these quantitative traits. Interestingly, the stand percentage exhibits a low correlation with other traits across locations, suggesting its sensitivity to external factors during the growing period.

The subsequent QTL mapping analysis uncovers significant loci associated with the studied traits across environments. These findings align with previous research, emphasizing the genetic complexity of traits. Notably, the study identifies specific QTLs for each trait in different locations, underlining the substantial impact of environmental factors on trait expression. Our analysis extends to QTL-environment interactions, revealing location-specific markers and significant interactions for pollen DAP, silk DAP, PH, EH, and yield. This underscores the importance of considering the environment when interpreting genetic influences on these traits.

Moreover, we identified significant epistatic interactions for each phenotypic trait. The following analysis introduces the concept of epistasis by environment interactions (EEI), illustrating pollen and silk DAP, PH, EH, and yield. EEI study comprehensively illustrates the genetic and environmental factors influencing agronomic traits in maize. Our research contributes valuable insights that can inform future breeding strategies and enhance our understanding of maize phenotypic traits' complex genetic and environmental interactions.

In the second study, we utilized two diverse seed sources of Epistasis Mapping Populations (EMPs), which aim to identify epistatic interactions between genomic regions, to investigate origin-of-seed effects on mRNA levels. By comparing phenotypic performances, EMP helps identify regions of the genome where epistasis occurs, contributing valuable insights into gene

regulation and species evolution. EMPs reduce the multiple-testing and provide a nuanced understanding of gene interactions. The study focused on the influence of origin-of-seed on maize seedlings at the V2 growth stage with EMPs. While complete sets of EMPs are traditionally required for studying epistatic interactions, the study uses a subset of genotypes in this part of the research question.

Our study found no differentially expressed genes (DEGs) between two seed sources at the V2 growing stage. Despite limitations, such as a small sample size and focusing solely on mRNA expression, the study suggests that greenhouse and field environmental conditions do not significantly impact the progeny's gene expression at the V2 growing stage. In addition, robust comparisons between different genotypes validate the reliability of the analysis, reinforcing the conclusion that origin-of-seed effects do not lead to significant expression. Overall, this research contributes valuable insights into the origin-of-seed effect on maize seedlings, pointing out a direction for further investigations.

Chapter 1: General Introduction

Background information

Throughout human history, the domestication and cultivation of plant species have been essential activities to ensure the survival and well-being of communities (Clement, 1999; Doebley et al., 2006; Janick, 2010). Improving the performance of such crops contributed to the gradual increase in the human population (Diepenbrock and Gore, 2015). A recent, major phase of agronomy as a driver in feeding the world's population was the so-called Green Revolution, led by Norman Borlaug, also widely considered the father of modern agriculture (Evenson and Gollin, 2003; Hobbs, 2007). We can describe the Green Revolution as a combination of several accomplishments that changed agriculture drastically: 1) agricultural mechanization, 2) availability of mineral nitrogen fertilizers, and 3) Borlaug's modern wheat and rice cultivars with short plant height and higher nitrogen usage efficiency (NUE), which utilized higher N doses more efficiently and thus provided significantly higher grain yield. This breakthrough set the stage for continued advancements in plant breeding techniques, leading to steady crop yields and food production improvement (Paddock, 1970; Rosset et al., 2000; Hedden, 2003; Patel, 2013). Today, plant breeding continues to play a critical role in modern agriculture, with scientists and breeders working together to develop improved varieties of crops that are better equipped to face the challenges posed by pests and diseases, are better adapted to changing climate conditions and hence have higher yields (Ronald, 2011; Begna, 2021). As a result, food and feed production has dramatically increased in recent decades, providing adequate nutrition for the world's growing population. The continued improvement in agriculture and food production is crucial in meeting the food needs of the growing human population (Borlaug, 1983; Serageldin, 2001; Kc et al., 2018).

Maize (corn, *Zea mays* L.) is one of the three most important cereal crops globally, alongside wheat (*Triticum aestifum* L.) and rice (*Oryza sativa* L.), in terms of production and consumption (Awika, 2011; Ranum et al., 2014; Macauley and Ramadjita, 2015). Maize is a staple crop that has played a significant role in human history and is an essential source of food, feed, and biofuel (Shiferaw et al., 2011; Datta et al., 2019; García-Lara and Serna-Saldivar, 2019). Maize was first domesticated in southern Mexico about 10,000 years ago from its wild ancestor, teosinte, and

introduced to Europe during the 15th century, quickly spread to the Old World (Pounds et al., 1979; Tenaillon and Charcosset, 2011). Since then, maize has become a staple crop in many countries and is widely cultivated for human consumption and animal feed. With the development of plant breeding techniques and the improvement of its genetic background, maize production has increased rapidly (Duvick, 1996; Ortiz-Monasterio et al., 2007). Today, maize is a prominent crop because it is versatile, adaptable to various environments, and can provide a high yield around the globe (Thornton et al., 2009, 2010). The rapid yield increase in the production of maize has made it a critical food source for humans and animals in many countries. Also, maize has become a hard-to-forsake feed ingredient for livestock and poultry (Palacios-Rojas et al., 2020; Grote et al., 2021; Wilson et al., 2021).

The increase in yield can be attributed to several other factors. One of these factors is the advancement in plant genetics, allowing plant breeders to identify desirable traits and incorporate them into hybrid cultivars (Hallauer et al., 2010; Breseghello and Coelho, 2013). For example, molecular markers and genomic selection have enabled the identification of genes controlling agronomic traits, such as drought tolerance, pest resistance, and earliness, allowing for the development of cultivars better adapted to specific environmental conditions (Agbicodo et al., 2009; Randhawa et al., 2013). Improved agronomic practices have also contributed to the increase in maize yields. Adopting proper sowing techniques, using high-quality seeds, and implementing practical pest and disease management strategies have further improved crop yields (Coakley et al., 1999; Amanullah and Khalid, 2020). Furthermore, modern technologies, such as precision agriculture and remote sensing, have allowed farmers to monitor crop growth and make more informed decisions regarding crop management, leading to even higher yields (Brisco et al., 1998; Robert, 2002; Liaghat and Balasundram, 2010). In short, the increase in maize yield results from several factors, including advances in plant genetics, plant nutrition, and agronomic practices (Laidig et al., 2014). By utilizing these various techniques and technologies, farmers and plant breeders have improved the performance of maize and increased its yield, providing a critical source of food, feed, and biofuel for the world's growing population.

Infinitesimal model and linkage disequilibrium (LD)

In the early 1900s, there were two predominant views on inheritance, the single-locus inheritance theory put forth by Mendel and the blending mechanism hypothesis proposed by the Biometricians of that time (Wright, 1984; Gluckman et al., 2016). These two hypotheses were gradually integrated over time, with Hardy-Weinberg introducing a formula in 1908 that showed how allele frequencies remain constant across generations in a closed population under the assumption that there is no mutation, selection, genetic drift, or random mating (Wittke-Thompson et al., 2005; Edwards, 2008). Fisher, regarded as the father of quantitative genetics, combined Mendel's single-locus inheritance theory with the idea that continuous traits are influenced by many genes with small effects in his "infinitesimal model," introduced in 1918 (Fisher, 1930; Toschi, 1960; Orr, 2005). In his model, Fisher explained that the visible variation of continuous traits observed in populations results from a complex interplay between genetic factors, environmental factors, and their interaction (G x E). Fisher's work marked a significant milestone in genetics, as it provided a framework for understanding that complex interplay between genetic and environmental factors that influence plant performance (Zondervan and Cardon, 2004; Atkinson and Urwin, 2012). This knowledge has enabled plant breeders to develop more effective breeding strategies, incorporating genetic and environmental factors into the breeding process. Today, the principles of single-locus inheritance, Hardy-Weinberg's formula, and Fisher's infinitesimal model form the foundation of modern genetics and plant breeding, providing the scientific community with a deep understanding of the genetic mechanisms that drive heritable changes in plant performance (Hill, 2010). Mendel's findings relied on absence of linkage between his monogenic traits under study, but it is agreed upon that genes on the same chromosome do not segregate independently (Ardlie et al., 2002; Flint-Garcia et al., 2003). LD became the foundation for genetic mapping, as genes that are physically closer together on a chromosome are more likely to be co-inherited into the next generation than otherwise (Kruglyak, 1997; Altshuler et al., 2008). The segregation ratio indicates the distance between two genes, expressed in Morgan units, with physically closer genes showing less likely recombination in the offspring than distant genes or genes on different chromosomes (Olson et al., 1999; Rafalski, 2002; Terwilliger and Göring, 2009).

Genetic markers and the importance of single nucleotide polymorphisms (SNPs)

There are several types of genetic markers, including single nucleotide polymorphisms (SNPs) (Chanock, 2001; Syvänen, 2001; Nasu, 2002), insertion/deletion polymorphisms (INDELs) (McCauley, 1995), microsatellites, and restriction fragment length polymorphisms (RFLPs) (Tanksley et al., 1989), each with its advantages and limitations for specific applications (Dodgson et al., 1997; Liu and Cordes, 2004; Kumar et al., 2009). SNP markers are the most common and applied markers by researchers and plant breeders (Poland and Rife, 2012; Mammadov et al., 2012; Nadeem et al., 2018). The application of genetic markers has become more widespread and accessible in the 21st century due to the rapid advancement of sequencing technology (Collard and Mackill, 2007; Kim et al., 2016). The decreasing cost of sequencing has allowed plant breeders to access the genomes of various plant species, providing valuable information that can be used to improve varieties of crop performance (Muir et al., 2016; Mardis, 2017). This information can be used to identify genes controlling important agronomic traits, such as drought tolerance, pest resistance, and yield, and then incorporate those desirable traits into cultivars (Forster et al., 2004; Jena and Mackill, 2008; Pandey et al., 2014). Moreover, sequencing technology has allowed the development of more accurate and efficient methods in plant breeding. For example, genomic selection has enabled plant breeders to predict which plant should be agronomically best based on a joint exploitation of genomic information and phenotypic information rather than based on phenotypic measurements alone (Heffner et al., 2009; Crossa et al., 2017). This results in shorter crop varieties' improvement and more rapid and efficient breeder work (Carvalho, 2006, 2017). In addition, sequencing technology has also led to the discovery of genes and pathways involved in plant growth, providing new insights into the underlying mechanisms that regulate agronomic traits. Additionally, new approaches such as CRISPR-Cas promise precise genetic modifications, which allow genetic modifications that can be used to improve plant performance for human needs (Chen et al., 2019).

Importance of G2F

Phenotypic observations of several genotypes in diverse environments must be obtained to assess any type of genotype by environment interactions (GxE or GEI), including epistasis by environment interaction (EEI). However, applying strictly common rules, definitions, and units

when phenotyping in different environments is challenging and requires a strict, agreed-on procedure to observe phenotypic traits. If analyses of GEI of quantitative traits should be performed with a focus on the single-loci contributions to such traits, genotyping and subsequent QTL analyses is a way to go. Both, phenotyping in diverse environments and genotyping of genotypes are challenging and costly. The “Genomes to field” (G2F) initiative created a maize population, and the collaborators of phenotyping across locations observed traits across the different locations (**Figure 1**). Also, genotyping was done by the G2F initiative. G2F aims to analyze and predict GxE with high accuracy, finding the effect of genes (single loci, QTLs) and epistatic interactions; and aims to discover candidate genes across environments or specific to environments. Several research teams exploited findings from these G2F experiments, especially with a focus on genomic prediction (Westhues et al., 2021; Kick et al., 2022; Rogers and Holland, 2022) and on high-throughput phenotyping (Sankaran et al., 2020).

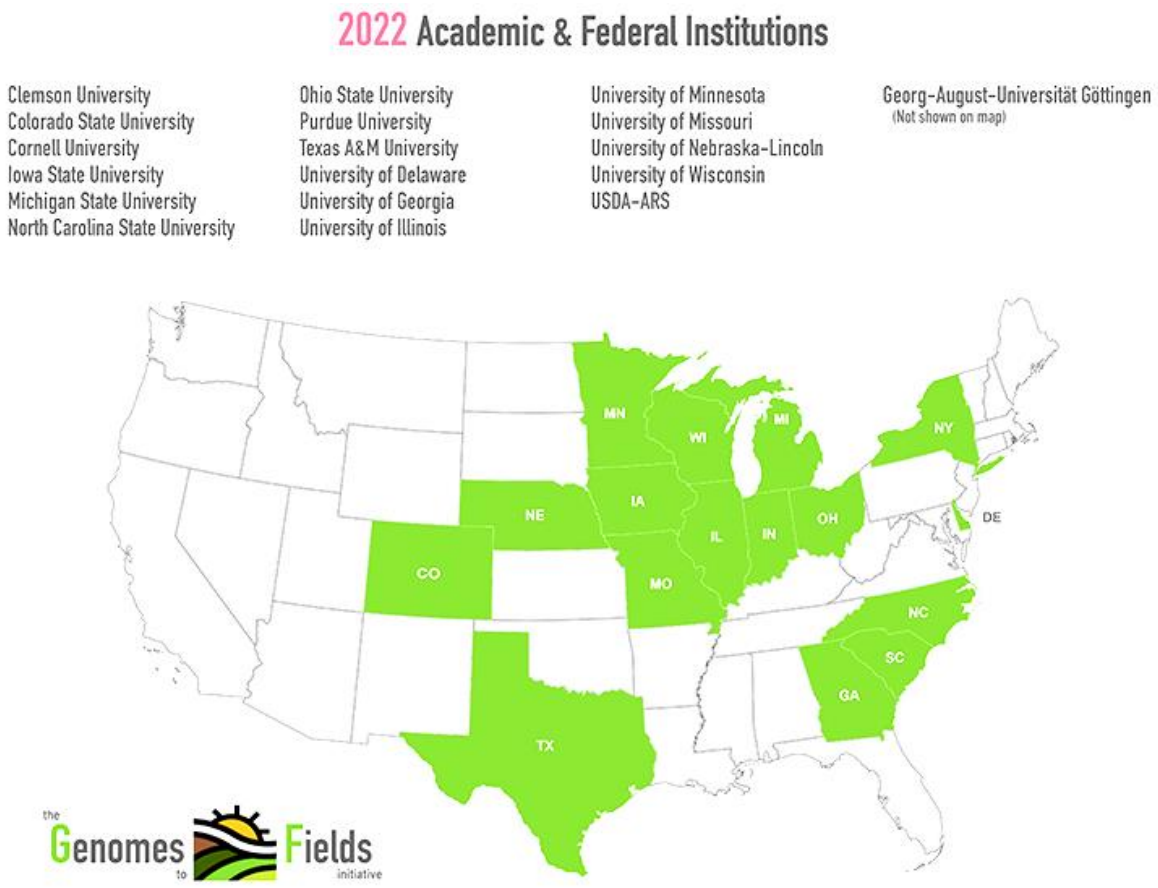


Figure 1: Genomes to Field (G2F) collaborators and the location of the genotypes tested (source: https://www.genomes2fields.org).

Importance of environment and QTL-by-environment interactions

Researchers and breeders have evaluated genotype-environment interactions, and one of the main goals of plant breeding is to identify genotypes that perform well under a wide range of environmental conditions (Kang, 1997; Annicchiarico, 2002). Hence, identifying the genes involved in a plant's phenotype variation and understanding how these genes interact with the environment is important to predict genotypes that perform well across environments. To achieve this, researchers have exploited multi-environment trials, hence testing the agronomic performance of different genotypes under diverse environmental conditions (Fox et al., 1997; Malosetti et al., 2013). The results of such trials have shed light on the complex interplay between genes and environments, and the concept of GEI is widely accepted and well-studied in the scientific community (Moffitt et al., 2005; Pigliucci, 2005; Laitinen and Nikoloski, 2019). In many

species, such as maize, wheat, and rice, researchers have identified location-specific QTLs, which are regions on chromosomes that play a role in determining the performance of a genotype. These QTLs are critical for breeders, as they provide valuable information about the performance of different genotypes in different locations, allowing breeders to make decisions according to the target environment (Weebadde et al., 2008; Barchi et al., 2012). Also, multi-environment trials have revealed stable QTLs consistent across different locations (Pilet et al., 2001; Hittalmani et al., 2003). These QTLs are valuable for breeders, as they can predict the phenotype accurately across different environmental conditions (Galiano-Carneiro et al., 2020).

Epistasis

Epistasis refers to the interaction between non-allelic genes that impact an organism's phenotype. Epistasis was first defined in the early 1900s by biologists and statisticians, with William Bateson defining it as the situation where genes at one locus mask the effect of genes at a different locus (Moore and Williams, 2009; de Visser et al., 2011; Mackay, 2014) (**Figure 3**). This concept is typically used to describe qualitative gene-gene interactions, such as determining eye color. Fisher, on the other hand, defined epistasis as deviation from additivity, meaning that the combined effect of genes at two loci on the phenotype is not identical with the sum of the individual effects of these loci (Fisher, 1919; Cordell, 2002; Moore, 2005). This concept of additivity should not be confused with intermediate gene action alleles at a locus, which is sometimes referred to as additive. The quantitative genetic perspective on epistatic interactions refers to whether and how the genes interact when causing a deviation from their expected joint effect on phenotype. Epistatic interactions have been debated among scientists for a long time, and they are easier to detect and observe in qualitative traits rather than quantitative traits (Whitlock et al., 1995; McKay et al., 2005; Phillips, 2008). As a result, epistatic interactions have often been overlooked or neglected in the study of quantitative traits. Detecting epistatic interactions in quantitative traits can require advanced statistical methods and a deeper investigation (Bender and Lange, 2001; Millstein et al., 2006). It is clear that epistatic interactions play a significant role in determining trait performance in plants, and therefore, scientific questions are built around epistatic interactions (Culverhouse et al., 2004; Carlborg and Haley, 2004).

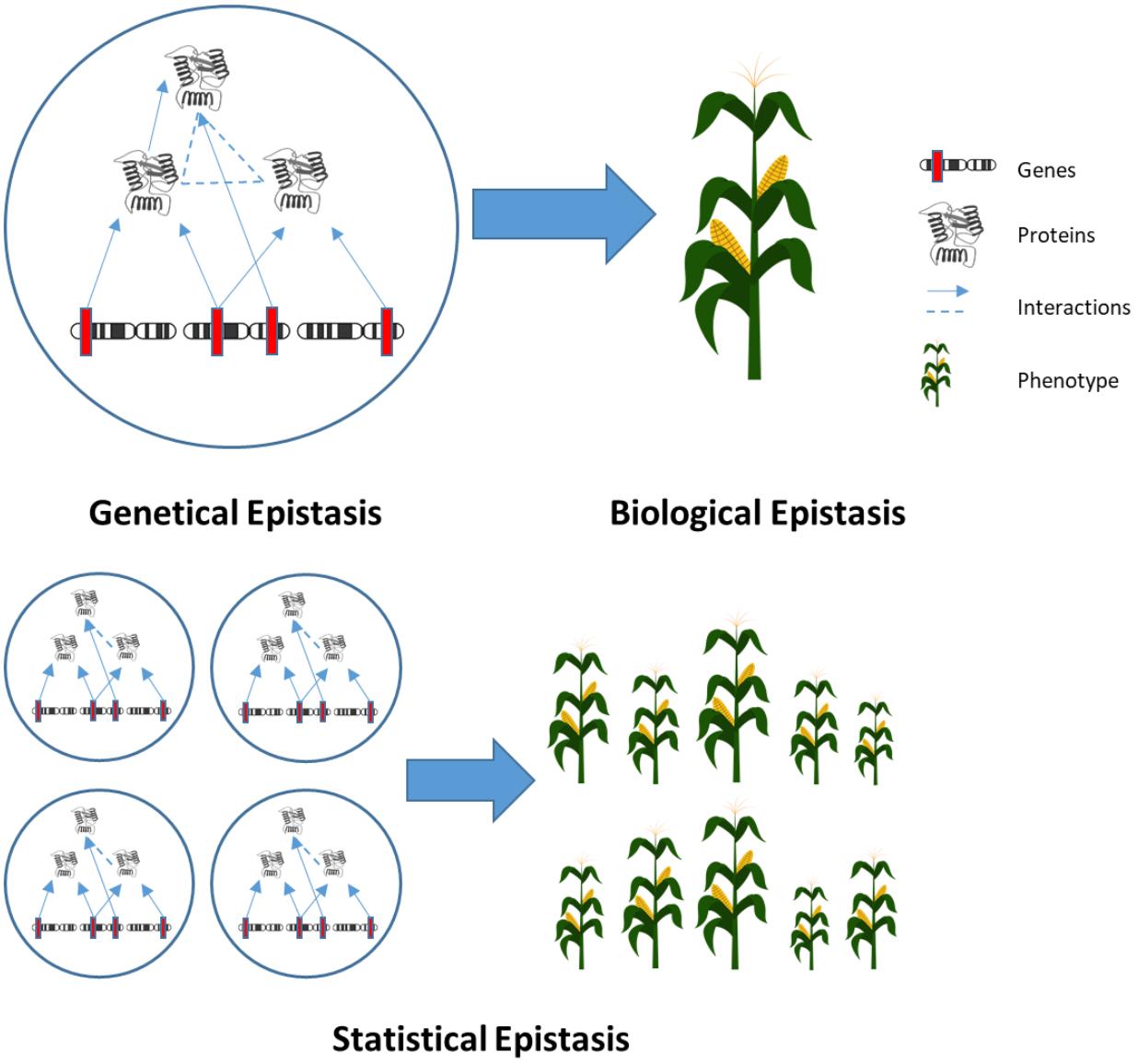


Figure 2: Epistasis encompasses genetic, biological, and statistical interactions influencing phenotypes. Genes undergo transcription and translation, directly impacting individual phenotypes (Genetic and biological epistasis). Additionally, interactions between proteins (shown as dashed lines) further shape phenotypic outcomes. Biological and genetic epistasis share a close relationship, involving molecular interactions within individuals. Statistical epistasis, however, operates at the population level, requiring variation among individuals to detect. It is important to note that epistasis can occur at the individual level without being statistically detectable (the figure is inspired by: (Moore, 2005)).

Epistasis is classified into three types based on its magnitude: negative, positive, and reciprocal sign epistasis. Negative epistasis occurs when the observed trait value resulting from the interaction between two mutations is lower than the expected sum of the gene effects. Conversely, positive epistasis is defined as a situation where the sum of the allele effects is greater than the interaction between the expected effects of the two genes. On the other hand, sign epistasis refers to a scenario in which the phenotypic value changes direction compared to the expected effects of the two genes (Phillips, 2008). Beneficial or deleterious alleles can impact the fitness of individuals, and genes can interact with each other. Positive and negative epistasis between two loci can alter the fitness effects in a population (de Visser et al., 2011; Bendixsen et al., 2017). For instance, even though two deleterious alleles individually many have detrimental effects, their joint occurrence and, hence, joint action could be beneficial in a competitive environment, and this surprise would hence be addressed as an epistatic interaction effect. Therefore, reciprocal sign epistasis can enhance the overall fitness of a population (Kvitek and Sherlock, 2011). In some cases, reciprocal sign epistasis can mitigate the adverse effects of deleterious mutations by suppressing their effects, a phenomenon known as genetic suppression (Mackay, 2014).

In evolutionary genetics, increased epistatic interactions can contribute to the unpredictability of a population's fitness and phenotype. Conversely, additivity could lead to a smoother fitness landscape, where the fitness of potential genotypes becomes predictable (**Figure 3**) (Kvitek and Sherlock, 2011). Consequently, a population in Hardy-Weinberg equilibrium in a specific environment will reach its maximum fitness potential. However, when complex environmental changes occur in a population, the chances of new mutations increase, which may exhibit negative epistasis between two loci that improve a trait of interest (Poelwijk et al., 2011). Nevertheless, individuals carrying such new mutations have the potential to surpass their local fitness maximum and outperform previous generations in a suitable environment (Steinberg and Ostermeier, 2016). Therefore, epistasis is another source of variation that can be exploited, and enhancing our understanding of epistasis and its interaction with the environment can aid in addressing challenges related to global warming and improving agricultural varieties.

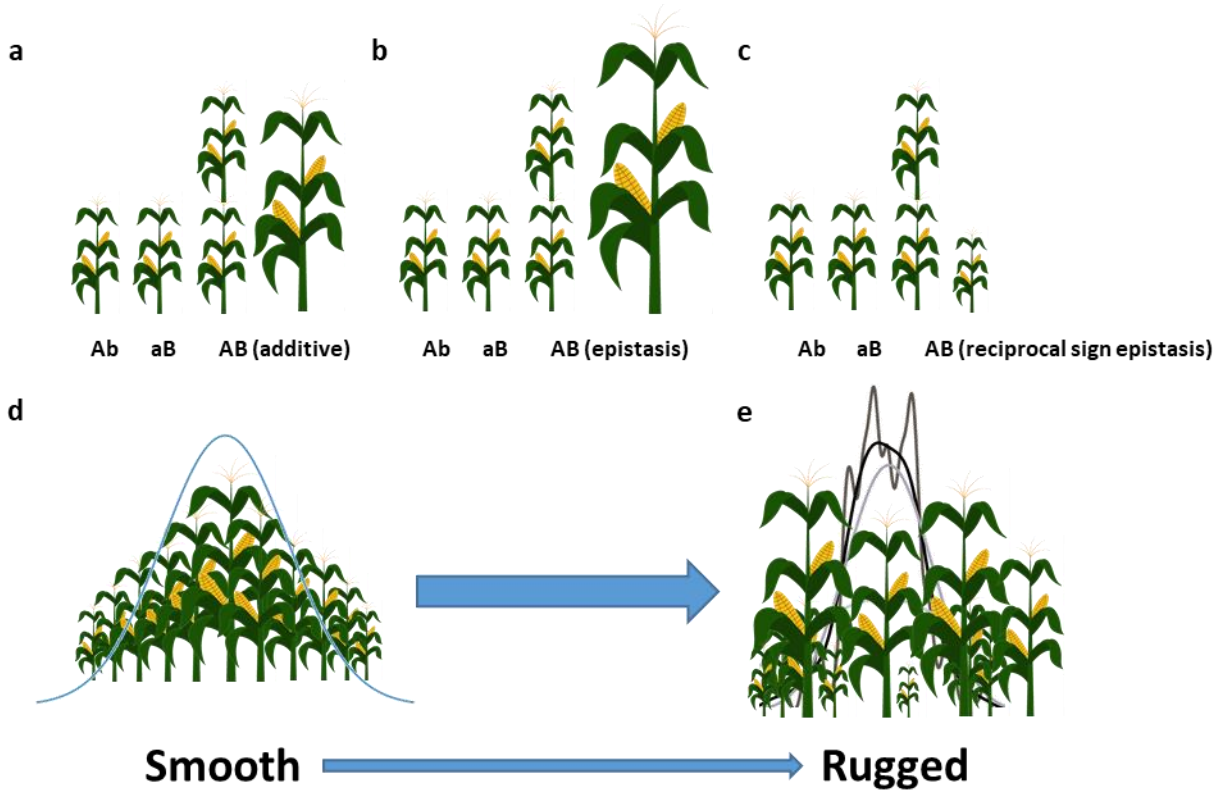


Figure 3: Interaction between two genes; additive (**a**), epistasis (positive) (**b**), reciprocal sign (**c**)). Below is a display of fitness landscapes between genes. When the number of additive interactions increases, the figure becomes smoother (**d**); with increased epistasis, the figure becomes rugged, and the fitness is unpredictable (**e**) (the figure is inspired by: (Shafee, 2014)).

Epistasis-by-environment interactions

The interaction between epistasis and environment is poorly understood, and only a few studies have been conducted on epistasis-environment interactions (EEI). In addition, testing a set of genotypes under diverse environmental conditions requires extensive resources and collaboration between multiple organizations. An excellent example of such collaboration is the Genomes2Field (G2F) program, a partnership between universities to test a set of genotypes in different locations. This program provides enormous data that allows researchers to study gene-

environment and gene-gene interactions. Furthermore, this data provides an excellent opportunity for EEI in more detail.

Epigenetics and epigenetic inheritance

Epigenetics involves studying changes in gene expression or cellular phenotypes that occur without altering the underlying DNA sequence. Epigenetic changes profoundly impact an organism's development, health, and disease susceptibility. Epigenetic mechanisms encompass a range of processes, including the chemical modification of DNA and its associated proteins, particularly histones. Through these modifications, epigenetics controls gene activation and deactivation, thereby influencing the synthesis of proteins and other essential molecules vital for various cellular functions. Prominent epigenetic modifications within plants include DNA methylation and histone modifications. These mechanisms are intricately interconnected, forming a complex network that collectively governs gene regulation and cellular responses to environmental cues. (Matzke and Mosher, 2014; Kim et al., 2015; Miryeganeh and Saze, 2020).

The term "epigenetic" first emerged in the 1940s, and it was redefined as "changes in a DNA sequence without altering the sequence," as officially established during the Cold Spring Harbor meeting in 2008 (Berger et al., 2009). While it is well known that DNA methylation undergoes a reset after germination in plants and during embryo development in humans and animals, some epigenetic markers can indeed be inherited by subsequent generations (Grossniklaus et al., 2013). The origins of these epigenetic changes are attributed to various factors, spanning from environmental influences such as stress and exposure to genetic and experiential influencers reverberating across generations. Epigenetic effects have also been observed in cloned plant progenies, revealing that successive cloning generations may not exhibit the same phenotype as the first generations. Researchers have pointed to an accumulation of methylation over later clonal generations to explain this phenomenon. A fundamental principle of epigenetics lies in the potential reversibility of these changes (Smulders and de Klerk, 2011). Nevertheless, there are cases where heritable epigenetic changes persist, exerting an enduring influence on mRNA expression in subsequent generations. Several types of epigenetic changes are summarized in the following sections:

DNA Methylation: DNA methylation is a fundamental epigenetic modification involving adding a methyl group to the cytosine base of a DNA molecule. This modification is pivotal in regulating gene expression by impeding the binding of transcription factors to the DNA sequence. In the context of plant genomes, DNA methylation primarily occurs in repetitive sequences that harbor CG dinucleotides; however, it can also extend to other sequence contexts such as CHG or CHH (where H represents A, C, or T), contingent upon the specific plant species (Bartels et al., 2018). In the intricate landscape of plants, DNA methylation serves multifaceted purposes. Notably, it silences transposable elements (TEs), mobile genetic entities that can potentially disrupt genomic integrity. This epigenetic silencing mechanism helps prevent the inappropriate activation and mobility of TEs, which can lead to genomic instability. Furthermore, DNA methylation orchestrates the precise control of nearby gene expression, exerting an influential role in fine-tuning various biological processes (Miryeganeh and Saze, 2020). One intriguing facet of DNA methylation is its responsiveness to environmental cues. Methylated TEs often maintain their silenced state unless they encounter specific environmental triggers. For instance, in response to pathogenic challenges like those posed by *Pseudomonas syringae*, certain promoter regions of TEs in the model plant Arabidopsis can undergo demethylation. This demethylation event is linked to the activation of immunity-related genes, effectively reprogramming the plant's defense responses against the pathogen (Yu et al., 2013).

Histone Modifications: Histones are pivotal proteins intricately involved in packaging DNA into a condensed structure known as chromatin. This packaging arrangement plays a crucial role in regulating the accessibility of DNA to various cellular processes, including transcription. Histone modifications can profoundly affect this accessibility, thus orchestrating the gene expression patterns essential for an organism's function and development. For instance, adding or removing chemical groups, such as acetyl or methyl groups, to specific histone residues can dictate whether a gene is actively transcribed or repressed. Precisely, histone acetylation typically corresponds to heightened gene expression. In contrast, histone methylation can have diverse outcomes—activation or repression—depending on the histone residue and its context within the chromatin structure.

Several notable examples have emerged as key regulators of gene transcription across various species as histone modifications. Acetylation events targeting histone H3 or H4, as well as methylation modifications involving histone H3 at lysine 4, 9, and 27 positions (abbreviated as H3K4me, H3K9me, and H3K27me), have garnered attention for their role in gene expression modulation.

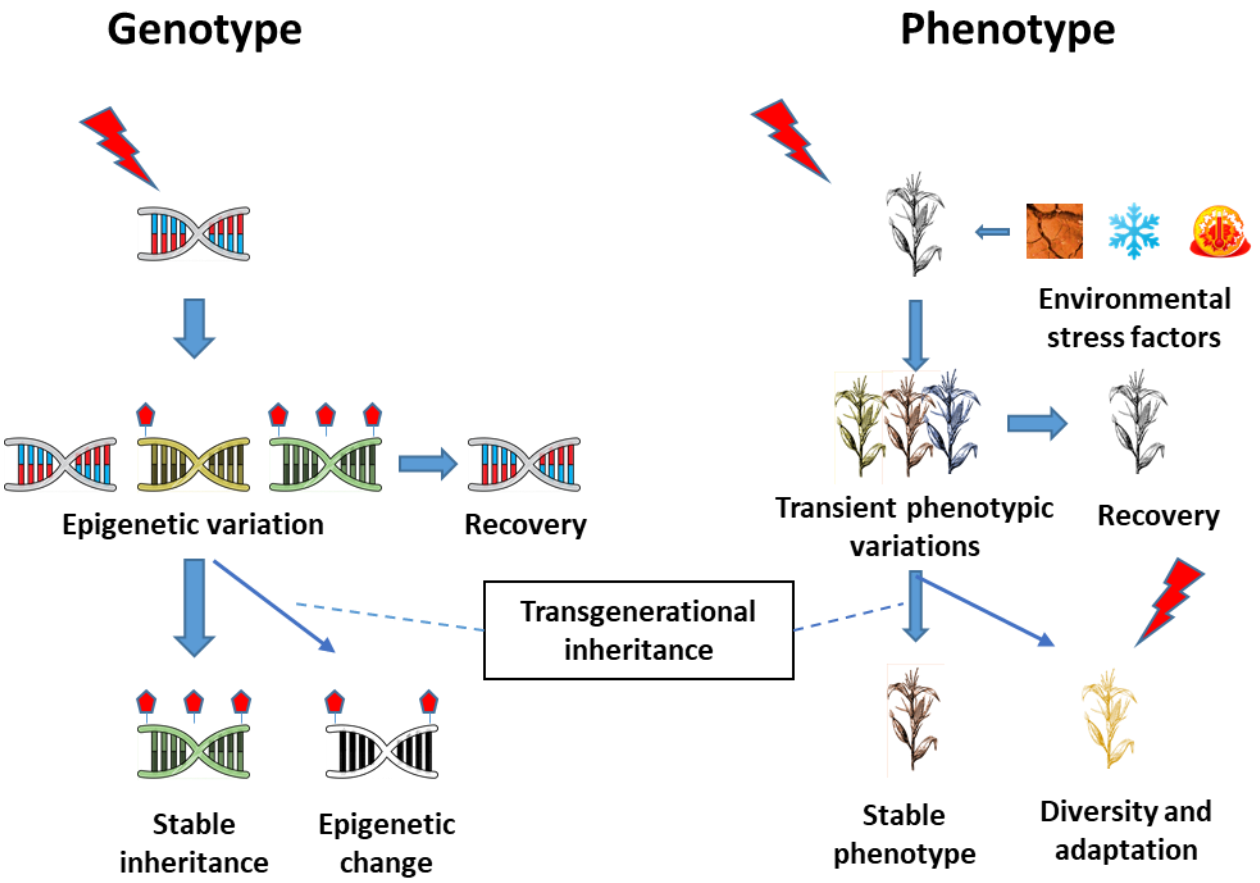


Figure 4: The model of epigenetic inheritance illustrates the relationship between genetic and epigenetic alterations on the left, resulting in changes in phenotype on the right, and the impact of environmental factors on phenotype displayed within the red box. While alterations in genotype/epigenotype and corresponding phenotype epigenetic changes can sometimes revert, leading to progeny with traits identical to their parents, subsequent generations might inherit these changes, a phenomenon referred to as transgenerational or epigenetic inheritance (the figure is inspired by: (Miryeganeh and Saze, 2020)).

Small RNAs: Small RNA molecules, such as microRNAs and small interfering RNAs (siRNA), are involved in RNA interference (RNAi). These molecules can target and degrade complementary mRNA molecules or inhibit their translation into proteins. RNAi plays a role in regulating gene expression and responding to environmental stress. Small RNAs play a particular role in preventing TE expression of TEs in the genome or silencing mRNAs (Zamore and Haley, 2005; McCue and Slotkin, 2012; Wheeler, 2013). Those small RNAs are part of a complex methylation process called RNA-directed DNA methylation (RdDM). In short, RNA polymerase IV transcribes a single-strand RNA, and RNA-dependent RNA polymerase 2 (RDR2) makes single-strand to double-strand RNAs, Dicer-Liker 3 (DCL3) cuts or shorten the double-stranded RNAs into sRNA duplexes, and those duplexes carry by Argonaute 4 (AGO4) to sequence-specific recognition target RNA or DNAs. In summary, small RNAs guide the sequence-specific recognition of target DNA, facilitating the recruitment of DNA methyltransferases and the establishment of DNA methylation patterns (Zhang and Zhu, 2011; Matzke and Mosher, 2014; Erdmann and Picard, 2020).

Transgenerational Epigenetic Inheritance: Plants can transmit epigenetic modifications to their offspring. These modifications can be stable over multiple generations. For example, stress conditions experienced by a plant can lead to epigenetic changes that prepare its offspring to tolerate similar stresses better (Ashe et al., 2021; Anastasiadi et al., 2021).

Epigenetic marks, such as DNA methylation and histone modifications, can accumulate over an individual's lifetime due to diverse environmental factors and experiences (Baulcombe and Dean, 2014). Notably, some of these marks can endure and be passed on to future generations, potentially impacting gene regulation and cellular processes in those descendants (Ashe et al., 2021). However, it is crucial to acknowledge that the mechanisms and extent of epigenetic inheritance remain topics of ongoing research and debate within the scientific community. While evidence supports the concept of transgenerational epigenetic effects, the precise mechanisms and the breadth of their impact are intricate and necessitate further investigation (Jablonka and Raz, 2009; Lacal and Ventura, 2018).

Importance of gene expression and origin-of-seed effect

Gene expression, the process in which genetic information is utilized to produce functional products such as proteins, is crucial for the optimal functioning of cells and organisms (Koch, 1996). It involves DNA transcription into RNA, RNA translation into a protein, and the expression of genes to control the quantity of gene products (Casassola et al., 2013) (**Figure 5**), which leads to specific functions and plays a significant role in the development, response to stress, and adaptation to changing environments (López-Maury et al., 2008). Additionally, environmental factors, which are heritable epigenetic changes, can impact gene expression and be passed on to future generations (Trerotola et al., 2015). Heritable epigenetic changes must be further investigated since this phenomenon could influence the phenotype of individuals (Gallusci et al., 2017).

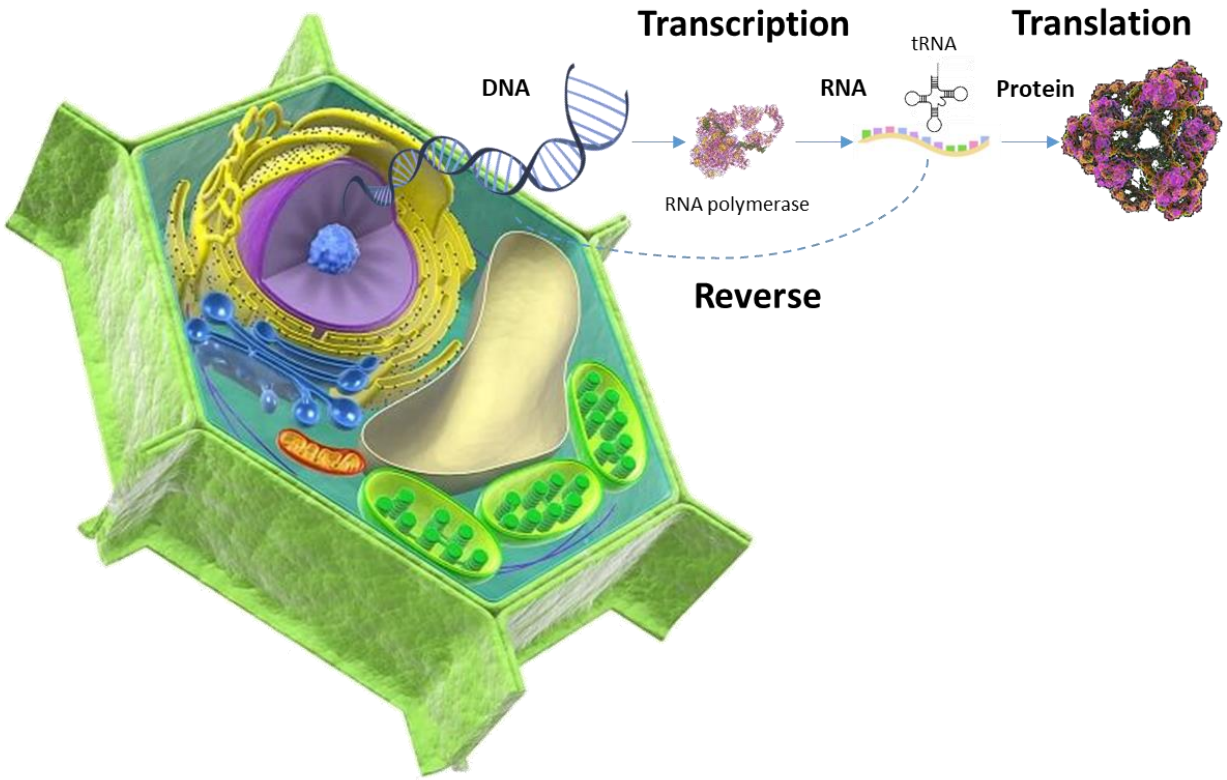


Figure 5: Central Dogma of molecular biology (Figure is inspired by: (Koonin, 2012)).

The impact of environmental factors on gene expression has been extensively studied, but the source material of EMP comes from two different sources. However, our understanding of the effects of the environment on seedlings remains incomplete. Therefore, we aim to investigate the origin-of-seed effects using genotypes raised under diverse environmental conditions, utilizing EMPs as a source material. This research investigates the differential expression of genes between two seed sources, a nursery (a field experiment) and a greenhouse, to determine whether material from different origins can be utilized in mRNA studies.

Aim of thesis

Previous research has emphasized the importance of environmental conditions on the expression of genes and genotypes. The performance of a genotype may vary significantly in different environments. While epistasis can significantly influence an individual's phenotype, limited information on how the environment impacts epistasis is currently available. Furthermore, the effects of origin-of-seed factors on differentially expressed genes (DEGs) on seedling stages from different seed sources remain to be fully explored. We aimed to shed light on these topics; hence, the thesis was structured according to several chapters, from “General Introduction” to “Conclusion,” and the following are the aims of the two research chapters of the thesis:

1. In the thesis's first research part, we aimed to identify significant QTLs and epistatic interactions across and between environments. This chapter investigated significant markers from the QTL analysis to identify QTL x environment and epistatic-by-environment interactions.
2. In the second research part, we examined the origin-of-seed effect of EMP and searched for expression differences between the two seed sources.

