DISSERTATION

INVESTIGATING THE ADAPTIVE GENETIC LANDSCAPE OF GLOBAL CROP SPECIES

Submitted by

Kirsten Marie Hein

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the Degree of Doctor in Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2024

Doctoral Committee:

Advisor: John McKay

Geoffrey Morris
Jeffrey Ross-Ibarra
Meagan Schipanski

ABSTRACT

INVESTIGATING THE ADAPTIVE GENETIC LANDSCAPE OF GLOBAL CROP SPECIES

Improving environmental adaptation in crops is essential for sustaining food security in the face of global climate change. Recent advances in high-throughput genomic sequencing and phenotyping technologies have enabled researchers to identify and validate the genetic factors shaping adaptation. In this dissertation, I investigated the genetic basis of environmental adaptation in global cereal crops, focusing on the staple crop, maize (*Zea mays* L.), and the orphan crop, tef (*Eragrostis tef* Zucc.). In Chapter 2, I employed a landscape genomics approach to identify the genetic and environmental drivers of adaptation in a georeferenced collection of Ethiopian tef. In Chapter 3, I utilized both forward and reverse genetic approaches to evaluate the precision of phenotype-genotype mapping across multiple phenotyping methods for quantifying root system architecture in field-excavated maize. In Chapter 4, I applied a functional genetics approach to characterize a novel gene model in maize predicted to regulate root system development and nitrogen capture under field conditions. Collectively, this work provides valuable insights into the complex relationships between phenotype, genotype, and environment, contributing to our understanding of adaptation in two distinct and vital crop systems.

ACKNOWLEDGEMENTS

First and foremost, I extend my heartfelt gratitude to my advisor, John McKay, whose unwavering support and belief in my potential have shaped my journey through and beyond this PhD. I am deeply thankful for his mentorship and the intellectually challenges he provided, which pushed me to grow beyond my own expectations.

To my collaborators at the Ethiopian Institute of Agricultural Research, the Donald Danforth Plant Science Center, and the University of Würzburg, thank you for supporting my ideas and encouraging my curiosity. Your contributions made this research possible.

I am also deeply grateful to the incredible members of the McKay lab throughout the last five years. Special thanks to Jack Mullen, Kevin Lehner, and Patrick Woods for your support, guidance, and camaraderie—I am grateful for our times at the Pickel Barrel and conference shenanigans. Also, thank you to Stephanie Goldin, Chad Soenksen, Lily Hermann, Anna Green, Jake Keenan, Dahlia Kopycienski, and Kiri Forster. Together, you made the lab feel like a family, and your support has been essential to both my research and my personal growth.

To my mentors, I owe you a great debt of gratitude. Dejene Girma, I cannot thank you enough for your unwavering guidance, kindness, and generosity throughout my PhD. Our collaboration and friendship have profoundly influenced my research values and appreciation for Ethiopian culture. I will always cherish my time at the EIAR and the support you showed by traveling to Colorado for my PhD defense. Your mentorship has been both a professional and personal gift, one I will carry forward in my career. To Chad Geater and Scott Stelpflug at Syngenta, thank you for your refreshing insights and industry perspectives, which played a huge role in shaping my career interests. Domenica Romagni, your friendship has been a grounding

presence during this journey, and I am thankful you came into my life when you did. I hope we can continue to have riveting conversations about *Buffy the Vampire Slayer*. A very special thank you to the faculty at the University of Tennessee at Chattanooga, especially Hill Craddock, Joanne Romagni, and Jose Barbosa, for igniting my passion for science and nurturing my early academic growth.

Of course, my deepest appreciation for my incredible support network, that is my friends and family. I am so fortunate to be surrounded by so many incredible individuals who continually inspire me to be great. Thank you for keeping me sane and challenging me to be adventurous. To my parents, Patti Evola, Joe and Jennifer Hein, thank you for your steadfast belief in me and for always encouraging me to reach for the stars. I am also profoundly grateful for the world's best sisters—Brooke Hein, Victoria Martin, Jenna Comella, and Emmaleigh Nix. Thank you for sharing in the good times and laughter that brings so much joy to my life.

To my remarkable husband and best friend, Tim Allen—thank you for your boundless patience, love and encouragement. Your unwavering belief in me has been a constant source of strength throughout this journey, and I am so grateful to have you by my side. You have been my rock through every high and low, and I could not have done this without you. I am truly looking forward to beginning our new chapter together. It would be remiss of me not to also mention our sweet pups, Buffy and Buster. Thank you for the comfort and endless joy you bring into my daily life. Your companionship has been a true gift over the years.

Lastly, in the words of the great Snoop Dogg: *"I want to thank me. I want to thank me for believing in me. I want to thank me for doing all this hard work…I want to thank me for never quitting…I want to thank me for just being me at all times."*

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Background

Theory and empirical study of adaptation

The study of evolution provides a unifying theoretical framework that offers critical insights into the history and future trajectories of life on Earth. The concepts of evolution hold fundamental theoretical and practical importance across the biological sciences. The five processes of evolution are: gene flow, genetic drift, mutation, recombination and selection allow a mechanistic understanding and prediction of phenotypic change, operating at different levels of biological organization (Caruso et al., 2017; Frankham, 2015; Kingsolver et al., 2001; Leinonen et al., 2008). In the biological sciences, evolutionary theory has emerged as an indispensable tool for understanding and forecasting phenotypic and functional changes in response to genetic and environmental variations over time (Reiskind et al., 2021).

Adaptation, the process by which organisms become better suited to a particular environment, has been a central focus of evolutionary research since the publication of *The Origin of Species by Means of Natural Selection* (Darwin 1809-1882, 1859). In this book, Darwin postulated that adaptation is the product of natural selection that acts on heritable variation within populations. Over the past century, numerous theoretical studies have expanded on this foundation, developing quantitative genetic theory to model the process of adaptation to environmental changes (Orr, 2005). Notably, R.A. Fisher's infinitesimal and geometric models of adaptation provided the initial theoretical basis for the genetic architecture of complex traits

(Fisher, 1919, 1930), suggesting that adaptation is likely driven by the small contributions from many independent, Mendelian loci in a continuous phenotypic space. In the latter half of the twentieth century, evolutionary biologist Motoo Kimura revisited these models, considering the size distribution of mutations and their contribution to adaptation. He concluded that adaptation involves few large effect mutations and many small effect mutations (Orr, 1999; Kimura, 1983; Orr, 1998). Sequence-based models of adaptive landscapes have refined this view, suggesting that molecular evolution through natural selection occurs in a discrete DNA sequence space, leaving a signature that differs markedly from that of neutrality (Gillespie, 1992; Maynard Smith, 1970; Wright, 1931). Historically, evolutionary theorists have built models of adaptation that describe qualitative patterns in genetic data based on individual mutations with specific effects; however, it remains uncertain whether current theory can fully capture the complexity of adaptation within a quantitative framework (Orr, 2005).

The theoretical basis for studying adaptation is supported by empirical evidence of local adaptation in plant populations. Reciprocal transplant and common garden experiments are classic approaches in ecological and evolutionary research that can test for the effects of genetics and environment on organism fitness and trait expression (Angert & Schemske, 2005; Falconer, 1990; Janzen et al., 2021; Kawecki & Ebert, 2004; Lovell et al., 2021; Price et al., 2018). Recent advances in high-throughput genomic sequencing and computational analysis have enabled researchers to extend these approaches by identifying genome-wide patterns of adaptive variation within plant populations. This is often achieved through genome scans of mapping populations to detect associations with environmental gradients or fitness-related traits, which can be validated experimentally through genomic prediction of genotype-by-environment interactions (G × E), common gardens, and mutant screens (De Villemereuil et al., 2016; Josephs

et al., 2019; Lasky et al., 2023; Rellstab et al., 2015). However, *post hoc* validation of gene function and its adaptive significance is rarely pursued.

Despite the extensive theoretical foundation on adaptation, largely built on the infinitesimal model, and experimental data on its genetic basis, bridging the gap between theory and empirical evidence remains challenging, with many questions still unanswered. This dissertation research, particularly in Chapter 2, integrates adaptive genetic theory with genome scans to test for sequence-based selection signals using traditional varieties and their local environments. In contrast, the research presented in Chapter 4 focuses on the functional validation of a putatively adaptive candidate gene.

Genotype-to-phenotype relationships: perspectives and applications in plant breeding

Quantitative genetics has provided valuable insights into plant breeding over the past century (Bernardo, 2020a). The basic model underlying quantitative variation continues to be that the observed phenotypic value (P) is the sum of the genotypic value (G), the environmental value (E), their interaction (G × E), and a residual value (*e*) that represents additional non-genetic factors not captured by the model (Lynch & Walsh, 1998). This relationship is expressed as: $P = G + E + (G \times E) + e$. As the effects of the non-genetic factors approaches zero, predictions of genotypic values from phenotypic measurements become more accurate. This model is particularly important for experimental studies of local adaptation or G × E, as the accurate characterization of the adaptive genetic effects relies on the effective control and measurement of environmental variation. Continuing the development of methods for more accurate and precise phenotyping will remain important to reduce the non-genetic factors that impact these predictions. The research presented in Chapter 3 was motivated by this need and explores

advanced phenotyping approaches to enhance the predictive power of adaptive belowground traits.

Over time, quantitative genetic tools have enhanced our understanding of trait inheritance and variation (Gardner, 1963; Matzinger, 1963), led to improved breeding methods for continuous traits (Dudley & Moll, 1969), and helped identify useful quantitative trait loci for crop improvement (Bernardo, 2020b). The classical concept of 'breeding values' in quantitative genetics is defined as the genetic contribution of an individual parent to the average phenotypic value of its offspring under random mating (Falconer, 1985; Fisher, 1918; Lynch & Walsh, 1998). However, they have limited relevance for plant breeders, as plant breeding typically violates the assumption of random mating and growers are primarily concerned with the performance of the cultivars they produce rather than their progeny (Bernardo, 2020a). Therefore, plant breeders place greater emphasis on selecting parents based on their performance in a target environment, reflecting their genotypic values, for a given set of continuous traits.

Summary of subsequent dissertation chapters

Chapter 2: Genetic diversity and environmental adaptation in Ethiopian tef

In Chapter 2, I resequenced a georeferenced collection of Ethiopian tef to test the hypothesis that traditional varieties are locally adapted along an abiotic environmental gradient. Using genotype-environment association methods and differentiation-based genome scans, I detected selection signatures within the genome, including a gene model known in other plant species for its role in regulating responses to diverse abiotic stresses. These findings offer novel adaptive alleles and broader regions under positive selection that expand the current knowledge of the adaptive genetic architecture and potential in tef. Furthermore, this The research presented

in this Chapter was conducted as a collaborative project with the Ethiopian Institute of Agricultural Research and has been submitted for publication to *G3*.

Chapter 3: Phenome-to-genome insights for evaluating root system architecture in maize

In Chapter 3, forward and reverse genetic approaches were used to test multiple phenotyping and analytical frameworks designed for quantifying root system architecture across three field experiments of maize lines. This research provides a comprehensive evaluation of various quantitative measurements of root architecture in their description of the overall root phenome and the significance of this information for mapping genotype-phenotype relationships. My findings support the use of complementary approaches to bridge the phenotype-to-genotype gap in field studies focused on belowground variation. The research presented in this Chapter was conducted in collaboration with researchers at the Donald Danforth Plant Science Center and will be submitted for publication to *The Plant Genome* in Fall 2024.

Chapter 4: S-type anion channel homologue (ZmSLAH) regulates root-to-shoot phenotypic plasticity and nitrogen capture in maize

In Chapter 4, I used reverse genetic approaches to test the function of a gene predicted to control variations in root system architecture in field-grown maize. A replicated nitrogen stress field experiment was conducted using a maize population segregating for a loss-of-function mutation to test the hypothesis that alternative states of the candidate gene mediate plant nutrient signaling, influencing distinct root growth patterns in the field. The field measurements provided functional evidence supporting this hypothesis, highlighting a promising gene target for developing low-input germplasm with increased nitrogen use efficiency in maize. As a result, the research presented in this chapter was used to file a provisional patent application in February 2024, and the resulting manuscript will be submitted for publication in Winter 2024. Further

functional characterization of the anion membrane transport phenotypes for this gene model is being performed in collaboration with researchers from the University of Würzburg.

REFERENCES

Angert, A. L., & Schemske, D. W. (2005). the Evolution of Species' Distributions: Reciprocal Transplants Across the Elevation Ranges of Mimulus Cardinalis and M. Lewisii. *Evolution*, 59(8), 1671. https://doi.org/10.1554/05-107.1

Bernardo, R. (2020a). Reinventing quantitative genetics for plant breeding: something old, something new, something borrowed, something BLUE. *Heredity*. https://doi.org/10.1038/s41437-020-0312-1

Bernardo, R. (2020b). *Breeding for quantitative traits in plants*, 3rd ed. Stemma Press, Woodbury, Minnesota.

Caruso, C. M., Martin, R. A., Sletvold, N., Morrissey, M. B., Wade, M. J., Augustine, K. E., Carlson, S. M., MacColl, A. D. C., Siepielski, A. M., & Kingsolver, J. G. (2017). What Are the Environmental Determinants of Phenotypic Selection? A Meta-analysis of Experimental Studies. *The American Naturalist*, 190(3), 363–376. https://doi.org/10.1086/692760

Darwin 1809-1882, C. (1859). *On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life*. London : John Murray, 1859. https://search.library.wisc.edu/catalog/9934839413602122

De Villemereuil, P., Gaggiotti, O. E., Mouterde, M., & Till-Bottraud, I. (2016). Common garden experiments in the genomic era: New perspectives and opportunities. *Heredity*, 116(3), 249–254. Nature Publishing Group. https://doi.org/10.1038/hdy.2015.93

Dudley, J. W., & Moll, R. H. (1969). Interpretation and Use of Estimates of Heritability and Genetic Variances in Plant Breeding. *Crop Science*, 9(3), https://doi.org/https://doi.org/10.2135/cropsci1969.0011183X000900030001x

Falconer, D. S. (1985). A note on Fisher's "average effect" and "average excess". *Genetical Research*, *46*(3), 337–347. https://doi.org/10.1017/s0016672300022825

Falconer, D. S. (1990). Selection in different environments: Effects on environmental sensitivity (reaction norm) and on mean performance. *Genetical Research*, 56(1), 57–70. https://doi.org/10.1017/S0016672300028883

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. https://doi.org/10.1017/S0080456800012163

Fisher, R. A. (1930). *The genetical theory of natural selection*. Clarendon Press. https://doi.org/10.5962/bhl.title.27468

Frankham, R. (2015). Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology*, 24(11), 2610–2618. https://doi.org/10.1111/mec.13139

Gardner, CO. (1963). Estimates of genetic parameters in cross-fertilizing plants and their implications in plant breeding. In: Hanson, WD., Robinson, HF. (eds) *Statistical genetics and plant breeding*. National Academy of Sciences—National Research Council, Washington, DC, p 225–252.

Gillespie, J. H. (1992). *The Causes of Molecular Evolution*. Oxford University Press. https://doi.org/10.1093/oso/9780195068832.001.0001 n

IPCC. (2023). Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. *Climate Change 2023: Synthesis Report*. [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, pp. 35-115, 10.59327/IPCC/AR6-9789291691647

Janzen, G. M., Aguilar-Rangel, M. R., Cíntora-Martínez, C., Blöcher-Juárez, K. A., González-Segovia, E., Studer, A. J., Runcie, D. E., Flint-Garcia, S. A., Rellán-Álvarez, R., Sawers, R. J. H., & Hufford, M. B. (2022). Demonstration of local adaptation in maize landraces by reciprocal transplantation. *Evolutionary Applications*, 15(5), 817–837. https://doi.org/https://doi.org/10.1111/eva.13372

Josephs, E. B., Berg, J. J., Ross-Ibarra, J., & Coop, G. (2019). Detecting adaptive differentiation in structured populations with genomic data and common gardens. *Genetics*, 211(3), 989–1004. https://doi.org/10.1534/genetics.118.301786

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7(12), 1225–1241. https://doi.org/10.1111/j.1461-0248.2004.00684.x

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press.

Kingsolver, J. G., Hoekstra, H. E., Hoekstra, J. M., Berrigan, D., Vignieri, S. N., Hill, C. E., Hoang, A., Gibert, P., & Beerli, P. (2001). The strength of phenotypic selection in natural populations. *The American Naturalist*, 157(3), 245–261. https://doi.org/10.1086/319193

Lasky, J. R., Josephs, E. B., & Morris, G. P. (2023). Genotype-environment associations to reveal the molecular basis of environmental adaptation. *Plant Cell*, 35(1), 125–138. https://doi.org/10.1093/plcell/koac26 7

Leinonen, T., O'Hara, R. B., Cano, J. M., & Merilä, J. (2008). Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology*, 21(1), 1–17. https://doi.org/10.1111/j.1420-9101.2007.01445.x

Lovell, J. T., MacQueen, A. H., Mamidi, S., Bonnette, J., Jenkins, J., Napier, J. D., Sreedasyam, A., Healey, A., Session, A., Shu, S., Barry, K., Bonos, S., Boston, L. B., Daum, C., Deshpande, S., Ewing, A., Grabowski, P. P., Haque, T., Harrison, M., … Schmutz, J. (2021). Genomic mechanisms of climate adaptation in polyploid bioenergy switchgrass. *Nature*, 590(2). https://doi.org/10.1038/s41586-020-03127-1

Lynch, M., & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer. https://books.google.com/books?id=UhCCQgAACAAJ

Matzinger, DF. (1963). Experimental estimates of genetic parameters and their applications in self-fertilizing plants. In: Hanson, WD., Robinson, HF. (eds) *Statistical genetics and plant breeding*. National Academy of Sciences—National Research Council, Washington,

Maynard Smith, J. (1970). Natural Selection and the Concept of a Protein Space. *Nature*, 225(5232), 563–564. https://doi.org/10.1038/225563a0

Orr, H. A. (1998). The Population Genetics of Adaptation: The Distribution of Factors Fixed during Adaptive Evolution. *Evolution*, 52(4), 935–949. https://doi.org/10.2307/2411226

Orr, H.A. (1999). The evolutionary genetics of adaptation: A simulation study. *Genetical Research*, 74(3), 207–214. https://doi.org/10.1017/S0016672399004164

Orr, H. A. (2005). The Genetic Theory of Adaptation: A Brief History. *Nature Reviews Genetics*. 6(2), 119–127. https://doi.org/10.1038/nrg1523

Price, N., Moyers, B. T., Lopez, L., Lasky, J. R., Monroe, J. G., Mullen, J. L., Oakley, C. G., Lin, J., Ågren, J., Schrider, D. R., Kern, A. D., McKay, J. K., Grey Monroe, J., Mullen, J. L., Oakley, C. G., Lin, J., Ågren, J., Schrider, D. R., Kern, A. D., & McKay, J. K. (2018). Combining population genomics and fitness QTLs to identify the genetics of local adaptation in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, 115(19), 5028–5033. https://doi.org/10.1073/pnas.1719998115

Reiskind, M. O. B., Moody, M. L., Bolnick, D. I., Hanifin, C. T., & Farrior, C. E. (2021). Nothing in Evolution Makes Sense Except in the Light of Biology. *BioScience*, 71(4), 370–382. https://doi.org/10.1093/biosci/biaa170

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, 24(17), 4348–4370. https://doi.org/10.1111/mec.13322

Walsh, B. (2014). Special issues on advances in quantitative genetics: introduction. *Heredity*, 112(1), 1–3. https://doi.org/10.1038/hdy.2013.115

Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, 16(2), 97–159. https://doi.org/10.1093/genetics/16.2.97

CHAPTER 2

GENETIC DIVERSITY AND ENVIRONMENTAL ADAPTATION IN ETHIOPIAN TEF

Summary

Orphan crops serve as essential resources for both nutrition and income in local communities and offer potential solutions to the challenges of food security and climate vulnerability. Tef [*Eragrostis tef* (Zucc.)], a small-grained allotetraploid, C4 cereal mainly cultivated in Ethiopia, stands out for its adaptability to marginal conditions and high nutritional value, which holds both local and global promise. Despite its significance, tef is considered an orphan crop due to limited genetic improvement efforts, reliance on subsistence farming, and its nutritional, economic, and cultural importance. Although pre-Semitic inhabitants of Ethiopia have cultivated tef for millennia between 4000 and 1000 BCE, our understanding of the current degree of local adaptation, the loci underlying it, and the potential for future adaptation remains limited. To address this issue, we resequenced traditional varieties of tef that were sampled across a range of production environments and tested for population structure and regions of the genome under selection. We identified 145 potential adaptive loci using Redundancy Analysis, complemented by single-site differentiation and LD-based methods to discern selection signatures. This study revealed putatively adaptive alleles within our sampling population that can be used for tef improvement, leading to the development of new varieties with enhanced adaptation to current and future conditions.

Introduction

Divergent selection pressures across a species range naturally drive adaptive differentiation among populations, leading to the maintenance of genomic and phenotypic variation that underlies local adaptation. Such local adaptation is expected and observed in natural populations, traditional crop varieties, and its wild relatives, serving as a source for novel traits, alleles, and donor lines for developing plant varieties that thrive in current and future climates. Understanding the genetics of such adaptation and characterizing the selective agents that drive adaptive evolution allows a mechanistic approach to designing and predicting crop suitability across current and future target populations and environments for a given crop.

Recent studies have described the genetic architecture of adaptation and identified environmental-associated loci by conducting extensive genome-wide analyses of wild and domesticated species, e.g. Arabidopsis [*Arabidopsis thaliana* (L.) Heynh.] (Agren et al., 2013; Fournier-Level et al., 2011; Hancock et al., 2011; Lasky et al., 2012), rice [*Oryza sativa* (L.)] (Huang et al., 2010), maize [*Zea mays* (L.)] (Hufford et al., 2012; Josephs et al., 2019), sorghum [*Sorghum bicolor* (L.) Moench] (Lasky et al., 2015; Morris et al., 2013), soybean [*Glycine soja* (Sieb. & Zucc.)] (Wang et al., 2022), bapolrley [*Hordeum vulgare* (L.)] (Abebe et al., 2015), pine [*Pinus taeda* (l.)] (Eckert et al., 2010a; Eckert et al., 2010b; Lu et al., 2019), and peach [*Prunus persica* (L.)] (Li et al., 2021). Research on crop species, especially traditional varieties, has shown their ability to adapt to diverse climates over generations of natural and artificial selection. However, in many crops, the significant reduction in effective population sizes during both domestication and crop improvement has led to the loss of adaptive variation and the accumulation of deleterious mutations (Moyers et al., 2018). In annual crops, there has been an observed loss of approximately 40% of diversity compared to wild relatives (Miller & Gross,

2011). Traditional varieties and their wild relatives offer abundant genetic diversity, making them invaluable assets for breeding programs (Bolger et al., 2014; Gao et al., 2019; Li et al., 2018; Wang et al., 2022).

Tef [*Eragrostis tef* (Zucc.)] is a small-grained C_4 cereal crop that is of significant importance for smallholder farmers in Eastern Africa. Belonging to the Poaceae family and genus *Eragrostis*, tef is a self-pollinated annual plant that is cultivated primarily as a grain for human consumption or as biomass for forage. As a short-day plant, tef is well adapted to altitudes ranging from 1800 to 3000 m, with optimal growth at mean temperatures of 15–21°C, annual rainfall between 400 and 750 mm, and in fertile soils with a pH range of 5.2–5.5 (Belay & Assen, 2023). Tef is of major cultural and economic importance in Ethiopia, where the grain is milled to produce injera, a nutrient-dense, gluten-free Ethiopian bread. In Western countries, tef is valued as a healthy food alternative for individuals with Celiac disease. Despite its nutritional value and ability to thrive under unfavorable conditions for other cereals, tef is considered an orphan crop and has experienced limited genetic improvement.

In the last decade, genomics research in tef has gained traction with the release of draft genome sequences (Cannarozzi et al., 2014) and a high-quality genome sequence (VanBuren et al., 2020). Tef has an allotetraploid origin and 20 chromosome pairs with a relatively small genome of approximately 622 Mega-base pairs (VanBuren et al., 2020). Studies investigating phenotypic variation in tef germplasm have found differences in days to maturity, harvest index, number of culm nodes, and diameters of first and second shoot culm internodes across collections from different elevations and regions (Assefa et al., 2000, 2002; Kefyalew et al., 2000; Tadesse, 1993; Woldeyohannes et al., 2020). Advancements in breeding techniques, complemented by insights into the diversity and inheritance of target agronomic traits, have been

key to the genetic improvement of tef and have contributed to the total release of 49 improved varieties from the Ethiopian Institute for Agricultural Research since the 1950s (Assefa et al., 2011; Ministry of Agriculture and Livestock Resources, 2019).

Ethiopia is one of the most agroecologically diverse countries in the world. As one of the twelve Vavilovian centers of diversity, it features a vast range of altitudinal clines spanning from 110 m below sea level to 4620 m above sea level (Abebe et al., 2015; Vavilov, 1951). These regions experience large variability in both seasonal and annual climatic conditions. Although, environmental heterogeneity may be a major driver of the richness of plant biodiversity across Ethiopia, climate change is predicted to lead to a major decrease in crop productivity through increased water and heat stress, the prevalence of diseases and pests, and shortened growing seasons (IPCC, 2014). To meet current and future agricultural demands across diverse agroecologies, it is essential to survey the existing allelic diversity of the traditional varieties grown across their range of cultivation.

In this study, we tested the hypothesis that traditional varieties of tef are locally adapted along an abiotic environmental gradient throughout its range of cultivation. To address this hypothesis, we sampled and resequenced a collection of tef germplasm across their current distribution to (i) characterize the extent and geographic distribution of the population structure; (ii) identify genetic variations associated with multivariate environmental gradients; (iii) isolate environmental factors contributing to population differentiation; (iv) detect genomic regions that have undergone recent selective sweeps; and (v) predict the suitability of existing germplasm to future climates. To achieve these aims, we employed multivariate genotype-environment association methods and differentiation-based genome scans.

Materials and Methods

Samples, Variant Calling, and Genotyping

The tef germplasm panel used in this study was sourced from the Ethiopian Biodiversity Institute (EBI) as part of a collaborative project with the Ethiopian Institute of Agricultural Research (EIAR). The panel consisted of 96 traditional tef varieties, each with known geographical coordinates from across major production regions. Seeds from each variety were sown in a greenhouse at EIAR, Ethiopia, and after 4 weeks, young leaves were sampled from a single plant representative of each variety. Genomic DNA was extracted from the plant leaf tissue in 96-well plate format using a modified version of the CTAB (Cetyltrimethylammonium bromide) method (Chua et al., 1990). DNA concentration and quality were determined using a 1.5% agarose gel and a Nano Drop Spectrophotometer. The extracted DNA was sent to the BioFrontiers Sequencing Core at the University of Colorado Boulder (Boulder, Colorado, USA) for unique dual indexing of 10-bp libraries. Genomic libraries were then sent to the Genomics and Microarray Core at the University of Colorado Anschutz Medical Campus (Aurora, Colorado, USA) for paired-end whole-genome sequencing (2 × 150 bp at ~ 10× coverage) using an Illumina NovaSeq 6000 platform.

The raw paired-end 150-nt reads were scanned for adapter sequences, low-quality base pairs, and reads shorter than 35 bp were trimmed using Fastp version 0.19.6 (Chen et al., 2018). Trimmed reads were aligned to the *E. tef* v3 genome (CoGe id50954) (VanBuren et al., 2020) using BWA-MEM version 0.7.17 with default settings (Li, 2013). The resulting Sequence Alignment Map files were sorted by chromosome, annotated for read group information, filtered for improperly paired reads, and reads with a mapping quality of less than 10, were converted to Binary Alignment Map format and indexed using SAMtools version 1.1 (Li et al., 2009) before

proceeding to SNP calling using GATK Best Practices recommendations (Poplin et al., 2017; van der Auwera & O'Connor, 2020). SNP calling was performed using GATK HaplotypeCaller with default settings, and the resulting Genomic Variant Call Format files were combined into a single database using GenomicDBImport for joint genotyping with GenotypeGVCF (McKenna et al., 2010). The resulting Variant Call Format (VCF) files were filtered to contain only bi-allelic single nucleotide polymorphisms (SNPs) that met quality criteria using R/vcfR (Knaus & Grünwald, 2017) and GATK SelectVariants and VariantFiltration tools (McKenna et al., 2010). Finally, TASSEL version 5.2.83 (Bradbury et al., 2007) was utilized to remove loci with a minor allele frequency ≤ 5% and heterozygous genotype calls ≥ 50%, as well as taxa that possessed a genotyping rate ≤ 80%. The filtered VCF file comprised 538,252 markers from 87 individuals and was used for subsequent analyses.

Climate Data

Statistical analyses of climatic variation were conducted on 87 georeferenced traditional tef varieties from 75 sampling locations to assess the distribution and magnitude of environmental conditions at each site. Bioclimatic (bio1-bio19) and altitudinal data were obtained through the latitude and longitude coordinates from each site using WorldClim version 2.1 layers at 10-minute resolution based on long-term averages for 1970-2000 (Fick & Hijmans, 2017). Descriptions and calculations for each WorldClim bioclimatic variable are presented in Table 2.1. The relationships among the 20 environmental variables across 75 locations were examined through principal component analysis (PCA) using R/FactoMineR (Lê et al., 2008) and R/factoextra (Kassambara & Mundt, 2020).

Table 2.1. Description of WorldClim version 2.1 climate and altitudinal variables (www.worldclim.org).

Predictor	Description
bio1	Annual Mean Temperature
bio2	Mean Diurnal Range (Mean of monthly (max – min temperature))
bio3	Isothermality (BIO2/BIO7) (×100)
bio4	Temperature Seasonality (standard deviation ×100)
bio5	Max Temperature of Warmest Month
bio6	Min Temperature of Coldest Month
bio7	Temperature Annual Range (BIO5-BIO6)
bio8	Mean Temperature of Wettest Quarter
bio9	Mean Temperature of Driest Quarter
bio10	Mean Temperature of Warmest Quarter
bio11	Mean Temperature of Coldest Quarter
bio12	Annual Precipitation
bio13	Precipitation of Wettest Month
bio14	Precipitation of Driest Month
bio15	Precipitation Seasonality (Coefficient of Variation)
bio16	Precipitation of Wettest Quarter
bio17	Precipitation of Driest Quarter
bio18	Precipitation of Warmest Quarter
bio19	Precipitation of Coldest Quarter
Altitude	Altitude (in m.a.s.l.)

We then used the Coupled Model Intercomparison Project (CMIP6), incorporating the Shared Socioeconomic Pathways (SSPs) developed by the Intergovernmental Panel on Climate Change (IPCC), to project future climate scenarios. The latest phase of this model integrates new emission and land-use scenarios compared to previous models (Eyring et al., 2016; O'Neill et al., 2016; Stouffer et al., 2017; Wyser et al., 2020). Our study focused on the WorldClim v2.1 bias-corrected SSP2-4.5 scenario, which is considered a median pathway among the five SSP scenarios (SSP1 to SSP5) that represent varying mitigation and adaptation challenges (Huang et al., 2019). Using a set of 12 CMIP6 global circulation models (listed in Table 2.2), we calculated the mean ensemble of projected climate change for the period 2061-2080 implementing R/raster (Hijmans, 2023).

Table 2.2. CMIP6 Global Circulation Models utilized for estimation of genomic offset for *E. tef* in Ethiopia.

Model Name	Source
ACCESS-CM2	(Dix et al., 2019)
CMCC-ESM2	(Lovato et al., 2023)
EC-Earth3-Veg	(EC-Earth Consortium (EC-Earth), 2019)
FIO-ESM-2-0	(Song et al., 2019)
GISS-E2-1-G	(NASA Goddard Institute for Space Studies (NASA/GISS), 2020)
HadGEM3-GC31-LL	(Good, 2019)
INM-CM5-0	(Volodin et al., 2019)
IPSL-CM6A-LR	(Boucher et al., 2019)
MIROC6	(Shiogama et al., 2019)
MPI-ESM1-2-HR	(Schupfner et al., 2019)
MRI-ESM2-0	(Yukimoto et al., 2019)
UKESM1-0-LL	(Good et al., 2019; Shim et al., 2020)

Inference of Population Structure

Accounting for population structure in genotype-to-phenotype association studies is essential for reducing the number of false positives (Coop et al., 2009; Excoffier et al., 2009; Excoffier & Ray, 2008; Kang et al., 2008; Pritchard et al., 2000; Rellstab et al., 2015; Yu et al., 2006). Principal component analysis (PCA) and genome-wide discriminant analysis of principal components (DAPC) were conducted to quantify population structure using R/adegenet (Jombart, 2008; Jombart & Ahmed, 2011). Individual genetic ancestries were estimated using the Q matrix to determine the appropriate number of genetic clusters (K) through cross-validation. K values for subpopulations ranging from K = 1 to K = 10 were tested, and K = 5 was selected for analyses. To estimate the degree to which marker data were structured by each of the environmental predictors, we examined the relationships between the ordinal axes generated by the genome-wide DAPC and the environmental predictors. The statistical dependence of each ordinal axis on each predictor was assessed using the corr.test function in R/psych (Revelle, 2023), with the method set to “spearman” and α level set to 0.05.

To assess the extent of isolation by distance (IBD) and isolation by environment (IBE) due to population structure, we implemented the Mantel test using the Mantel function in R/vegan based on Spearman's rank correlation. Spatial autocorrelation was performed along latitude and longitude using spherical coordinates to create a geographical distance matrix (in km) for sampling locations with two or more collected varieties, using the rdist.earth function in R/fields (Nychka et al., 2021). Similarly, environmental autocorrelation was carried out for these locations using Euclidean pairwise distances across 20 environmental variables, calculated using R/rdist (Blaser, 2020). Site-level genetic distances were estimated as pairwise FST using the genet.dist function in R/hierfstat (Goudet & Jombart, 2022), applying the "WC84" (Weir & Cockerham, 1984) method. Although our panel consisted of 96 individuals sampled from 75 distinct geographic locations, only 16 locations had more than one sample. As a result, the distance matrices for the IBD and IBE analyses were based on 39 individuals from these 16 locations. Negative FST estimates were converted to zero values for subsequent analyses. The significance of each Mantel statistic (r) was evaluated with a two-tailed 95% confidence interval centered on the null hypothesis of no spatial or environmental correlation with genetic structure ($r = 0$) with 9999 permutations.

Multivariate Genotype × Environment Association Analysis

Redundancy Analysis (RDA), a constrained ordination technique, was employed to detect adaptive processes and multilocus architectures associated with environmental variations using R/vegan (Oksanen et al., 2022). As a common multivariate genotype-environment association approach for exploring the relationship between genetic and environmental data, RDA has demonstrated promise for identifying sets of covarying loci that most strongly correlate with environmental predictors under varying selective strengths (Bourret et al., 2013, 2014; Forester

et al., 2018; Lasky et al., 2012; Legendre & Legendre, 2012; Rellstab et al., 2015). As RDA requires complete data frames, missing values identified within the genotype matrix (approximately 14.38% of the data) were imputed by using the most common genotype at each marker across all individuals before converting the data to allele counts for each locus and variety (Forester et al., 2018). Moreover, RDA is highly susceptible to problems when using highly correlated predictors. The use of highly correlated predictors in the RDA can increase its susceptibility to multicollinearity, thereby hindering the accurate assessment of individual predictor contributions (Dormann et al., 2013). To address this issue, we identified correlated environmental predictors of $|r| > 0.70$ using R/psych and excluded them from RDA model testing. Multicollinearity was further controlled by checking the variance inflation factors, with a threshold of 5 (Zuur et al., 2010).

Variance partitioning quantifies the proportion of variance attributed to a specific set of explanatory variables (e.g., climate) after accounting for the effects of other variables (e.g., population structure) (Legendre & Legendre, 2012). We used partial RDA-based variance partitioning to separate and evaluate the contributions of climate and population structure in explaining genetic variation. Formal tests for the analysis of variance were employed to assess the significance of all model combinations. Each model's statistical significance was evaluated using a permutation-based analysis of variance (ANOVA) with 999 permutations and a significance level (α) of 0.05 (Legendre & Legendre, 2012). The null hypothesis is that there is no linear relationship between genetic data and environmental predictors. The number of constrained axes retained for loci detection was visually determined using a screeplot of the canonical eigenvalues. We identified candidate adaptive loci as SNPs loading ±3.29 SD from the mean loading of these significant RDA axes. We then identified the covariates that were most

strongly correlated with each candidate locus (i.e., highest correlation coefficient) to group candidates by potential driving environmental variables. Loci loading in the tails on each RDA axis, with a two-tailed 99.9% confidence interval around the null hypothesis (±3.29 SD from the mean axis loading), were identified as putative adaptive loci (Capblancq & Forester, 2021; Forester et al., 2018).

Identification of Selection Signatures

We implemented two alternative approaches that incorporate intrapopulation (iHH12) and inter-population ($\overline{F}_{ST}$) statistics. Integrated Haplotype Homozygosity Pooled (iHH12) is a linkage disequilibrium (LD) based method that can identify genomic regions under positive selection by measuring the extent to which haplotypes are shared and homozygous in a population (Voight et al., 2006). For each site, we calculated the iHH12 score using the software selscan v2.0.0 (Szpiech, 2021; Szpiech & Hernandez, 2014), which offers better power than the integrated haplotype score (iHS) for soft sweep detection (Garud et al., 2015; Torres et al., 2018). These iHH12 scores were then normalized to Z-scores (standard normal scores), and 20-kb sliding windows were employed to identify potential target regions of positive selection, considering only the top 1% of the distribution of iHH12 values. By treating all samples from different locations as one population, we can identify genomic regions where recent selective sweeps have occurred without considering specific environmental factors. This approach can reveal general selection patterns that are advantageous across various environments, highlighting coordinated changes in allele frequency common to all tested tef varieties.

