THE FOUNDATION OF COMPUTATIONAL BIOLOGY IN PLANT IMMUNITY

By

Huan Chen

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Genetics and Genome Sciences Program – Doctor of Philosophy
Molecular Plant Sciences – Dual Major

2023

ABSTRACT

Plants constantly cope with various abiotic and biotic stress while interacting with the environment by utilizing their immunity. Computational biology has become one of the most important approaches in the study of plant biology with the assistance next-generation sequencing data analysis since mid-2000s. Although many core plant immune questions and functions of key elements in cytoskeleton have been revealed, there are still many foundational questions that remain mysterious. Plants may survive after experiencing biotic and abiotic stress, such as fungal infection and drought stress, but the detailed progress of the immune response is still not well-covered. Modern maize evolved in South America followed by key traits were selected under the influence of both nature and cultural preference, but the specific selection under culture influence is still unknown. While part of ACTIN DEPOLYMERIZING FACTOR (ADF), which is essential as one the actin-binding protein (ABP) of the plant cytoskeleton to regulate the plant's immune system's ability, is known, the multiple functions of ADFs remain unknown. As a result, this is a field of research where analysis through computational biology is necessary to answer questions about plant immunity, as well as to build biological models for it.

In this dissertation, I focus on using a combination of computation biology methods and plant biological knowledge to uncover the role of plant immunity in the response to biotic and abiotic stress, Incan cultural influence on selection of key traits of maize in South America and *ADF* gene family evolution and gene function prediction. First, I addressed the question of "why Fusarium virguliforme (*Fv*) is asymptomatic on monocot host and a symptomatic eudicot host" utilizing multiple methods, and I found that root senescence is a primary contributing factor underlying colonization and disease

progression in symptomatic versus asymptomatic host–fungal interactions. Furthermore, I identified the role of NDR1 (NON-RACE-SPECIFIC DISEASE RISISTANCE1) in drought response by analyzing the transcriptional data from different treatments at various times. The results revealed that NDR1 is required for signaling processes associated with plant cell drought stress responses. Lastly, I identified the key traits that were selected by under Inca culture, including shorter growing season, greater stress-resistant, and sweeter taste, between ancient and modern maize in South America. This was achieved by utilizing a combination of several computational methods, along with influence of Incan culture and maize diversity. This study provides insights into the complex biocultural role that Inca culture had in determining the direction of maize diversity in South America. As the supplement, I predicted the potential functions of ADF by identifying the evolutionary history of the *ADF* gene family through expression pattern. The results uncovered the potential function of ADFs and provided direction for future research in traditional biology. In summary, my work has provided significant insights into the plant immunity by using computational biology knowledge.

ACKNOWLEDGEMENTS

I would like first to thank my advisor Dr. Brad Day, for his generous support and illuminative guidance on my study and research. He provided me with multiple interesting projects, which helped train me to ask critical questions, think big and develop the ability to cooperate with other scientists successfully in my research areas. I also appreciate his respect for some of my naïve ideas on research and his provision of freedom to try and develop my abilities through practice. Without his support, I would not have grown into a proud young scientist.

I would like to thank the former and current members of the Day lab, Dr. Yi-Ju Lu, Dr. Alex Corrion, Dr. Amy Baetsen-Young, Dr. Masaki Shimono, Dr. Saroopa Samaradivakara, Dr. Brittni Kelley, Dr. Yongsig Kim, Dr. Yi Zhao, Dr. Pai Li and Mr. Ryan Sheick, for the supportive and enjoyable lab environment that we created. Special thanks go out to Dr. Yi-Ju Lu, Alex Corrion and Dr. Amy Baetsen-Young who provided me with the chance to cooperate with them.

I would like to thank my committee member Dr. Shin-han Shiu who provided me a seat in his lab for 1 year and 10 months, as well as help on my supplementary chapter. I also thank Dr. Melissa Lehti-Shiu for her advice on my supplementary chapter and all the members in Shiu lab. Special thanks go out to the current people of the Shiu lab: Kenia Segura Abá, Brianna Brown, Thilanka Ranaweera and Eleanor Siler, and former lab member Serena Lotreck, as well as Dr. Jamell Dacon and Harry Shomer from Data Science and Engineering (DSE) lab, who were very supportive, and I am fortunate to have all of them as both colleagues and as friends in both work and life. I would also like to

thank the Dr. Paulo Izquierdo and Brianna Brown's for their generous help on my internship application.

I would like to thank my Ph.D. Graduation Committee members Dr. Kevin Liu and Dr. Benjamin J Orlando for their critical guidance on the general direction and crucial details of my research. I also would like to thank my Ph.D. program chair, Dr. Vieille Claire, and Academic Program Coordinator, Alaina Burghardt, for their kind help towards my graduation.

I would like to thank my collaborator Dr. Gabriel Wrobel and Dr. William Lovis from the Department of Anthropology, Dr. Amy Baetsen-Young from Syngenta, Dr. Addie Thompson from Genetics & Genome Sciences Program, Dr. Emily Josephs from Ecology, Evolution & Behavior Program (EEB) on the collaboration project, for their significant help for my dissertation.

I would like to thank my parents, who have provided sustained support in my life and gave me enough freedom and respected the decisions that I have made. Although I am not the first person in our big family's history who earned a doctorate, I am proud that we now have one more. Finally, I would like to thank my husband, Peter, for his understanding, support, and companionship in this both happy and challenging time in my life.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

NDR1: NON-RACE-SPECIFIC DISEASE RISISTANCE1

F. virguliforme: *Fusarium virguliforme*

ADF(s): Actin Depolymerization Factor(s)

aBM: Archaeological Bolivian Maize

Chapter contributions

Chapter2 and 3 were completed with other collaborators. My major contributions for are as follows: In Chapter2, I contributed by incorporating senescence transcription factors' expression patterns, senescence-associated processes, significant differential gene expression (DEGs) in *Fv* and gene expression patterns in soybean and maize, shaping the overall manuscript design. To accomplish the manuscript, what I did was: 1) Reanalyzed datasets using shared expression values from the co-first author and corrected calculation errors in the previous manuscript (Table 2-1; Data set AA 2-2 and 2-5). 2) Generated expression patterns of significant DEGs (Figure 2-4G), positive and negative transcription factors (Figure 2-7B; Data Set AA 2-22), and senescence-associated biological processes (Figure 2-8; Data Set AA 2-21). Also, produced graphical representations for expressed genes and DEGs post-*Fv* inoculation (Figure AA 2-4), gene ontology enrichment of *Fv* DEGs (Figure AA 2-7), biological processes of *Fv* DEGs (Figure AA 2-8), gene ontology enrichment of putatively orthologous and non-orthologous genes (Figure AA 2-9), and biological processes of DEGs at each day after inoculation (Figure AA 2-10). 3) Reorganized and edited all original main figures. 4) Wrote the transcription factor part (Figure 2-7B), the *Fv*-promoted susceptibility via regulatory biological processes controlling senescence (Figure 2-8) and depicted changes in DEGs and biological processes (Figures AA 2-7, 2-8, 2-9, and 2-10). 5) Revised the manuscript according to specific suggestions from all the reviewers.

In Chapter 3, I analyzed RNA-Seq datasets and formulated hypotheses for downstream experiments based on gene transcription levels. To complete the manuscript, what I did was: 1) Conducted comprehensive data analysis of RNA-Seq

datasets (Data Set AA 3-1 – 3-15). 2) Analyzed biological processes of DEGs (Data Set AA 3-16 – 3-21, 3-23). 3) Completed co-expression and gene model analysis (Data Set AA 3-22). 4) Generated Figures 3-4, 3-5, and Figure AA 3-11 – 3-13-14. 5) Wrote the analysis section for RNA-Seq datasets and part of the discussion for the entire paper. 6) Uploaded RNA-Seq datasets to NCBI and obtained a BioProject ID.

CHAPTER 1: Literature review

The plant stress and computational biology

Introduction

Compared to animals, plants cannot move, so they must endure various abiotic and biotic stresses. These stresses greatly limited their growth, development, and distribution. To survive, plants use multiple processes, including sensing, signaling, transcription, transcript processing, translation, and post-translational protein modifications, to respond to these stresses on molecular mechanisms. In nature, these stresses negatively affect crop productivity and limit the utilization of land(Bailey-Serres et al., 2019; Zhang et al., 2022). Therefore, exploring and understanding the mechanisms of how plants respond to stress and adapt to the harsh environment is essential for global agriculture security.

The foundations of computation biology were started in the early 1960s, and it became essential for biology study because of the major improvements in sequencing technology in the 1990s through the 2000s, which led the costs reductions and data increases. As time goes by, computation biology has been critical for the arrival of the ‘Big Data’ (Gauthier et al., 2019). In this chapter, how plants respond to various stress through their immunity system be addressed by using computational biology will be mainly discussed.

Biotic and abiotic stress

Biotic stresses, including insects, viruses and microorganisms, fungi, competitive species, and animals. Biotic stresses are critical factors limiting soybean growth and development. Sudden death syndrome (SDS) is one of the most destructive diseases in soybean worldwide (Allen, 2017; Bandara et al., 2020; Wrather, 2010), which is caused by *Fusarium virguliforme* (*Fv*). Although soybean (*Glycine max*) uses complex regulatory

and defense mechanisms to respond to SDS, SDS still causes huge damage to the yield of soybean. Maize and soybean are often rotated on the same land to get rid of the SDS, but it still causes ~$274 million dollars annual loss in the U.S.(Allen, 2017; Bandara et al., 2020; Wrather, 2010). Recent research showed that SDS disease symptoms through root rot(Wang et al., 2019). The current knowledge provides the foundation to explore how pathogens with broad host ranges modulate immune signaling within diverged hosts.

Abiotic stresses, including drought, salinity, heat, cold, UV light, and flooding, have profound effects on plants, and they are major limiting factors affecting crop production both qualitatively and quantitatively(Bashir et al., 2019). It showed that abiotic stress can reduce yields by more than 60% on average(Boyer, 1982). Recently, the global human population is still increasing and the world's population could grow to around 9.7 billion by 2050 (Nations., 2023). However, the environmental changes made global warming become a big challenge for agriculture and its efforts to meet the growing population's need for food in the world. High temperatures are strongly associated with drought stress, and droughts or heat waves can lead to harvest failures and cause huge damage to the livelihoods of agricultural producers and the food security of communities worldwide. Drought is a condition in which water is either lacking or insufficient for a long time, and as a consequence of water shortage, plants are affected by drought. Drought stress is becoming one of the most significant issues with decreased availability and increasing demand for water in the whole world (Kopecka et al., 2023). Agriculture needs 70% of global clean water, but this percentage is expected to increase rapidly over the coming years (Boretti, 2019). It showed drought alone could even cause more annual loss in crop yield than all pathogens combined (Gupta et al., 2020).

To adapt to drought stress, plants use multiple processes, including sensing, signaling, transcription, transcript processing, translation, and post-translational protein modifications, to respond to these stresses on the molecular mechanisms level. For example, the key mechanism is the abscisic acid (ABA) hormone regulates the stomata closure(P. Wang et al., 2023). What's more, plants modify their root growth and architecture, and close stomata to reduce water loss on the physiology level (Gupta et al., 2020). The previous research showed that the *DEEPER ROOTING 1* (*OsDRO1*) gene in rice could promote deep rooting and maintain a high yield under drought conditions (Uga et al., 2013), and the expressing the *StDRO1* gene from potatoes in *Arabidopsis* has also shown positive effect of root architecture on drought resistance (Sun et al., 2022). These reports suggest that identifying key genes, that are important for drought resistance, could be used to preserve yields under drought stress.

Archaeological study of maize in South America

Mazie (*Zea mays* ssp. *mays*) is one of the most important and productive food crops in the modern world (Ranum et al., 2014). Maize originally evolved in Mexico from the flowering species of grass teosinte (*Zea mays* ssp. *parviglumus*) ~9,000 years before the present (cal. BP) (Doebley, 2004) (Matsuoka et al., 2002) and spread into South America by ~7,000 cal. BP (Grobman et al., 2012; Lombardo et al., 2020) from Eastern North America by ~2500 cal BP7. The dispersal of maize to the rest of the world followed European colonization of the Americas in the 15th and 16th centuries AD (Bonavia, 2013; Rebourg et al., 2003). As well known, the diversity of maize was shaped by natural selection. The evidence has also shown that cultural preferences were a factor in maintaining some key traits (Bellon, 1996; Knapp, 1997). For example, based on cultural

preferences of the time, sometimes at the cost of reduced cultivability or nutrition, maize seeds with certain color and texture still were selected and retained by farmers for use in the following growing year(s) to maintain or enhance specific morphological traits (Louette & Smale, 2000). Mazie was a sacred crop, the symbol of political power, and tightly associated with the life of high-status or elite social groups during the Inca Empire era (Bray, 2003). These reports suggest that the evolution and diversity of maize have been shaped by both Inca cultural and natural selection in a wide range of environmental and social contexts in South America.

Computation biology

In the 1990s through the 2000s, the major improvements in sequencing technology reduced costs and increased data in next-generation sequencing (NGS), contributing to the arrival of 'Big data' in the modern biology area, which helps biologists find answers to specific biology questions. Computer and specialized software have become an essential part of the biologist's toolkit for aiding answers finding from genomes, transcriptomics, proteomics, and metabolomics levels. Therefore, more expertise from computer science is needed in the field (Gauthier et al., 2019).

The earliest time of computational biology occurred more than 50 years ago, a time even earlier than the DNA could be sequenced. The foundation of computational biology started with the applications of computational methods to protein sequence analysis, and the first known bioinformatics software was developed to solve the assembling of the whole protein sequence from hundreds of small Edman peptide sequences(Edman & Begg, 1967). After that, the first multiple sequence alignment (MSA) algorithm which uses a scoring matrix whose dimensionality equals the number of sequences emerged in the

early 1980s(Murata et al., 1985), and the first software dedicated to analyzing Sanger sequencing reads was published (Staden, 1979) in 1979 since all of the 64 codons of the genetic code were deciphered in 1968 (Crick, 1968). Later on, the maximum likelihood (ML) method was developed to infer the phylogenetic tree from the DNA sequence (Felsenstein, 1981), and the first three-dimensional structure of the protein was determined experimentally in 1958(Kendrew et al., 1958). At the beginning of the 21st century, the publication of the human genome started the genomic era (Lander et al., 2001; Venter et al., 2001). At the same time, the sequencing of the *Arabidopsis thaliana* genome also ushered in the genomics era for plant research, and an incredible variety of computational tools were developed and used in large repositories of genomic, transcriptomic, proteomic, epigenomic, and other '-omic' data has accelerated our understanding of plant biology (Canto-Pastor et al., 2021). These suggest that we can find the answers to some key biological questions by using computational biology approaches.