Summary of material and method

1. In the first research chapter, we calculated Mega Environments (ME) and best linear unbiased estimation (BLUEs) of genetic values across MEs. We used these BLUEs for further analysis. Also, QTL mapping and identifying significant main effect QTLs was performed. In order to investigate epistasis-by-environment interactions (EEI), we used a linear mixed model approach. The significance of EEIs was assessed with permutation tests.
2. The second research chapter examined potential origin-of-seed effects by conducting differential expression analysis on our Epistasis Mapping Population (EMP). In addition, we compared distinct genotypes to check the validity of the analysis. The results of this study will be presented and discussed in the final sections.

Chapter 2: Identification of Significant Quantitative Trait Loci (QTL) in Diverse Mega Environments (MEs) and Interactions Between Epistasis by Environment in Maize

Abstract

Genes play a crucial role in determining the phenotype of organisms. Identifying the genes responsible for phenotypic traits is vital to improving these traits' performance through breeding programs. This study used a quantitative trait locus (QTL) mapping approach to identify the phenotypic traits contributing to phenotypic performance. The dataset used was from the Genomes to Field (G2F) program in maize, which created and distributed a mini nested association mapping (NAM) population across 29 locations. Phenotypic observations were made using a standardized operation procedure (SOP) by G2F collaborators. The following phenotypes were evaluated: pollen days after plantation (pollen DAP), silking DAP (silk DAP), plant height (PH), ear height (EH), stand percentage (stand %), and yield. The environments were compacted into seven multi-environment trials (MEs) to increase the power and simplify the data. The learnMET R package and the NASApower package with a k-means approach were used to create the MEs. The genotyping was done using skim-sequenced and single nucleotide polymorphic (SNP) markers called using the Practical Haplotype Graph (PHG) approach by the G2F initiative. The genotypic data were filtered based on a 10 kb window and minor allele frequency (MAF) > 0.02 criteria. A linkage map was created using Lep-MAP3 software with the Kosambi mapping function. Ten linkage groups (LG) were selected for further analysis, and a simple interval mapping (SIM) approach was applied to identify significant markers. This chapter aimed to identify significant QTLs associated with quantitative traits, and several significant QTLs, along with location-specific ones, were found.

Material & Methods

Population

In this study, we utilized a mini Nested Association Mapping (NAM) population of maize created through the G2F initiative in 2018. The NAM population consisted of three double haploid (DH) families, each comprising three inbred lines crossed with a common parent (**Figure 1**). The NAM population comprised 306 individual DH lines, excluding the parents. The DH families (Mo44, PHN11, and MoG) were genotyped and phenotyped, containing 100, 142, and 64 individual plants, respectively. The inbred lines were crossed with a location-specific tester (PHT69 or LH195), and the phenotypic performance of the resulting hybrids was evaluated in 29 diverse environments. Field trials were conducted separately in different locations across the US, Canada, and Germany by G2F collaborators, and each collaborator followed a Standard Operation Procedure (SOP) before, during, and after harvesting. The phenotypic observations for six traits were recorded: number of days after planting to pollen shed (pollen DAP), number of days after planting to silking (silk DAP), plant height (PH), ear height (EH), stand percentage in a plot (stand %), and yield. The measurements were taken according to the SOP, with pollen and silk DAP measured when 50% of the plot flowered and silked, respectively. PH and EH were calculated in centimeters (cm) from ground level to the flag leaf and ground to the primary ear, respectively. Stand percentage was calculated as a ratio of harvested plants to sowed plants, and yield was measured in bushels per acre (bu/Ac) by a harvester at each location and corrected for moisture content. All the phenotypic data collected are available on the G2F website (https://www.genomes2fields.org).

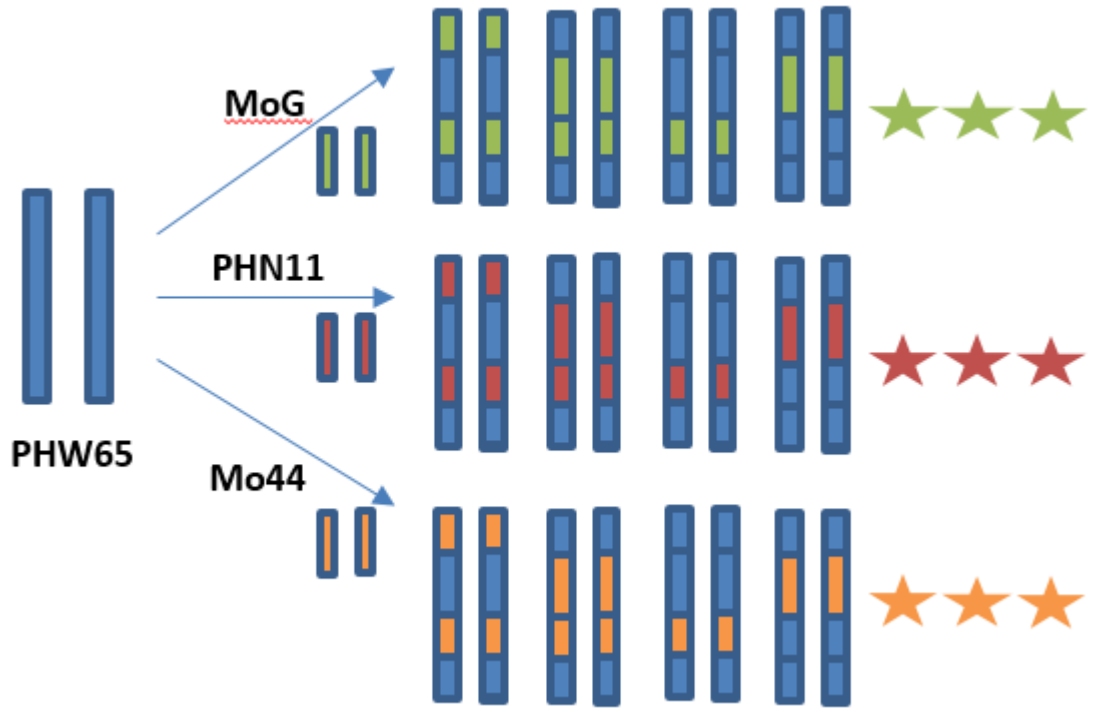


Figure 1: Illustration of Population of Genomes to Field (G2F). Three inbred lines (Mo44, PHN11, MoG) crossed with a common parent PHW65. Each arrow indicates a distinct DH family of 100, 142, and 64 genotypes.

Mega environments (MEs)

To reduce the number of QTL mapping tests per environment and increase statistical power, we adopted a strategy of clustering locations into Mega Environments (MEs) based on their environmental data. The environmental covariates (ECs) of the Genomes to Field (G2F) dataset were calculated using the daily environmental data from NASApower with ten fixed window days across the locations (Sparks, 2018). We used the learnMET/R package, which implements the NASApower/R tool, to perform this task (Westhues et al., 2021). We then employed the k-means algorithm on ECs to cluster the locations into seven different MEs, as shown in **Figure 2**. One ME was excluded from the statistical analysis due to missing observations on pollen DAP, silk DAP, PH, and EH.

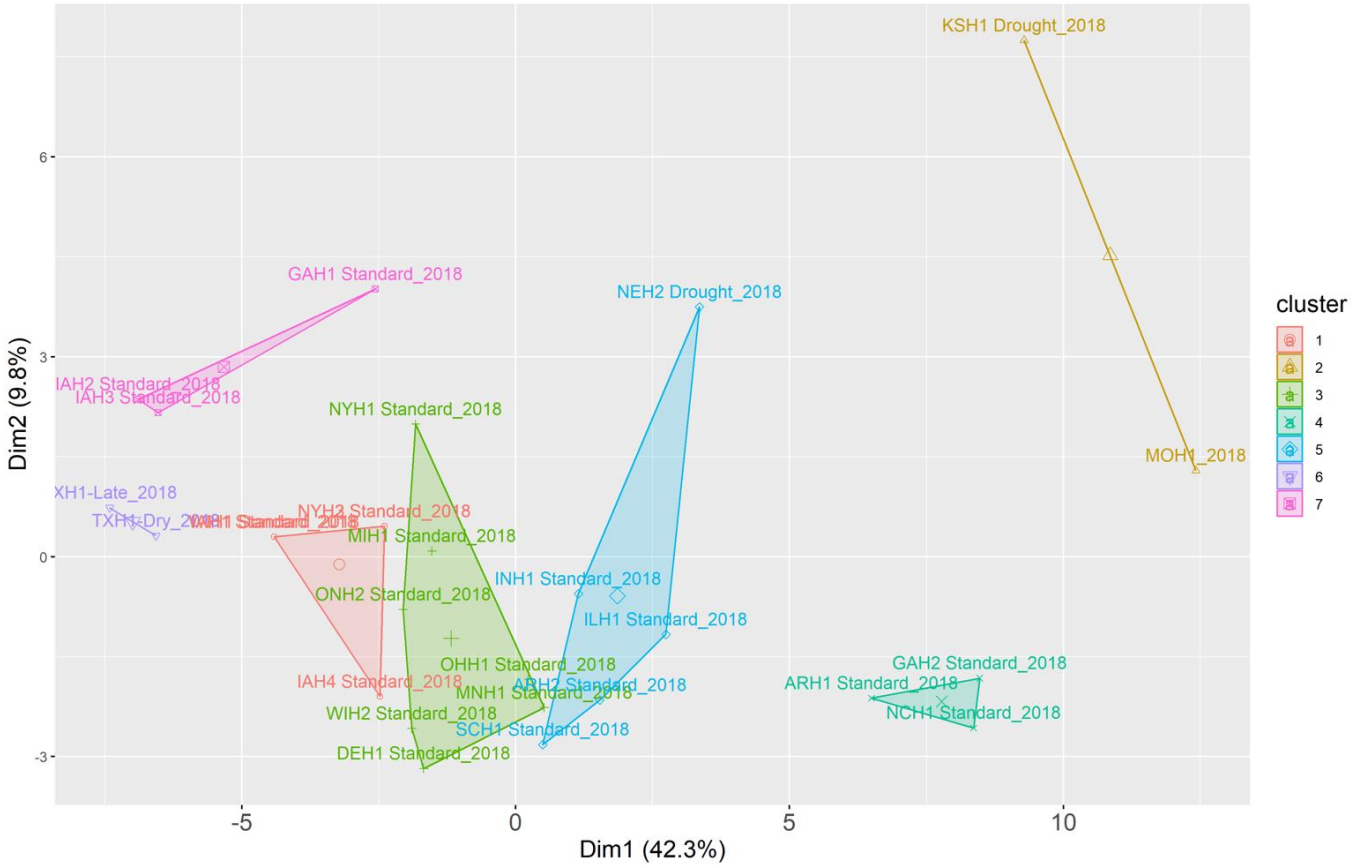


Figure 2: Creation of mega environments (MEs) with 29 locations of Genomes to field (G2F) with environmental covariates (ECs) from NASApower R package.

Genotyping

The G2F consortium employed the Practical Haplotype Graph (PHG) to genotype the NAM population in 2018. The inbreeds of G2F were sequenced with an average coverage of approximately 5x, and Novogene sequenced pools of 24 on a HiSeq X Ten lane. A total of 573 samples and 3.6 million SNPs were called. Initial filtering was performed by estimating the SNP call error of PHG through comparison with previous Genotype-by-sequencing (GBS) calls (Elshire et al., 2011). Additionally, mismatched SNPs coinciding with GBS calls (>2%) and minor allele frequency (MAF) below 0.000001 were filtered out. After these filtering steps, the final number of SNPs and individuals was 1.3 million and 312, respectively (Gage, 2018).

Principle Component Analysis (PCA)

Principal Component Analysis (PCA) is a widely used statistical method to investigate patterns and structures in large datasets. PCA reduces the dimensionality of data and identifies individuals' genetic structure. In this study, we performed a PCA using the SNPrelate R package on 1.3 million SNP markers from 312 individuals. The resulting plot in **Figure 3** allows us to visualize the samples' genetic relatedness and population structure. The PCA can further investigate the relationships between the individuals and identify any outliers or subpopulations.

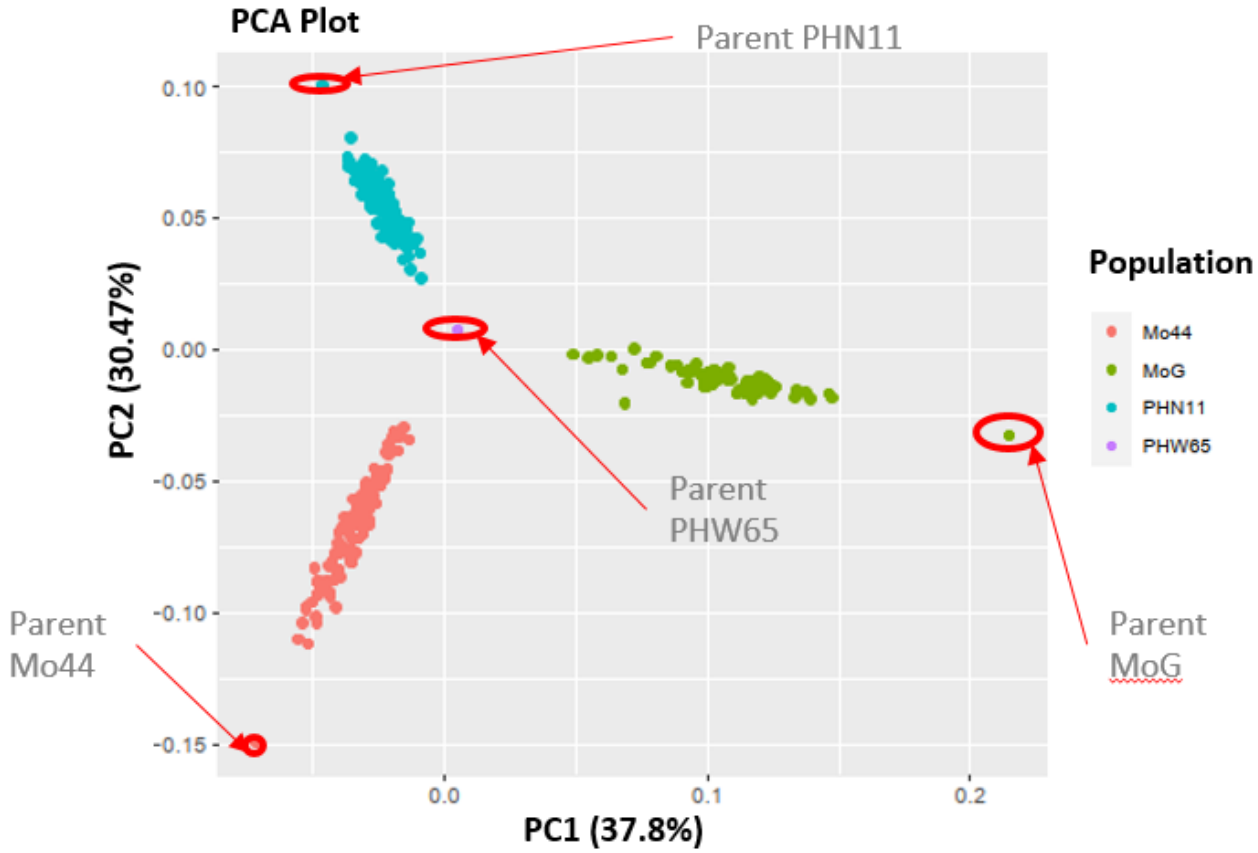


Figure 3: Distribution of the individuals in the first two principal components. Each dot represents an individual, and the dot's color represents a group of individuals.

Filtering, marker trimming, and linkage mapping

We applied additional filtering and trimming to the SNP markers based on the following criteria: SNPs with a minor allele frequency (MAF) below 0.02 and markers with a missing value rate higher than 0.1 were removed. Most markers fell within the MAF range of 0.00 to 0.02, as shown in

Figure 4. Therefore, we chose the MAF filtering threshold based on the allele frequency distribution in the NAM population. After applying the filters, the final number of SNPs used in the analysis was 1.1 million.

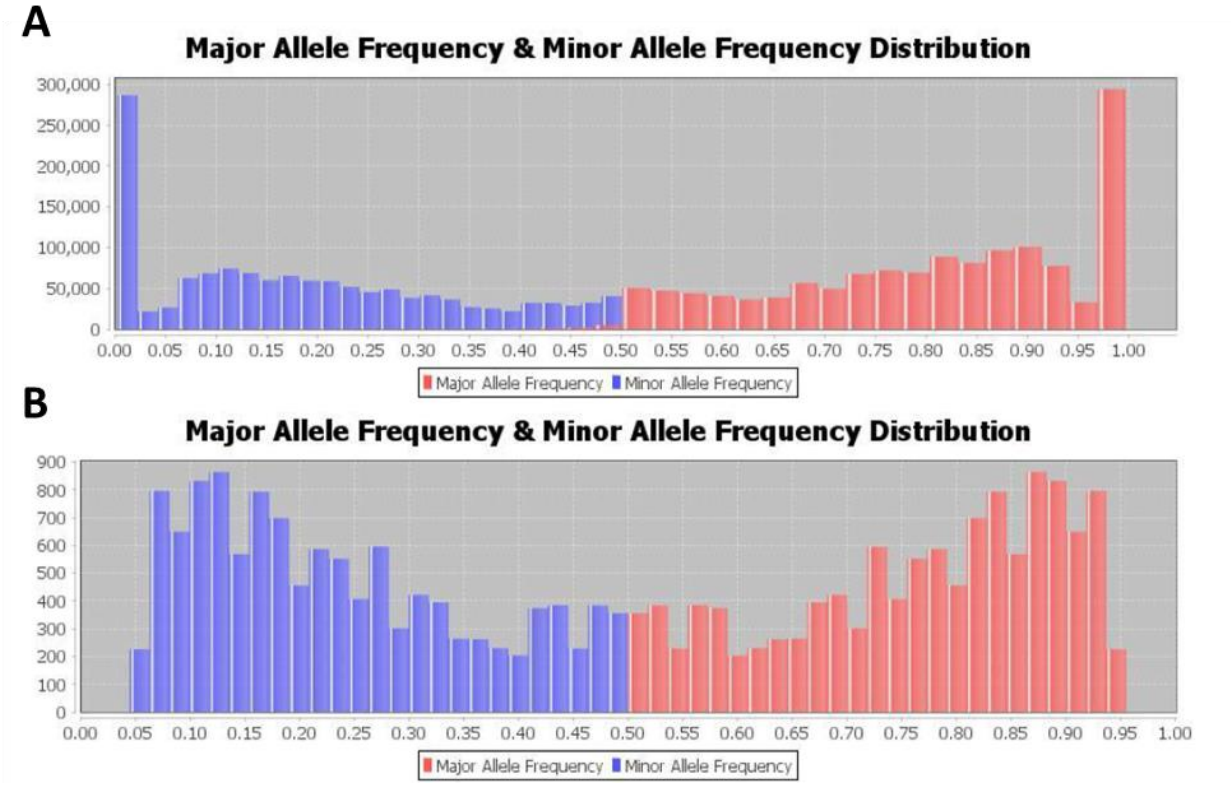


Figure 4: Distribution of allele frequencies before filtering minor allele frequency (MAF) < 0.02 (**A**) and after filtering MAF markers (**B**).

In addition to the initial filtering and trimming steps, we further trimmed the SNP markers with the TASSEL software (Bradbury et al., 2007). Specifically, we selected one marker for every 10-kb window. After trimming, the total number of markers that remained for linkage mapping was 14,559. To create a linkage map, we used Lep-MAP3 software (Rastas, 2017), which employs the MSTmap algorithm and can consider cofactors in the analysis, such as family. The following parameters were defined for the software: we used a LOD score of 18 to separate chromosomes with the SeparateChromosome2 function. Subsequently, we selected ten linkage groups (LG) out of 40 and used the OrderMarkers2 function to order the LGs, with the Kosambi mapping function applied. We manually matched the LGs with the SNP physical positions to find the chromosome numbers of the LGs. Co-segregating markers were discarded, resulting in 2,415 SNP markers

utilized for QTL mapping. Finally, we examined the recombination fraction distribution between the 2,415 SNP markers (**Figure 5**).

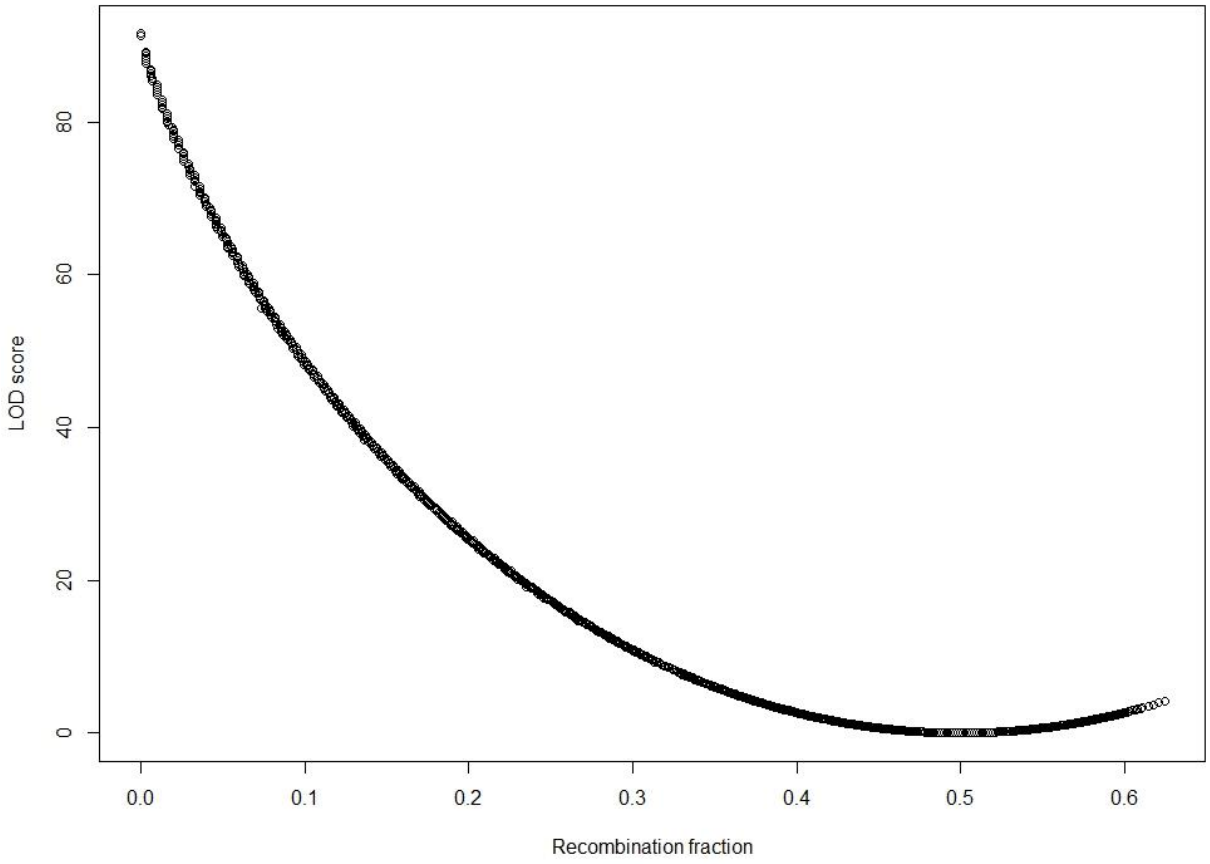


Figure 5: Recombination fraction (RF) of marker comparison between 2,415 SNP markers. If there is no linkage between markers, RF is 0.5. In a fully link situation, RF is 0.

Best linear unbiased estimation (BLUEs)

The best linear unbiased estimation (BLUE) values of the phenotypes were estimated using a linear mixed model and were then utilized in the QTL mapping. BLUE values were estimated separately for all locations and mega environments, and these values were used as phenotypic values in the QTL mapping. The linear mixed model used a k_{th} tester (T_k), p_{th} family (F_p), i_{th} genotypes (G_i), and r_{th} replication (R_r) as random effects. j_{th} mega environment (ME_j) is considered a fixed effect, and y is explained as the trait of interest (**Formula 1**). A similar linear mixed model was used for each mega environment separately. Genotypes were used as fixed effects. If the mega environment contained only one tester, the tester effect (T_k) was discarded

from the analysis. Additionally, the locations (E_j) were evaluated and used as random effects instead of the ME (ME_j) from model 1 (**Formula 2**). The linear mixed models were evaluated using the lme4 R package (Bates et al., 2015). The distribution of BLUE values for each trait is given in All locations (**Figure 6**) in all MEs (**Figures 7 to 12**).

Formula 1:

$$y = \mu + G_i + ME_j + T_k + F_p + R_r + \varepsilon$$

Formula 2:

$$y = \mu + G_i + E_j + T_k + F_p + R_r + \varepsilon$$

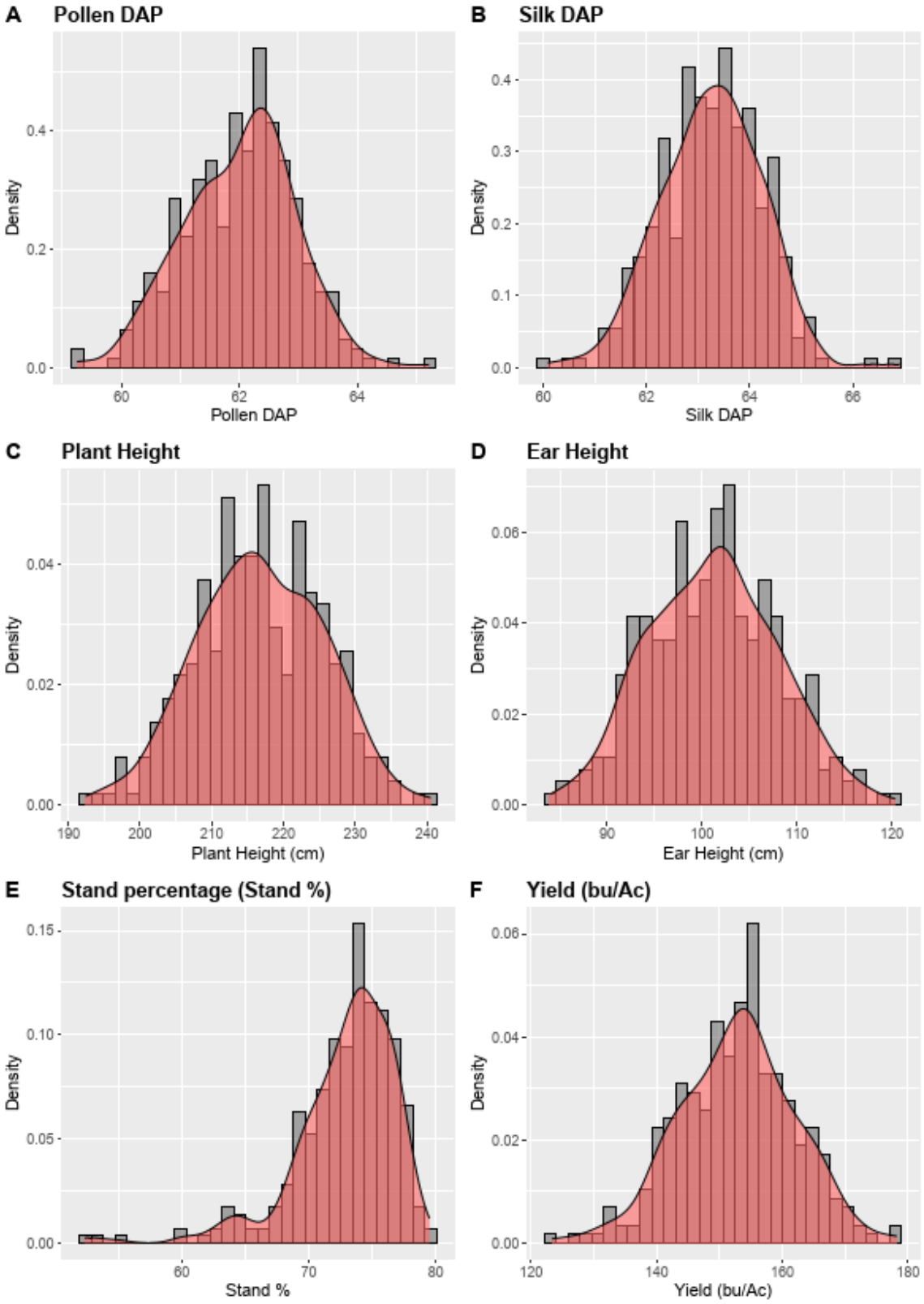


Figure 6: Density plot of phenotypic traits of pollen DAP (**A**), silk DAP (**B**), PH (**C**), EH (**D**), stand % (**E**), and yield (**F**) across locations.

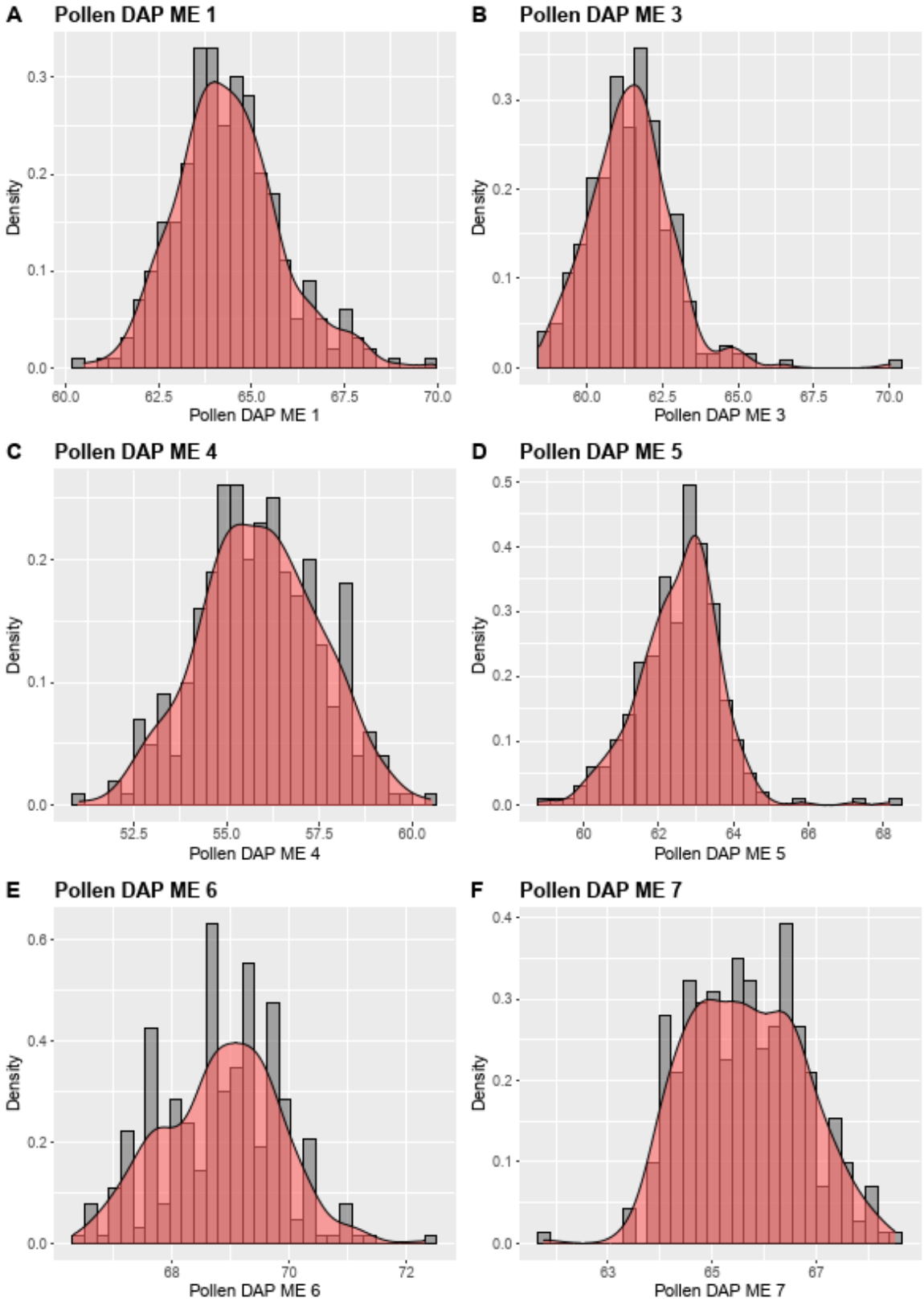


Figure 7: Density plot of phenotypic traits of pollen DAP ME 1 (**A**), ME 3 (**B**), ME 4 (**C**), ME 5 (**D**), ME 6 (**E**), and ME 7 (**F**) across mega environments.

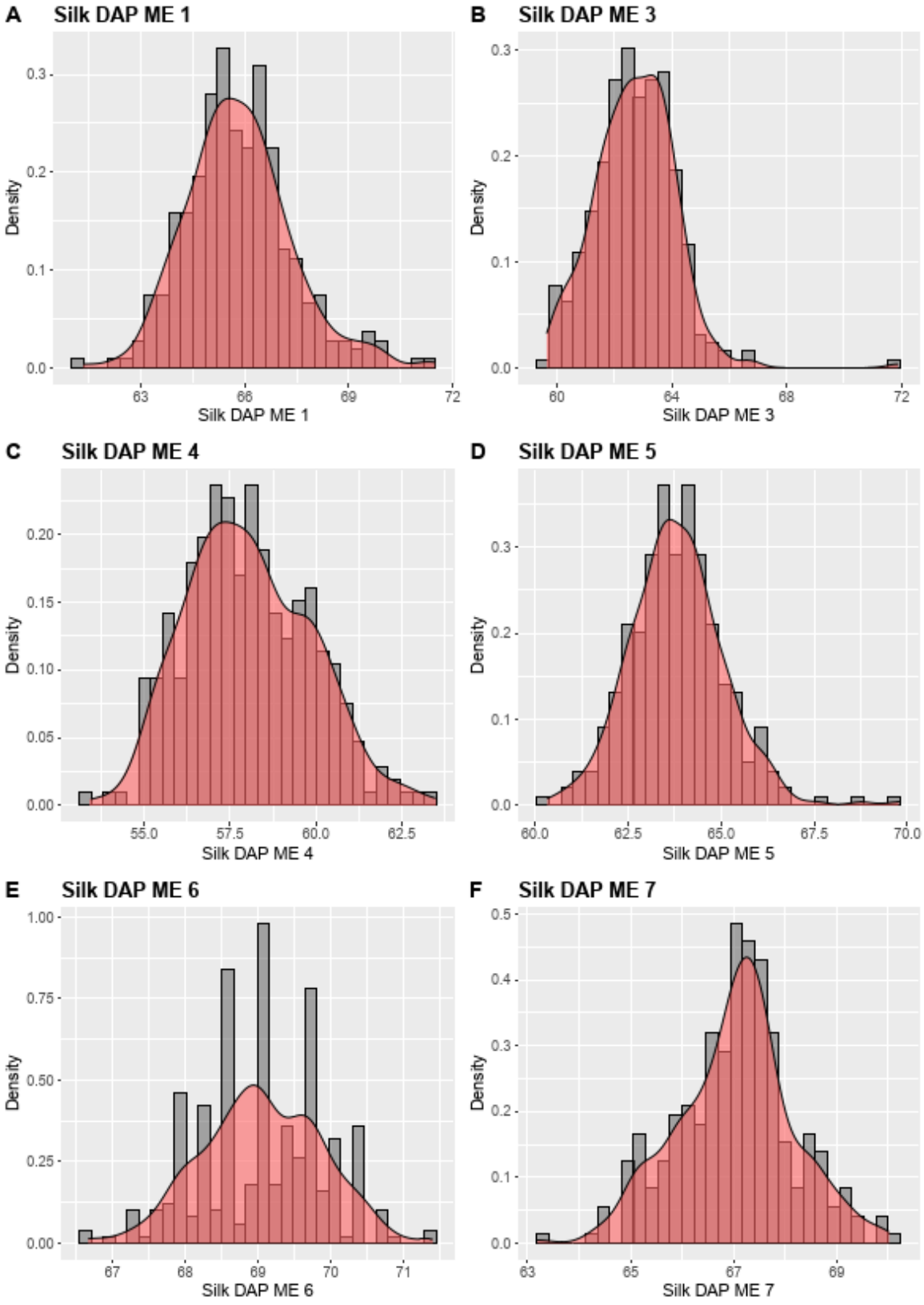


Figure 8: Density plot of phenotypic traits of silk DAP ME 1 (**A**), ME 3 (**B**), ME 4 (**C**), ME 5 (**D**), ME 6 (**E**), and ME 7 (**F**) across mega environments.

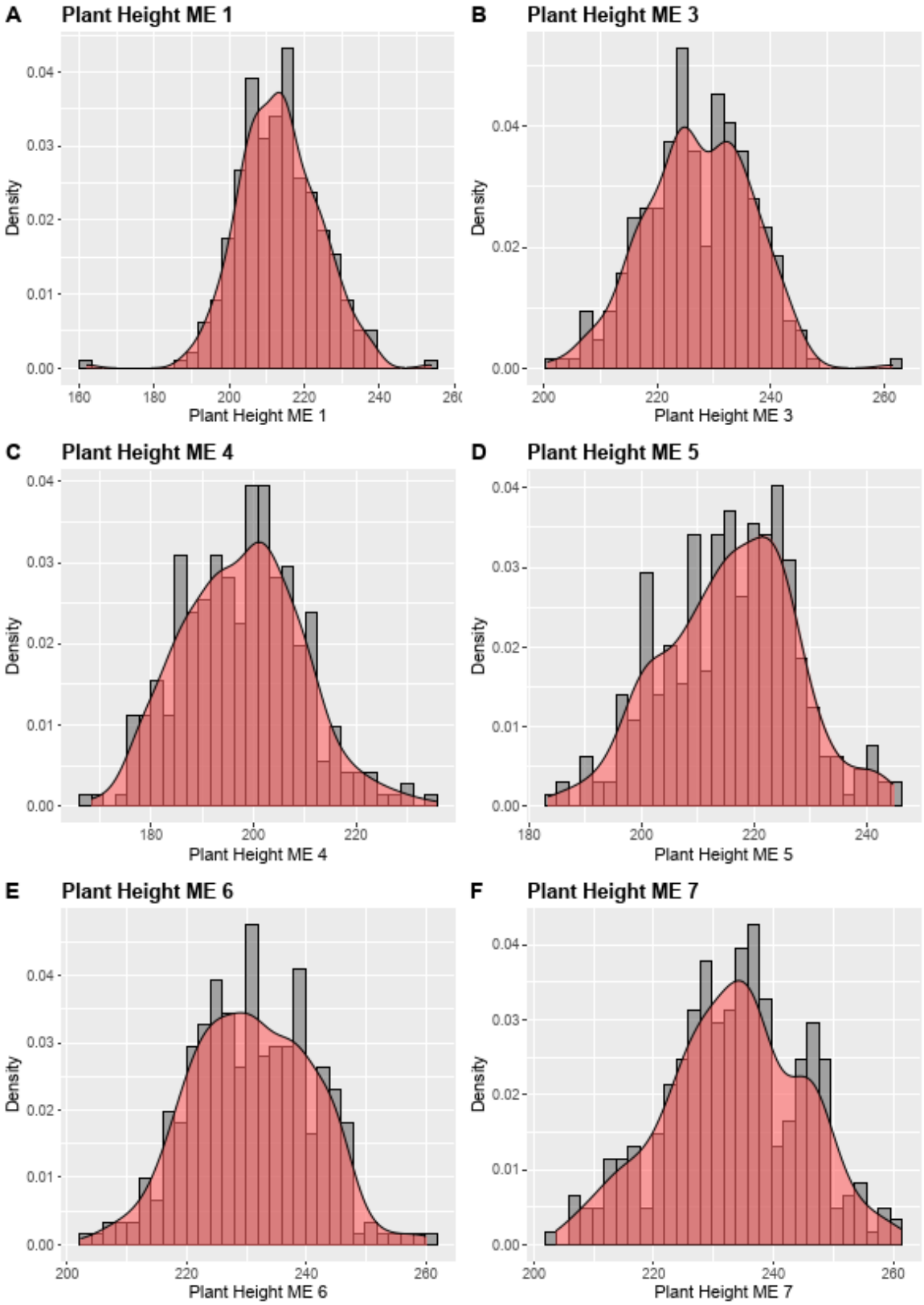


Figure 9: Density plot of phenotypic traits of plant height ME 1 (**A**), ME 3 (**B**), ME 4 (**C**), ME 5 (**D**), ME 6 (**E**), and ME 7 (**F**) across mega environments.

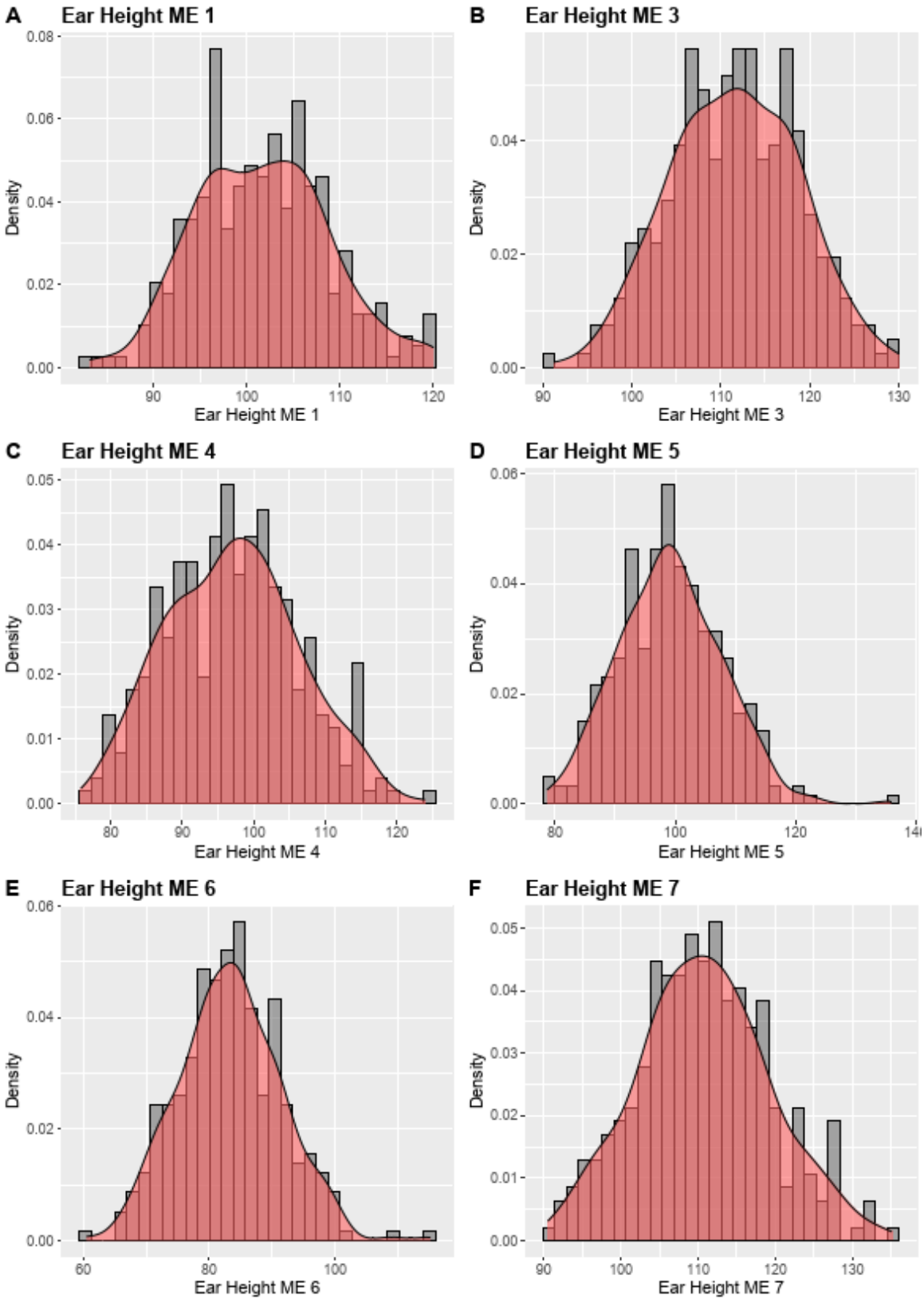


Figure 10: Density plot of phenotypic traits of ear height ME 1 (**A**), ME 3 (**B**), ME 4 (**C**), ME 5 (**D**), ME 6 (**E**), and ME 7 (**F**) across mega environments.