The concept of $\overline{F}_{ST}$, pioneered by Wright (1949), assesses differences in allele frequencies across populations and can be used to identify genomic regions with distinct fixation patterns. $\overline{F}_{ST}$ values can range from 0, which suggests no genetic differentiation, to 1, complete genetic

differentiation. Mean $\overline{F}_{ST}$ was estimated using a sliding window approach weighted by the number of sites, with a window size of 20-kb and a step size of 20-kb, using the software VCFtools version 0.1.16 (Danecek et al., 2011; Weir & Cockerham, 1984). To examine the degree of divergent selection within the distribution of select environmental variables, individuals were grouped based on the lower and upper quartiles of each explanatory variable used in the RDA. Altitude (in m.a.s.l) was included as an additional variable in the analysis due to the considerable variation in adaptive and farmer-preferred traits within and across altitudinal classes (Woldeyohannes et al., 2020). Genomic windows with $\overline{F}_{ST}$ values ≤ 0 were omitted from the analysis. The genomic windows with the highest 1% $\overline{F}_{ST}$ values were considered to represent signatures of divergent selection across environmental gradients. iHH12 scores and sliding window $\overline{F}_{ST}$ were visualized using R/ggplot2 (Wickham, 2016) and R/qqman (Turner, 2018), respectively.

Functional Annotation and Candidate Gene Analysis

Gene candidates were identified as gene IDs in the *E. tef* v3 genome located within 1-kb downstream or upstream of significant RDA loci. For the functional annotation of these gene models, we performed a blastp search against the National Center for Biotechnology Information (NCBI) nonredundant protein database combined with the Gene Ontology (GO) annotation database of genes based on putative rice orthologs (see Yu et al., 2023). The top homologous matches in the NCBI query were selected based on maximum sequence similarity (≥ 95%), maximum bit-score (≥ 50), and an E-value cutoff of 0.05. GO enrichment analysis of significant RDA loci was performed using R/clusterProfiler (Wu et al., 2021;Yu et al., 2012) following the methodology outlined by Yu et al. (2023). Furthermore, gene models were evaluated for the

presence of locus-specific non-synonymous sequence changes and evidence of overlap with regions under positive selection.

Predicting Adaptive Landscapes and Local Maladaptation

Following the methodologies described by Capblancq and Forester (2021) and Steane et al. (2014), we identified regions within the landscape where the adaptive allele frequencies in local gene pools align with the inferred fitness optima associated with local climates. Leveraging the genotype-environment relationship identified through RDA, we employed these candidate adaptive markers and their prospective environmental drivers to predict an index of genetic composition. This process involved employing loci as multivariate responses in an 'adaptively enriched' RDA, where bioclimatic variables were explanatory factors. Subsequently, by using environmental variable scores along the first two RDA axes, we calculated a genetic-based adaptive index for each environmental pixel using the following formula:

$$Adaptive\ Index = \sum_{i=1}^{n} a_i b_i$$

Here, 'a' represents the loading score of the bioclimatic variable along the RDA axis, 'b' denotes the standardized value specific to that variable at the focal pixel, and 'i' pertains to one of the n variables used in the RDA model. These climatic variable scores were calculated based on how each predictor influences adaptive genetic variation in terms of magnitude and direction along the RDA axes. Consequently, the adaptive index serves as an estimate of genetic similarity or dissimilarity among all pixels across the landscape considering the values of bioclimatic predictors at each location (Capblancq & Forester, 2021; Steane et al., 2014).

To predict future maladaptation, adaptive indices were calculated for each environmental pixel across the Ethiopian landscape considering both current and future environmental conditions. The contrast between these predictions indicates the necessary shift in adaptive index

to accommodate projected climates. Alternatively termed genetic/genomic offset (Fitzpatrick & Keller, 2015), risk of non-adaptedness (Rellstab et al., 2016), or genomic vulnerability (Bay et al., 2018), this proxy was estimated using the CMIP6 SSP2-4.5 projected climate changes for the period 2061-2080. The projections were then mapped across Ethiopia to identify areas potentially at risk of maladaptation.

Results

Geography does not influence genetic diversity observed in traditional tef varieties

A total of 538,252 markers were used to infer the population structure of 87 traditional tef varieties. Genome-wide DAPC and IBD across individuals identified high levels of gene flow across geography, suggesting a lack of spatial influence on tef genetic diversity. The K-means clustering algorithm identified K = 5 as the optimal number of clusters and grouped individuals based on their genetic similarity, with approximately 74% of the genetic variation explained by the first two principal components (Figure 2.1A). Cluster 1 consisted of 24 individuals, Cluster 2 consisted of 17 individuals, Cluster 3 consisted of 16 individuals, Cluster 4 consisted of 21 individuals, and Cluster 5 consisted of 9 individuals. Genetic clustering was based on an algorithm that minimizes variation within clusters while maximizing variation among clusters, which indicates a recent shared evolutionary history among individuals within clusters. Interestingly, genetically similar individuals did not cluster by geography (Figure 2.1B). Furthermore, there were no significant differences in environmental and geographical distribution among clusters. The statistical dependence of each DAPC axis on environmental predictors was weak, with LD1 and LD2 showing the highest correlation to mean temperature in the wettest quarter (LD1: $r = 0.224$, $p = 0.037$; LD2: $r = 0.280$, $p = 0.009$). Finally, we tested for spatial and environmental autocorrelation and found no evidence of IBD or IBE, where $\overline{F}_{ST}$

values were not significantly correlated to geographic distance (Mantel r = -0.006, p = 0.505) or environmental distance (Mantel r = 0.162, p = 0.071) (Figure 2.1C).

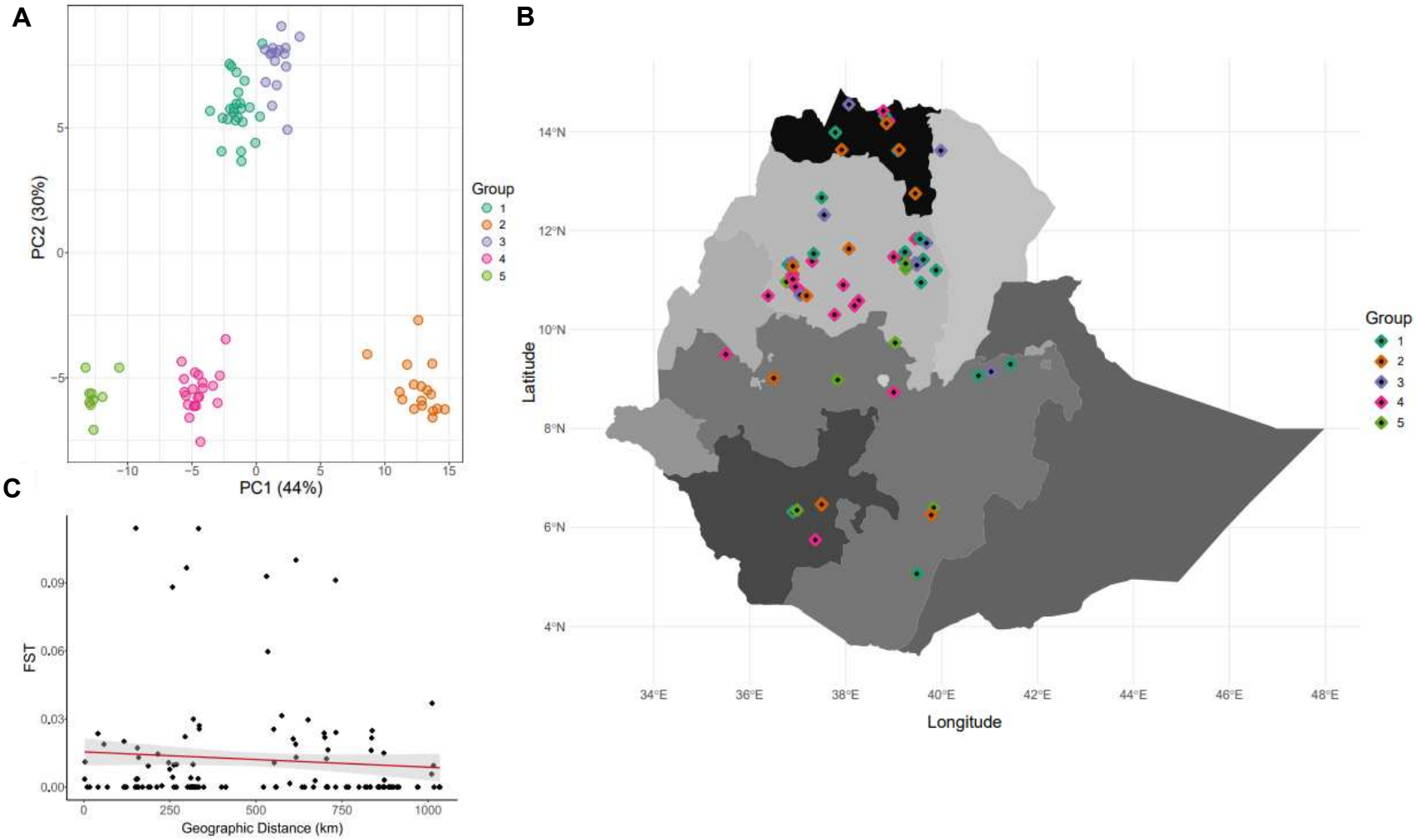

Figure 2.1 Geographic distribution and population structure of the 87 traditional tef varieties sampled across Ethiopia. (**A**) DAPC using genome-wide markers. Taxa are color-coded based on their respective DAPC genetic clusters, as defined in the legend. (**B**) Sampling distribution of tef varieties across Ethiopia, color-coded according to respective DAPC genetic clusters. (**C**) Linear regression of $\overline{F}_{ST}$ values as a function of geographic distance between locations with two or more varieties.

Low proportion of confounded genetic variance shared between climate and population structure

PCA biplots were created to explore the relationships among 20 bioclimatic variables across 75 locations. The first three principal components accounted for 81.70% of the total variance in the data (Figure 2.2). These biplots highlight a clear separation between temperature variables (43.20% variance explained), seasonality variables (26.50%), and precipitation variables (12%). After removing highly correlated predictors, six of the 20 bioclimatic variables were selected for RDA based on model significance ($p = 0.031$) and representation across the predictor types. This included annual mean temperature (bio1), temperature annual range (bio7), precipitation of the wettest month (bio13), precipitation of the driest month (bio14), precipitation of the warmest quarter (bio18), and precipitation of the coldest quarter (bio19). A Variance Inflation Factor check indicated insignificant multicollinearity among the six selected predictors (all values < 5).

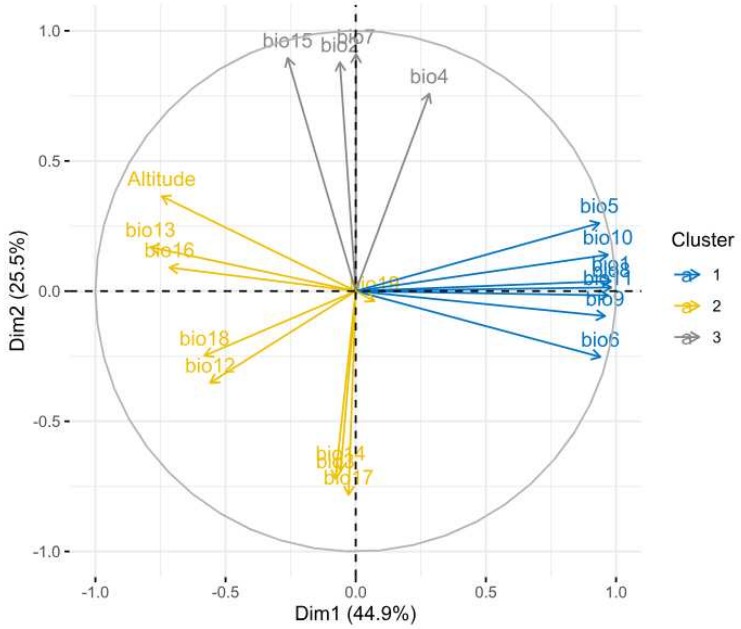

Figure 2.2. PCA biplot of environmental variables in Ethiopia. The biplot demonstrates environmental variation in Ethiopia, encompassing 19 bioclimatic variables and altitude. Loading vectors are grouped using k-means clustering (K = 3) to visualize similarities among variables.

Partial RDA-based variance partitioning was performed to disentangle the effects of climate and population structure on genetic variation. We used two sets of explanatory variables: (a) six bioclimatic variables ('climate') and (b) two proxies for genetic structure (population scores from the first two axes of a genetic PCA on 538,252 markers, termed 'structure'). Six models were analyzed, with individual allele counts as the response variable (Table 2.3). Together, climate and population structure significantly explained 12.03% ($p \leq 0.05$) of the total genetic variance across tef varieties. Climate alone accounted for 0.45% of total genetic variation. Variance partitioning indicated that a low proportion of confounded variance shared between climate and population structure explained 0.18% of the total genetic variation (Figure 2.3A; Table 2.3). This reflects a low degree of collinearity between the sets of explanatory variables. In this study, we evaluated RDA models both with and without controlling for population structure to minimize false negatives. Due to low spatial structure and collinearity between genetic structure and climate, we focused on the pure climate model (i.e., without population structure-based corrections) in downstream analyses.

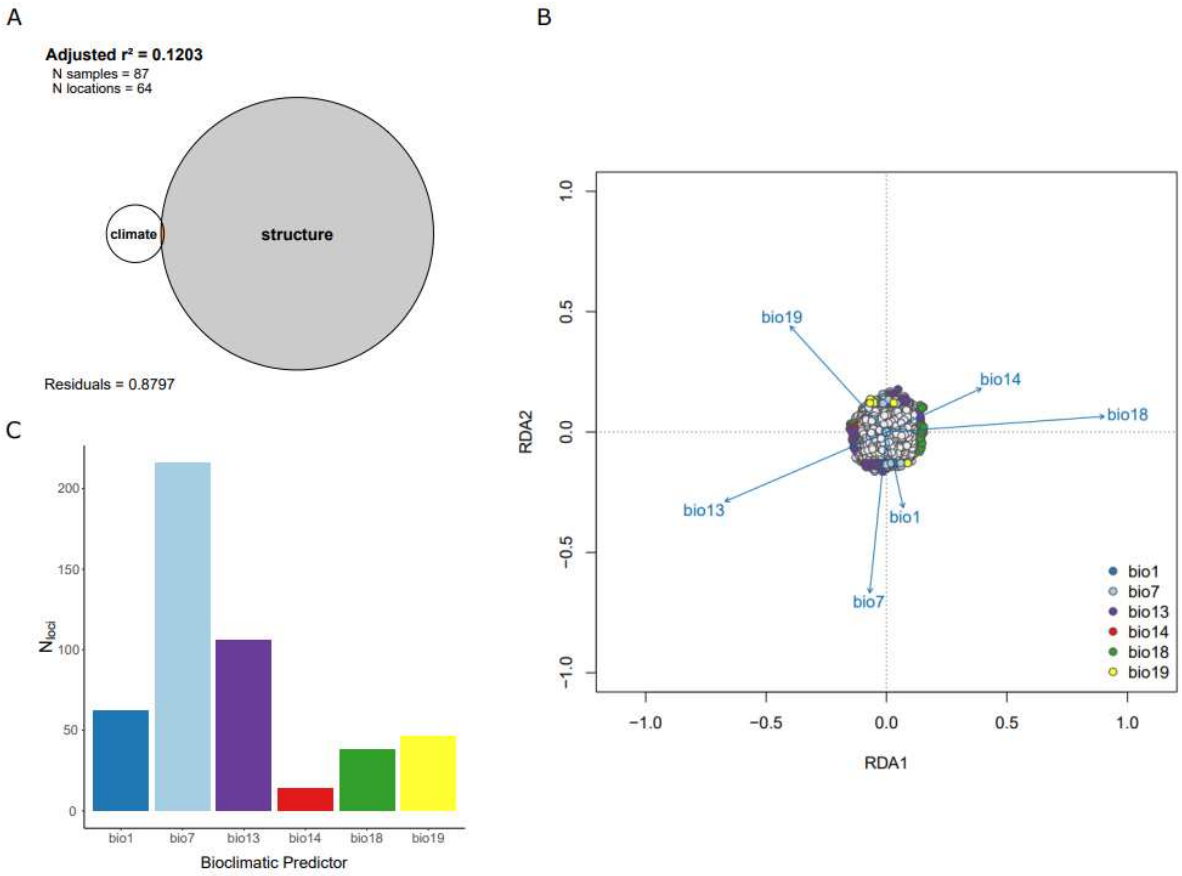

Figure 2.3. Genotype-environment association using RDA. (**A**) Variance partitioning of genomic diversity explained by a combination of the bioclimatic variables 'climate' and PC scores 'structure', represented as a Euler Plot. (**B**) RDA biplot illustrates the relative contribution of climate shaping the genetic structure of 538,252 markers across 87 traditional tef varieties, conditioned on six bioclimatic predictors (bio1: annual mean temperature, bio7: temperature annual range, bio13: precipitation of the wettest month, bio14: precipitation of the driest month, bio18: precipitation of the warmest quarter, bio19: precipitation of the coldest quarter). The outlier loci were grouped by color, reflecting their highest correlation with specific predictors. (**C**) Variation in the counts of significant RDA loci (N = 482) that were attributed to different bioclimatic predictors. The bars are color-coded according to the legend in panel (**B**).

Table 2.3. Effect of climate and genetic structure on genetic variation across RDA models.

RDA model		*Df*	R^2	*Adjusted* R^2	*p(>F)*
Full model:	G ~ climate + structure	8	0.202	0.120	0.001***
Pure climate:	G ~ climate	6	0.075	0.006	0.049*
Pure structure:	G ~ structure	2	0.115	0.115	0.001***
Partial climate:	G ~ climate | structure	6		0.004	0.028*
Partial structure:	G ~ structure | climate	2		0.114	0.001***
	Confounded climate/structure	0		0.002	
	Residual			0.880	

$*p \leq 0.05$; $***p \leq 0.001$
Note: G indicates genetic variation; climate, dataset of six bioclimate variables selected for downstream analyses.

RDA identifies loci associated with multivariate environmental gradients

The constrained ordination of the pure climate model (hereafter, RDA) explained approximately 0.63% of the 538,252 variants (r^2 = 0.076; adjusted r^2 = 0.006), which aligns with the expectation that most loci are likely to be neutral and not strongly associated with environmental predictors (Table 2.3). We then assessed the optimal number of constrained axes for locus detection by constructing a scree plot of the canonical eigenvalues. The first two constrained axes captured the majority of the variation in the model, with RDA1 explaining 25.92% of the variation and RDA2 explaining 18.77%, resulting in a cumulative variance of 44.69%. Consequently, we selected the first two constrained axes for further investigation. By examining histograms of the loadings for each axis, we observed a relatively normal distribution. Loci with loadings centered around zero lack a significant relationship with environmental predictors, whereas loci with loadings in the tails of the distribution are more likely to be under selection by these predictors or other factors correlated with them (Figure 2.4). We identified loci in the tails of these distributions below a significance cutoff of p = 0.001. This resulted in 482 candidate outlier loci, with 63 loci for RDA1 and 419 loci for RDA2.

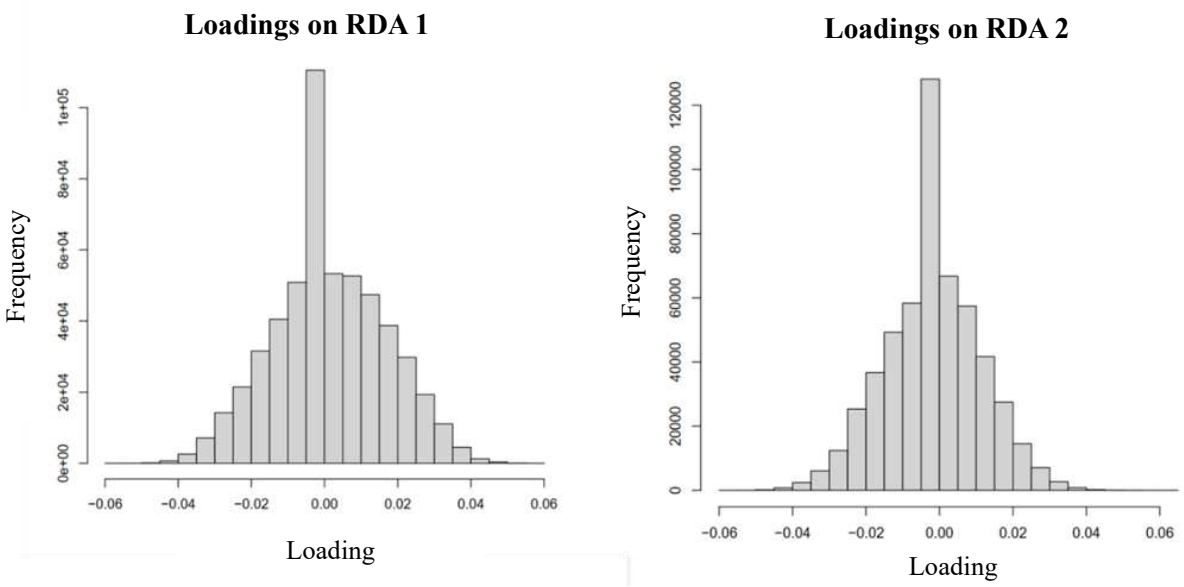

Figure 2.4. Distribution of locus loadings on the first two RDA axes. The position of loci within each distribution indicates potential relationships with environmental predictors. Loci positioned centrally within the distributions exhibit weaker associations, whereas those located at the tails indicate stronger potential relationships.

The majority of the 482 outlier loci (44.81%) exhibited the strongest correlations (r_{max} = 0.420) with the temperature annual range (bio7), followed by 12.86% of loci with the strongest correlations (r_{max} = 0.440) with the precipitation of the wettest month (bio13) (Figure 2.3C). To investigate the relationships between the outlier loci and predictors in more detail, we plotted their loadings against RDA1 and RDA2 (Figure 2.3B). The RDA biplot demonstrates the correlation between genetic variation and environmental predictors, highlighting how each RDA axis varies along adaptive gradients. Our analysis revealed that outlier loci on RDA1 represent multilocus haplotypes associated with precipitation of the wettest month (bio13), precipitation of the driest month (bio14), and precipitation of the warmest month (bio18), while outlier loci on RDA2 represent haplotypes associated with annual mean temperature (bio1), temperature annual range (bio7), and precipitation of the coldest quarter (bio19). These findings introduce novel candidate regions and multilocus haplotypes to better understand tef adaptation across Ethiopia.

High $\overline{F}_{ST}$ regions reveal patterns of genetic divergence across environmental distributions

Genome-wide analysis of genetic differentiation demonstrated a low to moderate level of genetic divergence across environmental variables, with the average $\overline{F}_{ST}$ values across 20-kb sliding windows 0.047 ± 0.050 (Figure 2.5). Among all environmental variables (annual mean temperature – bio1, temperature annual range – bio7, precipitation of the wettest month – bio13, precipitation of the driest month – bio14, precipitation of the warmest quarter – bio18, precipitation of the coldest quarter – bio19, and altitude), we identified 622 genomic regions (comprised of 99 unique regions) with $\overline{F}_{ST}$ values 0.281 ± 0.072, derived from the top 1% of the distribution. Our analysis revealed prominent genetic signals on chromosomes 1A [18,880,001-19,140,001 bp], 1B [31,000,001-31,180,000 bp], and 4B [13,600,000-13,940,000 bp] that were uniquely associated with differences in precipitation of the warmest quarter (bio18), annual mean temperature (bio1), and precipitation of the wettest month (bio13), respectively (Figure 2.5). Furthermore, we detected genetic signals that were colocalized among environmental variables. Specially, we found associations between annual mean temperature (bio1), precipitation of the warmest quarter (bio18), and altitude, spanning chromosomes 3A [5,600,001-6,100,000 bp] and 7A [14,260,001- 14,500,000 bp]. We also identified two regions on chromosome 5A [1,502,000 – 17,500,000 bp; 23,100,001 – 23,410,000 bp] that were shared between annual mean temperature (bio1), precipitation of the warmest quarter (bio18), and altitude, indicating a shared genetic response to these environments.

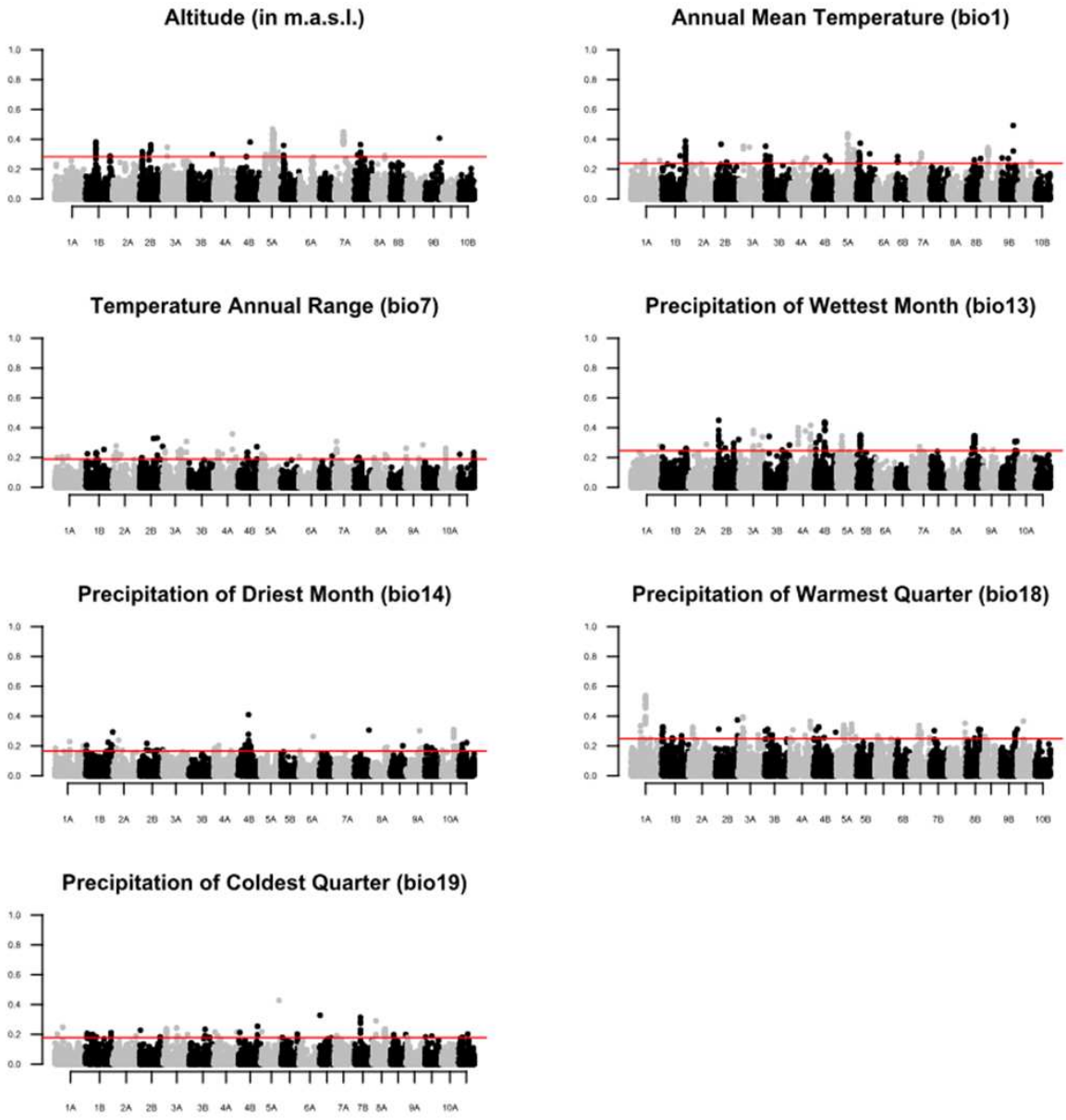

Figure 2.5. Manhattan plots of genome-wide $\overline{F}_{ST}$ within 20-kb windows between populations under divergent environmental conditions. Genetic markers are positioned along the 20 chromosomes for each of six bioclimatic variables and altitude. The red line denotes the top 1% threshold for $\overline{F}_{ST}$ for each climate variable.

Lack of overlap between recent selective sweeps and climate-correlated loci

Genome-wide selection scans using the haplotype-based statistic iHH12 identified 192 genomic windows representing the top 1% of the distribution of iHH12 values, encompassing 5,137 loci and spanning 3.84-Mb across 20 chromosomes. We conducted these selection scans to

detect genomic regions under positive selection within the tef sampling population. Notably, chromosomes 1A, 1B, and 2A exhibited the highest enrichment for selective sweeps. These regions spanned from 1A [33,140,001 - 40,300,001 bp], 1B [33,140,001 – 35,480,001 bp], and 2A [28,760,001 – 35,420,001 bp], with peak signals around 1A:21,220,001, 1B:27,900,001, and 2A:33,220,001. Further analysis revealed a lack of colocalization between these regions and the RDA-identified loci. Additionally, we observed an inverse relationship between iHH12 scores and statistical significance (log-transformed p-values) of the RDA-identified loci. As illustrated in Figure 2.6A, as the transformed RDA p-values increased, there was a concurrent decrease in iHH12 scores. This decay pattern suggests that the RDA-identified adaptive loci have not been subject to recent selective sweeps. Furthermore, a similar genome-wide pattern emerged between iHH12 scores and aggregated $\overline{F}_{ST}$ values , where highly divergent alleles were predominantly associated with lower iHH12 values (Figure 2.7). This finding aligns with our expectations that loci with higher $\overline{F}_{ST}$ values are maintained at intermediate frequencies, whereas loci associated with higher iHH12 scores occur at higher frequencies within a single common haplotype.

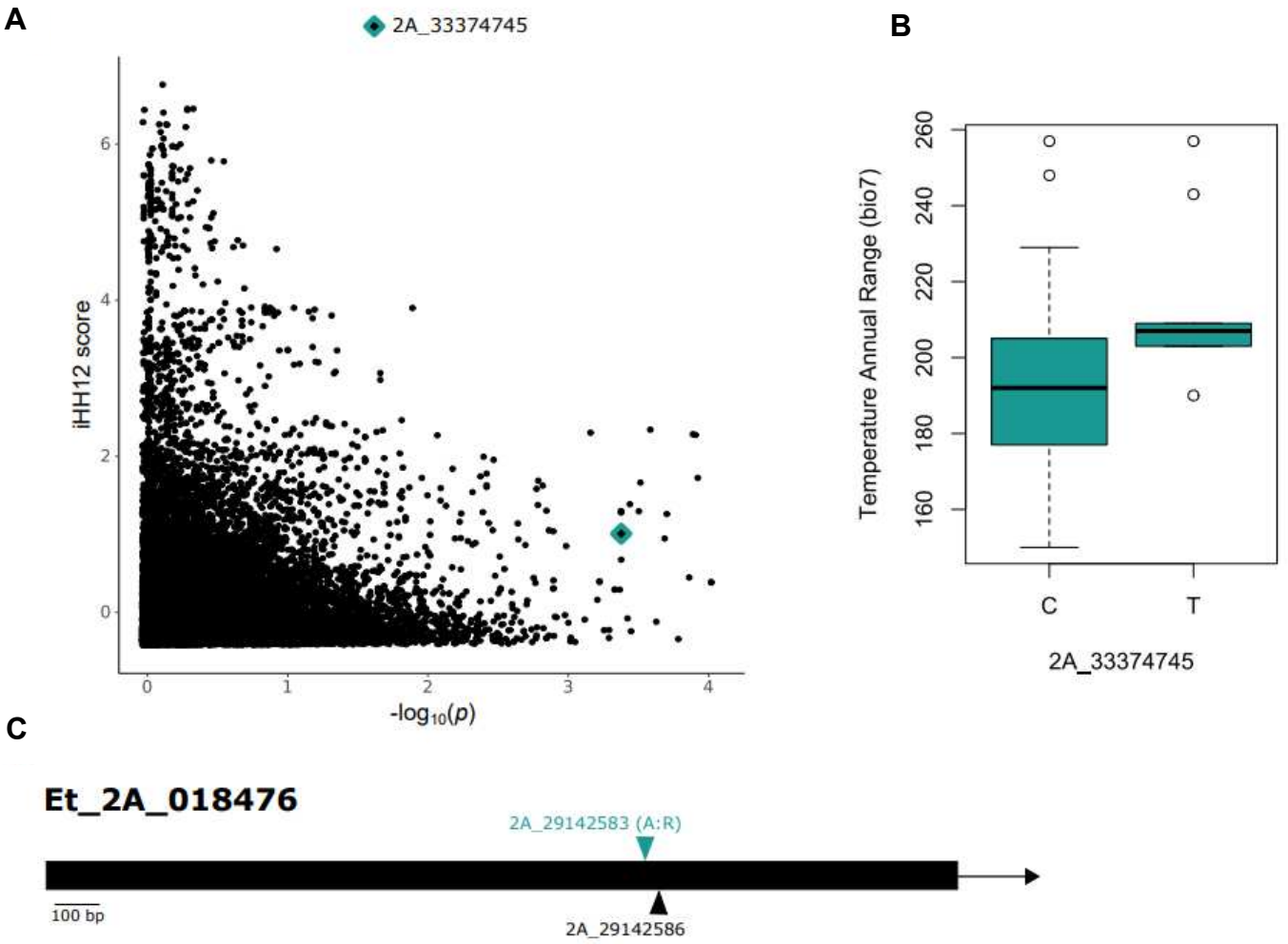

Figure 2.6. Characterization of candidate gene Et_2A_015154. (**A**) Log-transformed p-values from the pure climate RDA model plotted against iHH12 scores on chromosome 2A. Highlighted locus 2A_33374745 confers a non-synonymous nucleotide substitution within its respective gene model. (**B**) Variation in temperature annual range (bio7) based on allelic state at locus 2A_33374745. (**C**) Gene model Et_2A_015154 displaying RDA-significant loci with non-synonymous (green triangle) and synonymous (black triangle) nucleotide substitutions. Amino acid substitutions are represented by letter codes within the label.

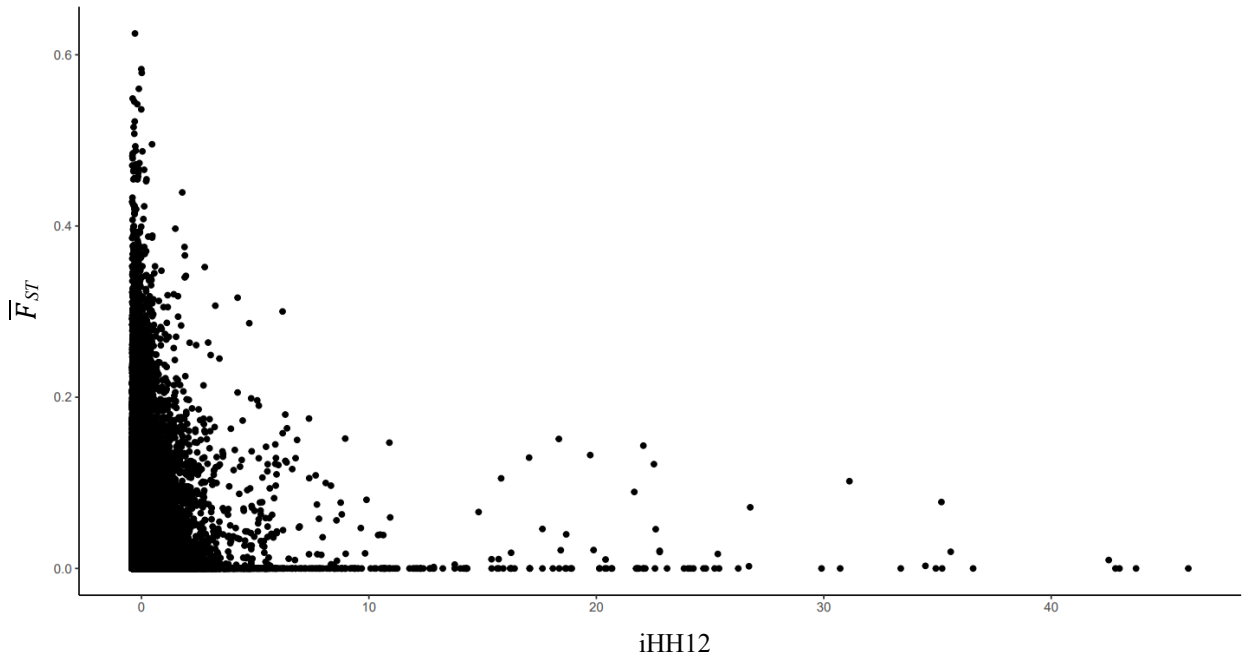

Figure 2.7. Genome-wide patterns of environmental divergence and selective sweeps. The figure illustrates a subset of 100,000 loci sampled without replacement, focusing on $\overline{F}_{ST}$ values related to annual mean temperature (bio1), temperature annual range (bio7), precipitation of the wettest month (bio13), precipitation of the driest month (bio14), precipitation of the warmest quarter (bio18), precipitation of the coldest quarter (bio19), and altitude against iHH12 values.