Aims of the thesis research

This thesis is aimed at using computational biology methods to explore the mechanisms of how plants respond to abiotic and biotic stress, what key traits were selected under culture preference in ancient plants, and how this contributed to plant diversity. In Chapter 2, I found that root senescence is a primary contributing factor underlying colonization and disease progression in a symptomatic eudicot, soybean, versus an asymptomatic monocot, maize, host–fungal interactions. In Chapter 3, I identified that NDR1 associates with the plasma membrane-localized H+-ATPases AHA5, and in this association, negatively regulates the activity of the guard cell-localized ATPase

to play a role in response to drought stress. In Chapter 4, I identified the key traits that were selected in maize under Incan culture, including shorter growing seasons, greater stress-resistant, and sweeter taste, in South America. In supplement, identifying the evolutionary history of the *ADF* gene family and partly identifying the *ADF* gene expression pattern.

CHAPTER 2: Contrasting transcriptional responses to *Fusarium virguliforme* colonization in symptomatic and asymptomatic hosts

The work presented in this chapter has been published:

from Amy Baetsen-Young*, Huan Chen* et al., 2021

Plant Cell, Volume 33, Issue 2, February 2021, Pages 224–247

Chapter abstract

The broad host range of *Fusarium virguliforme* represents a unique comparative system to identify and define differentially induced responses between an asymptomatic monocot host, maize (*Zea mays*), and a symptomatic eudicot host, soybean (*Glycine max*). Using a temporal, comparative transcriptome-based approach, we observed that early gene expression profiles of root tissue from infected maize suggest that pathogen tolerance coincides with the rapid induction of senescence dampening transcriptional regulators, including *ANACs (Arabidopsis thaliana NAM/ATAF/CUC protein)* and *Ethylene-Responsive Factors*. In contrast, the expression of senescence-associated processes in soybean was coincident with the appearance of disease symptom development, suggesting pathogen-induced senescence as a key pathway driving pathogen susceptibility in soybean. Based on the analyses described herein, we posit that root senescence is a primary contributing factor underlying colonization and disease progression in symptomatic versus asymptomatic host–fungal interactions. This process also supports the lifestyle and virulence of *F. virguliforme* during biotrophy to necrotrophy transitions. Further support for this hypothesis lies in comprehensive co-expression and comparative transcriptome analyses, and in total, supports the emerging concept of necrotrophy-activated senescence. We propose that *F. virguliforme* conditions an environment within symptomatic hosts, which favors susceptibility through transcriptomic reprogramming, and as described herein, the induction of pathways associated with senescence during the necrotrophic stage of fungal development.

Introduction

Analyses of contrasting phenes have illuminated our understanding of the complexity of processes in plants, examples of which include the regulation and interaction between plant development and the environment (York et al., 2013) and response to abiotic stress (Lynch et al., 2014). Indeed, “genotype phenotype” interaction studies have aided in the identification of a multitude of genetic interactions underpinning host susceptibility and resistance to pathogens. For example, comparative studies evaluating phenotypically distinct responses within a single plant to a single pathogen have provided insight at physiological-to-genetic/genomics-scale resolution (Lorrain et al., 2018; O'Connell et al., 2012). Indeed, when coupled with functional analyses of gene networks, comparative transcriptome-based approaches have provided insight into a wide breadth of immune signaling processes. Key examples include processes governing (1) the activation of the hypersensitive response, a hallmark of effector-triggered immunity (Chisholm et al., 2006) and the activation of resistance (R) genes (Etalo et al., 2013), (2) pattern-triggered immunity (Bagnaresi et al., 2012; Zhang et al., 2018a), and (3) the role of antimicrobial signaling and metabolism. Taken together, the array of these processes illustrates the temporal colonization dynamics between susceptible and resistant plant cultivars during their interaction(s) with a variety of pathogenic organisms (Burkhardt & Day, 2016; Kong et al., 2015).

While numerous processes required for the regulation of plant immunity have been described through the use of pairwise genomics- and transcriptomics-based approaches, most studies to date have focused on comparison(s) within cultivars from a single species, or species within a single genera (Burkhardt & Day, 2016; Chen et al., 2016;

Chowdhury et al., 2017; Lanubile et al., 2015). Thus, knowledge gaps still remain in our understanding of how pathogens with broad host ranges modulate immune signaling within diverged hosts. This is particularly limiting in the case of fungal pathogens, many of which colonize a broad range of agronomically important crops, yet only a few are true pathogens (Banerjee et al., 2019; Gdanetz, 2017). For example, phytopathogenic Ascomycetes found within the species complexes of Fusarium solani and F. oxysporum colonize more than 200 plant species, including both monocots and eudicots (Michielse & Rep, 2009; Sisic et al., 2018). We posit that a comparison of host responses across highly diverged plant species colonized by the same fungus will provide insight(s) into immunity-induced pathways relevant to both resistance and tolerance. Moreover, such comparisons would be highly relevant and translatable to monoculture-based agroecosystems, particularly those comprising rotational crops, which are typically exposed to similar microbial community compositions each growing season (Gdanetz, 2017; Katan, 2017; Leandro et al., 2018). Relevant to the work described herein, rotation-based management practices are proving to be unsustainable as many field crops support the asymptomatic persistence of numerous pathogenic fungi (Kolander et al., 2012; Lofgren et al., 2018; Malcolm et al., 2013).

A mechanistic understanding of the molecular-genetic fundamentals of plant immunity has been largely built upon the framework of the model plant *Arabidopsis thaliana*, and from this, a robust understanding of eudicot immune signaling exists. However, immune signaling in monocots remains largely enigmatic (Balmer et al., 2013b; Balmer et al., 2013a), and whilst defense gene homologs in monocots and eudicots have been described, there are many examples where well-defined eudicot pathways are not

consistently conserved, nor do they function the same in monocots (Balmer et al., 2013b; Humphry et al., 2010; Lu et al., 2011). Furthermore, attempting to infer the function and activity of below-ground host immune responses based on knowledge generated through the analysis of above-ground colonization processes and disease symptom development is unrealistic (Chuberre et al., 2018; De Coninck et al., 2015). In this regard, the work described herein offers insight into the contrasting signaling processes that exist between an asymptomatic monocot and a symptomatic eudicot following fungal pathogen colonization.

Differential host responses are prominent among crops colonized by Fusarium virguliforme, the causal agent of soybean (Glycine max) sudden death syndrome (SDS), a devastating disease yielding annual losses in the United States in excess of $274 million dollars (Allen, 2017; Wrather, 2010). As one example of management strategies to reduce yield losses, soybean and maize (Zea mays) are typically grown in rotation (Katan, 2017). However, a recent study showed that maize can serve as an asymptomatic host for F. virguliforme, thereby providing a potential reservoir for this devastating soybean pathogen (Kolander et al., 2012). In the current study, we conducted a temporal transcriptome-based analysis to identify, and compare, genetic signatures underlying asymptomatic and symptomatic responses in maize and soybean roots during fungal pathogen colonization. Through this, we have defined the early temporal interactions between F. virguliforme and two hosts: maize (asymptomatic) and soybean (symptomatic) to generate a transcriptional atlas of pathogen-induced responses in each crop species. Using this approach, we asked if orthologous transcriptional responses are induced in a temporally independent manner between soybean and maize. Based on the output of the analysis

presented herein, we suggest that pathogen-induced host senescence plays a predominant role in the establishment of fungal necrotroph colonization, a process we hypothesize distinguishes symptomatic (i.e. pathogen-induced) senescence in soybean from asymptomatic colonization and persistence of the pathogen in maize.

Results

Fusarium virguliforme colonizes soybean and maize

We recently reported a role for fungal transcriptome plasticity in the regulation and adaptation of F. virguliforme in association with asymptomatic (maize) and symptomatic (soybean) hosts (Baetsen-Young et al., 2020). To extend this analysis, and to identify induced host defense signaling responses associated with fungal pathogen colonization, we conducted a comprehensive analysis of fungal colonization of soybean and maize roots over a 2-week period of colonization. Following pathogen inoculation, soybean radical elongation was observed from 0 to 4 days after inoculation (DAI), with subsequent development of lateral and secondary lateral roots at 7 and 10 DAI, respectively (Figure 2-1A; Figure AA 2-1A). By 14 DAI, lateral and tap root growth was abundant, with full expansion of cotyledons in soybean hosts following mock- and pathogen-inoculation. This represented a vegetative cotyledon (Fehr WR, 1977) stage for soybean.

Consistent with the hemibiotrophic lifestyle of F. virguliforme, no symptoms were apparent on soybean roots from 0 to 4 DAI, illustrative of the biotrophic stage of the pathogen(Ngaki et al., 2016). At 7 DAI, discoloration of the lower tap root of F. virguliforme-inoculated samples was observed, which subsequently developed into necrotrophic streaking at 10 DAI. Beyond 10 DAI, complete necrosis of the lower hypocotyl was observed, with subsequent disease symptom spread to the lateral roots

adjoining the tap root. Both mock- and pathogen-inoculated maize showed comparable growth and development patterns to those observed in soybean (Figure 2-1A; Figure AA 2-1B). In brief, maize radical and seminal roots slowly emerged from 0 to 3 DAI, and by 4 DAI, had expanded in length. At 7 DAI, lateral roots initialized along the primary root, followed by crown root emergence (Figure AA 2-1B). By 14 DAI, the seminal, crown, and primary roots had further elongated, with parallel expansion of leaves to the V1 growth stage (i.e. full extension of the first leaf). Given that we did not observe visual symptoms on F. virguliforme-inoculated maize roots, we next employed quantitative real-time polymerase chain reaction (qPCR) to evaluate the in planta fungal growth in both maize and soybean. Using this approach, we observed that F. virguliforme DNA concentrations increased on both soybean and maize roots over the timecourse of analysis, with 0 and 2 DAI displaying significantly lower levels of F. virguliforme DNA quantities than 14 DAI ($P < 0.0001$; Figure 2-1B). Interestingly, both inoculated soybean and maize exhibited similar levels of F. virguliforme, as determined by qPCR at most timepoints, with notable increases between soybean and maize between 10 and 14 DAI. All sampled timepoints after fungal inoculation were significantly different from mock treatments (Tukey's test with $P < 0.001$; Table AA 2-1). As previously observed, we detected nominal background levels of DNA in tissue samples from mock-inoculated roots using F. virguliforme-specific DNA primers, an artifact that is associated with the non-specific nature of the qPCR assay at detection limits below 10 fg of DNA resulting in modest signals associated with nonspecific DNA (plant or fungal) binding (Wang et al., 2015).

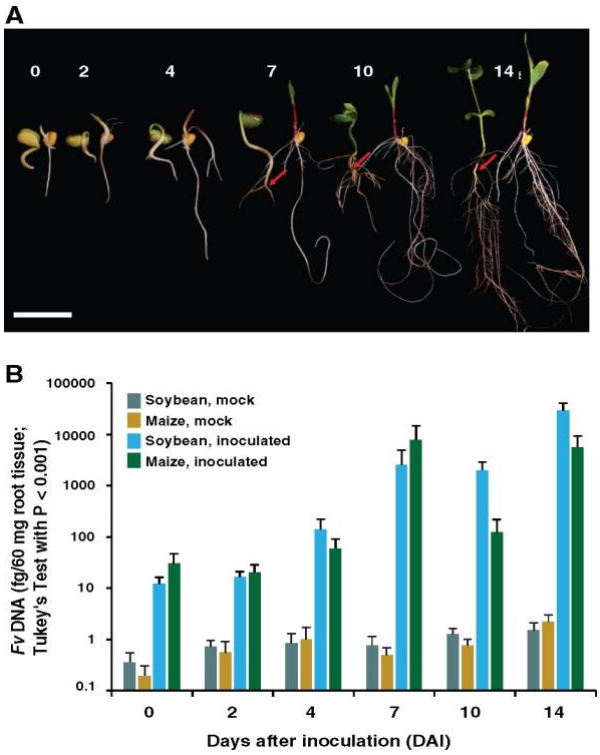

Figure 2-1: *Fusarium virguliforme* growth on maize and soybean. (A) Plant growth and development over 14 days (numbered) post inoculation with *F. virguliforme.* Soybean is on the left and maize is on the right in each pair. Bar = 4 cm. (B) F. virguliforme DNA in planta quantification from mock and inoculated soybean and maize roots. *F. virguliforme* levels were determined using a TaqMan-based quantitative PCR assay. Values shown are the average of three biological replicates, each of which contained two plants, ± SEM (n = 6).

Temporal expression of induced defense genes in soybean and maize

Having established that F. virguliforme colonizes both hosts, we conducted a comparative RNA-seq experiment using samples from six selected timepoints over a 0- to 14-day inoculation period. Timepoints were selected based on preliminary colonization timecourse experiments (Figure 2-1; Figure AA 2-1). The impetus for this approach was to (1) determine if the timing of defense responses underpin symptomatic and asymptomatic phenotype development and (2) identify pathway rewiring in host plants by

the pathogen which potentially function in modulating tolerance and susceptibility (Chen et al., 2016; Chowdhury et al., 2017).

Mock- and fungal-inoculated maize and soybean roots from 0, 2, 4, 7, 10, and 14 DAI were subjected to RNA-seq. Each of these biological samples were pooled from six independent plants, as well as repeated in three independent growth chamber runs. Sampling of plant tissues from both mock- and fungal-inoculated events at each timepoint allowed us to discover genes specifically related fungal colonization. In addition, all sampling and phenotypic observations were performed from the same inoculation site within the root, thereby enabling a comprehensive evaluation of host response to fungal inoculation, including fungal adhesion, development and colonization, and penetration and proliferation. Sequencing of each sample to an average of 70 million reads yielded uniquely mapped reads (after adapter trimming and quality thresholding) in the range of 76%– 83% for soybean and 77%–81% for maize (Figure AA 2-2 and Data Set AA 2-1).