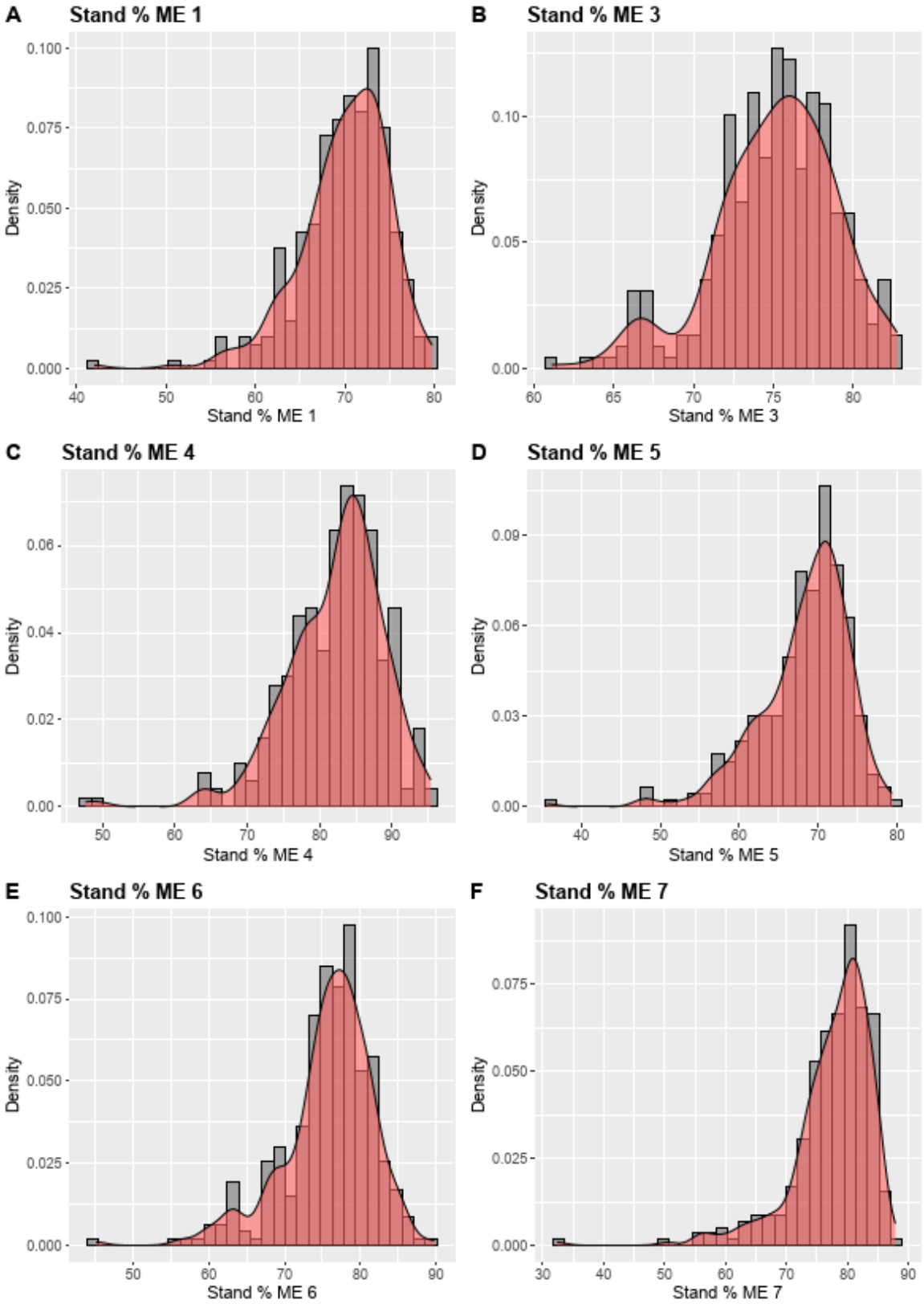


Figure 11: Density plot of phenotypic traits of stand percentage ME 1 (**A**), ME 3 (**B**), ME 4 (**C**), ME 5 (**D**), ME 6 (**E**), and ME 7 (**F**) across mega environments.

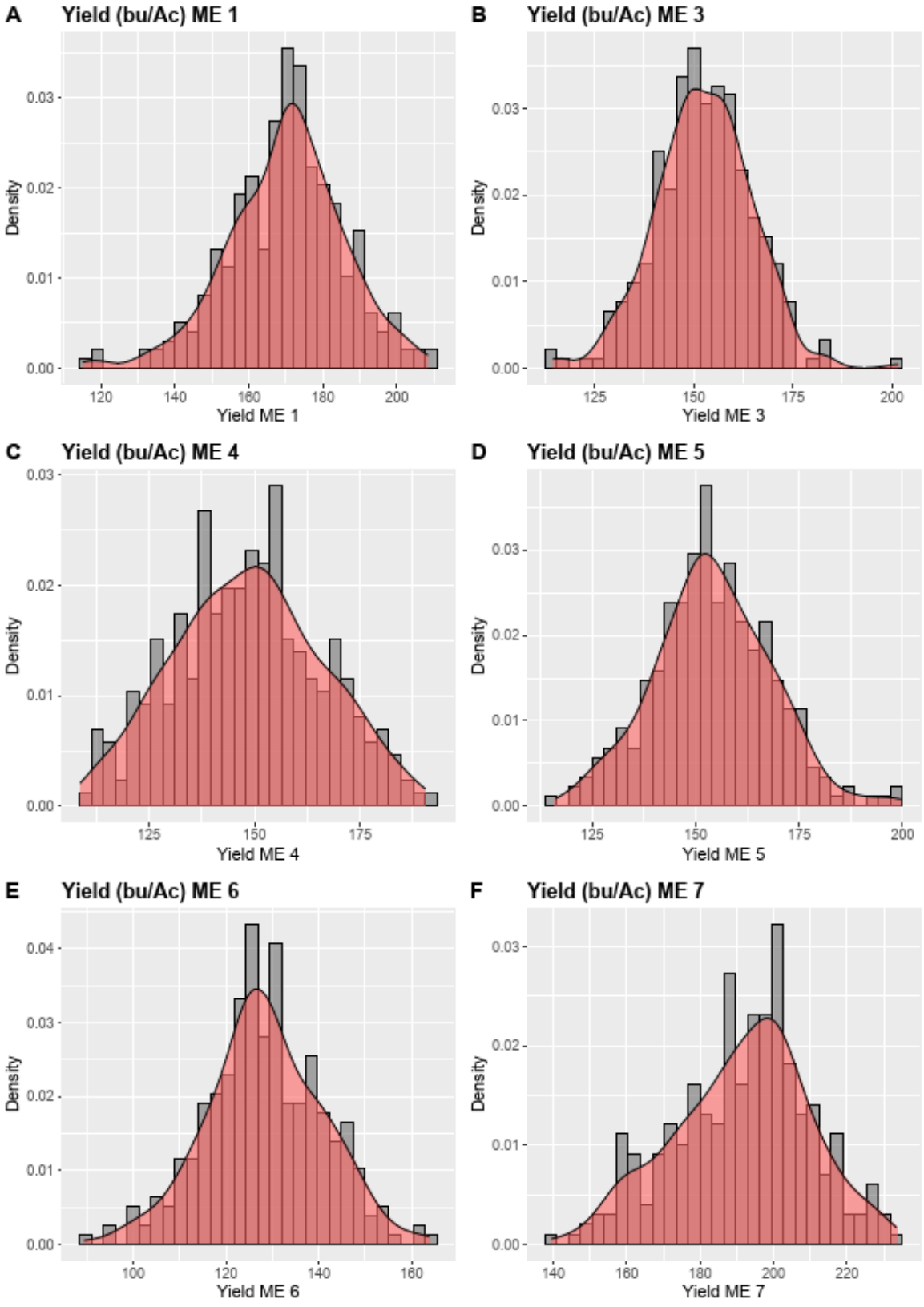


Figure 12: Density plot of phenotypic traits of yield (bu/Ac) ME 1 (**A**), ME 3 (**B**), ME 4 (**C**), ME 5 (**D**), ME 6 (**E**), and ME 7 (**F**) across mega environments.

Broad sense heritability

Broad sense heritability (H^2) is an essential measure of the extent to which phenotypic variation in a specific population is attributable to genetic variation. We used the formula James B. Holland and colleagues described in 2003 (Holland et al., 2003). It is calculated using the following formula:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_\varepsilon^2}{p_h}}$$

where σ_g^2 is the genotypic variance, σ_ε^2 is the error variance, and p_h is the number of replications per environment. The genotypic variance (σ_g^2) can be estimated as the variance among genotypes within a population. It is a measure of the total genetic variation in the population that contributes to phenotypic differences in the trait of interest. The error variance (σ_ε^2) represents the variation not explained by genetic factors, such as environmental effects, measurement error, and other sources of variability.

The number of replications per environment (p_h) is an essential factor in calculating H^2, as it influences the precision and accuracy of the estimates. The more replications performed, the more precise the estimates of genotypic and error variances will be and the more reliable the estimate of H^2 will be.

Broad sense heritability is an essential parameter in plant breeding and genetics, as it measures the potential for selection to improve the trait of interest. High heritability indicates that a large proportion of the phenotypic variation is attributed to genetic factors, and thus, selection is likely to be effective in improving the trait. Conversely, low heritability indicates that environmental factors are more prominent in determining the phenotype, and selection may be less effective. Therefore, an accurate estimation of H^2 is crucial for optimizing breeding strategies and improving crop yields.

QTL mapping

The R/qtl package (Broman et al., 2003) analyzed 2,415 SNP markers and the Best Linear Unbiased Estimation (BLUE) values of the observed phenotypic traits. Simple Interval Mapping (SIM) was

performed using the three DH families of the NAM population as interactive and additive covariates with Haley and Knott (HK) regression method (Haley and Knott, 1992). This approach allows loci detection with both main effects and epistatic interactions with other loci.

Epistasis QTL mapping:

The scan-two function was utilized to calculate the LOD scores of epistatic interactions across locations and MEs. Significance thresholds were determined with 1000 permutations. To evaluate epistasis by environment interactions (EEI), we selected the markers that exceeded the significance threshold of SIM and scan-two tests. Recorded markers were utilized for further statistical analysis for each specific phenotypic trait, and recorded markers are shared in **Figure 13**. In order to avoid multiple testing problems, not all significant markers from the scan-two analysis were included in the linear mixed model. Thus, a few significant markers from the scan two analysis were selected for further analysis. The number of selected markers for EEI analysis were 11 markers for pollen and silk DAP, 11 and 14 markers for plant and ear height, 12 for stand %, and 14 for yield.

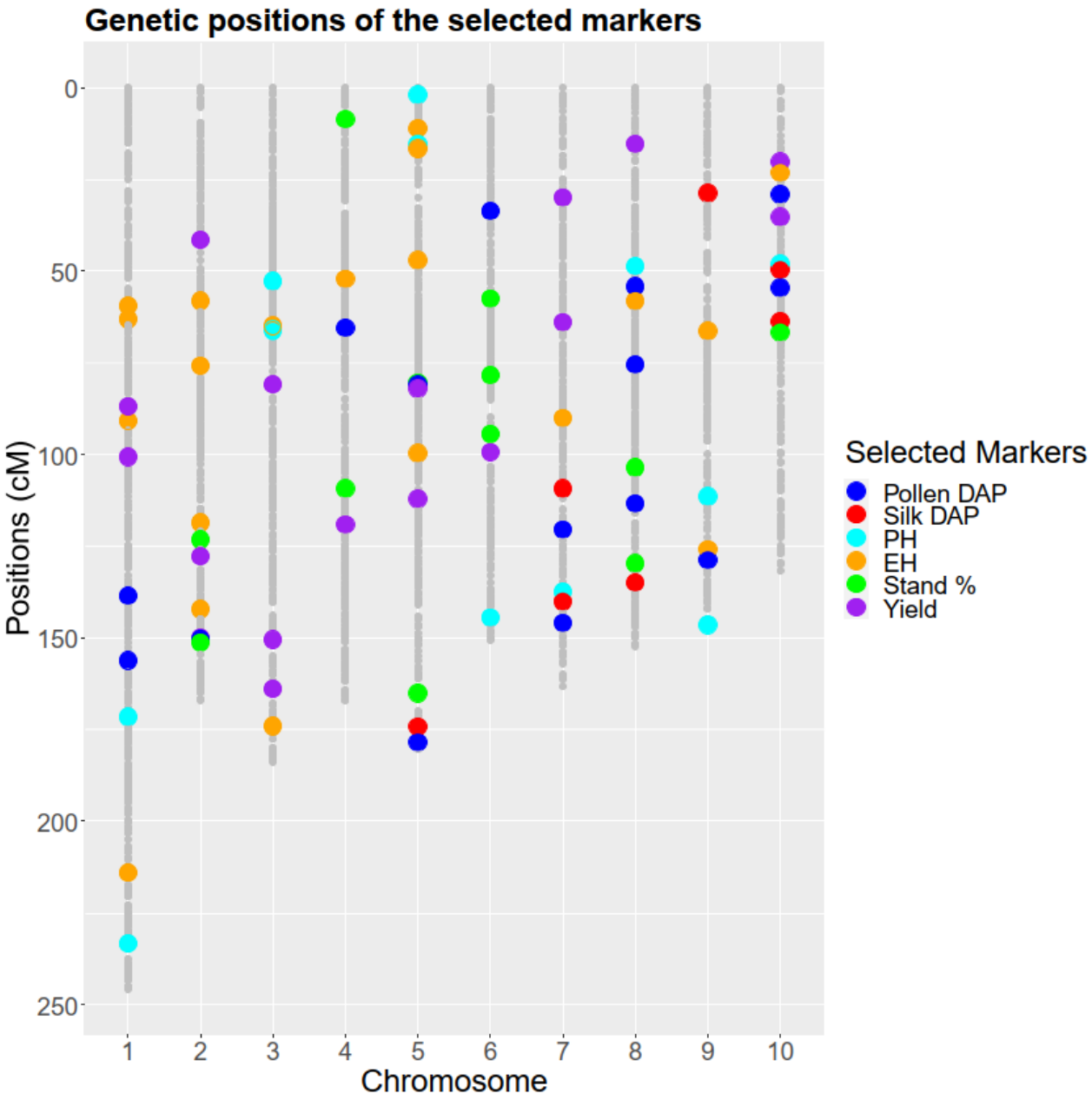


Figure 13: Selected markers for EEI analysis. Significant markers from QTL mapping and epistasis search were selected for EEI analysis.

Epistasis by environment interactions:

A linear mixed model has been created with the significant peak markers from the SIM and epistasis QTL mapping for each phenotype separately. The linear mixed model is described in **Formula 3**. Y is the phenotypic value for each ME, μ is the grand mean, βj is the effect of the jth marker, αik effect of marker i and k, qj, qi, qk are the effect of j,k, and kth markers, δjn is the

effect of the jth marker in the nth location, ηikn is the effect of i and kth marker in the nth location, and ε is the error term.

Formula 3:

$$y = \mu + G_i + ME_j + F_p + \sum_j \beta_j q_j + \sum_{\mathrm{i,k}} \alpha_{ik} q_i q_k + \sum_{\mathrm{j,n}} \delta_{jn} q_i e_n + \sum_{\mathrm{i,k,n}} \eta_{ikn} q_i q_k e_n + \varepsilon$$

Restricted maximum likelihood (REML) was used instead of maximum likelihood (ML) due to high dimensionality because of interaction terms in the model. We used false discovery rate (FDR) correction (Benjamini and Hochberg, 1995) to correct p-values for multiple testing. The accuracy of the linear mixed model was tested with a permutation test for each trait and ME. The phenotypic value is permuted for each genotype, and we select markers 1000 times randomly. The underlying reason for the permutation test is to shuffle the data and test it several times under the assumption of a null hypothesis. Because the expectation is the null hypothesis, an even distribution of p-values from 0 to 1 under this assumption is expected. However, the distribution might have false positive values under a problematic model, and the distribution of p-values might skew to the left or right of a figure. In the analysis, we selected three to eight markers for each permutation test and ran the same model 1000 times. Permutation tests aimed to answer any false positives depending on different complex models. Lastly, we created a p-value distribution for the permutation tested of each ME (**Figure 14**).

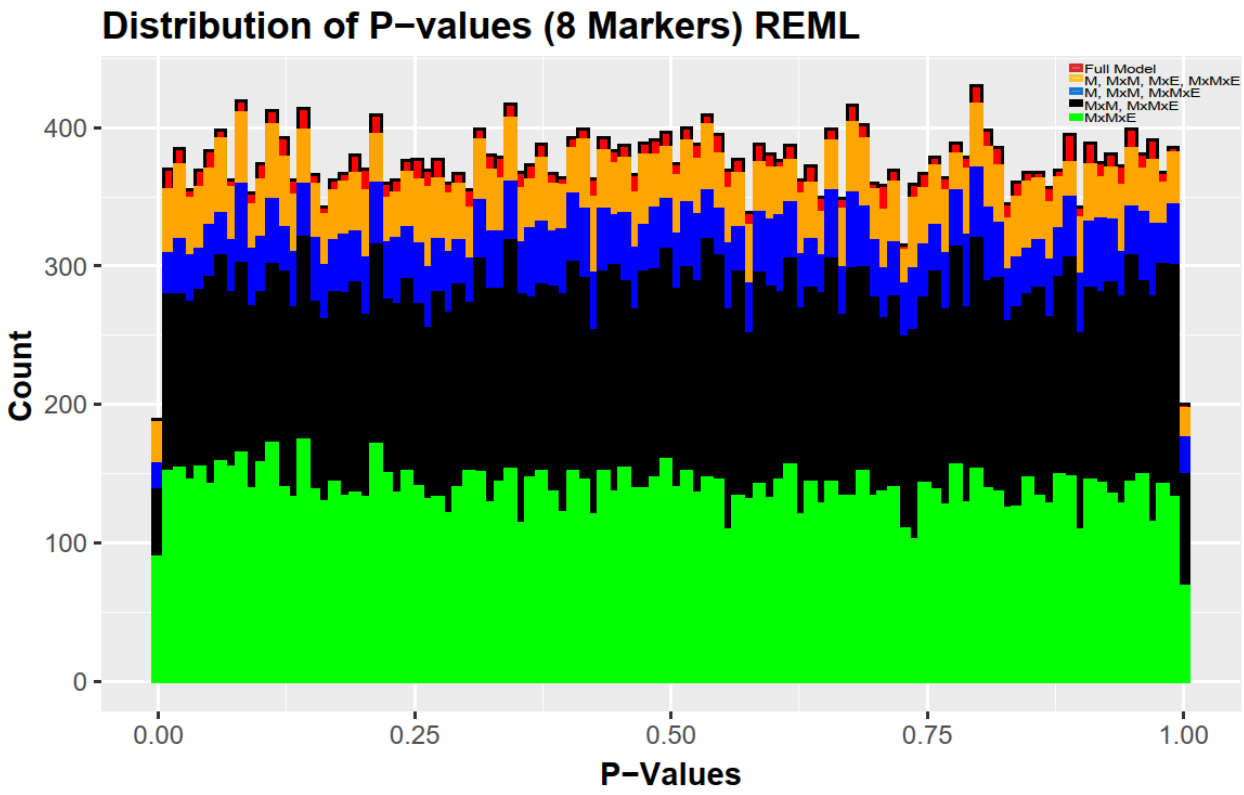


Figure 14: Distribution of p-values of 1000 permutation test results. P-value results were divided into five categories: Full model (**red**), QTL, QTL-environment interactions epistasis, and EEI (**orange**), QTL, epistasis, EEI (**blue**), epistasis and EEI (**black**), and EEI (**green**). P-values are distributed evenly for each model across the agronomic traits.

Results

Linkage mapping

The genetic data used in this study consisted of 2,415 markers. The distribution of markers across the ten chromosomes was not uniform, with chromosome 1 having the highest number of markers (365) and chromosome 10 having the lowest (142). The genetic distance between markers varied across chromosomes, with chromosome 1 having the longest total length (245.810 cM) and chromosome 10 having the shortest (131.755 cM). Each chromosome has 365, 263, 263, 275, 271, 196, 201, 252, 182, and 142 SNP markers, respectively, and the length of the chromosome from 1 to 10 in order of 245.810, 167.030, 183.892, 167.193, 180.433, 150.694, 163.312, 152.307, 147.454, 131.755 centimorgans. The markers were somewhat evenly distributed across the genetic map (**Figure 15**). We compared the physical positions to the genetic positions and found that the relationship between the two maps was mostly linear, except around the centromere, as shown in **Figure 16**. To ensure the accuracy of our genetic map, we estimated recombination frequencies est.rf(), and simulating the genotypes when the observed marker data was given with sim.geno(), and calculated the conditional genotype probabilities with calc.genoprob() functions in the qtl/r package. The markers were evenly distributed across the genetic map, although some regions had a higher density of markers. For example, chromosome 3 had a concentration of markers around 50 cM, while chromosome 5 had a concentration between 50 and 100 cM. The minimum and maximum distances between adjacent markers were 0.326 and 6.241 centimorgans, respectively. The power of a QTL mapping analysis is unaffected by the abundant markers as well as the distance between adjacent markers between 0 and 10 cM (Rebai et al., 1995; Semagn et al., 2006). Therefore, we did not apply further filtering steps to remove redundant markers.

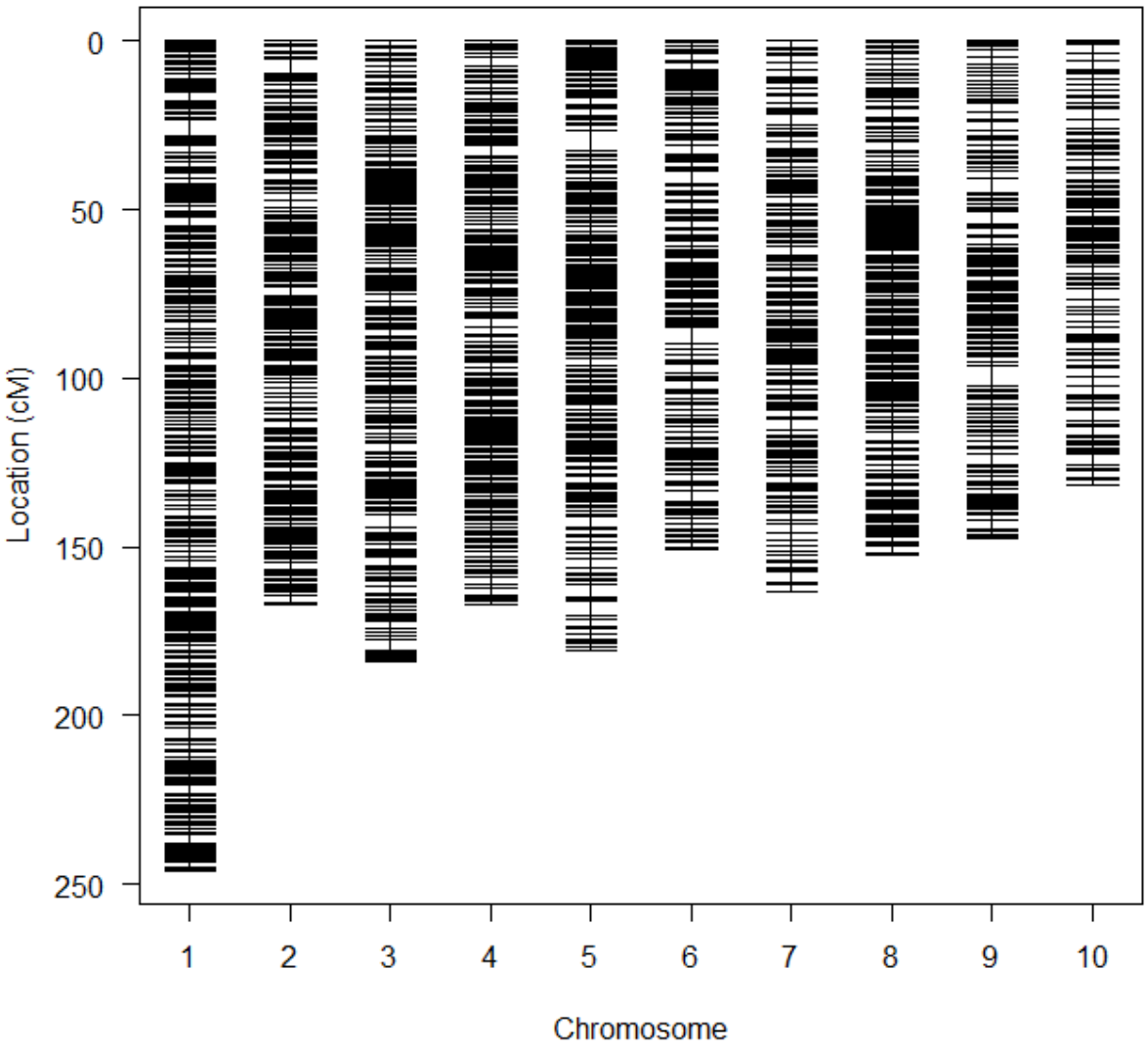
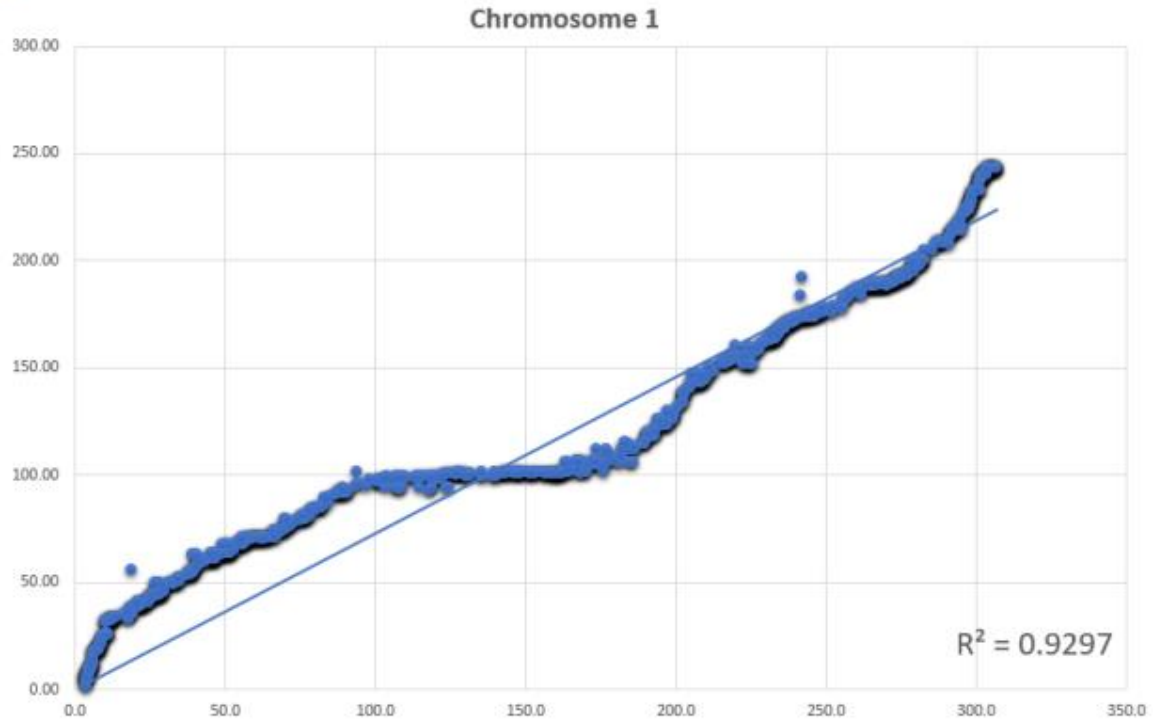
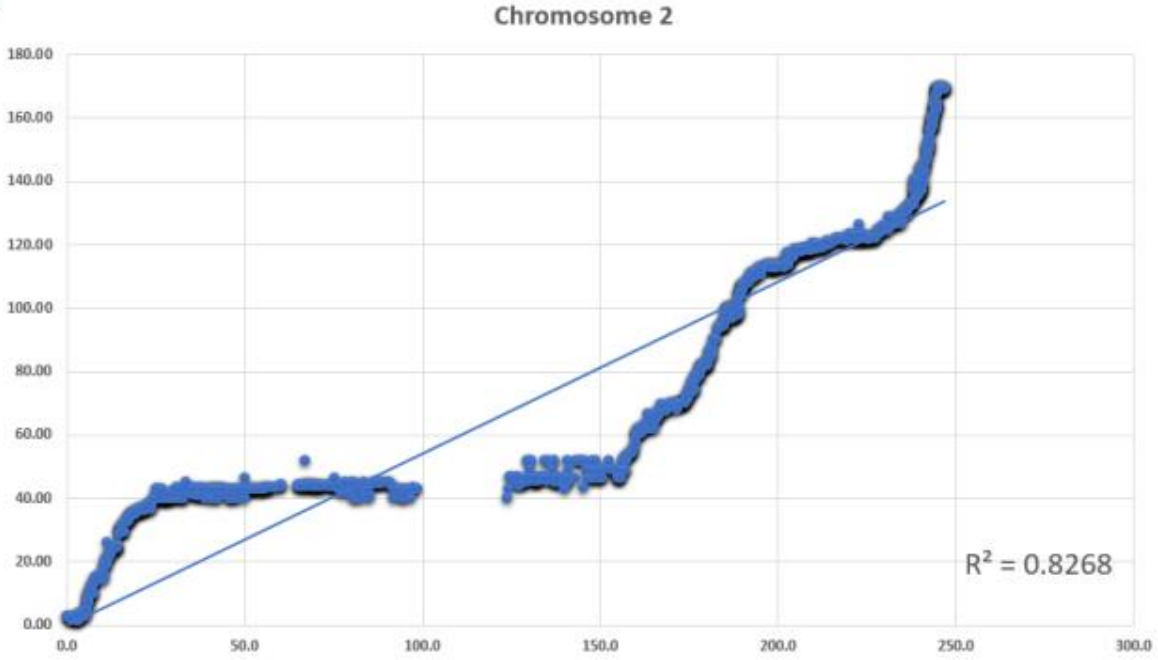


Figure 15: Genetic map of G2F NAM population. X-axes indicate the chromosomes, and the y-axes indicate the genetic location of the markers.

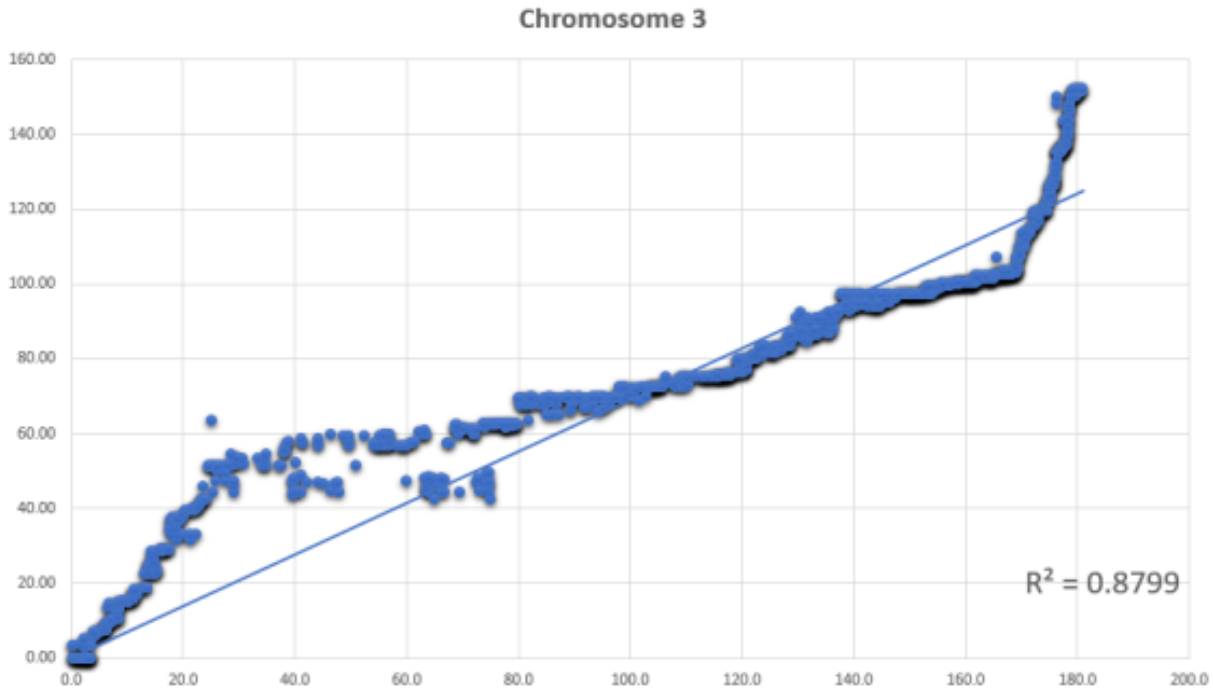
A


Chromosome 1
300.00
250.00
200.00
150.00
100.00
50.00
0.00
0.0
50.0
100.0
150.0
200.0
250.0
300.0
350.0
R² = 0.9297

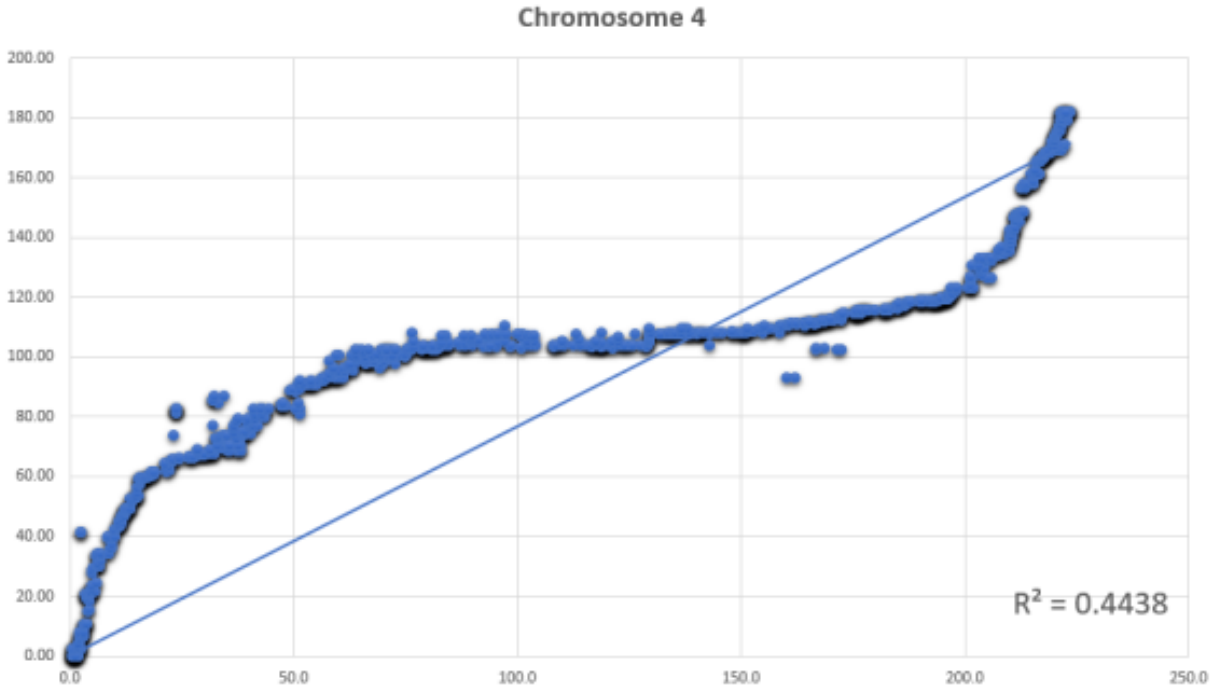
B


Chromosome 2
180.00
160.00
140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
0.0
50.0
100.0
150.0
200.0
250.0
300.0
R² = 0.8268

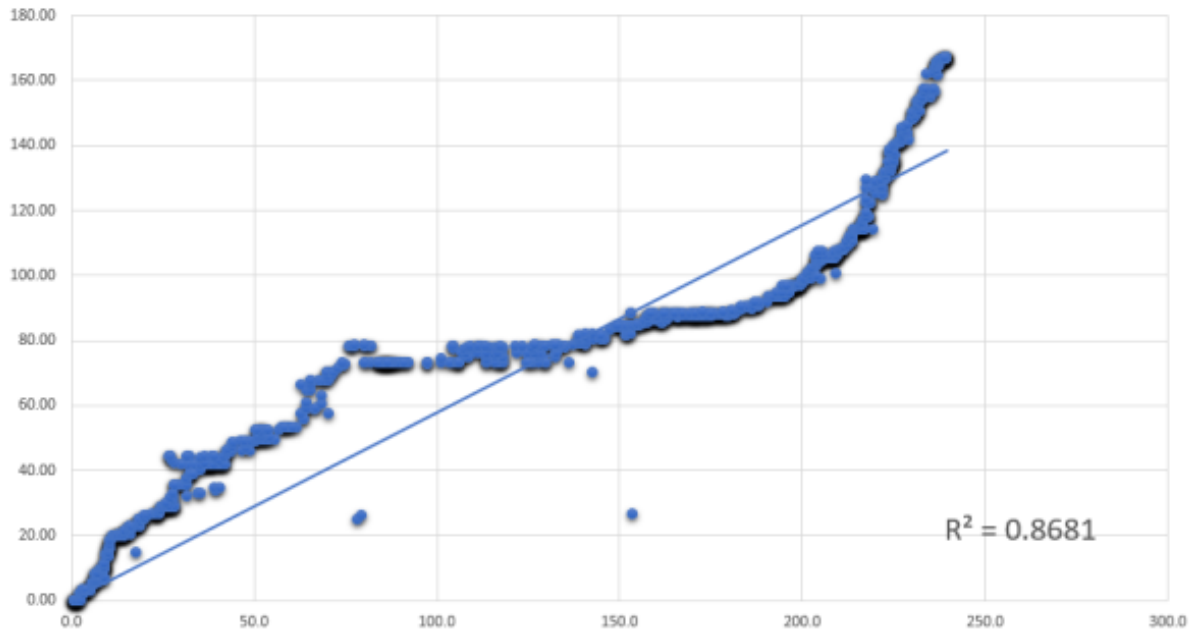
C


Chromosome 3
160.00
140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
0.0
20.0
40.0
60.0
80.0
100.0
120.0
140.0
160.0
180.0
200.0
R² = 0.8799

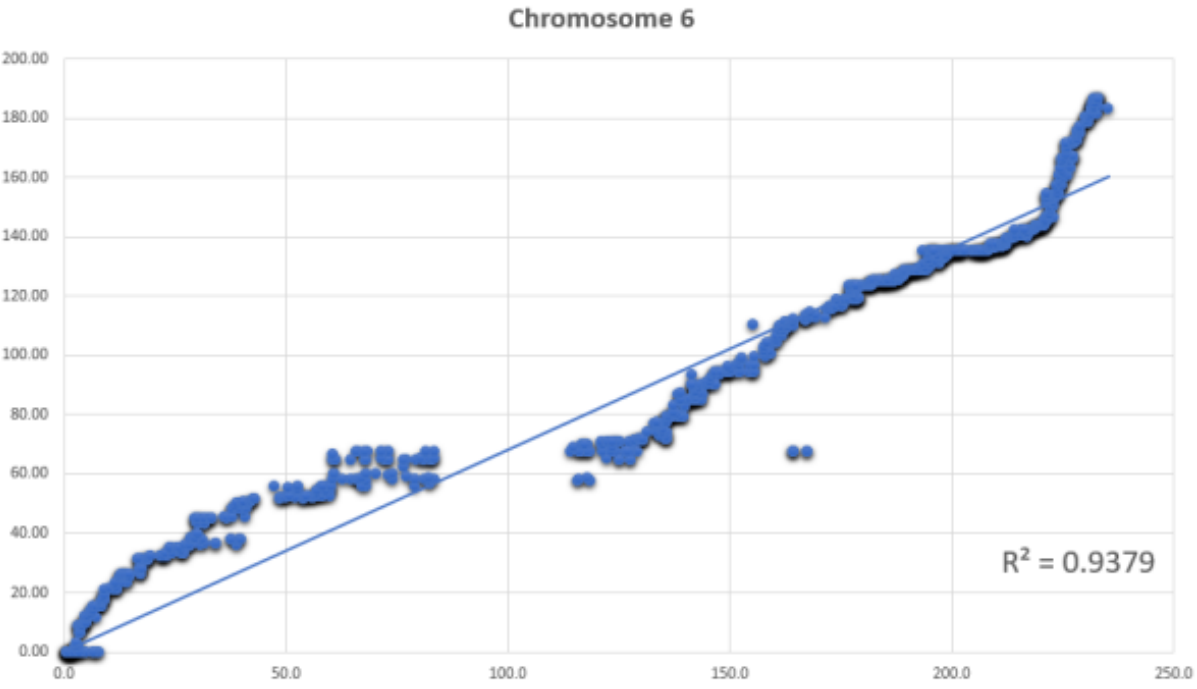
D


Chromosome 4
200.00
180.00
160.00
140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
0.0
50.0
100.0
150.0
200.0
250.0
R² = 0.4438