GO enrichment analysis highlights two candidate genes associated with stress response

GO enrichment analyses of 145 RDA-identified adaptive loci revealed 31 enriched biological process GO terms linked to 92 genes (FDR $p < 0.05$) across five bioclimatic variables (Figure 2.8). The annual mean temperature (bio1) was excluded from the analysis due to a lack of gene sets with associated gene models in the GO annotation database. We identified five prominent functional categories represented by the top 19 enriched GO terms involving 38 genes. These categories included allantoin gene RNA, organelle component organization, protein stress conjugation, intracellular blue light, and cotyledon cytokinin stimulus (Figure 2.8). The GO functional categories exhibited minimal overlap among different bioclimatic variables, except for temperature annual range (bio7), precipitation of the wettest month (bio13), and precipitation of the warmest quarter (bio18), which shared terms in the allantoin gene RNA

functional group. Moreover, the temperature annual range (bio7) and precipitation of the wettest month (bio13) shared the GO term for the cellular response to salt stress (GO:0071472). The largest cluster was related to stress responses and protein modification regulation, featuring 18 genes. Within this cluster, three RDA loci were identified with non-synonymous changes in the coding sequences of Et_1B_010994: c.694G>C (p.Pro232Arg), Et_1B_010994: c.652C>G (p.Gly218Ala), and Et_1B_010993: c.301A>T (p.Glu99Val) (GO: 0071470; GO: 0071472).

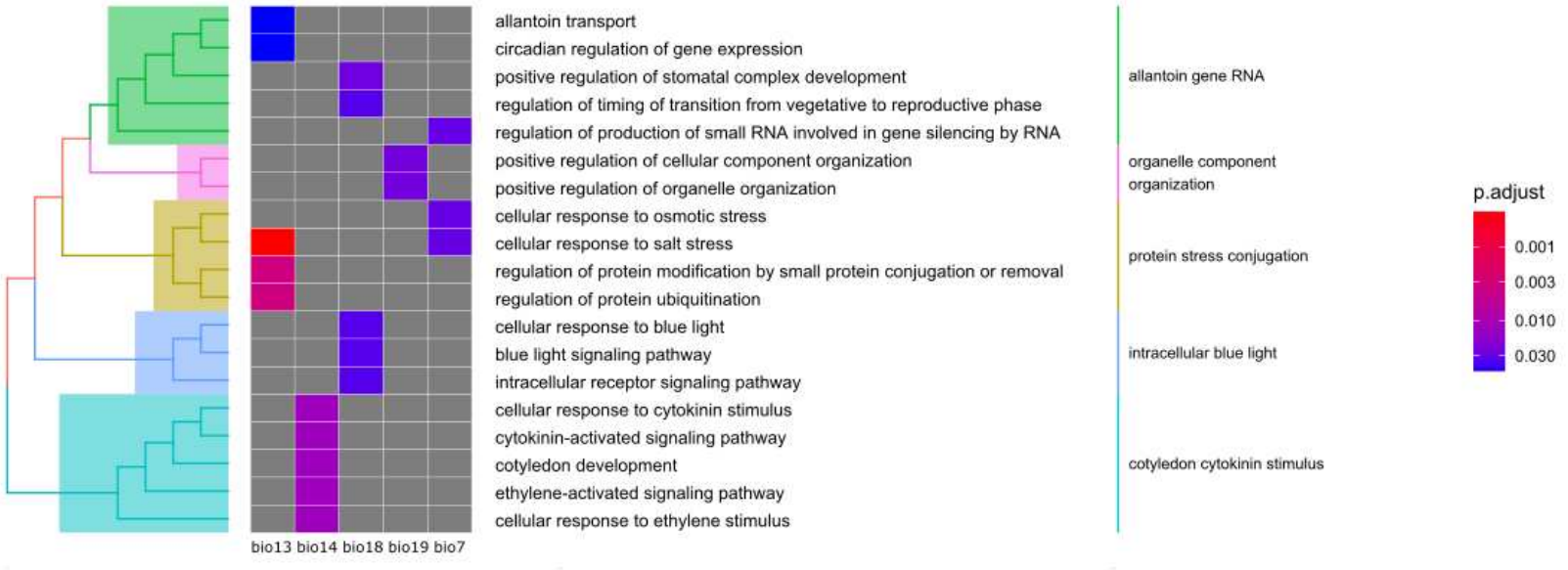

Figure 2.8. Functional grouping tree diagram for GO enrichment result from gene set enrichment analysis. Red and blue to represent smaller and larger p-values, respectively. The analysis involved 145 loci, revealing 92 genes (FDR $p < 0.05$), across five bioclimatic variables gene IDs. Precipitation of the wettest month (bio13), precipitation of the driest month (bio14), precipitation of the warmest quarter (bio18), precipitation of the coldest quarter (bio19), and temperature annual range (bio7), were enriched with 23, 1, 5, 12, and 51 genes, respectively. The top 19 enriched GO terms revealed five distinct functional categories involving 38 genes.

The integrated the results of multivariate genotype-environment association methods and $\overline{F}_{ST}$ genome scans revealed two loci on chromosome 7A proximal to a genetic signal shared among three environmental variables: annual mean temperature (bio1), precipitation of the warmest quarter (bio18), and altitude. These loci, positioned at 14,652,594 bp and 14,764,385 bp, were approximately 390 and 500-kb from the signal, respectively. Furthermore, these two loci showed positive correlations with

temperature annual range (bio7) in the RDA model ($r = 0.300$ and $r = 0.268$). Based on the v3 *E. tef* genome annotations, one locus was located within an intron of the gene model Et_7A_051114. The NCBI blastp search revealed that the closest homologous match for Et_7A_051114 was found in Heller's rosette grass [*Dichanthelium oligosanthes* (Schult.)] and encoded the GTP-binding protein YPTM1 [OEL32215]. YPTM1 belongs to the Rab family of small GTPases, which play essential roles in vesicle transport and membrane trafficking processes in plants (Bischoff et al., 1999). GO enrichment analysis revealed three significant (FDR $p < 0.05$) biological processes related to this gene: pollen tube development, lipoprotein metabolism, and endoplasmic reticulum to Golgi vesicle-mediated transport ($p = 0.0138$).

While there was minimal overlap between the RDA-identified loci and the putative iHH12 sweep regions, several RDA-identified candidate genes were located in proximity. This included one locus identified on chromosome 2A at 33,374,745 bp, approximately 45-kb from the nearest sweep region. This locus was predicted to cause a non-synonymous substitution within the coding sequence of the gene Et_2A_015154: c.702C>T (p.Asp235Asn) and showed a positive correlation with temperature annual range ($r = 0.237$) (Figure 2.6). Although there were three other RDA candidates detected within the coding regions on chromosome 2A, they did not alter the protein sequences. Therefore, a one-way ANOVA was conducted as a *post hoc* analysis to assess the effects of environment and spatial structure on Et_2A_015154. The results showed significant effects of temperature annual range (bio7) and latitude ($p \leq 0.05$) on its allele state. The NCBI blastp results indicated that Et_2A_015154 resembled the mitogen-activated protein kinase kinase kinase 1 in green-millet [XP_034591805], a member of the mitogen activated protein kinase (MAPK) cascade which is involved in the regulation of diverse cellular processes and stress responses (Colcombet & Hirt, 2008; Li et al., 2021; Liu et al., 2019; Liu et al., 2015;

Rodriguez et al., 2010; Šamajová et al., 2013; Sinha et al., 2011; Smékalová et al., 2014; Vadovič et al., 2019). Further functional analysis of Et_2A_015154 revealed significant enrichment for GO terms related to kinase regulation and phosphorylation ($p = 0.013$).

Existing genetic diversity in traditional tef varieties predicts a regional genomic offset under future climate scenario

The 'adaptively enriched genetic space' reveals the strong correlations between genetic markers and diverse bioclimatic variables, as well as the interrelationships among these variables (Figure 2.9). Notably, most of the observed variation was aggregated on RDA1 (67% of the variance), to which allele frequencies were associated with either high temperature annual range (bio7) and high precipitation of the wettest month (bio13) (positive RDA1 scores) or low precipitation in the coldest quarter (bio19) (negative RDA1 scores). This adaptive gradient, when projected onto the landscape, delineates the contrasting conditions between low and high elevation areas, as indicated by the positive and negative RDA1 scores (Figure 2.10A). Additionally, the adaptive gradients related to RDA2 (17% of the variance) distinguished two distinct regions (i.e., Amhara and Benshangul-Gumaz) on the landscape. These regions exhibited lower precipitation levels, particularly during the colder months (negative RDA2 scores), whereas other regions were characterized by higher precipitation levels, especially during the warmer months (positive RDA2 scores). To better understand the potential risks of future maladaptation, we used our enriched RDA model to predict the optimal adaptive indices for current and future environmental conditions across Ethiopia. We estimated shifts between these two predictions (i.e., genomic offset), predicting a contrasting pattern for 2080 (Figure 2.10B). Our model predicted a risk of maladaptation in northern Ethiopia (genomic offset ≥ 4) compared to the rest of the region currently suitable for tef cultivation.

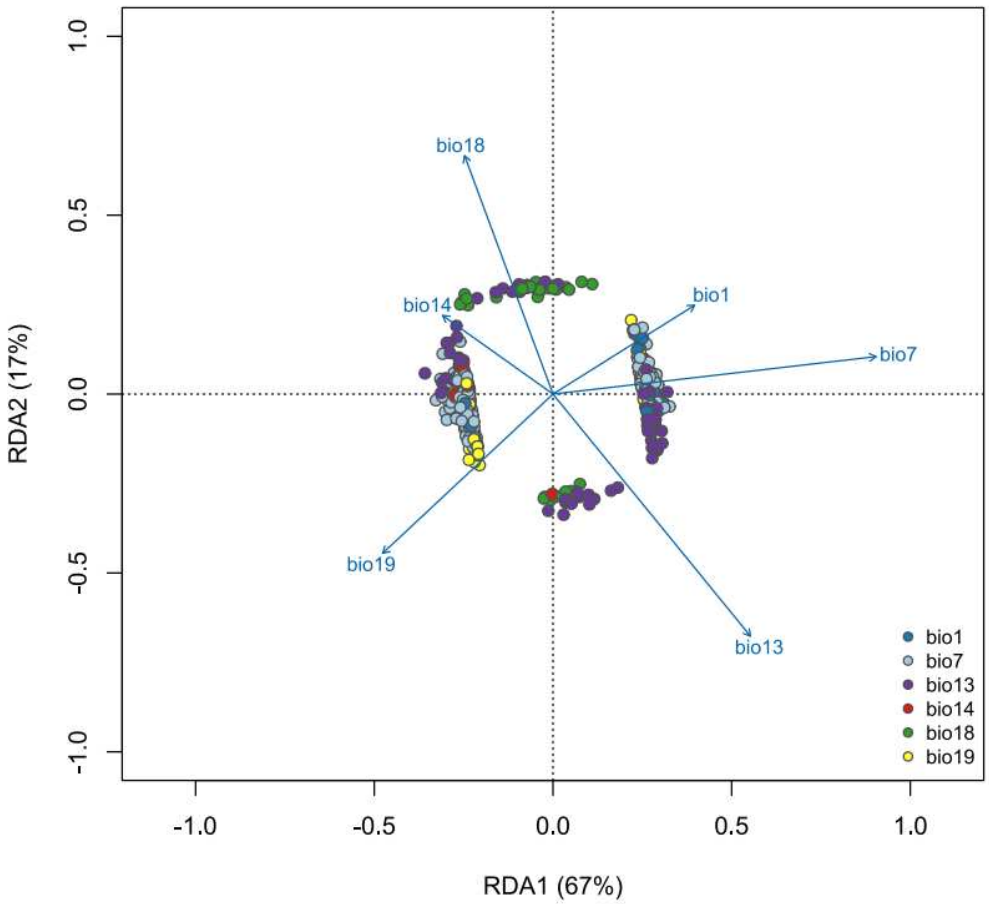

Figure 2.9. Biplot projection of 482 significant RDA loci and six bioclimatic variables into the adaptively enriched genetic space.

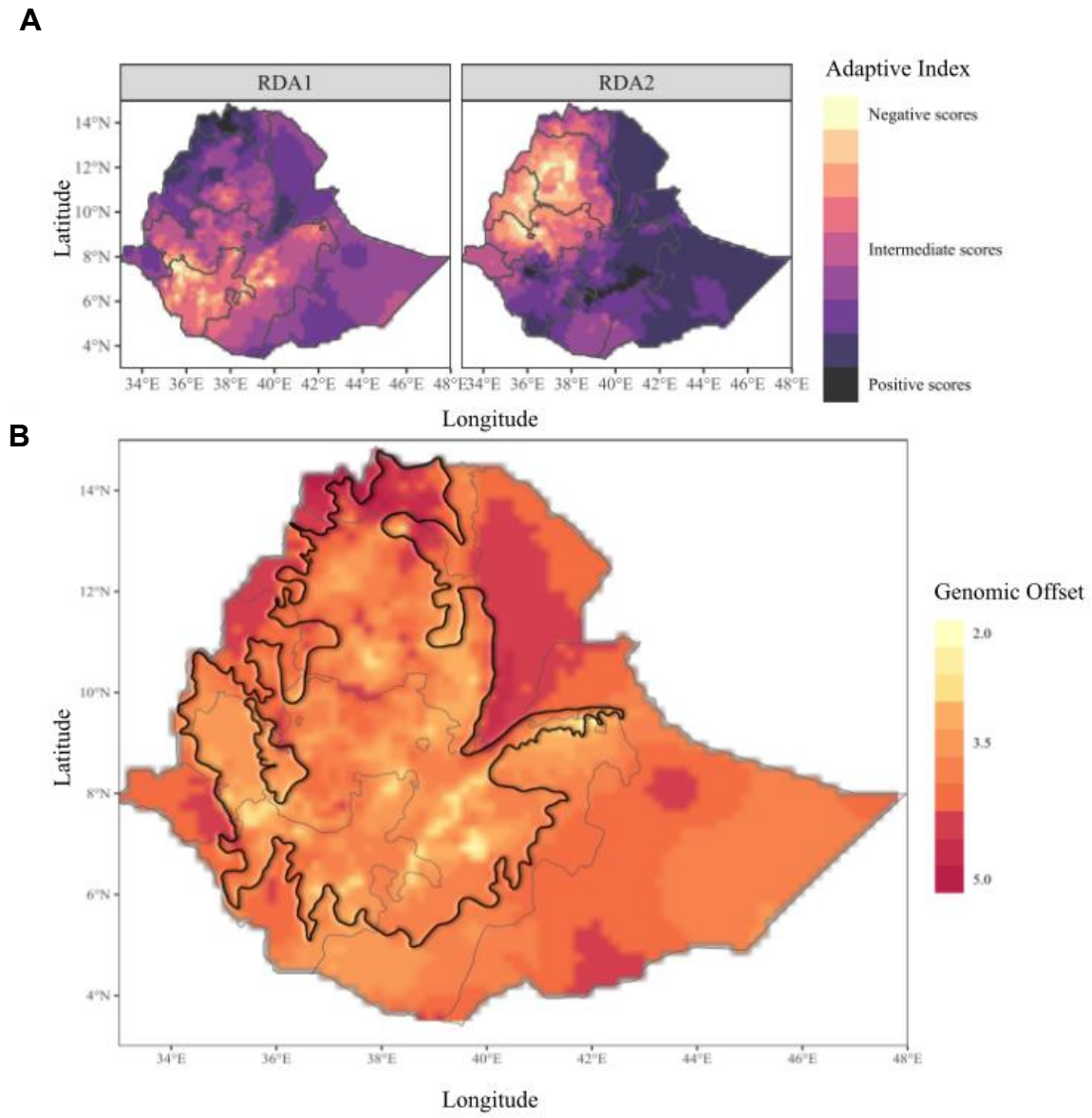

Figure 2.10. Current and predicted adaptive landscape of *E. tef* in Ethiopia. (**A**) Spatial extrapolation of RDA1 and RDA2 adaptive genetic gradients under current climatic conditions. (**B**) Average predicted genomic offset by the year 2080, estimated using the SSP2-4.5 climate scenario and the multi-model ensemble projection of 12 CMIP6 Global Circulation Models (refer to Table 2.2). The bolded black outline within the country perimeters highlights a region of current land suitability for *E. tef*, based on the spatial projection of the landscape metadata collected by and provided through the EIAR.

Discussion

In this study, we conducted a large-scale genome-environment association analysis of a previously unexamined, diverse collection of Ethiopian tef. We employed whole-genome

sequencing to test the hypothesis that traditional varieties of tef are locally adapted along an abiotic environmental gradient throughout its range of cultivation. Our analysis revealed that geographic distance had a minimal influence on existing genetic diversity. Using a combination of multivariate genotype-environment association methods and differentiation-based genome scans, we identified genetic variations exhibiting signatures of both diversifying and positive selection. Finally, we used these results to predict the spatial distribution of adaptive genetic variation across Ethiopia to highlight areas at risk of maladaptation under future climatic conditions.

Geographic distance often limits gene flow, and combined with local genetic drift, it leads to spatially structured differences in allele frequencies, resulting in a pattern of IBD (Charlesworth et al., 2003; Wright, 1943). Our analysis revealed five distinct genetic clusters among the sampled individuals, with the first axis explaining nearly half of the genetic variation. However, genetic diversity did not reflect geographically-based population structure (Figure 2.1). Traditional tef varieties were sampled from across five regions (Amhara, Tigray, Benishangul-Gumuz, Oromia, and the Southern Nations Nationalities and Peoples), with representatives from all regions appearing in each DAPC cluster, except for Benishangul-Gumuz, which was represented by a single sample. Although PC1 and PC2 explained approximately 74% of the genetic variation in the DAPC by maximizing group separation, they accounted for only 11.5% in the RDA when used as proxies for population structure (Table 2.3). This discrepancy highlights that, while these components are effective for distinguishing genetic clusters, their explanatory power for overall genetic variation is limited when assessed as predictors in the RDA.

Furthermore, spatial and environmental autocorrelation analyses showed no evidence of IBD or IBE, as $\overline{F}_{ST}$ values did not strongly correlate with geographic distance (Mantel r = -0.006, p = 0.505) and or environmental distance (Mantel r = 0.162, p = 0.071) (Figure 2.1C). These findings were consistent with previous studies documenting the extensive seed trade and genetic admixture within informal seed systems among Ethiopian small-holder farmers (Abebe et al., 2013; Ferede, 2011; Girma et al., 2018; Negassa, 1985). Although socio-cultural factors, such as kinship systems and ethnolinguistic groups (Labeyrie et al., 2016; Samberg et al., 2013), and environmental factors, such as elevation (Woldeyohannes et al., 2020), have been identified as influential in seed-mediated gene flow, our data does not allow us to separate the effects of spatial and environmental from social intervention. Woldeyohannes et al. (2022) showed that small-holder farmers consistently preferred the same tef ideotype regardless of gender, genetic background, or locality, favoring high-yielding, high-biomass, and fast-maturing varieties. Further research is warranted to achieve a comprehensive understanding of the role of social processes in shaping spatial patterns of genetic diversity in tef.

Outlier detection methods are commonly used to identify genomic regions under positive selection, providing insights into the genetic basis of adaptation, as previously reviewed by Biwas & Akey (2006) and Saravanan et al. (2021). In our study, we applied three outlier detection methods to investigate adaptive genetic variation in Ethiopian tef, focusing primarily on the results of RDA. RDA, a multivariate genotype-environment association approach, is particularly effective in identifying multilocus architectures associated with environmental variables across diverse landscapes (Bourret et al., 2013, 2014; Forester et al., 2018; Hoban et al., 2016; Lasky et al., 2012; Legendre & Legendre, 2012; Rellstab et al., 2015). Additionally, we conducted genome scans using iHH12 statistics to detect recent selection signatures by

identifying high-frequency haplotypes with an extended linkage disequilibrium. This method is particularly suited for domesticated species, where artificial selection pressures are stronger than those in natural populations. Finally, we estimated the genetic differentiation along environmental gradients using $\overline{F}_{ST}$, a measure of inter-population differences in allele frequencies, which is more suitable for detecting ancestral selective events (Maiorano et al., 2018). These diverse methods collectively provide a comprehensive toolkit for investigating genomic signatures of adaptation across various temporal and spatial scales.

Partial RDA models indicated a small proportion of shared variance between climate and population structure, with only 0.18% of genetic variation being explained by this overlap (Figure 2.3A; Table 2.3). These models have been employed to estimate the relative contributions of climate and space to explain genetic variation in species such as Arabidopsis (Lasky et al., 2012), *Eucalyptus globulus* (Labill.) (Butler et al., 2022), and barley (Abebe et al., 2015). Correcting for population structure can eliminate statistical signals from both factors, which leads to false negatives. However, neglecting population structure can increase the rate of false positives if the neutral structure follows selective gradients (Excoffier et al., 2009). Forester et al. (2018) evaluated the statistical power of detecting selective events through simulation and empirical data and identified an optimal balance between low false positive and high true positive rates across all selection levels. They found that false positive rates, which were already minimal for RDA, increased only slightly when accounting for population structure. These findings were consistent across demographic histories, sampling designs, sample sizes, and population structures. Therefore, we decided to not apply population structure-based corrections for RDA due to low spatial structure, as IBD often serves as the main driver (or a reliable proxy) of intraspecific genetic structure (Lasky et al., 2015) (Figure 2.1C; Table 2.1).

Climate variables in the RDA model accounted for approximately 0.63% of the genetic variation, identifying 482 climate-associated loci, with 145 of these loci near or within the *E. tef* v3 gene sequences. Though the observed proportion of the genetic variation explained by climate variables in our sample population was relatively low, this result aligns with the expectation that most loci are likely neutral and unaffected by environmental predictors, consistent with the neutral theory of molecular evolution. This theory posits that most genetic variation within a species is neutral or nearly neutral with respect to selection (Kimura, 1983). Therefore, only a small subset of loci would be expected to show significant relationships with environmental factors, reflecting adaptive genetic variation. Further analysis showed that 46 of these loci were situated within the coding sequence regions, with 31 inducing non-synonymous sequence changes. Interestingly, the majority of these non-synonymous sites (23 of the 31; 74.19%) were identified in the B subgenome. Variance partitioning among the bioclimatic variables revealed temperature annual range (bio7) and precipitation of the wettest month (bio13) as the main climate drivers of the detected genetic variation (Figure 2.3).

While the three outlier detection methods used in our study showed limited genome-wide overlap, their integrated results provided valuable insights into the genetic basis of local adaptation in our sample population of Ethiopian tef. As expected, our dataset showed an inverse relationship between genome-wide putative sweep regions and environment-associated divergent alleles, consistent with local adaptation, which maintains allele frequencies at an intermediate level (Figure 2.7; Price et al., 2018). $\overline{F}_{ST}$ analysis further revealed distinct fixation patterns across environmental predictors, with two high $\overline{F}_{ST}$ genomic regions on chromosome 5A colocalized between three variables: annual mean temperature (bio1), precipitation of the warmest quarter (bio18), and altitude (Figure 2.5). The climate-associated loci identified through RDA did not

overlap with the genomic regions highlighted by the other two outlier detection methods (Figure 2.6A). Notably, the relationship between RDA-identified loci and putative sweep regions mirrored a pattern similar to that of $\overline{F}_{ST}$, suggesting that these climate-associated regions may be subject to different selective pressures and are likely unrelated to recent selective events in the sample population.

GO enrichment analysis of *E. tef* v3 candidate genes revealed a diverse array of biological processes implicated in tef adaptation to environmental variables. The top 19 enriched GO terms revealed five distinct functional categories, involving 38 of the 92 candidate genes. Among these gene sets, the GO terms related to responses to stress and regulation of protein modification were the most significant (FDR $p \leq 0.05$), particularly in relation to the variables of precipitation of the wettest month (bio13) and temperature annual range (bio7) (Figure 2.8). Of the 31 non-synonymous sites identified, we highlighted one locus on chromosome 2A (Et_2A_015154) that demonstrated a positive relationship with temperature annual range (bio7) and was proximal to a putative sweep region (Figure 2.6). Our finding that XP_034591805 (mitogen-activated protein kinase kinase kinase 1) was the closest homolog to Et_2A_015154 was notable for several reasons. MAPKs are crucial in plant responses to abiotic stress such as cold, salt, heat, and UV radiation. In addition, they serve as convergence points for transmitting environmental stress signals to downstream targets. They have been extensively studied in different plant species, including Arabidopsis, rice, tomato [*Lycopersicon esculentum* (Miller.)], and cotton [*Gossypium hirsutum* (L.)], showing cross-talk with other signaling pathways (Sinha et al., 2011).

The distribution of adaptive genetic variation across a landscape shapes how species respond to climate change (Aitken et al., 2008; Capblancq et al., 2020; Capblancq & Forester,

2021; Jump & Peñuelas, 2005; Steane et al., 2014). As the climate shifts, populations adapted to current local climates may struggle to achieve optimal fitness levels, creating a mismatch between existing genotypes and the new environment (Aitken et al., 2008; Capblancq et al., 2020). This mismatch represents an 'adaptation debt', which requires changes in local genotype frequencies to minimize maladaptation in the new fitness landscape. In our study, the greatest genomic-adaptive offset between the current and future climate scenarios was predicted in the Tigrayan region of tef cultivation (Figure 2.10B), which has historically experienced relatively higher climatic variability (Figure 2.9; Figure 2.10A). Further analysis of the candidate gene Et_2A_015154 revealed a significant association ($p < 0.05$) between the alternate allele and latitude, with the alternate allele increasing in frequency within 12–14°N. These findings suggest a connection between the biological role of MAPKs in stress signaling and temperature-associated variation at Et_2A_015154, particularly at higher latitudes, indicating that positive selection for the alternate allele reflects the indirect selection of Ethiopian tef with increased environmental plasticity in northern Ethiopia.

We propose that future controlled experiments to focus on the functional validation of candidate genes (e.g., common gardens and experiments in silico) to provide empirical evidence of the effect size of these genetic variants on plant performance and fitness across climatic gradients. Specifically, the next critical step in confirming the role of these genes involves developing and testing near-isogenic lines (NILs) or creating genetically engineered lines that isolate the specific variants. Conducting fitness trade-off experiments in contrasting environments, ideally in field conditions representative of the environmental gradient with which the allelic variation at those sites is strongly associated, would provide further evidence to

determine whether these alleles contribute to adaptation. This process is essential to establish the validity of RDA findings and to effectively utilize these markers for tef breeding.

While this study offers valuable insights, several potential limitations warrant consideration. First, the assumption of linearity in RDA may overlook the non-linear relationships between environmental gradients and adaptive genetic variation. Second, genomic prediction of climate maladaptation relies on statistical-based associations and is contingent on several key assumptions, including the correct identification and inclusion of the adaptive genetic component in the prediction model, the proportional relationship between genetic offset and expected decrease in fitness, and the reflection of the genotype × environment interaction in the current optima between standing genetic variation and prevailing climatic conditions (Capblancq et al., 2020). Finally, tef's predominantly self-pollinating mating system (0.1% to 1% outcrossing) presents significant implications for adaptation and allele introgression (Berhe et al., 2000). In comparison to outcrossing species, self-pollinating systems, such as tef, exhibit increased linkage disequilibrium, reduced effective recombination rates, and lower genetic diversity, which may constrain the genetic response to selection and adaptation (Charlesworth, 2003). These characteristics may diminish the efficacy of adaptive allele testing and introgression in germplasm development.

Tef, along with other orphan crops, plays a key role in the subsistence of local communities and presents promising solutions to food insecurity and climate change vulnerability. However, despite their advantages, orphan crops have lagged behind major crops in terms of improvement efforts, largely due to limited research investment and limited selective pressures by subsistence farmers, resulting in undesirable agronomic traits reminiscent of wild relatives (Lemmon et al., 2018). With 95% of the crop area in Ethiopia being rainfed, limited

water availability coupled with rising temperatures poses a significant concern regarding its projected impact on crop development and production, especially in lowland regions (Administration, 2024). Leveraging indigenous germplasm, alongside advanced breeding techniques like genome editing, presents opportunities for broadening genetic variation and developing improved tef varieties that can withstand future environmental challenges (Assefa et al., 2011; Beyene et al., 2022).

REFERENCES

Abebe, Tiegist D, Naz, A. A., & Léon, J. (2015). Landscape genomics reveal signatures of local adaptation in barley (Hordeum vulgare L.). *Frontiers in Plant Science*, 6(10), 813. https://doi.org/10.3389/fpls.2015.00813

Abebe, Tiegist Dejene, Mathew, B., & Léon, J. (2013). Barrier analysis detected genetic discontinuity among Ethiopian barley (Hordeum vulgare L.) landraces due to landscape and human mobility on gene flow. *Genetic Resources and Crop Evolution*, 60(1), 297–309. https://doi.org/10.1007/s10722-012-9834-6

Administration, I. T. (2024). Ethiopia - Country Commercial Guide. Agricultural Sectors. https://www.trade.gov/country-commercial-guides/ethiopia-agricultural-sectors

Agren, J., Oakley, C. G., McKay, J. K., Lovell, J. T., & Schemske, D. W. (2013). Genetic mapping of adaptation reveals fitness tradeoffs in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, 110(52), 21077–21082. https://doi.org/10.1073/pnas.1316773110

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, 1(1), 95–111. https://doi.org/10.1111/j.1752-4571.2007.00013.x

Assefa, K., Ketema, S., Tefera, H., Kefyalew, T., & Chundera, F. (2000). Trait diversity, heritability and genetic advance in selected germplasm lines of tef [Eragrostis tef (Zucc.) Trotter]. *Hereditas*, 133(1), 29–37. https://doi.org/10.1111/j.1601-5223.2000.t01-1-00029.x

Assefa, K., Tefera, H., & Merker, A. (2002). Variation and inter-relationships of quantitative traits in tef (Eragrostis tef (Zucc.) Trotter) germplasm from western and southern Ethiopia. *Hereditas*, 136(2), 116–125. https://doi.org/10.1034/j.1601-5223.2002.1360205.x

Assefa, K., Yu, J. K., Zeid, M., Belay, G., Tefera, H., & Sorrells, M. E. (2011). Breeding tef [Eragrostis tef (Zucc.) trotter]: Conventional and molecular approaches. *Plant Breeding*, 130(1), 1–9. https://doi.org/10.1111/j.1439-0523.2010.01782.x

Bay, R. A., Harrigan, R. J., Underwood, V. Le, Gibbs, H. L., Smith, T. B., & Ruegg, K. (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science*, 359(6371), 83–86. https://doi.org/10.1126/science.aan4380

Belay, M., & Assen, M. (2023). 7 - Land suitability evaluation for cereals and coffee in Cheleleka Bobe (Wolonkomi area), Ethiopia. In S. K. Meena, A. D. O. Ferreira, V. S. Meena, A. Rakshit, R. P. Shrestha, C. S. Rao, & K. H. M. Siddique (Eds.), *Agricultural*

Soil Sustainability and Carbon Management (pp. 181–206). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-323-95911-7.00006-2

Berhe, T., Nelson, L.A., Morris, M.R., Schmidt, J.W. The genetics of qualitative traits in tef. In Proceedings of the International Workshop on Tef Genetics and Improvement, Debre Zeit, Ethiopia, 16–19 October 2000; pp. 79–85.

Beyene, G., Chauhan, R. D., Villmer, J., Husic, N., Wang, N., Gebre, E., Girma, D., Chanyalew, S., Assefa, K., Tabor, G., Gehan, M., McGrone, M., Yang, M., Lenderts, B., Schwartz, C., Gao, H., Gordon-Kamm, W., Taylor, N. J., & MacKenzie, D. J. (2022). CRISPR/Cas9-mediated tetra-allelic mutation of the 'Green Revolution' SEMIDWARF-1 (SD-1) gene confers lodging resistance in tef (Eragrostis tef). *Plant Biotechnology Journal*, 20(9), 1716–1729. https://doi.org/10.1111/pbi.13842

Bischoff, F., Molendijk, A., Rajendrakumar, C. S., & Palme, K. (1999). GTP-binding proteins in plants. *Cellular and Molecular Life Sciences* : CMLS, 55(2), 233–256. https://doi.org/10.1007/s000180050287

Biswas, S., & Akey, J. M. (2006). Genomic insights into positive selection. *Trends in Genetics*, 22(8), 437–446. https://doi.org/10.1016/j.tig.2006.06.005

Blaser, N. (2020). rdist: Calculate Pairwise Distances. https://cran.r-project.org/package=rdist

Bolger, A., Scossa, F., Bolger, M. E., Lanz, C., Maumus, F., Tohge, T., Quesneville, H., Alseekh, S., Sørensen, I., Lichtenstein, G., Fich, E. A., Conte, M., Keller, H., Schneeberger, K., Schwacke, R., Ofner, I., Vrebalov, J., Xu, Y., Osorio, S., … Fernie, A. R. (2014). The genome of the stress-tolerant wild tomato species Solanum pennellii. *Nature Genetics*, 46(9), 1034–1038. https://doi.org/10.1038/ng.3046

Bourret, V., Dionne, M., & Bernatchez, L. (2014). Detecting genotypic changes associated with selective mortality at sea in Atlantic salmon: polygenic multilocus analysis surpasses genome scan. *Molecular Ecology*, 23(18), 4444–4457. https://doi.org/10.1111/mec.12798

Bourret, V., Dionne, M., Kent, M. P., Lien, S., & Bernatchez, L. (2013). Landscape genomics in Atlantic salmon (Salmo salar): searching for gene-environment interactions driving local adaptation. *Evolution*; *International Journal of Organic Evolution*, 67(12), 3469–3487. https://doi.org/10.1111/evo.12139

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

Butler, J. B., Harrison, P. A., Vaillancourt, R. E., Steane, D. A., Tibbits, J. F. G., & Potts, B. M. (2022). Climate Adaptation, Drought Susceptibility, and Genomic-Informed Predictions of Future Climate Refugia for the Australian Forest Tree Eucalyptus globulus. *Forests*, 13(4), 1–25. https://doi.org/10.3390/f13040575

Cannarozzi, G., Plaza-Wüthrich, S., Esfeld, K., Larti, S., Wilson, Y. S., Girma, D., de Castro, E., Chanyalew, S., Blösch, R., Farinelli, L., Lyons, E., Schneider, M., Falquet, L., Kuhlemeier, C., Assefa, K., & Tadele, Z. (2014). Genome and transcriptome sequencing identifies breeding targets in the orphan crop tef (Eragrostis tef). *BMC Genomics*, 15(1). https://doi.org/10.1186/1471-2164-15-581

Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M., & Keller, S. R. (2020). Genomic Prediction of (Mal)Adaptation across Current and Future Climatic Landscapes. *Annual Review of Ecology, Evolution, and Systematics*, 51, 245–269. https://doi.org/10.1146/annurev-ecolsys-020720-042553

Capblancq, T., & Forester, B. R. (2021). Redundancy analysis: A Swiss Army Knife for landscape genomics. *Methods in Ecology and Evolution*, 12(12), 2298–2309. https://doi.org/10.1111/2041-210X.13722

Carlson, C. S., Thomas, D. J., Eberle, M. A., Swanson, J. E., Livingston, R. J., Rieder, M. J., & Nickerson, D. A. (2005). Genomic regions exhibiting positive selection identified from dense genotype data. *Genome Research*, 15(11), 1553–1565. https://doi.org/10.1101/gr.4326505

Charlesworth, B., Charlesworth, D., & Barton, N. H. (2003). The Effects of Genetic and Geographic Structure on Neutral Variation. Annual Review of Ecology, *Evolution, and Systematics*, 34(1), 99–125. https://doi.org/10.1146/annurev.ecolsys.34.011802.132359

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chua, K. Y., Doyle, C. R., Simpson, R. J., Turner, K. J., Stewart, G. A., & Thomas, W. R. (1990). Isolation of cDNA coding for the major mite allergen Der p II by IgE plaque immunoassay. *International Archives of Allergy and Applied Immunology*, 91(2), 118–123. https://doi.org/10.1159/000235101

Colcombet, J., & Hirt, H. (2008). Arabidopsis MAPKs: a complex signalling network involved in multiple biological processes. *The Biochemical Journal*, 413(2), 217–226. https://doi.org/10.1042/BJ20080625

Coop, G., Pickrell, J. K., Novembre, J., Kudaravalli, S., Li, J., Absher, D., Myers, R. M., Cavalli-Sforza, L. L., Feldman, M. W., & Pritchard, J. K. (2009). The role of geography in human adaptation. *PLoS Genetics*, 5(6). https://doi.org/10.1371/journal.pgen.1000500

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & Group, 1000 Genomes Project Analysis. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1), 27–46. https://doi.org/https://doi.org/10.1111/j.1600-0587.2012.07348.x

Eckert, A. J., Bower, A. D., GonzÁlez-MartÍnez, S. C., Wegrzyn, J. L., Coop, G., & Neale, D. B. (2010a). Back to nature: Ecological genomics of loblolly pine (Pinus taeda, Pinaceae). *Molecular Ecology*, 19(17), 3789–3805. https://doi.org/10.1111/j.1365-294X.2010.04698.x

Eckert, A. J., Van Heerwaarden, J., Wegrzyn, J. L., Nelson, C. D., Ross-Ibarra, J., González-Martínez, S. C., & Neale, D. B. (2010b). Patterns of population structure and environmental associations to aridity across the range of loblolly pine (Pinus taeda L., Pinaceae). *Genetics*, 185(3), 969–982. https://doi.org/10.1534/genetics.110.115543

Excoffier, L, Hofer, T., & Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity*, 103(4), 285–298. https://doi.org/10.1038/hdy.2009.74

Excoffier, Laurent, & Ray, N. (2008). Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution*, 23(7), 347–351. https://doi.org/10.1016/j.tree.2008.04.004

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Ferede, S. (2011). Tef seed system in smallholder agriculture: A case study in selected major tef growing districts of Ethiopia.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315. https://doi.org/https://doi.org/10.1002/joc.5086

Fitzpatrick, M. C., & Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, 18(1), 1–16. https://doi.org/10.1111/ele.12376

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology*, 27(9), 2215–2233. https://doi.org/10.1111/mec.14584

Fournier-Level, A., Korte, A., Cooper, M. D., Nordborg, M., Schmitt, J., & Wilczek, A. M. (2011). A map of local adaptation in Arabidopsis thaliana. *Science*, 334(6052), 86–89. https://doi.org/10.1126/science.1209271

Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D. M., Burzynski-Chang, E. A., Fish, T. L., Stromberg, K. A., Sacks, G. L., Thannhauser, T. W., Foolad, M. R., Diez, M. J., Blanca, J., Canizares, J., Xu, Y., van der Knaap, E., Huang, S., Klee, H. J., … Fei, Z. (2019). The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nature Genetics*., 51(6), 1044–1051.