As a key step in demonstrating the relatedness of gene expression profiles across sampling as a function of host and time, we observed a high correlation (496%) of gene expression patterns among biological replicates (Figure AA 2-3), illustrating consistency among independent replicates and within treatments. To explore expression patterns between mock- and F. virguliforme-infected plants, samples were grouped using principal component analysis, the output of which revealed that gene expression patterns in maize were primarily separated from each other by time, with distinct groupings at 0, 2–7, 10, and 14 DAI (Figure AA 2-2B). This is in contrast to groupings by treatment over time, indicating that root development had a greater impact on gene expression than did fungal colonization. Conversely, transcripts from soybean were primarily grouped by treatment

(Figure AA 2-2C). By 4 DAI, mock- and fungal-inoculated treatments had separated into distinct groups, and the remaining samples were separated by treatment as the time course proceeded. The large separation between treatments suggests large global expression changes within soybean roots, as the F. virguliforme induced disease symptoms.

To further catalog gene expression changes as those that were identified as differentially expressed following *F. virguliforme* inoculation, we first filtered the dataset for genes that were significantly up- or downregulated (multiple-testing adjusted $P < 0.05$, $|\log2(\text{fold-change}| > 1$) between fungal inoculated samples and mock-inoculated treatments within each host species at each timepoint. Among the 28,956 expressed maize genes identified in at least one timepoint, 600 unique responsive [i.e. differentially expressed genes (York et al.)] genes were detected across all timepoints. Among the 600 significantly responsive genes identified in maize in response to *F. virguliforme*, 266 were upregulated and 336 were downregulated with two significantly up and downregulated at different timepoints, with the majority of the identified differential expression changes occurring at 0 and 14 DAI (Figure 2-2, A and Table 2-1; Table AA 2-2, Figure AA 2-4 and Data Set AA 2-2). In contrast, among 43,308 soybean genes expressed at one or more timepoints, 10,898 were significantly differentially expressed following *F. virguliforme* colonization (Table AA 2-3 and Data Set AA 2-3). Interestingly, there was an almost equal number of up- and downregulated genes from soybean (5,643 and 5,325, respectively), with a majority of DEGs identified at 7 DAI and beyond (Figure 2-2B and Table 2-1). This contrasts with the temporal pattern of response changes observed in maize, an observation we posit further highlights that differential host gene expression may underpin

the divergence in symptomatic versus asymptomatic phenotypes in soybean and maize, respectively.

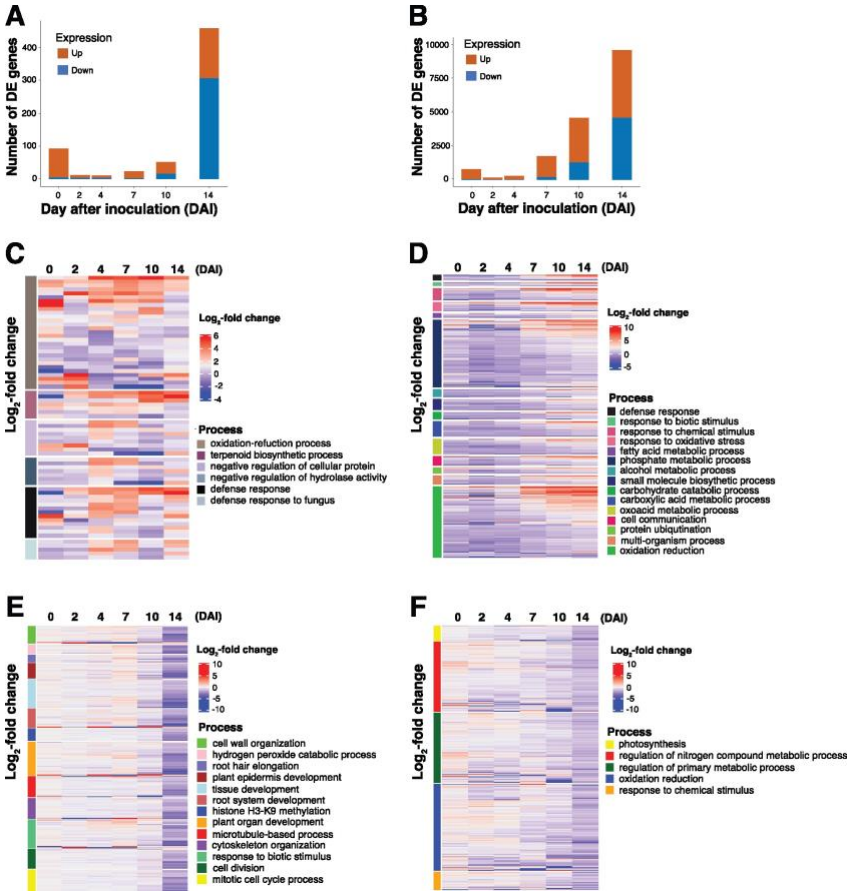

Figure 2-2: Temporal pattern of defense gene expression in maize and soybean. (A and B) The number of significantly DEGs between mock and inoculated maize with log2(FC) > 1, respectively, of early differentially regulated genes in maize or soybean from six timepoints over the colonization timecourse. (C and D) Heatmap of log2-fold gene expression changes among transcripts cataloged in ontologies significantly enriched in upregulated genes across pooled timepoints for maize (n = 266) and soybean (n = 5,643). (E and F) Heatmap of significantly enriched gene ontologies from *F. virguliforme*-induced downregulated genes across pooled timepoints for maize (n = 336) and soybean (n = 5,325).

Previous studies have demonstrated that fungi elicit an array of pro-immune transcriptional responses in their host(s) roots, including the activation of metabolic

processes such as the generation of reactive oxygen species (ROS) and the release of secondary metabolites (Pusztahelyi, 2016; Zhang et al., 2017). To gain insight into processes that underpin the observed phenotypic differences between symptomatic and asymptomatic host responses, we first conducted a gene ontology (GO) enrichment analysis of upregulated DEGs. We identified defense-related transcriptional response categories as the most enriched processes from those within upregulated genes over the timecourse of colonization of maize (Table AA S1-4). Indeed, and as illustrated in Figure 2-2C, we observed an enrichment of genes from maize that were associated with defense response processes following *F. virguliforme* colonization, including those associated with oxidation–reduction and terpenoid biosynthesis (Boba et al., 2020). This suggests that maize, while phenotypically asymptomatic, recognized *F. virguliforme* and activated defense-associated processes following inoculation. Based on the observation that *F. virguliforme* inoculation elicited a broader alteration of gene expression in soybean than in maize, we predicted that a greater diversity of enriched processes would be evident in soybean as compared to maize. Specifically, 12 significantly (multiple testing adjusted $P < 0.05$) enriched biological process categories were identified in maize (Table AA 2-4), including defense response signaling and the regulation of oxidation reduction processes. In contrast, 46 significantly (multiple-testing adjusted $P < 0.05$) enriched biological process categories were identified in soybean (Table AA 2-5). Among these were numerous categories such as phosphate metabolism, protein ubiquitination, and cell-to-cell communication (Figure 2-2D). The increased number of identified categories not only points to broader disruption in soybean following pathogen colonization, but in the case of altered metabolic functions, is associated with the alteration of soybean root

physiological processes and associated *F. virguliforme* colonization. When we compared the same enriched categories across hosts, we observed that defense-related ontologies were upregulated much earlier (i.e. 0–4 DAI) in samples from inoculated maize roots than in soybean (Figure 2-2, C and D). We hypothesize that this may underpin *F. virguliforme* tolerance in maize and is further supportive of a diverged temporal response between the two hosts. Indeed, we observed that the same gene categories that were also enriched early (i.e. induced) in maize were induced at later points in the colonization of soybean (7–14 DAI).

Table 2-1: Transcriptomic landscape of soybean and maize during colonization by *F. virguliforme* over a 2-week timecourse of colonization.

Transcriptome	Soybean	Maize
Total expressed genes	43,308	28,956
DEGs	10,898	600
Orthologous genes	23,274	14,542
Orthologs, N (%)		
Orthologous DEGs	4,528 (41.5)	221 (36.8)
Orthologous DEGs, up	2,296 (50.7)	86 (38.9)
Orthologous DEGs, down	2,257 (49.8)	137 (62)
Non-orthologs, N (%)		
Nonorthologous DEGs	6,370 (58.5)	379 (63.2)
Nonorthologous DEGs, up	3,347 (52.5)	180 (47.5)
Nonorthologous, DEGs, down	3,068 (48.2)	199 (52.5)

The total percent of orthologous DEGs that were up- and downregulated in soybean reflects the percentage assignment(s) resulting from 25 genes in soybean being both up- and downregulated over the timecourse. The total percent of orthologous DEGs that were up- and downregulated in soybean reflects the percentage assignment(s) resulting from two genes in soybean being both up- and downregulated over the timecourse. The total percent of nonorthologous DEGs that were up- and downregulated in soybean reflects the percentage assignment(s) resulting from 45 genes in soybean being both up- and downregulated over the timecourse.

To identify biological processes that were repressed in each host over the duration of the colonization timecourse, we conducted an enrichment analysis of downregulated

genes and their associated ontologies. As shown in Figure 2-2E, we identified 336 genes in maize, enriched among 13 biological process categories, that were significantly (multiple-testing adjusted $P < 0.05$) downregulated over the timecourse of *F. virguliforme* colonization (Table 2-1). Notably, among these were those associated with plant growth and development, as well as several key defense processes (Data Set AA 2-2). In contrast, in soybean, 5,325 downregulated genes and their enriched categories—especially at later timepoints of the infection—were identified (Data Set AA 2-3). Among these, many have previously been described as host processes essential for fungal growth and development, and nitrogen metabolism (Figure 2-2F; Fagard et al., 2014). For example, the downregulation of photosynthesis during pathogen infection has been demonstrated to not only coincide with the activation of robust defense signaling (Su et al., 2018), but is also associated with pathogen-induced downregulation of photosynthesis as a mechanism to induce ROS accumulation in chloroplasts, which leads to reduced levels of photosystems I and II (Kretschmer et al., 2019).

As a final step in this stage of our analysis, we conducted a comparative co-expression analysis of each host's transcriptome using a differential gene correlation analysis (DGCA). The impetus for this was to discover significant changes within gene pairs caused by the treatment, by employing the median z-score difference of the gene pair correlations in the first condition compared to the second condition, and then compared to all gene pairs (McKenzie et al., 2016). Using this approach, we identified 128 gene pairs (ca. 0.00003%) in maize that were significantly differentially correlated; that is, these gene pairs were observed to undergo a change in correlated expression in comparison of mock versus pathogen inoculation (Data Set AA 2-4). Among these, two

transcripts from maize with homologs to *Arabidopsis* R-genes were identified as having significant differential correlations. The first, *RESISTANCE TO PSEUDOMONAS SYRINGAE pv. MACULICOLA (RPM1*; Zm00001d014099), showed an expression profile that was positively correlated with *XYLOGLUCAN GALACTOSYLTRANSFERASE* (Zm00001d029862) in response to fungal inoculation, but negatively correlated in mock treatments. Specifically, expression of the putative maize RPM1 ortholog was downregulated at 0 DAI (ca. 19.5-fold change), but expression gradually increased over the time course to an approximate two-fold change. The second, *RESISTANCE TO PERONOSPORA PARASITICA 13-LIKE-4* [*RPP13L4*; Zm00001d018786, containing a Rho-N domain {Zm0001d051967], showed a negative expression correlation during colonization, yet was positively correlated under mock-inoculation conditions. The former suggests possible fungal modulation of R-gene expression during colonization, a scenario supported by previous work which identified this *R*-gene as a candidate for SDS resistance (Zhang et al., 2015). Additionally, the observed co-regulation of *R*-gene expression with genes associated with cell wall-associated remodeling processes (i.e. xyloglucan galactosyltransferase) further illustrates the connectivity of immunity with pathogen-induced changes in host cell architecture (Day & Graham, 2007).

In soybean, many more gene pairs (7,950,000) were identified as significantly differentially correlated when we compared samples from mock- and *F. virguliforme*-inoculated roots. To determine how these identified gene pairs were co-expressed, we constructed a planar filtered network (PFN), dissecting the resultant data into 1,160 modules (Supplementary Data Modules). Using this, we next explored the enrichment of co-regulated defense genes in the above modules and found that while they were

significantly enriched in “metabolic process” and “primary metabolic process” (Figure AA 2-5 and Data Set AA 2-5), we did not identify enrichment in defense-related processes. This further supports our assertion that biotrophic fungal pathogens suppress plant cell death and manipulate plant metabolism in soybean (Doehlemann et al., 2017).

Orthologous host gene analysis reveals differential patterns of induced defense responses

Previous studies have described temporal changes in gene expression during compatible and/or incompatible pathogen interactions, an observation that associates with the onset of disease symptom development (Kong et al., 2015). Here, we observed the enrichment of defense related GOs at distinct timepoints in soybean and maize. To address whether defense-induced responses in soybean and maize occur in a temporally similar, or diverged, manner, we first identified the orthologous genes between soybean and maize. We reasoned that this would enable us to compare the temporal expression pattern changes between orthologs, which are descended from a single gene stemming from the last common ancestor. First, we clustered the predicted protein sequences from the maize and soybean genomes using OrthoFinder, which yielded 10,700 orthogroups, containing 23,273 and 14,542 expressed genes in soybean and maize, respectively (Data Set AA 2-6). For each orthogroup, if orthologs contain more than a one-to-one relationship (e.g. two maize orthologs to one soybean ortholog), the median of the log2-fold changes value for each DEG was calculated at each timepoint. The variation in DE expression between soybean and maize orthologs was assessed to capture the greatest differential by the median transformation of gene expression values, as represented by significant correlations at several timepoints (Figure AA 2-6).