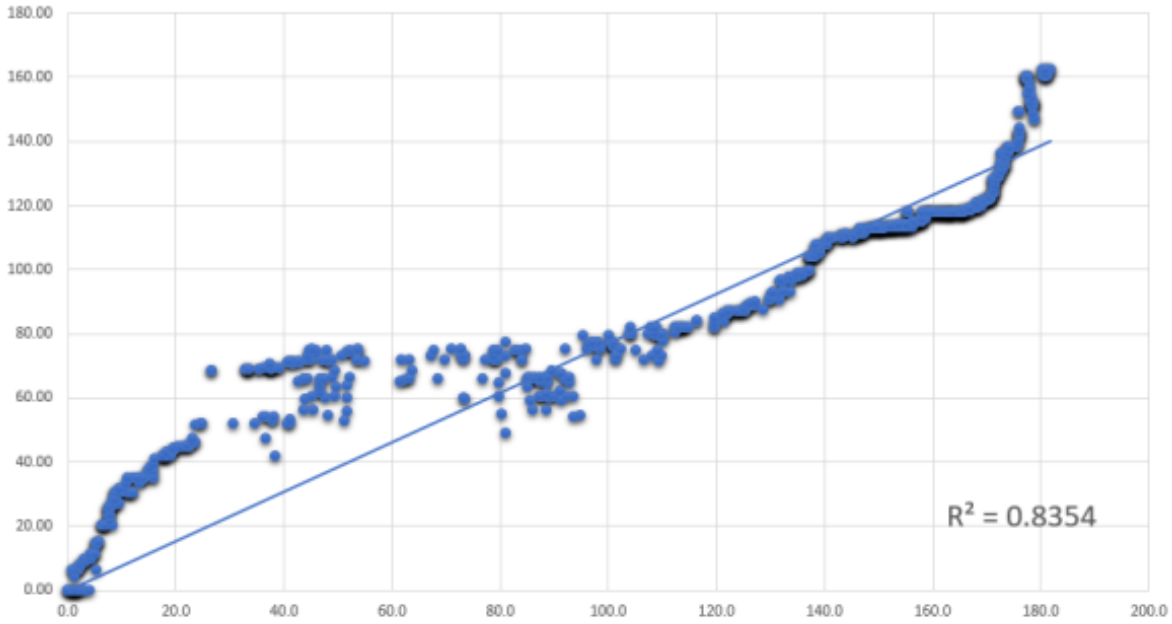
E


Chromosome 5
180.00
160.00
140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
0.0
50.0
100.0
150.0
200.0
250.0
300.0
R² = 0.8681

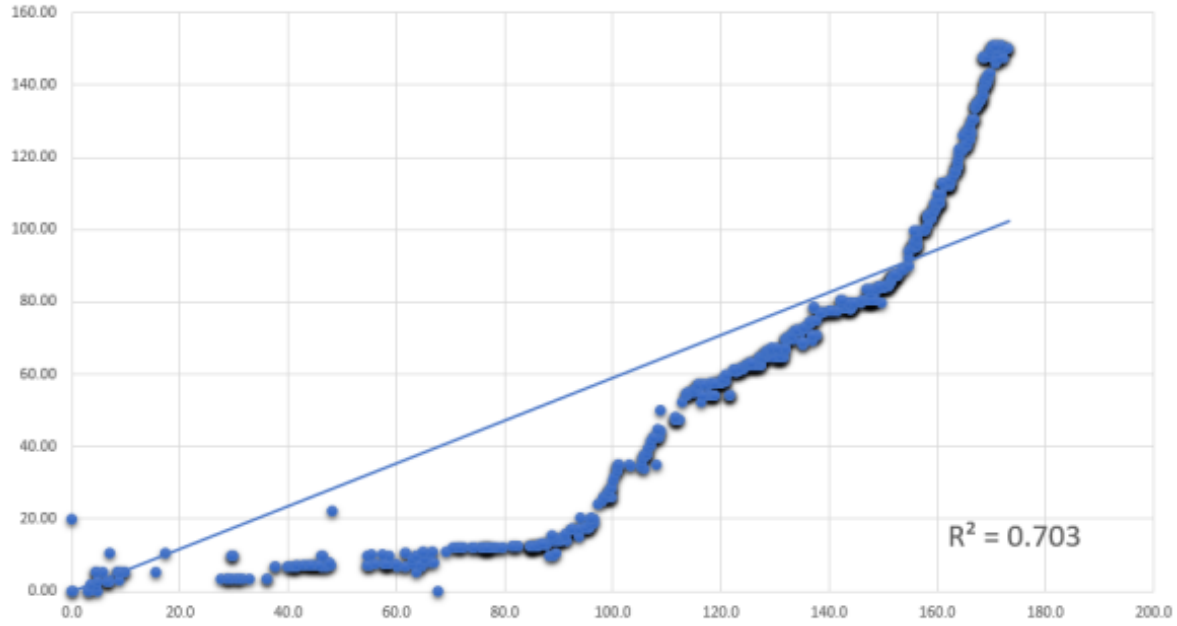
F


Chromosome 6
200.00
180.00
160.00
140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
0.0
50.0
100.0
150.0
200.0
250.0
R² = 0.9379

G


Chromosome 7
180.00
160.00
140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
0.0
20.0
40.0
60.0
80.0
100.0
120.0
140.0
160.0
180.0
200.0
R² = 0.8354

H


Chromosome 8
160.00
140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
0.0
20.0
40.0
60.0
80.0
100.0
120.0
140.0
160.0
180.0
200.0
R² = 0.703

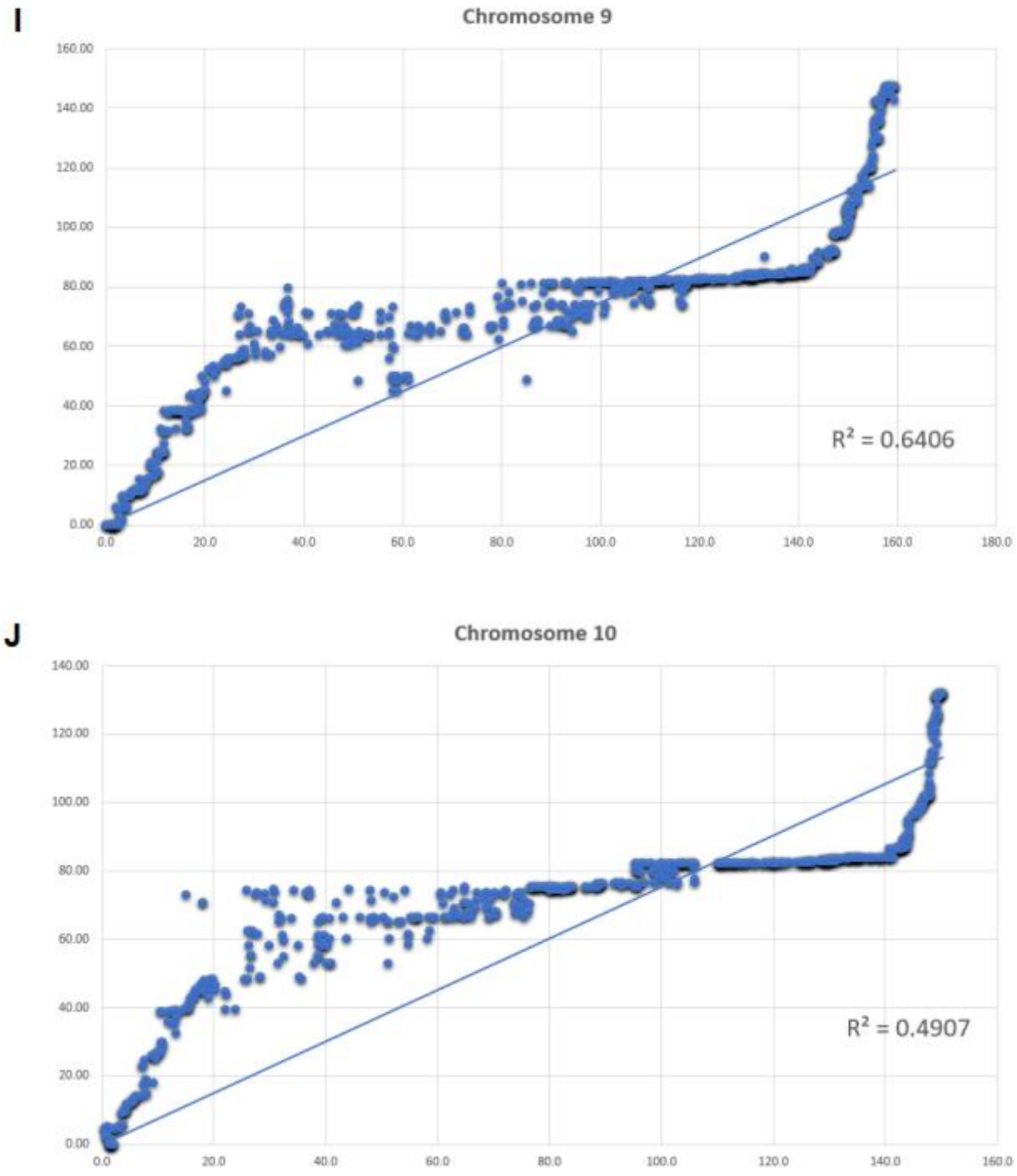


Figure 16: Physical positions and genetic positions of the SNP markers for each chromosome in maize. X-axes represent the physical positions (megabase pairs (Mbp)) of SNPs, and Y-axes represent the genetic distances of SNPs (centiMorgan (cM)) (**from A to J**).

Phenotype

Our study aimed to investigate the following phenotypes: days to anthesis for pollen (pollen DAP), days to anthesis for silk (silk DAP), plant height (PH), ear height (EH), stand percentage (stand %), and grain yield. The observations ranged from 45 to 82 days for pollen DAP, 46 to 82 days for silk DAP, 106 to 359 cm for PH, 49 to 234 cm for EH, 3 to 97% for stand %, and 13.35 to 370.70 bushels per acre for grain yield. The data for each phenotype was normally distributed. A linear mixed model was applied, and best linear unbiased estimates (BLUEs) were calculated for new phenotypic observations across all locations and mega environments, as done in previous publications (Zhang et al., 2014; Bloom, 2015). Our study observed deviations from the expected and fitted values; however, these deviations did not violate the normality assumption. The calculated BLUE values for each genotype ranged from 59.25 to 65.22 days for pollen DAP, 60.09 to 66.91 days for silk DAP, 192.4 to 240.5 cm for plant height, 83.99 to 120.36 cm for ear height, 52.36 to 79.50% for stand percentage, and 123.3 to 178.2 bushels per acre for grain yield. The pollen DAP for each ME ranged from 60.48 to 69.94 days for ME1, 58.42 to 70.05 days for ME3, 51.02 to 60.51 days for ME4, 58.79 to 68.19 days for ME5, 66.32 to 72.32 days for ME6, and 61.75 to 68.51 days for ME7. The BLUE values for each ME for silk DAP, plant height, ear height, stand percentage, and grain yield ranged from 61.34 to 71.48 days for silk DAP, 161.9 to 253.8 cm for plant height, 83.17 to 120.06 cm for ear height, 41.90 to 79.78% for stand percentage, and 115.1 to 208.4 bushels per acre for grain yield for ME1, 59.64 to 71.88 days for silk DAP, 200.6 to 261.4 cm for plant height, 91.34 to 130.03 cm for ear height, 61.13 to 82.74% for stand %, and 114.5 to 201.6 bushels per acre for grain yield for ME3, 53.45 to 63.48 days for silk DAP, 168.4 to 235.7 cm for plant height, 75.79 to 123.95 cm for ear height, 47.66 to 95.33% for stand %, and 108.7 to 190.5 bushels per acre for grain yield for ME4, 60.33 to 69.77 days for silk DAP, 183.4 to 244.7 cm for plant height, 78.64 to 135.92 cm for ear height, 35.64 to 79.37% for stand %, and 116.1 to 199.5 bushels per acre for grain yield for ME5, 66.67 to 71.40 days for silk DAP, 202.3 to 260.0 cm for plant height, 60.52 to 115.19 cm for ear height, 45.10 to 89.78% for stand %, and 89.45 to 163.82 bushels per acre for grain yield for ME6, and 63.19 to 70.02 days for silk DAP, 203.6 to 261.3 cm for plant height, 90.53 to 135.11 cm for ear height, 32.33 to 87.88% for stand %, and 139.5 to 233.5 bushels per acre for grain yield for ME7, respectively (**Figures 6 to 12**).

Correlations

The correlation analysis revealed strong associations between the pollen and silk DAP across all locations and MEs, with correlation coefficients ranging from 0.71 to 0.94 (**Figure 17**). The correlations between pollen DAP and other traits, including PH, EH, stand percentage (%), and yield, were moderate, ranging from 0.29 to 0.59, 0.35 to 0.55, -0.22 to 0.02, and -0.18 to 0.21, respectively. Similarly, the correlations between silk DAP and the other traits were moderate and consistent with those observed for pollen DAP. Furthermore, PH and EH showed a strong positive correlation across all locations and MEs, with correlation coefficients ranging from 0.61 to 0.83. The correlation between PH and yield was moderate, ranging from -0.01 to 0.47, whereas the correlation between PH and stand percentage was weak, ranging from -0.25 to 0.08. The correlation between yield, PH, EH, and stand percentage varied depending on the MEs, but generally, the relationships were moderately correlated. The correlation between yield and PH ranged from 0.21 to 0.40, whereas the correlation between yield and EH ranged from 0.07 to 0.44, and the correlation between yield and stand percentage ranged from 0.22 to 0.49. The lowest correlation was observed between yield, PH, EH, and stand percentage in ME 3, with coefficients of -0.08, -0.12, and 0.23, respectively.

Interestingly, the correlation between stand % and the other traits, including pollen DAP, silk DAP, PH, EH, and yield, was the weakest among all correlations. These findings suggest that the stand % may be influenced by factors other than the traits evaluated in this study. Overall, these results provide valuable insights into the correlations between various traits in maize and can be used to guide future breeding efforts for optimizing maize production.

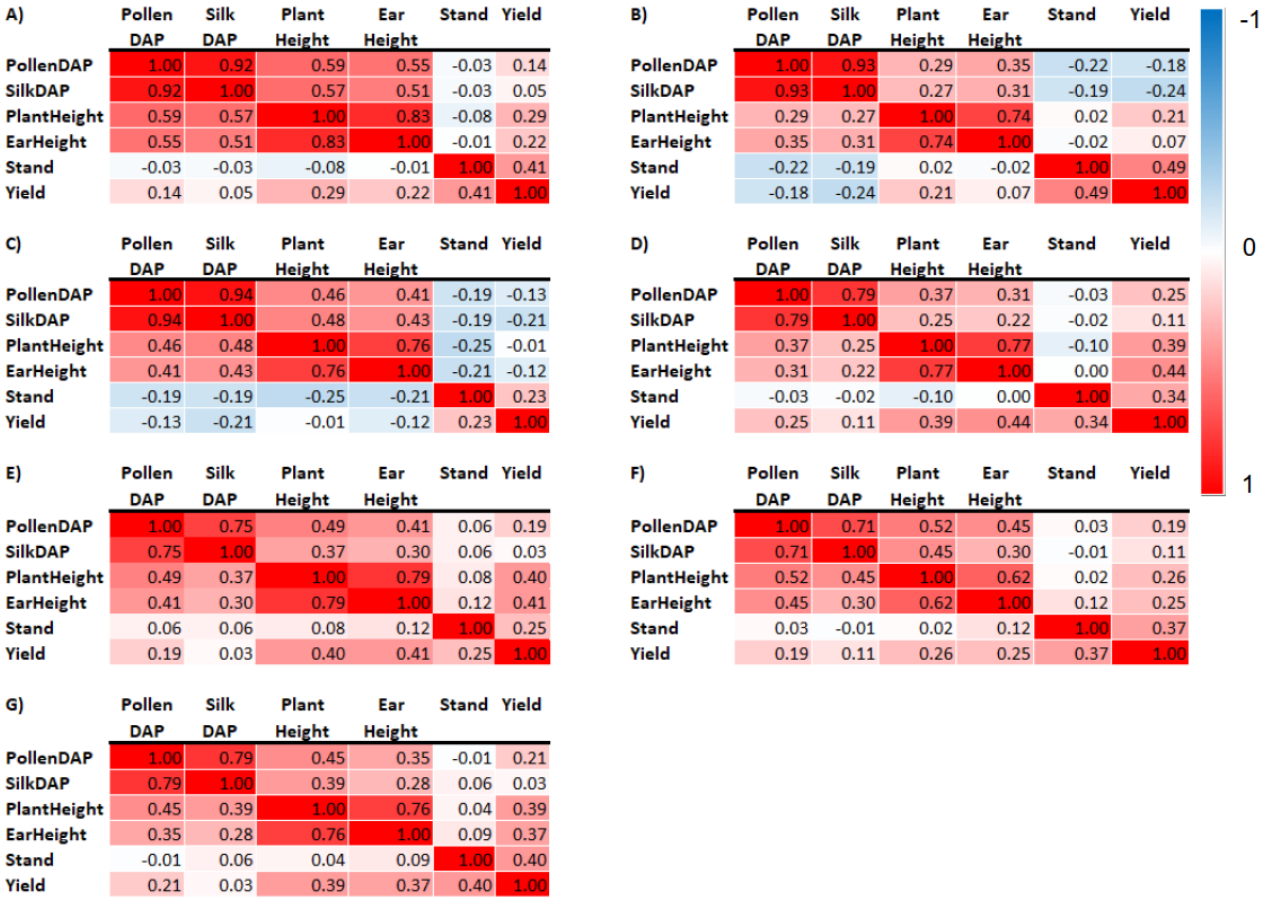


Figure 17: Correlations between the phenotypic traits across locations and mega environments (MEs). Across locations (**A**), ME1 (**B**), ME3 (**C**), ME4 (**D**), ME5 (**E**), ME6 (**F**), and ME7 (**G**).

Heritability

Table 1 presents the heritability estimates of all phenotypic traits across multiple environments (MEs). The heritability of pollen DAP ranged from 0.140 to 0.89, with the highest estimate observed in ME6 and the lowest in ME3. Moderate heritability (0.504-0.546) was observed for pollen DAP in ME2 and ME6, while relatively high heritability (0.741) was observed in ME1. Silk DAP showed a similar trend, with heritability ranging from 0.149 to 0.785 across locations and MEs, except for ME6, where the heritability was relatively low. PH and EH exhibited the highest heritability estimates among all traits, with the highest heritability of PH observed across all locations (0.93) and the lowest in ME4 (0.622). For ME1, ME3, ME5, ME6, and ME7, the heritability estimates ranged from 0.709 to 0.826 for PH. Similarly, the heritability of EH was highest across all locations (0.922) and the lowest in ME4 (0.686). The heritability estimates for

EH in ME1, ME3, ME5, ME6, and ME7 ranged from 0.651 to 0.76. Heritability of Stand % ranged between 0.482 - 0.882 across all locations, with the highest estimate observed in ME1 and the lowest in ME4. The heritability of yield ranged from 0.362 to 0.833, with the highest estimate observed in all locations and the lowest in ME4. The heritability estimates for yield varied across MEs, ranging from 0.515 to 0.656.

Table 1: Heritability of the phenotypic traits across the locations and mega environments (MEs) for pollen DAP, silk DAP, PH, EH, stand %, and yield.

	Pollen DAP	Silk DAP	Plant Height	Ear Height	Stand %	Yield
Location	**H^2**	**H^2**	**H^2**	**H^2**	**H^2**	**H^2**
All L.	0.79	0.78	0.93	0.92	0.88	0.83
ME 1	0.78	0.79	0.76	0.74	0.58	0.61
ME 3	0.14	0.15	0.71	0.69	0.48	0.58
ME 4	0.56	0.56	0.62	0.65	0.55	0.36
ME 5	0.64	0.69	0.83	0.76	0.67	0.62
ME 6	0.89	0.26	0.76	0.66	0.75	0.51
ME 7	0.84	0.78	0.73	0.68	0.63	0.66

QTL mapping

We conducted SIM separately in all locations and MEs to identify significant QTLs for each phenotype. Significant markers were identified for each trait, and the -log10 p-value of the peak markers LOD was compared across locations. We applied a 1000 permutation test with a 5% alpha value to determine the significance threshold for SIM. We calculated a 95% confidence interval (95% CI) for each analysis and selected one marker from each CI for further statistical analysis.

We used SIM to identify significant QTLs for each phenotypic trait across all locations of MEs, as shown in **Figure 18**. For pollen DAP, we applied a threshold of 4.25-4.66 LOD scores across all locations and MEs and found two significant markers in all locations: one on chromosome 1 at 138.9 cM with a LOD score of 4.68 and the other on chromosome 5 at 81.1 cM with a LOD score of 19.5. The second confidence interval (CI) of the QTL position showed significant results across

MEs, with LOD scores ranging from 10.1 to 20.49, except for ME4. In addition, one and two additional QTLs were found on chromosomes 9, 4, and 8, respectively, in ME6 and ME7 (**Table 2**).

Table 2: LOD and phenotypic variation explained by the marker (R^2) score of significant markers for each phenotypic trait and each ME given under R2 and LOD columns. 95 % confidence interval (CI) values and the peak cM of the significant markers were given under the Distance column. In addition, the name of the marker (Marker) under which chromosome (Chr) is also specified.

R2 (%)							LOD							Distance (cM)				
Joint	ME7	ME6	ME5	ME4	ME3	ME1	Joint	ME7	ME6	ME5	ME4	ME3	ME1	Interval	Peak	Marker	Chr	Trait
6.80	3.23	6.11	3.62	2.62	4.60	2.15	4.68	2.18	4.19	2.45	1.8	3.13	1.44	129.110 - 165.062	138.92	S1_203118552	1	Pollen DAP
4.45	6.65	4.11	2.61	4.05	1.24	4.32	2.52	4.57	2.5	1.59	2.2	0.73	2.3	49.035 - 90.219	57.861	S4_159938092	4	
25.44	26.38	21.78	13.79	5.08	13.96	12.21	19.5	20.5	17.2	17.2	10	10.4	11.7	68.013 - 82.066	81.085	S5_142437768	5	
2.14	6.88	5.50	3.66	1.96	1.93	3.54	3.09	5.43	1	2.48	2.6	1.3	2.39	50 - 61.448	52.624	S8_63479448	8	
3.47	1.57	9.30	3.58	2.18	2.16	0.12	2.35	4.37	6.49	6.49	1.5	1.45	0.08	34.327 - 63.098	60.484	S9_48378188	9	
1.83	1.51	1.51	1.86	7.28	0.18	1.40	1.22	1.01	1.03	1.19	3.1	0.12	0.94	113.094 - 192.517	187.29	S1_267003964	1	Silk DAP
19.84	13.91	13.91	5.48	2.97	15.25	10.23	14.7	9.95	16.3	9.86	3.5	11	7.17	67.686 - 85.661	81.085	S5_144805011	5	
6.98	3.44	3.44	5.00	3.43	3.63	3.12	4.81	2.33	3.85	2.38	1.4	2.45	2.11	24.846 - 58.850	28.769	S9_10972540	9	
9.69	2.05	2.05	11.44	5.72	1.96	6.07	6.77	1.38	1.14	2.62	1.4	1.32	4.16	45.442 - 69.302	63.744	S10_125572892	10	
12.21	10.16	10.10	7.46	6.67	9.96	9.97	8.66	7.12	7.08	5.15	4.6	6.97	6.98	165.062 - 171.272	168.33	S1_236244055	1	Plant Height
16.66	10.50	6.68	17.20	6.60	9.52	11.57	12.1	7.37	4.59	12.5	4.5	6.65	8.17	71.773 - 77.002	74.061	S3_157440022	3	
7.53	4.92	6.80	4.61	6.83	4.64	5.61	5.2	3.35	4.68	3.14	4.7	3.15	3.84	0 - 102.003	1.634	S5_1564701	5	
3.28	2.74	3.63	7.08	2.57	1.85	0.79	2.21	1.85	2.45	4.88	1.7	1.24	0.52	66.701 - 118.668	113.44	S7_163323962	7	
6.14	8.46	6.82	4.54	3.74	2.62	5.58	4.21	5.87	4.69	3.09	2.5	1.76	3.82	50.005 - 65.365	60.463	S8_108101571	8	
1.81	1.29	7.17	0.38	1.59	1.58	1.07	1.22	0.86	4.94	0.25	1.1	1.06	0.72	44.461 - 142.531	62.771	S9_106737481	9	
6.28	3.49	5.78	2.46	3.62	9.70	5.52	4.31	2.36	3.95	1.65	2.5	6.78	3.77	41.52 - 54.593	48.383	S10_64147115	10	
6.02	3.66	3.76	3.36	3.26	4.63	6.98	4.12	2.48	2.54	2.27	2.2	3.15	4.81	154.277 - 205.921	167.68	S1_235652231	1	Ear Height
12.15	7.90	3.83	10.82	6.26	5.80	14.91	8.61	5.46	2.6	7.61	4.3	3.97	10.7	62.622 - 90.730	72.1	S3_152447071	3	
7.43	4.87	2.71	7.11	7.30	4.08	4.35	5.13	3.32	1.83	4.9	5	2.77	2.96	0 - 120.307	98.734	S5_178308576	5	
4.51	5.59	5.55	2.00	8.30	1.98	5.13	3.07	3.82	3.79	1.34	5.8	1.33	3.5	39.231 - 57.216	47.729	S9_19527846	9	
6.45	6.89	7.79	4.66	5.77	3.50	3.75	4.43	4.74	5.39	3.17	4	2.36	2.54	46.095 - 68.001	66.366	S9_104861090	9	
4.06	1.57	2.32	1.68	3.46	3.84	5.44	2.75	1.05	1.56	1.12	2.3	2.6	3.72	0 - 52.632	8.829	S10_3652415	10	
7.80	7.62	3.95	5.26	5.00	0.58	5.99	5.4	5.27	2.68	3.59	3.4	0.38	4.1	89.882 - 147.078	95.111	S2_180671323	2	Yield
5.86	8.58	2.50	8.11	5.21	0.41	3.17	4.01	5.96	1.68	5.62	3.6	0.28	2.14	113.742 - 131.063	121.91	S2_201906553	2	
6.07	5.58	1.09	1.22	2.32	1.98	7.91	4.16	3.82	0.73	0.82	1.6	1.33	5.47	147.601 - 182.905	173.75	S3_225468828	3	
4.86	5.72	1.34	3.78	7.65	0.39	2.45	3.31	3.91	0.9	2.56	5.3	0.26	1.65	66.379 - 108.868	82.066	S5_154061900	5	
7.55	4.47	0.71	1.11	1.77	4.06	2.57	5.22	3.04	0.47	0.74	1.2	2.75	1.73	55.59 - 134.361	68.663	S7_128747685	7	
3.76	2.05	0.73	2.41	1.63	9.16	2.45	2.54	1.37	0.49	1.62	1.1	6.39	1.65	67.326 - 95.759	91.183	S8_155128571	8	

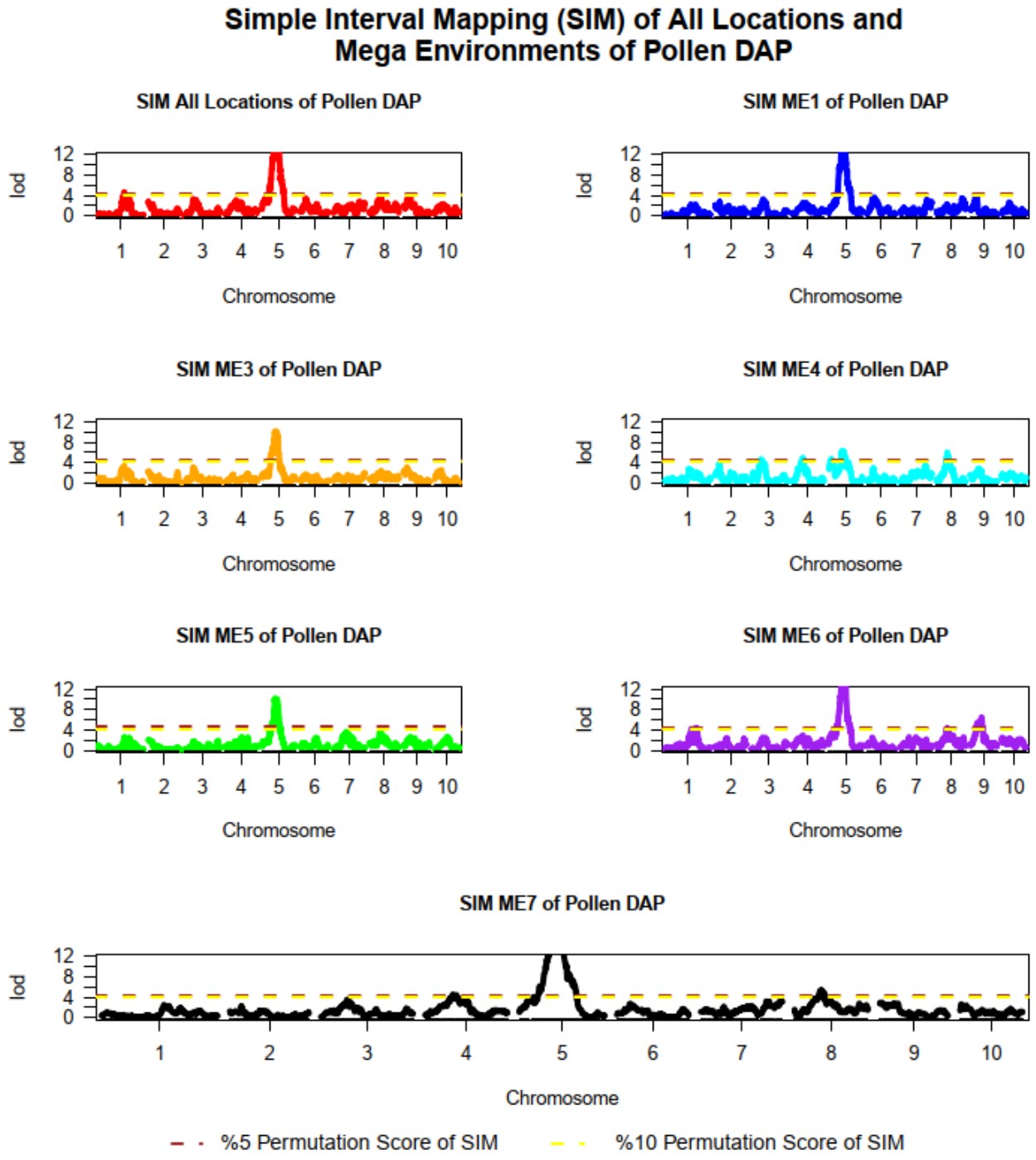


Figure 18: Simple interval mapping (SIM) of pollination days after plantations (pollen DAP) across locations and for each mega environment (MEs). The significant threshold is calculated by permutation test for each analysis separately.

Significant markers and interactions between markers were identified for the phenotypic trait of silk DAP (**Figure 19, Table 2**) across locations and MEs. Three key markers were detected on

chromosomes 5, 9, and 10 at positions 81.1 cM, 28.8 cM, and 63.7 cM, respectively, across all locations. The LOD scores of each marker were 14.7, 4.81, and 6.77, respectively. The 90% CI of each significant marker on chromosomes 5, 9, and 10 intersect with other MEs. For example, the marker on chromosome 5 was also identified in ME1, ME3, ME6, and ME7. Additionally, the marker on chromosome 10 across all locations was found in ME1, ME4, and ME5. A marker on chromosome 1, at position 187 cM, was only identified in ME4. The SIM LOD threshold for silk DAP for all locations and each ME was 4.36, 4.45, 4.33, 4.74, 4.94, 4.40, and 4.52, respectively.

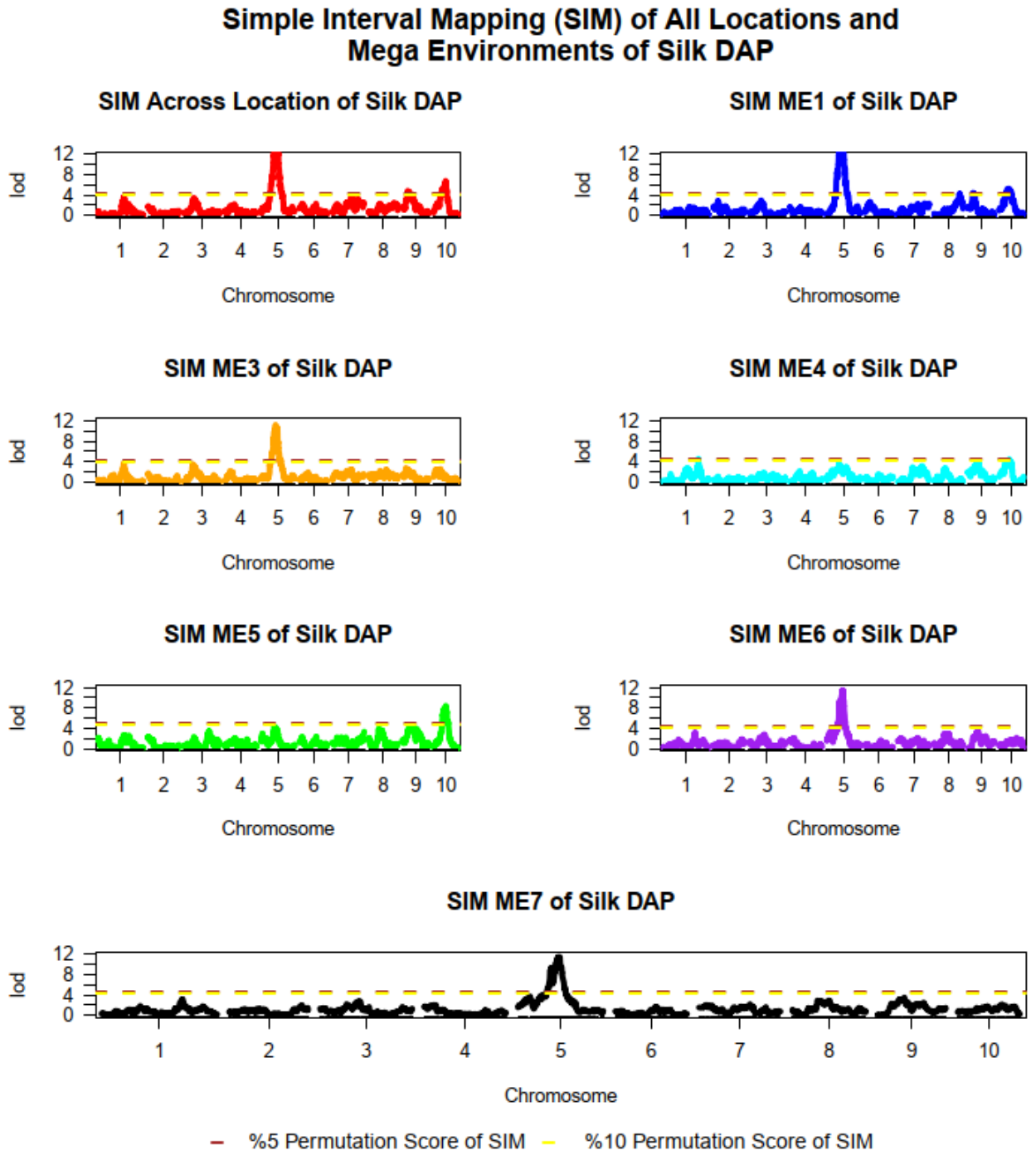


Figure 19: Simple interval mapping (SIM) of silk days after plantations (silk DAP) across locations and for each mega environment (MEs). The significant threshold is calculated by permutation test for each analysis separately.

Several significant markers were identified in chromosomes 1, 3, and 5 for PH at positions 168.33 cM, 74.06 cM, and 1.63 cM across locations (**Figure 20, Table 2**). Similar significant markers were

found in the same regions for chromosomes 1 and 3 across different MEs, except for chromosome 5, which showed significance for ME4 and ME6. Location-specific QTLs were found in ME3, ME5, ME6, and ME7, with chromosomes 9 and 10 on ME3, chromosome 7 on ME7, chromosomes 8 and 9 on ME6, and chromosome 8 on ME7.

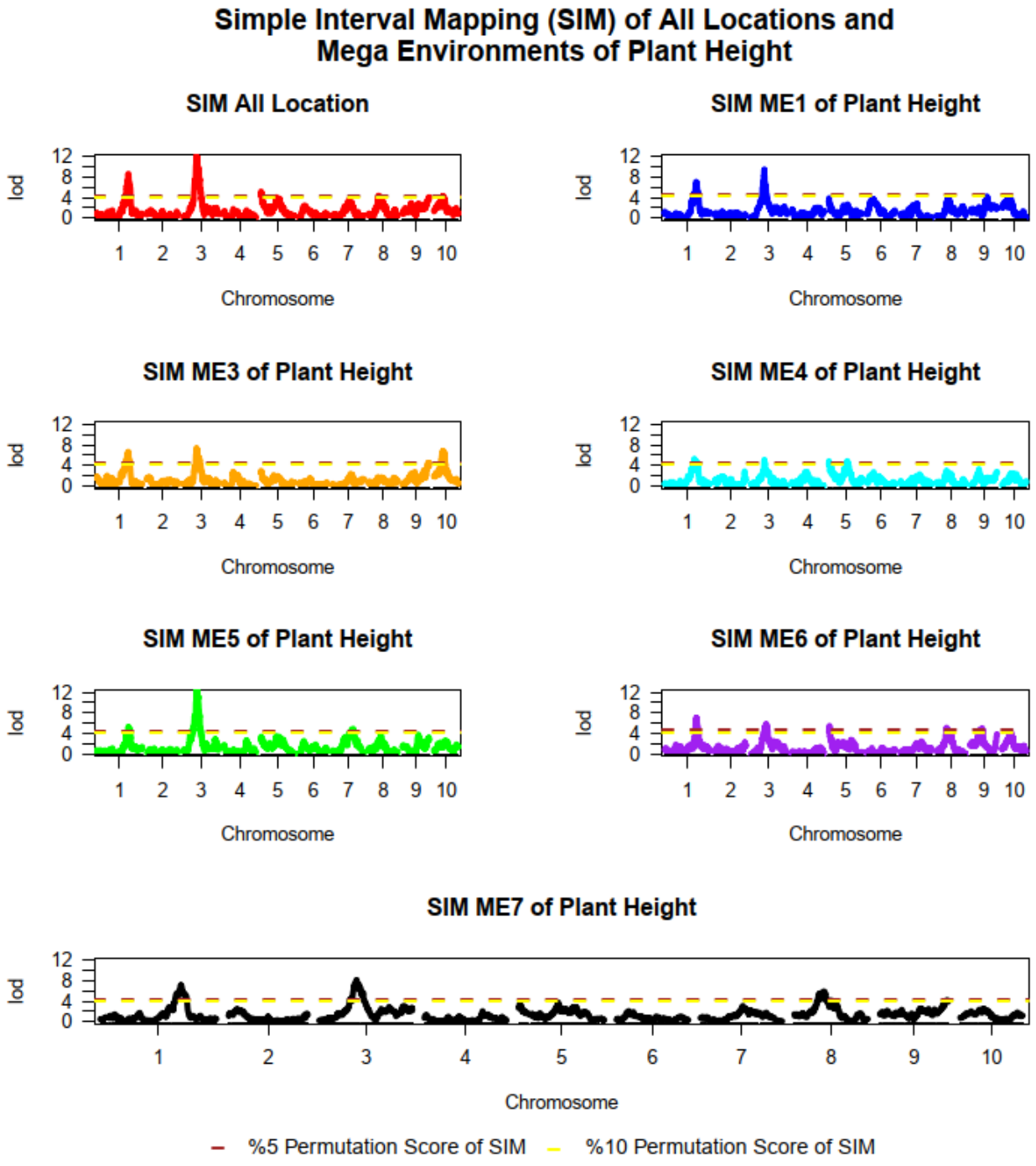


Figure 20: Simple interval mapping (SIM) of plant height (PH) across locations and for each mega environment (MEs). The significant threshold is calculated by permutation test for each analysis separately.

Two significant markers on chromosomes 3 and 5 at positions 72.1 and 98.7 cM were found for the EH observations across all locations (**Figure 21, Table 2**). Significant QTLs were identified in

each ME as follows: chromosomes 1, 3, and 5 for ME1; chromosomes 3 and 10 for ME3; chromosomes 5 and 9 for ME4; chromosomes 3 and 5 for ME5; chromosomes 1 and 9 for ME6; and chromosomes 3 and 9 for ME7.

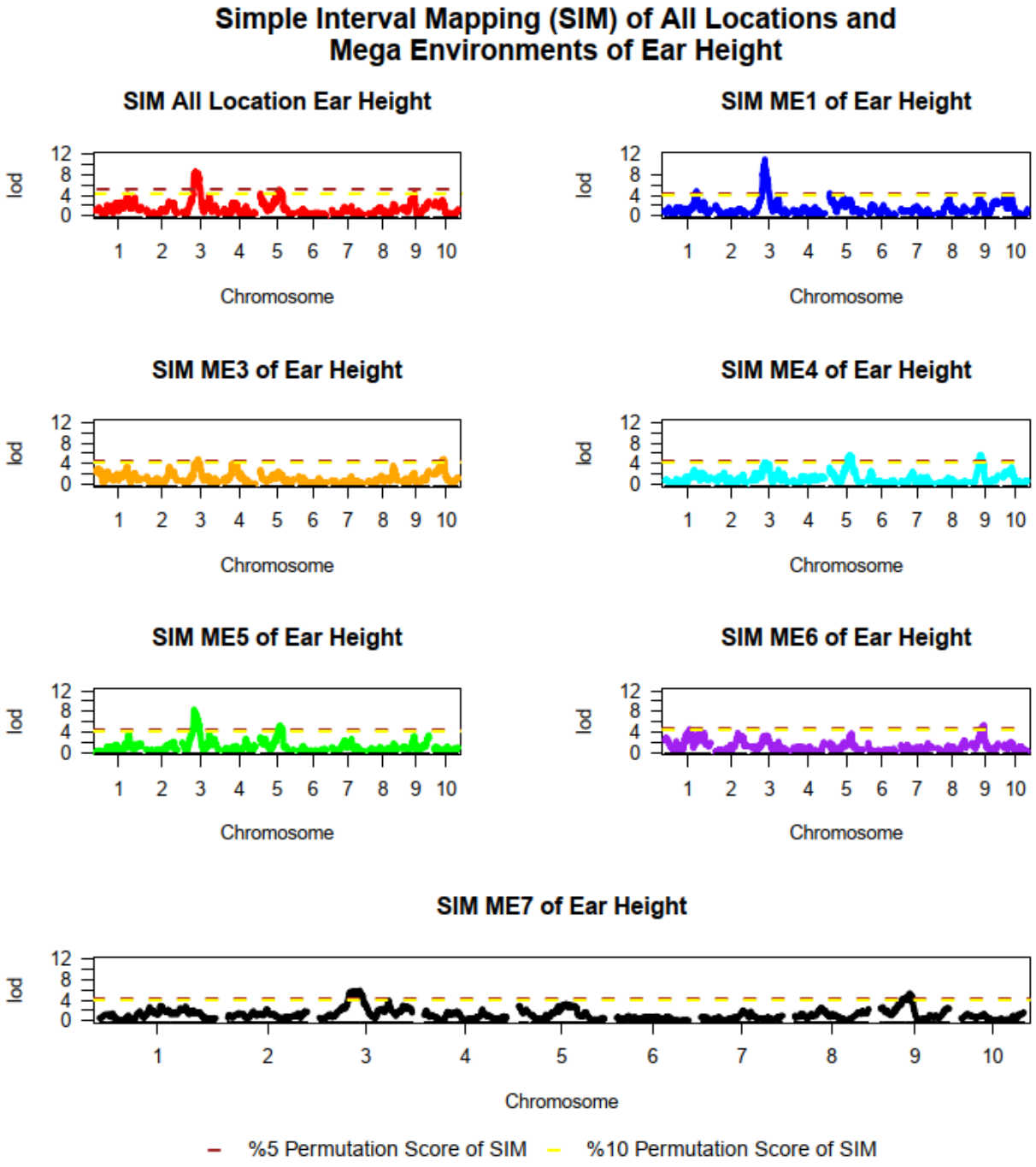


Figure 21: Simple interval mapping (SIM) of ear height (EH) across locations and for each mega environment (MEs). The significant threshold is calculated by permutation test for each analysis separately.