Garud, N. R., Messer, P. W., Buzbas, E. O., & Petrov, D. A. (2015). Recent selective sweeps in North American Drosophila melanogaster show signatures of soft sweeps. *PLoS Genetics*, 11(2), e1005004. https://doi.org/10.1371/journal.pgen.1005004

Girma, D., Cannarozzi, G., Weichert, A., & Tadele, Z. (2018). Genotyping by sequencing reasserts the close relationship between tef and its putative wild eragrostis progenitors. *Diversity*, 10(2). https://doi.org/10.3390/d10020017

Goudet, J., & Jombart, T. (2022). hierfstat: Estimation and Tests of Hierarchical F-Statistics. https://cran.r-project.org/package=hierfstat

Hancock, A. M., Jarymowycz, L. B., Sperone, F. G., Toomajian, C., Roux, F., & Bergelson, J. (2011). *Arabidopsis thaliana Genome*. 83(10), 83–87.

Hijmans, R. J. (2023). raster: Geographic Data Analysis and Modeling. https://cran.r-project.org/package=raster

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M. L., Reed, L. K., Storfer, A., & Whitlock, M. C. (2016). Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *The American Naturalist*, 188(4), 379–397. https://doi.org/10.1086/688018

Hohenlohe, P. A., Phillips, P. C., & Cresko, W. A. (2010). Using population genomics to detect selection in natural populations: key concepts and methodological considerations. *International Journal of Plant Sciences*, 171(9), 1059–1071. https://doi.org/10.1086/656306

Huang, J., Qin, D., Jiang, T., Wang, Y., Feng, Z., Zhai, J., Cao, L., Chao, Q., Xu, X., Wang, G., & Su, B. (2019). Effect of Fertility Policy Changes on the Population Structure and Economy of China: From the Perspective of the Shared Socioeconomic Pathways. *Earth's Future*, 7(3), 250–265. https://doi.org/https://doi.org/10.1029/2018EF000964

Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., Zhu, C., Lu, T., Zhang, Z., Li, M., Fan, D., Guo, Y., Wang, A., Wang, L., Deng, L., Li, W., Lu, Y., Weng, Q., … Han, B. (2010). Genome-wide asociation studies of 14 agronomic traits in rice landraces. *Nature Genetics*, 42(11), 961–967. https://doi.org/10.1038/ng.695

Hufford, M. B., Xu, X., Van Heerwaarden, J., Pyhäjärvi, T., Chia, J. M., Cartwright, R. A., Elshire, R. J., Glaubitz, J. C., Guill, K. E., Kaeppler, S. M., Lai, J., Morrell, P. L., Shannon, L. M., Song, C., Springer, N. M., Swanson-Wagner, R. A., Tiffin, P., Wang, J., Zhang, G., … Ross-Ibarra, J. (2012). Comparative population genomics of maize domestication and improvement. *Nature Genetics*, 44(7), 808–811. https://doi.org/10.1038/ng.2309

IPCC, I. P. on C. C. (2014). Climate Change 2014 – Impacts, Adaptation and Vulnerability: Part B: Regional Aspects: Working Group II Contribution to the IPCC Fifth Assessment Report: Volume 2: Regional Aspects. Cambridge University Press. https://doi.org/10.1017/CBO9781107415386

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics, 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics. https://doi.org/10.1093/bioinformatics/btr521

Josephs, E. B., Berg, J. J., Ross-Ibarra, J., & Coop, G. (2019). Detecting adaptive differentiation in structured populations with genomic data and common gardens. *Genetics*, 211(3), 989–1004. https://doi.org/10.1534/genetics.118.301786

Jump, A. S., & Peñuelas, J. (2005). Running to stand still: adaptation and the response of plants to rapid climate change. *Ecology Letters*, 8(9), 1010–1020. https://doi.org/10.1111/j.1461-0248.2005.00796.x

Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., & Eskin, E. (2008). Efficient control of population structure in model organism association mapping. *Genetics*, 178(3), 1709–1723. https://doi.org/10.1534/genetics.107.080101

Kassambara, A., & Mundt, F. (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. https://cran.r-project.org/package=factoextra

Kefyalew, T., Tefera, H., Assefa, K., & Ayele, M. (2000). Phenotypic diversity for qualitative and phenologic characters in germplasm collections of tef (Eragrostis tef). *Genetic Resources and Crop Evolution*, 47(1), 73–80. https://doi.org/10.1023/A:1008781022038

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press.

Knaus, B. J., & Grünwald, N. J. (2017). {VCFR}: a package to manipulate and visualize variant call format data in {R}. Molecular Ecology Resources, 17(1), 44–53. https://dx.doi.org/10.1111/1755-0998.12549

Labeyrie, V., Thomas, M., Muthamia, Z. K., & Leclerc, C. (2016). Seed exchange networks, ethnicity, and sorghum diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 113(1), 98–103. https://doi.org/10.1073/pnas.1513238112

Lasky, J. R., Des Marais, D. L., McKay, J. K., Richards, J. H., Juenger, T. E., & Keitt, T. H. (2012). Characterizing genomic variation of Arabidopsis thaliana: The roles of geography and climate. *Molecular Ecology*, 21(22), 5512–5529. https://doi.org/10.1111/j.1365-294X.2012.05709.x

Lasky, J. R., Upadhyaya, H. D., Ramu, P., Deshpande, S., Hash, C. T., Bonnette, J., Juenger, T. E., Hyma, K., Acharya, C., Mitchell, S. E., Buckler, E. S., Brenton, Z., Kresovich, S., & Morris, G. P. (2015). Genome-environment associations in sorghum landraces predict adaptive traits. *Science Advances*, 1(6), 1–14. https://doi.org/10.1126/sciadv.1400218

Lê, S., Josse, J., & Husson, F. (2008). {FactoMineR}: A Package for Multivariate Analysis. Journal of Statistical Software, 25(1), 1–18. https://doi.org/10.18637/jss.v025.i01

Legendre, P., & Legendre, L. (2012). Numerical Ecology. *Elsevier*.

Lemmon, Z. H., Reem, N. T., Dalrymple, J., Soyk, S., Swartwood, K. E., Rodriguez-Leal, D., Van Eck, J., & Lippman, Z. B. (2018). Rapid improvement of domestication traits in an orphan crop by genome editing. Nature Plants, 4(10), 766–770. https://doi.org/10.1038/s41477-018-0259-x

Lewontin, R. C., & Hubby, J. L. (1966). A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of Drosophila pseudoobscura. *Genetics*, 54(2), 595–609. https://doi.org/10.1093/genetics/54.2.595

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 00(00), 1–3. http://arxiv.org/abs/1303.3997

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, Yangping, Colleoni, C., Zhang, J., Liang, Q., Hu, Y., Ruess, H., Simon, R., Liu, Y., Liu, H., Yu, G., Schmitt, E., Ponitzki, C., Liu, G., Huang, H., Zhan, F., Chen, L., Huang, Y., Spooner, D., & Huang, B. (2018). Genomic Analyses Yield Markers for Identifying Agronomically Important Genes in Potato. *Molecular Plant*, 11(3), 473–484. https://doi.org/10.1016/j.molp.2018.01.009

Li, Yong, Cao, K., Li, N., Zhu, G., Fang, W., Chen, C., Wang, X., Guo, J., Wang, Q., Ding, T., Wang, J., Guan, L., Wang, J., Liu, K., Guo, W., Arús, P., Huang, S., Fei, Z., & Wang, L. (2021). Genomic analyses provide insights into peach local adaptation and responses to climate change. *Genome Research*, 31(4), 592–606. https://doi.org/10.1101/GR.261032.120

Liu, X., Zhang, X., Sun, B., Hao, L., Liu, C., Zhang, D., Tang, H., Li, C., Li, Y., Shi, Y., Xie, X., Song, Y., Wang, T., & Li, Y. (2019). Genome-wide identification and comparative analysis of drought-related microRNAs in two maize inbred lines with contrasting drought tolerance by deep sequencing. *PLoS ONE*, 14(7), 449–464. https://doi.org/10.1371/journal.pone.0219176

Liu, Y., Zhou, M., Gao, Z., Ren, W., Yang, F., He, H., & Zhao, J. (2015). RNA-Seq Analysis Reveals MAPKKK Family Members Related to Drought Tolerance in Maize. *PloS One*, 10(11), e0143128. https://doi.org/10.1371/journal.pone.0143128

Lu, M., Loopstra, C. A., & Krutovsky, K. V. (2019). Detecting the genetic basis of local adaptation in loblolly pine (Pinus taeda L.) using whole exome-wide genotyping and an integrative landscape genomics analysis approach. *Ecology and Evolution*, 9(12), 6798–6809. https://doi.org/10.1002/ece3.5225

Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, 4(12), 981–994. https://doi.org/10.1038/nrg1226

Maynard Smith, J., & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetics Research*, *23*, 23–35. https://doi.org/10.1017/S0016672308009579

Maiorano, A. M., Lourenco, D. L., Tsuruta, S., Toro Ospina, A. M., Stafuzza, N. B., Masuda, Y., Filho, A. E. V., Dos Santos Goncalves Cyrillo, J. N., Curi, R. A., & De Vasconcelos Silva, J. A. (2018). Assessing genetic architecture and signatures of selection of dual purpose Gir cattle populations using genomic information. *PLoS ONE*, 13(8), 1–24. https://doi.org/10.1371/journal.pone.020069423

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *GENOME RESEARCH*, 20(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Miller, A. J., & Gross, B. L. (2011). From forest to field: Perennial fruit crop domestication. *American Journal of Botany*, 98(9), 1389–1414. https://doi.org/10.3732/ajb.1000522

Ministry of Agriculture and Livestock Resources. (2019). Plant Variety Release, Protection and Seed Quality Control Directorate. *Crop Variety Register* (22).

Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., Riera-Lizarazu, O., Brown, P. J., Acharya, C. B., Mitchell, S. E., Harriman, J., Glaubitz, J. C., Buckler, E. S., & Kresovich, S. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences of the United States of America*, 110(2), 453–458. https://doi.org/10.1073/pnas.1215985110

Moyers, B. T., Morrell, P. L., & McKay, J. K. (2018). Genetic costs of domestication and improvement. *Journal of Heredity*, 109(2), 103–116. https://doi.org/10.1093/jhered/esx069

Negassa, M. (1985). Patterns of phenotypic diversity in an Ethiopian barley collection, and the Arussi-Bale Highland as a center of origin of barley. *Hereditas*, 102(1), 139–150. https://doi.org/10.1111/j.1601-5223.1985.tb00474.x

Nychka, D., Furrer, R., Paige, J., & Sain, S. (2021). fields: Tools for spatial data. https://github.com/dnychka/fieldsRPackage

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., & Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, 9(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., … Weedon, J. (2022). vegan: Community Ecology Package. https://cran.r-project.org/package=vegan

Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T., Carneiro, M. O., der Auwera, G. A. Van, Kling, D. E., Gauthier, L. D., Levy-Moonshine, A., Roazen, D., Shakir, K., Thibault, J., Chandran, S., Whelan, C. W., Lek, M., Gabriel, S., Daly, M. J., Neale, B., MacArthur, D. G., & Banks, E. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv*. https://api.semanticscholar.org/CorpusID:89880024

Price, N., Moyers, B. T., Lopez, L., Lasky, J. R., Monroe, J. G., Mullen, J. L., Oakley, C. G., Lin, J., Ågren, J., Schrider, D. R., Kern, A. D., McKay, J. K., Grey Monroe, J., Mullen, J. L., Oakley, C. G., Lin, J., Ågren, J., Schrider, D. R., Kern, A. D., & McKay, J. K. (2018). Combining population genomics and fitness QTLs to identify the genetics of local adaptation in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(19), 5028–5033. https://doi.org/10.1073/pnas.1719998115

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*. 24(17), 4348–4370. John Wiley & Sons, Ltd. https://doi.org/10.1111/mec.13322

Rellstab, C., Zoller, S., Walthert, L., Lesur, I., Pluess, A. R., Graf, R., Bodénès, C., Sperisen, C., Kremer, A., & Gugerli, F. (2016). Signatures of local adaptation in candidate genes of oaks (Quercus spp.) with respect to present and future climatic conditions. *Molecular Ecology*, 25(23), 5907–5924. https://doi.org/10.1111/mec.13889

Revelle, W. (2023). psych: Procedures for Psychological, Psychometric, and Personality Research. https://cran.r-project.org/package=psych

Rodriguez, M. C. S., Petersen, M., & Mundy, J. (2010). Mitogen-activated protein kinase signaling in plants. *Annual Review of Plant Biology*, 61, 621–649. https://doi.org/10.1146/annurev-arplant-042809-112252

Šamajová, O., Komis, G., & Šamaj, J. (2013). Emerging topics in the cell biology of mitogen-activated protein kinases. *Trends in Plant Science*, 18(3), 140–148. https://doi.org/10.1016/j.tplants.2012.11.004

Samberg, L. H., Shennan, C., & Zavaleta, E. (2013). Farmer Seed Exchange and Crop Diversity in a Changing Agricultural Landscape in the Southern Highlands of Ethiopia. *Human Ecology*, 41(3), 477–485. https://doi.org/10.1007/s10745-013-9579-7

Saravanan, K. A., Panigrahi, M., Kumar, H., Parida, S., Bhushan, B., Gaur, G. K., Dutt, T., Mishra, B. P., & Singh, R. K. (2021). Genomic scans for selection signatures revealed candidate genes for adaptation and production traits in a variety of cattle breeds. *Genomics*, 113(3), 955–963. https://doi.org/10.1016/j.ygeno.2021.02.009

Sinha, A. K., Jaggi, M., Raghuram, B., & Tuteja, N. (2011). Mitogen-activated protein kinase signaling in plants under abiotic stress. *Plant Signaling & Behavior*, 6(2), 196–203. https://doi.org/10.4161/psb.6.2.14701

Smékalová, V., Doskočilová, A., Komis, G., & Samaj, J. (2014). Crosstalk between secondary messengers, hormones and MAPK modules during abiotic stress signalling in plants. *Biotechnology Advances*, 32(1), 2–11. https://doi.org/10.1016/j.biotechadv.2013.07.009

Steane, D. A., Potts, B. M., McLean, E., Prober, S. M., Stock, W. D., Vaillancourt, R. E., & Byrne, M. (2014). Genome-wide scans detect adaptation to aridity in a widespread forest tree species. *Molecular Ecology*, 23(10), 2500–2513. https://doi.org/10.1111/mec.12751

Storfer, A., Patton, A., & Fraik, A. K. (2018). Navigating the interface between landscape genetics and landscape genomics. *Frontiers in Genetics*. 9(3), 68. Frontiers Media S.A. https://doi.org/10.3389/fgene.2018.00068

Stouffer, R. J., Eyring, V., Meehl, G. A., Bony, S., Senior, C., Stevens, B., & Taylor, K. E. (2017). CMIP5 Scientific Gaps and Recommendations for CMIP6. *Bulletin of the American Meteorological Society*, 98(1), 95–105. https://doi.org/https://doi.org/10.1175/BAMS-D-15-00013.1

Szpiech, Z. A. (2021). selscan v2.0.0 User Manual.

Szpiech, Z. A., & Hernandez, R. D. (2014). selscan: An Efficient Multithreaded Program to Perform EHH-Based Scans for Positive Selection. *Molecular Biology and Evolution*, 31(10), 2824–2827. https://doi.org/10.1093/molbev/msu211

Tadesse, D. (1993). Study on genetic variation of landraces of teff (Eragrostis tef (Zucc.) Trotter) in Ethiopia. *Genetic Resources and Crop Evolution*, 40(2), 101–104. https://doi.org/10.1007/BF00052640

Torres, R., Szpiech, Z. A., & Hernandez, R. D. (2018). Human demographic history has amplified the effects of background selection across the genome. *PLoS Genetics*, 14(6), e1007387. https://doi.org/10.1371/journal.pgen.1007387

Turner, S. (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. The Journal of Open Source Software. https://doi.org/10.21105/joss.00731

Vadovič, P., Šamajová, O., Takáč, T., Novák, D., Zapletalová, V., Colcombet, J., & Šamaj, J. (2019). Biochemical and Genetic Interactions of Phospholipase D Alpha 1 and Mitogen-Activated Protein Kinase 3 Affect Arabidopsis Stress Response. *Frontiers in Plant Science*, 10, 275. https://doi.org/10.3389/fpls.2019.00275

van der Auwera, G., & O'Connor, B. D. (2020). Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly Media, Incorporated. https://books.google.com/books?id=wwiCswEACAAJ

VanBuren, R., Man Wai, C., Wang, X., Pardo, J., Yocca, A. E., Wang, H., Chaluvadi, S. R., Han, G., Bryant, D., Edger, P. P., Messing, J., Sorrells, M. E., Mockler, T. C., Bennetzen, J. L., & Michael, T. P. (2020). Exceptional subgenome stability and functional divergence in the allotetraploid Ethiopian cereal teff. *Nature Communications*, 11(1). https://doi.org/10.1038/s41467-020-14724-z

Vavilov, N. I. (1951). The Origin, Variation, Immunity and Breeding of Cultivated Plants (Translated by S.K. Chestitee). *Chronica Botonica*, 13, 1–366.

Voight, B. F., Kudaravalli, S., Wen, X., & Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLoS Biology*, 4(3), e72. https://doi.org/10.1371/journal.pbio.0040072

Wang, E. T., Kodama, G., Baldi, P., & Moyzis, R. K. (2006). Global landscape of recent inferred Darwinian selection for Homo sapiens. *Proceedings of the National Academy of Sciences of the United States of America*, 103(1), 135–140. https://doi.org/10.1073/pnas.0509691102

Wang, J., Hu, Z., Liao, X., Wang, Z., Li, W., Zhang, P., Cheng, H., Wang, Q., Bhat, J. A., Wang, H., Liu, B., Zhang, H., Huang, F., & Yu, D. (2022). Whole-genome resequencing reveals

signature of local adaptation and divergence in wild soybean. *Evolutionary Applications*, 15(11), 1820–1833. https://doi.org/10.1111/eva.13480

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org

Woldeyohannes, A. B., Accotto, C., Desta, E. A., Kidane, Y. G., Fadda, C., Pè, M. E., & Dell'Acqua, M. (2020). Current and projected eco-geographic adaptation and phenotypic diversity of Ethiopian teff (Eragrostis teff) across its cultivation range. *Agriculture, Ecosystems and Environment*, 300. https://doi.org/10.1016/j.agee.2020.107020

Wright, S. (1943). Isolation by Distance. *Genetics*, 28(2), 114–138. https://doi.org/10.1093/genetics/28.2.114

Wright, Sewall. (1949). The genetical structure of populations. *Annals of Eugenics*, 15(4), 323–354. https://doi.org/10.1111/j.1469-1809.1949.tb02451.x

Wright, S. I., Bi, I. V., Schroeder, S. G., Yamasaki, M., Doebley, J. F., McMullen, M. D., & Gaut, B. S. (2005). The Effects of Artificial Selection on the Maize Genome. *Science*, 308(5726), 1310–1314. https://doi.org/10.1126/science.1107891

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., & Yu, G. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, 2(3), 100141. https://doi.org/10.1016/j.xinn.2021.100141

Wyser, K., Kjellström, E., Koenigk, T., Martins, H., & Döscher, R. (2020). Warmer climate projections in EC-Earth3-Veg: The role of changes in the greenhouse gas concentrations from CMIP5 to CMIP6. *Environmental Research Letters*, 15(5). https://doi.org/10.1088/1748-9326/ab81c2

Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS: A Journal of Integrative Biology*, 16(5), 284–287. https://doi.org/10.1089/omi.2011.0118

Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., & Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, 38(2), 203–208. https://doi.org/10.1038/ng1702

Yu, Y., Beyene, G., Villmer, J., Duncan, K. E., Hu, H., Johnson, T., Doust, A. N., Taylor, N. J., & Kellogg, E. A. (2023). Grain shattering by cell death and fracture in Eragrostis tef. *Plant Physiology*, 192(1), 222–239. https://doi.org/10.1093/plphys/kiad079

Zuur, A., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol*, 314.

CHAPTER 3

PHENOME-TO-GENOME INSIGHTS FOR EVALUATING ROOT SYSTEM ARCHITECTURE IN FIELD STUDIES OF MAIZE

Summary

Identifying the genetic basis of root system architecture (RSA) in crops requires innovative approaches that enable high-throughput and precise phenotyping in field conditions. In this study, we evaluated multiple phenotyping and analytical frameworks for quantifying RSA in mature, field-grown maize in three field experiments. We used forward and reverse genetic approaches to evaluate > 1,700 root crowns sampled from a genetically diverse sample of maize including a GWAS population, a biparental mapping population, and mutants and wild-types at two known RSA genes, Deeper Rooting 1 and Rootless 1. Here, we demonstrate the utility of increasing the dimensionality of traditional 2D techniques, '2D multi-view' method, to improve trait heritability estimation and precision of genotype-to-phenotype mapping. Comparison of univariate and multivariate genome-wide association study (GWAS) approaches revealed that multivariate traits were effective at dissecting complex RSA phenotypes and identifying pleiotropic quantitative trait loci (QTL). Overall, 3D root models generated from X-ray computed tomography and digital phenotyping captured a larger proportion of RSA variations, as evidenced by genome-wide and single-gene analyses. However, we found that root pulling force (RPF) was a highly heritable estimate of RSA that demonstrated strong genotypic correlations and identified the largest number of shared QTL with 3D phenotypes. Our study provides further

evidence that integrating complementary phenotyping technologies and statistical frameworks can offer deeper insights into the genetic architecture of the global RSA in field-grown maize.

Introduction

Root systems play critical roles in acclimation in plants by modulating growth timing and directionality in response to water and nutrient availability. Variation in root systems is also important for adaptation across species, germplasm, and localities and underscores heritable root traits as breeding targets for enhancing crop resiliency. Despite their importance in plant development and fitness, the quantitative genetic control of root-related traits, particularly in field-grown crops, remains largely underexplored. This lack of understanding of the traits and gene models underlying adaptive variation in roots is due largely to the challenges of obtaining high-quality measurements from hundreds or thousands of samples. With the emergence of various high-throughput phenotyping technologies for roots of field grown crop genotypes, we seek to investigate the relative value of these methods to expand our current understanding of the overall root system phenome and its underlying genetic architecture.

Root system architecture (RSA) is a term that describes the spatial organization of root systems, integrating developmental, anatomical, and topological aspects of root-environment interactions (Fitter 1987). Concerted efforts have been made to develop standardized terminology and classification schemes for root systems, led by the International Society of Root Research and Plant Ontology Consortium (Feix et al., 2002; Ilic et al., 2007; Zobel & Waisel, 2010). These classifications provide precise descriptions of root function and structure, accommodating the inherent complexity and plasticity of root systems, across species and environments. RSA demonstrates notable plasticity in response to environmental factors such as nutrient and water availability, soil composition, and rhizosphere biotic interactions (Clarke et

al., 2019; Johnson & Rasmann, 2015; Karlova et al., 2021; Lareen et al., 2016). The regulation of specific root traits, such as root angle and crown root number, has been shown to improve mobile and immobile nutrient acquisition in response to different environmental scenarios (Bonser et al., 1996; de Jong et al., 2019; Gaudin et al., 2011; Lynch & Brown, 2001; Saengwilai et al., 2014; Sun et al., 2018; Trachsel et al., 2013; York & Lynch, 2015). By integrating systematic classifications with ongoing research, our understanding of RSA can facilitate innovative breeding strategies that optimize plant productivity and adaptation to specific production environments (Schneider et al., 2020).

Current root phenotyping platforms can be broadly categorized into laboratory- and field-based methods (York & Lobet, 2017; York & Lynch, 2015). Lab-based technologies can offer high-resolution measurements but have lower throughput and limited applicability to agronomic environments. In contrast, field-based methods produce data that are more applicable to actual crop production scenarios, despite their lower resolution and destructive nature. Historically, field studies have sampled root crown architecture through manual scoring methods, such as Shovelomics (Trachsel et al., 2011) and root pulling force (RPF) (Ortman et al., 1968; Woods et al. 2022). RPF is a manual measurement of the amount of vertical force required to extract a root system from the soil and has since been scaled to high throughput via mechanization (McKay et al., 2023). RPF is advantageous in that the sampling event itself creates data on a heritable phenotype (Shao et al., 2021; Woods et al., 2022); however, similar to Shovelomics, it is a destructive method of root sampling.

The development of high-throughput phenotyping technologies has advanced our understanding of RSA by enabling the detailed measurement of these complex traits over space and time (Lynch et al., 2014; York et al., 2013). Once a root system is destructively sampled and

cleaned, imaging analysis pipelines, such as automated two-dimensional (2D) image analysis software and X-ray tomography (XRT), have significantly enhanced the quality of data obtained from mature root crowns, providing deeper insights into root morphology and function (Bray & Topp, 2018; Das et al., 2015; Liu et al., 2021; Lobet, 2017; Topp et al., 2013; Zurek et al., 2015). Nevertheless, these methods frequently encounter trade-offs between the degree to which the experimental setup captures phenotypic variations that reflect realistic agronomic conditions and the throughput and detail of the data (York & Lobet, 2017). For example, while XRT can facilitate detailed *in situ* or post-extraction insights into plant structure-function relationships across development (Duncan et al., 2019; Helliwell et al., 2017; Ju et al., 2024; Mooney et al., 2012; Zeng et al., 2021), the application of XRT is limited owing to its relatively high cost and low throughput. Given the array of available phenotyping methods, no single approach fully captures the complete phenotypic reality of root architecture, highlighting the need for using diverse technologies to achieve a comprehensive biological understanding as it relates to agronomically applicable conditions.

Bridging the gap between genotype and phenotype remains a significant challenge in root biology research. RSA is quantitative in nature, controlled by many genes and environmental factors (Rogers & Benfey, 2015; Stuber, 1995). While the dicot model organism *Arabidopsis thaliana* has provided a foundation for understanding root development, it does not directly translate to the genetic architecture of monocot crops, such as maize, rice, and sorghum, which differ in the development and morphology of their root systems (Jung & McCouch, 2013; Péret et al., 2009; Petricka et al., 2012; Scheres et al., 1996; Ueda et al., 2005). The most extensive research on the genetic regulation of RSA in cereal crops has been conducted in rice, where two gene underlying phenotypic variations in roots have been identified: Deeper Rooting 1 (*DRO 1*;

Uga et al., 2013), which promotes a steeper root angle, resulting in deeper root systems, and Phosphorus-Starvation Tolerance 1 (*PSTOL 1*; Gamuyao et al., 2012), which confers early crown root formation under low phosphorus conditions. This underscores the need for further research to identify the key genes and genomic regions that influence root architecture, development, and plasticity in other major cereal crops.

To date, there have been several cloned maize root genes that have exhibited phenotypic abnormalities in root anatomy and development of seedlings (reviewed thoroughly by Feix et al., 2002). While numerous RSA QTL have been identified in field studies, only a few of these genes have been cloned in maize. These more recently include the root-expressed CBL-interacting serine/threonine-protein kinase 15 (*ZmCIPK15*) (Schneider et al., 2022) and the auxin-related *ZmRSA3.1* and *ZmRSA3.2* (Ren et al., 2022), which were shown through reverse genetic approaches to contribute to the regulation of rooting angle and depth. Rootless 1 (*rt1*), first identified in a segregating population by Jenkins (1930), has been shown to have a large effect on shoot-borne and lateral root development across development. One reason for the gap between identified RSA QTLs and functionally validated genes from field studies is the inherent difficulty in capturing the whole phenotypic space. Employing advanced multi-view 2D and 3D imaging techniques could improve the resolution and accuracy of root phenotype data, potentially enabling more effective detection and characterization of key RSA QTL. In this study, we utilize these approaches to provide additional, more comprehensive annotations of the mutant phenotypes for allelic variants of *DRO 1* (‘*dro1-1*’) and *rt1* (‘*rt1-2*’) in mature maize root crowns.

In this study, we systematically evaluate phenome-to-genome variation in field-excavated maize root crowns. Our goal is to understand how these high-throughput measurements capture

the phenotypic space of maize root system architecture, and to advance our understanding of the genetic architecture underlying quantitative root traits through genome-wide and single-gene analyses across three independent field experiments. To achieve this, we employ a diverse set of phenotyping methods, including quantitative measurements of root biomass (g) and RPF (kg), which are collectively referred to as 'physical measurements', as well as measurements of RSA obtained through 2D and 3D imaging and feature extraction. Through this integrative approach, we aim to provide insights into using high-throughput root phenotyping technologies to survey phenome-to-genome variation of RSA for future field studies.

Materials and Methods

Germplasm and Experimental Design

We investigated phenome and genome-wide variation using RSA phenotypes obtained from three field experiments. These materials were grown at the Colorado State University Agricultural Research Development and Education Center in Fort Collins, CO, USA (40.649 N, -105.000 W) during the summers of 2019, 2022, and 2023.

In the first experiment (2019) we evaluated the variations in quantitative root traits measured in 312 inbred lines from the shoot apical meristem (SAM) maize diversity panel (Leiboff et al., 2015). Seeds were sown in May using a split-plot design with two treatment levels of factor irrigation: full irrigation (wet) or limited irrigation (dry), each consisting of three field replicates per genotype. Roots were harvested and phenotyped in both irrigation treatments when approximately 30% of the lines reached anthesis, hereafter referred to as mid-flowering.

In the second experiment (2022) we employed a maize recombinant inbred line (RIL) population that segregated for root trait variations to compare RSA heritability estimates derived from single and multi-view imaging and analysis of 2D maize root crowns. The RIL population

was obtained from a cross between two nested association mapping (NAM) parental lines, B73 and CML69, as described by Buckler et al. (2009). Seeds were sown in May in replicate under full irrigation. These root crowns were harvested and phenotyped prior to anthesis at 11-weeks post-planting. Detailed information about the 2019/2022 field conditions, plant-to-plant spacing, and irrigation differentials was described by Woods et al. (2022).

In the third experiment (2023) we utilized two maize root development mutants, *dro1-1* and *rt1-2*, and their respective wild-type allele lines, W22 and T43, to evaluate the effect of single-gene mutations on quantitative RSA phenotypes. Planting occurred on May 24, with seeds sown at 22-cm spacing in single row, 6.1-meter plots with a row spacing of 76 cm. The field was fertilized according to the recommendations for a targeted yield of 200 bu/ac, amounting to 170 lb/ac N prior to planting. Throughout the growing season, the plots received approximately 2.5 cm of weekly irrigation and were regularly weeded by hand. Phenotyping for root system size was conducted post-anthesis at 21-weeks post-planting.

Field Phenotyping

In the 2019 field experiment, maize root crowns were extracted from soil using manual root pulls. During this process, the maximum force (kg) required to extract the root system from the ground was measured using handheld force gauges and designated as the root pulling force (RPF) (Woods et al., 2022). The throughput of this method was improved in 2022 and 2023 field experiments, where root crown excavations were conducted using a custom root pulling device (CZero Inc., Fort Collins, CO, USA) mounted onto a high clearance tractor (McKay et al., 2023). This mechanical device utilizes a camera to identify maize stalks and securely grasps them by closing a gripper around each stalk and then vertically extracting the plants from the soil. Samples that did not result in successful pull events, which constituted numerous samples from

the 2023 experiment, were subsequently shoveled out manually. The excavated root systems were washed with water and air-dried by hanging, yielding a total sample size of N_{SAM} = 1,546, N_{RIL} = 193, and $N_{MU/WT}$ = 60 cleaned root crowns for 2019, 2022, and 2023, respectively. Root biomass measurements (g) were taken across all three field experiments after the root crowns were air-dried to a constant weight.

Root Imaging and Feature Extraction

We performed image-based analyses of the maize root systems to quantify and compare field-grown variations in root architectural traits using three imaging pipelines. 2D images were captured by suspending the root samples vertical in front of a dark background using a DSLR camera. These images were then processed and analyzed using the 2D DIRT image analysis software (Das et al., 2015). Three-dimensional (3D) root imaging and analysis were conducted using XRT and a custom image processing and feature extraction pipeline (Donald Danforth Plant Science Center, St. Louis, MO, USA), as previously described (Shao et al. 2021)), referred to here as the Root Crown Analysis Pipeline (RCAP). Finally, we reproduced a cost-effective multi-view 'Octagon' chamber and imaging system, as described in Tovar et al. (2018), to simultaneously capture the features of root crowns from three angles (2D multi-view). Raspberry Pi computers controlled the individual 6 MP cameras, which were spaced at 135° and securely attached to the Octagon using industrial-grade VELCRO and facilitated data transfer to a remote host via rsync. For detailed information and protocols regarding software and hardware implementation, refer to Tovar et al. (2018). 2D mult-view RSA measurements were obtained through 2D feature extraction of maize root crowns using DIRT image analysis software (Das et al., 2015) and were computed as the aggregate genotype averages based on three rotated viewpoints.

Descriptive Statistics of Phenotypic Space

Statistical analyses were conducted using R version 4.3.0 (R Core Team, 2023). In the 2019 experiment, we examined root architectural variations in 312 inbred lines from the SAM maize diversity panel. Outliers within each treatment and genotype group were identified and excluded using the Bonferroni outlier test in R/car (Fox & Weisberg, 2019). Observations with Bonferroni-adjusted p-values < 0.05 were considered mean-shift outliers.

In the 2019 and 2022 experiments, broad-sense heritability (H^2) was estimated as the proportion of genotypic variance to total variance using a random-effects one-way analysis of variance (ANOVA) model. Specifically, the phenotypic data were partitioned by experiment and then analyzed using the "lmer" function in R/lme4 (Bates et al., 2015), treating genotype as a random effect in the model. These results were visualized using R/ggplot2 (Wickham, 2016).

Genotype averages for the 2019 experiment were calculated as best linear unbiased predictors (BLUPs), treating genotype as a random effect within treatments using the R/lme4 package (Bates et al., 2015). Spearman's rank genetic correlations were then computed from the BLUPs using R/cor (R Core Team, 2023). RSA phenotypes that demonstrated negligible correlations between phenotyping methods were excluded from further analysis.

Principal component analysis (PCA), an effective statistical method for reducing the dimensional space of data, was performed on the remaining trait variables by employing the function "PCA" in R/FactoMineR (Lê et al., 2008). Axis selection was visually determined by analyzing a scree plot of the canonical eigenvalues. Clustering analyses were visualized using R/factoextra software (Kassambara & Mundt, 2020).

Quantitative Genetic Control of RSA in SAM Maize Genotypes

We implemented a genome-wide association study (GWAS) to identify genomic regions that covary with RSA in 312 SAM lines. Prior to analysis, the phenotypic data were partitioned by irrigation treatment and BLUP values were subsequently calculated for each trait by treating genotype as a random effect using the "lmer" function in R/lme4 (Bates et al., 2015). These values were used as the univariate GWAS response variables. PCA was then used to generate multivariate RSA phenotypes, referred to as "PCs," using the scaled BLUP values from each set of imaging traits (2D and 3D) in R/prcomp (R Core Team, 2023). The data derived from the genotypes on the first two PCA axes were used as GWAS multivariate response variables in the subsequent analyses. These variables were denoted as '2D PC1', '2D PC2', '3D PC1', and '3D PC2' for each environment.

GWAS was performed on 115 BLUP traits and four PC traits across full and limited irrigation using the GAPIT implementation of FarmCPU (Lipka et al., 2012; Liu et al., 2016), following an approach similar to that described by Woods et al. (2022). The initial HapMap genotype matrix for the SAM panel contained 1.2 million single nucleotide polymorphisms (SNPs) (Leiboff et al., 2015), which was refined to 860,000 after imposing a minor allele frequency filter of 5%. The first three principal components and a kinship matrix calculated using GAPIT were used as covariates to control for population structure. Only traits with Q-Q plots displaying p-value distributions that approximated normality were included in the FarmCPU results. SNPs whose significance passed the Benjamini—Hochberg false discovery rate threshold (0.10) were further investigated to determine the presence of co-located QTL across methods, including the list of 2019 RSA-associated SNPs previously identified by Woods et al. (2022).

The RSA-associated SNPs were fitted as random effects to estimate the SNP-attributed and residual variance components using the R package lme4 v1.1-34 (Bates et al., 2015). The proportion of phenotypic variance explained (PVE) by each SNP was calculated as the ratio of their corresponding variance to the total variance. Candidate regions of SNP colocalization were identified using 1-Mb bins within each chromosome. Bin indices were calculated based on the number of complete bins that occur before the current SNP position within each chromosome, using the following formula:

$$Bin\ Index_{Chromsome} = \left|\frac{Position - 1}{Bin\ Size}\right|$$

Here, 'Position' represents the SNP position mapped to the v2 B73 reference genome, and 'Bin Size' refers to the desired bin size, set at 1 Mb (in base pairs, bp). The results were visualized using R/ComplexHeatmap, R/factoextra, and R/VennDiagram (Chen, 2022; Gu, 2022; Gu et al., 2016; Kassambara & Mundt, 2020).