To understand the significance of the soybean–maize orthologous gene set in light of the host response to *F. virguliforme* colonization, we next explored whether the DEGs were contained in the orthologs. Overall, a similar number of DEGs were identified as orthologous in both soybean and maize (41.5% versus 36.8%). However, of the orthologous genes cataloged from soybean and maize, an apparent difference was observed in the percentage of genes that were either up- or downregulated within each host. For example, in soybean, we identified a similar number of up and downregulated genes from among the list of orthologous DEGs (50.7% up and 49.8% down), while in maize, 38.9% were upregulated and 62% were downregulated (Table 2-1; Figure AA 2-4 and Data Set AAs 2-7 and 2-8).

To determine if the host response to fungal colonization was unique to each host, we cataloged each orthogroup depending on whether significant DEGs, or groups of genes, were maize-specific, soybean-specific, or occurred in both hosts. As shown in Figure 3A, we identified 3,003 significantly upregulated orthogroups, whereby 93.3% of the orthogroups classified as significantly upregulated were uniquely upregulated in soybean. In contrast, only 3.3% of the orthogroups were uniquely upregulated in maize. In addition, 3.4% were upregulated in both hosts following *F. virguliforme* inoculation (Figure 3A; Data Set AA 2-9), indicating that more than half of the upregulated genes in maize were also upregulated in soybean. To further explore the processes contained within the list of identified orthologous DEGs, we developed a merged gene ontologies annotation containing gene annotations from maize and soybean. This new GO annotation contained 4,685 orthogroups with GO annotations from maize, and an additional 6,015 orthogroups were supported by both soybean and maize annotations.

Interestingly, no orthogroups were solely associated with soybean (Data Set AA 2-10). Orthogroups that were uniquely upregulated in soybean were highly enriched for responses to oxygen containing compounds, defense, and hormone metabolism (Data Set AA 2-11). Orthogroups that were upregulated in both hosts in response to *F. virguliforme* colonization were enriched for response to organic substance, stimulus, and defense (Data Set AA 2-11). Orthogroups that were upregulated and unique to maize were enriched with cytoskeleton associated processes and DNA methylation (Data Set AA 2-12).

In addition, we identified 1,660 significantly downregulated orthogroups; as shown in Figure 2-3B, 92.2% of the orthogroups classified as significantly downregulated were uniquely downregulated in soybean, while only 5.8% of the orthogroups were uniquely downregulated in maize. Approximately 2.0% were downregulated in both hosts (Data Set AA 2-9). Specifically, orthogroups that were uniquely downregulated in soybean were enriched for cell wall organization, root development, and auxin transport (Data Set AA 2-11). Surprising, we observed that the most significantly enriched process was photosynthesis, a finding that is agreement with previous studies demonstrating that the rate of photosynthesis decreases during leaf senescence (Wojciechowska et al., 2018). While roots and leaves utilize numerous distinct processes during growth, development, and environment interactions, the process of cellular and tissue senescence in each follow similar patterns of gene process expression (Liu et al., 2019). Leveraging this similarity, we exploited our root-specific transcriptome datasets as a guide to discover

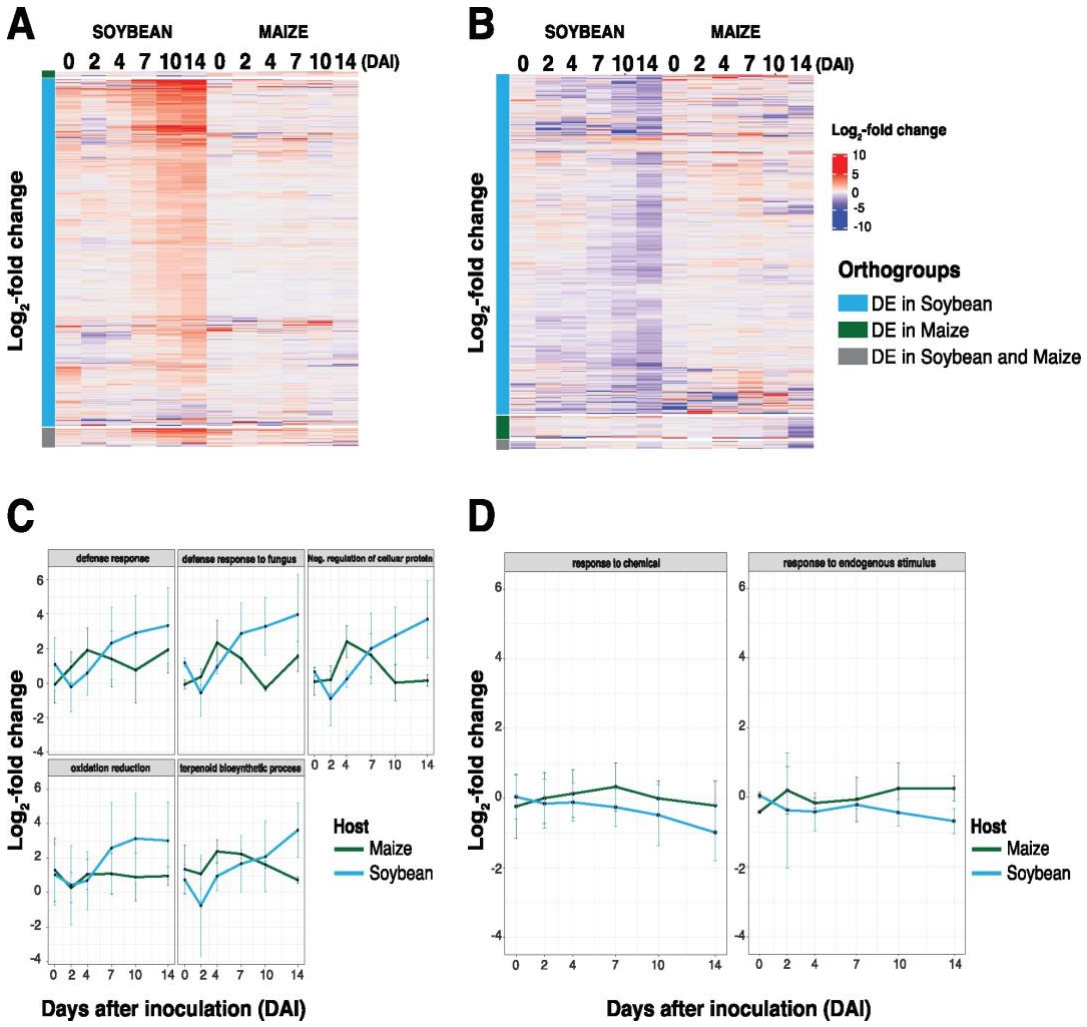

Figure 2-3: An analysis of putative orthologous host processes reveals differential patterns of induced defense responses. (A and B) Heatmap of expression of significantly (log2(FC) > 1) upregulated (n = 3,003) and downregulated (n = 1,660) genes from mock- and fungal-inoculated samples. Gray indicates orthologous genes from soybean and maize are significant, blue indicates orthologous genes from only soybean are significant, and green illustrates orthologous genes from maize that were uniquely significantly DE. (C and D) Mean expression patterns of log2(FC) > 1 gene expression profiles from among significantly up- and downregulated orthologous genes, respectively, at a single timepoint in both hosts over the colonization timecourse. Error bars indicate one standard deviation of the mean.

and further characterize the processes that are associated with pathogen-induced senescence in soybean. Using this approach, we observed that a small proportion of genes downregulated in both soybean and maize was highlighted by only 19 processes

being significantly enriched (Data Set AA 2-11). Not surprisingly, these included response to stimulus, root development, and response to external stimulus. Orthogroups that were uniquely downregulated in maize were enriched for cell cycle processes, the regulation of reproduction, and the negative regulation of metabolic process. Additionally, we evaluated *F. virguliforme* GO enrichments during colonization of both maize and soybean (Figure AAs 1-7 and 1-8). We did not observe an increase in the expression of *F. virguliforme* genes involved in chitin metabolic processes (a key indicator of fungal growth and development), yet we did identify a downregulation of *F. virguliforme* genes associated with ROS production from samples collected from colonized maize roots. Based on this, we surmised that these observations provide a basis from which we could proceed to compare and identify orthologous responses to colonization between the two hosts (Figure AA 2-9, A and B, Figure AA 2-10, A and B). Surprisingly, we did not observe an enrichment in soybean genes associated with defense response (e.g. “defense response to fungus”, “immune response”, and “root development”). Taken together, these data highlight the different host responses in soybean and maize in response to *F. virguliforme*.

To address how the temporal patterns of orthogroup induction differed between soybean and maize, we next focused on the defense-relevant GO categories that were enriched in maize during the early timepoints of the colonization (0–4 DAI; Data Set AA 2-11). To do this, we first identified upregulated orthogroups in biological processes associated with host–pathogen interactions, including: (1) defense response; (2) defense response to fungus; (3) negative regulation of cellular protein; (4) oxidation reduction; and (5) terpenoid biosynthetic process (Figure 2-3C). As shown, we observed a shift in the

induced patterns of expression between maize and soybean, with a notable upregulation of orthologous genes at 2 and 4 DAI in maize. Interestingly, and consistent with our hypothesis of a delayed response to pathogen colonization being one of the underlying factors driving susceptibility, these same responses were not observed as enriched/induced until 4 to 7 DAI in soybean. Genes within these upregulated orthogroups encoded *PATHOGENESIS-RELATED PROTEIN 10 (PR10), STRESSINDUCED PROTEIN 1, TERPENE SYNTHASE*, and *BROWMANBIRK TYPE TRYPSIN INHIBITOR*. Based on this, it is tempting to hypothesize that while similar host defense-associated genes are activated during the course of *F. virguliforme* colonization in both soybean and maize, the induced expression of these genes at earlier time intervals in maize associates with the absence of disease symptom progression, as compared to soybean.

Using an approach similar to the above, we also identified downregulated defense-relevant process groups (Figure 2-3D; Data Set AA 2-11). Among these, we observed that the expression of downregulated genes associated with response to chemical and endogenous stimulus were slightly diminished between maize and soybean; specifically, we observed an upregulation of orthologous genes in maize, but a downregulation in soybean and 7–10 DAI. We posit that this further supports the hypothesis that a delayed response to pathogen in soybean underpins, yet is not wholly responsible for, susceptibility to *F. virguliforme* colonization. Indeed, if one considers the *in planta* lifestyle transition (i.e. biotrophy to necrotrophy) of *F. virguliforme*, pathogen co-option of host processes at discrete stages of this transition in a host-specific manner (Figure AAs 2-7 and 2-8) likely serves as a critical virulence strategy. Thus, it is reasonable to hypothesize that the pathogen's ability to alter the timing and activation sequence of host defense

processes underpins host transitions from pathogen recognition, to defense, to either susceptibility, resistance, and/or tolerance (Figure AA 2-9, A and B, Figure AA 2-10, A and B; (Doehlemann et al., 2017).

Nonorthologous defense processes within maize and soybean

While the identification of orthologous defense-related processes in soybean and maize identified temporal divergence in host response(s) to fungal colonization, an analysis of the novel host-specific defense signaling processes from each was not apparent. To resolve this, we surveyed the specific processes unique to each host associated with upregulated and downregulated DEGs as a comparison between maize and soybean. As shown in Figure 2-4, A and B, we identified non-orthologous genes that were significantly upregulated (180 and 3,347 in maize and soybean, respectively) and downregulated (199 and 3,068 in maize and soybean, respectively) following fungal colonization (Table 2-1; Data Set AAs 2-12 and 2-13, Table AA 2-6 and 2-7).

We next asked if there were induced defense responses derived from defense gene categories that were unique to each host (i.e. genes that did not have an apparent ortholog). As shown in Figure 2-4C, upregulated DEGs in maize following *F. virguliforme* inoculation were associated with GO enrichment of processes including defense response to fungus, immune signaling, the hypersensitive response, and host programmed cell death (Table AA 2-8 and Data Set AA 2-14). Among these, and consistent with the initial biotic interaction lifestyle of this fungus, processes in maize were primarily associated with host programmed cell death, a mechanism likely associated with localized defenses preventing fungal proliferation and further invasion (Deshmukh et al., 2006; Diamond et al., 2013; Jones, 2001; Schafer et al., 2007). Conversely, in soybean,

we identified upregulated genes associated with processes such as glucose catabolic and response to carbohydrate catabolic process, previously identified as metabolic signatures of the onset of host–fungal interactions, host susceptibility, and disease (Figure 2-4D; Table AA 2-8, Data Set AA 2-15;(Askew et al., 2009; Deveau et al., 2008; Divon & Fluhr, 2007). In addition, we explored the significantly upregulated nonorthologous genes pattern changes between maize and soybean within the same biological processes. As shown in Figure 2-5A, processes associated with “defense response”, “defense response to fungus”, “oxidation reduction”, and “terpenoid biosynthetic process” showed similar responses to fungal colonization in maize (ca. 2–7 DAI) than observed in soybean. In maize, the majority of downregulated genes identified were differentially regulated at later points in the colonization (i.e. 14 DAI; Figure 2-4E; Data Set AA 2-16). In soybean, the host processes known to be associated with survival during fungal colonization were enriched at early timepoints in the colonization; specifically, as early as 2 DAI (Figure 2-4F; Data Set AA 2-17).