We found one ME-specific significant peak for stand percentage, chromosome 8, positions 98.4 on ME4 (**Figure 22, Table 2**).

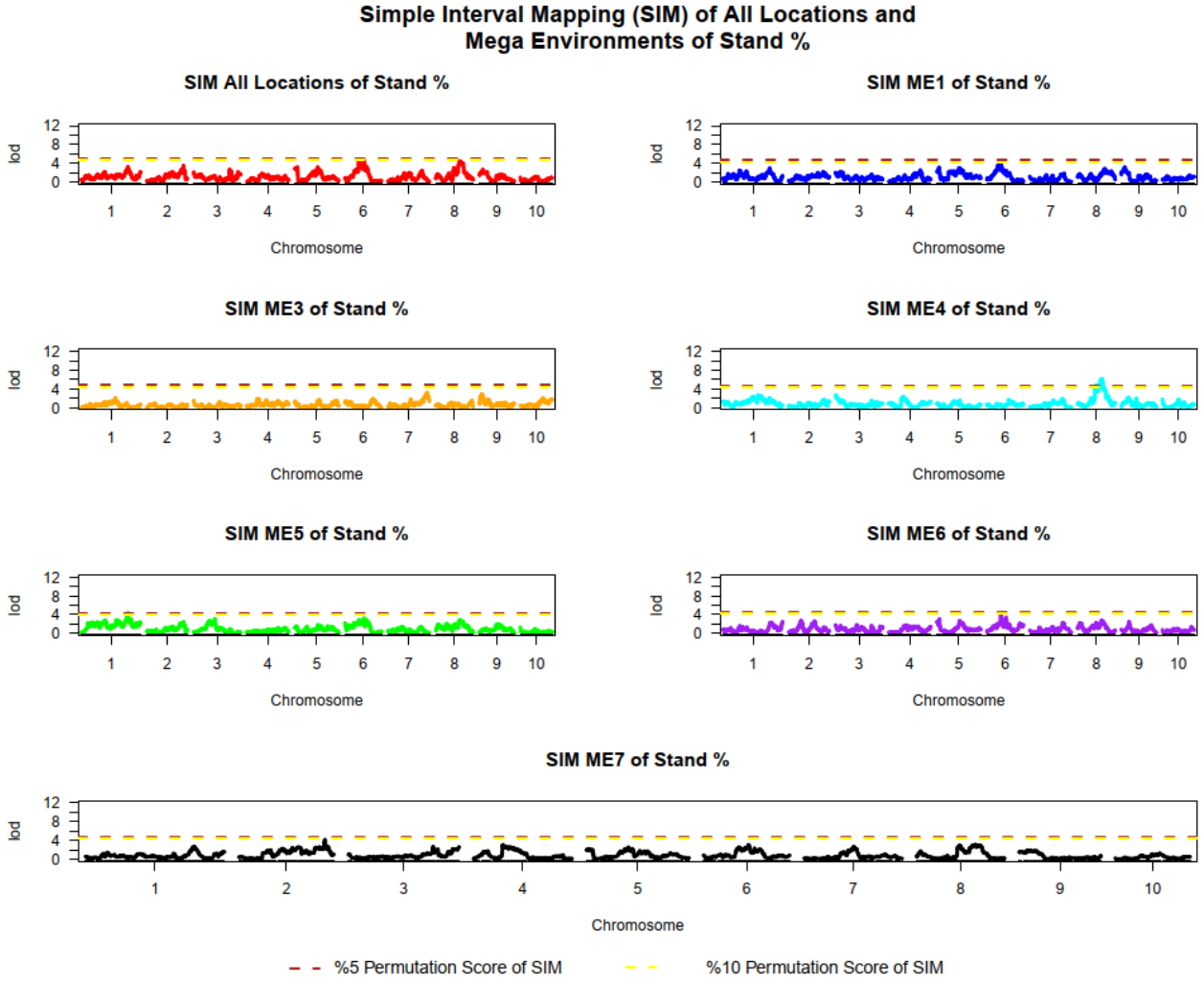


Figure 22: Simple interval mapping (SIM) of stand percentage (stand %) across locations and for each mega environment (MEs). The significant threshold is calculated by permutation test for each analysis separately.

Lastly, we investigated yield and identified several significant peaks across locations and MEs. Two markers were specific to MEs: on chromosome 3 at position 174 and chromosome 8 at position 91.2 for ME1 and ME3 (**Figure 23, Table 2**). Moreover, a marker on chromosome 2 at position 95.1 was significant across all locations and in ME5 and ME7.

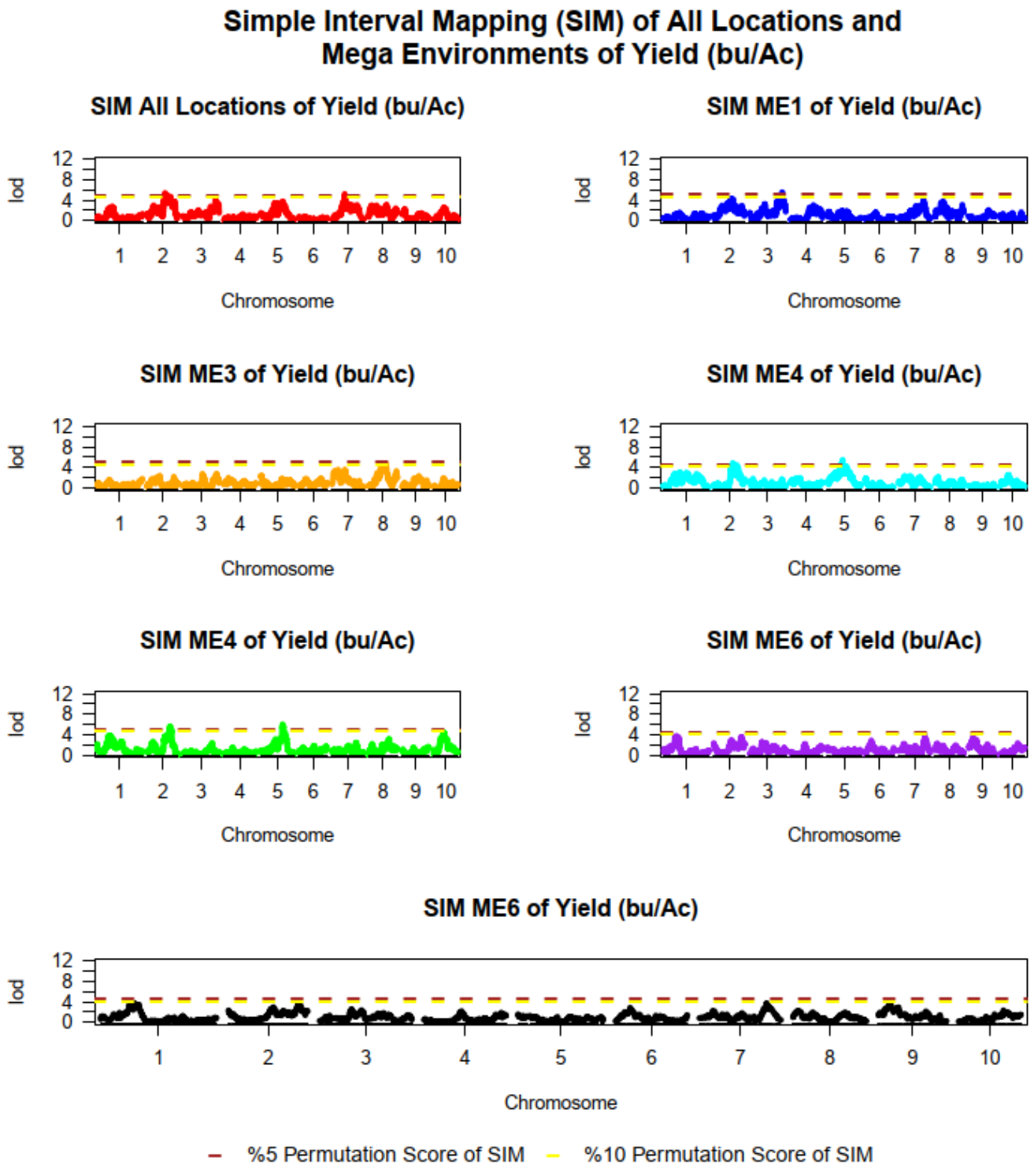


Figure 23: Simple interval mapping (SIM) of yield across locations and for each mega environment (MEs). The significant threshold is calculated by permutation test for each analysis separately.

Epistasis QTL mapping

We performed a 1000 permutation test and applied a 5% significance threshold for SIM and the two-dimensional QTL search. The threshold for significant two-dimensional QTL for epistatic interactions was 5.29-5.44 for all locations and MEs. Significant epistatic interactions were found between chromosomes 7 and 10 for all locations, chromosomes 2 and 8 for ME1, chromosomes 2 and 9 for ME4, and chromosomes 2 and 7, and chromosomes 7 and 10 for ME6 (**Table 3**). No significant epistatic interactions were found for ME3, ME5, and ME7.

Table 3: Significant epistatic interactions between markers on pollen DAP, silk DAP, PH, EH, stand %, and yield. LOD full is the full model of two QTLs and the interaction between the QTLs, fv1 is the conditional-interactive LOD score which indicates evidence for a second QTL allowing for the possibility for epistasis, LOD interactions is the indication of epistasis QTL, LOD add is the additive QTL on both chromosomes. LOD av1 is the evidence for the second QTL position assuming no epistatic interactions.

Pollen DAP	**chr**	**position 1**	**position 2**	**lod full**	**lod fv1**	**lod interaction**	**lod add**	**lod av1**
All Loc.	7 and 10	128	53.9	7.36	4.5	**5.32**	2.04	-0.82
ME 1	2 and 8	19.9	77.1	7.67	4.95	**6.06**	1.62	-1.11
ME 4	2 and 9	149	80.1	6.88	4.52	**5.84**	1.05	-1.32
ME 6	2 and 7	135.3	64.4	6.52	3.89	**5.46**	1.05	-1.57
	7 and 10	133.1	53.9	9.42	6.79	**6.24**	3.18	0.56
Silk DAP	**chr**	**position 1**	**position 2**	**lod full**	**lod fv1**	**lod interaction**	**lod add**	**lod av1**
All Loc.	1 and 3	18	14.6	6.11	2.48	**5.53**	0.58	-3.06
ME 1	1 and 3	53.28	96.6	6.55	4.52	**6.33**	0.22	-1.81
	3 and 8	3.66	147.1	8.31	5.47	**5.63**	2.68	-0.16
ME 7	1 and 2	27.8	133.7	6.57	3.25	**5.28**	1.29	-2.03
	1 and 7	162.4	154	8.99	4.08	**5.7**	3.28	-1.62
	2 and 6	93.8	93.5	7.43	4.11	**5.52**	1.91	-1.41
	2 and 9	15	128.5	7.89	4.32	**5.55**	2.35	-1.23
	6 and 9	68.6	22.6	9.25	5.67	**6.04**	3.21	-0.36
	8 and 9	30.4	20.9	9.18	5.6	**5.92**	3.26	-0.32
Plant Height	**chr**	**position 1**	**position 2**	**lod full**	**lod fv1**	**lod interaction**	**lod add**	**lod av1**
ME 3	1 and 2	220.3	75.8	6.7	-0.29	**5.47**	1.24	-5.76
	3 and 7	131.6	46.4	7.22	-0.06	**5.44**	1.78	-5.5
	4 and 6	101.7	58.2	8.54	6.13	**7.1**	1.44	-0.98

	8 and 10	31.7	25.8	9.59	2.81	**5.64**	3.95	-2.83
ME 4	1 and 8	95.4	111.1	6.05	0.7	**5.57**	0.48	-4.87
	5 and 10	123.9	85	8.37	3.26	**5.85**	2.51	-2.59
	6 and 7	143.2	145.8	7.24	4.57	**6.13**	1.11	-1.56
ME 5	1 and 7	208.9	120.6	10.36	5.13	**6.07**	4.29	-0.94
	2 and 7	141.8	104	11.13	6.25	**5.97**	5.16	0.28
ME 6	3 and 4	125.7	86	8.35	2.52	**6.27**	2.08	-3.75
ME 7	1 and 2	198.7	36.3	7.27	0.15	**5.48**	1.79	-5.33
	5 and 8	16.7	49	11.51	5.64	**5.44**	6.07	0.19
Ear Height	**chr**	**position 1**	**position 2**	**lod full**	**lod fv1**	**lod interaction**	**lod add**	**lod av1**
ME 3	3 and 10	6.1	118.7	7.8	2.75	**6.01**	1.78	-3.27
ME 6	1 and 4	113.7	47.4	10.01	5.45	**6.19**	3.82	-0.74
	8 and 10	49.4	72.6	6.98	4.57	**5.6**	1.37	-1.03
ME 7	1 and 7	9.15	138.6	7.88	4.94	**6.19**	1.69	-1.26
Stand %	**chr**	**position 1**	**position 2**	**lod full**	**lod fv1**	**lod interaction**	**lod add**	**lod av1**
All Loc.	2 and 6	123.5	93.5	12.09	7.53	**7.98**	4.12	-0.45
	2 and 7	131.1	105.9	9.99	6.52	**7.75**	2.24	-1.23
ME 3	1 and 6	141.86	111.1	10.46	7.7	**6.83**	3.63	0.88
	2 and 5	90.21	82.1	8.88	7.22	**7.06**	1.82	0.15
ME 5	1 and 9	52.63	0.654	7.3	4.75	**6.08**	1.22	-1.33
	6 and 9	84.65	95.459	7.98	4.02	**6.16**	1.81	-2.14
ME 7	1 and 6	76.16	7.19	9.86	5.38	**7.3**	2.56	-1.93
	3 and 9	155.45	101.34	10.04	6.36	**6.66**	3.38	-0.3
	4 and 6	164.74	57.85	12.36	7.87	**8.93**	3.43	-1.06
	4 and 7	0	170.66	7.09	3.71	**6.41**	0.67	-2.7
Yield (bu/Ac)	**chr**	**position 1**	**position 2**	**lod full**	**lod fv1**	**lod interaction**	**lod add**	**lod av1**
All Loc.	2 and 5	9.48	68	8.85	3.45	**5.5**	3.35	-2.05
	6 and 8	99.36	58.5	7.31	4.2	**5.56**	1.76	-1.36
	7 and 8	30.1	15.4	10.87	5.65	**7.05**	3.81	-1.41
ME 1	1 and 6	19.3	28.1	7.26	5.64	**6.18**	1.08	-0.53
	4 and 8	114.4	66.7	8.36	4.7	**5.67**	2.69	-0.98
	5 and 7	25.9	127.8	9.79	5.66	**6.2**	3.59	-0.54
	6 and 8	14.4	67.7	8.99	5.32	**8.02**	0.97	-2.7
	7 and 8	27.8	24.2	7.42	3.29	**5.5**	1.92	-2.21
ME 3	1 and 8	209.8	82.4	12.27	5.88	**5.7**	6.57	0.18
	1 and 9	19.6	128.1	8.01	6.14	**6.65**	1.36	-0.51
	3 and 4	108.7	109.2	9.34	7.14	**5.9**	3.43	1.24
	3 and 7	76.7	57.9	8.43	4.69	**5.49**	2.95	-0.79
	4 and 9	169.3	59.8	7.36	5.17	**6.24**	1.12	-1.07
	6 and 8	99.4	83.7	12.21	5.83	**5.77**	6.45	0.06

	7 and 8	51.3	139.2	8.87	2.48	**6.53**	2.34	-4.04
ME 4	4 and 10	89.6	130.11	6.85	4.3	**5.56**	1.29	-1.26
ME 5	3 and 7	143	98.4	9.84	7.37	**7.12**	2.72	0.26
	7 and 8	45.8	15.4	8.11	4.97	**5.9**	2.21	-0.94
ME 6	1 and 5	1.96	92.52	7.63	3.97	**6.46**	1.17	-2.5
	1 and 8	25.83	61.12	8.08	4.41	**5.87**	2.21	-1.46
	5 and 8	103.96	138.58	7.3	5.05	**6.15**	1.14	-1.1
	7 and 10	94.48	108.86	8.51	4.76	**6.03**	2.48	-1.27
ME 7	1 and 2	129.1	91.5	11.48	5	**6.34**	5.13	-1.34
	2 and 8	127.5	52.9	12.15	5.68	**5.96**	6.2	-0.28
	4 and 7	0	160.2	9.24	6.05	**6.84**	2.4	-0.79

Several significant epistatic interactions were found for silk DAP in all locations and ME1 between chromosomes 1 & 3 and 3 & 8. No significant interactions were found in other locations except ME7, where six significant epistatic interactions were identified between chromosomes 1 & 2, 1 & 7, 2 & 6, 2 & 9, 6 & 9, and 8 & 9 (**Table 3**). The significant epistatic thresholds were 5.38, 5.36, 8.45, 5.53, 5.87, 5.37, and 5.26 across locations and MEs.

Significant epistatic interactions were found for all locations except for ME2 and ME4, following a two-dimensional search for PH. Four, three, two, one, and two significant epistatic interactions were found for ME3, ME4, ME5, ME6, and ME7, respectively, between chromosomes 1 and 2, 3 and 7, 4 and 6, 8 and 10 on ME3, 1 and 8, 5 and 10, 6 and 7 on ME4, 1 and 7, 2 and 7 on ME5, 3 and 4 on ME6, and 1 and 2, 5 and 8 on ME7 (**Table 3**).

Significant epistatic interactions were found for EH across locations ME3, ME5, and ME7. Especially for ME3 and ME6, we observed two epistatic interactions between the chromosomes for all locations 2 & 6 and 2 & 7, ME3, 1 & 6 and 2 & 5, and ME5 1 & 9 and 6 & 9, and Four significant epistatic interactions on chromosomes 1 & 6, 3 & 9, 4 & 6, and 4 & 7 (**Table 3**).

A two-dimensional scan for all locations has successfully identified epistatic interactions for yield. We observed three, five, seven, one, two, four, and three epistatic interactions on all locations, ME1, ME3, ME4, ME5, ME6, and ME7 (**Table 3**).

Epistasis by environment interactions

A linear mixed model was created separately for the six phenotypic observations (pollen DAP, silk DAP, PH, EH, stand %, and yield). The models considered genotype, environment, family, epistasis, QTL, QTL-environment interaction, and EEI effects, and FDR correction was applied to detect significant p-values (**Table 4**). Sixteen, eight, twelve, twenty, fourteen, and eighteen significant markers from SIM and two-dimensional epistasis searches were selected for the linear mixed models of pollen DAP, silk DAP, PH, EH, stand percentage, and yield, respectively. Significant EEI effects were observed for pollen DAP, silk DAP, PH, and yield but not for EH. Environment and family effects were statistically significant for each phenotypic trait. In the analysis of pollen DAP, two significant QTLs, two QTL-environment interactions, and two EEI were identified. For silk DAP, two significant QTLs, two QTL-environment interactions, one epistasis interaction, and one EEI were detected. Analysis of PH revealed six significant QTLs, two QTL-environment interactions, and one EEI. In the analysis of EH, five significant QTLs, one QTL-environment interaction, and three epistatic interactions were found. For stand percentage, two significant QTLs and one EEI were observed. Lastly, two QTLs, three QTL-environment interactions, and one EEI were found for yield (**Figure 24 A to F, Figure 25**).

Table 4: Significant terms for each phenotypic trait on pollen DAP, silk DAP, PH, EH, stand %, and yield. Adjusted P-values (Padj) is the P-values of the FDR corrected P-values (Pr(>F)).

Pollen DAP	**Sum Sq**	**Mean Sq**	**NumDF**	**DenDF**	**F value**	**Pr(>F)**	**Padj**
Location	19048.6	2721.23	7	21.73	6444.25	0	0
S5_144805011	21.2	21.25	1	237.2	50.32	0	0
S4_146407719	5.2	5.16	1	237.17	12.23	0	0.01
Location:S7_169202944	10.2	1.69	6	1434	4.01	0	0.01
Location:S2_9498230	8.6	1.43	6	1434	3.39	0	0.05
Location:S9_48378188:S2_224615471	8.9	1.49	6	1434	3.53	0	0.04
Location:S2_9498230:S2_224615471	10.2	1.7	6	1434	4.03	0	0.01
Silk DAP	**Sum Sq**	**Mean Sq**	**NumDF**	**DenDF**	**F value**	**Pr(>F)**	**padj**
Location	47927	6846.8	7	2	11532.9	0	0
S5_144805011	37	37.5	1	237.52	63.11	0	0
S10_125572892	26	26	1	237.1	43.84	0	0
Location:S10_125572892	23	3.9	6	1434	6.51	0	0
Location:S1_266472728	24	4	6	1434	6.71	0	0
S6_132320681:S9_8224784	7	6.9	1	237.13	11.67	0	0.01
Location:S9_10972540:S3_4836798	14	2.4	6	1434	3.99	0	0.01
Plant Height	**Sum Sq**	**Mean Sq**	**NumDF**	**DenDF**	**F value**	**Pr(>F)**	**padj**
Location	464859	66408	7	2	2021.03	0	0.01
S1_236244055	613	613	1	236.14	18.67	0	0
S3_157440022	1034	1034	1	236.13	31.47	0	0
S5_1564701	569	569	1	236.77	17.32	0	0
S9_157493113	373	373	1	236.16	11.36	0	0.01
S10_64147115	406	406	1	236.02	12.37	0	0.01
S8_33490583	530	530	1	236.17	16.14	0	0
Location:S1_236244055	833	139	6	1428	4.22	0	0.01
Location:S8_33490583	706	118	6	1428	3.58	0	0.02
Location:S3_157440022:S8_33490583	649	108	6	1428	3.29	0	0.04

Ear Height	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)	padj
Location	180503	25786.1	7	2	1510.7	0	0.01
S3_152447071	1019	1019.3	1	198.04	59.71	0	0
S5_178308576	541	541.1	1	198.27	31.7	0	0
S10_84124156	207	206.6	1	198.06	12.11	0	0.01
S5_1890045	341	341.1	1	198.3	19.98	0	0
S9_104861090	603	603.3	1	198.25	35.35	0	0
S8_23383882	212	211.8	1	198.13	12.41	0	0.01
Location:S1_5106842	619	103.2	6	1200	6.04	0	0
S5_178308576:S9_104861090	222	221.6	1	198.02	12.98	0	0.01
S1_189290778:S4_23889254	365	364.8	1	198	21.37	0	0
S1_5106842:S7_172473946	197	196.8	1	198.07	11.53	0	0.02
Stand %	**Sum Sq**	**Mean Sq**	**NumDF**	**DenDF**	**F value**	**Pr(>F)**	**padj**
Location	1032825	147546	7	2	9958.63	0	0.02
S7_161220674:S1_38255817	155	155	1	198.46	10.47	0	0.05
S1_38255817:S9_331601	181	181	1	198.95	12.18	0	0.04
Location:S6_142400130:S9_144139888	339	56	6	1194	3.81	0	0.04
Yield (bu/Ac)	**Sum Sq**	**Mean Sq**	**NumDF**	**DenDF**	**F value**	**Pr(>F)**	**padj**
Location	1351892	193127	7	2	1565.1	0	0.02
S2_180671323	3201	3201	1	199.9	25.94	0	0
S7_128747685	1335	1335	1	199.74	10.82	0	0.03
Location:S2_180671323	2929	488	6	1206	3.96	0	0.02
Location:S3_225468828	4470	745	6	1206	6.04	0	0
Location:S2_201906553	3870	646	6	1206	5.23	0	0
Location:S6_93911476:S8_118250825	3130	522	6	1206	4.23	0	0.02

Pollen DAP		B_1B_1	B_2B_2
All L.	A_1A_1	62.1	61.8
	A_2A_2	62.1	61.9
ME 1	A_1A_1	64.5	64.2
	A_2A_2	64.6	64.2
ME 3	A_1A_1	61.7	61.2
	A_2A_2	61.4	61.5
ME 4	A_1A_1	56.1	55.4
	A_2A_2	56.0	56.1
ME 5	A_1A_1	62.6	62.3
	A_2A_2	62.7	62.6
ME 6	A_1A_1	68.9	68.6
	A_2A_2	69.1	68.8
ME 7	A_1A_1	65.7	65.4
	A_2A_2	65.8	65.6

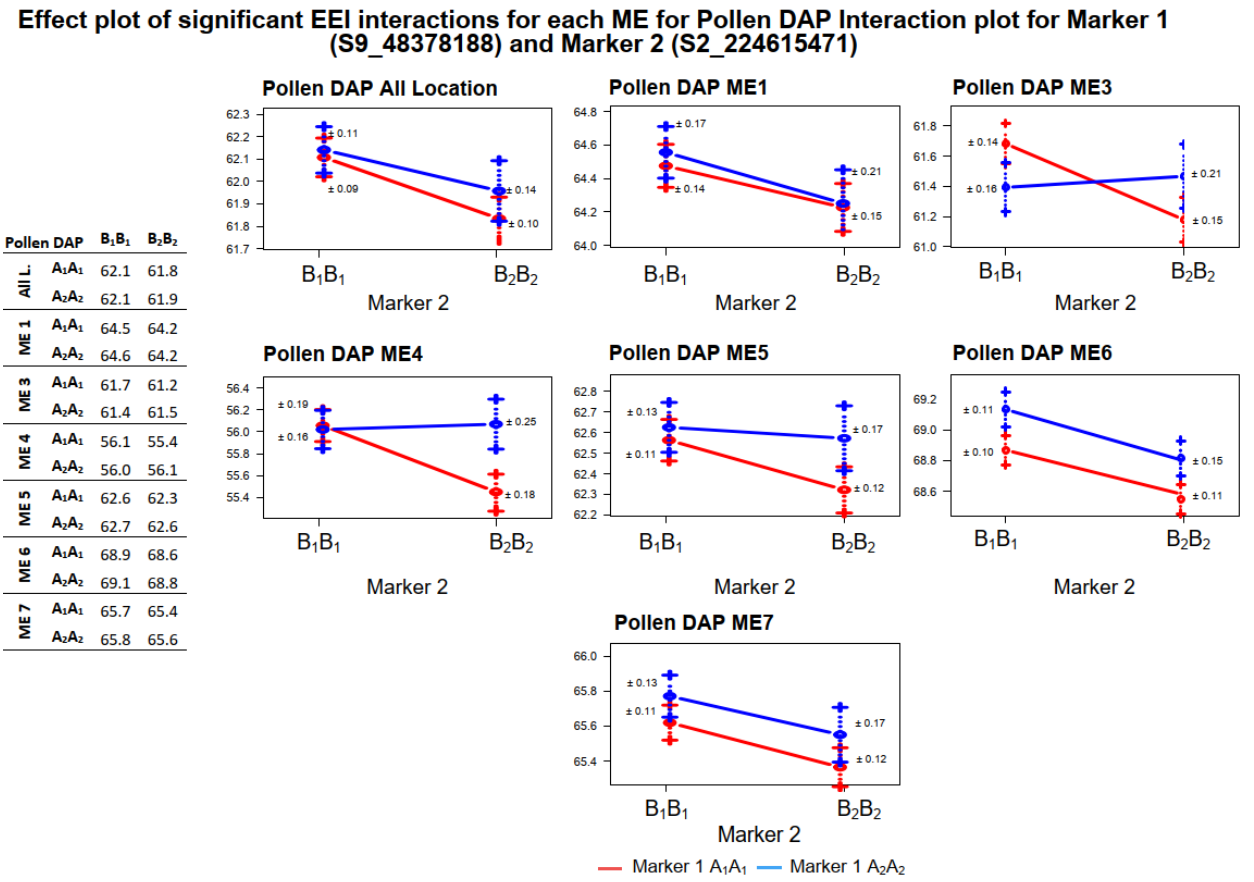
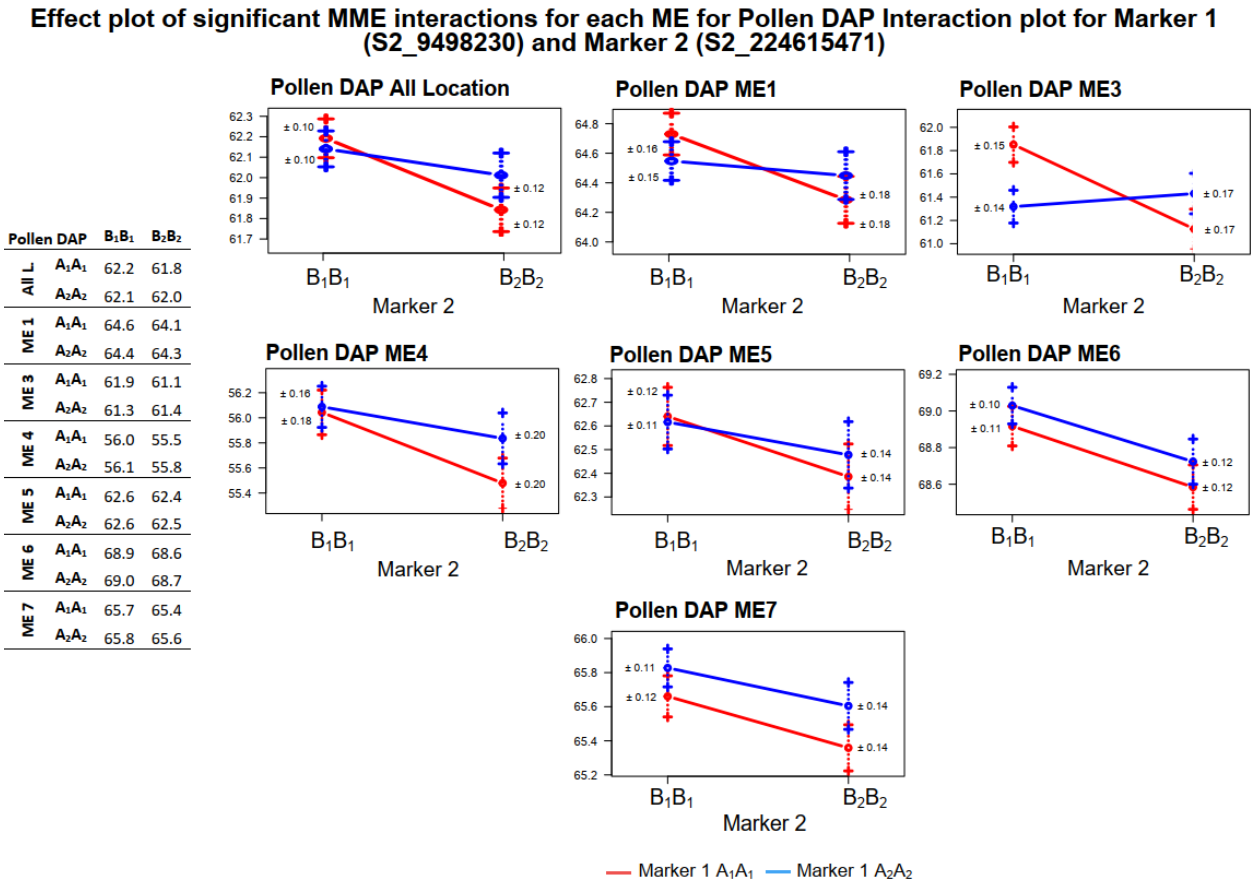
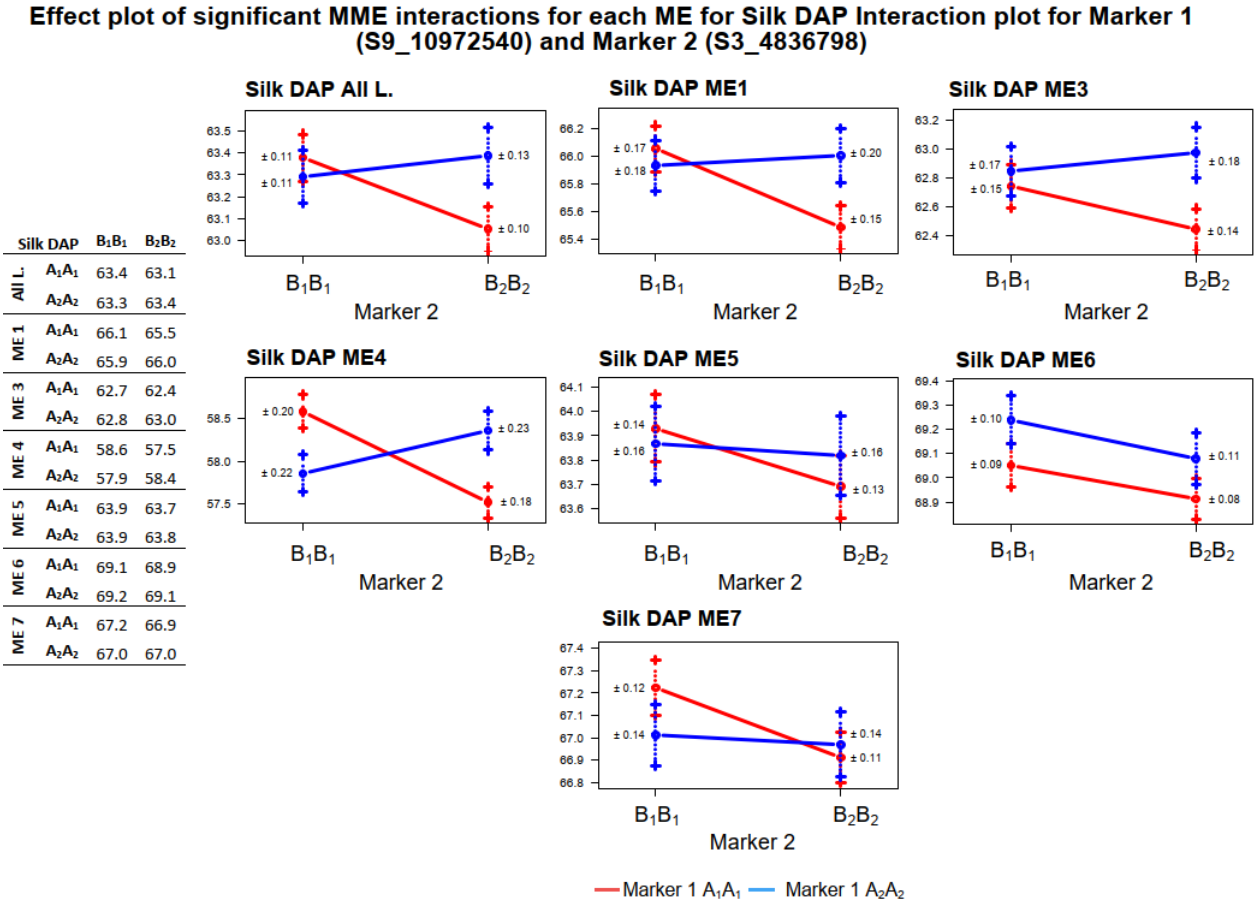


Figure 24 – A: Significant interaction plots for each mega environment (MEs) for pollen DAP between two markers (S9_48378188 and S2_224615471).



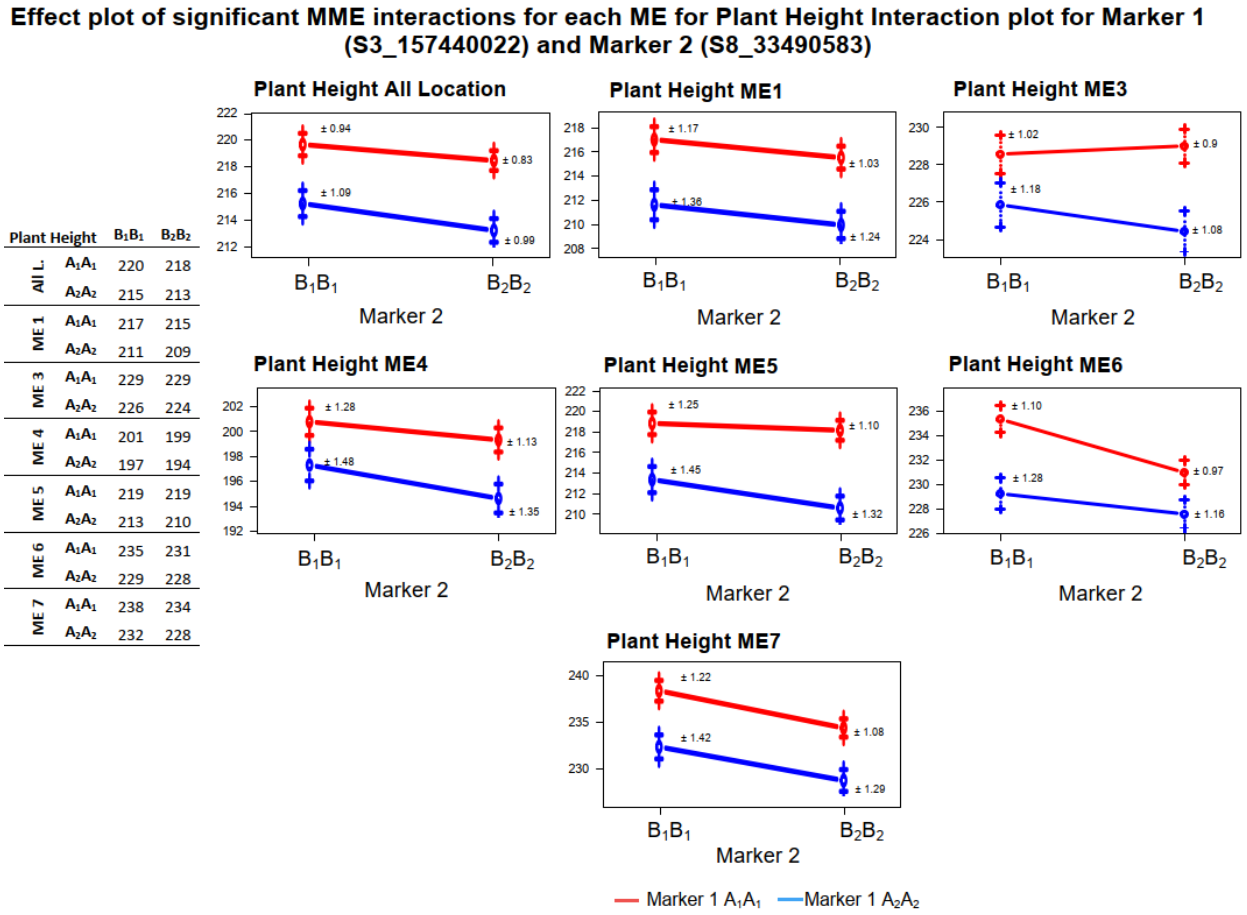
Pollen DAP	B_1B_1	B_2B_2
All L	A_1A_1	62.2	61.8
	A_2A_2	62.1	62.0
ME 1	A_1A_1	64.6	64.1
	A_2A_2	64.4	64.3
ME 3	A_1A_1	61.9	61.1
	A_2A_2	61.3	61.4
ME 4	A_1A_1	56.0	55.5
	A_2A_2	56.1	55.8
ME 5	A_1A_1	62.6	62.4
	A_2A_2	62.6	62.5
ME 6	A_1A_1	68.9	68.6
	A_2A_2	69.0	68.7
ME 7	A_1A_1	65.7	65.4
	A_2A_2	65.8	65.6

Figure 24 – B: Significant interaction plots for each mega environment (MEs) for pollen DAP between two markers (S2_9498230 and S2_224615471).



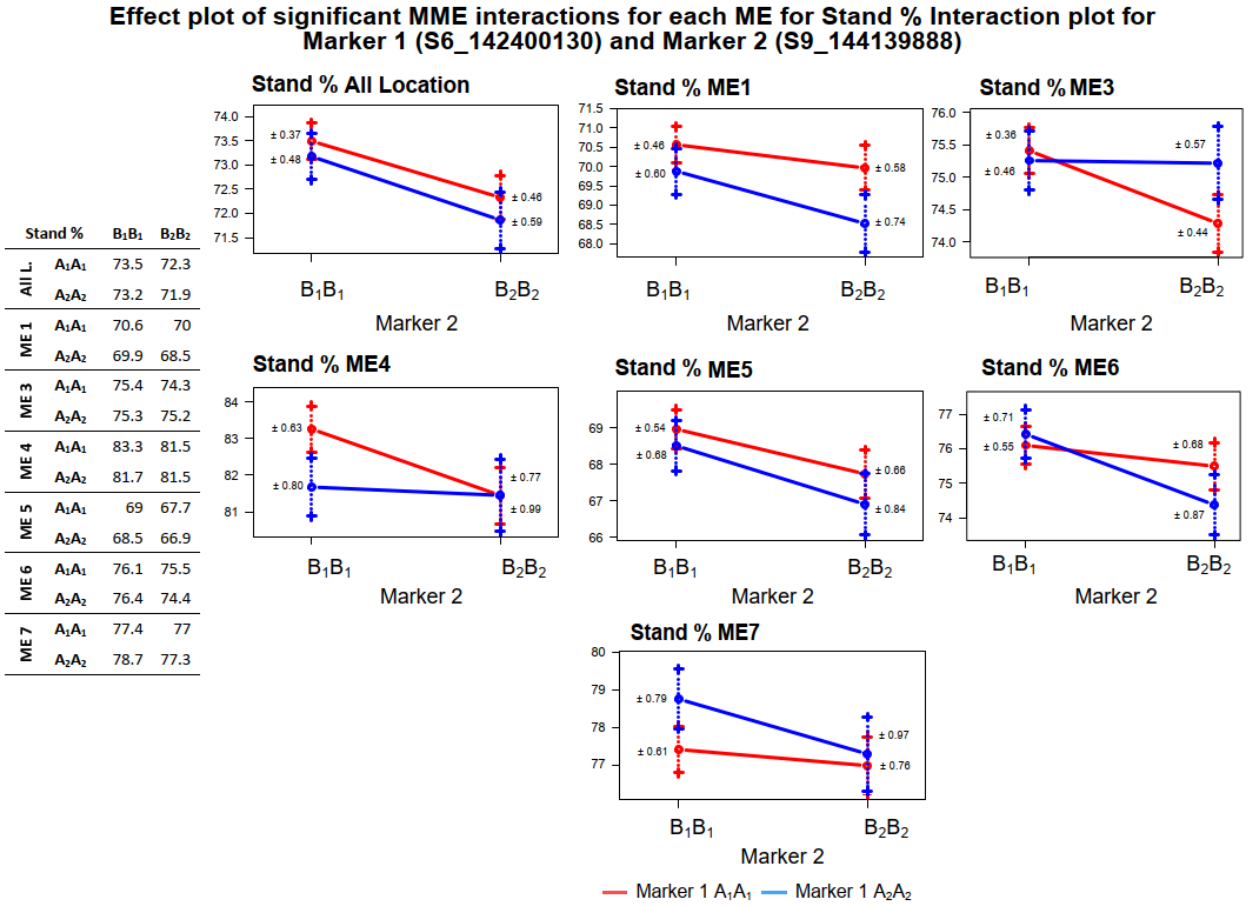
Silk DAP		B_1B_1	B_2B_2
All L.	A_1A_1	63.4	63.1
	A_2A_2	63.3	63.4
ME 1	A_1A_1	66.1	65.5
	A_2A_2	65.9	66.0
ME 3	A_1A_1	62.7	62.4
	A_2A_2	62.8	63.0
ME 4	A_1A_1	58.6	57.5
	A_2A_2	57.9	58.4
ME 5	A_1A_1	63.9	63.7
	A_2A_2	63.9	63.8
ME 6	A_1A_1	69.1	68.9
	A_2A_2	69.2	69.1
ME 7	A_1A_1	67.2	66.9
	A_2A_2	67.0	67.0

Figure 24 – C: Significant interaction plots for each mega environment (MEs) for silk DAP between two markers (S9_10972540 and S3_4836798).