Evaluation of Single-Gene Perturbed Maize Genotypes

In the 2023 experiment, two maize root development mutants and their corresponding wild-type allele lines were used to examine the effects of single-gene mutations on individual RSA phenotypes and to determine patterns of among-trait covariance across 2D, 2D multi-view, and 3D imaging approaches. Differences in average phenotype response between mutant and wild-type alleles were assessed using a two-sample t-test in R/t.test (R Core Team, 2023), with a significance level (α) of 0.05. Normality and equal variance were assessed prior to statistical testing to confirm the data met the test assumptions. When these assumptions were not met, appropriate transformations were applied using R/bestNormalize (Peterson, 2021; Peterson & Cavanaugh, 2020; Peterson, 2023).

Pleiotropy is a common phenomenon where a single gene can influence multiple traits. In such cases, selection on mutations that affect the expression and inheritance of multiple traits can shape existing covariance structures. To investigate this phenomenon in the two maize root mutant genotypes, we conducted a Pearson correlation analysis using R/cor.test (R Core Team, 2023) to calculate the genetic correlation between RSA phenotypes within and among imaging methods. In addition, we applied a multivariate analysis of variance (MANOVA) to test for pairwise interactions among all trait combinations in individual models. We evaluated the results of the MANOVA models at a significance level (α) of 0.05.

Results

Multi-view 2D imaging improves H^2 estimation of RSA by reduction of sampling bias in RILs

Higher estimates of broad-sense heritability were observed for 17 of the 33 RSA traits in the CML69 x B73 NAM RIL population when utilizing the 2D multi-view imaging approach compared to 2D single-view imaging (Figure 3.1). Among these traits, the average difference between the two approaches was 0.13 ± 0.09, with the highest increases observed for average root density (2D AvgDensity; ± 0.36) and skeleton root width (2D SklWidth; ± 0.24). Overall, phenotype variances were higher in all the traits measured using 2D single-view imaging, except for 2D AvgDensity, top root angle (2D AngTop), and bottom root angle (2D AngBtm). Root crowns are rarely symmetrical and are typically positioned with the widest part orthogonal to the camera; this is likely to cause sampling bias, and consequently, less consistent RSA measurements during feature extraction using the DIRT software. The remaining 14 phenotypes showed marginally higher average differences in heritability when utilizing single-view 2D imaging (0.17 ± 0.09). This group primarily comprised individual root distribution metrics, including percentage of root width accumulation (2D D10-D90) and rate of accumulation (2D

DS10, DS80-DS90) at different depths, demonstrating an inverse relationship between heritability and depth (i.e., root width accumulation measurements at shallower depths were more heritable). These findings indicate that using as few as three rotational viewpoints in 2D imaging and feature extraction can improve the accuracy and consistency of broad-sense heritability estimates for the coarser trait characteristics of field-grown maize root crowns.

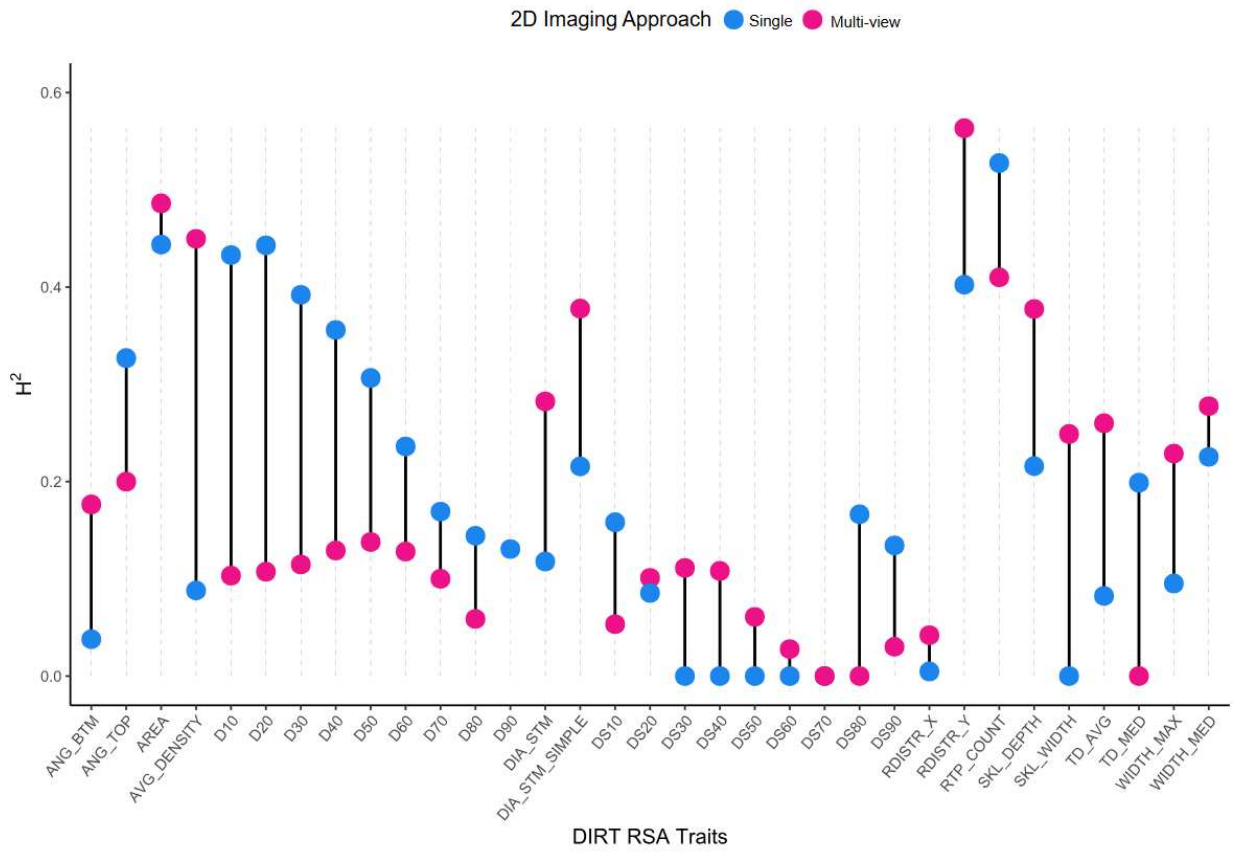

Figure 3.1. Trait-specific differences in broad-sense heritability between 2D single and multi-view approaches for the 2022 CML69 x B73 RIL population.

Belowground variation is largely driven by root system size and distribution in SAM maize panel

Spearman's correlation method revealed a broad distribution of genotypic correlations among the different phenotyping methods in the SAM maize panel. RPF showed the greatest positive correlations ($r \leq 0.70$) with traits associated with overall root system size, including root

biomass, 2D Area, 3D Volume, and fractal dimension side-view (3D FractalDimensionS). These results were consistent with the findings of Shao et al. (2021) and Woods et al. (2022).

In addition, there were 41 RSA traits that displayed weaker genetic correlations across phenotyping methods. Specifically, 3D measurements of root model solidity at various depths (vhist 1-20) showed weaker correlations on average (r = -0.01 ± 0.12) with all other traits. 2D RSA traits, including median root tip diameter (2D TdMed), horizontal distribution of the root shape (2D RDistrX), 2D D10-D90, and 2D DS10-DS90 also showed weaker correlations on average (r = -0.02 ± 0.08) with other traits. These results suggest that the aforementioned 2D traits, which exhibit greater phenotypic variance when assessed using traditional 2D imaging techniques and lower heritability estimates when evaluated using a multi-view approach, are less useful in approximating key phenotypes in maize root crowns. To effectively infer trait relationships among different phenotyping methods, we excluded these 41 traits from further analysis.

PCA was used to investigate the phenotypic relationships among the remaining 115 RSA phenotypes. The first two principal components accounted for 32.30% and 9.90% of the total variation, respectively, with no discrete separation among individuals between treatment groups (Figure 3.2). We found a clear separation between the types of RSA phenotypes that were strongly loaded on PC1 compared to PC2 (Figure 3.3A). PC1 predominantly captured variation in overall root system size and distribution, with higher scores positively correlated with 2D Area, 3D ConvexVolume, and 3D biomass distributions (vhist 13-16), indicative of larger root sizes. In contrast, PC2 captured differences in root density (3D Density S2, S6, T2, T4-T6), and in maximum root width (3D HorEqDiameter). Further evidence of these relationships was depicted by Figure 3.3B, which illustrated a sample of the variations in maize root crown size,

shape, and distribution through XRT scans of four individuals from genotypes that were representative of each quadrant of the multidimensional space.

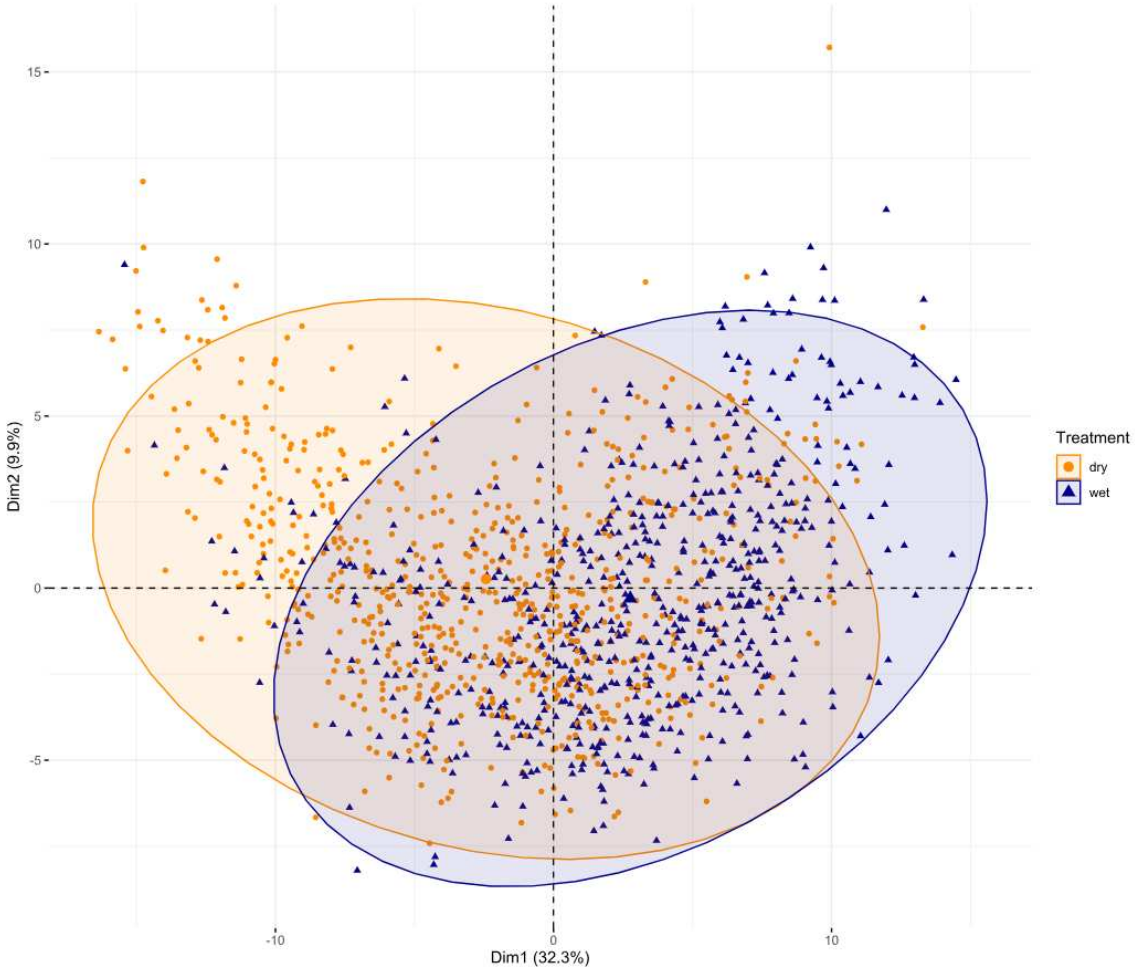

Figure 3.2. PCA of RSA phenotypes partitioned by irrigation treatment measured in the 2019 SAM maize panel.

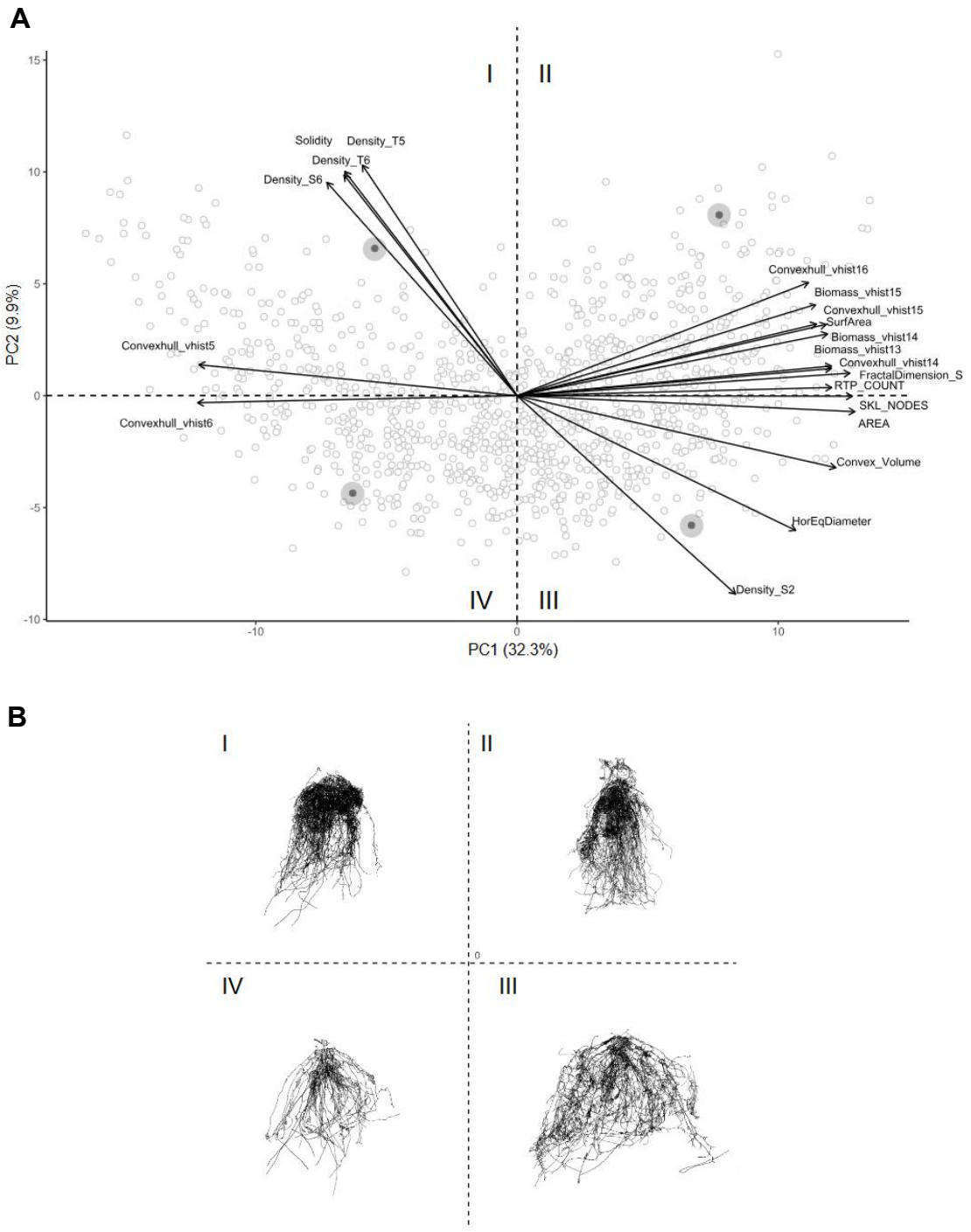

Figure 3.3. Relationship of 115 RSA phenotypes across 312 genotypes from the 2019 SAM maize panel. **(A)** PCA biplot of the scaled first two principal components for roots excavated at mid-flowering. Open circles represent individual genotypes, and vectors indicate the top 20 RSA variables with the highest loadings on PC1 and PC2. Highlighted individuals are further illustrated in panel (B). **(B)** Comparison of maize root crown features through X-ray computed tomography (XRT) scans of four distinct genotypes, contrasting the physical attributes of genotypes collected of the same replicate from each quadrant (I-IV) in the biplot.

Estimates of broad-sense heritability in the SAM lines were generally higher for all evaluated RSA phenotypes under wet than under dry conditions (Figure 3.4). This supports the prediction that low-stress environments, such as irrigated environments, can increase the range of observable phenotypic variation, facilitating better distinction between genotypes (reviewed by Visscher et al., 2008). Within the wet treatment, heritability estimates ranged from 0.33 to 0.43 for 2D RSA phenotypes that were highly correlated with PC1, with 2D Area exhibiting the highest estimate. The 3D RSA phenotypes associated with PC1 exhibited a broader range (H^2 = 0.19 to 0.45), with FractalDimensionS exhibiting the highest estimate. The phenotypes associated with PC2, which encompass a range of root density traits, resulted in heritability estimates ranging from 0.32 to 0.41, with the highest estimation attributed to 3D Density T5. Although RPF and root biomass were not strongly correlated with PC1 and PC2, these measurements demonstrated higher average heritability than 2D and 3D phenotypes (Figure 3.4).

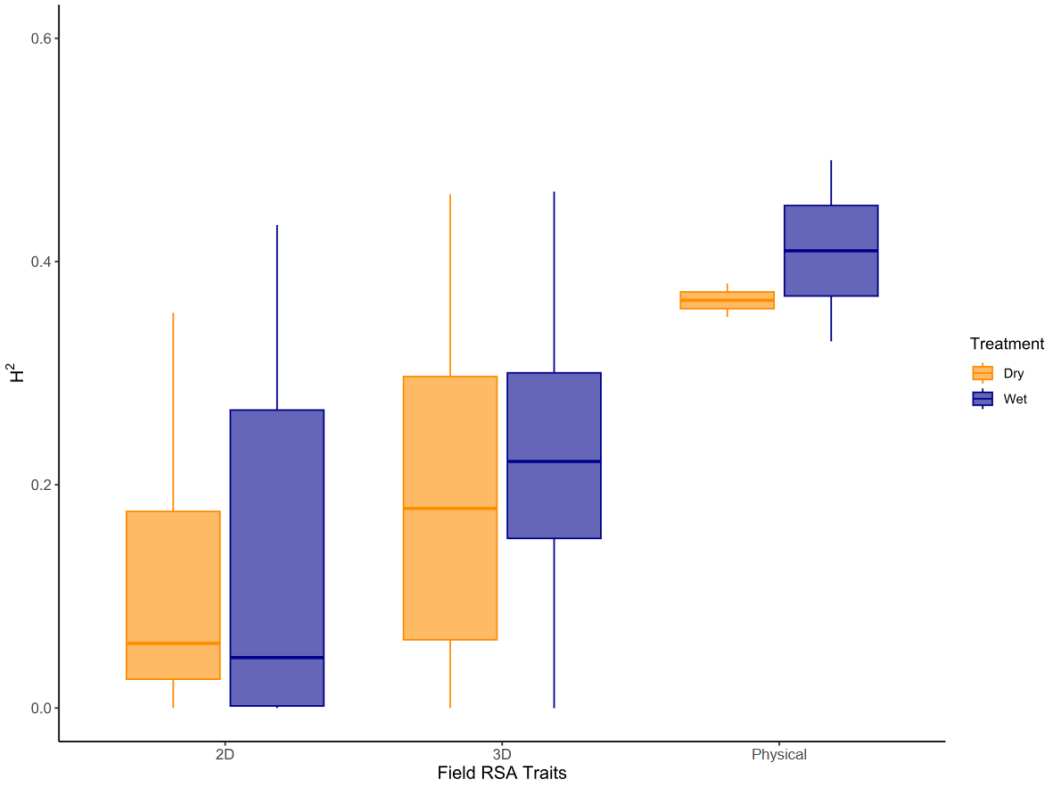

Figure 3.4. Distribution of broad-sense heritability estimates of RSA traits among phenotyping schemes and irrigation treatments.

Multivariate analysis identifies genome-wide relationships among individual RSA phenotypes

The FarmCPU GWAS models identified 405 significant SNPs across 288 1-Mb bins associated with multivariate (PC) and univariate RSA phenotypes, in addition to the 22 SNPs previously reported by Woods et al. (2022) (Figure 3.5). The simultaneous use of multivariate and univariate RSA phenotypes as response variables in GWAS models enabled better detection of trait-specific and pleiotropic genomic signals. The combined set of significant SNPs was predominantly identified by 3D phenotypes (83.84%), with 3D biomass distributions and 3D convex hull distributions accounting for the majority of these associations. This was followed by the RPF (7.03%), 2D (5.62%), and PC (3.51%) phenotypes (Figure 3.5). A greater proportion of SNPs associated with individual RSA phenotypes corresponded to RPF (7.03%), 3D Solidity (4.45%), and 3D Density S6 (3.75%). Notably, RPF shared the largest proportion of 1-Mb bins

with SNPs associated with other phenotypes. RPF measurements were highly heritable in both environments (wet treatment: $H^2 = 0.35$; dry treatment: $H^2 = 0.49$) and predominantly colocalized with lower heritable 3D phenotypes. Specifically, these SNPs were associated with phenotypes representative of root system density, and biomass and convex hull distributions within the upper seven centimeters of the root crown, which cumulatively displayed higher values of PVE than RPF (Figure 3.6). Overall, 3D RSA phenotypes accounted for the largest proportion of phenotypic variation and the corresponding genetic loci.

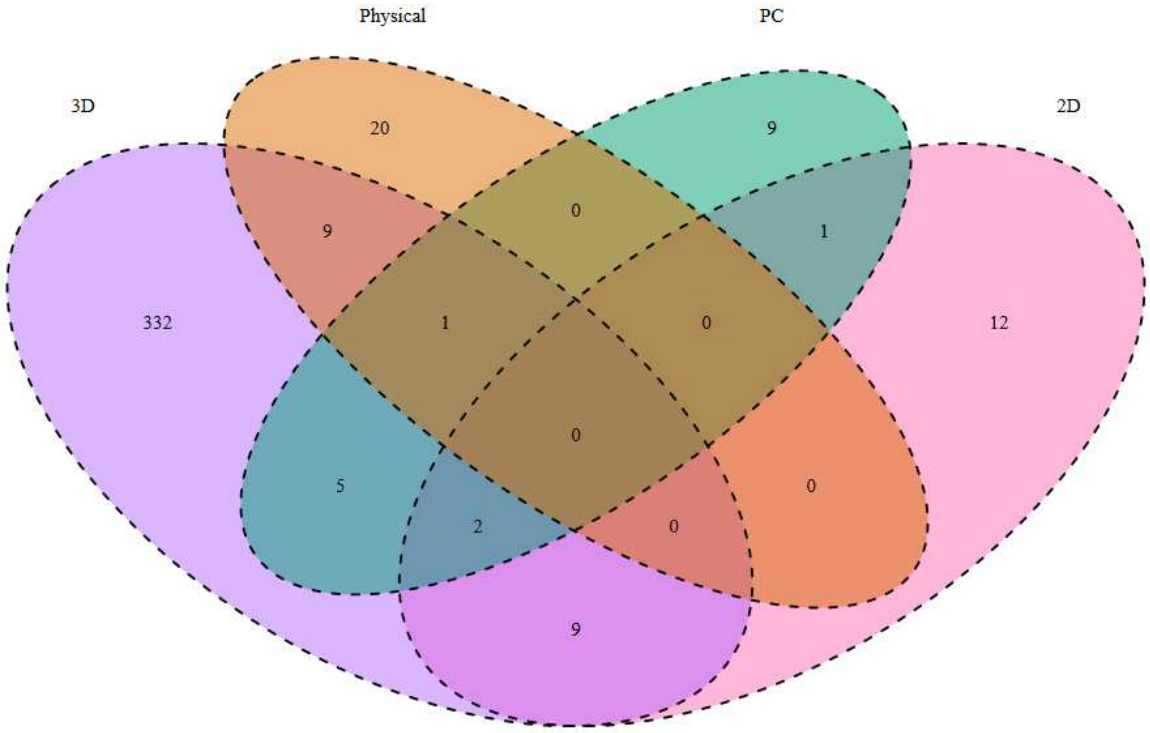

Figure 3.5. Venn Diagram of GWAS trait-associated 1-Mb bins partitioned by RSA phenotyping methods. Overlapping regions indicate 1-Mb bins that were enriched in two or more comparisons.

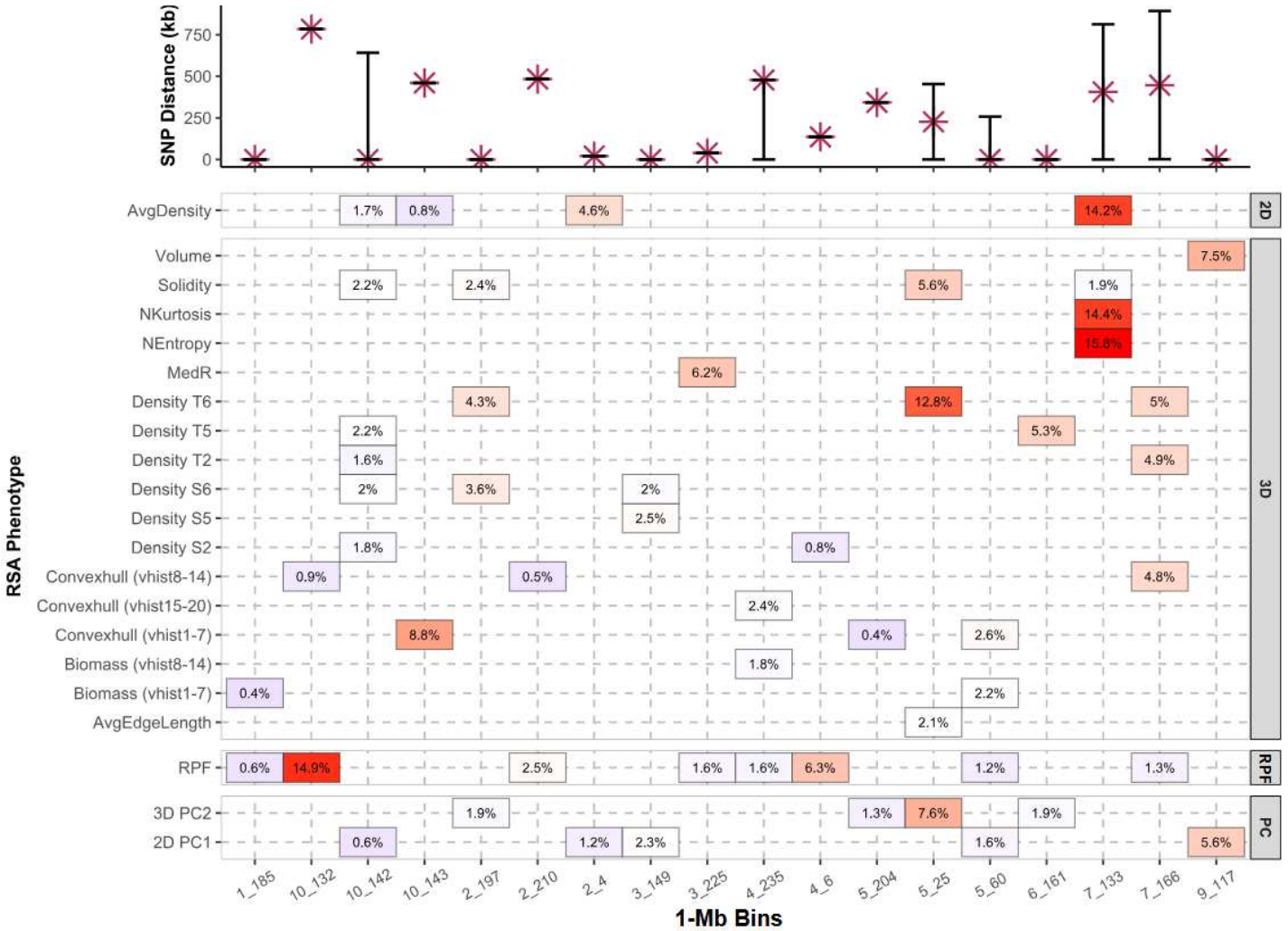

Figure 3.6. Heatmap of the percentage of phenotypic variance explained (PVE) by 1-Mb genomic regions associated with traits colocalized across imaging, RPF, and multivariate approaches. The color intensity reflects the percentage of PVE attributed to significant GWAS SNPs identified by this study for each 1-Mb region. The marginal plot above the heatmap displays the genomic distances between SNPs within each 1-Mb region, with error bars representing the minimum and maximum distances (in kb) and stars indicating the median distances.

The use of multivariate RSA phenotypes were instrumental in investigating key phenotype-to-genotype relationships. Approximately 23.08% (nine SNPs) of the SNPs associated with the multivariate traits were co-located with several of the univariate phenotypes. PC1 and PC2 accounted for substantial portions of the phenotypic variation, with approximately 51% explained in the 2D phenotypes and 32% explained in the 3D phenotypes, under both treatments. We decomposed the trait space of the eight multivariate phenotypes to identify phenotypic relationships among genotypes along each principal axis. Among the 2D multivariate phenotypes, PC1 predominantly captured variations in root size and spatial characteristics, such as root area and lateral extent, while PC2 largely described variations in root angle and depth. The 3D PC1 captured variations in root distribution, including 3D biomass and convex hull distributions in the lower 50% (vhist 13-16) of the root crown, while PC2 described variations in root density and uniformity. These relationships were captured on the genome-wide level as SNPs associated with 2D PC1 were identified in multiple 1-Mb bins that contained SNPs associated with individual 2D and 3D root density, area, and volume related phenotypes. Similarly, the SNPs identified by 3D PC2 colocalized with individual 3D density and solidity phenotypes (Figure 3.6). We estimated the phenotypic correlations between the multivariate phenotypes and the breeding values for root biomass to obtain a biological ground truth for data interpretation. Spearman's correlation analysis revealed significant positive correlations for root biomass with PC1, with PC1 2D ($r = 0.69$; $p < 0.05$) and PC1 3D ($r = 0.67$; $p < 0.05$). In contrast, weaker correlations were observed with PC2 , where PC2 2D ($r = -0.04$; $p = 0.30$) and PC2 3D ($r = -0.28$; $p < 0.05$) (Figure 3.7). These results indicate that PC1 for both 2D and 3D RSA phenotypes was a strong predictor of belowground biomass.

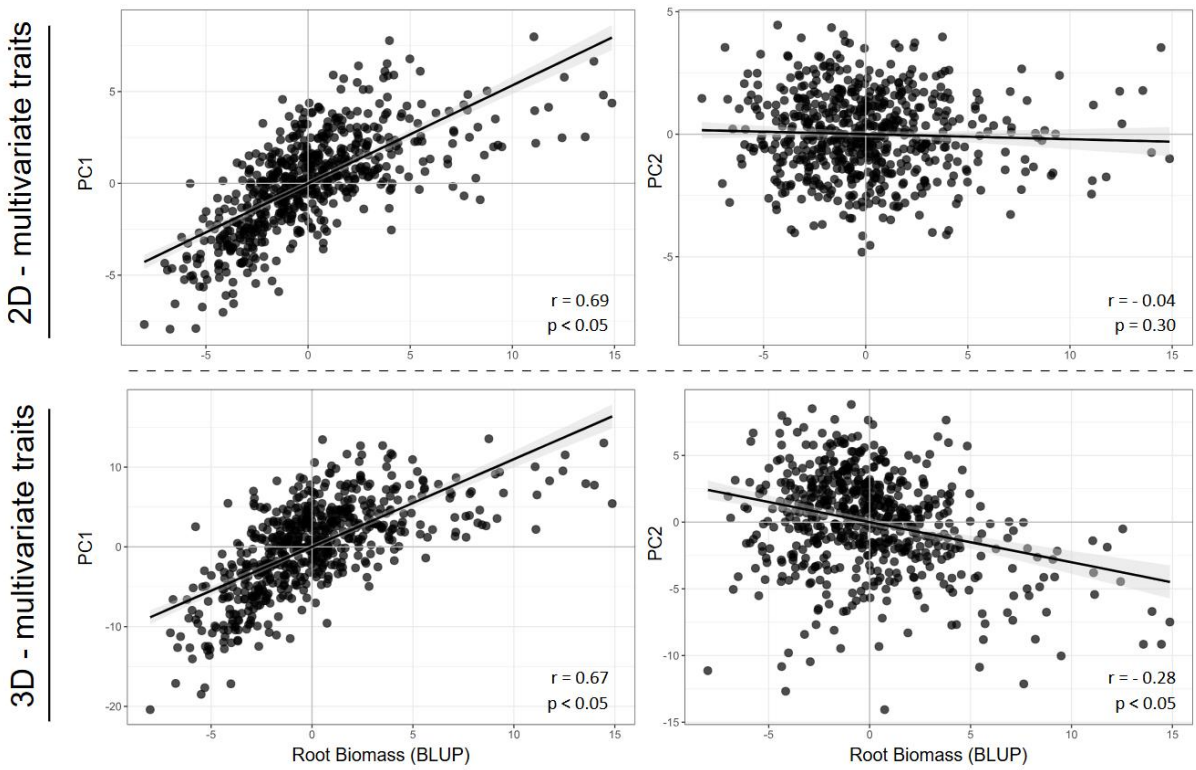

Figure 3.7. Correlation between genotypic values (BLUPs) of 2D and 3D GWAS multivariate root traits (PC1 and PC2) and root biomass in the 2019 SAM maize panel. Abbreviations: BLUP = best linear unbiased predictor; *r* = Spearman's rank correlation coefficient; *p* = probability value.

Complementary phenotyping detects mutational variance-covariances in maize Deeper Rooting 1 and Rootless 1

The two maize mutants selected for altered root architectural phenotypes exhibited variations in several RSA phenotypes. The Rootless 1 mutation in maize has been documented to reduce the number of lateral and crown roots in maize (Jenkins, 1930). The *rt1-2* maize mutant and wild-type allele lines of the 2023 field experiment demonstrated significant differences ($p < 0.05$) in the genotype averages for several heritable root phenotypes under normal conditions (Table 3.1). We observed that the mutation reduced the root volume, area, and width in both 2D and 3D imaging. The 3D RSA features calculated in the RCAP pipeline provided additional evidence that was not captured by the 2D RSA data alone. For example, multiple measurements

representative of root system branching, such as the estimated number of branching points (3D NumberBifCl), maximum number of lateral roots (3D MaxR), and number of segments between branching points (3D EdgeNum), were substantially reduced in the mutant lines. In this study, we measured field-grown mature maize root crowns using imaging-based phenotyping technologies and obtained results in the *rt1-2* mutant that were consistent with the qualitative descriptions of seedling root phenotypes of maize rootless monogenic mutants (reviewed by Hochholdinger et al., 2018).

Table 3.1. Differences in heritable RSA phenotypes between *dro1-1* and *rt1-2* maize mutant lines and their corresponding wild-type alleles. Broad-sense heritability estimated for the SAM maize lines subjected to full-irrigation are presented for each phenotype, with additional estimates derived from the CML69 x B73 RIL population presented in parentheses.

		dro1-1					*rt1-2*				
	H^2	M_{wt}	M_{mu}	*df*	*t*	*p*	M_{wt}	M_{mu}	*df*	*t*	*p*
Area[a]	0.43 (0.44)	8533	10086	11	1.60	0.137	10095	6917	19	-3.30	**0.004***
Area[ac]	0.43 (0.49)	11600	11879	12	0.25	0.809	-	13449	-	-	**-**
WidthMed[a]	0.39 (0.23)	99.1	101	18	0.21	0.837	98	72.2	21	-2.62	**0.016***
WidthMed[ac]	0.39 (0.28)	79.3	92.5	16	1.32	0.205	-	84.5	-	-	**-**
SklWidth[a]	0.36 (0)	173.7	187.4	14	0.84	0.414	176.7	131.5	19	-2.42	**0.025***
SklWidth[ac]	0.36 (0.25)	152	183	13	2.99	**0.010***	-	158	-	-	-
RTPCount[a]	0.38 (0.53)	407	552	22	2.42	0.024	586	373	15	-2.34	**0.033***
RTPCount[ac]	0.38 (0.41)	105	150	16	3.54	**0.002***	-	88.1	-	-	-
FractalDimensionT[b]	0.44	1.6	1.7	21	3.11	**0.005***	1.7	1.6	11	-4.53	**≤ 0.001***
FractalDimensionS[b]	0.43	1.7	1.7	16	2.14	**0.048***	1.7	1.7	11	-5.59	**≤ 0.001***
HorEqDiameter[b]	0.41	401	521	8	3.59	**0.007***	466	364	9	-5.50	**≤ 0.001***
NumberTips[b]	0.46	1637	2328	9	3.06	**0.013***	2386	1225	11	-6.32	**≤ 0.001***
NumberBifCl[b]	0.41	4485	6284	6	2.49	**0.042***	6326	4154	10	-4.22	**≤ 0.001***
TotalLength[b]	0.43	90338	117418	9	2.53	**0.031***	127778	83140	11	-4.52	**≤ 0.001***

Abbreviations: H^2, SAM Heritability Estimate (full-irrigation); M, Mean; Wt, Wild-type; Mu, Mutant; WidthMed, Median Width; RTPCount, Root Tip Count; SklWidth, Skeleton Width; FractalDimensionT, Fractal Dimension (top-view); FractalDimensionS, Fractal Dimension (side-view); HorEqDiameter, Maximum Width; NumberTips, Number Tips; NumberBifCl, Number Skeleton Branching Points; TotalLength, Skeleton Root Length.
***Welch test is statistically significant at $p \leq 0.05$
[a]DIRT
[b]RCAP

Mutant screens of the *dro1-1* maize lines demonstrated variation in the root system size and horizontal spread (Table 3.1). Across these individuals, the mutant exhibited significantly

higher trait averages ($p < 0.05$) for 2D and 3D root branching characteristics, including root tip count (2D RtpCount), 3D NumberBifCl, and 3D EdgeNum. The mutant allele also conferred a higher width-to-depth ratio compared to the wild-type, with significant differences ($p < 0.05$) in the genotype averages for 3D HorEqDiameter, root length (3D TotalLength), and width-to-depth ratio (3D WDRatio). Although the differences in 2D SklWidth were not statistically significant between the mutant and wild-type lines, increasing the number of root crown images from single to multi-view enhanced the ability to detect significant associations with this trait. As further evidence of the utility of the 2D multi-view imaging as a phenotyping method, Figure 3.8 depicts the genotype averages of root width associated with the 2D, 2D multi-view, and 3D imaging approaches and feature extraction, wherein 2D and multi-view imaging were represented by 2D SklWidth (DIRT) while 3D imaging was represented by 3D HorEqDiameter (RCAP). All three estimates of root width demonstrated similar relationships among genotypic averages. We found that capturing more of the phenotypic space improved the genotypic estimations in monogenic maize mutants. These trait patterns correspond to established phenotypic characterizations of root gravitropism associated with the Deeper Rooting 1 gene found in maize (Feng et al., 2022) and rice cultivars (Uga et al., 2013).