Next, we explored the nonorthologous response to *F. virguliforme* between maize and soybean within each timepoint. The impetus for this was to identify the various, and differentially induced, biological processes utilized by each host in their response to fungal colonization, including as a function of the timing of fungal lifestyle transitions (Figure AA 2-9, C and D, Figure AA 2-10, C and D). Additionally, we queried if defense processes identified using this approach were relevant to the observed host phenotypes (Figure AA

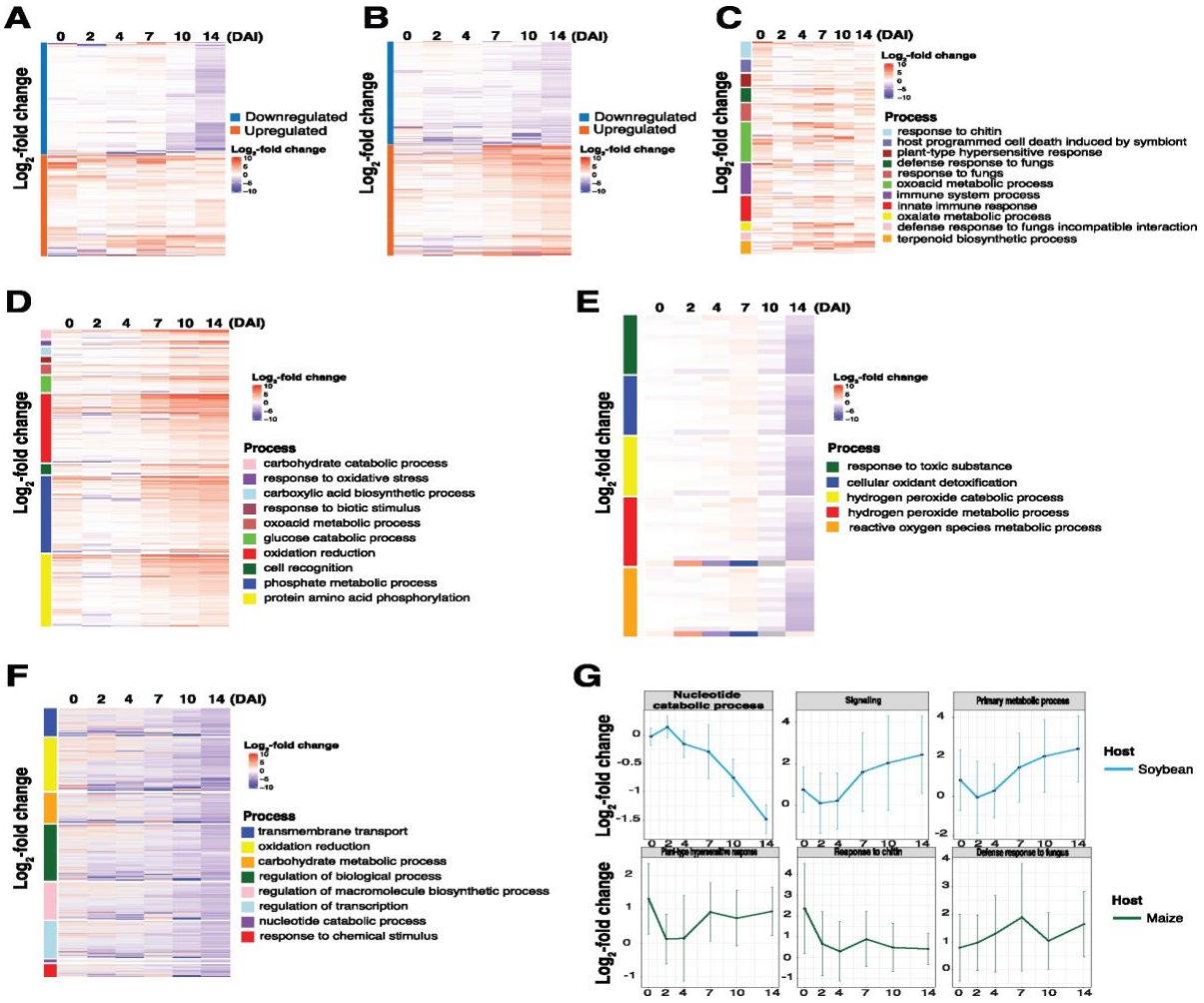

Figure 2-4: Analysis of nonorthologous host processes reveals host-specific patterns of induced defense responses that are consistent with disease and pathogen tolerance. (A and B) Heatmap of expression of log2(FC) > 1 genes from a list of significantly upregulated (n = 180) and downregulated (n = 199) genes at a single timepoint from maize, and significantly upregulated genes (n = 3,347) and downregulated genes (n = 3,068) from soybean at a single timepoint. Orange indicates significantly upregulated DE nonorthologous genes and blue indicates significantly downregulated DE nonorthologous genes. (C and D) Heatmap of significantly enriched gene ontologies from upregulated genes across pooled timepoints for maize and soybean. (E and F) Heatmap of significantly enriched gene ontologies from downregulated genes across pooled timepoints for maize and soybean, respectively, following F. virguliforme inoculation. (G) Mean expression patterns of log2(FC) > 1 of significantly upregulated and downregulated genes, respectively, at a single timepoint in both hosts, over the colonization timecourse. Error bars indicate one standard deviation of the mean.

2-1, Table AA 2-6 and 2-7), which identified the enrichment of several key biological

processes that presented as potentially host-specific (Figure 2-4G; Table AA 2-8 and 2-9, Data Set AAs 2-14 – 2-17). In soybean, enriched processes included “primary metabolite biosynthesis” and “signaling”, as well as “nucleotide catabolic processes”. In maize, this also included well-defined processes, such as those associated with the elicitation of the hypersensitive response, as well as signaling associated with response to chitin and fungus. Similar to orthologous defense gene comparisons described above, the GO process of “defense response” containing processes related to antifungal activity were upregulated earlier in maize than in soybean and include *PATHOGENESISRELATED MAIZE SEED (PRms)*, *TIFY10B* (containing *JASMONATE-ZIM DOMAIN* repressors), and hevein-like (Figure 2-5A). Interestingly, both *PRms* and hevein were previously identified as being induced following fungal colonization (Majumdar et al., 2017; Wong et al., 2017). Similarly, *TIFY10B* was also upregulated in maize, and further analysis revealed that key marker genes within the jasmonic acid (JA) biosynthesis pathway were minimally impacted (mock versus treated).

Previous studies have demonstrated that the JA pathway plays a crucial role in protecting plants against both pathogen colonization and wounding (Zhang et al., 2018b). Based on the observation that JA precursor biosynthesis gene expression was not altered by *F. virguliforme* colonization of maize (Figure 2-5B; Figure AA 2-11), we were curious as to how maize perceives and responds to *F. virguliforme* colonization. To address this, we next interrogated the expression patterns of several defense marker genes previously demonstrated as specific to maize root fungal colonization (Balmer et al., 2013b; Chuberre et al., 2018). As shown in Figure 2-5B, pathogenesis-related (*PR*) genes were induced over the timecourse of colonization. At the same time, while we detected the

induction of several *PR* genes, *NONEXPRESSER OF PR GENES 1 (NPR1)* was not upregulated in maize. While salicylic acid (SA)-dependent defense signaling does not specifically require the induced expression of *NPR1*, we hypothesize an additional bifurcation in SA-associated signaling, and that the observed upregulation of *PR* in response to F*. virguliforme*, herein, occurs in a SA independent manner (Balmer et al., 2013a). Additional hormone signaling processes, such as auxin and gibberellic acid biosynthesis, were also not significantly altered as a function of transcriptional expression throughout the timecourse of colonization. Interestingly, however, ethylene biosynthesis was rapidly upregulated (0 DAI) in maize roots (Figure AA 2-11), as was the expression of *ETHYLENE RESPONSIVE FACTOR 105-LIKE* (*ERF-105-like*). This is significant, as ethylene biosynthesis was recently reported to be induced in SDS-resistant soybean cultivars (Abdelsamad et al., 2019), a broader defense mechanism hypothesized to be associated with resistance to necrotrophic pathogens (Laluk & Mengiste, 2010).

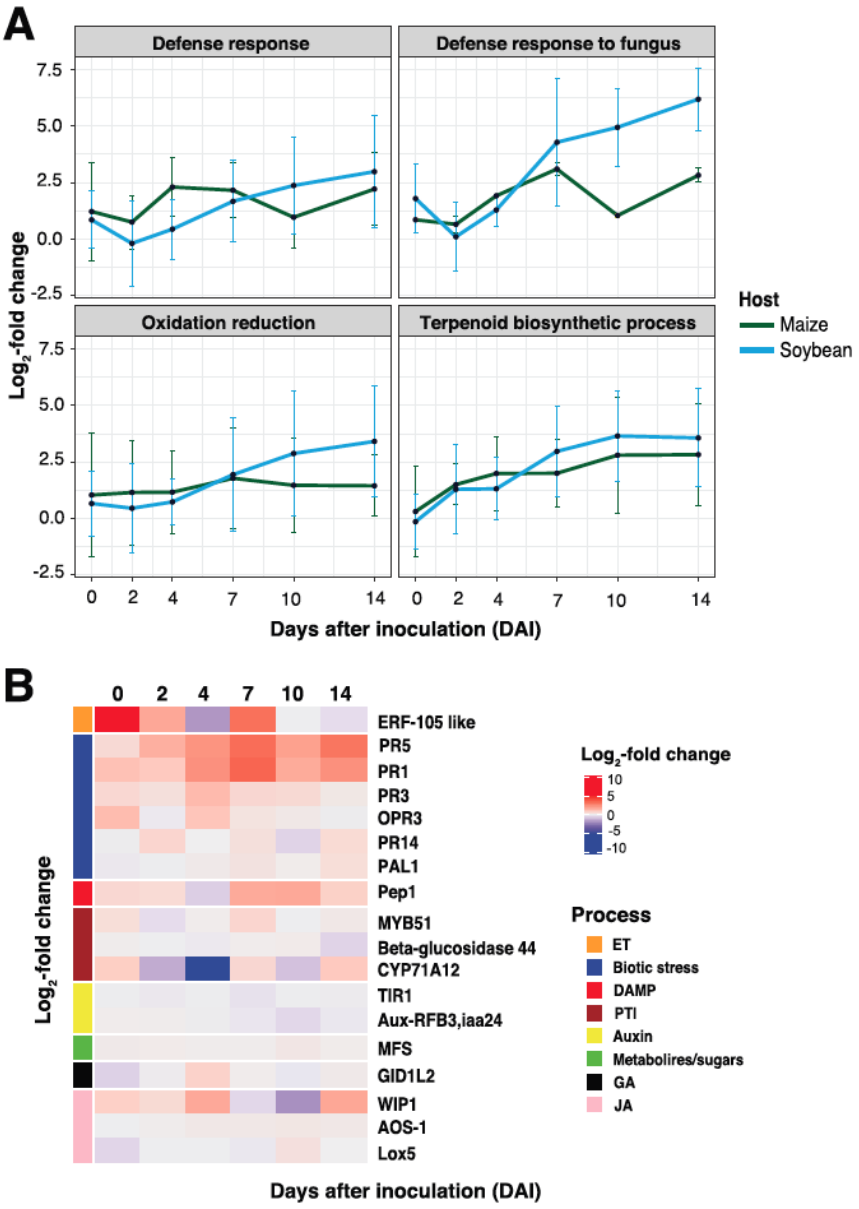

Figure 2-5: Processes unique to maize are associated with host immune responses following *F. virguliforme* inoculation. (A) Mean expression patterns of log2(FC) 41 of significantly upregulated nonorthologous genes from at least a single timepoint in both hosts over the colonization timecourse. Genes for analysis were selected based on those identified within the same gene ontology categories, but themselves are nonorthologous. Error bars indicate one standard deviation of the mean. (B) Gene expression profile of maize root defense markers that do not share a soybean orthologous group across the colonization timecourse.

Constitutive expression of maize genes within orthologous defense

To date, numerous studies have focused on induced plant defense responses following fungal colonization (van der Does & Rep, 2017). However, the role of constitutive defense-associated responses, particularly those in maize in response to *F. virguliforme* colonization, remains largely unaddressed. To determine if the duration and amplitude of constitutive orthologous gene expression underpins these responses, we next compared soybean and maize orthogroups following mock-inoculation, as opposed to fungal-treatment, detailed above. Using this approach, we successfully identified 182 orthogroups that were significantly upregulated in maize (compared to soybean) over the mock-inoculation timecourse (Figure 2-6; Data Set AA 2-18). Based on the output of this analysis (i.e. 182 orthogroups), we hypothesize that the function of these genes might provide insight into host tolerance to *F. virguliforme* (i.e. maize) through a mechanism associated with constitutive, sustained, elevated defense gene expression. Additionally, we hypothesized that the expression of these orthogroups will decrease over the course of fungal colonization in maize (i.e. tolerant, asymptomatic host), in contrast to a robust induction of defense-associated processes in soybean (i.e. susceptible, symptomatic host). Surprisingly, we identified a similar pattern of expression of constitutively upregulated orthogroups at 2–14 DAI in both fungal-inoculated hosts. Indeed, constitutively induced genes in maize were on average four-fold higher in the F. virguliforme treatment when compared to the constitutive expression of the same genes following mock-inoculation at 0 DAI. In summary, these expression profiles support our hypothesis that the rapid upregulation of a core group of defense associated transcripts

during the initial interaction between maize and *F. virguliforme* is associated with the establishment of the asymptomatic interaction.

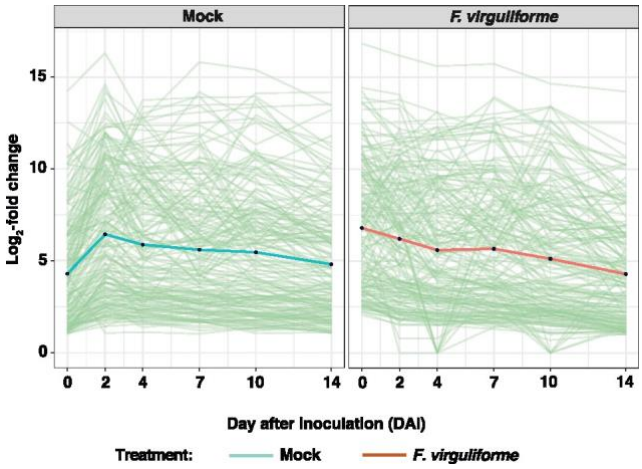

Figure 2-6: Conservation of defense gene expression patterns preceding inoculation with *F. virguliforme.* Line graph of expression patterns of orthogroups that are uniquely upregulated in maize when compared to soybean log2(FC) > 1, and corresponding regulation when maize and soybean are colonized by *F. virguliforme*. Individual green lines represent individual orthogroups. Solid green or brown lines represent the mean of orthogroups from mock- or *F. virguliforme*-inoculated.

Defense expression patterns of orthologous transcription factors is diverged across soybean and maize

The expression patterns of defense-induced orthologous genes in response to *F. virguliforme* inoculation varied by host, suggesting potential differences in host pathway activation, or repression. To define the regulation of these expression patterns within our orthologous soybean and maize datasets, we next compared the patterns of transcription factor (TF) expression (i.e. induced or upregulated) in each host following *F. virguliforme* inoculation. Similar to the orthologous genes identified as upregulated by *F. virguliforme*, the vast majority (93.3%) of the DE TFs were uniquely expressed within soybean (Data

Set AAs 2-19 and 2-20). Specifically, our analyses revealed that only 4 TF orthogroups were uniquely induced in maize (i.e. *bHLH-157, MYB-41, Homeobox-transcription factor 29*, and *GRAS-transcription factor 11*), while 11 were induced in both soybean and maize in response to *F. virguliforme* (Data Set AA 2-19).