Plant Height		B_1B_1	B_2B_2
All L.	A_1A_1	220	218
	A_2A_2	215	213
ME 1	A_1A_1	217	215
	A_2A_2	211	209
ME 3	A_1A_1	229	229
	A_2A_2	226	224
ME 4	A_1A_1	201	199
	A_2A_2	197	194
ME 5	A_1A_1	219	219
	A_2A_2	213	210
ME 6	A_1A_1	235	231
	A_2A_2	229	228
ME 7	A_1A_1	238	234
	A_2A_2	232	228

Figure 24 – D: Significant interaction plots for each mega environment (MEs) for plant height (PH) between two markers (S3_157440022 and S8_33490583).



	Stand %	B_1B_1	B_2B_2
All L.	A_1A_1	73.5	72.3
	A_2A_2	73.2	71.9
ME 1	A_1A_1	70.6	70
	A_2A_2	69.9	68.5
ME 3	A_1A_1	75.4	74.3
	A_2A_2	75.3	75.2
ME 4	A_1A_1	83.3	81.5
	A_2A_2	81.7	81.5
ME 5	A_1A_1	69	67.7
	A_2A_2	68.5	66.9
ME 6	A_1A_1	76.1	75.5
	A_2A_2	76.4	74.4
ME 7	A_1A_1	77.4	77
	A_2A_2	78.7	77.3

Figure 24 – E: Significant interaction plots for each mega environment (MEs) for stand percentage (stand %) between two markers (S6_142400130 and S9_144139888).

Yield (bu/Ac)		B_1B_1	B_2B_2
All L.	A_1A_1	156	151
	A_2A_2	152	153
ME 1	A_1A_1	179	167
	A_2A_2	166	172
ME 3	A_1A_1	156	152
	A_2A_2	154	149
ME 4	A_1A_1	147	146
	A_2A_2	152	150
ME 5	A_1A_1	155	153
	A_2A_2	153	155
ME 6	A_1A_1	130	129
	A_2A_2	126	129
ME 7	A_1A_1	197	187
	A_2A_2	190	193

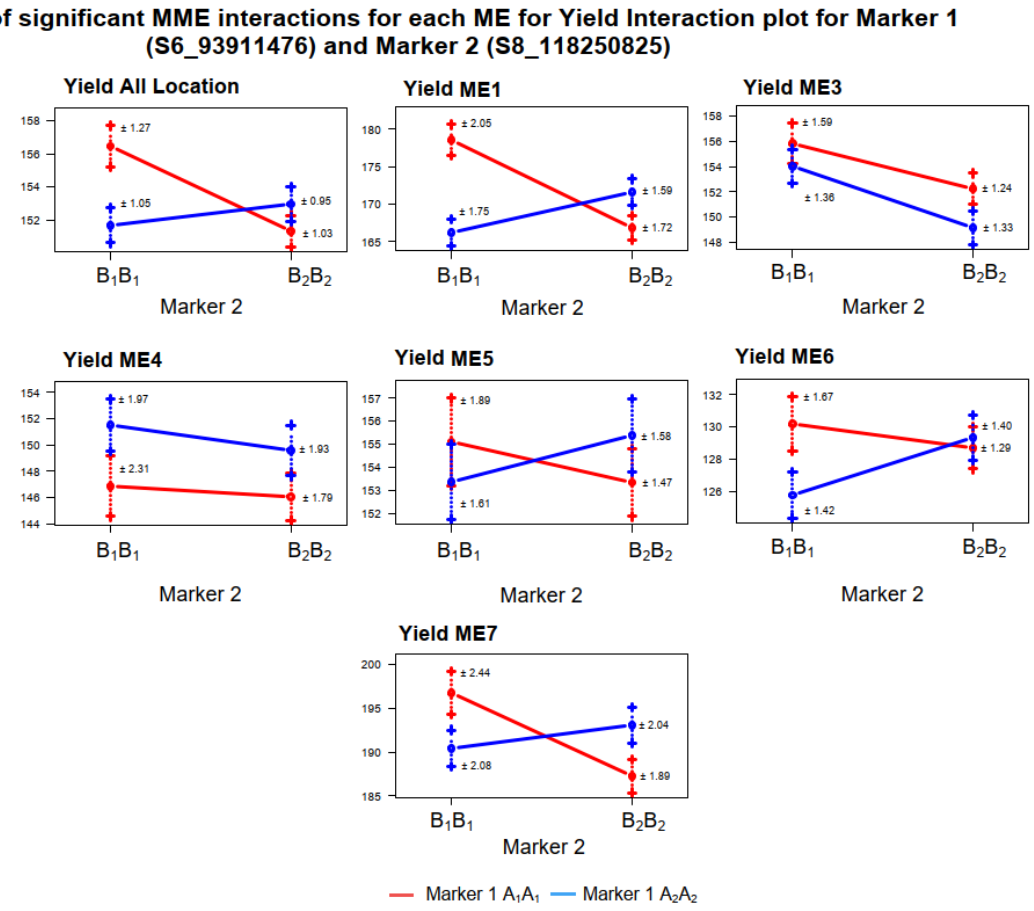


Figure 24 – F: Significant interaction plots for each mega environment (MEs) for yield (yield (bu/Ac)) between two markers (S6_93911476 and S8_118250825).

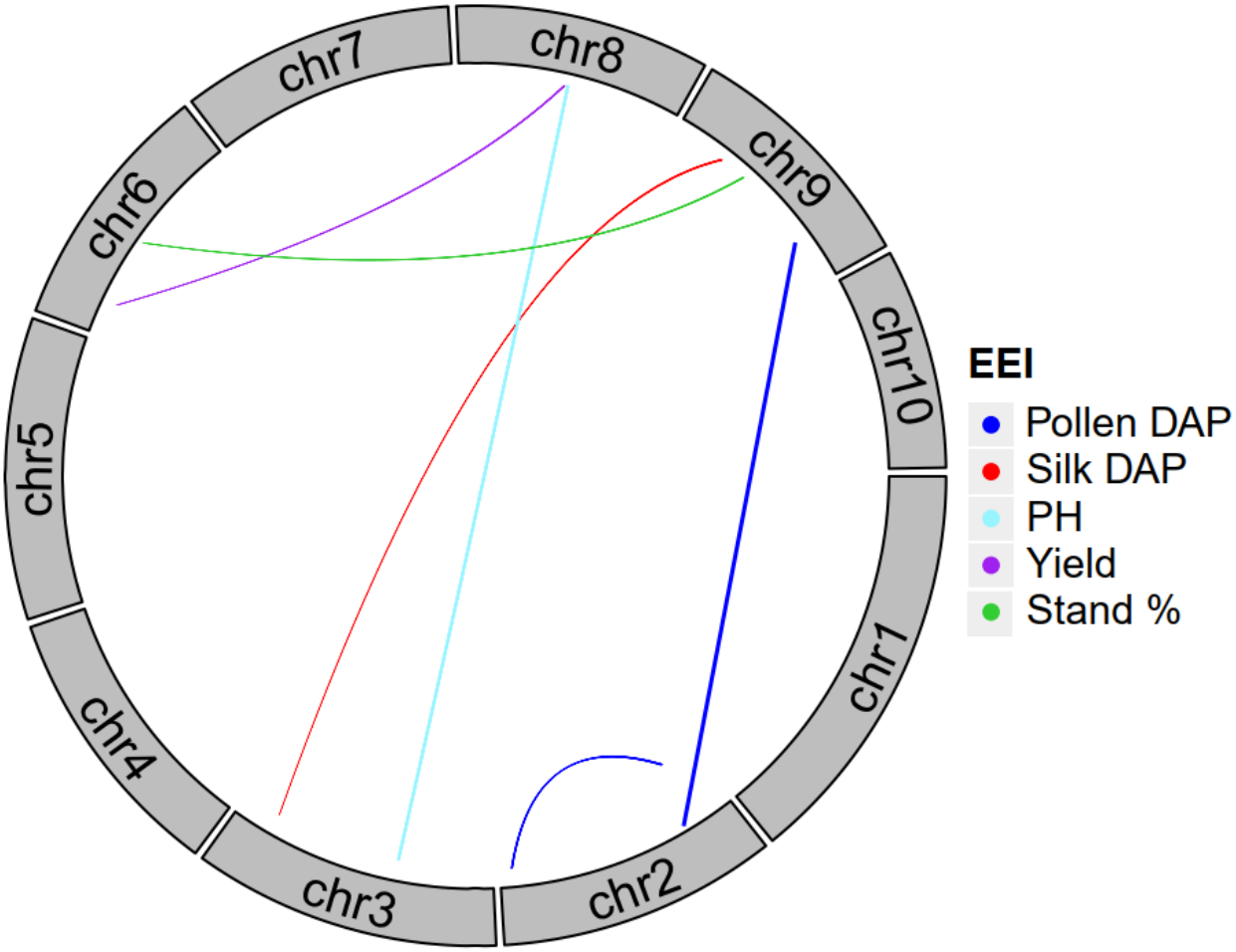


Figure 25: Significant EEI plot for each trait (pollen and silk DAP, plant height, ear height, stand %, and yield (bu/Ac)).

Discussion

Correlations & heritability

It is well known that there is a correlation between agronomic traits, and correlations can account for genetic and non-genetic relationships between two or more traits (Silva et al., 2016). In previous studies, A strong correlation between pollen DAP and silk DAP has been found (Ribaut et al., 1996). This is expected since the inflorescence time for maize tassel and silk is consecutive. Expectedly, we found a strong correlation between the pollen and silk DAP. As previous studies found, PH and EH, with which we found a strong correlation between the two traits, are other highly correlated quantitative traits (Malik et al., 2005).

Furthermore, moderate correlations between PH, EH, pollen DAP, and silk DAP are reported since one quantitative trait influences another, as previous studies demonstrated (Messmer et al., 2009). We found a similar interaction pattern between the PH, EH, pollen DAP, and silk DAP. A correlation between stand % and other agronomic traits is expected to be close to zero in a fully germinated experiment. However, during a growing period, external influencer factors might affect and eliminate plants in an experiment, which stand % might correlate with other traits indirectly (Zhao et al., 2018). Across locations, we observed no correlation with stand %, but we found low negative correlations in ME1 for pollen DAP and silk DAP and all agronomic traits in ME3, except for yield. As expected, stand % and yield correlated low to moderate across MEs. Previous research indicates that yield is a highly quantitative agronomic trait influenced by external factors, such as environments (Kravchenko and Bullock, 2000). Therefore, yield correlated poorly with all traits in the experiment.

H^2 across locations was high for each phenotypic trait, especially for PH and EH. Also, H^2 is calculated in previous publications for PH and EH, which was reported as a highly heritable trait (Li et al., 2017). In addition, heritability calculation could be affected by external factors, such as the environment and the missing data (Stirling et al., 2002; Nakagawa and Freckleton, 2008). H^2 was high across MEs as well. However, H^2 was low in ME3 for pollen and silk DAP and ME6 for silk DAP due to the high error variance. Phenotyping errors or missing values might cause high error variance. The missingness for ME3 for pollen and silk DAP was one environment out of five. For ME6, all genotypes within environments in ME6 were observed except for some missing values

for silk DAP. It was observed that phenotypic variation between environments within MEs was high for the corresponding trait. As stated previously, low heritability has low power to detect minor QTLs (Viana et al., 2017). Our analysis with low heritability MEs detected major QTLs, but minor effect QTLs and significant epistatic interactions were not observed in low H^2; therefore, for EEI statistical analysis, there were no significant QTLs from low H^2.

Linkage and QTL mapping

QTL mapping analysis revealed significant QTLs for pollen and silk DAP, PH, EH, stand %, and yield across environments or location-specific QTLs in our analysis. Similar results were revealed in previous publications on each phenotypic trait. A study worked on maize and identified QTLs male and female flowering time with 234 individuals. They genotyped F2 populations on 142 loci, and F3 families were evaluated. QTLs on chromosomes 1, 2, 4, 5, 8, 9, and 10 accounted for 48% of the phenotypic variation on anthesis and silking DAP (Ribaut et al., 1996). Another study found several significant QTLs on PH in all chromosomes with four populations of maize (Beavis et al., 1991). Other studies discovered identical results on pollen and silk DAP, PH, EH, stand %, and yield (Yan, 2003; Messmer et al., 2009; Cai et al., 2012; Leng et al., 2022). Various QTLs were found across locations for each corresponding trait. However, location is a substantial factor in identifying QTLs. Each publication indicated several stable QTLs across locations.

Genetics controls flowering and silking time; it has also been known that the environment affects naturally evolved populations. It has been discovered that one large effect of QTL controlling the flowering time is called Vgt1, located on chromosome 8 (Salvi et al., 2007). However, follow-up studies indicated that the flowering time is quantitative and controlled with many minor effects QTLs. A study with over 4000 maize landraces across 22 environments discovered significant SNPs associated with flowering time. In addition, the study revealed that more than 50% of the identified SNPs for the flowering time were also associated with altitude, which indicates environmental effects on flowering time (Romero Navarro et al., 2017). A study with two near-isogenic lines (NILs) of F3 maize lines identified significant QTLs on flowering time on chromosomes 1, 8, and 10 (Koester et al., 1993). Also, Our study found significant QTLs on chromosomes 1, 4, 5, 8, 9, and 1, 5, 9, and 10 for pollen and silk DAP, respectively. As indicated

previously, the correlation between the two traits is high. Therefore, some of the QTLs colocalized.

Plant height and ear height are complex traits influenced by genetic and environmental factors. Several studies have identified significant quantitative trait loci (QTLs) associated with maize plants and ear height. Multiple factors, including hormones such as gibberellins and environmental effects such as light and temperature, regulate the expression of genes that control plant height and ear height. Understanding the genetic basis of these traits can help breeders develop maize varieties with improved agronomic performance and yield potential (Cai et al., 2012). For PH, on chromosomes 1, 3, 5, 7, 8, 9, and 10, significant QTLs were found. In one study, significant QTLs for PH in maize were found on chromosomes 1, 3, 5, 8, and 9. The study was conducted in different stages of the PH and concluded that QTLs related to the PH differently express in various stages of maize growth (Yan, 2003). Another study worked with 103 microsatellite markers, and the F2:4 family found QTLs on PH and EH on chromosomes 2, 3, 4, 8, and 9. The QTLs explained 41.5% of EH and 78.27% variance for PH (Zhang et al., 2006). Another large effect of QTL on PH (32.3% variation explained) on chromosome 3 within four cM intervals was found. Furthermore, they created Segmental isolines and narrowed the interval of the QTL to 12.6 kb (Teng et al., 2013).

As for EH, significant QTLs were found on chromosomes 1, 3, 5, 9, and 10. EH and PH are strongly correlated traits. Therefore, in some cases, our experiment and the literature share the co-located QTLs. A study with F2:3 and RIL maize populations found shared QTL locations on chromosomes 1, 2, and 5 for the F2:3 population and chromosomes 2, 3, and 9 for the RIL population (Li et al., 2016). Another study was conducted on 14 diverse locations to identify QTLs on PH and EH on maize with F1 maize hybrid lines. Researchers successfully identified stable QTLs across environments for both traits (Li et al., 2017).

Stand % in a plot indicates how the genotypes resist environmental factors. Green snap and stalk lodging are two factors that could influence agronomic traits, such as yield. Thus, stand % is one factor that points out how the plants are resilient to those external factors. This study identified only one significant QTL on chromosome 8 in ME 4. In a study focusing on maize, researchers evaluated stalk bending strength (SBS) using parameters like maximum load to breaking (F max),

breaking moment (M max), and critical stress (σ max) in a population of 216 recombinant inbred lines. The heritability of these SBS traits was high, with values ranging from 0.75 to 0.81. Through genetic analysis, they identified multiple quantitative trait loci (QTLs) responsible for significant portions of the genotypic variance in these SBS traits, concentrated in specific genomic regions. The study also explored using genomic prediction methods, such as GBLUP and BayesB, which showed higher predictive accuracy than QTL mapping (Hu et al., 2013).

Many studies have identified significant QTLs for yield in maize in every chromosome. In one study, a significant QTL for yield on chromosome 1 was found under drought and well-water regimes with 101 SSR markers with a RIL population (Guo et al., 2008). Another study on yield components in maize has discovered several QTLs related to ear length, diameter, ear number, ear weight, and prolificacy, respectively, 5, 5, 2, 2, and 3 QTLs identified. They identified significant QTLs on each chromosome except chromosome 6 (Karen Sabadin et al., 2008). Another study indicated that plant density could influence yield components in maize. Researchers sow the plants under two diverse plant densities with F2:3 genotypes derived from two elite inbred lines to test their results. According to their results, 30 QTLs were detected with high-density sowing. Also, they indicated that the high plant density affects some yield component traits, such as ear length (Guo et al., 2011). Our study identified several significant QTLs on yield on chromosomes 2, 3, 5, 7, and 8.

In conclusion, numerous studies on PH, EH, pollen and silk DAP, stand %, and yield have identified significant QTLs contributing to each trait of interest. However, the specific significant QTLs can vary depending on the population, environment, and the number of individuals included in the study. The examples provided for each trait showed both similar and different results. Our analysis also found significant QTLs consistent across different environments or specific to certain locations. Furthermore, the subsequent chapter revealed significant QTL-environment interactions and epistasis effects, as well as significant epistasis by environment interactions, which were evaluated using a mixed linear model approach.

QTL environment interactions

Although the main effect of QTLs had been found in the previous analysis, many indicated location-specific QTLs. For instance, we observed a significant marker on chromosome 5, position 81.1, for Pollen DAP across the location (**Table 2**). However, a marker located on chromosome 9, position 60.5, was only observed on ME6. One reason for the location-specific marker is an interaction between genetics and the environment (QTL x environment). Previous studies observed QTL x environment interactions, as we observed in our analysis. Fériani et al. investigated barley with previously mapped SSR markers under two diverse locations in Tunisia. Although the number of the environment is limited, they found significant QTL x environment interactions (Fériani et al., 2020). Wang et al. studied introgressed rice lines under six environments with yield-related phenotypic traits with SNP markers. They found significant markers for each tested trait. In addition, they pointed out that the majority of the significant markers are environment-specific (Wang et al., 2014). Despite the obstacles, such as few markers and environments to test the genotypes, previous researchers indicated significant QTL x environment interactions. In this research, we were able to find significant QTL x environment interactions: two for pollen, silk DAP, and PH, one for EH, and three for yield traits among 2415 markers and six MEs. For pollen DAP, two significant interactions were found between the environments and chromosomes 7 and 2. For silk DAP, chromosomes 1 and 10 interact with environments. Also, two significant QTL and environment interactions were found for PH between chromosomes 1 and 8. Only one significant QTL environment interaction was found on chromosome 1 for EH. No significant QTL x environment interactions were found for stand %. Lastly, three significant interactions with environments were found for chromosomes 2 and 3 for yield.

Epistasis

The interaction between genetics and environment attracted researchers' attention to understand the underlying reasons. Yu et al. have identified epistatic interactions on plant height in rice. They used 240 F2:3 families from a cross between two elite inbred lines. Significant epistatic interactions for plant height and heading date had been identified (Yu et al., 2002). Another study with maize RIL population derived between the cross B73 and By804 high-oil line

identified 42 major QTLs as well as 36 epistasis loci related to palmitic acid, stearic acid, oleic acid, linoleic acid, and oil. Researchers indicated that additive major QTLs play a crucial role in epistatic interactions, contributing to the oil content in maize (Yang et al., 2010).

In our research, significant epistatic interactions were found for each phenotypic trait. Five epistatic interactions have been found for pollen DAP. For silk DAP, nine significant epistatic interactions were found, and six of the significant epistatic interactions were found on ME7. Twelve and four significant epistatic interactions were found for PH and EH. Ten epistatic interactions were found for stand %, and twenty-six significant epistatic interactions were found for yield.

Epistasis by environment interactions

Epistasis may vary depending on the environmental conditions observed and explained in the previous chapters. This can complicate the identification and characterization of QTLs for a given trait, as the effect of epistasis may be indistinct in different environments. Previous studies have proven that environmental factors are crucial in genetics, and researchers identified QTL x environment interactions. However, epistasis by the environment is largely unknown. In this study, we found significant epistasis by environment interactions for pollen and silk DAP, PH and EH between chromosomes 2 and 9, 2 and 2 for pollen DAP, 3 and 9 for silk DAP, 3 and 8 for PH, and lastly, 6 and 8 for yield. Only a few studies pointed out the EEI. A study of the Chinese rapeseed DH population with 282 lines in four locations revealed significant epistasis by environment interactions with a linear mixed model on the oil seed content. Their study used the analysis with 125 SSR markers (Zhao et al., 2005). Another study focused on identifying the EEI with association mapping study on cotton cultivars with 323 accessions with 651 SSR markers under nine diverse environmental conditions on yield-related traits with an LMM approach. They indicated that the heritability of significant EEI is higher than additive and epistasis for two yield-related traits, namely, lint yield and boll number (Jia et al., 2014). Limitations of the previous studies were the low number of environments and the low number of markers evaluated in the experiments.

In conclusion, main QTLs are identified for each corresponding phenotypic trait; also, environmental conditions affect phenotypic observations of identical genotypes. Thus location-

specific QTLs have been identified in addition to the main effect QTLs. The main effect of QTLs was detected in previous studies alongside environment-specific QTLs. The effect of QTL-environment interactions is illustrated in different studies. However, there are few studies focused on EEI. This study focused on identifying the main effect of QTLs, interactions of QTL-environment, and EEI with a linear mixed model. The results indicate that there are interactions between epistasis and environment. Permutation tests indicated that the linear mixed model results of EEI are trustworthy. Therefore, researchers should consider this factor when they analyze their data. Further investigation is needed to understand the underlying reasons for EEI.

Chapter 3: Gene Expression Analysis to Detect the Effect of the Origin of Seeds for an Epistasis Mapping Population Experiment in Maize

Abstract

Epigenetic modifications are an important factor that could impact the phenotype of individual plants, and the next generations could inherit those modifications. Therefore, understanding the epigenetic effects on a gene expression level between genotypes could improve the underlying reason for phenotypic variation.

In this study, we used a unique maize population called the Epistasis Mapping Population (EMP), created to identify epistatic interactions efficiently by reducing the number of tests required to detect epistasis. This population was generated by crossing near-isogenic lines (NILs) with the recurrent parent B73 and each other. Crossing only 20 NIL founders in all 190 pairwise combinations, the population encompasses approximately 75% of the genome to test for pairwise interactions.

We conducted a preliminary study to evaluate whether there is an origin-of-seed effect on gene expression between two distinct environments (Nursery (field) and greenhouse (GH)). We grew a subset of the EMP, which included two NILs (B73 and the F1NILs) in a controlled environment. We harvested the above-ground tissues at the end of the V2 growth stage and performed 3'-tag Digital Gene Expression (DGE) analysis. BrAD-seq protocol for the library preparation was utilized, and the libraries were sequenced by Illumina sequencing technology. Our bioinformatics pipeline included quality control, mapping the reads to the B73v5 reference transcriptome, and using the R package DESeq2 to detect significantly differently expressed (DE) genes.

Although we observed DE genes between all the genotypes tested, we did not detect any DE genes between seeds of different origins. Our results suggest an insignificant origin-of-seed effect on transcription in above-ground tissue at the V2 stage.

Material & Method

Population and growth condition

Our research uses an Epistasis Mapping Population (EMP). EMPs were created by backcrossing near-isogenic lines (NIL) to B73 (BC-NILs) and crossing them with each other to create F1-NILs. The NIL founders were created by Eichten et al. by crossing B73 to Mo17, backcrossing to B73 for two generations, and self-crossing for four generations (Eichten et al., 2011). The main goal of the EMPs is to decrease the number of tests required to detect epistatic interactions with 20 NIL founders, 20 BC-NILs, and 190 F1-NIL hybrids. The effect of BC-NIL1 and BC-NIL2 is α1 and α2, which can be calculated compared to the introgressed region of BC-NIL1 and BC-NIL2 to B73. If the sum of α1 and α2 is not equal, we can conclude that there are epistatic interactions between the introgressed regions. Otherwise, there are no epistatic interactions (**Figure 1**).

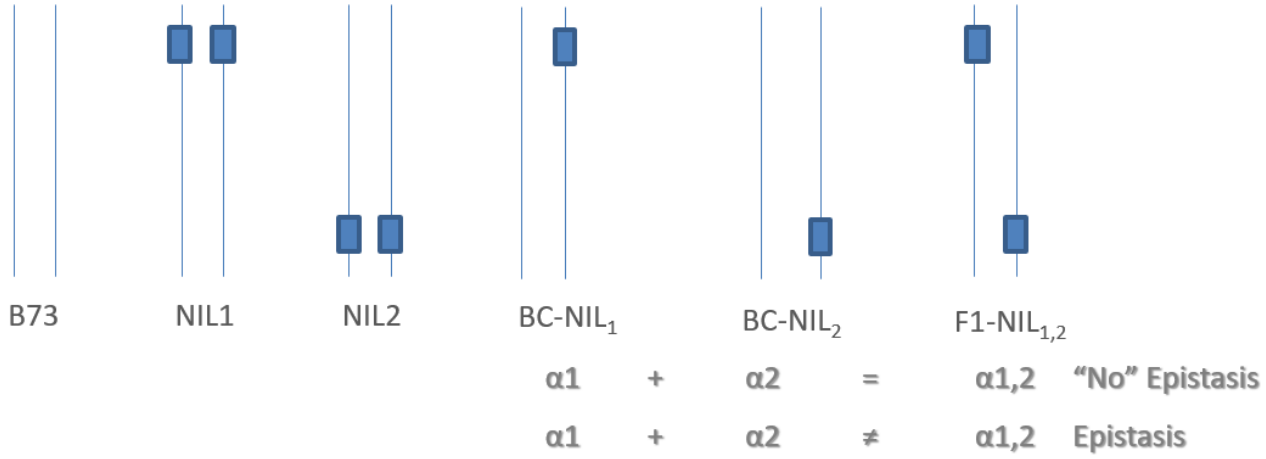


Figure 1: Design of the Epistasis Mapping Population (EMPs). α1 represents the effect of the introgressed region between BC-NIL1 and NIL1, while α2 represents the effect between BC-NIL2 and NIL2. α1,2 is the effect of F1-NIL1,2. The impact of the hybrid is equal to the sum of α1 and α2, indicating no epistasis. If there is epistasis between the introgressed regions, the impact of the hybrid will differ.

For this study, we selected B73, 2 NILs (b169 and b175), and an F1 NIL hybrid of the selected NILs (b169 x b175) from two seed sources, which one seed source was from a field experiment (TB_nursery) in 2016, and the second was from a greenhouse (GH) in 2017. We had two different

seed sources because the missing genotypes of EMP were re-created in the 2017 GH experiment. The selected genotypes for the experiment were grown in a controlled environment (greenhouse). Above-ground parts of the maize plants (leaves and stems) were harvested, and the roots of the plants were not included in RNA extraction. Plants were collected at the end of the V2 stage (**Figure 2**). Four genotypes were replicated five times, and each replicated genotype was pooled twice from each seed source. Each pool was homogenized, and homogenized samples were kept at -80 °C for further analysis.

Figure 2: Photos of the experiment

RNA isolation, gene expression analysis

The cDNA libraries were prepared following the BrAD-seq (Breath Adapter Directional sequencing) protocol (Townsley et al., 2015). Strand-specific DGE was used with mRNA fragmentation using oligo L-3ILL-20TV.2. mRNA fragmentation was carried out at 25 °C for 1 second, 94 °C for 3 minutes, 30 °C for 1 minute, and 20 °C for 4 minutes. The reverse transcriptase (RT) program for cDNA synthesis was as follows: 25 °C for 10 minutes, 42 °C for 50 minutes, 50 °C for 10 minutes, and 70 °C for 10 minutes. The PCR enrichment protocol consisted of 98 °C for 30 seconds, 12 cycles of 98 °C, 65 °C, and 72 °C for 10, 30, and 30 seconds, respectively, and 72 °C for 5 minutes. The libraries were then checked by DNA electrophoresis, and the cDNA libraries were sent to an external facility for sequencing.

Quality control, RNA alignment, data collection

The quality of the raw sequencing data for each sample was assessed using FastQC ("FastQC," 2015). Cutadapt was then used to remove adapter sequences and low-quality nucleotides, with parameters -q, -m, -l, and --max-n (-q 5 -m 20 -l 50 --max-n 3) (Martin, 2011). Between 40.04 to 61.11 % of the raw reads for each sample were removed.

The high-quality reads were aligned to the B73v5 reference transcriptome using Kallisto, which the data is obtained from the maizeDGB database (Bray et al., 2016; Portwood et al., 2019). Kallisto pipeline was configured with a fragment length mean of 180, a standard deviation of 20, and a k-mer length of 31 for each sample. 62.98 – 69.57 % of the raw reads for each sample were aligned to the B73v5 reference transcriptome. The report for alignment for each sample can be seen in (**Table 1**).

Table 1: Kallisto runs for each sample describe raw read before and after quality control, the percentage of removed reads, the number of aligned sequences (NAS), and the percentage of aligned to the reference transcriptome.

Run	Sample (34)	Raw reads BeforeQC	Raw reads AfterQC (RRAQC)	Percentage of remove cleaning	Number of Aligned Sequence (NAS)	Percentage Align (RRAGQ/NAS)
s111	b169A	4450496	2248933	50.53	1478170	65.73
s112	b175A	5360644	2703651	50.44	1830513	67.71
s113	b73A	5952645	3637491	61.11	2427390	66.73
s114	b169/b175A	7346277	2941662	40.04	1965041	66.80
s121	b169B	6868029	3422706	49.84	2275053	66.47
s122	b175B	5226325	2583760	49.44	1627139	62.98
s123	b73B	9455119	4724712	49.97	3120545	66.05
s124	b169/b175B	9431741	4855705	51.48	3227706	66.47
s211	b169A	9832830	4836143	49.18	3264010	67.49
s212	b175A	6412962	3251022	50.69	2130968	65.55
s213	b73A	7703771	3810105	49.46	2650794	69.57
s214	b169/b175A	17960068	9062340	50.46	5987319	66.07
s221	b169B	3351868	1644454	49.06	1076298	65.45
s222	b175B	3004551	1522723	50.68	1008403	66.22
s223	b73B	8692789	4355562	50.11	2953864	67.82
s224	b169/b175B	5551977	2811051	50.63	1936105	68.87
s311	b169A	6264826	3215286	51.32	2168759	67.45
s312	b175A	10979046	5176741	47.15	3366794	65.04
s314	b169/b175A	3714961	1886668	50.79	1246110	66.05
s321	b169B	3747956	1935181	51.63	1297044	67.02
s322	b175B	7060394	3661156	51.85	2458562	67.15
s323	b73B	6631023	3348640	50.50	2305474	68.85
s324	b169/b175B	2663031	1335463	50.15	886879	66.41
s411	b169A	4021699	1950649	48.50	1298522	66.57
s412	b175A	2609937	1328625	50.91	876192	65.95
s421	b169B	3897133	1988464	51.02	1320688	66.42
s422	b175B	4442197	2283557	51.41	1539027	67.40
s423	b73B	2627651	1364637	51.93	919939	67.41
s424	b169/b175B	4615663	2356390	51.05	1599055	67.86
s511	b169A	5656513	2853807	50.45	1921680	67.34
s512	b175A	4157546	2143117	51.55	1423055	66.40
s513	b73A	4190794	2150508	51.32	1475452	68.61
s514	b169/b175A	4679979	2417193	51.65	1651525	68.32
s521	b169B	7007271	3551138	50.68	2420640	68.17
s522	b175B	2816127	1450801	51.52	960362	66.20
s523	b73B	4000212	2016125	50.40	1320413	65.49
s524	b169/b175B	6514683	3247326	49.85	2164467	66.65

Outlier, principle component analysis (PCA), and statistical analysis

Differential expression genes (DEG) were identified using the R package DESeq2 (Love et al., 2014). We filtered the raw and normalized gene counts according to the following criteria: first, we discarded any gene with a non-normalized count of less than 10 across all 34 samples. This step discards outliers and improves normalizations. Second, we normalized the remaining genes using the **estimateSizeFactors** command implemented in the DESeq2 package and discarded any genes with a normalized count of less than five across less than 30 of the 34 samples. This step aimed to discard genes with outliers if only a few samples had high normalized DEGs. Before filtering, the initial gene count was 39,756, and after the two filtering steps, the count was reduced to 14,410 genes.

Next, we examined Cook's distances of the samples and performed Principal Component Analysis (PCA) with the top 100 highly expressed genes. DESeq2 was used to perform differential expression analysis based on the Negative Binomial Distribution with the DESeq function implemented in the package. We selected log fold change ($|log2FC| > |2|$) to determine significance and adjusted p-values with FDR < 0.05 as thresholds.

Results

Cook's distances and PCA

The Cook's distance is a statistical analysis used to identify outliers by examining the influential points in a least-squares regression model. The Cook's distances for each sample can be seen in **Figure 3**, where the A and B symbols next to the genotype names represent two different seed sources: GH and TB_Nursery. The Cook's distances indicated no outlier genes in the samples. Additionally, we compared the samples using principal component analysis (PCA) performed on the post-filtered top 100 highly expressed genes (100 out of 14,410 genes). **Figure 4** illustrates the PCA of the genotypes with five replications. Principal Components 1 (PC1) and Principal Components 2 (PC2) explained 18% and 12% of the variation, respectively. Two samples marked as b169 genotypes were clustered with b175 genotypes in PC analysis and thus discarded from the analysis.

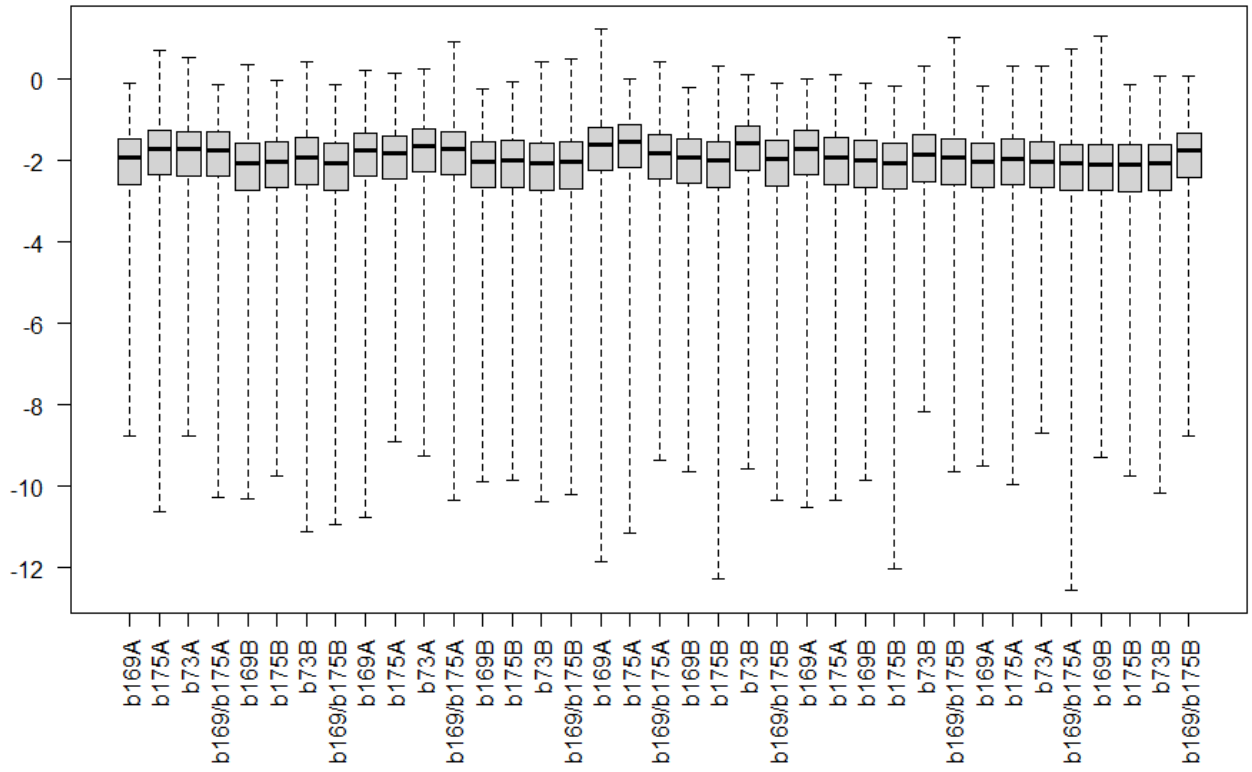


Figure 3: Cook's distances of the DE genes from the sequenced samples. The X-axis indicates the influence points in a linear model, while the Y-axis shows the sequenced samples.

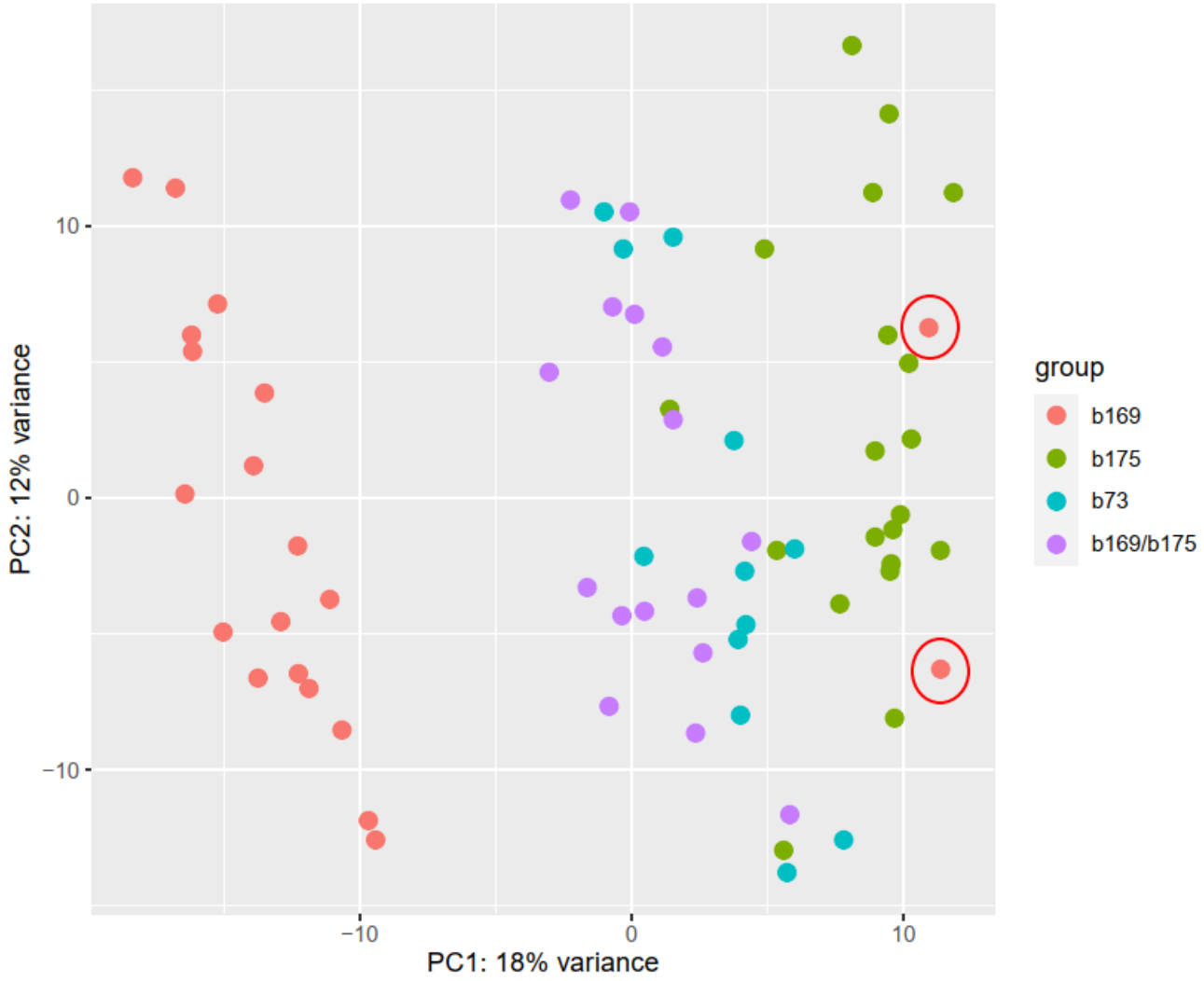


Figure 4: The principal component analysis (PCA) plot illustrates the tested genotypes. The X and Y-axes represent the first and second principal components (PC1 and PC2). The conditions are denoted by two separate seed sources, TB_Nursery and greenhouse (GH). The genotypes are labeled with different symbols. Red circles in the figures are marked as outliers.

DEG analysis

First, we compared different seed sources, and the results are presented in the volcano plot in **Figure 5**. The dots in the figure represent the DEGs, and we assessed the significance based on |log2FC| > |2| and adj p-values < .05 thresholds, which are marked in red. However, we found no significant DEGs between the two seed sources.