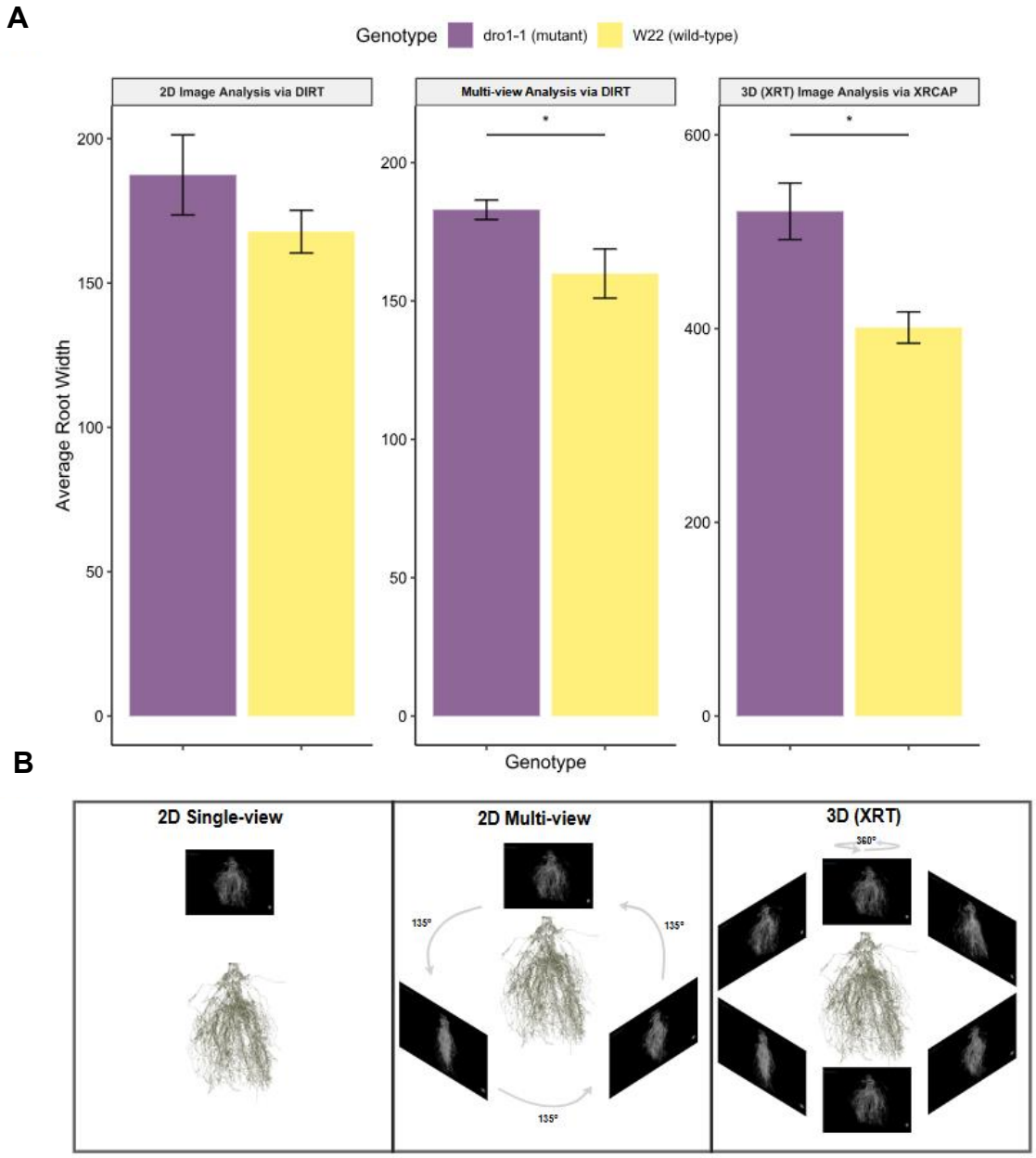

Figure 3.8. Comparison of average RSA trait values for a single-gene perturbation across three distinct imaging and analysis pipelines. **(A)** Average root width ± SE between the maize mutant line *dro1-1* and its corresponding wild-type allele line "W22." The feature extraction of root width via DIRT for 2D and octagon platforms was "SklWidth" and via RCAP for the 3D platform was "HorEqDiameter." Significant differences between the mutant and wild-type comparisons are indicated by asterisks (*), with $p \leq 0.05$. **(B)** A simplified schematic of the relative amount of root crown data captured by each imaging platform.

Our analysis of mutational covariance in the *rt1-2* and *dro1-1* maize lines revealed strong evidence for pleiotropy among the 3D RSA phenotypes. A two-way MANOVA demonstrated a

significant effect of genotype [Pillai's Trace = 0.50, $F(2, 68) = 5.63$, $p < 0.001$] on 3D HorEqDiameter and Depth, accounting for genetic background. Comparison of mutants between the gene models showed different patterns of covariation, with *dro1-1* showing a significant negative correlation between the two traits ($r = -0.81$, $p < 0.05$), whereas *rt1-2* displayed no evidence of phenotypic correlation ($r = 0.02$, $p = 0.96$). This trade-off between root width and depth aligns with our prediction for the *dro1-1* mutant phenotype, which is characterized by a shallower root system with greater horizontal spread (Uga et al., 2013). There was also a significant effect of genotype [Pillai's Trace = 0.56, $F(4, 68) = 6.65$, $p < 0.001$] on 3D NumberBifCl and 3D Volume. The *rt1-2* mutant allele lines showed a significant positive correlation ($r = 0.90$, $p < 0.01$), indicating that as root size increases, the number of branching points also increases. In contrast, the *dro1-1* mutant allele lines exhibited a nonsignificant negative correlation ($r = -0.34$, $p = 0.65$). Given that root volume can be a good proxy for "biomass" in digital phenotyping (Shao et al. 2021), we estimated the correlation between root biomass (g) and 3D NumBifCl to assess the precision of 3D measurements in estimating mutational covariances. The results were consistent, with the *rt1-2* showing a strong positive correlation between root biomass and 3D NumBifCl ($r = 0.83$, $p < 0.05$), whereas the *dro1-1* mutants exhibited a weak negative correlation ($r = -0.87$, $p = 0.73$).

Discussion

In this study, we evaluated multiple phenotyping and analytical frameworks for quantifying RSA in mature maize root crowns across three field experiments. Our primary objectives were to assess how different quantitative measurements of RSA describe the broader root phenome and to determine the significance of this information for mapping genotype-phenotype relationships. Specifically, we examined their roles in the genome-wide identification

of genetic loci controlling RSA and in phenotypic characterization of genes known to influence root development. Our results demonstrated that increasing the number of rotational viewpoints in 2D imaging improved RSA heritability estimates and genotype-to-phenotype mapping, compared to single images. Additionally, our findings support the use of multivariate traits as an effective method to quantify complex RSA phenotypes and identify putative pleiotropic loci that are often missed by canonical univariate approaches. Overall, our phenomic and genome-wide analyses revealed that the 3D RSA phenotypes accounted for the largest proportion of phenotypic variation and the corresponding genetic loci. This finding was further supported by the analysis of maize root development mutant lines, *dro1-1* and *rt1-2*, whose phenotypic annotations aligned with previous qualitative descriptions that were not captured by the 2D single-view approach. However, measurements of RPF from the 2019 maize panel also exhibited higher heritability estimates and demonstrated strong correlations with 3D phenotypes—suggesting that this method may serve as a more accessible alternative for large-scale field phenotyping

The field of phenomics aims to enhance our understanding of genotype-phenotype maps through the systematic capture of whole-organism phenotype information and multidimensional data analysis (Houle et al., 2010). Research in this domain has expanded the plant phenome and identified new loci and gene models that underlie it (Li et al., 2018; Li et al., 2024; Ubbens et al., 2020). Our findings contribute to these broader efforts by demonstrating how diverse RSA phenotyping methods under different environmental conditions inform the understanding of genotype-phenotype relationships. Heritability estimates, as shown by previous research, can fluctuate based on measurement method and environmental conditions, with lower heritability often observed in less favorable environments (Hill et al., 1983; Hoffmann et al., 1999; Macgregor et al., 2006; Roff et al., 1997). Our study using the SAM maize lines revealed higher

broad-sense heritability for RSA phenotypes in low-stress environments, alongside observable differences between 2D and 3D phenotyping approaches. Specifically, 2D imaging and feature extraction, while commonly utilized as an accessible and relatively straightforward approach, captured less variation and resulted in lower heritability estimates under irrigated conditions ($H^2 = 0.12 \pm 0.15$) (Figure 3.4). Despite its limitations, the 2D multi-view approach represents a readily implementable improvement for addressing genetic hypotheses, especially when investigating RSA traits implicated in asymmetry, such as root width, angle, and branching patterns. However, additional work is needed to optimize the integration of multi-view data into consolidated phenotypic values, as simple averages across images may not accurately represent the 3D root system.

Advances in 3D imaging technology have significantly enhanced our ability to interrogate root systems and capture the complexity of RSA in agricultural systems (Li et al., 2023; Liu et al., 2021; Jung et al., 2013; Shao et al., 2021; Topp et al., 2013; Morris et al., 2017; Dowd et al., 2022; Mattupalli et al., 2018). Our study reinforced this by demonstrating that 3D measurements not only provided detailed insights at the phenome level, but also revealed a greater capacity to detect significant variations in both genome-wide and single-gene analyses. PCA of SAM maize lines revealed substantial phenotypic variation driven by 3D phenotypes, particularly those linked to root size, distribution, and density (Figure 3.3A). These results were further validated by XRT scans of the root crowns, as the root models visually reflected size variations along PC1 and density variations along PC2, consistent with the vectors in the PCA biplot (Figure 3.3B). Our genome-wide association analyses identified the largest number of RSA-associated loci using 3D phenotypes, with these loci showing higher average PVE values than 2D, RPF, and multivariate traits (Figures 3.5 and 3.6). Notably, mutant lines such as *dro1-1*

and *rt1-2* highlighted the ability of 3D phenotyping to capture root allocation trade-offs, consistent with their previous qualitative descriptions, revealing putative pleiotropic effects of the perturbed genes (Figure 3.8; Table 3.1) (Feng et al., 2022; Hochholdinger et al., 2018; Jenkins, 1930; Uga et al., 2013).

Although 3D phenotyping offers substantial advantages for genome-wide studies of RSA in maize, its widespread application is often constrained by its associated overhead costs and training. In contrast, RPF measurements, which demonstrate high heritability and strong correlations with 3D phenotypes, may serve as a more accessible method for quantifying RSA for large-scale field phenotyping (Shao et al., 2021; Woods et al., 2022). However, the suitability of its application varies depending on the system, root archetype, and experimental design. Contrary to our expectations, genome-wide association analyses revealed numerous SNPs associated with lower heritable 3D phenotypes that co-located with RPF yet explained a higher proportion of phenotypic variance for each bin (Figure 3.6). Based on the quantitative genetic theory, we hypothesized that the genetic architecture of RPF may be more polygenic in nature, with many genes contributing to small effects (Fisher, 1919, 1930). In contrast, the larger PVE values found among information-rich 3D phenotypes could be controlled by fewer loci with larger effects. Genome-wide associations with RPF could have also been confounded by non-additive genetic variation, as estimates of broad-sense heritability reflect the total genetic variation comprising additive, dominance, and epistatic effects (Lynch & Walsh, 1998).

Complex phenotypes in biological organisms emerge from interactions and contributions of individual traits. A common approach to studying the genetic underpinnings of these phenotypes involves using data reduction techniques, such as principal components (PCs), to generate ‘composite traits’ (Hotelling, 1933; Klei et al., 2008; Li et al., 2023; Rice et al., 2020;

Topp et al., 2013). Our study utilized 2D and 3D multivariate phenotypes derived from PCA to test for genome-wide associations. We found that these phenotypes, particularly PC1, had strong positive associations with root biomass accumulation (Figure 3.7). Additionally, the loci identified by multivariate phenotypes were primarily located within regions that were also associated with individual traits (Figure 3.5). Notably, the combination of individual traits for these regions reflected both the correlated (e.g., 3D PC2 and 3D root density traits) and orthogonal (e.g., 2D PC1 and 2D AvgDensity) phenotypic relationships observed in the PCA (Figure 3.3A; Figure 3.6). One explanation for these patterns is that similar and distinct elements of maize root architecture are controlled by the same genes, which is indicative of a pleiotropic genetic architecture. Although multivariate phenotypes do not directly provide biological insights into pleiotropic loci, they are advantageous for elucidating the relationships between low- and highly heritable traits at a given locus, which are often missed by univariate approaches (Rice et al., 2020). These results further demonstrate that the multivariate approach has the power to replicate univariate results, reinforcing its value in forward genetic studies (Rice et al., 2020; Topp et al., 2013).

Understanding the genetic and molecular mechanisms that regulate root system development and architecture is critical for enhancing plant resilience to environmental stress, improving resource use efficiency, and supporting ecosystem services such as soil stabilization and carbon sequestration. Historically, much of our understanding of this area has been limited to studies in controlled environments, focusing on seedlings or model systems. This has created a gap in our ability to study complex, mature root systems in agronomic environments, primarily because of the lack of appropriate phenotyping tools and analytical methods. Our research addressed this by applying advanced RSA phenotyping frameworks to mature maize root

crowns, revealing that capturing multidimensional data, such as through 2D multi-view, 3D phenotyping, and multivariate approaches, significantly improves the precision of phenotype-genotype mapping. These findings underscore the importance of employing complementary approaches to fully explore the phenotype-genotype gap. However, it is worth noting that the relationship between phenotyping methods and their effectiveness is not necessarily static. As feature extraction technologies—such as AI-based tools, topological data analysis, and advanced 3D graphics—continue to evolve, the potential for capturing more comprehensive and accurate root phenotypes will grow. Therefore, future research should not only build on current methods but also remain adaptable, integrating cutting-edge technologies to further refine our understanding of the root phenome and its genetic drivers under diverse agricultural conditions.

REFERENCES

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bonser, A. M. Y. M., Lynch, J., & Snapp, S. (1996). Effect of phosphorus deficiency on growth angle of basal roots in Phaseolus vulgaris. *New Phytologist*, *132*(2), 281–288. https://doi.org/https://doi.org/10.1111/j.1469-8137.1996.tb01847.x

Bray, A. L., & Topp, C. N. (2018). The Quantitative Genetic Control of Root Architecture in Maize. *Plant and Cell Physiology*, *59*(10), 1919–1930. https://doi.org/10.1093/pcp/pcy141

Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., Ersoz, E., Flint-Garcia, S., Garcia, A., Glaubitz, J. C., Goodman, M. M., Harjes, C., Guill, K., Kroon, D. E., Larsson, S., Lepak, N. K., Li, H., Mitchell, S. E., Pressoir, G., … McMullen, M. D. (2009). The genetic architecture of maize flowering time. *Science*, *325*(5941), 714–718. https://doi.org/10.1126/science.1174276

Casci, T. (2004). Pleiotropy — what a drag. *Nature Reviews Genetics*, 5(4), 248. https://doi.org/10.1038/nrg1335

Chen, H. (2022). *VennDiagram: Generate High-Resolution Venn and Euler Plots*. https://cran.r-project.org/package=VennDiagram

Clarke, C. R., Timko, M. P., Yoder, J. I., Axtell, M. J., & Westwood, J. H. (2019). Molecular Dialog Between Parasitic Plants and Their Hosts. *Annual Review of Phytopathology*, *57*, 279–299. https://doi.org/https://doi.org/10.1146/annurev-phyto-082718-100043

Das, A., Schneider, H., Burridge, J., Ascanio, A. K. M., Wojciechowski, T., Topp, C. N., Lynch, J. P., Weitz, J. S., & Bucksch, A. (2015). Digital imaging of root traits (DIRT): A high-throughput computing and collaboration platform for field-based root phenomics. *Plant Methods*, 11(1), 1–12. https://doi.org/10.1186/s13007-015-0093-3

de Jong, M., Tavares, H., Pasam, R. K., Butler, R., Ward, S., George, G., Melnyk, C. W., Challis, R., Kover, P. X., & Leyser, O. (2019). Natural variation in Arabidopsis shoot branching plasticity in response to nitrate supply affects fitness. *PLoS Genetics*, 15(9). https://doi.org/10.1371/journal.pgen.1008366

Dowd, T. G., Li, M., Bagnall, G. C., Johnston, A., & Topp, C. N. (2022). Root system architecture and environmental flux analysis in mature crops using 3D root mesocosms. *Frontiers in Plant Science*, 13(12), 1–19. https://doi.org/10.3389/fpls.2022.1041404

Duncan, K. E., Bray, A. L., Dowd, T. G., & Topp, C. N. (2019). Using 3D X-ray Microscopy to Study Crown Root Development and Primary Root Tip Growth in Diverse Maize (Zea mays L.) Lines. *Microscopy and Microanalysis*, 25(2), 1032–1033. https://doi.org/10.1017/S1431927619005890

Feix, G., Hochholdinger, F., & June Park, W. (2002). Maize Root System and Genetic Analysis of Its Formation. *Plant Roots*, 239–248. https://doi.org/10.1201/9780203909423.pt3

Feng, X., Jia, L., Cai, Y., Guan, H., Zheng, D., Zhang, W., Xiong, H., Zhou, H., Wen, Y., Hu, Y., Zhang, X., Wang, Q., Wu, F., Xu, J., & Lu, Y. (2022). ABA-inducible DEEPER ROOTING 1 improves adaptation of maize to water deficiency. *Plant Biotechnology Journal*, 20(11), 2077–2088. https://doi.org/10.1111/pbi.13889

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. https://doi.org/DOI: 10.1017/S0080456800012163

Fisher, R. A. (1930). The genetical theory of natural selection. In The genetical theory of natural selection. Clarendon Press. https://doi.org/10.5962/bhl.title.27468

Fitter, A. H. (1987). an Architectural Approach To the Comparative Ecology of Plant Root Systems. *New Phytologist*, *106*, 61–77. https://doi.org/10.1111/j.1469-8137.1987.tb04683.x

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third). Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Gamuyao, R., Chin, J. H., Pariasca-Tanaka, J., Pesaresi, P., Catausan, S., Dalid, C., Slamet-Loedin, I., Tecson-Mendoza, E. M., Wissuwa, M., & Heuer, S. (2012). The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. *Nature*, *488*(7412), 535–539. https://doi.org/10.1038/nature11346

Gaudin, A. C. M., McClymont, S. A., Holmes, B. M., Lyons, E., & Raizada, M. N. (2011). Novel temporal, fine-scale and growth variation phenotypes in roots of adult-stage maize (Zea mays L.) in response to low nitrogen stress. *Plant, Cell & Environment*, *34*(12), 2122–2137. https://doi.org/https://doi.org/10.1111/j.1365-3040.2011.02409.x

Gu, Z. (2022). Complex Heatmap Visualization. *IMeta*. https://doi.org/10.1002/imt2.43

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btw313

Helliwell, J. R., Sturrock, C. J., Mairhofer, S., Craigon, J., Ashton, R. W., Miller, A. J., Whalley, W. R., & Mooney, S. J. (2017). The emergent rhizosphere: imaging the development of the porous architecture at the root-soil interface. *Scientific Reports*, *7*(1), 14875. https://doi.org/10.1038/s41598-017-14904-w

Hill, W. G., Edwards, M. R., Ahmed, M.-K. A., & Thompson, R. (1983). Heritability of milk yield and composition at different levels and variability of production. *Animal Science*, 36(1), 59–68. https://doi.org/10.1017/S0003356100039933

Hochholdinger, F., & Feix, G. (1998). Early post-embryonic root formation is specifically affected in the maize mutant lrt1. *The Plant Journal For Cell and Molecular Biology*, 16(2), 247–255. https://doi.org/10.1046/j.1365-313x.1998.00280.x

Hochholdinger, F., Yu, P., & Marcon, C. (2018). Genetic Control of Root System Development in Maize. *Trends in Plant Science*, 23(1), 79–88. https://doi.org/10.1016/j.tplants.2017.10.004

Hoffmann, A. A., & Merilä, J. (1999). Heritable variation and evolution under favourable and unfavourable conditions. *Trends in Ecology & Evolutio*n, 14(3), 96–101. https://doi.org/https://doi.org/10.1016/S0169-5347(99)01595-5

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417–441. https://doi.org/10.1037/h0071325

Houle, D., Govindaraju, D. R., & Omholt, S. (2010). Phenomics: The next challenge. *Nature Reviews Genetics*, 11(12), 855–866. https://doi.org/10.1038/nrg2897

Ilic, K., Kellogg, E. A., Jaiswal, P., Zapata, F., Stevens, P. F., Vincent, L. P., Avraham, S., Reiser, L., Pujar, A., Sachs, M. M., Whitman, N. T., McCouch, S. R., Schaeffer, M. L., Ware, D. H., Stein, L. D., & Rhee, S. Y. (2007). The plant structure ontology, a unified vocabulary of anatomy and morphology of a flowering plant. *Plant Physiology*, *143*(2), 587–599. https://doi.org/10.1104/pp.106.092825

Jenkins, M. (1930). Heritable characters of maize. *Journal of Heredity*, 21(2), 79–80. https://doi.org/10.1093/oxfordjournals.jhered.a102019

Johnson, S. N., & Rasmann, S. (2015). Root-Feeding Insects and Their Interactions with Organisms in the Rhizosphere. *Annual Review of Entomology*, 60, 517–535. https://doi.org/https://doi.org/10.1146/annurev-ento-010814-020608

Ju, Y., Liu, A. E., Oestreich, K., Wang, T., Topp, C. N., & Ju, T. (2024). TopoRoot+: computing whorl and soil line traits of field-excavated maize roots from CT imaging. *Plant Methods*, 20(1), 132. https://doi.org/10.1186/s13007-024-01240-0

Jung, J. K. H., & McCouch, S. (2013). Getting to the roots of it: Genetic and hormonal control of root architecture. In *Frontiers in Plant Science*, 4(6), 186. https://doi.org/10.3389/fpls.2013.00186

Karlova, R., Boer, D., Hayes, S., & Testerink, C. (2021). Root plasticity under abiotic stress. *Plant Physiology*, 187(3), 1057–1070. https://doi.org/10.1093/plphys/kiab392

Kassambara, A., & Mundt, F. (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. https://cran.r-project.org/package=factoextra

Klei, L., Luca, D., Devlin, B., & Roeder, K. (2008). Pleiotropy and principal components of heritability combine to increase power for association analysis. *Genetic Epidemiology*, 32(1), 9–19. https://doi.org/10.1002/gepi.20257

Lareen, A., Burton, F., & Schäfer, P. (2016). Plant root-microbe communication in shaping root microbiomes. *Plant Molecular Biology*, 90(6), 575–587. https://doi.org/10.1007/s11103-015-0417-8

Lê, S., Josse, J., & Husson, F. (2008). {FactoMineR}: A Package for Multivariate Analysis. *Journal of Statistical Software*, 25(1), 1–18. https://doi.org/10.18637/jss.v025.i01

Leiboff, S., Li, X., Hu, H. C., Todt, N., Yang, J., Li, X., Yu, X., Muehlbauer, G. J., Timmermans, M. C. P., Yu, J., Schnable, P. S., & Scanlon, M. J. (2015). Genetic control of morphometric diversity in the maize shoot apical meristem. *Nature Communications*, 6, 1–10. https://doi.org/10.1038/ncomms9974

Li, M., Frank, M. H., Coneva, V., Mio, W., Chitwood, D. H., & Topp, C. N. (2018). The persistent homology mathematical framework provides enhanced genotype-to-phenotype associations for plant morphology. *Plant Physiology*, 177(4), 1382–1395. https://doi.org/10.1104/pp.18.00104

Li, M., Liu, Z., Jiang, N., Laws, B., Tiskevich, C., Moose, S. P., & Topp, C. N. (2023). Topological data analysis expands the genotype to phenotype map for 3D maize root system architecture. *Frontiers in Plant Science*, 14(1), 1–13. https://doi.org/10.3389/fpls.2023.1260005

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., & Zhang, Z. (2012). GAPIT: Genome association and prediction integrated tool. *Bioinformatics*, 28(18), 2397–2399. https://doi.org/10.1093/bioinformatics/bts444

Liu, S., Barrow, C. S., Hanlon, M., Lynch, J. P., & Bucksch, A. (2021). DIRT/3D: 3D root phenotyping for field-grown maize (Zea mays). *Plant Physiology*, 187(2), 739–757. https://doi.org/10.1093/plphys/kiab311

Liu, X., Huang, M., Fan, B., Buckler, E. S., & Zhang, Z. (2016). Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLOS Genetics*, 12(2), 1–24. https://doi.org/10.1371/journal.pgen.1005767

Lobet, G. (2017). Image Analysis in Plant Sciences: Publish Then Perish. *Trends in Plant Science*, 22(7), 559–566. https://doi.org/10.1016/j.tplants.2017.05.002

Lynch, J. P., & Brown, K. M. (2001). Topsoil foraging – an architectural adaptation of plants to low phosphorus availability. *Plant and Soil*, 237(2), 225–237. https://doi.org/10.1023/A:1013324727040

Lynch, J. P., Chimungu, J. G., & Brown, K. M. (2014). Root anatomical phenes associated with water acquisition from drying soil: targets for crop improvement. *Journal of Experimental Botany*, 65(21), 6155–6166. https://doi.org/10.1093/jxb/eru162

Lynch, M., & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer. https://books.google.com/books?id=UhCCQgAACAAJ

Macgregor, S., Cornes, B. K., Martin, N. G., & Visscher, P. M. (2006). Bias, precision and heritability of self-reported and clinically measured height in Australian twins. *Human Genetics*, 120(4), 571–580. https://doi.org/10.1007/s00439-006-0240-z

Mattupalli, C., Seethepalli, A., York, L. M., & Young, C. A. (2018). Digital imaging to evaluate root system architectural changes associated with soil biotic factors. *BioRxiv*. https://api.semanticscholar.org/CorpusID:92292738

McKay, J. K., Alvarado, C. A., Babbitt, G. R., Palmiscno, K. J., Turner, C. W., & Whitehill, B. C. (2023). *Device for automated crop root sampling*. https://www.osti.gov/biblio/2221983

Mooney, S. J., Pridmore, T. P., Helliwell, J., & Bennett, M. J. (2012). Developing X-ray Computed Tomography to non-invasively image 3-D root systems architecture in soil. *Plant and Soil*, 352(1), 1–22. https://doi.org/10.1007/s11104-011-1039-9

Morris, E. C., Griffiths, M., Golebiowska, A., Mairhofer, S., Burr-Hersey, J., Goh, T., von Wangenheim, D., Atkinson, B., Sturrock, C. J., Lynch, J. P., Vissenberg, K., Ritz, K., Wells, D. M., Mooney, S. J., & Bennett, M. J. (2017). Shaping 3D Root System Architecture. *Current Biology*, 27(17), R919–R930. https://doi.org/10.1016/j.cub.2017.06.043

Ortman, E. E., Peters, D. C., & Fitzgerald, P. J. (1968). Vertical-Pull Technique for Evaluating Tolerance of Corn Root Systems to Northern and Western Corn Rootworms123. *Journal of Economic Entomology*, 61(2), 373–375. https://doi.org/10.1093/jee/61.2.373

Péret, B., De Rybel, B., Casimiro, I., Benková, E., Swarup, R., Laplaze, L., Beeckman, T., & Bennett, M. J. (2009). Arabidopsis lateral root development: an emerging story. *Trends in Plant Science*, 14(7), 399–408. https://doi.org/10.1016/j.tplants.2009.05.002

Peterson, Ryan A. (2021). Finding Optimal Normalizing Transformations via bestNormalize. *The R Journal*, 13(1), 310–329. https://doi.org/10.32614/RJ-2021-041

Peterson, Ryan A, & Cavanaugh, J. E. (2020). Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. *Journal of Applied Statistics*, 47(13–15), 2312–2327. https://doi.org/10.1080/02664763.2019.1630372

Peterson, Ryan Andrew. (2023). *bestNormalize: Normalizing Transformation Functions*. https://cran.r-project.org/package=bestNormalize

Petricka, J. J., Winter, C. M., & Benfey, P. N. (2012). Control of Arabidopsis Root Development. *Annual Review of Plant Biology*, 63, 563–590. https://doi.org/https://doi.org/10.1146/annurev-arplant-042811-105501

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Ren, W., Zhao, L., Liang, J., Wang, L., Chen, L., Li, P., Liu, Z., Li, X., Zhang, Z., Li, J., He, K., Zhao, Z., Ali, F., Mi, G., Yan, J., Zhang, F., Chen, F., Yuan, L., & Pan, Q. (2022). Genome-wide dissection of changes in maize root system architecture during modern breeding. *Nature Plants*, 8(12), 1408–1422. https://doi.org/10.1038/s41477-022-01274-z

Rice, B. R., Fernandes, S. B., & Lipka, A. E. (2020). Multi-trait genome-wide association studies reveal loci associated with maize inflorescence and leaf architecture. *Plant and Cell Physiology*, 61(8), 1427–1437. https://doi.org/10.1093/pcp/pcaa039

Rogers, E. D., & Benfey, P. N. (2015). Regulation of plant root system architecture: Implications for crop advancement. *Current Opinion in Biotechnology*, 32, 93–98. https://doi.org/10.1016/j.copbio.2014.11.015

Saengwilai, P., Tian, X., & Lynch, J. P. (2014). Low Crown Root Number Enhances Nitrogen Acquisition from Low-Nitrogen Soils in Maize. *Plant Physiology*, 166(2), 581–589. https://doi.org/10.1104/pp.113.232603

Scheres, B., McKhann, H. I., & Van Den Berg, C. (1996). Roots Redefined: Anatomical and Genetic Analysis of Root Development. *Plant Physiology*, 111(4), 959–964. https://doi.org/10.1104/pp.111.4.959

Schneider, H. M., Lor, V. S. N., Hanlon, M. T., Perkins, A., Kaeppler, S. M., Borkar, A. N., Bhosale, R., Zhang, X., Rodriguez, J., Bucksch, A., Bennett, M. J., Brown, K. M., & Lynch, J. P. (2022). Root angle in maize influences nitrogen capture and is regulated by calcineurin B-like protein (CBL)-interacting serine/threonine-protein kinase 15 (ZmCIPK15). *Plant Cell and Environment*, 45(3), 837–853. https://doi.org/10.1111/pce.14135

Schneider, H. M., & Lynch, J. P. (2020). Should Root Plasticity Be a Crop Breeding Target? *Frontiers in Plant Science*, 11(5), 1–16. https://doi.org/10.3389/fpls.2020.00546

Shao, M. R., Jiang, N., Li, M., Howard, A., Lehner, K., Mullen, J. L., Gunn, S. L., McKay, J. K., & Topp, C. N. (2021). Complementary phenotyping of maize root system architecture by root pulling force and X-ray imaging. *Plant Phenomics*, 2021. https://doi.org/10.34133/2021/9859254

Stuber, C. W. (1995). Mapping and manipulating quantitative traits in maize. *Trends in Genetics*, 11(12), 477–481. https://doi.org/https://doi.org/10.1016/S0168-9525(00)89156-8

Sun, B., Gao, Y., & Lynch, J. P. (2018). Large Crown Root Number Improves Topsoil Foraging and Phosphorus Acquisition. *Plant Physiology*, 177(1), 90–104. https://doi.org/10.1104/pp.18.00234

Topp, C. N., Bray, A. L., Ellis, N. A., & Liu, Z. (2016). How can we harness quantitative genetic variation in crop root systems for agricultural improvement? *Journal of Integrative Plant Biology*, 58(3), 213–225. https://doi.org/10.1111/jipb.12470

Topp, C. N., Iyer-Pascuzzi, A. S., Anderson, J. T., Lee, C. R., Zurek, P. R., Symonova, O., Zheng, Y., Bucksch, A., Mileyko, Y., Galkovskyi, T., Moore, B. T., Harer, J., Edelsbrunner, H., Mitchell-Olds, T., Weitz, J. S., & Benfey, P. N. (2013). 3D phenotyping and quantitative trait locus mapping identify core regions of the rice genome controlling root architecture. *Proceedings of the National Academy of Sciences of the United States of America*, 110(18). https://doi.org/10.1073/pnas.1304354110

Tovar, J. C., Hoyer, J. S., Lin, A., Tielking, A., Callen, S. T., Elizabeth Castillo, S., Miller, M., Tessman, M., Fahlgren, N., Carrington, J. C., Nusinow, D. A., & Gehan, M. A. (2018). Raspberry Pi–powered imaging for plant phenotyping. *Applications in Plant Sciences*, 6(3), 1–12. https://doi.org/10.1002/aps3.1031

Trachsel, S, Kaeppler, S. M., Brown, K. M., & Lynch, J. P. (2013). Maize root growth angles become steeper under low N conditions. *Field Crops Research*, 140, 18–31. https://doi.org/https://doi.org/10.1016/j.fcr.2012.09.010

Trachsel, Samuel, Kaeppler, S. M., Brown, K. M., & Lynch, J. P. (2011). Shovelomics: high throughput phenotyping of maize (Zea mays L.) root architecture in the field. *Plant and Soil*, 341(1), 75–87. https://doi.org/10.1007/s11104-010-0623-8

Ubbens, J., Cieslak, M., Prusinkiewicz, P., Parkin, I., Ebersbach, J., & Stavness, I. (2020). Latent space phenotyping: Automatic image-based phenotyping for treatment studies. *Plant Phenomics*. https://doi.org/10.34133/2020/5801869

Ueda, M., Koshino-Kimura, Y., & Okada, K. (2005). Stepwise understanding of root development. *Current Opinion in Plant Biology*, *8*(1), 71–76. https://doi.org/10.1016/j.pbi.2004.11.014

Uga, Y., Yamamoto, E., Kanno, N., Kawai, S., Mizubayashi, T., & Fukuoka, S. (2013). A major QTL controlling deep rooting on rice chromosome 4. *Scientific Reports*, 3. https://doi.org/10.1038/srep03040

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era - Concepts and misconceptions. *Nature Reviews Genetics*, 9(4), 255–266. https://doi.org/10.1038/nrg2322

Wei, T., & Simko, V. (2021). *R package "corrplot": Visualization of a Correlation Matrix* (Version 0.92). https://github.com/taiyun/corrplot

Wen, T.-J., Hochholdinger, F., Sauer, M., Bruce, W., & Schnable, P. S. (2005). The roothairless1 gene of maize encodes a homolog of sec3, which is involved in polar exocytosis. *Plant Physiology*, 138(3), 1637–1643. https://doi.org/10.1104/pp.105.062174

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Woods, P., Lehner, K. R., Hein, K., Mullen, J. L., & McKay, J. K. (2022). Root Pulling Force Across Drought in Maize Reveals Genotype by Environment Interactions and Candidate Genes. *Frontiers in Plant Science*, 13(4), 1–14. https://doi.org/10.3389/fpls.2022.883209

York, L. M., & Lobet, G. (2017). Phenomics of Root System Architecture : Measuring and Analyzing Root Phenes. *The Plant Cell*, (9), 1–9. https://doi.org/10.1105/tpc.17.tt0917

York, L. M., & Lynch, J. P. (2015). Intensive field phenotyping of maize (Zea mays L.) root crowns identifies phenes and phene integration associated with plant growth and nitrogen acquisition. *Journal of Experimental Botany*, 66(18), 5493–5505. https://doi.org/10.1093/jxb/erv241

York, L. M., Nord, E. A., & Lynch, J. P. (2013). Integration of root phenes for soil resource acquisition. *Frontiers in Plant Science*, 4(9), 1–15. https://doi.org/10.3389/fpls.2013.00355

Zeng, D., Li, M., Jiang, N., Ju, Y., Schreiber, H., Chambers, E., Letscher, D., Ju, T., & Topp, C. N. (2021). TopoRoot: a method for computing hierarchy and fine-grained traits of maize roots from 3D imaging. *Plant Methods*, 17(1), 127. https://doi.org/10.1186/s13007-021-00829-z

Zobel, R. W., & Waisel, Y. (2010). A plant root system architectural taxonomy: A framework for root nomenclature. *Plant Biosystems - An International Journal Dealing with All Aspects of Plant Biology*, 144(2), 507–512. https://doi.org/10.1080/11263501003764483

Zurek, P. R., Topp, C. N., & Benfey, P. N. (2015). Quantitative trait locus mapping reveals regions of the maize genome controlling root system architecture. *Plant Physiology*, 167(4), 1487–1496. https://doi.org/10.1104/pp.114.251751

CHAPTER 4

S-TYPE ANION CHANNEL HOMOLOGUE (*ZMSLAH*) REGULATES ROOT-TO-SHOOT PHENOTYPIC PLASTICITY AND NITROGEN CAPTURE IN MAIZE

Summary

Root systems are complex, consisting of numerous individual root phenotypes, each influenced by distinct genetic factors and interactions with the surrounding environment. Therefore, measurements of root systems grown under field conditions are essential for detecting genotype-by-environment interactions (G × E) in mapping studies, as root phenotyping under controlled conditions may yield adaptive values and effect sizes that are not translatable to agronomic environments. Genome-wide association mapping identified Zm00001eb159490 as a putative S-type anion channel homologue (*ZmSLAH*) and gene regulator of maize (*Zea mays* L.) root system size. Using 3D phenotyping techniques, we found that the allelic variation at *ZmSLAH* exhibited constitutive differences in root architecture among genotypes in low-stress environments across the 2018 and 2019 field seasons. A mutator (Mu) transposon was identified in the coding sequence of *ZmSLAH* and used to generate a maize population that segregates for a loss-of-function mutation in the W22 genetic background. Field evaluations of wild-type and mutant plants grown under high- and low-nitrogen conditions revealed contrasting patterns in root architecture, nitrogen capture, and biomass accumulation. We further found that the wild-type plants had greater biomass accumulation and leaf nitrogen content under nitrogen-limited conditions compared to the mutant plants. Additional functional genetics will be used to characterize the molecular basis of *ZmSLAH* through heterologous gene expression and

electrophysiological analyses in *Xenopus* oocytes. Based on these findings, we report *ZmSLAH* as a gene modulator of plant responses to edaphic stress in maize.