To narrow our focus to the analysis of those TFs whose expression patterns contained altered gene expression in both hosts, we next evaluated the expression patterns of TFs that were upregulated in both maize and soybean following fungal colonization. As shown in Figure 2-7A (dashed black boxes), among the selected TF orthogroups upregulated in both species, we identified three TFs that were expressed at least four-fold higher in maize than soybean at 0 DAI. Among these was a C2H2-zinc finger TF, ZINC FINGER OF ARABIDOPSIS THALIANA12, previously identified as being induced in response to elevated hydrogen peroxide generation in maize (Mittler et al., 2006), and a NAC042 TF, which exhibited the largest expression change in comparison to mock versus fungal-inoculated maize. This is noteworthy, as *NAC042* was also shown to be responsive to hydrogen peroxide generation (Wu et al., 2012), illustrating that ROS production in maize may underlie the more than four-fold difference in expression compared with soybean. The third TF that was identified as induced was DEHYDRATION-RESPONSIVE ELEMENT-BINDING PROTEIN 1A, a gene regulated by *NAC042*, and hypothesized to function in the activation of oxidative stress in tomato (Thirumalaikumar et al., 2018). Based on the sum of these analyses, we hypothesize that maize rapidly attenuates ROS production following defense activation to dampen cellular stress. Of significance to this process is the potential role of *NAC042*, which has been shown to function as a negative regulator of senescence (Wu et al., 2012). Taken

together, we propose a model wherein *F. virguliforme* activates ethylene defense responses in maize, and that prolonged expression of *NAC042* leads to the promotion of cell longevity via dampening of cell stress responses and the inhibition of cell death (Figure 2-7C). This hypothesis is consistent with the observation that *NAC042* expression in soybean did not approach expression levels of more than four-fold change until 7 DAI, a timepoint that coincides with symptom development. Based on these data, we hypothesize that host senescence plays a key role in SDS development.

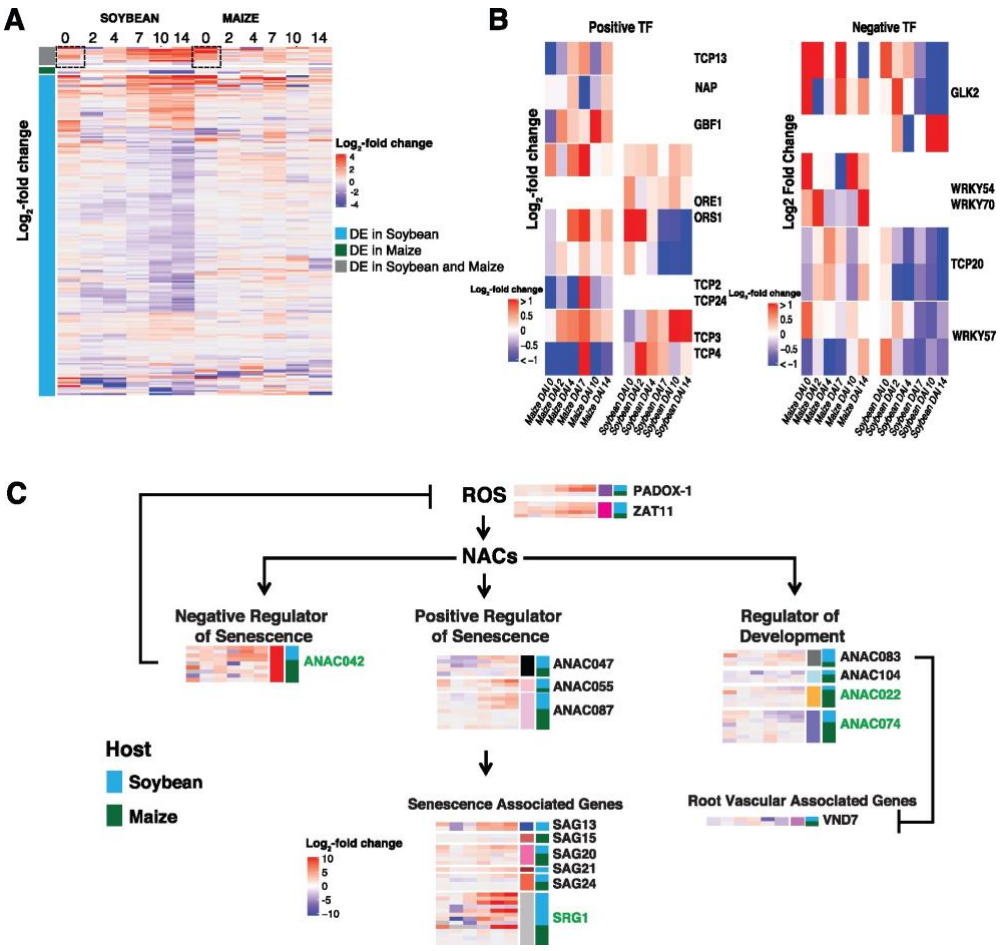

Figure 2-7: Divergence of defense expression patterns of orthologous transcription factors. (A) Heatmap of log2(FC) > 1 of significantly upregulated genes at a single timepoint in at least one host between mock and inoculated (n = 215). Gray indicates orthologous genes from soybean and maize were significant, blue indicates orthologous genes from only soybean were significant, and green indicates that maize orthologous genes were uniquely significantly DE. Black dashed boxes highlight changes in transcription factor expression between hosts. (B) Heatmap of positive-regulatory TFs in maize (n = 9) and soybean (n = 4), and negative regulatory TFs in maize (n = 5) and soybean (n = 3). (C) Representation of reactive oxygen species induced NO APICAL MERISTEM (NAM), ATAF1/2, CUP-SHAPED COTYLEDON-2 (CUC2) pathways in soybean that exhibited differential expression patterns of log_2(FTAbleC) > 1 between mock and pathogen-inoculated samples. The corresponding orthologs or homologs in maize associated with senescence and root vascular development. Each gene heatmap illustrates temporal changes from 0 to 14 DAI in soybean (blue) and maize (green). Green font indicates orthologous genes.

Fusarium virguliforme promotes susceptibility via regulatory biological processes that control senescence

Previous studies indicated that root senescence follows an intrinsically regulated developmental program that is regulated by multiple fundamental signaling pathways (Liu et al., 2019). As such, we hypothesized that pathogen-induced senescence is a key virulence strategy, which supports pathogen colonization and disease progression. To test this, we compared gene expression patterns in related GO processes between maize and soybean previously identified as required for senescence-associated signaling (e.g. fatty acid, chromatin, hydrolase process, response to water, JA, ethylene, pigment, carbohydrate, and ion relative biological process; (Li et al., 2019; Liu et al., 2019; Woo et al., 2013; Yang & Ohlrogge, 2009). As shown in Figure 2-8, A and B, we observed that some of the orthologs assigned to these GO terms were DE in each host (Data Set AA 2-21). For example, protein metabolic processes were generally unchanged in the early timepoints of colonization in maize yet were downregulated at later stages of the colonization. In contrast, in soybean, more than half of the DEGs were upregulated at the later stages of fungal colonization. In addition, genes involved in response to water deprivation were upregulated at early timepoints in maize. In soybean, transcripts assigned to these processes were generally not DE over the course of colonization. Indeed, and in further support of pathogen manipulation of senescence-associated signaling processes, we identified an increase in the expression of genes associated with pigment biosynthesis processes in maize (Zhao et al., 2018), yet in soybean, the DE of these processes was absent.

Based on the analysis above detailing the differential patterns of expression of defense associated orthologs from soybean and maize, we next queried the list of

differentially regulated nonorthologous genes. As shown in Figure 2-8, C and D, processes associated with protein metabolism, ROS metabolism, response to water derivation, and JA biosynthesis were upregulated in maize, while no significant changes were observed in soybean. Interestingly, the expression pattern of nonorthologous genes associated with these biological processes was similar to the pattern of orthologous genes (Figure 2-8, A and B; Data Set AA 2-21). Taken together, these results demonstrate that processes related to senescence were altered in maize, yet not affected in soybean during symptom development.

The data presented above highlight the importance of early, rapid, changes in host gene expression of *F. virguliforme* recognition and colonization, a process we hypothesize underscores the differential response, and outcome, between soybean and maize. To investigate the role of pathogen manipulation of host senescence-associated transcriptional reprogramming of growth and development, as well as the downregulation of host defenses, we first investigated the expression of TFs known to regulate senescence associated processes, including those previously identified as positive and/or negative regulators of leaf senescence (Koyama, 2014; Koyama et al., 2017). Using this published dataset as a template to guide our discovery of root associated senescence processes, we identified 9 and 5 previously characterized positive and negative regulators, respectively, of leaf senescence. As shown in Figure 2-7B, we observed a general trend of downregulation of maize gene expression profiles at early timepoints (i.e. 0–4 DAI) of TFs associated with the activation of senescence-associated processes (Data Set AA 2-22). In support of our hypothesis that pathogen-induced host senescence is associated with disease symptom development in soybean, we observed

that the same positive regulatory TFs previously identified (i.e. (Koyama et al., 2017) were significantly upregulated in soybean, supporting the hypothesis for a mechanism of host transcriptome reprogramming in favor of induced host senescence. Specifically, among the identified induced TFs identified was the senescence-positive factor Teosinte Branched1/Cycloidea/Proliferating cell factors, which promotes the expression of JA biosynthetic enzyme genes (Koyama, 2014; Schommer et al., 2008). As predicted, TFs that negatively regulate senescence-associated processes, such as WRKY54 were significantly induced in maize, while in soybean, their induction was not observed. Taken together, these results suggest that the regulation of host senescence during the early stages of pathogen colonization is a key process associated with susceptibility and disease development.

As described above, we observed that the upregulation of positive-senescence-associated TFs decreased in maize compared to soybean at early timepoints of the colonization, while the pattern of expression of negative-senescence associated TFs increased in comparison with soybean. To further test the hypothesis that host senescence plays a key role in SDS development, we interrogated the soybean transcriptome for significant expression changes in *ANAC* (*ABSCISIC-ACID-RESPONSIVE NAC* (*NAM, ATAF1,2,* and *CUC2*); (Podzimska-Sroka et al., 2015) genes. Using this approach, we identified an additional seven *ANACs* that were DE in soybean but not in maize. Excitingly, and in further agreement with our hypothesis, previous work demonstrated that these ANACs play a role in both the negative and positive regulation of senescence, as well as in root development (reviewed in (Majid I, 2019); Figure 2-7C; Data Set AA 2-23). Specifically, three *ANACs—ANAC047*, *ANAC055*, and *ANAC087—*

have established roles in senescence-associated signaling in *Arabidopsis*, and two, *ANAC055* and *ANAC087*, have been demonstrated to be abscisic acid (ABA) responsive (Huysmans et al., 2018). In the context of the current work, we further observed that the expression of *9-CIS-EXPOSYCAROTENOID DIOXYGENASE*, the product of which is required for the first step in ABA production (Xiong & Zhu, 2003), increased at 2 DAI in soybean, and was followed by a 16-fold increase in expression of *ANAC055* and *ANAC087* (Supplemental Figure 2-10). Interestingly, these genes were not upregulated in maize. Along this same theme, several senescence-associated genes (*SAGs*) were also upregulated in soybean, but not in maize, as early as 4–7 DAI. This is noteworthy, as previous work showed that expression of *SENESCENCE-ASSOCIATED GENE 13* (*SAG13*) is associated with the cell death response (Pell, 2004), with additional work revealing that *SAG21* and *SAG24* are upregulated during early senescence (Salleh et al., 2012). Interestingly, *SAG20* expression was found to be upregulated by necrosis- and ethylene-inducing peptides from *Fusarium* spp., as well as by phytotoxins produced by *F. virguliforme* (Chang et al., 2016).

As a potential link to the development-associated gene networks in each of the hosts, we observed that *NAC* TFs (specifically those associated with cell development) were also induced in soybean roots infected with *F. virguliforme*. This is interesting, as previous studies showed that *NAC022* plays a significant role in lateral root growth (Xie et al., 2000); herein, *NAC022* was downregulated following *F. virguliforme* inoculation. In a similar manner, *NAC074*, which positively regulates xylem development (Xia, 2018), was observed as being downregulated in soybean in response to *F. virguliforme* colonization. Likewise, *NAC083*, which plays a role in vascular development (Yamaguchi

et al., 2010), was upregulated over the timecourse of colonization. In total, the expression profiles of key development associated regulators identified in the current study support a role for the alteration of root development and the concomitant acceleration of senescence by *F. virguliforme* as a potential mechanism of disease symptom development and susceptibility in soybean.