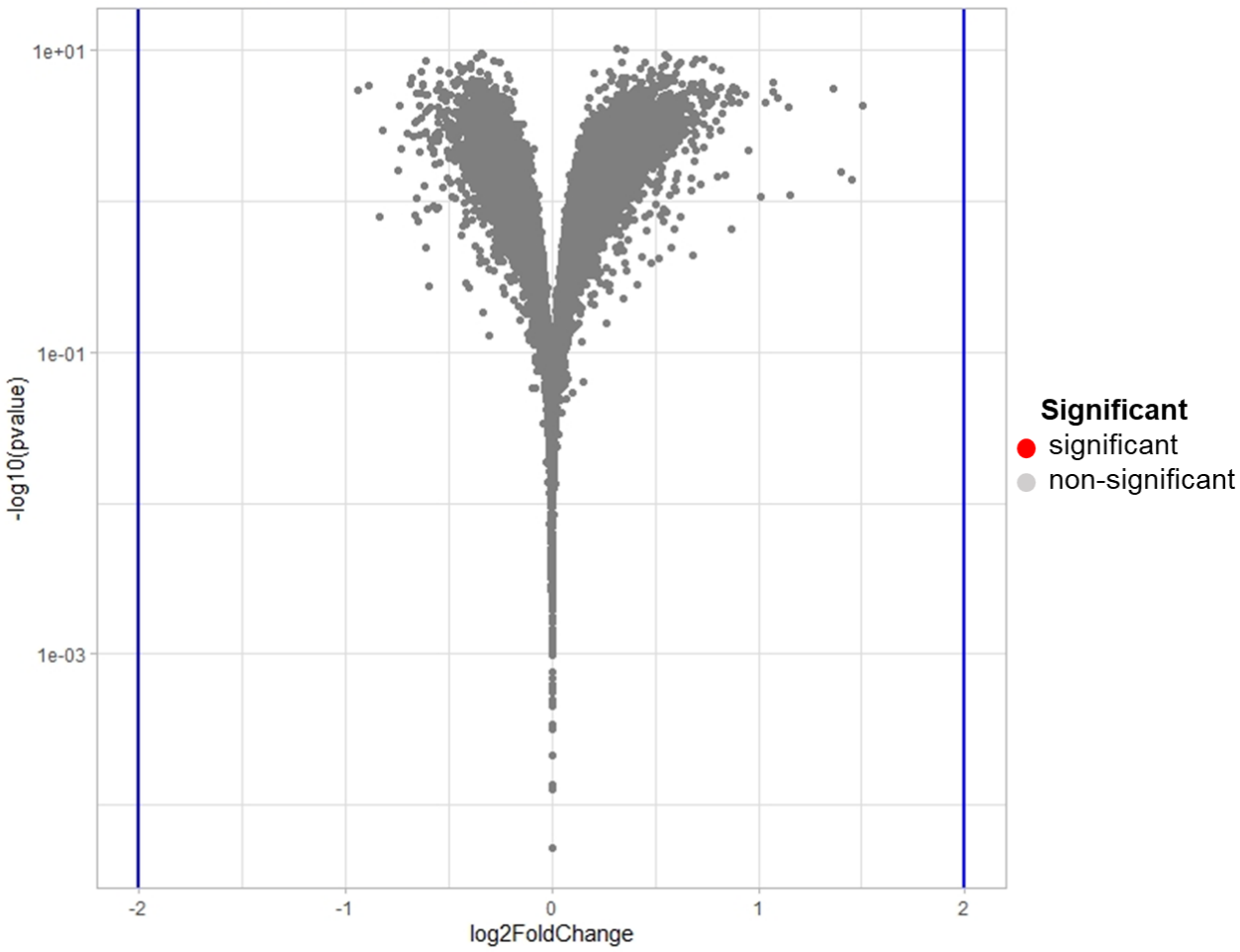


Figure 5: Volcano plot of differently expressed genes (DEGs) between the two seeds of origin. X-axes illustrate the log fold changes between the seed sources (log2FoldChange), and the Y-axes demonstrate the p-values of each data point. Significance is determined if the data point exceeds the -2 or +2 log fold change and, in addition, if the adjusted p-value is smaller than 0.05 (Padj < 0.05).

Second, we compared seeds of origins within the groups, and the comparison was with a low sample; DE genes were not found in the two comparisons, but we found one DEG for B73 and b175 and no DEG for b169 and b169/b175.

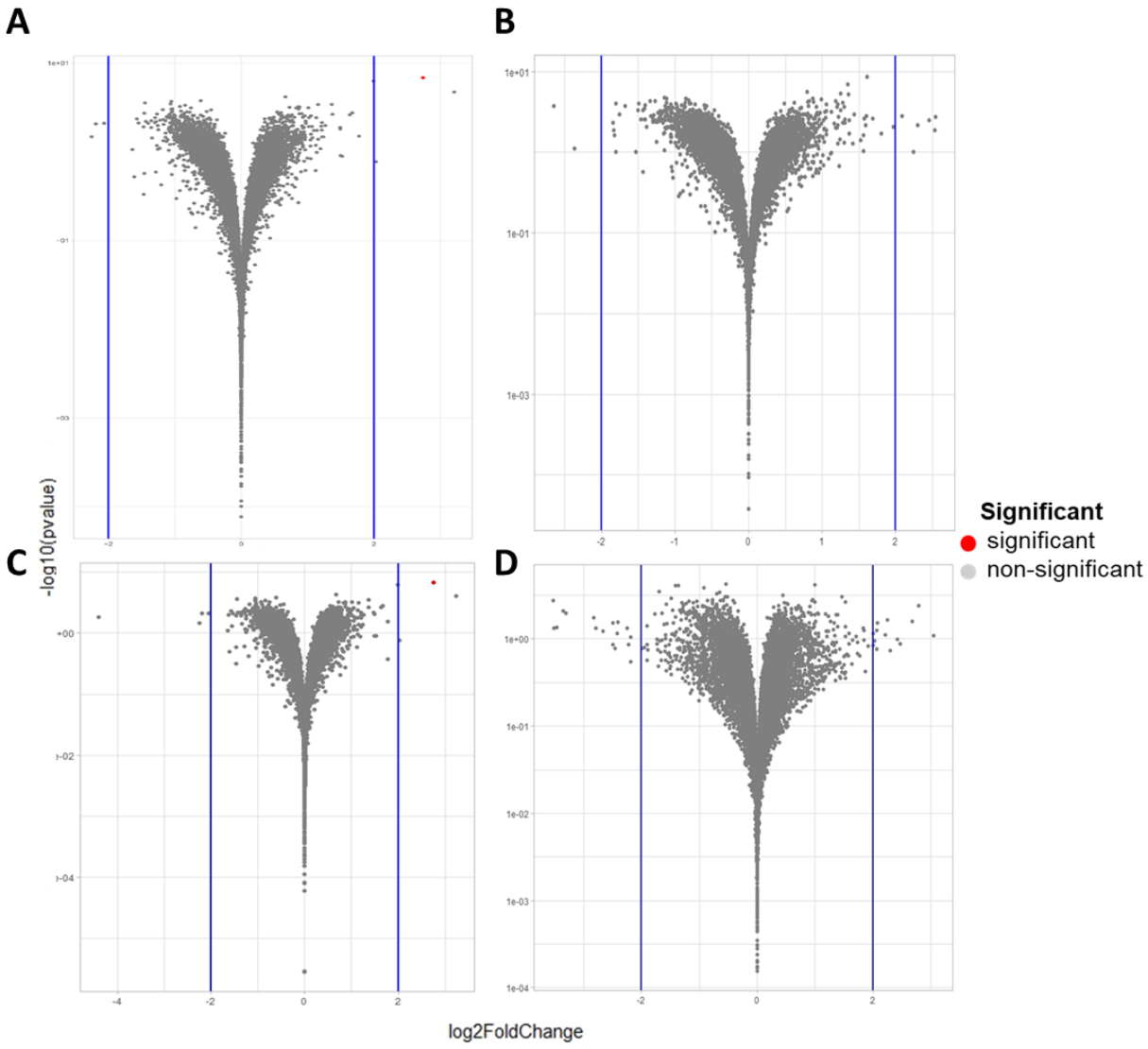


Figure 6: Volcano plot of differently expressed genes (DEGs) between the two seeds of origin within different groups of genotypes. b169 (**A**), b175 (**B**), B73 (**C**), b169/b175 (**D**). X-axes illustrate the log fold changes between the seed sources (log2FoldChange), and the Y-axes demonstrate the p-values of each data point. Significance is determined if the data point exceeds the -2 or +2 log fold change and, in addition, if the adjusted p-value is smaller than 0.05 (Padj < 0.05).

Next, we compared different genotypes with each other. **Figure 7** shows the DEGs for each genotype comparison (A, B, C, D, E, F). We identified 28, 24, 17, 71, 30, and 20 upregulated or downregulated DEGs between b169/b175 & B73, b169/b175 & b169, b169/b175 & b175, B73 & b169, B73 & b175, and b169 & b175, respectively. **Figure 8** displays a heatmap of each comparison.

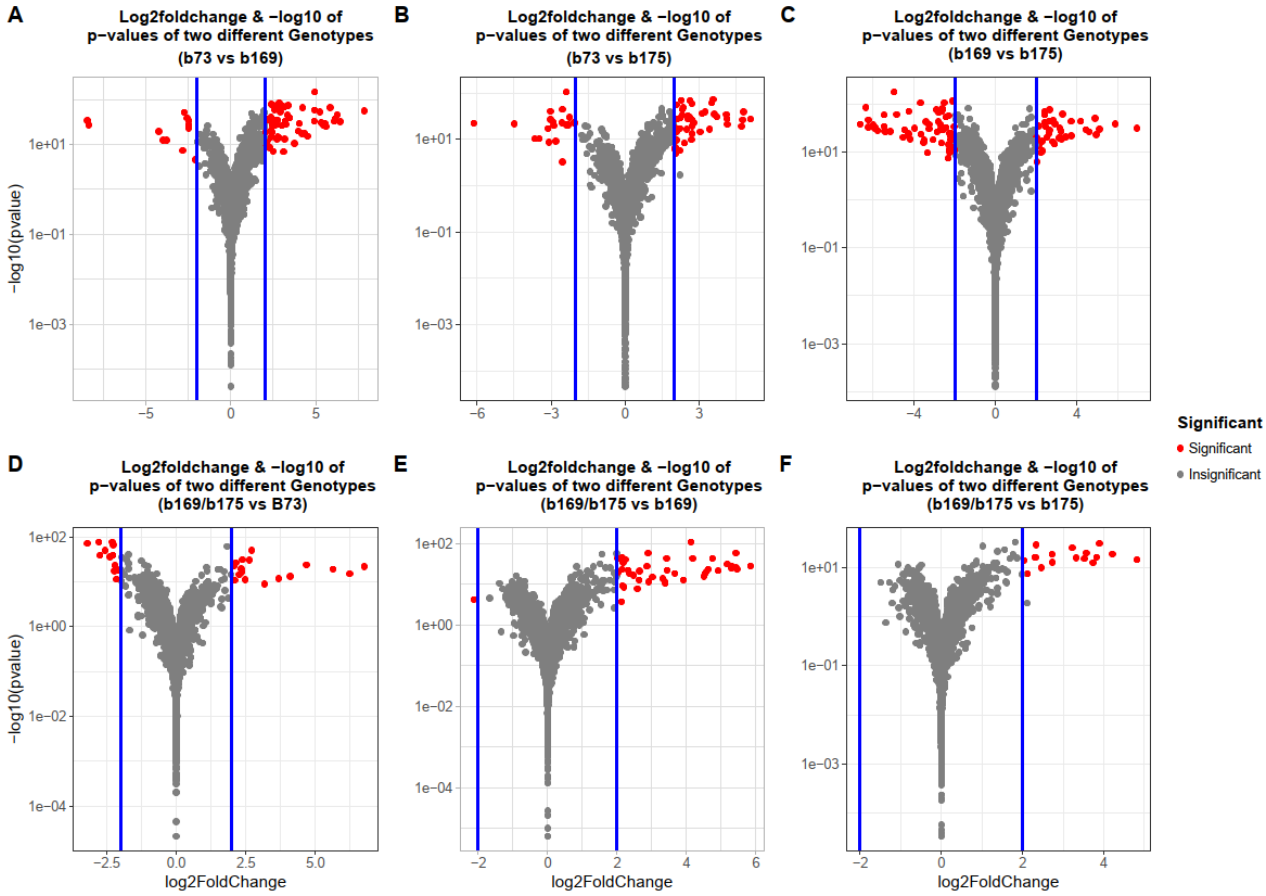


Figure 7: Volcano plot of differently expressed genes (DEGs) between the two genotypes. B73 vs. b169 (**A**), B73 vs. b175 (**B**), b169 vs. b175 (**C**), b169/b175 vs. B73 (**D**), b169/b175 vs. b169 (**E**), b169/b175 vs. b175 (**F**). and the Y-axes demonstrate the p-values of each data point. Significance is determined if the data point exceeds the -2 or +2 log fold change and, in addition, if the adjusted p-value is smaller than 0.05 (Padj < 0.05).

For example, the DEG Zm00001eb017120 on chromosome 1 was reported in previous studies to be significant between the genotypes b169/b175 & b169 and B73 & b169, and it is involved in managing the Terponoid defense compound (Fontana et al., 2011; Block et al., 2019). Another DEG, Zm00001eb098980, on chromosome 2 was found to be significant between B73 & b175 and b169 & b175 and is known to play a role in encoding plastid RNA polymerase complex, DNA binding, and protein binding (Majeran et al., 2012; Williams-Carrier et al., 2014).

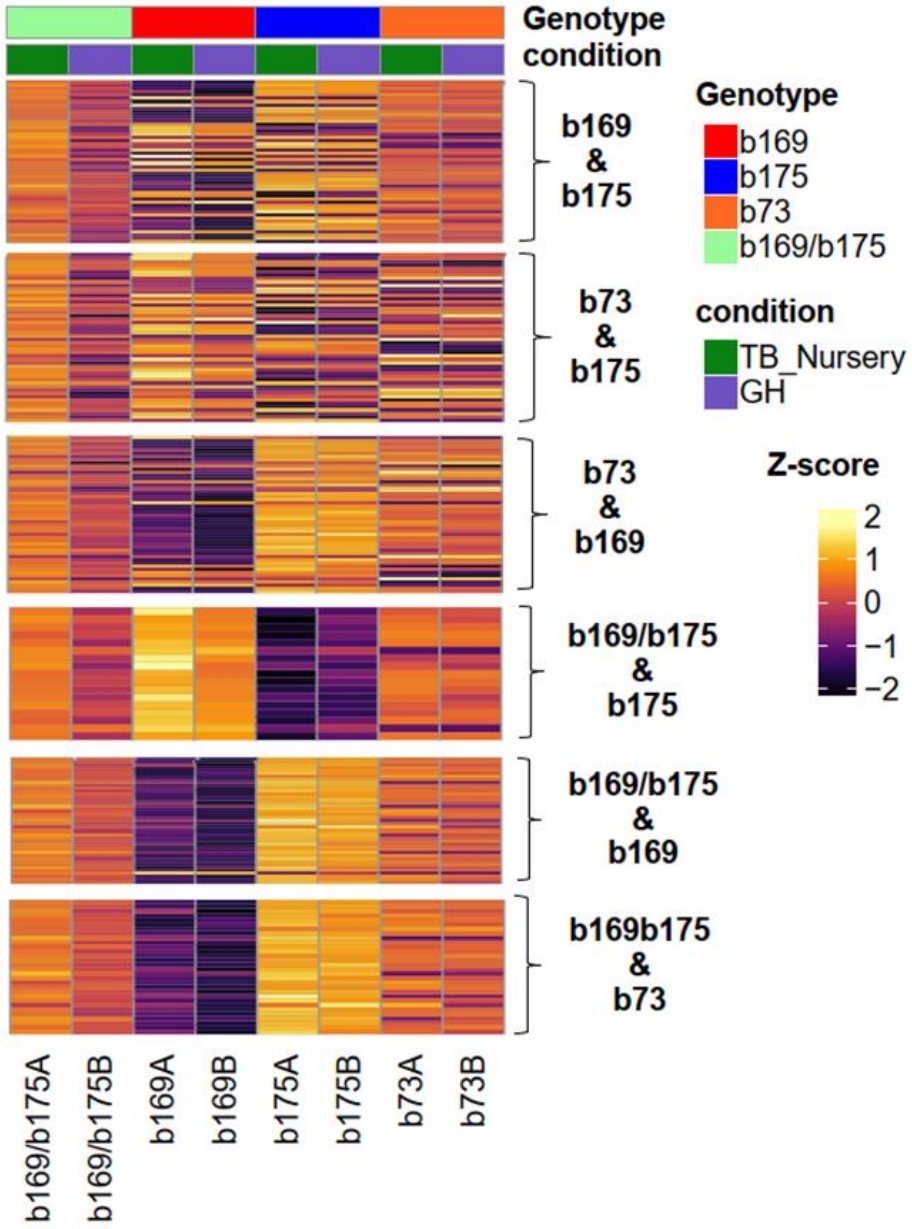


Figure 8: Heatmap of differently expressed genes (DEGs) between genotypes. DEGs are centered according to Z-scores (Right side of the figure). Genotypes and seed sources are indicated with four and two colors (red, blue, orange, and light-green for genotypes. Dark-green and purple for origin of seed), respectively. b169 vs. b175, B73 vs. b175, B73 vs. b169, b169/b175 vs. b175, b169/b175 vs. b169, b169/b175 vs. B73, first, second, third, fourth, fifth, and sixth categories, respectively. Each box represents DEGs between the genotypes and conditions.

Discussion

Epistasis mapping population (EMP)

Beissinger and his group created an epistasis mapping population (EMP) to study the epistatic interactions affecting quantitative traits. EMP is designed to identify the epistatic interactions between genetic loci contributing to the expression of a quantitative trait. This was achieved by crossing near-isogenic lines (NILs) with each other and with the B73 reference line (Eichten et al., 2011). NILs were created by backcrossing the original inbred line Mo17 to a reference inbred line B73 multiple times, resulting in lines (NILs) that are, in their genetic makeup, nearly identical to B73, except for specific chromosomal regions of interest. Such NILs then were crossed with B73 a last time, creating backcrossed NILs. EMPs aim to compare the phenotypic performance of the backcrossed NILs with B73; and with F1-hybrids between two backcrossed NILs. The trait differences between two backcrossed NILs and B73 should add up to the difference between their F1-hybrid and B73. If this was the case, no epistatic interactions were detected; otherwise, epistasis occurs between the chromosomal regions on the genome that are represented by any two focal backcrossed NILs and their F1-hybrid. Utilizing EMPs will help reduce the multiple-testing problem in such epistasis research, will hence increase the ability to detect epistatic interactions, and will reduce false positive results. In addition to identifying epistatic interactions, EMPs may provide insight into the underlying mechanisms of gene regulation and the role of epistasis in the evolution of species. By understanding the genetic basis of quantitative traits, EMPs can contribute to increasing the efficiency of plant breeding methods and strategies. In conclusion, EMPs are a useful tool in studying the genetics of quantitative traits, providing valuable information about the role of epistasis in trait determination and enabling the development of more effective breeding strategies.

Origin of seed

Previous studies have investigated the effect of environments in following generations on late-growth stage phenotypic differentiation (Herman and Sultan, 2011; Kou et al., 2011). Dyer et al., 2010, investigated Aegilops (*Aegilops triuncialis*) offsprings with parents grown in different soil conditions, and they observed early flowering of the Aegilops offspring of parents grown in serpentine soil (Dyer et al., 2010). In our study, we focused on the effect of origin-of-seed on

maize seedlings at the V2 stage, and we used a subset of genotypes from EMPs from two different seed sources. Our main objective in this study was to test the origin of the seed effect. Since there is a lack of a full EMP set from one origin, we did not investigate epistatic interactions, which require a complete set of NILs, BC-NILs, and F1NIL hybrids. Although desirable, adtually obtaining the complete set from only one seed source was challenging. Therefore, to inprect this item, we set up a preliminary experiment to determine whether gene expression differed between the two here-used seed sources (origins of the seed).

Investigating gene expression differentiation between genotypes is time-consuming due to the computational challenges of aligning mRNAs to the reference genome. However, software options are available to align libraries to the reference genome, such as Bowtie2 and STAR. To overcome these challenges, we used Kallisto and aligned the transcripts to the reference transcriptome. This approach significantly reduced the alignment time, taking approximately five minutes per sample and three hours for all samples combined. Additionally, Kallisto is known for its high accuracy compared to other software options (Du et al., 2020).

Our study focused on EMPs to examine DEGs in the progeny from two distinct seed sources, aiming to understand whether diverse environmental conditions affect the progeny of parents producing seed in those environments. In contrast to previous studies, upregulated and downregulated genes were insignificant between the two seed sources. However, it is worth noting that previous studies examined different growth stages of plants compared to our study. Additionally, our study only analyzed mRNA expression, and further statistical analyses are needed to evaluate potential methylation differences between genetically identical genotypes to fully understand the epigenetic effects on the next generation. Furthermore, the specific environmental conditions to which the parents of the progeny were exposed remain unknown. As previous studies have indicated, stress factors can influence progeny, allowing for the inheritance of epigenetic factors to subsequent generations. Although parents of the genotypes grew in distinct environmental conditions in our experiment, since the environmental conditions are greenhouse and field, the expected environmental conditions exposed were not extreme. Therefore, the observed results match our null hypothesis.

Although the comparison of mRNA levels between different seed sources showed no differentially expressed (DE) genes, a single DE gene was found within the B73 and b175 groups. On the other hand, the expected distribution of DGE distribution between two comparisons is not one but many genes, as shown in the genotype comparison in **Figure 7**. In addition, in this study, the sample size for comparing genotypes is lower than for origin-of-seed comparison with two seed sources. Therefore, previous studies have shown that low sample sizes can decrease the precision and power of differential expression analysis, leading to false positives (Stretch et al., 2013; Ching et al., 2014). Hence, comparing the seed sources within or between the genotypes in this study may not have been entirely accurate due to the low sample size, and false positives may have been expected. Despite the low sample size between each genotype comparison, several DE genes were observed for each different genotype comparison.

After analyzing the Differentially Expressed Genes (DEGs), we found no significant expression differences between the seed sources from the field and the greenhouse in maize seedlings at the V2 growth stage. In addition, to validate the reliability of our analysis, we compared different genotypes, and as expected, we observed DEGs among all possible genotype combinations. Therefore, we can conclude that DEGs between the seed sources are not large since we would have observed them frequently. Therefore, the probability of interfering with the origin-of-seed effects for further analysis with EMPs is low.

Chapter 4: General Discussion

Importance of epistasis and previous findings

Genetic variation is one of the primary drivers of phenotypic variation within and between populations. Genetic variation can arise be steered via various sources, including mutations, gene flow (migration), random genetic drift, and natural and artificial selection (Barton et al., 1997). Additionally, genetic variation can be influenced by the history of a given population, including its effective population size, its genetic structure, and by past migration events (Epperson, 1993; Gomez et al., 2014). Environmental factors, such as soil quality and temperature, usually also markedly impact the phenotypes of plants (Chesson, 1986; Hill and Mulder, 2010). Moreover, environmental factors can interact with genetic factors to create complex and often unpredictable phenotypic outcomes (Hill, 1975; Via and Lande, 1985). Therefore, a plant genotype performing better-than-others in one environment may perform inferior in another setting.

It is widely accepted that the phenotypic performance of genotype is – for most agronomically relevant traits – to a high extent determined by genetics. Sir Ronald Fisher previously described genetic variation as a heritable contribution to phenotypic variation, which he subdivided into variation due to additive effects, dominance effects, and epistasis (Fisher, 1930; Meredith Jr., 1984; Barton and Turelli, 1989). Additive genetic variation is that part of the genetic variation that is usually inherited in future (sexually produced) generations and has been extensively studied in animal and plant breeding (Falconer, 1965; Moll and Stuber, 1974; Leary et al., 1985). Many studies have produced similar results, with additive variation accounting for most of the genetic variation in a population. However, some studies have emphasized the importance of dominance and epistasis variation (Willis and Orr, 1993). In a study on *Arabidopsis thalia*, researchers found that epistatic interactions between two loci affected flowering time (Koornneef et al., 2004). Another study on rice indicated that epistatic interactions influence grain yield (Li et al., 1997). According to classical Quantitative Genetic theory (Falconer, 1996) and a simulation study by Carter et al. (2005), epistatic interactions can be relevant and even increase the additive genetic variation. Thus, with positive directional epistasis, the population has a higher local maximum compared to only additive interactions (Carter et al., 2005). Gambe, in 1962 studied 15 maize

crosses and concluded that dominance and epistasis effects are more important for inheritance than additive effects. Also, epistasis variation was larger than additive effects (Gamble, 1962). In his study, Gamble utilized a mean comparison of the progenies and the parents. Although means can indicate epistatic interactions, negative and positive epistatic interactions can cancel each other out. Therefore, estimating the genetic parameters based on variation can improve the detection of epistasis interaction, and a well-known design of this approach is the North Carolina design (Hohls, 1996; Abdallah and McDaniel, 2000). However, one drawback of this approach is that it requires large population size. Another study compared hybrid maize lines, F1, F2, F3, back-crosses, second back-crosses, and selfed back-cross generations in acid and non-acid soils and found low variation caused by epistasis in non-acid soils but significant epistasis in acid soils. In addition, they notably indicated increasing epistasis interaction with GxE interactions (Ceballos et al., 1998). As we think in the simulation study of Carter et al. (2005) and the study of Ceballos et al. (1998), epistasis plays a crucial role in evolving a population in a particular environment. In self-fertilizing plants such as wheat, additive genetic variation can explain most of the variation in a population. Ketata et al. (1976) found that additive genetic variation was more important than dominance and epistatic genetic variants in wheat; in addition, they cautioned that a model without epistatic interactions could lead to biased results (Ketata et al., 1976). A study on maize found significant epistatic interactions under different environmental conditions, well-watered, and drought. This suggests epistatic interactions can significantly determine how plants respond to changing environmental conditions, which is especially relevant given climate conditions (Messmer et al., 2009).

Additive genetic variation and, to some extent, dominance-related genetic variation have traditionally been the focus of plant breeding efforts for many years, but researchers have increasingly recognized epistatic interactions' importance. Taking epistasis into account can lead to a more accurate understanding of the inheritance of traits and could ultimately help to develop more effective and sustainable plant breeding strategies (Moore and Williams, 2009).

QTL by environment and epistasis, and epistasis by environment interactions

QTL mapping studies' power and resolution are closely tied to the specific population and environment being studied. In our research, we have identified significant QTLs and compared

them to similar and dissimilar peaks on chromosomal locations reported in previous studies in the literature. Notably, we observed LOD score changes in QTLs across different environmental conditions, suggesting an interaction between QTLs and the environment. Such interactions have been documented earlier in various studies in the literature.

As demonstrated in our analysis, environmental conditions can influence certain QTLs' effects. These interactions have been captured using a linear mixed model, as discussed in the first chapter of our study. Given the ongoing impact of global warming, understanding these QTL-environment interactions is crucial for the breeding goal of developing adapted plant materials for specific local conditions.

Furthermore, it is essential not to overlook nor neglect existing epistatic gene-to-gene interactions, as they can modulate and enhance genotype performance in future generations. As previous studies have indicated, detecting agronomically favorable epistatic interactions among QTLs may be just as essential as identifying QTLs related to the trait of interest.

Moreover, epistasis-environment interactions can play a decisive role in enhancing genotypes that may have reached their potential trait performance. Our study provides valuable insights into the field of plant breeding, shedding light on the interactions between QTLs and the environment, as well as the potential for epistasis-environment interactions in shaping future breeding programs.

Epistasis mapping population (EMP) and differential expression genes (DEG)

Detecting the epistatic interaction between loci is challenging and requires a large sample size of a research population. In addition, statistical challenges to detect significant epistatic interaction, such as multiple testing problems, sample size limitation, and computational challenges, must not be forgotten. Multiple testing problems arise when many statistical tests are conducted simultaneously. Multiple testing increases the occurrence of false positives in statistical analysis, and p-values need correction in order to avoid false positives. In association studies, researchers test thousands or millions of genetic markers or other variables for their association with a particular outcome, such as for association between marker genotypes and phenotypic differences between the pertinent genotype groups. Particular methods have been developed to

address this issue. One and most used method is to adjust the significance level to reduce false positives in an experiment. Researchers use the Bonferroni correction and false discovery rate (FDR) test as popular methods (Verhoeven et al., 2005). Both methods aim to reduce type I errors due to multiple testing to detect true positive p-values. FDR is less stringent with false positives than Bonferroni at the cost of increasing type I errors. Either way, adjusting the p-values does not solve the fundamental problem of detecting the epistatic interactions because both methods also eliminate statistically true epistatic interactions. Therefore, minor epistatic interactions are undetected with statistical analysis. Other approaches to reduce this dilemma are increasing the sample size or decreasing the number of markers used in the statistical analysis (such reduction could be, e.g., based on linkage disequilibrium between used and kicked-out markers). However, the first method increases the costs of phenotyping and genotyping. The second method is decreasing the computational time to detect the epistatic interactions, but the precision and resolution of the interactions depend on the existing and exploited LD and multiple testing problems are not completely eradicated.

Epistasis mapping populations (EMPs) were specifically created to address the issue of multiple testing. EMPs utilize near-isogenic lines (NILs). In our case, these NILs were developed by Eichter and his group (Eichten et al., 2011). EMPs aim to test randomly introgressed chromosomal regions instead of testing all (i.e., very many) marker loci that can be displayed by the available assays, residing more or less anywhere in the genome. The advantage of testing a relatively small number of regions comes with the disadvantage that there will be several or many QTLs within any of these regions. Their small, large, positive, and negative epistatic interactions with the several or many QTLs on any other region may cancel out. Hence, if no epistasis is detected, this does not propose that there is none. On the other side, if epistatic interactions are indeed detected with that approach, they are not less real and convincing as if accruing with alternative approaches.

In this preliminary study, we examined the expression differences between two seed sources of EMPs: seed sources were greenhouse and field environment. Differentially Expressed Genes (DEGs) between these origins of the used seed would refer to variations in gene expression patterns resulting from differences in these two environmental conditions. Our analysis did not identify significantly upregulated or downregulated DEGs between the two distinct seed origins.

As expected, the up or downregulation of DEGs between different genotypes confirmed the reliability of the analysis.

Some studies have suggested that environmental effects, possibly through epigenetic mechanisms, can influence seed dormancy levels and phenotypic observations in genetically identical genotypes under diverse environmental conditions (Fenner, 1991; Penfield and MacGregor, 2017). However, the inheritance of these epigenetic effects from parents to offspring remains a question. For instance, Elwell et al. (2011), investigated the effect of seed size on the next generation of *Arabidopsis thaliana*. They sowed and harvested seeds under six diverse environmental conditions and subsequently observed traits such as primary root growth, gravitropism, and flowering time based on seed size. They concluded that historical effects on parents could influence progeny phenotype (Elwell et al., 2011). Molinier et al. (2006) with *Arabidopsis thaliana* plants treated with ultraviolet-C or flagellin observed increased somatic homologous recombination of a transgenic reporter in subsequent generations (Molinier et al., 2006). Lira-Medeiros et al. (2010) focused on phenotypic differences between SM and RS locations and found distinct observations for quantitative traits such as tree diameter, leaf width, and leaf area. The researchers later discovered that one location (SM) exhibited hypomethylation compared to the other (RS) (Lira-Medeiros et al., 2010). However, genetic investigation of the inheritance of epigenetic factors to the next generations remains limited.

General conclusion

Here, genetic and breeding research work on maize with a focus on epistasis and epistasis-environment interactions was presented. QTL mapping was conducted for several traits analyzed either across all locations or specific to mega-environments. Significant QTLs were identified for each trait, with varying numbers detected in different mega environments. The analysis provided valuable information about the genomic regions associated with these traits. It is crucial to understand how the impact of segregating QTLs on trait variation changes if environmental parameters change, and for the following reasons: firstly, cultivar development unavoidably depends on our ability to predict the agronomic value of here-and-now tested candidates in farmers' field condition, there-and-then; and secondly, this level of difficulty of this task is increased by the ongoing climate change because this makes our historical data less informative (Hansen et al., 2010; Aerni, 2023).

Estimating the QTL effects on different environments can increase our analytic understanding and, hence, the usefulness of such QTL data in a breeding program. Therefore, estimates of the interaction effects of relevant QTLs with meaningful environments will probably help improve any breeding program.

Epistasis QTL mapping revealed significant epistatic interactions between various sections of chromosomes for several pertinent traits across different environments. These interactions can provide insights into the complex genetic interactions underlying these traits. Breeders often neglect epistatic interactions in a breeding program. The main focus so far has been to estimate and employ additive genetic variation. Meanwhile, the global challenges are severe enough to put us on the lookout even for secondary and tertiary important sources of genetic variation. So, unquestionably, epistasis should not be ignored anymore because of its potential to improve breeding populations. Complex interaction can help populations overcome current metabolic and adaptation limitations, and breeders might discover individual and genotypes that surpass all sofar known genotypes.

Lastly, our study explored epistasis by environment interactions (EEI) to understand how different environments might influence genetic interactions. Significant EEI effects were observed for

pollen days to anthesis, silk days to anthesis, plant height, and yield. These findings indicate that the genetic control of these traits can be influenced by the specific environment in which they are assessed. Our study indicates that effects of environments on epistatic interactions exist. We are pointing out that the breeding programs that consider epistatic interactions should be aware of the potential environmental effects on epistasis, and these interactions might alter the observed effect of epistatic interactions. Therefore, future breeding programs should evaluate epistatic interactions across a range of diverse environmental conditions or consider environmental factors when advertising their products.

The origin-of-seed study comprehensively analyzed genetic data and gene expression patterns in various maize genotypes and seed sources. The analysis included an outlier assessment using Cook's distances, principal component analysis (PCA) for data visualization, and a comparison of differentially expressed genes (DEGs) to uncover genetic variations. The study aimed to investigate the impact of origin-of-seed on maize seedlings at the V2 growth stage, with a specific focus on understanding epigenetic interactions using gene expression data. The analysis was conducted with the available resources and a subset of genotypes consisting of EMP parents, BC-NILs, and F1NIL hybrids from different seed sources.

While no significant differences in gene expression were observed between different seed sources, the DEG analysis revealed specific genes that were significantly differentially expressed when comparing different maize genotypes. This suggests that, in our experimental conditions, diverse environmental conditions did not lead to detectable gene expression differences in the progeny of parents that produced the seed in these environments. It is important to note that our study focused solely on mRNA expression and did not explore methylation differences, essential for a comprehensive understanding of epigenetic effects, constituting a complex aspect of epigenetics.

Overall, The EEI study provided a comprehensive analysis of genetic and phenotypic data for various traits in maize. The results contribute to our understanding of the genetic architecture and relationships between these traits, which can be invaluable for optimizing maize production in future breeding efforts. Identifying significant QTLs and epistatic interactions across different

environments opens new avenues for targeted genetic improvement in maize breeding programs.

References

Abdallah, J.M., and B.T. McDaniel. 2000. Genetic Parameters and Trends of Milk, Fat, Days Open, and Body Weight After Calving in North Carolina Experimental Herds1. Journal of Dairy Science 83(6): 1364–1370. doi: 10.3168/jds.S0022-0302(00)75004-1.

Aerni, P. 2023. COP-27: A great opportunity to address the double crisis of food security and climate change–and for the EU to re-align its farm to fork strategy. Frontiers in Environmental Economics 1. https://www.frontiersin.org/articles/10.3389/frevc.2022.1082869 (accessed 19 November 2023).

Agbicodo, E.M., C.A. Fatokun, S. Muranaka, R.G.F. Visser, and C.G. Linden van der. 2009. Breeding drought tolerant cowpea: constraints, accomplishments, and future prospects. Euphytica 167(3): 353–370. doi: 10.1007/s10681-009-9893-8.

Altshuler, D., M.J. Daly, and E.S. Lander. 2008. Genetic Mapping in Human Disease. Science 322(5903): 881–888. doi: 10.1126/science.1156409.

Amanullah, D., and S. Khalid. 2020. Agronomy: Climate Change. BoD – Books on Demand.

Anastasiadi, D., C.J. Venney, L. Bernatchez, and M. Wellenreuther. 2021. Epigenetic inheritance and reproductive mode in plants and animals. Trends in Ecology & Evolution 36(12): 1124–1140. doi: 10.1016/j.tree.2021.08.006.

Annicchiarico, P. 2002. Genotype X Environment Interactions: Challenges and Opportunities for Plant Breeding and Cultivar Recommendations. Food & Agriculture Org.

Ardlie, K.G., L. Kruglyak, and M. Seielstad. 2002. Patterns of linkage disequilibrium in the human genome. Nat Rev Genet 3(4): 299–309. doi: 10.1038/nrg777.

Ashe, A., V. Colot, and B.P. Oldroyd. 2021. How does epigenetics influence the course of evolution? Philosophical Transactions of the Royal Society B: Biological Sciences 376(1826): 20200111. doi: 10.1098/rstb.2020.0111.

Atkinson, N.J., and P.E. Urwin. 2012. The interaction of plant biotic and abiotic stresses: from genes to the field. Journal of experimental botany 63(10): 3523–3543.

Awika, J.M. 2011. Major Cereal Grains Production and Use around the World. Advances in Cereal Science: Implications to Food Processing and Health Promotion. American Chemical Society. p. 1–13

Barchi, L., S. Lanteri, E. Portis, G. Valè, A. Volante, et al. 2012. A RAD Tag Derived Marker Based Eggplant Linkage Map and the Location of QTLs Determining Anthocyanin Pigmentation. PLOS ONE 7(8): e43740. doi: 10.1371/journal.pone.0043740.

Bartels, A., Q. Han, P. Nair, L. Stacey, H. Gaynier, et al. 2018. Dynamic DNA Methylation in Plant Growth and Development. International Journal of Molecular Sciences 19(7): 2144. doi: 10.3390/ijms19072144.

Barton, N.H., J. Mallet, B.C. Clarke, and P.R. Grant. 1997. Natural selection and random genetic drift as causes of evolution on islands. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 351(1341): 785–795. doi: 10.1098/rstb.1996.0073.

Barton, N.H., and M. Turelli. 1989. Evolutionary Quantitative Genetics: How Little Do We Know? Annual Review of Genetics 23(1): 337–370. doi: 10.1146/annurev.ge.23.120189.002005.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using **lme4**. J. Stat. Soft. 67(1). doi: 10.18637/jss.v067.i01.

Baulcombe, D.C., and C. Dean. 2014. Epigenetic Regulation in Plant Responses to the Environment. Cold Spring Harb Perspect Biol 6(9): a019471. doi: 10.1101/cshperspect.a019471.

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

Begna, T. 2021. Global role of plant breeding in tackling climate change. International Journal of Agricultural Science and Food Technology: 223–229. doi: 10.17352/2455-815X.000111.

Bender, R., and S. Lange. 2001. Adjusting for multiple testing—when and how? Journal of Clinical Epidemiology 54(4): 343–349. doi: 10.1016/S0895-4356(00)00314-0.

Bendixsen, D.P., B. Østman, and E.J. Hayden. 2017. Negative Epistasis in Experimental RNA Fitness Landscapes. J Mol Evol 85(5): 159–168. doi: 10.1007/s00239-017-9817-5.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 57(1): 289–300. doi: https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Berger, S.L., T. Kouzarides, R. Shiekhattar, and A. Shilatifard. 2009. An operational definition of epigenetics. Genes & Development 23(7): 781. doi: 10.1101/gad.1787609.

Block, A.K., M.M. Vaughan, E.A. Schmelz, and S.A. Christensen. 2019. Biosynthesis and function of terpenoid defense compounds in maize (Zea mays). Planta 249(1): 21–30. doi: 10.1007/s00425-018-2999-2.

Bloom, J.S. 2015. Genetic interactions contribute less than additive effects to quantitative trait variation in yeast. NATURE COMMUNICATIONS: 6.

Borlaug, N.E. 1983. Contributions of Conventional Plant Breeding to Food Production. Science 219(4585): 689–693. doi: 10.1126/science.219.4585.689.

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, et al. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23(19): 2633–2635. doi: 10.1093/bioinformatics/btm308.

Bray, N.L., H. Pimentel, P. Melsted, and L. Pachter. 2016. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol 34(5): 525–527. doi: 10.1038/nbt.3519.

Breseghello, F., and A.S.G. Coelho. 2013. Traditional and Modern Plant Breeding Methods with Examples in Rice (Oryza sativa L.). J. Agric. Food Chem. 61(35): 8277–8286. doi: 10.1021/jf305531j.

Brisco, B., R.J. Brown, T. Hirose, H. McNairn, and K. Staenz. 1998. Precision Agriculture and the Role of Remote Sensing: A Review. Canadian Journal of Remote Sensing 24(3): 315–327. doi: 10.1080/07038992.1998.10855254.

Broman, K.W., H. Wu, S. Sen, and G.A. Churchill. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19(7): 889–890. doi: 10.1093/bioinformatics/btg112.

Cai, H., Q. Chu, R. Gu, L. Yuan, J. Liu, et al. 2012. Identification of QTLs for plant height, ear height and grain yield in maize (Zea mays L.) in response to nitrogen and phosphorus supply. Plant Breeding 131(4): 502–510. doi: 10.1111/j.1439-0523.2012.01963.x.

Carlborg, Ö., and C.S. Haley. 2004. Epistasis: too often neglected in complex trait studies? Nat Rev Genet 5(8): 618–625. doi: 10.1038/nrg1407.

Carter, A.J.R., J. Hermisson, and T.F. Hansen. 2005. The role of epistatic gene interactions in the response to selection and the evolution of evolvability. Theoretical Population Biology 68(3): 179–196. doi: 10.1016/j.tpb.2005.05.002.

Carvalho, F.P. 2006. Agriculture, pesticides, food security and food safety. Environmental Science & Policy 9(7): 685–692. doi: 10.1016/j.envsci.2006.08.002.

Carvalho, F.P. 2017. Pesticides, environment, and food safety. Food and Energy Security 6(2): 48–60. doi: 10.1002/fes3.108.

Casassola, A., S.P. Brammer, M. Soares Chaves, J.A. Martinelli, M.F. Grando, et al. 2013. Gene Expression: A Review on Methods for the Study of Defense-Related Gene Differential Expression in Plants. AJPS 04(12): 64–73. doi: 10.4236/ajps.2013.412A3008.

Ceballos, H., S. Pandey, L. Narro, and J.C. Perez-Velázquez. 1998. Additive, dominant, and epistatic effects for maize grain yield in acid and non-acid soils: Theor Appl Genet 96(5): 662–668. doi: 10.1007/s001220050786.

Chanock, S. 2001. Candidate genes and single nucleotide polymorphisms (SNPs) in the study of human disease. Disease Markers 17(2): 89–98.

Chen, K., Y. Wang, R. Zhang, H. Zhang, and C. Gao. 2019. CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. Annual Review of Plant Biology 70(1): 667–697. doi: 10.1146/annurev-arplant-050718-100049.

Chesson, P.L. 1986. Environmental variation and the coexistence of species. Community ecology 240: 54.

Ching, T., S. Huang, and L.X. Garmire. 2014. Power analysis and sample size estimation for RNA-Seq differential expression. RNA 20(11): 1684–1696. doi: 10.1261/rna.046011.114.

Clement, C.R. 1999. 1942 and the Loss of Amazonian Crop Genetic Resources. I. The Relation between Domestication and Human Population Decline. Economic Botany 53(2): 188–202.

Coakley, S.M., H. Scherm, and S. Chakraborty. 1999. Climate Change and Plant Disease Management. Annual Review of Phytopathology 37(1): 399–426. doi: 10.1146/annurev.phyto.37.1.399.

Collard, B.C.Y., and D.J. Mackill. 2007. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philosophical Transactions of the Royal Society B: Biological Sciences 363(1491): 557–572. doi: 10.1098/rstb.2007.2170.

Cordell, H.J. 2002. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Human Molecular Genetics 11(20): 2463–2468. doi: 10.1093/hmg/11.20.2463.

Crossa, J., P. Pérez-Rodríguez, J. Cuevas, O. Montesinos-López, D. Jarquín, et al. 2017. Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. Trends in Plant Science 22(11): 961–975. doi: 10.1016/j.tplants.2017.08.011.

Culverhouse, R., T. Klein, and W. Shannon. 2004. Detecting epistatic interactions contributing to quantitative traits. Genetic Epidemiology 27(2): 141–152. doi: 10.1002/gepi.20006.

Datta, A., A. Hossain, and S. Roy. 2019. An overview on biofuels and their advantages and disadvantages. Asian Journal of Chemistry 8(31): 1851–1858. doi: 10.14233/ajchem.2019.22098.

Diepenbrock, C.H., and M.A. Gore. 2015. Closing the Divide between Human Nutrition and Plant Breeding. Crop Science 55(4): 1437–1448. doi: 10.2135/cropsci2014.08.0555.

Dodgson, J., H. Cheng, and R. Okimoto. 1997. DNA marker technology: a revolution in animal genetics. Poultry Science 76(8): 1108–1114. doi: 10.1093/ps/76.8.1108.