Introduction

The global population is projected to reach 9.7 billion by 2050, placing unprecedented pressure on agricultural systems to meet rising food demands (UN DESA/POP, 2024). Simultaneously, arable land is diminishing due to urbanization, soil degradation, and climate change, necessitating innovative approaches to enhance crop productivity (FAO, 2023; Pretty, 2008). Maize (*Zea mays* L.), as one of the world's most important cereal crops, plays a crucial role in addressing these challenges. Root traits are major targets for the second green revolution because of their potential to improve crop productivity, increase drought tolerance, and increase capture of carbon in the soil (Kell, 2013; Lynch, 1998). Therefore, a deeper understanding of the genetic and environmental drivers of maize resource utilization in root systems and the genetic and environmental drivers is essential for future breeding efforts aimed at improving crop resiliency in production environments.

Phenotypic plasticity is generally referred to as the ability to which individuals sense and respond to the environment (Falconer, 1990; Pigliucci, 2001). This variation at the population-level, known as genotype-by-environment interactions (G × E), is heritable and can evolve through selection (Agrawal, 2001; Bradshaw, 2006; DeWitt et al., 1998; Lynch & Walsh, 1998; Sultan, 2000). Large-scale phenomic efforts have greatly improved our ability to map phenotype-genotype relationships and identify G × E in maize. For example, the Maize G × E Project, part of the Genomes to Fields (G2F) initiative, has measured maize hybrid phenotypes across diverse North American environments (www.genomes2fields.org). As an ongoing experiment, it provides a unique platform for investigating phenotypic plasticity and adaptation (Gage et al.,

2017). However, these efforts have largely neglected belowground phenotypic variation, as field phenotyping remains a major bottleneck in scaled experiments.

Root system architecture (RSA) and function are key determinants of plant fitness, serving as the primary site for soil resource acquisition, anchorage, and microbial interactions. Specific root traits, such as root growth angle, root biomass, and depth, have been shown across crop species to facilitate the capture of mobile nutrients, such as water and nitrate (NO_3^-), as well as immobile nutrients, such as phosphorus, potassium, and micronutrients, in distinct soil domains (Bonser et al., 1996; Lynch et al., 2013; Saengwilai et al., 2014; Schneider & Lynch, 2020; Schneider et al., 2022; Trachsel et al., 2013). For example, shallower root systems with more horizontal growth angles are better suited for accessing immobile nutrients that are more abundant in the upper soil layers, whereas deeper root systems with steep growth angles are advantageous for capturing mobile nutrients that tend to accumulate in deeper soil layers over time. When these phenotypic changes occur in response to resource availability, this can be described by the functional equilibrium theory (Poorter et al., 2012). The theory of functional equilibrium, also known as optimal partitioning theory, postulates that plants regulate the distribution of carbon between roots and leaves based on soil resource availability. This process ensures that plant resources are allocated to organs responsible for acquiring the limiting resource, demonstrating active plasticity (Brouwer, 1962; Lambers, 1983). Through active plasticity, plants modulate their morphology or physiology to optimize resource utilization in their environment.

Forward genetic studies have identified numerous quantitative trait loci (QTL) associated with RSA in maize, but validation of candidate gene function under agronomic conditions is still needed. A more recent example is the work by Schneider et al. (2022), who identified *ZmCIPK15*

as a regulator of maize root crown angle under field conditions. Their study showed that differences in the growth angle of *ZmCIPK15* mutant plants grown under nitrogen-limited conditions led to greater nitrogen acquisition and increased plant biomass compared with wild-type plants, suggesting an adaptive response aimed at improving nitrogen capture (Trachsel et al., 2013). Previously, we identified the gene Zm00001eb159490, which encodes a putative slow-type anion channel (SLAC), as a potential regulator of root system size through field-based measurements of root pulling force (RPF). This candidate was detected to have a significant effect of G × E in a diverse population of maize inbred lines subjected to differential irrigation in 2018. SLAC1 (slow anion channel-associated 1) and its homologues (SLAH2, SLAH3, SLAH4) are primarily responsible for controlling the transport of anions, particularly chloride (Cl^-) and NO_3^-. SLAC1 regulates these ions in guard cells, facilitating stomatal closure, while SLAH channels are involved in NO_3^- and Cl^- uptake in the roots and their translocation to the shoots (Dreyer et al., 2012; Geiger et al., 2011; Negi et al., 2008; Vahisalu et al., 2008). Notably, research has shown that SLAC1 coordinates signals from CO_2, abscisic acid (ABA) , water stress, and nutrient availability via distinct protein kinase pathways (Geiger et al., 2009; Müller et al., 2017; Sun et al., 2016). Nevertheless, our understanding of S-type anion channel functionality in maize remains limited.

In this study, we posited that the alternative states of the candidate gene (*ZmSLAH*) lead to varying signals of nutrient deprivation and supplementation, which would subsequently influence distinct root growth patterns in field-grown maize. The goal of this study was to provide functional evidence for *ZmSLAH* as a gene regulator of root system size and nitrogen capture in maize. We sought to achieve this by utilizing (i) a maize mapping population and 3D

phenotyping techniques to identify constitutive RSA QTLs and (ii) a maize population segregating for a loss-of-function mutation in the *ZmSLAH* gene model.

Materials and Methods

GWAS Follow-up Study

We followed up on the 2019 genome-wide association study (GWAS) conducted in Chapter 3 to identify constitutive root QTLs co-located with our candidate gene using 3D root models. To review, GWAS was performed using 99 3D RSA phenotypes and 312 inbred lines from the shoot apical meristem (SAM) maize diversity panel using the GAPIT implementation of FarmCPU (Leiboff et al., 2015; Lipka et al., 2012; Liu et al., 2016). The 3D RSA measurements, reported as genotype averages, were obtained using X-ray computed tomography (XRT) and a custom image processing and feature extraction pipeline (Donald Danforth Plant Science Center, St. Louis, MO, USA), as previously described (Shao et al., 2021), referred to as the Root Crown Analysis Pipeline (RCAP).

Our study focused on a 1-Mb genomic region centered at the position of the original v2 SNP (chr3:217,665,275) that identified the candidate gene. QTLs that were identified within this genomic region were considered to be co-located and were further evaluated in parallel with the polymorphism at our candidate gene for their *post hoc* effects on RPF and 3D RSA in the 2019 field experiment. Linear models were fitted for treating SNP, treatment, and SNP-by-treatment interactions as fixed effects. The population structure among genotypes was controlled in each model by treating the first three principal components of genotype PCA as individual fixed effects. Statistical testing was performed using a type II two-way analysis of variance (ANOVA), followed by *post hoc* Tukey's honestly significant (HSD) test to assess multiple comparisons of means at a significance threshold of 0.05 using R/car (Fox & Weisberg, 2019).

Plant Materials

Seeds of sib-pollinated F3 plants for the UFmu-01876 line, from the UniformMu transposon insertion population (McCarty et al., 2005) were obtained from the Maize Genetics Cooperation Stock Center (MGCSC, Urbana, IL). Plants were backcrossed to the W22 wild-type (MGCSC stock number X17EA), and the resulting plants were self-pollinated to develop a BCF3 population of lines segregating for the insertion mu1014502.

Characterization of the Insertion Location

Genomic DNA was extracted from lyophilized leaf tissue of individual plants across generations to determine the presence and insertion locations of mutator (Mu) elements in the W22 gene model Zm00004b019600. PCR was conducted using gene-specific primers (5'-GTCTCTATACCCCTGCGTGC-3') and (5'-AGATCTTGGAGTCCGGGGAT-3'), in combination with the Mu terminal inverted repeat specific primer TIR6 described by Settles et al (2004). Primer oligos were procured from Integrated DNA Technologies, Inc. (Coralville, IA, USA). Sanger sequencing of the amplicons was performed by Azenta, Inc. (Burlington, MA, USA) to determine the location of the insertion. Preliminary protein analyses involved querying homology using blastp at NCBI (http://www.ncbi.nlm.nih.gov/blast/) and predicting domains using InterProScan (Paysan-Lafosse et al., 2023). Phylogenetic tree construction for Zm00004b019600 was performed using the Ensembl Compara Gene Tree tool in the Gramene database (Tello-Ruiz et al., 2016; http://www.gramene.org; Gramene database release 4; August 13, 2024).

Experimental Design

Maize mutant and wild-type allele lines were evaluated at the Colorado State University Agricultural Research Development and Education Center in Fort Collins, CO, USA (40.653 N, -

104.993 W) over two consecutive years (2023-2024). Each year, the seeds were sown in mid-May at a 22-cm spacing in replicated single row, 6.1-meter plots with a row spacing of 76 cm. Throughout the growing season, the plots received approximately 2.5 cm of weekly irrigation and were regularly hand-weeded.

A preliminary field study was conducted in the summer of 2023 to evaluate the effects of genotype on maize RSA under normal conditions. Fertilizer was applied according to the recommendations for a targeted yield of 200 bu/ac, amounting to 170 lb/ac N prior to planting. Plants were harvested and phenotyped during the mid-flowering stage, approximately 14-weeks after planting.

A replicated nitrogen stress experiment was conducted in the summer of 2024 with three field replicates of mutant and wild-type genotypes per treatment (high N and low N). Plots in the high N treatment were fertilized at the same rate as in the previous year, receiving 170 lb/ac N prior to planting. For the low N treatment, fertilizer application was withheld for the duration of the experiment. Plants were harvested and phenotyped during the mid-flowering stage, approximately 12-weeks after planting.

Mutant Phenotyping

Intact whole-root systems of mutant and wild-type maize genotypes were extracted from the soil using a combination of nondestructive excavation methods. In the 2023 field study, we used a custom-built, specialized root-pulling device (CZero Inc., Fort Collins, CO, USA) and attached it to a high-clearance tractor. This mechanical pulling device uses a camera to identify maize stalks and then vertically extract the plants from the soil. During the extraction process, the load cell measured the applied vertical force, and the maximum force was denoted as the RPF. Similar to the handheld force gauge measurements, the RPF recorded by the mechanical

root puller was measured in kg. Concurrently, additional roots were extracted manually by shoveling. Following excavation, root crowns collected from both studies were hand-washed with water and air-dried by hanging. After air-drying to a constant weight, shoot and root biomass were recorded in grams.

Image-based RSA phenotyping was performed on excavated maize root crowns to obtain a comprehensive dataset of 154 RSA measurements. In both years, two-dimensional (2D) root imaging was conducted by vertically suspending the root samples in front of a black backdrop and capturing images using a DSLR camera. These images were processed and analyzed for RSA features using 2D DIRT image analysis software (Das et al., 2015). In the following field season, we introduced a 3D root imaging and analysis pipeline using XRT and RCAP to analyze RSA variation in mutant and wild-type maize lines (Donald Danforth Plant Science Center, St. Louis, MO, USA), as previously described (Galkovskyi et al., 2012; Topp et al., 2013).

We performed nitrogen elemental analysis to determine the effect of genotype and environmental nitrogen on leaf nitrogen allocation across individuals during mid-flowering in the 2024 experiment. Leaf samples were collected from three representative plants in each plot. A tissue punch was used to obtain 10 individual ¼-inch disks from the topmost fully expanded leaf of each plant. These leaf samples were placed in 2-ml microcentrifuge tubes and transported on ice, where they were immediately lyophilized for 48 h. Following lyophilization, the samples were pulverized, and the nitrogen concentration was measured using an elemental analyzer (LECO TruSpec CN).

Statistical Analysis

Statistical analyses were performed using R version 4.3.0 (R Core Team, 2023). Phenotypic data were collected from genotyped individuals across homozygous wild-type and

mutant genotypes, two nitrogen treatments (high N and low N), and two field seasons (2023 and 2024). For each year, trait outliers were identified and excluded based on visual comparison of the mean and median values within each treatment/genotype combination. In the 2023 field study, average genotype values were estimated across replicates. Statistical comparisons among average genotype values were performed using two-sample t-tests at a significance threshold of 0.05 in R/car (Fox & Weisberg, 2019). In the 2024 experiment, we performed a type III two-way ANOVA to test the effects of genotype, treatment, and their interactions on average trait values using a significance threshold of 0.05 in R/car (Fox & Weisberg, 2019). The data structure was assessed for normality using diagnostic plots to ensure that it met the ANOVA assumptions before testing. For each model containing significant fixed effects, *post hoc* Tukey's tests were performed to assess the significance of differences among the genotypes at each treatment level.

Results

Mu transposon induces loss-of-function mutation in maize SLAC1-like domain

Woods et al. (2022) identified Zm00001eb159490 as a candidate gene for regulating root system size. This RPF-associated polymorphism was a synonymous polymorphism (C/T) located within an exon of the v2 (chr3:217,665,275) and v5 (chr3:223,191,889) B73 gene models. In our study, further evaluation of this gene candidate was conducted using the W22 genetic background (Zm00004b019600), which differs from the B73 reference. This candidate gene was one of 14 W22 maize paralogs and showed protein sequence similarities to Arabidopsis [*Arabidopsis thaliana* (L.) Heynh.] genes *AtSLAH2* and *AtSLAH3*, at 57.54% and 60.91%, respectively (Figure 4.1A). Protein sequence analysis predicted that this region in maize is a highly conserved Tellurite-resistance/Dicarboxylate Transporter (TDT) SLAC1-like domain

(residues 256-558; *cd09323*) within the SLAC/SLAH protein family (residues 173-584; *IPR030183*) (Figure 4.1B).

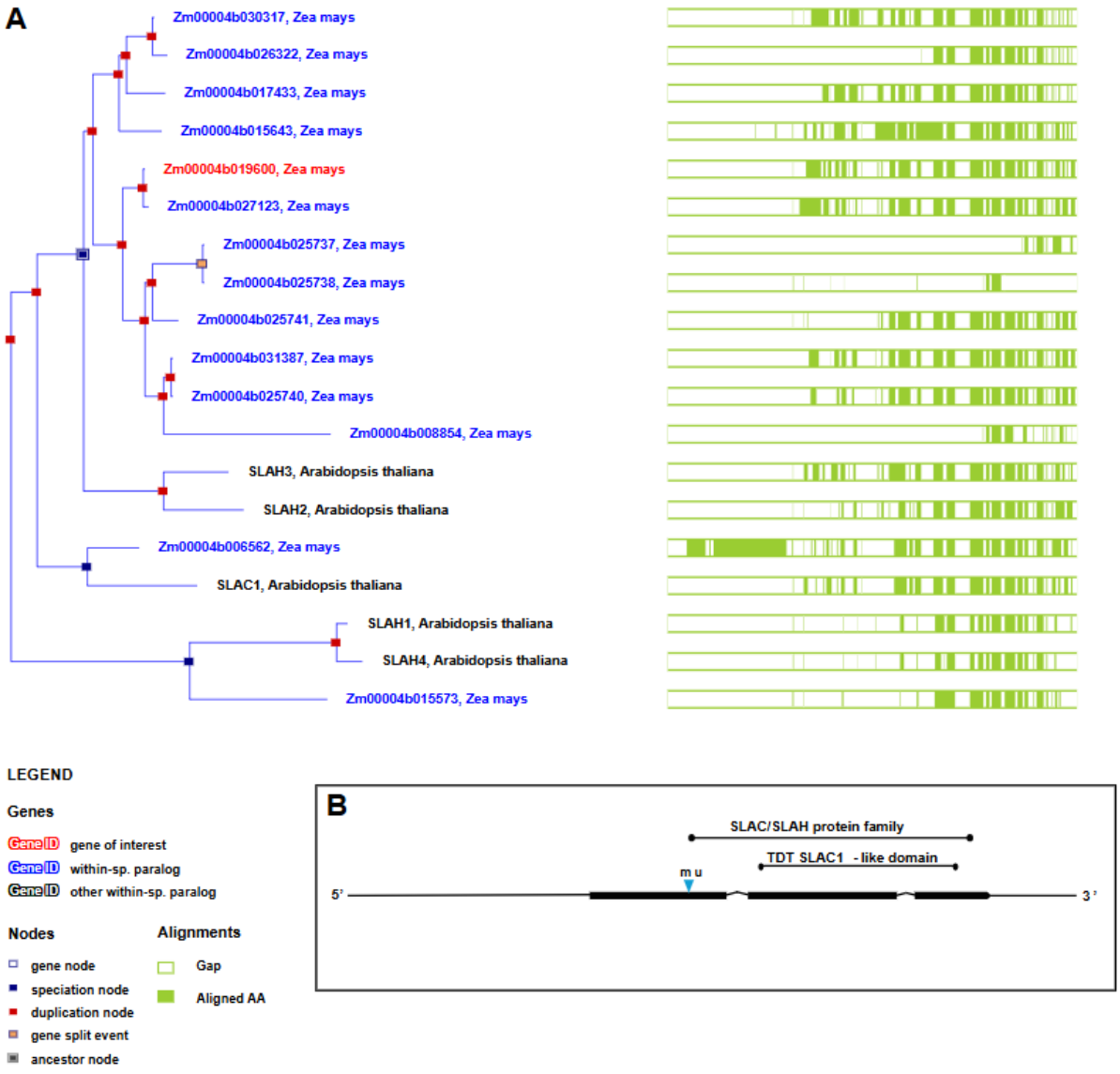

Figure 4.1. *ZmSLAH* gene model and sequence homology. (**A**) Phylogenetic tree showing the relationship between the maize gene Zm00004b019600 (highlighted in red), its closely related paralogs in the W22 genetic background (highlighted in blue), and homologs in *Arabidopsis thaliana*. The amino acid alignment to Zm00004b019600 is displayed to the right of the gene IDs, color-coded with green bars to indicate aligned regions. (**B**) Schematic representation of the exon-intron structure of *ZmSLAH* (Zm00004b019600) and the conserved TDT-SLAC1-like domain on chromosome 3. The conserved SLAC1/SLAH protein domain and protein family alignment are shown above. The Mu transposon insertion site (mu1014502) is marked by a blue triangle.

We obtained a maize mutant line, *ZmSLAH* (UFmu-01876), using a Mu transposon insertional mutation in Zm00004b019600, to generate a population of maize mutant and wild-type lines homozygous for *ZmSLAH*. This population was used to study the effects of the insertion on root development and its biological function in relation to its Arabidopsis homologs under field conditions. The presence of the insertion was confirmed by PCR and subsequently verified by Sanger sequencing. The size and position of the insertion were found to be approximately 3,000 bp in length, located within the first exon of the v2 W22 genome (chr3:226,299,935 - chr3:226,299,965; see Figure 4.1B) and within the second exon of the v5 B73 genome(chr3:223,191,719 - chr3:223,191,749). Genotyping and selection of homozygous *ZmSLAH* lines were performed in each generation until field testing.

GWAS analysis identifies a QTL co-located with ZmSLAH using 3D root features

Our FarmCPU GWAS model identified a single SNP positioned at chr3:217,841,675 that was significantly associated with 3D Density S6, a measure of root system density, under full irrigation. This SNP was located within an intergenic region approximately 176-kb upstream of the RPF-associated polymorphism at chr3:217,665,275. For both SNPs, *post hoc* ANOVA was conducted to assess the effects of SNP, treatment, and SNP-by-treatment interactions on RPF and 3D Density S6. The alternate allele at chr3:217,665,275 was found to be significantly associated ($p < 0.05$) with a higher RPF in the wet treatment, accounting for 12.32% of the phenotypic variance. Similarly, the polymorphism at chr3:217,841,675 was observed to have a significant effect of SNP on 3D Density S6, where the effect size of the alternate allele was greater in the wet environment, explaining 11.80% of the phenotypic variance. Although RPF and 3D Density S6 displayed weak phenotypic correlations (Spearman's $r = -0.24$), our analyses were able to

detect evidence of phenotypic plasticity at two proximal SNPs associated with RSA variation (Figure 4.2).

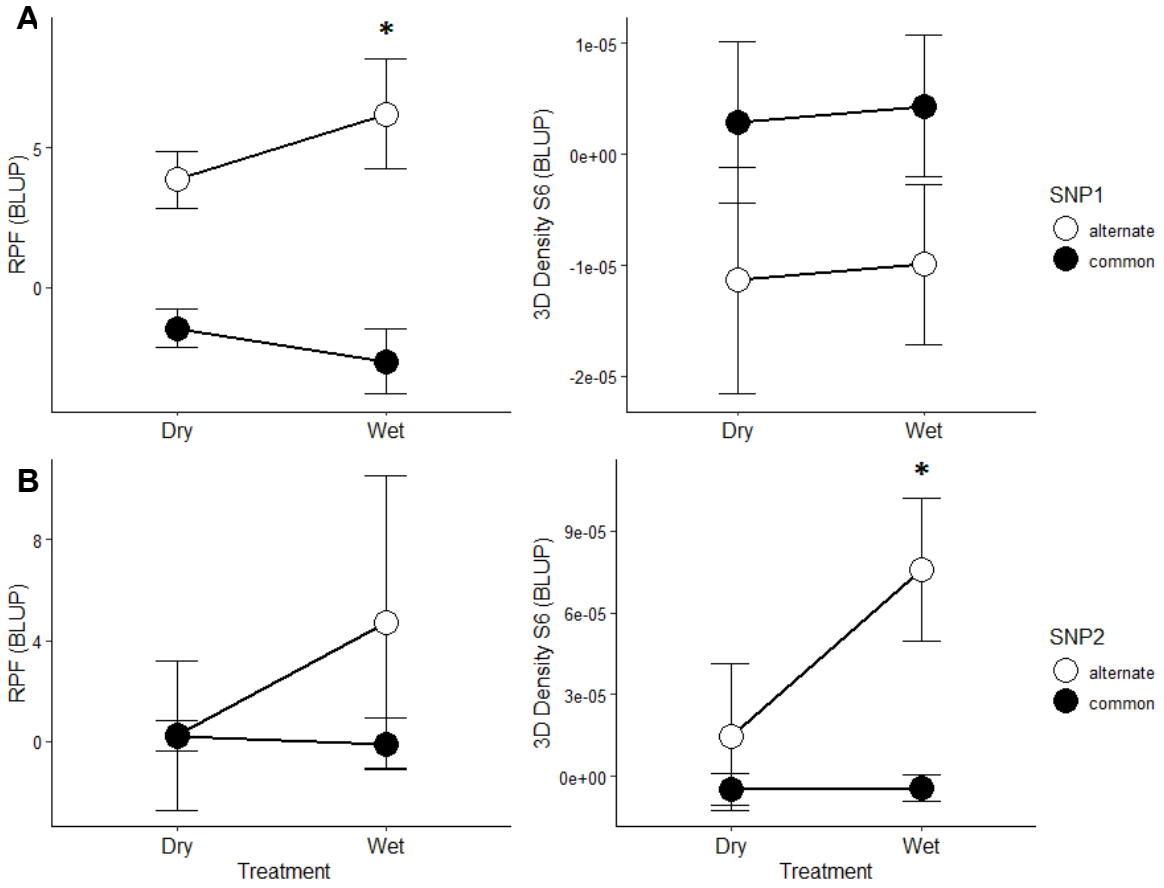

Figure 4.2. Reaction norms of RPF and 3D Density S6 BLUPs from 2019 mid-flowering measurements for lines with contrasting alleles at (**A**) SNP1 [v2 chr3:217,665,275] and (**B**) SNP2 [v2 chr3:217,841,675] across differential irrigation (mean ± SE). *Indicates a significant difference, $p < 0.05$.

ZmSLAH affects whole-plant-level phenotypic plasticity in response to environmental nitrogen

We evaluated the phenotypes of maize homozygous mutant and wild-type allele lines in a replicated nitrogen stress experiment under field conditions. Variance analysis revealed significant genotype effects on leaf nitrogen content ($p < 0.05$) and a nearly significant genotype-by-environment (G × E) interaction [$F(3, 68) = 3.42$, $p = 0.069$]. In the low N treatment, wild-type plants had a 7.61% higher leaf nitrogen content than mutant plants (Figure 4.3A).

Additionally, significant genotype effects were observed for both root and shoot biomass across nitrogen treatments ($p < 0.05$). On average, wild-type plants exhibited approximately 25.2% greater root biomass and 21.3% greater shoot biomass than mutant plants across treatments (Figure 4.3C-D). However, both genotypes showed a non-significant reduction in root-mass-ratio (RMR) when shifting from high to low N conditions (Figure 4.3B). Tukey's *post hoc* multiple comparisons confirmed that the genotypes differed significantly for root and shoot biomass traits ($p < 0.05$). These results suggest that *ZmSLAH* contributes to systemic biomass accumulation and nitrogen capture, particularly under low N conditions.

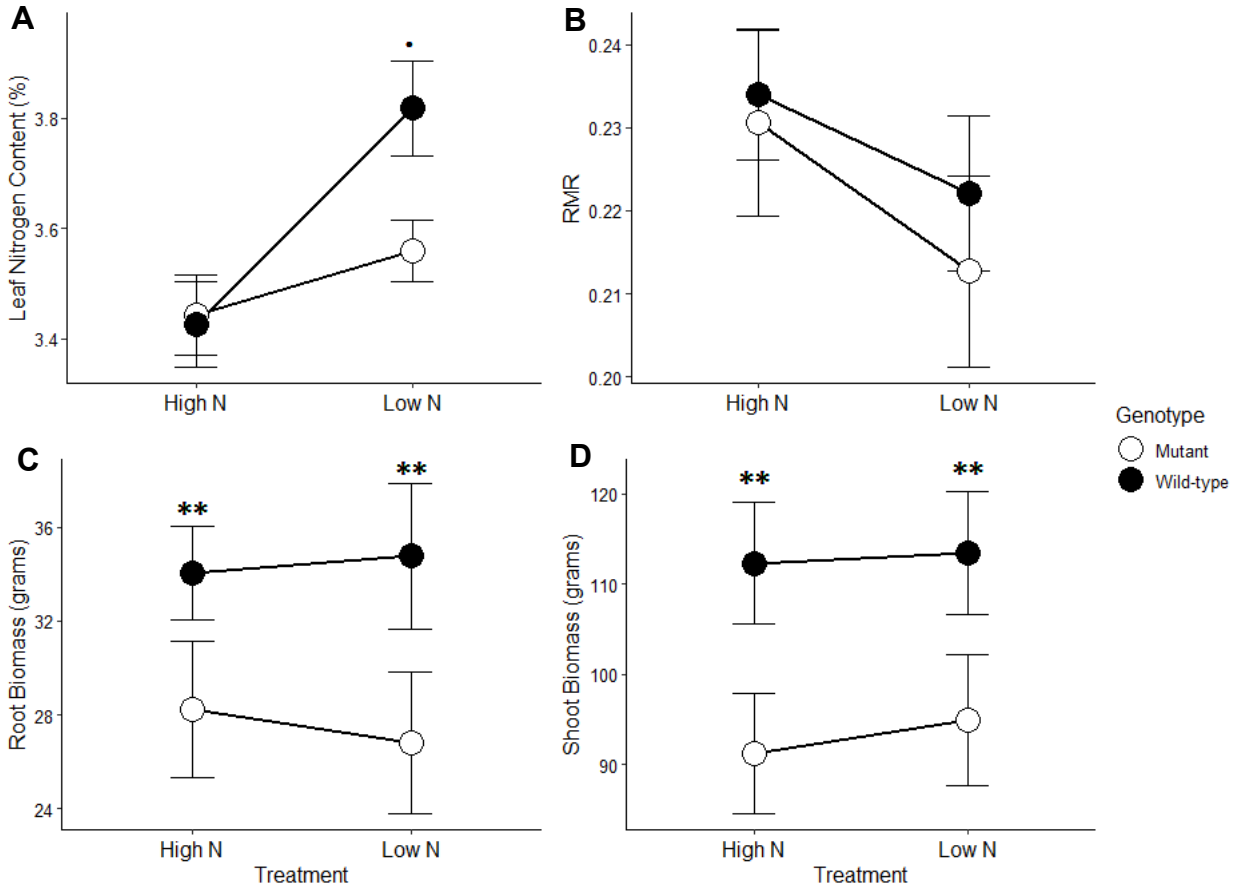

Figure 4.3. Phenotypic differences between *ZmSLAH* mutant and wild-type allele lines from the 2024 mid-flowering measurements of (**A**) leaf nitrogen content (%), (**B**) root-mass-ratio (RMR), (**C**) root biomass (grams), and (**D**) shoot biomass (grams). Significance levels: '.'$p < 0.10$, *$p < 0.05$, **$p < 0.01$

Measurements of RSA in wild-type and mutant plants revealed contrasting root growth patterns under high and low N conditions (Figure 4.4A). Mutant plants had higher global RSA trait values than wild-type plants under high N conditions. These results aligned with preliminary data from the previous field season, in which mutant plants showed higher average RPF (mean = 137 kg, SD = 31 kg) than wild-type plants (mean = 101 kg, SD = 32.3 kg) in the same environment. This variation was reflected in the 2D and 3D measurements of the RSA, where the mutant plants under high N conditions exhibited on average increased root width and number of root tips compared to the wild-type plants (Table 4.1). A significant genotype effect ($p < 0.05$) was detected for 2D root tip count, but this finding was not corroborated by 3D imaging data. The limited sample size and unequal genotype representation in the preliminary study likely confounded these results. Phenotypic variation in RSA between genotypes was larger in the low N treatment compared with the high N treatment. In Figure 4.4B, mutant plants demonstrated a consistent reduction in standardized average trait values (Z-scores) from high to low N conditions, whereas wild-type plants displayed a consistent increase in trait values, with less variability across treatments. Although variance analyses did not identify statistically significant differences in genotype or G × E effects ($p > 0.05$), they were indicative of root system plasticity in response to nitrogen stress.

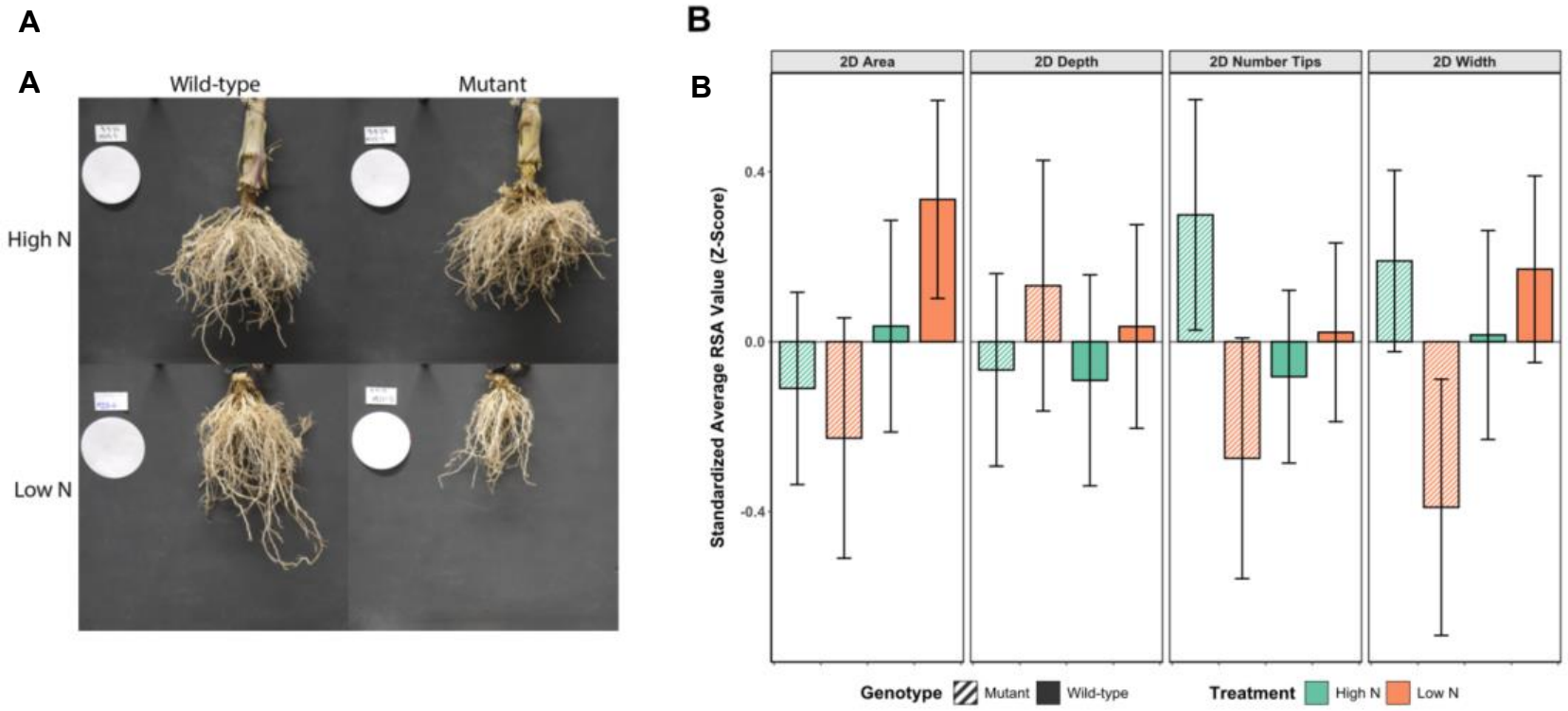

Figure 4.4. RSA plasticity among *ZmSLAH* mutant and wild-type allele lines for plants harvested at mid-flowering in the 2024 nitrogen stress field experiment. (**A**) Images of representative maize root crowns across genotypes and nitrogen treatments. (**B**) Bar plot comparing average standardized (Z-score) RSA trait measurements obtained through 2D imaging and DIRT feature extraction between genotypes and nitrogen treatments (mean ± SE). DIRT abbreviations: 2D Area, Area; 2D Depth, Skeleton Depth; 2D Number Tips, Root Tip Count; 2D Width, Maximum Width.

Table 4.1. Genotypic differences in root system architecture (RSA) phenotypes between *ZmSLAH* maize mutant and wild-type lines grown under normal field conditions in 2023.

		Wild-type		Mutant				
		M	SD	M	SD	*df*	*t*	*p*
Physical Traits	Root Biomass (g)	18.8	8.15	22.3	7.51	25	1.29	0.208
	RPF (kg)	100.57	38.89	137.04	29.38	7	2.09	0.071
2D[a] Imaging Traits	Area	8074	1991	9286	1978	21	-1.76	0.093
	Width	162.54	23.4	173.17	36.7	32	-1.09	0.285
	Depth	227.80	50.1	236.99	49.4	20	-0.53	0.601
	Number Tips	272	108	410	199	27	-2.44	0.021*
3D[b] Imaging Traits	Area	855.43	101	786.52	97.6	27	2.00	0.055
	Width	387.66	76	425.07	52.3	21	-1.61	0.123
	Depth	2111068	505245	1994004	454351	25	0.70	0.499
	Number Tips	1687.36	380	1703.76	423	30	-0.12	0.906

Abbreviations: M, Mean; SD, Standard Deviation; RPF, Root Pulling Force; 2D Width, Skeleton Width (SklWidth); 2D Depth, Skeleton Depth (SklDepth); 2D Number Tips, Root Tip Count (RTPCount), 3D Area, Surface Area (SurfArea); 3D Width, Maximum Width (HorEqDiameter)
*Welch test is statistically significant at $p \leq 0.05$
[a]DIRT
[b]RCAP

Heterologous gene expression and electrophysiological analysis of ZmSLAH paralogs in Xenopus oocytes: Further investigation of the function of ZmSLAH in maize nutrient signaling

In Winter 2024, we will report a novel characterization of the molecular function of *ZmSLAH* using *in vitro* assays in the laboratory. Oligo DNA strings were synthesized from the cDNA sequences of v2 B73 (Zm00001eb159490) and W22 (Zm00004b019600). The cDNAs were extended with 5'-GGCTTAAT-3' before the start codon and 5'-ATTAAACC-3' after the stop codon. These sequences were cloned into *Xenopus* oocyte expression vectors (pGEM) using a uracil-excision-based cloning technique, as described by Nour-Eldin et al. The dependence of channel properties were tested through the co-expression of W22/B73 *ZmSLAH* with different activating kinases, such as calcineurin B–like (CBL) calcium sensors and their interacting protein serine-threonine protein kinases (CIPKs), and *AtSLAH1* (At1G62280). For the double-electrode voltage-clamp studies, oocytes were perfused with Tris/Mes-based buffers composed of

NaCl and $NaNO_3$, testing for anion selectivity at different concentrations and cytosolic pH values. Electrophysiological measurements were recorded according to protocols established in the literature (Lehmann et al., 2021; Maierhofer, Diekmann, et al., 2014; Maierhofer, Lind, et al., 2014). Data collection, analysis and interpretation are currently in progress and will not be included as a result in this chapter.