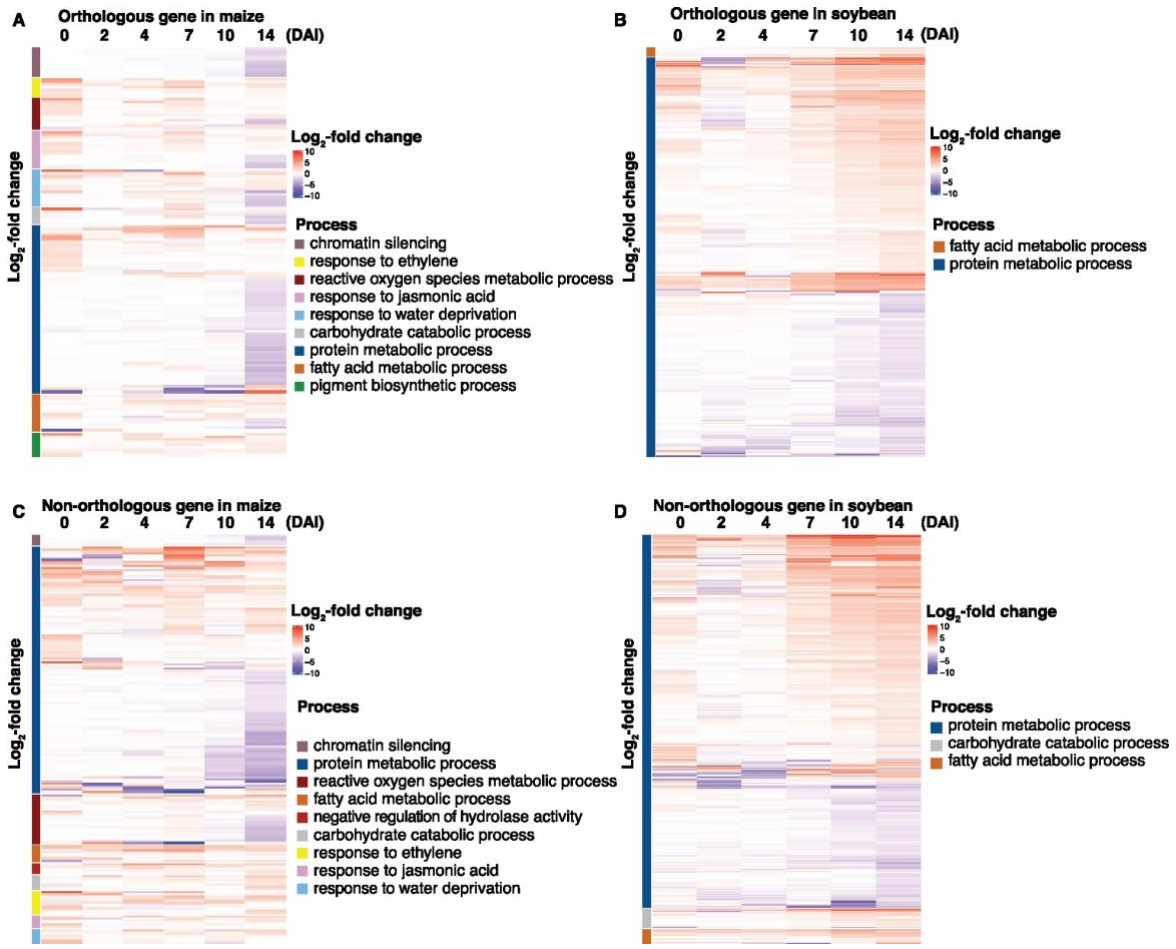

Figure 2-8: Gene expression patterns of regulatory pathways that control senescence-associated processes. (A and B) Gene expression pattern of orthologous genes in maize (103) and soybean (376). (C and D) Gene expression pattern of nonorthologous genes in maize (159) and soybean (668).

Discussion

Whole-transcriptome analyses of resistant and susceptible host responses to phytopathogens are a common approach to evaluate host–pathogen dynamics during colonization. In total, the sum of these studies provided a foundation to comparatively explore how immune pathways across diverged hosts respond to a single, common, pathogen. In the current study, we exploited the broad host range of *F. virguliforme* as a comparative system to probe differentially induced root responses between a monocot and a eudicot to uncover immune regulatory responses encoding tolerance and susceptibility. To do this, we generated 72 transcriptomes from mock and inoculated hosts to identify host responses specific to fungal root colonization.

Soybean and maize have distinct phenotypic responses to *F. virguliforme* colonization. For example, soybean shows marked signs of symptom development as early as 7 DAI, with chlorosis manifesting at 10 DAI, and tap root necrosis at 14 DAI. In maize, disease symptoms do not develop in response to *F. virguliforme* colonization, yet as noted above, maize supports similar levels of fungal colonization as observed in soybean. In total, this apparent dichotomy suggests a difference in the underlying genetic and/or physiological interactions between the host and pathogen that yields asymptomatic versus symptomatic colonization. For example, an analysis of maize revealed that root growth and development were unaltered in response to *F. virguliforme* colonization. However, and contrary to the lack of visible changes, we did identify significant changes in gene expression following colonization, most of which included immune response-associated genes. Conversely, an analysis of soybean revealed a massive transcriptomic reprogramming of defense signaling responses concomitant with

symptom development. These observations are in agreement with a divergence in the magnitude of host responses as previously noted in the analysis of a single host independently colonized by both pathogenic and nonpathogenic isolates (Lanubile et al., 2015).

While the number of host genes induced by *F. virguliforme* varied substantially, both hosts exhibited gene expression changes associated with marked enrichments in defense related ontologies, indicating that the temporal induction of immunity plays a predominant role in the tolerance of *F. virguliforme* by maize. Indeed, when we compared the expression of orthologous defense-related processes, we observed a strikingly different response between soybean and maize, with very few orthologous genes identified as uniquely DE in both hosts. This observation supports our hypothesis of a significant divergence in the transcriptome of orthologous genes. Indeed, orthologs encoding defense responses identified as induced by *F. virguliforme* exhibited disparate temporal patterns of induction. This is significant, as this timeframe coincides with the recognition of the fungus and the subsequent initiation of host defense responses. This also coincides with the developmental transition of the pathogen from a biotrophic phase to a necrotrophic phase in soybean. Similar shifts in host gene expression as a function of fungal lifestyle changes during colonization have been described previously in the case of hemibiotrophic interactions (Chowdhury et al., 2017; V. N. Njiti, 1996). Coincident with developmental transitions in pathogen lifestyle, defense-related processes typically upregulated in infected hosts are commonly associated with processes, for example, associated with hydrogen peroxide production and the release of anti-microbial compounds.

In the current study, our analyses reveal that as fungal colonization develops over time, numerous orthologous defense genes in maize were downregulated, indicating a dampening of host defense signaling by the fungus. While this observation is consistent with previous work evaluating gene expression change in resistant cultivars (Chen et al., 2016), there are limited examples of this occurring in an asymptomatic host. Not unexpectedly, we observed that soybean defense responses increased to maximal induced levels at 10–14 DAI, further highlighting a delayed host response to F. virguliforme colonization. Taken together, these data point to the temporal activation of defense signaling as the underlying mechanism distinguishing symptomatic versus asymptomatic responses. For example, nonorthologous defense-induced responses exhibited equivalent trends between hosts for a majority of defense-associated processes. Again, these data support a role for transcriptional control (i.e. TF expression) as a critical process underpinning *F. virguliforme* tolerance in maize.

The analyses described herein point to a key role for the early recognition and activation of defenses, including both the timing and amplitude of expression. For example, our data suggest that the early activation of defense responses in maize may stem from the specific induction of ethylene and ROS production, a process that was previously identified as a process associated with reduced disease severity in soybean in response to *F. virguliforme* colonization (Abdelsamad et al., 2019). The analyses presented herein highlight a potential role for ERF-105 (Bolt et al., 2017) and NAC042 (Thirumalaikumar et al., 2018) as key regulators of symptomatic versus asymptomatic signaling during fungal colonization of soybean and maize, respectively. Indeed, as *F. virguliforme* transitions from a biotrophic to necrotrophic lifestyle on soybean, a

substantial transcriptomic rewiring occurs. Necrotrophic plant pathogens promote plant cell death, which results in dramatic shifts in host gene expression and cellular metabolism. Through an analysis of host gene expression changes over the timecourse of colonization, we identified changes in the regulation of several senescence-related TFs in soybean following pathogen colonization. As a result of these changes in gene expression, downstream signaling through additional senescence-related genes within soybean occurred, the outcome of which is the inhibition of protein synthesis, the hydrolysis of macromolecules, and the degeneration of cells (Podzimska-Sroka et al., 2015). As a function of pathogen proliferation and host cell death, the observed onset of senescence in soybean is associated with the timing of fungal pathogen necrotrophy and the appearance of disease symptoms. Similar processes have previously been described during pathogen infection and defense and include the production of ROS—a defense mechanism that is concomitant with the induction of NAC-associated host senescence (reviewed in (Haffner et al., 2015).

As a further link to the work described herein, senescence-associated processes which signal through NAC TFs are also induced by abscisic acid. This is noteworthy, as we also identified the induced expression of genes associated with abscisic acid biosynthesis from samples collected from soybean roots at 2 DAI. This also includes the observation of *NAC055* and *NAC087* induction at 4 and 7 DAI, respectively. We posit that F. virguliforme susceptibility in soybean is mediated in part by pathogen-induced senescence processes, similar to those previously described in the case of Sclerotinia sclerotium (Williams et al., 2011). Based on the data presented herein, we predict that future research in this area will help clarify previous observations suggesting that plant

senescence and defense share many overlapping, critical, signaling pathways. These include, for example, the modulation of expression of key NAC TFs, most of which have demonstrable roles in immunity and susceptibility (Bu et al., 2008; Saga et al., 2012; Xi Yuan, 2019) What remains unclear is if gene expression is an indicator of pathogen-triggered senescence or host-induced susceptibility, or both. However, the mounting evidence presented herein, and previously seems to suggest that necrotrophic pathogen manipulation of genetic pathways leading to a favorable host environment are evident (Chowdhury et al., 2017; Haffner et al., 2010; Podzimska-Sroka et al., 2015; Williams et al., 2011). In conclusion, the data presented in the current study demonstrate that the intersection of pathogen lifestyle, host genetics, and plant development play a key role in the interaction of defense, virulence, and processes of pathogen-accelerated host senescence.

Materials and Methods

Plant and F. virguliforme assay

Soybean (*Glycine max* cv. Sloan; gift from Martin Chilvers, Michigan State University) and maize (*Zea mays*) hybrid E13022S (Epley Brothers Hybrids Inc., gift from Martin Chilvers) were surface sterilized in 70% ethanol for 30 s, followed by 10% bleach for 20 min, and then rinsed three times with sterile distilled water for 1 min. Surface sterilized soybean seeds were placed between two sheets of sterile Whatman filter paper soaked with 5 mL of sterile water, which was then placed inside a sterile Petri dish. Surface sterilized maize seeds were incubated in the dark in sterile water for 24 h and were then placed between two sheets of sterile 100-mm Whatman filter paper with 5 mL

of sterile water inside a Petri dish. All seeds were incubated in the dark for 5 days at 21°C before transplanting.

The *F. virguliforme* Mont-1 isolate was propagated for 7 weeks on potato dextrose agar (Difco, Thermo-Fisher). Asexual macroconidia spores were collected, diluted to 1 X 10^5 macroconidia mL–1, and sprayed onto 5-day-old maize or soybean seedlings using a 3-oz travel spray bottle. Twenty-five sprays were applied to the seedlings at angles of 0°, 90°, 180°, and 270° to ensure that seedlings were thoroughly inoculated. For mock-inoculated samples, sterile water was sprayed onto the seedlings. Seedlings were incubated for 30 min with the inoculum (including mock-inoculated), after which time the excess inoculum was removed, and seedlings were incubated for an additional 1 h. Three seedlings (maize and soybean) were placed into CYG germination pouches (Mega International; CYG-19LB) containing 25 mL of sterile distilled water, and the pouches were transferred to a BioChambers Bigfoot Series Model AC-60 growth chamber with 140 $\mu E\ m^{-2}\ sec^{-1}$ and 14:10 h light/dark cycle at 12°C for 7 d and then 25°C for 7 d. Plants were watered as needed with sterile water. Tap roots from soybean, or radicals from maize, were collected at the same time of day (i.e., 16:00 h) from the original 4-cm inoculation site. The 2-week timecourse was repeated three times (independent biological replicates) in the same growth chamber, with sampling of six plants for RNA isolation and three plants for DNA isolation at 0, 2, 4, 7, 10, and 14 DAI for each biological repeat. Sampling of timepoint 0 was performed after completion of mock and fungal inoculations at 1-h post-fungal inoculum application. Plant growth and disease symptomology was recorded at each timepoint by photography with a D50 Nikon camera.

Analysis of fungal colonization: visual phenotypes and quantification

To visualize fungal growth on plant roots, microscopy analyses of maize and soybean were conducted at each timepoint for all treatments. Roots (n = 6) were cleared in 100% ethanol, followed by staining in a 0.05% trypan blue solution containing equal parts of water, glycerol, and lactic acid (Savory et al., 2012). Fungal structures were observed using a MZ16 dissecting microscope (Leica).

To determine the amount of fungal biomass present in inoculated samples over the duration of the colonization timecourse, genomic DNA was extracted from flash-frozen root tissue and used for real-time quantitative polymerase chain reaction (qPCR) analysis. A total of 60 mg of ground root tissue was extracted from individual maize or soybean plants from each timepoint. DNA was extracted using the NucleoSpin Plant II Kit (Macherey-Nagel) according to the manufacturer's protocol, with an additional incubation of 1 h at 65°C for the first step in DNA isolation (i.e., cell lysis step). DNA samples for qPCR-based detection of *F. virguliforme* were prepared following the method of Wang et al. (2015). The 5′-end of the *F. virguliforme* TaqMan probe (prb) was labeled with 6-fluorescein (6-FAM; Life Technologies), whereas the 3′-end was modified with a minor groove binder nonfluorescent quencher (MGBNFQ; Thermo-Fisher). PrimeTime dual-labeled probes were labeled with 5′ 6-FAM, internal ZEN quencher and 3′ Iowa Black FQ quencher (3IABkFQ; IDT Technologies). The exogenous control HHIC (Haudenshield and Hartman Internal Control) assay primers and probe were purchased from IDT Technologies and are described in (Haudenshield & Hartman, 2011). All DNA primers for *F. virguliforme* quantification are listed in Table AA 2-9 Real-time qPCR reactions were performed using the Applied Biosystems 7500 Fast Real-Time PCR System (Thermo-

Fisher). At least 3 biological replicates containing 3 technical repeats were performed for each timepoint. Reaction mixtures consisted of 10 μL of TaqMan Universal real-time PCR master mix (2X) (Applied Biosystems), 2 μL of DNA, 0.5 μL of FvPrb-3 (*Fv* TaqMan probe; 10 μM), 0.5 μL of F6-3 and R6 primers (20 μM, each), 0.6 μL of HHIC-F (forward primer; 20 μM) primer, 0.2 μL of HHIC-R (reverse primer 20 μM), 0.4 μL of HHIC-prb (10 μM), 0.5 μL of linearized HHIC DNA plasmid (10 fg/μL; Haudenshield and Hartman (2011), 0.4 μL of bovine serum albumin (BSA; New England BioLabs, catalog # B9000S) at 10 mg/mL, and 4.4 μL of distilled water (Thermo-Fisher). Cycling conditions were as follows: 1 cycle at 50°C for 2 min, 1 cycle at 95°C for 10 min, and 40 cycles at 95°C for 15 s, and a final step at 60°C for 1 min. Fluorescence data was collected during the annealing and extension stages of the program.