Doebley, J.F., B.S. Gaut, and B.D. Smith. 2006. The Molecular Genetics of Crop Domestication. Cell 127(7): 1309–1321. doi: 10.1016/j.cell.2006.12.006.

Duvick, D.N. 1996. Plant Breeding, an Evolutionary Concept. Crop Science 36(3): cropsci1996.0011183X003600030001x. doi: 10.2135/cropsci1996.0011183X003600030001x.

Dyer, A.R., C.S. Brown, E.K. Espeland, J.K. McKay, H. Meimberg, et al. 2010. SYNTHESIS: The role of adaptive trans-generational plasticity in biological invasions of plants. Evolutionary Applications 3(2): 179–192. doi: 10.1111/j.1752-4571.2010.00118.x.

Edwards, A.W.F. 2008. G. H. Hardy (1908) and Hardy–Weinberg Equilibrium. Genetics 179(3): 1143–1150. doi: 10.1534/genetics.104.92940.

Eichten, S.R., J.M. Foerster, N. de Leon, Y. Kai, C.-T. Yeh, et al. 2011. B73-Mo17 Near-Isogenic Lines Demonstrate Dispersed Structural Variation in Maize. Plant Physiol. 156(4): 1679–1690. doi: 10.1104/pp.111.174748.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, et al. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species (L. Orban, editor). PLoS ONE 6(5): e19379. doi: 10.1371/journal.pone.0019379.

Elwell, A.L., D.S. Gronwall, N.D. Miller, E.P. Spalding, and T.L. Durham Brooks. 2011. Separating parental environment from seed size effects on next generation growth and development in Arabidopsis. Plant, Cell & Environment 34(2): 291–301. doi: 10.1111/j.1365-3040.2010.02243.x.

Epperson, B.K. 1993. Recent Advances in Correlation Studies of Spatial Patterns of Genetic Variation. In: Hecht, M.K., MacIntyre, R.J., and Clegg, M.T., editors, Evolutionary Biology. Springer US, Boston, MA. p. 95–155

Erdmann, R.M., and C.L. Picard. 2020. RNA-directed DNA Methylation. PLoS Genet 16(10): e1009034. doi: 10.1371/journal.pgen.1009034.

Evenson, R.E., and D. Gollin. 2003. Assessing the Impact of the Green Revolution, 1960 to 2000. Science 300(5620): 758–762. doi: 10.1126/science.1078710.

Falconer, D.S. 1965. The inheritance of liability to certain diseases, estimated from the incidence among relatives. Annals of human genetics 29(1): 51–76.

Falconer, D.S. 1996. Introduction to quantitative genetics. Pearson Education India.

FastQC. 2015. https://qubeshub.org/resources/fastqc.

Fenner, M. 1991. The effects of the parent environment on seed germinability. Seed Science Research 1(2): 75–84. doi: 10.1017/S0960258500000696.

Fériani, W., S. Rezgui, and M. Cherif. 2020. Detection of QTL and QTL × environment interaction for scald resistance in a two-row × six-row cross of barley. CEREAL RESEARCH COMMUNICATIONS 48(2): 187–193. doi: 10.1007/s42976-020-00024-1.

Fisher, R.A. 1919. XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. Transactions of the Royal Society of Edinburgh 52(2): 399–433. doi: 10.1017/S0080456800012163.

Fisher, R.A. 1930. The Genetical Theory of Natural Selection JH Bennett, ed. Clarendon Press. INTRODUCTION.

Flint-Garcia, S.A., J.M. Thornsberry, and E.S. Buckler. 2003. Structure of Linkage Disequilibrium in Plants. Annual Review of Plant Biology 54(1): 357–374. doi: 10.1146/annurev.arplant.54.031902.134907.

Fontana, A., M. Held, C.A. Fantaye, T.C. Turlings, J. Degenhardt, et al. 2011. Attractiveness of Constitutive and Herbivore-Induced Sesquiterpene Blends of Maize to the Parasitic Wasp Cotesia marginiventris (Cresson). J Chem Ecol 37(6): 582–591. doi: 10.1007/s10886-011-9967-7.

Forster, B.P., R.P. Ellis, J. Moir, V. Talamè, M.C. Sanguineti, et al. 2004. Genotype and phenotype associations with drought tolerance in barley tested in North Africa. Annals of Applied Biology 144(2): 157–168. doi: 10.1111/j.1744-7348.2004.tb00329.x.

Fox, P.N., J. Crossa, and I. Romagosa. 1997. Multi-environment testing and genotype × environment interaction. In: Kempton, R.A., Fox, P.N., and Cerezo, M., editors, Statistical Methods for Plant Variety Evaluation. Springer Netherlands, Dordrecht. p. 117–138

Gage, J.L. 2018. tassel_selection. doi: 10.7946/P2K04H.

Galiano-Carneiro, A.L., B. Kessel, T. Presterl, D.S. Gaikpa, M.B. Kistner, et al. 2020. Multi-parent QTL mapping reveals stable QTL conferring resistance to Gibberella ear rot in maize. Euphytica 217(1): 2. doi: 10.1007/s10681-020-02748-x.

Gallusci, P., Z. Dai, M. Génard, A. Gauffretau, N. Leblanc-Fournier, et al. 2017. Epigenetics for Plant Improvement: Current Knowledge and Modeling Avenues. Trends in Plant Science 22(7): 610–623. doi: 10.1016/j.tplants.2017.04.009.

Gamble, E.E. 1962. Gene effects in corn (zea mays l.): i. separation and relative importance of gene effects for yield. Can. J. Plant Sci. 42(2): 339–348. doi: 10.4141/cjps62-048.

García-Lara, S., and S.O. Serna-Saldivar. 2019. Chapter 1 - Corn History and Culture. In: Serna-Saldivar, S.O., editor, Corn (Third Edition). AACC International Press, Oxford. p. 1–18

Gluckman, P., A. Beedle, T. Buklijas, F. Low, and M. Hanson. 2016. Principles of Evolutionary Medicine. Oxford University Press.

Gomez, F., J. Hirbo, and S.A. Tishkoff. 2014. Genetic Variation and Adaptation in Africa: Implications for Human Evolution and Disease. Cold Spring Harb Perspect Biol 6(7): a008524. doi: 10.1101/cshperspect.a008524.

Grossniklaus, U., W.G. Kelly, A.C. Ferguson-Smith, M. Pembrey, and S. Lindquist. 2013. Transgenerational epigenetic inheritance: how important is it? Nat Rev Genet 14(3): 228–235. doi: 10.1038/nrg3435.

Grote, U., A. Fasse, T.T. Nguyen, and O. Erenstein. 2021. Food Security and the Dynamics of Wheat and Maize Value Chains in Africa and Asia. Frontiers in Sustainable Food Systems 4. https://www.frontiersin.org/articles/10.3389/fsufs.2020.617009 (accessed 8 February 2023).

Guo, J., Z. Chen, Z. Liu, B. Wang, W. Song, et al. 2011. Identification of genetic factors affecting plant density response through QTL mapping of yield component traits in maize (Zea mays L.). Euphytica 182(3): 409–422. doi: 10.1007/s10681-011-0517-8.

Guo, J., G. Su, J. Zhang, and G. Wang. 2008. Genetic analysis and QTL mapping of maize yield and associate agronomic traits under semi-arid land condition. African Journal of Biotechnology 7(12). doi: 10.4314/ajb.v7i12.58813.

Haley, C.S., and S.A. Knott. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69(4): 315–324. doi: 10.1038/hdy.1992.131.

Hallauer, A.R., M.J. Carena, and J.B.M. Filho. 2010. Quantitative Genetics in Maize Breeding. Springer Science & Business Media.

Hansen, J., R. Ruedy, M. Sato, and K. Lo. 2010. Global Surface Temperature Change. Reviews of Geophysics 48(4). doi: 10.1029/2010RG000345.

Hedden, P. 2003. The genes of the Green Revolution. Trends in Genetics 19(1): 5–9. doi: 10.1016/S0168-9525(02)00009-4.

Heffner, E.L., M.E. Sorrells, and J.-L. Jannink. 2009. Genomic Selection for Crop Improvement. Crop Science 49(1): 1–12. doi: 10.2135/cropsci2008.08.0512.

Herman, J., and S. Sultan. 2011. Adaptive Transgenerational Plasticity in Plants: Case Studies, Mechanisms, and Implications for Natural Populations. Frontiers in Plant Science 2. https://www.frontiersin.org/articles/10.3389/fpls.2011.00102 (accessed 2 December 2023).

Hill, J. 1975. Genotype-environment interaction – a challenge for plant breeding. The Journal of Agricultural Science 85(3): 477–493. doi: 10.1017/S0021859600062365.

Hill, W.G. 2010. Understanding and using quantitative genetic variation. Philosophical Transactions of the Royal Society B: Biological Sciences 365(1537): 73–85. doi: 10.1098/rstb.2009.0203.

Hill, W.G., and H.A. Mulder. 2010. Genetic analysis of environmental variation. Genetics Research 92(5–6): 381–395. doi: 10.1017/S0016672310000546.

Hittalmani, S., N. Huang, B. Courtois, R. Venuprasad, H.E. Shashidhar, et al. 2003. Identification of QTL for growth- and grain yield-related traits in rice across nine locations of Asia. Theor Appl Genet 107(4): 679–690. doi: 10.1007/s00122-003-1269-1.

Hobbs, P.R. 2007. PAPER PRESENTED AT INTERNATIONAL WORKSHOP ON INCREASING WHEAT YIELD POTENTIAL, CIMMYT, OBREGON, MEXICO, 20–24 MARCH 2006 Conservation agriculture: what is it and why is it important for future sustainable food production? J. Agric. Sci. 145(02): 127. doi: 10.1017/S0021859607006892.

Hohls, T. 1996. Setting confidence limits to genetic parameters estimated by restricted maximum likelihood analysis of North Carolina design II experiments. Heredity 77(5): 476–487. doi: 10.1038/hdy.1996.174.

Holland, J.B., W.E. Nyquist, C.T. Cervantes-Martínez, and J. Janick. 2003. Estimating and interpreting heritability for plant breeding: an update. Plant breeding reviews 22.

https://www.genomes2fields.org. https://www.genomes2fields.org/ (accessed 3 January 2023).

Hu, H., W. Liu, Z. Fu, L. Homann, F. Technow, et al. 2013. QTL mapping of stalk bending strength in a recombinant inbred line maize population. Theor Appl Genet 126(9): 2257–2266. doi: 10.1007/s00122-013-2132-7.

Jablonka, E., and G. Raz. 2009. Transgenerational Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution. The Quarterly Review of Biology 84(2): 131–176. doi: 10.1086/598822.

Janick, J. 2010. Plant Breeding Reviews, Volume 24, Part 2: Long-term Selection: Crops, Animals, and Bacteria. John Wiley & Sons.

Jena, K.K., and D.J. Mackill. 2008. Molecular Markers and Their Use in Marker-Assisted Selection in Rice. Crop Science 48(4): 1266–1276. doi: 10.2135/cropsci2008.02.0082.

Jia, Y., X. Sun, J. Sun, Z. Pan, X. Wang, et al. 2014. Association Mapping for Epistasis and Environmental Interaction of Yield Traits in 323 Cotton Cultivars under 9 Different Environments (B. Zhang, editor). PLoS ONE 9(5): e95882. doi: 10.1371/journal.pone.0095882.

Kang, M.S. 1997. Using Genotype-by-Environment Interaction for Crop Cultivar Development. In: Sparks, D.L., editor, Advances in Agronomy. Academic Press. p. 199–252

Karen Sabadin, P., C. Lopes de Souza Júnior, A. Pereira de Souza, and A. Augusto Franco Garcia. 2008. QTL mapping for yield components in a tropical maize population using microsatellite markers. Hereditas 145(4): 194–203. doi: 10.1111/j.0018-0661.2008.02065.x.

Kc, K.B., G.M. Dias, A. Veeramani, C.J. Swanton, D. Fraser, et al. 2018. When too much isn't enough: Does current food production meet global nutritional needs? PLOS ONE 13(10): e0205683. doi: 10.1371/journal.pone.0205683.

Ketata, H., E.L. Smith, L.H. Edwards, and R.W. McNew. 1976. Detection of Epistatic, Additive, and Dominance Variation in Winter Wheat (Triticum aestivum L. em Thell.)1. Crop Science 16(1): cropsci1976.0011183X001600010001x. doi: 10.2135/cropsci1976.0011183X001600010001x.

Kick, D.R., J.G. Wallace, J.C. Schnable, J.M. Kolkman, B. Alaca, et al. 2022. Yield prediction through integration of genetic, environment, and management data through deep learning. bioRxiv: 2022.07. 29.502051.

Kim, C., H. Guo, W. Kong, R. Chandnani, L.-S. Shuang, et al. 2016. Application of genotyping by sequencing technology to a variety of crop breeding programs. Plant Science 242: 14–22. doi: 10.1016/j.plantsci.2015.04.016.

Kim, J.-M., T. Sasaki, M. Ueda, K. Sako, and M. Seki. 2015. Chromatin changes in response to drought, salinity, heat, and cold stresses in plants. Frontiers in Plant Science 6. https://www.frontiersin.org/articles/10.3389/fpls.2015.00114 (accessed 21 August 2023).

Koch, K.E. 1996. Carbohydrate-Modulated Gene Expression in Plants. Annual Review of Plant Physiology and Plant Molecular Biology 47(1): 509–540. doi: 10.1146/annurev.arplant.47.1.509.

Koester, R.P., P.H. Sisco, and C.W. Stuber. 1993. Identification of Quantitative Trait Loci Controlling Days to Flowering and Plant Height in Two Near Isogenic Lines of Maize. Crop Science 33(6): cropsci1993.0011183X003300060020x. doi: 10.2135/cropsci1993.0011183X003300060020x.

Koonin, E.V. 2012. Does the central dogma still stand? Biol Direct 7(1): 27. doi: 10.1186/1745-6150-7-27.

Koornneef, M., C. Alonso-Blanco, and D. Vreugdenhil. 2004. Naturally Occurring Genetic Variation in Arabidopsis Thaliana. Annual Review of Plant Biology 55(1): 141–172. doi: 10.1146/annurev.arplant.55.031903.141605.

Kou, H.P., Y. Li, X.X. Song, X.F. Ou, S.C. Xing, et al. 2011. Heritable alteration in DNA methylation induced by nitrogen-deficiency stress accompanies enhanced tolerance by progenies to the stress in rice (Oryza sativa L.). Journal of Plant Physiology 168(14): 1685–1693. doi: 10.1016/j.jplph.2011.03.017.

Kravchenko, A.N., and D.G. Bullock. 2000. Correlation of Corn and Soybean Grain Yield with Topography and Soil Properties. Agronomy Journal 92(1): 75–83. doi: 10.2134/agronj2000.92175x.

Kruglyak, L. 1997. The use of a genetic map of biallelic markers in linkage studies. Nat Genet 17(1): 21–24. doi: 10.1038/ng0997-21.

Kumar, P., V.K. Gupta, A.K. Misra, D.R. Modi, and B.K. Pandey. 2009. Potential of molecular markers in plant biotechnology. Plant omics 2(4): 141–162.

Kvitek, D.J., and G. Sherlock. 2011. Reciprocal Sign Epistasis between Frequently Experimentally Evolved Adaptive Mutations Causes a Rugged Fitness Landscape. PLOS Genetics 7(4): e1002056. doi: 10.1371/journal.pgen.1002056.

Lacal, I., and R. Ventura. 2018. Epigenetic Inheritance: Concepts, Mechanisms and Perspectives. Frontiers in Molecular Neuroscience 11. https://www.frontiersin.org/articles/10.3389/fnmol.2018.00292 (accessed 1 December 2023).

Laidig, F., H.-P. Piepho, T. Drobek, and U. Meyer. 2014. Genetic and non-genetic long-term trends of 12 different crops in German official variety performance trials and on-farm yield trends. Theor Appl Genet 127(12): 2599–2617. doi: 10.1007/s00122-014-2402-z.

Laitinen, R.A.E., and Z. Nikoloski. 2019. Genetic basis of plasticity in plants. Journal of Experimental Botany 70(3): 739–745. doi: 10.1093/jxb/ery404.

Leary, R.F., F.W. Allendorf, and K.L. Knudsen. 1985. Inheritance of Meristic Variation and the Evolution of Developmental Stability in Rainbow Trout. Evolution 39(2): 308–314. doi: 10.1111/j.1558-5646.1985.tb05668.x.

Leng, P., S.U. Khan, D. Zhang, G. Zhou, X. Zhang, et al. 2022. Linkage Mapping Reveals QTL for Flowering Time-Related Traits under Multiple Abiotic Stress Conditions in Maize. International Journal of Molecular Sciences 23(15): 8410. doi: 10.3390/ijms23158410.

Li, Z., S.R.M. Pinson, W.D. Park, A.H. Paterson, and J.W. Stansel. 1997. Epistasis for Three Grain Yield Components in Rice (*Oryxa sativa L.*). Genetics 145(2): 453–465. doi: 10.1093/genetics/145.2.453.

Li, H., Q. Yang, N. Fan, M. Zhang, H. Zhai, et al. 2017. Quantitative trait locus analysis of heterosis for plant height and ear height in an elite maize hybrid zhengdan 958 by design III. BMC Genet 18(1): 36. doi: 10.1186/s12863-017-0503-9.

Li, X., Z. Zhou, J. Ding, Y. Wu, B. Zhou, et al. 2016. Combined Linkage and Association Mapping Reveals QTL and Candidate Genes for Plant and Ear Height in Maize. Frontiers in Plant Science 7. https://www.frontiersin.org/articles/10.3389/fpls.2016.00833 (accessed 1 March 2023).

Liaghat, S., and S. Balasundram. 2010. A Review: The Role of Remote Sensing in Precision Agriculture. American Journal of Agricultural and Biological Science 5. doi: 10.3844/ajabssp.2010.50.55.

Lira-Medeiros, C.F., C. Parisod, R.A. Fernandes, C.S. Mata, M.A. Cardoso, et al. 2010. Epigenetic Variation in Mangrove Plants Occurring in Contrasting Natural Environment (P. Meyer, editor). PLoS ONE 5(4): e10326. doi: 10.1371/journal.pone.0010326.

Liu, Z.J., and J.F. Cordes. 2004. DNA marker technologies and their applications in aquaculture genetics. Aquaculture 238(1): 1–37. doi: 10.1016/j.aquaculture.2004.05.027.

López-Maury, L., S. Marguerat, and J. Bähler. 2008. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. Nat Rev Genet 9(8): 583–593. doi: 10.1038/nrg2398.

Love, M.I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12): 550. doi: 10.1186/s13059-014-0550-8.

Macauley, H., and T. 2015 Ramadjita. 2015. Cereal crops: rice, maize, millet, sorghum, wheat. http://ir.ucc.edu.gh/jspui/handle/123456789/4510 (accessed 8 February 2023).

Mackay, T.F.C. 2014. Epistasis and quantitative traits: using model organisms to study gene–gene interactions. Nat Rev Genet 15(1): 22–33. doi: 10.1038/nrg3627.

Majeran, W., G. Friso, Y. Asakura, X. Qu, M. Huang, et al. 2012. Nucleoid-Enriched Proteomes in Developing Plastids and Chloroplasts from Maize Leaves: A New Conceptual Framework for Nucleoid Functions. Plant Physiology 158(1): 156–189. doi: 10.1104/pp.111.188474.

Malik, H.N., S.I. Malik, M. Hussain, S.U.R. Chughtai, and I. Javed. 2005. Genetic Correlation among Various Quantitative Characters in Maize (Zea mays L.) Hybrids. 1(3).

Malosetti, M., J.-M. Ribaut, and F.A. van Eeuwijk. 2013. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. Frontiers in Physiology 4. https://www.frontiersin.org/articles/10.3389/fphys.2013.00044 (accessed 9 February 2023).

Mammadov, J., R. Aggarwal, R. Buyyarapu, and S. Kumpatla. 2012. SNP Markers and Their Impact on Plant Breeding. International Journal of Plant Genomics 2012: 1–11. doi: 10.1155/2012/728398.

Mardis, E.R. 2017. DNA sequencing technologies: 2006–2016. Nat Protoc 12(2): 213–218. doi: 10.1038/nprot.2016.182.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal; Vol 17, No 1: Next Generation Sequencing Data AnalysisDO -

10.14806/ej.17.1.200. http://journal.embnet.org/index.php/embnetjournal/article/view/200/458.

Matzke, M.A., and R.A. Mosher. 2014. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet 15(6): 394–408. doi: 10.1038/nrg3683.

McCauley, D.E. 1995. The use of chloroplast DNA polymorphism in studies of gene flow in plants. Trends in Ecology & Evolution 10(5): 198–202. doi: 10.1016/S0169-5347(00)89052-7.

McCue, A.D., and R.K. Slotkin. 2012. Transposable element small RNAs as regulators of gene expression. Trends in Genetics 28(12): 616–623. doi: 10.1016/j.tig.2012.09.001.

McKay, J.K., C.E. Christian, S. Harrison, and K.J. Rice. 2005. "How Local Is Local?"—A Review of Practical and Conceptual Issues in the Genetics of Restoration. Restoration Ecology 13(3): 432–440. doi: 10.1111/j.1526-100X.2005.00058.x.

Meredith Jr., W.R. 1984. Quantitative Genetics. Cotton. John Wiley & Sons, Ltd. p. 131–150

Messmer, R., Y. Fracheboud, M. Bänziger, M. Vargas, P. Stamp, et al. 2009. Drought stress and tropical maize: QTL-by-environment interactions and stability of QTLs across environments for yield components and secondary traits. Theor Appl Genet 119(5): 913–930. doi: 10.1007/s00122-009-1099-x.

Millstein, J., D.V. Conti, F.D. Gilliland, and W.J. Gauderman. 2006. A Testing Framework for Identifying Susceptibility Genes in the Presence of Epistasis. The American Journal of Human Genetics 78(1): 15–27. doi: 10.1086/498850.

Miryeganeh, M., and H. Saze. 2020. Epigenetic inheritance and plant evolution. Population Ecology 62(1): 17–27. doi: 10.1002/1438-390X.12018.

Moffitt, T.E., A. Caspi, and M. Rutter. 2005. Strategy for Investigating Interactions Between Measured Genes and Measured Environments. Archives of General Psychiatry 62(5): 473–481. doi: 10.1001/archpsyc.62.5.473.

Molinier, J., G. Ries, C. Zipfel, and B. Hohn. 2006. Transgeneration memory of stress in plants. Nature 442(7106): 1046–1049. doi: 10.1038/nature05022.

Moll, R.H., and C.W. Stuber. 1974. Quantitative Genetics—Empirical Results Relevant to Plant Breeding11Paper No. 4236 of the Journal Series of the North Carolina State University Agricultural Experiment Station, Raleigh, North Carolina. In: Brady, N.C., editor, Advances in Agronomy. Academic Press. p. 277–313

Moore, J.H. 2005. A global view of epistasis. Nat Genet 37(1): 13–14. doi: 10.1038/ng0105-13.

Moore, J.H., and S.M. Williams. 2009. Epistasis and Its Implications for Personal Genetics. The American Journal of Human Genetics 85(3): 309–320. doi: 10.1016/j.ajhg.2009.08.006.

Muir, P., S. Li, S. Lou, D. Wang, D.J. Spakowicz, et al. 2016. The real cost of sequencing: scaling computation to keep pace with data generation. Genome Biology 17(1): 53. doi: 10.1186/s13059-016-0917-0.

Nadeem, M.A., M.A. Nawaz, M.Q. Shahid, Y. Doğan, G. Comertpay, et al. 2018. DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. Biotechnology & Biotechnological Equipment 32(2): 261–285. doi: 10.1080/13102818.2017.1400401.

Nakagawa, S., and R.P. Freckleton. 2008. Missing inaction: the dangers of ignoring missing data. Trends in Ecology & Evolution 23(11): 592–596. doi: 10.1016/j.tree.2008.06.014.

Nasu, S. 2002. Search for and Analysis of Single Nucleotide Polymorphisms (SNPs) in Rice (Oryza sativa, Oryza rufipogon) and Establishment of SNP Markers. DNA Research 9(5): 163–171. doi: 10.1093/dnares/9.5.163.

Olson, J.M., J.S. Witte, and R.C. Elston. 1999. Genetic mapping of complex traits. Statistics in Medicine 18(21): 2961–2981. doi: 10.1002/(SICI)1097-0258(19991115)18:21<2961::AID-SIM206>3.0.CO;2-U.

Orr, H.A. 2005. The genetic theory of adaptation: a brief history. Nat Rev Genet 6(2): 119–127. doi: 10.1038/nrg1523.

Ortiz-Monasterio, J.I., N. Palacios-Rojas, E. Meng, K. Pixley, R. Trethowan, et al. 2007. Enhancing the mineral and vitamin content of wheat and maize through plant breeding. Journal of Cereal Science 46(3): 293–307. doi: 10.1016/j.jcs.2007.06.005.

Paddock, W.C. 1970. How Green Is the Green Revolution? BioScience 20(16): 897–902. doi: 10.2307/1295581.

Palacios-Rojas, N., L. McCulley, M. Kaeppler, T.J. Titcomb, N.S. Gunaratna, et al. 2020. Mining maize diversity and improving its nutritional aspects within agro-food systems. Comprehensive Reviews in Food Science and Food Safety 19(4): 1809–1834. doi: 10.1111/1541-4337.12552.

Pandey, M.K., H.D. Upadhyaya, A. Rathore, V. Vadez, M.S. Sheshshayee, et al. 2014. Genomewide Association Studies for 50 Agronomic Traits in Peanut Using the 'Reference Set' Comprising 300 Genotypes from 48 Countries of the Semi-Arid Tropics of the World. PLOS ONE 9(8): e105228. doi: 10.1371/journal.pone.0105228.

Patel, R. 2013. The Long Green Revolution. The Journal of Peasant Studies 40(1): 1–63. doi: 10.1080/03066150.2012.719224.

Penfield, S., and D.R. MacGregor. 2017. Effects of environmental variation during seed production on seed dormancy and germination. Journal of Experimental Botany 68(4): 819–825. doi: 10.1093/jxb/erw436.

Phillips, P.C. 2008. Epistasis — the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9(11): 855–867. doi: 10.1038/nrg2452.

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

Pilet, M. l., G. Duplan, M. Archipiano, P. Barret, C. Baron, et al. 2001. Stability of QTL for Field Resistance to Blackleg across Two Genetic Backgrounds in Oilseed Rape. Crop Science 41(1): 197–205. doi: 10.2135/cropsci2001.411197x.

Poelwijk, F.J., S. Tănase-Nicola, D.J. Kiviet, and S.J. Tans. 2011. Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. Journal of Theoretical Biology 272(1): 141–144. doi: 10.1016/j.jtbi.2010.12.015.

Poland, J.A., and T.W. Rife. 2012. Genotyping-by-Sequencing for Plant Breeding and Genetics. The Plant Genome 5(3). doi: 10.3835/plantgenome2012.05.0005.

Portwood, J.L. 2nd, M.R. Woodhouse, E.K. Cannon, J.M. Gardiner, L.C. Harper, et al. 2019. MaizeGDB 2018: the maize multi-genome genetics and genomics database. Nucleic Acids Res 47(D1): D1146–D1154. doi: 10.1093/nar/gky1046.

Pounds, N.J.G., Pounds, and rman J.G. Pounds. 1979. An Historical Geography of Europe, 1500-1840. Cambridge University Press.

Rafalski, J.A. 2002. Novel genetic mapping tools in plants: SNPs and LD-based approaches. Plant Science 162(3): 329–333. doi: 10.1016/S0168-9452(01)00587-8.

Randhawa, H.S., M. Asif, C. Pozniak, J.M. Clarke, R.J. Graf, et al. 2013. Application of molecular markers to wheat breeding in Canada. Plant Breeding 132(5): 458–471. doi: 10.1111/pbr.12057.

Ranum, P., J.P. Peña-Rosas, and M.N. Garcia-Casal. 2014. Global maize production, utilization, and consumption. Annals of the New York Academy of Sciences 1312(1): 105–112. doi: 10.1111/nyas.12396.

Rastas, P. 2017. Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data (B. Berger, editor). Bioinformatics 33(23): 3726–3732. doi: 10.1093/bioinformatics/btx494.

Rebai, A., B. Goffinet, and B. Mangin. 1995. Comparing Power of Different Methods for QTL Detection. Biometrics 51(1): 87. doi: 10.2307/2533317.

Ribaut, J.-M., D.A. Hoisington, J.A. Deutsch, C. Jiang, and D. Gonzalez-de-Leon. 1996. Identification of quantitative trait loci under drought conditions in tropical maize. 1. Flowering parameters and the anthesis-silking interval. Theoret. Appl. Genetics 92(7): 905–914. doi: 10.1007/BF00221905.

Robert, P.C. 2002. Precision agriculture: a challenge for crop nutrition management. In: Horst, W.J., Bürkert, A., Claassen, N., Flessa, H., Frommer, W.B., et al., editors, Progress in Plant Nutrition: Plenary Lectures of the XIV International Plant Nutrition Colloquium: Food security and sustainability of agro-ecosystems through basic and applied research. Springer Netherlands, Dordrecht. p. 143–149

Rogers, A.R., and J.B. Holland. 2022. Environment-specific genomic prediction ability in maize using environmental covariates depends on environmental similarity to training data. G3 Genes|Genomes|Genetics 12(2): jkab440. doi: 10.1093/g3journal/jkab440.

Romero Navarro, J.A., M. Willcox, J. Burgueño, C. Romay, K. Swarts, et al. 2017. A study of allelic diversity underlying flowering-time adaptation in maize landraces. Nat Genet 49(3): 476–480. doi: 10.1038/ng.3784.

Ronald, P. 2011. Plant Genetics, Sustainable Agriculture and Global Food Security. Genetics 188(1): 11–20. doi: 10.1534/genetics.111.128553.

Rosset, P., J. Collins, and F.M. Lappé. 2000. Lessons from the green revolution. Third World Resurgence: 11–14.

Salvi, S., G. Sponza, M. Morgante, D. Tomes, X. Niu, et al. 2007. Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. Proceedings of the National Academy of Sciences 104(27): 11376–11381. doi: 10.1073/pnas.0704145104.

Sankaran, S., C. Zhang, J.P. Hurst, A. Marzougui, A.N. Veeranampalayam-Sivakumar, et al. 2020. Investigating the potential of satellite imagery for high-throughput field phenotyping applications. Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping V. SPIE. p. 1141402

Semagn, K., Å. Bjørnstad, and M.N. Ndjiondjop. 2006. Principles, requirements and prospects of genetic mapping in plants.

Serageldin, I. 2001. Assuring Water for Food: The Challenge of the Coming Generation. International Journal of Water Resources Development 17(4): 521–525. doi: 10.1080/07900620120094154.

Shafee, T. 2014. Evolvability of a viral protease: experimental evolution of catalysis, robustness and specificity. https://www.repository.cam.ac.uk/handle/1810/245207 (accessed 17 July 2023).

Shiferaw, B., B.M. Prasanna, J. Hellin, and M. Bänziger. 2011. Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. Food Sec. 3(3): 307–327. doi: 10.1007/s12571-011-0140-5.

Silva, T.N., G.V. Moro, F.V. Moro, D.M.M. dos Santos, and R. Buzinaro. 2016. Correlation and path analysis of agronomic and morphological traits in maize. REVISTA CIÊNCIA AGRONÔMICA 47(2). doi: 10.5935/1806-6690.20160041.

Smulders, M.J.M., and G.J. de Klerk. 2011. Epigenetics in plant tissue culture. Plant Growth Regul 63(2): 137–146. doi: 10.1007/s10725-010-9531-4.

Sparks, A. 2018. nasapower: A NASA POWER Global Meteorology, Surface Solar Energy and Climatology Data Client for R. JOSS 3(30): 1035. doi: 10.21105/joss.01035.

Steinberg, B., and M. Ostermeier. 2016. Environmental changes bridge evolutionary valleys. Sci Adv 2(1): e1500921. doi: 10.1126/sciadv.1500921.

Stirling, D.G., D. Réale, and D.A. Roff. 2002. Selection, structure and the heritability of behaviour. Journal of Evolutionary Biology 15(2): 277–289. doi: 10.1046/j.1420-9101.2002.00389.x.

Stretch, C., S. Khan, N. Asgarian, R. Eisner, S. Vaisipour, et al. 2013. Effects of Sample Size on Differential Gene Expression, Rank Order and Prediction Accuracy of a Gene Signature. PLOS ONE 8(6): e65380. doi: 10.1371/journal.pone.0065380.

Syvänen, A.-C. 2001. Accessing genetic variation: genotyping single nucleotide polymorphisms. Nat Rev Genet 2(12): 930–942. doi: 10.1038/35103535.

Tanksley, S.D., N.D. Young, A.H. Paterson, and M.W. Bonierbale. 1989. RFLP Mapping in Plant Breeding: New Tools for an Old Science. Nat Biotechnol 7(3): 257–264. doi: 10.1038/nbt0389-257.

Tenaillon, M.I., and A. Charcosset. 2011. A European perspective on maize history. Comptes Rendus Biologies 334(3): 221–228. doi: 10.1016/j.crvi.2010.12.015.

Teng, F., L. Zhai, R. Liu, W. Bai, L. Wang, et al. 2013. ZmGA3ox2, a candidate gene for a major QTL, qPH3.1, for plant height in maize. The Plant Journal 73(3): 405–416. doi: 10.1111/tpj.12038.

Terwilliger, J.D., and H.H.H. Göring. 2009. Gene Mapping in the 20th and 21st Centuries: Statistical Methods, Data Analysis, and Experimental Design. hbio 81(5/6): 663–728. doi: 10.3378/027.081.0615.

Thornton, P.K., P.G. Jones, G. Alagarswamy, and J. Andresen. 2009. Spatial variation of crop yield response to climate change in East Africa. Global Environmental Change 19(1): 54–65. doi: 10.1016/j.gloenvcha.2008.08.005.

Thornton, P.K., P.G. Jones, G. Alagarswamy, J. Andresen, and M. Herrero. 2010. Adapting to climate change: Agricultural system and household impacts in East Africa. Agricultural Systems 103(2): 73–82. doi: 10.1016/j.agsy.2009.09.003.

Toschi, A. 1960. Fisher, R. A. - the Genetical Theory of Natural Selection. Scientia 54(95): 202.

Townsley, B.T., M.F. Covington, Y. Ichihashi, K. Zumstein, and N.R. Sinha. 2015. BrAD-seq: Breath Adapter Directional sequencing: a streamlined, ultra-simple and fast library preparation protocol for strand specific mRNA library construction. Front. Plant Sci. 6. doi: 10.3389/fpls.2015.00366.

Trerotola, M., V. Relli, P. Simeone, and S. Alberti. 2015. Epigenetic inheritance and the missing heritability. Hum Genomics 9(1): 17. doi: 10.1186/s40246-015-0041-3.

Verhoeven, K.J., K.L. Simonsen, and L.M. McIntyre. 2005. Implementing false discovery rate control: increasing your power. Oikos 108(3): 643–647.

Via, S., and R. Lande. 1985. GENOTYPE-ENVIRONMENT INTERACTION AND THE EVOLUTION OF PHENOTYPIC PLASTICITY. Evolution 39(3): 505–522. doi: 10.1111/j.1558-5646.1985.tb00391.x.

Viana, J.M.S., F.F. e Silva, G.B. Mundim, C.F. Azevedo, and H.U. Jan. 2017. Efficiency of low heritability QTL mapping under high SNP density. Euphytica 213(1): 13. doi: 10.1007/s10681-016-1800-5.

de Visser, J.A.G.M., T.F. Cooper, and S.F. Elena. 2011. The causes of epistasis. Proceedings of the Royal Society B: Biological Sciences 278(1725): 3617–3624. doi: 10.1098/rspb.2011.1537.

Wang, X., Y. Pang, J. Zhang, Q. Zhang, Y. Tao, et al. 2014. Genetic background effects on QTL and QTL×environment interaction for yield and its component traits as revealed by reciprocal introgression lines in rice. The Crop Journal 2(6): 345–357. doi: 10.1016/j.cj.2014.06.004.

Weebadde, C.K., D. Wang, C.E. Finn, K.S. Lewers, J.J. Luby, et al. 2008. Using a linkage mapping approach to identify QTL for day-neutrality in the octoploid strawberry. Plant Breeding 127(1): 94–101. doi: 10.1111/j.1439-0523.2007.01430.x.

Westhues, C.C., H. Simianer, and T.M. Beissinger. 2021. learnMET: an R package to apply machine learning methods for genomic prediction using multi-environment trial data. Genetics.

Wheeler, B.S. 2013. Small RNAs, big impact: small RNA pathways in transposon control and their effect on the host stress response. Chromosome Res 21(6): 587–600. doi: 10.1007/s10577-013-9394-4.

Whitlock, M.C., P.C. Phillips, F.B.-G. Moore, and S.J. Tonsor. 1995. Multiple Fitness Peaks and Epistasis. Annual Review of Ecology and Systematics 26(1): 601–629. doi: 10.1146/annurev.es.26.110195.003125.

Williams-Carrier, R., R. Zoschke, S. Belcher, J. Pfalz, and A. Barkan. 2014. A Major Role for the Plastid-Encoded RNA Polymerase Complex in the Expression of Plastid Transfer RNAs. Plant Physiology 164(1): 239–248. doi: 10.1104/pp.113.228726.

Willis, J.H., and H.A. Orr. 1993. Increased Heritable Variation Following Population Bottlenecks: The Role of Dominance. Evolution 47(3): 949–957. doi: 10.2307/2410199.

Wilson, W.C., M. Slingerland, F.P. Baijukya, H. van Zanten, S. Oosting, et al. 2021. Integrating the soybean-maize-chicken value chains to attain nutritious diets in Tanzania. Food Sec. 13(6): 1595–1612. doi: 10.1007/s12571-021-01213-4.

Wittke-Thompson, J.K., A. Pluzhnikov, and N.J. Cox. 2005. Rational Inferences about Departures from Hardy-Weinberg Equilibrium. The American Journal of Human Genetics 76(6): 967–986. doi: 10.1086/430507.

Wright, S. 1984. Evolution and the Genetics of Populations, Volume 1: Genetic and Biometric Foundations. University of Chicago Press.

Yan, J. 2003. Dynamic analysis of QTL for plant height at different developmental stages in maize (Zea mays L.). Chinese Sci Bull 48(23): 2601. doi: 10.1360/03wc0044.

Yang, X., Y. Guo, J. Yan, J. Zhang, T. Song, et al. 2010. Major and minor QTL and epistasis contribute to fatty acid compositions and oil concentration in high-oil maize. Theor Appl Genet 120(3): 665–678. doi: 10.1007/s00122-009-1184-1.

Yu, A., G. Lepère, F. Jay, J. Wang, L. Bapaume, et al. 2013. Dynamics and biological relevance of DNA demethylation in Arabidopsis antibacterial defense. Proceedings of the National Academy of Sciences 110(6): 2389–2394. doi: 10.1073/pnas.1211757110.

Yu, S.B., J.X. Li, C.G. Xu, Y.F. Tan, X.H. Li, et al. 2002. Identification of quantitative trait loci and epistatic interactions for plant height and heading date in rice. Theor Appl Genet 104(4): 619–625. doi: 10.1007/s00122-001-0772-5.

Zamore, P.D., and B. Haley. 2005. Ribo-gnome: The Big World of Small RNAs. Science 309(5740): 1519–1524. doi: 10.1126/science.1111444.

Zhang, J., J. Chen, C. Chu, W. Zhao, J. Wheeler, et al. 2014. Genetic Dissection of QTL Associated with Grain Yield in Diverse Environments. Agronomy 4(4): 556–578. doi: 10.3390/agronomy4040556.

Zhang, Z.M., M.J. Zhao, H.P. Ding, T.Z. Rong, and G.T. Pan. 2006. QTL mapping analysis of plant height and ear height of maize (Zea mays L.). Genetika 42(3): 391–396.

Zhang, H., and J.-K. Zhu. 2011. RNA-directed DNA methylation. Current Opinion in Plant Biology 14(2): 142–147. doi: 10.1016/j.pbi.2011.02.003.

Zhao, J., H.C. Becker, D. Zhang, Y. Zhang, and W. Ecke. 2005. Oil Content in a European x Chinese Rapeseed Population: QTL with Additive and Epistatic Effects and Their Genotype–Environment Interactions. CROP SCIENCE 45: 9.

Zhao, B., J. Zhang, C. Yang, G. Zhou, Y. Ding, et al. 2018. Rapeseed Seedling Stand Counting and Seeding Performance Evaluation at Two Early Growth Stages Based on Unmanned Aerial Vehicle Imagery. Front. Plant Sci. 9: 1362. doi: 10.3389/fpls.2018.01362.

Zondervan, K.T., and L.R. Cardon. 2004. The complex interplay among factors that influence allelic association. Nat Rev Genet 5(2): 89–100. doi: 10.1038/nrg1270.

Appendix

Acknowledgments

I would like to thank

Prof. Dr. Timothy M Beissinger, I thank him for supporting me, investing his time, and being patient during my education. It was a fantastic experience to discuss articles together during the Journal Club or our weekly discussion. I express my deepest appreciation for your commitment to my personal growth.

Prof. Dr. Stefan Scholten, I thank him for his support, patience, and feedback during my Ph.D. progress. His feedback improved my perspective and understanding of the research field.

Prof. Dr. Reimund P. Rötter, I thank him for agreeing to be my third thesis committee member and for his valuable comments on my project.

apl. Prof. Dr. Wolfgang Link, I would like to thank him for his constant support and discussions of terminology as well as his contribution to my thesis. I would especially like to thank him for his support towards the last year of my Ph.D.

Everyone in the Division of Plant Breeding Methodology and Crop Plant Genetics. Especially **Dr. Antje Schierholt** for her constant support and for asking me about my thesis and progress; Ph.D. students **Mila Tost**, **Leke Victor Aiyesa**, **Bright Enogieru Osatohanmwen**, for sharing information and discussing the articles in the Journal club; Technical assistants **Dietrich Kaufmann** for his help in the field, **Bianka Wehr** and **Katharina Kubon** for their teaching and help in the laboratory; and my former colleagues and mentors **Dr. Stefanie Griebel**, **Dr. Medhat Mahmoud**, **Dr. Cathy Westhues**.

My family for supporting me financially and mentally throughout my lifetime and being patient during my education. Also, I would like to thank them for supporting me even though they cannot visit me for long distances.

General Directorate of Education Abroad (Yurt Dışı Eğitim Genel Müdürlüğü) Ministry of Education, Republic of Türkiye, for supporting my Ph.D. education financially.

Declaration of Academic Integrity

I confirm that I have composed the present scientific treatise (thesis) independently using no other sources and resources than those stated. I have accepted the assistance of third parties only in a scope that is scientifically justifiable and compliant with the legal statutes of the examinations. In particular, I have completed all parts of the dissertation myself; I have neither, nor will I, accept unauthorised outside assistance either free of charge or subject to a fee. Furthermore, I have not applied for an equivalent doctoral examination elsewhere and submitted the present thesis as a whole or in parts at another university. I am aware of the fact that untruthfulness with respect to the above declaration repeals the admission to complete the doctoral studies and/or subsequently entitles termination of the doctoral process or withdrawal of the title attained.

Candidate's signature: _____

(Barış Alaca)