Discussion

In Chapter 3, we demonstrate that 3D RSA phenotypes and low-stress environments capture the largest proportions of phenotypic and genotypic variation in our maize mapping population. Building on this, we sought to determine whether the QTL associated with RPF in 2018 (i.e., *ZmSLAH*) was constitutive across years and treatments using 3D imaging techniques. Although the genome-wide analysis in 2019 did not reveal a significant association between *ZmSLAH* and RPF, *post hoc* tests found a significant genetic effect on the wet treatment, which was consistent with our 2018 findings. Additionally, a SNP located approximately 176-kb upstream of *ZmSLAH*, which segregated for 3D Density S6 (representing the densest root on a scale from S1 to S6), also showed a significant association in the wet environment. Shao et al. (2021) reported that variations in root density distributions, particularly for 3D Density S5, accounted for the largest proportions of both environmental and genotypic variance, whereas RPF, though moderately heritable, explained a comparatively lower proportion of variance. Consistent with the findings of Shao et al., we found weak correlations between RPF and 3D root density traits. However, the observation that both RPF and 3D Density S6 exhibited allele-specific root plasticity at the same QTL suggests a shared genetic mechanism influencing distinct aspects of root architecture across differential irrigation (Figure 4.2). These results indicate that the additive genetic variation at this QTL confers significant changes in mature maize RSA in

low-stress environments. Additional field evaluations using diverse maize lines will be essential to further investigate the additive genetic variation at this QTL and its role in modulating root growth responses to varying nitrogen environments.

In the present study, we aimed to provide functional evidence of *ZmSLAH* as a regulator of root system size and nitrogen capture in maize. Protein sequence analysis confirmed sequence identities similar to *AtSLAH2* and *AtSLAH3* and a conserved SLAC1-like domain. The 2019 GWAS identified a significant SNP located 176-kb upstream of *ZmSLAH* that was associated with 3D root system density. Field phenotypic evaluations demonstrated that the *ZmSLAH* insertional mutation led to functional differences in root architecture, nitrogen capture, and biomass accumulation. Notably, we observed G × E patterns for leaf nitrogen content in the nitrogen stress experiment. These results contribute novel single-gene insights into the genetic architecture underlying field-driven plasticity of maize RSA. Furthermore, this work will advance the biological characterization of *ZmSLAH* through the systematic capture of whole-plant phenotype data.

At anthesis, field evaluations of the *ZmSLAH* wild-type and mutant plants grown under high and low N conditions revealed contrasting patterns of biomass accumulation and nitrogen capture. Although both genotypes showed increased shoot biomass under low N conditions, only wild-type plants had greater root biomass and RMR (Figure 4.3). This variation was captured by the 2D measurements of root area, width, and number of tips (Figure 4.4). Across environments, shoot and root biomass were significantly higher ($p < 0.05$) in wild-type plants compared to mutant plants; however, no significant differences in RMR were detected. Interestingly, we observed a near significant G × E for leaf nitrogen content, where the wild-type plants showed greater leaf nitrogen capture under low N conditions compared to the mutant plants (Figure

4.3A). This was likely a consequence of the greater nutrient-uptake capacity of plants with a functional copy of *ZmSLAH*, allowing them to meet the increased shoot demand for nitrogen in the low N treatment as the plants grew to maturity (Coque et al., 2008; Garnett et al., 2013; Lu et al., 2024; Ma & Dwyer, 1998; Ning et al., 2013; Wang et al., 2006; Yan et al., 2011). Our results demonstrate that *ZmSLAH* induces systemic phenotypic changes and alters the relative sink activity in response to nitrogen stress. Therefore, we conclude that *ZmSLAH* plays a role in mediating active plastic responses in field-grown maize. Additional field sampling for grain nitrogen content will be conducted at the end of the growing season to assess the downstream effects on nitrogen remobilization during grain development.

In recent years, several genes involved in NO_3^- transport and assimilation have been identified and characterized *in planta*. For example, 97 members of the NO_3^- transporter/peptide transporter family (NRTs/NPFs) have been identified in maize (Plett et al. 2010, Léran et al. 2014). Among these, only members of the nitrate transporter 2 (NRT2) gene family are classified as high-affinity transporters that function under low-environmental NO_3^- conditions. Notably, *ZmNRT2.2* has been shown to play a significant role in genotype tolerance to low levels of NO_3^- in controlled environments (Garnett et al. 2015; Yan et al., 2020). The regulation of NO_3^- and Cl^- acquisition and transport by SLAC1/SLAH channels is conserved across plant lineages and has been studied in multiple species (Dreyer et al., 2012; Lind et al., 2015). These include moss [*Physcomitrella patens* (Hedw.) Bruch & Schimp.] (Caine et al., 2016; Chater et al., 2016; Dreyer et al., 2012;), dicot species like Arabidopsis and cotton [*Gossypium hirsutum* L.] (Demir et al., 2013; Geiger et al., 2010; Geiger et al., 2009; Hedrich & Geiger, 2017; Lehmann et al., 2021; Maierhofer, Diekmann, et al., 2014; Maierhofer, Lind, et al., 2014), and monocot species such as rice [*Oryza sativa* L.], date palm [*Phoenix dactylifera* L.], and maize (Geiger et al., 2009;

Li et al., 2022; Müller et al., 2017; Negi et al., 2008; Qi et al., 2018; Sun et al., 2016; Vahisalu et al., 2008). However, although SLAC1 has been the focus of extensive research, the roles and functions of its homologues have remained comparatively underexplored.

S-type anion channel homologues in Arabidopsis are divided into two groups based on their structure and function: SLAH1/4 and SLAH2/3 (Dreyer et al., 2012). The SLAH1/4 group includes SLAH1, a silent regulatory channel subunit, and SLAH4, whose function remains unknown (Dreyer et al., 2012; Hedrich & Geiger, 2017; Negi et al., 2008; Vahisalu et al., 2008). In contrast, the SLAH2/3 group consists of two ion-conducting channels: SLAH2, expressed in the root stele, and SLAH3, found in both guard cells and mesophyll cells, both contributing to the xylem loading of NO_3^- (Geiger et al., 11; Maierhofer, Lind, et al., 2014). Despite being close homologues, the SLAH2 and SLAH3 anion channels require unique properties for activation. For instance, while SLAH2 and SLAH3 both require phosphorylation by calcium-dependent kinases and the presence of extracellular NO_3^-, only the SLAH2 channel is strictly NO_3^- selective. SLAH3 can be activated by the heteromerization with SLAH1, even in the absence of extracellular NO_3^-, leading to increased Cl^- conductance (Cubero-Font et al., 2016; Jaborsky et al., 2016). Lehmann et al. (2021) reported that SLAH3-mediated root anion efflux can be activated by cytosolic acidification in Arabidopsis seedlings. While the biophysiology and properties of S-type anion channels in Arabidopsis have been well studied, more research is needed on economically important crop species.

In conclusion, we report *ZmSLAH* as an active regulator of root system size and nitrogen capture in field-grown maize. Specifically, our findings demonstrate that *ZmSLAH* has the capacity to regulate root-to-shoot NO_3^- transport under nitrogen-limited conditions. Although our findings provide insights into the genetic architecture of maize RSA and its plasticity to varying

environmental nitrogen, these experiments did not address its function at the molecular level. This was achieved through a collaborative project with a research group at the University of Würzburg that has demonstrated expertise in characterizing plant molecular-signaling networks in Arabidopsis and, notably, participated in the functional characterization of *AtSLAC1*, *AtSLAH1*, *AtSLAH2*, and *AtSLAH3* (Demir et al., 2013; Geiger et al., 2010; Geiger et al., 2009; Hedrich & Geiger, 2017; Lehmann et al., 2021; Maierhofer, Diekmann, et al., 2014; Maierhofer, Lind, et al., 2014). Additional experiments using heterologous gene expression and electrophysiological measurements in *Xenopus* oocytes will provide insights into the genetic identity and anion selectivity of *ZmSLAH* in maize. Overall, this study highlights *ZmSLAH* as a promising target for low-input germplasm development aimed at improving nitrogen capture in maize.

REFERENCES

Agrawal, A. A. (2001). Ecology: Phenotypic plasticity in the interactions and evolution of species. *Science*, *294*(5541), 321–326. https://doi.org/10.1126/science.1060701

Anderson, J. T., Wagner, M. R., Rushworth, C. A., Prasad, K. V. S. K., & Mitchell-Olds, T. (2014). The evolution of quantitative traits in complex environments. *Heredity*, *112*(1), 4–12. https://doi.org/10.1038/hdy.2013.33

Baythavong, B. S., & Stanton, M. L. (2010). Characterizing selection on phenotypic plasticity in response to natural environmental heterogeneity. *Evolution; International Journal of Organic Evolution*, *64*(10), 2904–2920. https://doi.org/10.1111/j.1558-5646.2010.01057.x

Bonser, A. M., Lynch, J., & Snapp, S. (1996). Effect of phosphorus deficiency on growth angle of basal roots in Phaseolus vulgaris. *New Phytologist*, *132*(2), 281–288. https://doi.org/https://doi.org/10.1111/j.1469-8137.1996.tb01847.x

Brouwer, R. (1962). Nutritive influences on the distribution of dry matter in the plant. *Netherlands Journal of Agricultural Science*, *10*(5), 399–408. https://doi.org/10.18174/njas.v10i5.17581

Burgess, S., & Thompson, S. G. (2021). *Mendelian Randomization* (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9780429324352

Caine, R. S., Chater, C. C., Kamisugi, Y., Cuming, A. C., Beerling, D. J., Gray, J. E., & Fleming, A. J. (2016). An ancestral stomatal patterning module revealed in the non-vascular land plant Physcomitrella patens. *Development*, *143*(18), 3306–3314. https://doi.org/10.1242/dev.135038

Chater, C. C., Caine, R. S., Tomek, M., Wallace, S., Kamisugi, Y., Cuming, A. C., Lang, D., MacAlister, C. A., Casson, S., Bergmann, D. C., Decker, E. L., Frank, W., Gray, J. E., Fleming, A., Reski, R., & Beerling, D. J. (2016). Origin and function of stomata in the moss Physcomitrella patens. *Nature Plants*, *2*(11). https://doi.org/10.1038/nplants.2016.179

Coque, M., Martin, A., Veyrieras, J. B., Hirel, B., & Gallais, A. (2008). Genetic variation for N-remobilization and postsilking N-uptake in a set of maize recombinant inbred lines. 3. QTL detection and coincidences. *Theoretical and Applied Genetics*, *117*(5), 729–747. https://doi.org/10.1007/s00122-008-0815-2

Cubero-Font, P., Maierhofer, T., Jaslan, J., Rosales, M. A., Espartero, J., Díaz-Rueda, P., Müller, H. M., Hürter, A. L., AL-Rasheid, K. A. S., Marten, I., Hedrich, R., Colmenero-Flores, J. M., & Geiger, D. (2016). Silent S-Type Anion Channel Subunit SLAH1 Gates SLAH3 Open for Chloride Root-to-Shoot Translocation. *Current Biology*, *26*(16), 2213–2220. https://doi.org/10.1016/j.cub.2016.06.045

Das, A., Schneider, H., Burridge, J., Ascanio, A. K. M., Wojciechowski, T., Topp, C. N., Lynch, J. P., Weitz, J. S., & Bucksch, A. (2015). Digital imaging of root traits (DIRT): A high-throughput computing and collaboration platform for field-based root phenomics. *Plant Methods*, *11*(1), 1–12. https://doi.org/10.1186/s13007-015-0093-3

de Mendiburu, F. (2023). *agricolae: Statistical Procedures for Agricultural Research*. https://cran.r-project.org/package=agricolae

Demir, F., Horntrich, C., Blachutzik, J. O., Scherzer, S., Reinders, Y., Kierszniowska, S., Schulze, W. X., Harms, G. S., Hedrich, R., Geiger, D., & Kreuzer, I. (2013). Arabidopsis nanodomain-delimited ABA signaling pathway regulates the anion channel SLAH3. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(20), 8296–8301. https://doi.org/10.1073/pnas.1211667110

Des Marais, D. L., Hernandez, K. M., & Juenger, T. E. (2013). Genotype-by-Environment Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic Environment. *Annual Review of Ecology, Evolution, and Systematics*, *44*(1), 5–29. https://doi.org/10.1146/annurev-ecolsys-110512-135806

DeWitt, T. J. (1998). Costs and limits of phenotypic plasticity: Tests with predator-induced morphology and life history in a freshwater snail. *Journal of Evolutionary Biology*, *11*(4), 465–480. https://doi.org/10.1046/j.1420-9101.1998.11040465.x

Dreyer, I., Gomez-Porras, J. L., Riaño-Pachón, D. M., Hedrich, R., & Geiger, D. (2012). Molecular evolution of slow and quick anion channels (SLACs and QUACs/ALMTs). *Frontiers in Plant Science*, *3*(11), 1–12. https://doi.org/10.3389/fpls.2012.00263

Dudley, S. A., & Schmitt, J. (1996). Testing the Adaptive Plasticity Hypothesis: Density-Dependent Selection on Manipulated Stem Length in Impatiens capensis. *The American Naturalist*, *147*(3), 445–465. https://doi.org/10.1086/285860

El-Soda, M., Malosetti, M., Zwaan, B. J., Koornneef, M., & Aarts, M. G. M. (2014). Genotype × environment interaction QTL mapping in plants: Lessons from Arabidopsis. *Trends in Plant Science*, *19*(6), 390–398. https://doi.org/10.1016/j.tplants.2014.01.001

Falconer, D. S. (1952). The Problem of Environment and Selection. *The American Naturalist*, *86*(830), 293–298. https://doi.org/10.1086/281736

Falconer, D. S. (1990). Selection in different environments: Effects on environmental sensitivity (reaction norm) and on mean performance. *Genetical Research*, *56*(1), 57–70. https://doi.org/10.1017/S0016672300028883

FAO, 2023. FAOSTAT: Crops and livestock products. [Accessed on [16 August 2024]]. [https://www.fao.org/faostat/en/#data/QCL/metadata] License: CC-BY-4.0.

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third). Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Gage, J. L., Jarquin, D., Romay, C., Lorenz, A., Buckler, E. S., Kaeppler, S., Alkhalifah, N., Bohn, M., Campbell, D. A., Edwards, J., Ertl, D., Flint-Garcia, S., Gardiner, J., Good, B., Hirsch, C. N., Holland, J., Hooker, D. C., Knoll, J., Kolkman, J., … De Leon, N. (2017). The effect of artificial selection on phenotypic plasticity in maize. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-01450-2

Galkovskyi, T., Mileyko, Y., Bucksch, A., Moore, B., Symonova, O., Price, C. A., Topp, C. N., Iyer-Pascuzzi, A. S., Zurek, P. R., Fang, S., Harer, J., Benfey, P. N., & Weitz, J. S. (2012). GiA Roots: Software for the high throughput analysis of plant root system architecture. *BMC Plant Biology*, *12*. https://doi.org/10.1186/1471-2229-12-116

Garnett, T., Conn, V., Plett, D., Conn, S., Zanghellini, J., Mackenzie, N., Enju, A., Francis, K., Holtham, L., Roessner, U., Boughton, B., Bacic, A., Shirley, N., Rafalski, A., Dhugga, K., Tester, M., & Kaiser, B. N. (2013). The response of the maize nitrate transport system to nitrogen demand and supply across the lifecycle. *New Phytologist*, *198*(1), 82–94. https://doi.org/10.1111/nph.12166

Garnett, T., Plett, D., Conn, V., Conn, S., Rabie, H., Rafalski, J. A., Dhugga, K., Tester, M. A., & Kaiser, B. N. (2015). Variation for N uptake system in maize: Genotypic response to N supply. *Frontiers in Plant Science*, *6*(11), 1–13. https://doi.org/10.3389/fpls.2015.00936

Gates, D. J., Runcie, D., Janzen, G. M., Navarro, A. R., Willcox, M., Sonder, K., Snodgrass, S. J., Rodríguez-Zapata, F., Sawers, R. J. H., Rellán-Álvarez, R., Buckler, E. S., Hearne, S., Hufford, M. B., & Ross-Ibarra, J. (2019). Single-gene resolution of locally adaptive genetic variation in Mexican maize. *BioRxiv*, *July*, 706739. https://doi.org/10.1101/706739

Geiger, D., Scherzer, S., Mumm, P., Marten, I., Ache, P., Matschi, S., Liese, A., Wellmann, C., Al-Rasheid, K. A. S., Grill, E., Romeis, T., & Hedrich, R. (2010). Guard cell anion channel SLAC1 is regulated by CDPK protein kinases with distinct Ca2+ affinities. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(17), 8023–8028. https://doi.org/10.1073/pnas.0912030107

Geiger, Dietmar, Maierhofer, T., Al-Rasheid, K. A. S., Scherzer, S., Mumm, P., Liese, A., Ache, P., Wellmann, C., Marten, I., Grill, E., Romeis, T., & Hedrich, R. (2011). Stomatal closure by fast abscisic acid signaling is mediated by the guard cell anion channel SLAH3 and the receptor RCAR1. *Science Signaling*, *4*(173), ra32. https://doi.org/10.1126/scisignal.2001346

Geiger, Dietmar, Scherzer, S., Mumm, P., Stange, A., Marten, I., Bauer, H., Ache, P., Matschi, S., Liese, A., Al-Rasheid, K. A. S., Romeis, T., & Hedrich, R. (2009). Activity of guard cell anion channel SLAC1 is controlled by drought-stress signaling kinase-phosphatase pair. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(50), 21425–21430. https://doi.org/10.1073/pnas.0912021106

Hedrich, R., & Geiger, D. (2017). Biology of SLAC1-type anion channels – from nutrient uptake to stomatal closure. *New Phytologist*, *216*(1), 46–61. https://doi.org/10.1111/nph.14685

Jaborsky, M., Maierhofer, T., Olbrich, A., Escalante-Pérez, M., Müller, H. M., Simon, J., Krol, E., Cuin, T. A., Fromm, J., Ache, P., Geiger, D., & Hedrich, R. (2016). SLAH3-type anion channel expressed in poplar secretory epithelia operates in calcium kinase CPK-autonomous manner. *New Phytologist*, *210*(3), 922–933. https://doi.org/10.1111/nph.13841

Jia, X., Liu, P., & Lynch, J. P. (2018). Greater lateral root branching density in maize improves phosphorus acquisition from low phosphorus soil. *Journal of Experimental Botany*, *69*(20), 4961–4970. https://doi.org/10.1093/jxb/ery252

Kell, D. B. (2012). Large-scale sequestration of atmospheric carbon via plant roots in natural and agricultural ecosystems: why and how. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1595), 1589–1597. https://doi.org/10.1098/rstb.2011.0244

Lambers, H. (1983). The "functional equilibrium": Nibbling on the edges of a paradigm. *Netherlands Journal of Agricultural Science*, *31*, 305–311. https://doi.org/10.18174/njas.v31i4.16935

Lasky, J. R., Upadhyaya, H. D., Ramu, P., Deshpande, S., Hash, C. T., Bonnette, J., Juenger, T. E., Hyma, K., Acharya, C., Mitchell, S. E., Buckler, E. S., Brenton, Z., Kresovich, S., & Morris, G. P. (2015). Genome-environment associations in sorghum landraces predict adaptive traits. *Science Advances*, *1*(6), 1–14. https://doi.org/10.1126/sciadv.1400218

Lehmann, J., Jørgensen, M. E., Fratz, S., Müller, H. M., Kusch, J., Scherzer, S., Navarro-Retamal, C., Mayer, D., Böhm, J., Konrad, K. R., Terpitz, U., Dreyer, I., Mueller, T. D., Sauer, M., Hedrich, R., Geiger, D., & Maierhofer, T. (2021). Acidosis-induced activation of anion channel SLAH3 in the flooding-related stress response of Arabidopsis. *Current Biology*, *31*(16), 3575-3585.e9. https://doi.org/10.1016/j.cub.2021.06.018

Leiboff, S., Li, X., Hu, H. C., Todt, N., Yang, J., Li, X., Yu, X., Muehlbauer, G. J., Timmermans, M. C. P., Yu, J., Schnable, P. S., & Scanlon, M. J. (2015). Genetic control of morphometric diversity in the maize shoot apical meristem. *Nature Communications*, *6*, 1–10. https://doi.org/10.1038/ncomms9974

Léran, S., Varala, K., Boyer, J. C., Chiurazzi, M., Crawford, N., Daniel-Vedele, F., David, L., Dickstein, R., Fernandez, E., Forde, B., Gassmann, W., Geiger, D., Gojon, A., Gong, J. M., Halkier, B. A., Harris, J. M., Hedrich, R., Limami, A. M., Rentsch, D., … Lacombe, B. (2014). A unified nomenclature of nitrate transporter 1/peptide transporter family members in plants. *Trends in Plant Science*, *19*(1), 5–9. https://doi.org/10.1016/j.tplants.2013.08.008

Li, X. D., Gao, Y. Q., Wu, W. H., Chen, L. M., & Wang, Y. (2022). Two calcium-dependent protein kinases enhance maize drought tolerance by activating anion channel ZmSLAC1 in

guard cells. *Plant Biotechnology Journal*, *20*(1), 143–157. https://doi.org/10.1111/pbi.13701

Lind, C., Dreyer, I., López-Sanjurjo, E. J., von Meyer, K., Ishizaki, K., Kohchi, T., Lang, D., Zhao, Y., Kreuzer, I., Al-Rasheid, K. A. S., Ronne, H., Reski, R., Zhu, J.-K., Geiger, D., & Hedrich, R. (2015). Stomatal Guard Cells Co-opted an Ancient ABA-Dependent Desiccation Survival System to Regulate Stomatal Closure. *Current Biology*, *25*(7), 928–935. https://doi.org/10.1016/j.cub.2015.01.067

Lowry, D. B., Lovell, J. T., Zhang, L., Bonnette, J., Fay, P. A., Mitchell, R. B., Lloyd-Reilley, J., Boe, A. R., Wu, Y., Rouquette, F. M., Wynia, R. L., Weng, X., Behrman, K. D., Healey, A., Barry, K., Lipzen, A., Bauer, D., Sharma, A., Jenkins, J., … Juenger, T. E. (2019). QTL × environment interactions underlie adaptive divergence in switchgrass across a large latitudinal gradient. *Proceedings of the National Academy of Sciences*, *116*(26), 12933–12941. https://doi.org/10.1073/pnas.1821543116

Lu, J., Lankhost, J. A., Stomph, T. J., Schneider, H. M., Chen, Y., Mi, G., Yuan, L., & Evers, J. B. (2024). Root plasticity improves maize nitrogen use when nitrogen is limiting - an analysis using 3D plant modelling. *Journal of Experimental Botany*, *July*. https://doi.org/10.1093/jxb/erae298

Lynch, J. (1998). The Role of Nutrient-Efficient Crops in Modern Agriculture. *Journal of Crop Production*, *1*(2), 241–264. https://doi.org/10.1300/J144v01n02_10

Lynch, J. P. (2013). Steep, cheap and deep: An ideotype to optimize water and N acquisition by maize root systems. *Annals of Botany*. 112(2). 347–357. https://doi.org/10.1093/aob/mcs293

Lynch, J. P., & Brown, K. M. (2001). Topsoil foraging – an architectural adaptation of plants to low phosphorus availability. *Plant and Soil*, *237*(2), 225–237. https://doi.org/10.1023/A:1013324727040

Lynch, M., & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer. https://books.google.com/books?id=UhCCQgAACAAJ

Ma, B. L., & Dwyer, L. M. (1998). Nitrogen uptake and use of two contrasting maize hybrids differing in leaf senescence. *Plant and Soil*, *199*(2), 283–291. https://doi.org/10.1023/A:1004397219723

Maierhofer, T., Diekmann, M., Offenborn, J. N., Lind, C., Bauer, H., Hashimoto, K., Al-Rasheid, K. A. S., Luan, S., Kudla, J., Geiger, D., & Hedrich, R. (2014). Site-and kinase-specific phosphorylation-mediated activation of SLAC1, a guard cell anion channel stimulated by abscisic acid. *Science Signaling*, *7*(342), 1–12. https://doi.org/10.1126/scisignal.2005703

Maierhofer, T., Lind, C., Hüttl, S., Scherzer, S., Papenfuß, M., Simon, J., Al-Rasheid, K. A. S., Ache, P., Rennenberg, H., Hedrich, R., Müller, T. D., & Geiger, D. (2014). A single-pore

residue renders the Arabidopsis root anion channel SLAH2 highly nitrate selective. *Plant Cell*, *26*(6), 2554–2567. https://doi.org/10.1105/tpc.114.125849

McCarty, D. R., Mark Settles, A., Suzuki, M., Tan, B. C., Latshaw, S., Porch, T., Robin, K., Baier, J., Avigne, W., Lai, J., Messing, J., Koch, K. E., & Curtis Hannah, L. (2005). Steady-state transposon mutagenesis in inbred maize. *Plant Journal*, *44*(1), 52–61. https://doi.org/10.1111/j.1365-313X.2005.02509.x

Monroe, J. G., Powell, T., Price, N., Mullen, J. L., Howard, A., Evans, K., Lovell, J. T., & McKay, J. K. (2018). Drought adaptation in arabidopsis thaliana by extensive genetic loss-of-function. *ELife*, *7*, 1–18. https://doi.org/10.7554/eLife.41038

Müller, H. M., Schäfer, N., Bauer, H., Geiger, D., Lautner, S., Fromm, J., Riederer, M., Bueno, A., Nussbaumer, T., Mayer, K., Alquraishi, S. A., Alfarhan, A. H., Neher, E., Al-Rasheid, K. A. S., Ache, P., & Hedrich, R. (2017). The desert plant Phoenix dactylifera closes stomata via nitrate-regulated SLAC1 anion channel. *New Phytologist*, *216*(1), 150–162. https://doi.org/https://doi.org/10.1111/nph.14672

Negi, J., Matsuda, O., Nagasawa, T., Oba, Y., Takahashi, H., Kawai-Yamada, M., Uchimiya, H., Hashimoto-Sugimoto, M., & Iba, K. (2008). CO2 regulator SLAC1 and its homologues are essential for anion homeostasis in plant cells. *Nature*, *452*, 483–486. https://doi.org/10.1038/nature06720

Ning, P., Li, S., Yu, P., Zhang, Y., & Li, C. (2013). Post-silking accumulation and partitioning of dry matter, nitrogen, phosphorus and potassium in maize varieties differing in leaf longevity. *Field Crops Research*, *144*, 19–27. https://doi.org/https://doi.org/10.1016/j.fcr.2013.01.020

Nour-Eldin, H. H., Hansen, B. G., Nørholm, M. H. H., Jensen, J. K., & Halkier, B. A. (2006). Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Research*, *34*(18), e122–e122. https://doi.org/10.1093/nar/gkl635

Paterson, A. H., Saranga, Y., Menz, M., Jiang, C.-X., & Wright, R. (2003). QTL analysis of genotype × environment interactions affecting cotton fiber quality. *Theoretical and Applied Genetics*, *106*(3), 384–396. https://doi.org/10.1007/s00122-002-1025-y

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., Bileschi, M. L., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunić, I., Marchler-Bauer, A., Mi, H., Natale, D. A., Orengo, C. A., Pandurangan, A. P., Rivoire, C., … Bateman, A. (2023). InterPro in 2022. *Nucleic Acids Research*, *51*(D1), D418–D427. https://doi.org/10.1093/nar/gkac993

Pigliucci, M. (2001). *Phenotypic plasticity: beyond nature and nurture*. JHU Press.

Piperno, D. R., Ranere, A. J., Holst, I., Iriarte, J., & Dickau, R. (2009). Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley,

Mexico. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(13), 5019–5024. https://doi.org/10.1073/pnas.0812525106

Plett, D., Toubia, J., Garnett, T., Tester, M., Kaiser, B. N., & Baumann, U. (2010). Dichotomy in the NRT gene families of dicots and grass species. *PLoS ONE*, *5*(12). https://doi.org/10.1371/journal.pone.0015289

Poorter, H., Niklas, K. J., Reich, P. B., Oleksyn, J., Poot, P., & Mommer, L. (2012). Biomass allocation to leaves, stems and roots: Meta-analyses of interspecific variation and environmental control. *New Phytologist*, *193*(1), 30–50. https://doi.org/10.1111/j.1469-8137.2011.03952.x

Postma, J. A., Dathe, A., & Lynch, J. P. (2014). The optimal lateral root branching density for maize depends on nitrogen and phosphorus availability. *Plant Physiology*, *166*(2), 590–602. https://doi.org/10.1104/pp.113.233916

Pretty, J. (2008). Agricultural sustainability: Concepts, principles and evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1491), 447–465. https://doi.org/10.1098/rstb.2007.2163

Qi, G.-N., Yao, F.-Y., Ren, H.-M., Sun, S.-J., Tan, Y.-Q., Zhang, Z.-C., Qiu, B.-S., & Wang, Y.-F. (2018). The S-Type Anion Channel ZmSLAC1 Plays Essential Roles in Stomatal Closure by Mediating Nitrate Efflux in Maize. *Plant and Cell Physiology*, *59*(3), 614–623. https://doi.org/10.1093/pcp/pcy015

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Rogers, A. R., Dunne, J. C., Romay, C., Bohn, M. O., Buckler, E. S., Ciampitti, I. A., Edwards, J. W., Ertl, D., Flint-Garcia, S. A., Gore, M. A., Graham, C., Hirsch, C. N., Hood, E. E., Hooker, D. C., Knoll, J. E., Lee, E. C., Lorenz, A. J., Lynch, J. P., McKay, J. K., … Holland, J. B. (2021). The importance of dominance and genotype-by-environment interactions on grain yield variation in a large-scale public cooperative maize experiment. *G3*, *11(2)*. https://api.semanticscholar.org/CorpusID:231927024

Saengwilai, P., Tian, X., & Lynch, J. P. (2014). Low Crown Root Number Enhances Nitrogen Acquisition from Low-Nitrogen Soils in Maize. *Plant Physiology*, *166*(2), 581–589. https://doi.org/10.1104/pp.113.232603

Schneider, H. M., Lor, V. S. N., Hanlon, M. T., Perkins, A., Kaeppler, S. M., Borkar, A. N., Bhosale, R., Zhang, X., Rodriguez, J., Bucksch, A., Bennett, M. J., Brown, K. M., & Lynch, J. P. (2022). Root angle in maize influences nitrogen capture and is regulated by calcineurin B-like protein (CBL)-interacting serine/threonine-protein kinase 15 (ZmCIPK15). *Plant Cell and Environment*, *45*(3), 837–853. https://doi.org/10.1111/pce.14135

Schneider, H. M., & Lynch, J. P. (2020). Should Root Plasticity Be a Crop Breeding Target? *Frontiers in Plant Science*, *11*(5), 1–16. https://doi.org/10.3389/fpls.2020.00546

Settles, A. M., Latshaw, S., & McCarty, D. R. (2004). Molecular analysis of high-copy insertion sites in maize. *Nucleic Acids Research*, *32*(6), 1–12. https://doi.org/10.1093/nar/gnh052

Shao, M. R., Jiang, N., Li, M., Howard, A., Lehner, K., Mullen, J. L., Gunn, S. L., McKay, J. K., & Topp, C. N. (2021). Complementary phenotyping of maize root system architecture by root pulling force and X-ray imaging. *Plant Phenomics*, *2021*. https://doi.org/10.34133/2021/9859254

Smith, J. S. C., Duvick, D. N., Smith, O. S., Cooper, M., & Feng, L. (2004). Changes in pedigree backgrounds of pioneer brand maize hybrids widely grown from 1930 to 1999. *Crop Science*, *44*(6), 1935–1946. https://doi.org/10.2135/cropsci2004.1935

Stelpflug, S. C., Sekhon, R. S., Vaillancourt, B., Hirsch, C. N., Buell, C. R., de Leon, N., & Kaeppler, S. M. (2016). An expanded maize gene expression atlas based on RNA sequencing and its use to explore root development. *Plant Genome*, *9*(1). https://doi.org/10.3835/plantgenome2015.04.0025

Sultan, S. E. (2000). Phenotypic plasticity for plant development, function and life history. *Trends in Plant Science*, *5*(12), 537–542. https://doi.org/10.1016/S1360-1385(00)01797-0

Sun, S. J., Qi, G. N., Gao, Q. F., Wang, H. Q., Yao, F. Y., Hussain, J., & Wang, Y. F. (2016). Protein kinase OsSAPK8 functions as an essential activator of S-type anion channel OsSLAC1, which is nitrate-selective in rice. *Planta*, *243*(2), 489–500. https://doi.org/10.1007/s00425-015-2418-x

Tello-Ruiz, M. K., Stein, J., Wei, S., Youens-Clark, K., Jaiswal, P., & Ware, D. (2016). Gramene: A Resource for Comparative Analysis of Plants Genomes and Pathways. *Methods in Molecular Biology*, *1374*, 141–163. https://doi.org/10.1007/978-1-4939-3167-5_7

Topp, C. N., Iyer-Pascuzzi, A. S., Anderson, J. T., Lee, C. R., Zurek, P. R., Symonova, O., Zheng, Y., Bucksch, A., Mileyko, Y., Galkovskyi, T., Moore, B. T., Harer, J., Edelsbrunner, H., Mitchell-Olds, T., Weitz, J. S., & Benfey, P. N. (2013). 3D phenotyping and quantitative trait locus mapping identify core regions of the rice genome controlling root architecture. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(18). https://doi.org/10.1073/pnas.1304354110

Trachsel, S., Kaeppler, S. M., Brown, K. M., & Lynch, J. P. (2013). Maize root growth angles become steeper under low N conditions. *Field Crops Research*, *140*, 18–31. https://doi.org/https://doi.org/10.1016/j.fcr.2012.09.010

United Nations, Department of Economic and Social Affairs, Population Division (2024). *World Population Prospects 2024*: Data Sources.

Vahisalu, T., Kollist, H., Wang, Y.-F., Nishimura, N., Chan, W.-Y., Valerio, G., Lamminmäki, A., Brosché, M., Moldau, H., Desikan, R., Schroeder, J. I., & Kangasjärvi, J. (2008). SLAC1 is required for plant guard cell S-type anion channel function in stomatal signalling. *Nature*, *452*(7186), 487–491. https://doi.org/10.1038/nature06608

van Heerwaarden, J., Doebley, J., Briggs, W. H., Glaubitz, J. C., Goodman, M. M., de Jesus Sanchez Gonzalez, J., & Ross-Ibarra, J. (2011). Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Sciences*, *108*(3), 1088–1092. https://doi.org/10.1073/pnas.1013011108

Wang, H., Inukai, Y., & Yamauchi, A. (2006). Root Development and Nutrient Uptake. *Critical Reviews in Plant Sciences*, *25*(3), 279–301. https://doi.org/10.1080/07352680600709917

Woods, P., Lehner, K. R., Hein, K., Mullen, J. L., & McKay, J. K. (2022). Root Pulling Force Across Drought in Maize Reveals Genotype by Environment Interactions and Candidate Genes. *Frontiers in Plant Science*, *13*(4), 1–14. https://doi.org/10.3389/fpls.2022.883209

Yan, H., Li, K., Ding, H., Liao, C., Li, X., Yuan, L., & Li, C. (2011). Root morphological and proteomic responses to growth restriction in maize plants supplied with sufficient N. *Journal of Plant Physiology*, *168*(10), 1067–1075. https://doi.org/https://doi.org/10.1016/j.jplph.2010.12.018

Yan, M., Luo, J., Liang, L., Chen, L., Cao, Y. Y., Chen, Y. H., Zhu, X. Y., & Yu, C. M. (2020). Comparative analysis of nitrogen metabolism-related genes in two inbred maize lines with different low-nitrogen tolerance. *Biologia Plantarum*, *64*, 289–301. https://doi.org/10.32615/bp.2019.093

York, L. M., Galindo-Castañeda, T., Schussler, J. R., & Lynch, J. P. (2015). Evolution of US maize (Zea mays L.) root architectural and anatomical phenes over the past 100 years corresponds to increased tolerance of nitrogen stress. *Journal of Experimental Botany*, *66*(8), 2347–2358. https://doi.org/10.1093/jxb/erv074

York, L. M., Nord, E. A., & Lynch, J. P. (2013). Integration of root phenes for soil resource acquisition. *Frontiers in Plant Science*, *4*(9), 1–15. https://doi.org/10.3389/fpls.2013.00355