Analysis of variance was calculated for DNA quantities using the lme4 (Bates, 2015) and Car (Fox, 2011) packages in R (v3.4.1; (Team., 2010); Table AA 2-1). Fungal DNA means were separated by Tukey's least significant difference test using the multicomp package (Hothorn et al., 2008) at $P \leq 0.05$.

RNA extraction

Root tissue from mock and infected samples from 0, 2, 4, 7, 10, and 14 DAI was used for RNA-sequencing. Biological samples were pooled from six independent plants, and for RNA-sequencing, three independent replicates were included from three independent growth chamber experiments. Sampling of plant tissue from both mock and inoculated roots at each timepoint allowed us to discover genes specifically related to fungal colonization. Additionally, all tissue was sampled from the same inoculation site within the root, further enabling us to explore responses within the inoculation site

stemming from fungal adhesion, colonization, and proliferation. mRNA was prepared using the miRNeasy Mini Kit (Qiagen) from a total RNA isolation derived from 200 mg of either ground flash frozen germinating macroconidia or plant root sample. Contaminating genomic DNA was removed using the TURBO DNase Free kit (Invitrogen). RNA quality was evaluated by gel electrophoresis using the 2100 Bioanalyzer (Agilent) in combination with the Agilent RNA 6000 Pico kit, according to the manufacturer's instructions.

Library preparation and sequencing

mRNA libraries were prepared from three biological repeats of each timepoint of *F. virguliforme* or mock inoculated maize or soybean or germinating macroconidia sample using the Illumina TruSeq mRNA Library Preparation Kit by the Michigan State Research Technology and Support Facility. Samples were pooled and sequenced on the Illumina HiSeq 4000 (single end 50-bp mode). Base calling was performed using the Illumina real-time analysis (RTA; v2.7.7), and the output of RTA was demultiplexed and converted to FastQ format using Illumina Bcl2fastq (v2.19.1). Sequencing of each sample was performed to an average yield of 70 million reads.

Quantification of RNA-seq expression and differential analysis

Reads were adapter-trimmed and evaluated for read quality using Trimmomatic (v0.33; (Bolger et al., 2014). After trimming for adapters and quality thresholding, uniquely mapped soybean reads ranged from 76% to 83%, while maize reads were in the range of 77%–81% (Figure AA 2-2A and Data Set AA 2-1). Resultant trimmed reads were mapped to the corresponding reference genomes of soybean (Wm82.a2.v1) and maize (B73 RefGen_v4, AGPv4) using HISAT2 (v 2.1.0; (Kim et al., 2015). The following parameters were applied: –dta–rna-strandness F.

Hits from HISAT2 were converted from SAM to BAM format by Picard (v2.18.1; http://broadinstitute.github.io/picard/). Sequence alignments were counted using HTSeq (v0.6.1; (Anders et al., 2015) with the following options: –minaqual 50-m intersection-strict-s reverse–idattr=gene_id. Gene counts were imported into DESeq2 (v1.22.2; (Love et al., 2014), executed in R (Team., 2010), normalized for library size, and log2 transformed to determine the correlation of biological replicates at each timepoint.

To determine differential gene expression, DESeq2 (v1.22.2), executed in R with raw HTSeq counts, was used. Gene counts of less than 10 were excluded. DESeq2 was used to identify DE transcripts which met the requirement of an adjusted $P \leq 0.05$ and greater than $|1\text{-}\log_2|$ fold difference in expression between mock- and *F. virguliforme*-inoculated samples. Pairwise comparisons were evaluated at each timepoint and within each host, including mock- and *F. virguliforme*-inoculated samples.

Differential gene co-expression network analysis

Gene counts were filtered for differential gene correlation analysis (DGCA; implemented in R) (McKenzie et al., 2016) for 90% of genes with less than 10 reads across all samples. The resultant 43,308 soybean and 28,956 maize genes were variance stabilized, transformed for importation, and the Pearson correlation of individual gene pairs within each treatment was calculated and compared to mock- or *F. virguliforme*-inoculated treatments. To classify differential correlation significance, DGCA transforms correlation coefficients to z-scores and uses differences in z-scores to calculate P-values of differential between gene pairs (e.g. $\rho < 0$, adjusted $P < 0.05$ for negative correlation, and $\rho > 0$, adjusted $P < 0.05$ for positive correlation), with a Bonferroni correction of P-value (i.e., (McKenzie et al., 2016). Significant gene pairs from differentially induced

classes (e.g., +/0, -/0, 0/-, +/0, +/-, and -/+) were weighted by the z-score difference between treatments to convert into a planar filtered network (PFN) assembled upon unique links between significantly differentially correlated gene pairs within a spherical surface to derive with the most relevant information from a similarity matrix based on a topological sphere. Gene pairs were imported into MEGENA and multiscale modules and hubs were identified through partitioning of the parent PFN by k-splitting via evaluation of network compactness through multi-scale embedded gene co-expression analysis, with a hub detection significance threshold of $P < 0.05$, module significance threshold of $P < 0.05$, network permutations of 100, and module size greater than 10 (Song & Zhang, 2015). Differential gene expression correlation analysis of samples from maize identified 128 significant gene pairs from 419,152,582 total gene pairs between mock- and pathogen-inoculated. Analysis of the correlation between soybean gene expression identified 5,526,057 significant gene pairs from 904,379,186 gene pairs between treatments. Soybean differential gene pairs were clustered into 1,161 modules using MEGENA (https://cran.r-project.org/web/packages/MEGENA/index.html). Modules were visualized in R using ggplot2 (v3.1.1; (Wickham, 2016).

Identification of orthologous genes

Protein sequences were retrieved from Ensembl Plants (http://plants.ensembl.org/index.html). The longest protein sequences for genes from soybean and maize were analyzed by OrthoFinder using program default settings (v2.2.7; (Emms & Kelly, 2019). A total of 23,274 soybean and 14,542 maize genes were discovered and subsequently classified into 10,700 orthogroups. This dataset was next analyzed to filter log2-fold changes between mock and inoculated maize or soybean at

the orthogroup level. Because many orthogroups contained more than a one-to-one relationship, the median of log2-fold changes of all genes within one host orthogroup was evaluated. The median transformation of genes within an orthogroup captured the most variation between maize and soybean (Data Set AA 2-9).

Differential expression analysis of orthogroups

To determine differential orthogroup expression, DESeq2 (v1.22.2; executed in R) was employed using the median transformed gene HTSeq counts within each orthogroup. Orthogroup counts with a value less than 10 were filtered from the analysis. Significant differential expression patterns were determined using an adjusted $P \leq 0.05$. A greater than one-fold difference between mock treatments between maize and soybean, as well as inoculated treatments between maize and soybean, at each timepoint was used as a cutoff for analyses.

GO enrichment analysis

GO was retrieved from the GFF file of the soybean genome annotation (Wm82.a2.v1) and the revised GO annotation for maize (Wimalanathan, 2018). The GO annotation varies in completeness between soybean and maize. For example, the defense response term (GO:0006952) has 374 annotated genes; however, in soybean, only 86 genes are associated with this GO annotation. To generate an ontology that is similar for an ortholog comparison, we merged gene ontologies from maize and soybean within each orthogroup and removed redundant ontology terms. In total, all 10,700 orthogroups had GO assignments.

Orthogroup lists from differential analyses were analyzed by TopGO (2.34.0) conducted in R (Alexa A., 2018). Fisher's exact test, set at an adjusted $P \leq 0.05$, was

conducted on each orthogroup to determine the relative significance of enrichments across all orthogroups. Additionally, lists of genes from differential analysis or gene differential clustering modules were analyzed for GO term enrichment by the singular enrichment analysis using the AgriGO (v2) website (Tian et al., 2017) with the Fisher statistical adjustment method set at a significance level of 0.05.

Orthologous TF analysis

TF sequences from maize were retrieved from: http://planttfdb.cbi.pku.edu.cn/. Maize genome v3 gene IDs were converted to v4 IDs for the 2,290 TFs and orthologous genes in soybean; this included 508 orthogroups. Orthogroups were filtered for significant defense induction over the timecourse.

Positive and negative senescence TFs analysis

Positive and negative senescence TFs were selected from previous studies (Koyama, 2014), with a total of nine positive TFs and nine negative TFs chosen for further expression analysis herein. The gene expression value of each TFs was derived from mock treatments between maize and soybean, as well as inoculated treatments between maize and soybean, at each timepoint after quantification of differential gene expression.

Simplify GO enrichment analysis

GO enrichments from *F. virguliforme* (Baetsen-Young et al., 2020), maize, and soybean were analyzed by using a simplified enrichGo method in the clusterProfiler package of R (Yu et al., 2012) to reduce redundant GO terms from the generated output of the enrichGO function. All user-defined parameters were set to the program default mode.

GO analysis

Related gene ontologies of *F. virguliforme* (Baetsen-Young et al., 2020) were selected based on its biotrophy–necrotrophy switch in pathogen lifestyle(Doehlemann et al., 2017). Related gene ontologies of maize and soybean were selected based on the biological processes related to response to *F. virguliforme* colonization (Vargas et al., 2012). Gene count numbers were calculated based on gene numbers that sorted into each selected GO category and as a function of up- or downregulated expression changes.

Accession numbers

The RNA-seq data generated herein are contained within the National Center for Biotechnology Information (NCBI) Short Read Archive (SRA) in BioProject under project ID: PRJNA549915.

CHAPTER 3: NDR1 and the *Arabidopsis* plasma membrane ATPase AHA5 are required for processes that converge on drought tolerance and immunity

The work presented in this chapter:

from Yi-Ju Lu*, Huan Chen*, Alex Corrion* et al. 2021

In Revision; Plant Physiology

Chapter abstract

NON-RACE-SPECIFIC DISEASE RISISTANCE1 (NDR1) is a key component of plant immune signaling, required for defense against the bacterial pathogen *Pseudomonas syringae*. Plant stress responses have overlapping molecular, physiological, and cell biology signatures, and given the central role of NDR1 during biotic stress perception and signaling, we hypothesized that NDR1 also functions in abiotic stress responses, including in a role that mediates signaling at the plasma membrane (PM) - cell wall (CW) continuum. Here, we demonstrate that NDR1 is required for the induction of drought stress responses in plants, a role that couples stress signaling in an abscisic acid-dependent manner. We show that NDR1 physically associates with the PM-localized H+ ATPases AHA1, AHA2, and AHA5 and is required for proper regulation of H+-ATPase activity and stomatal guard cell dynamics, providing a mechanistic function of NDR1 during drought responses. In the current study, we demonstrate that NDR1 functions in signaling processes associated with both biotic and abiotic stress response pathways, a function we hypothesize represents NDR1's role in the maintenance of cellular homeostasis during stress. We propose a role for NDR1 as a core transducer of signaling between cell membrane processes and intercellular stress response activation.

Introduction

Crop productivity is impacted by a myriad of abiotic and biotic stresses, and among these, water availability and disease pressure are the two primary factors limiting performance and yield (Atkinson & Urwin, 2012; Fahad et al., 2017). In recent years, many of the genetic and physiological processes that underpin plant resilience in response to both abiotic and biotic stress, as well as the convergence of signaling

mediating growth, development, and response to the environment, have been defined(Des Marais et al., 2013). Indeed, work in this area has demonstrated that plants rely on highly conserved genetic mechanisms that underscore growth and development (Guo et al., 2018), and environmental stress tolerance (Saijo & Loo, 2020) during immune signaling. Particularly noteworthy is the innate ability of plants to respond to, and often overcome, the onset of multiple, simultaneous, stress events (Rejeb et al., 2014). Related to the work described herein, several recent studies have identified substantial signaling overlap between a battery of stress signaling processes, including the convergence of processes governing immunity, adaptation to changes in temperature, and water and nutrient availability(Atkinson & Urwin, 2012).

Among the best-characterized mechanisms of water stress tolerance and adaptation are signaling processes linked to reductions in transpiration, a process mediated by the rapid closure of stomata on the leaf surface via the action of the plant hormone abscisic acid (ABA; (Sussmilch et al., 2017; Sussmilch & McAdam, 2017)). Additionally, and as a mechanism broadly classified as protection and maintenance of cellular homeostasis, the modulation of membrane dynamics is a primary physiological process associated with dehydration-associated responses in plants(Jarzyniak & Jasinski, 2014). These include, for example, the regulation of signaling associated with the maintenance of cellular membrane homeostasis and metabolic dysfunction(Garcia-Gomez et al., 2000; Mahajan & Tuteja, 2005). As a second, downstream layer of stress response signaling during water deficit, plants also activate numerous dehydration responsive genes(Nakashima et al., 2014; Yamaguchi-Shinozaki & Shinozaki, 2006), activate the production of LEA/dehydrin-type proteins(Singh et al., 2005) and various

molecular chaperones(Jacob et al., 2017), as well as rely on cell detoxification mechanisms for the removal of reactive oxygen species (ROS; (Noctor et al., 2014)). Not surprisingly, disease resistance signaling also engages similar cellular processes and signaling mechanisms.

Abscisic acid is one of several common denominators linking immunity and abiotic stress signaling(Qi et al., 2018). Indeed, several studies in this area have demonstrated that application of exogenous ABA enhances disease susceptibility(Fan et al., 2009; Xu et al., 2013; Yasuda et al., 2008), while insensitivity to ABA enhances disease resistance(Cao et al., 2011; Mang et al., 2012). As a regulator of response to conditions of limited water availability, the accumulation of abscisic acid (ABA) is required for the activation of abiotic stress responses, including downstream gene activation and cellular physiological changes. Among the most important downstream cellular changes in response to rapid accumulation of ABA are rapid turgor presser changes to close stomata. In support of the research described herein, several recent studies have identified significant overlap between water stress response signaling and the regulation of plant defense and immunity(Davila Olivas et al., 2016; Gupta et al., 2017).

In the current study, we demonstrate that NDR1 (NON RACE-SPECIFIC DISEASE RESISTANCE-1) is required for the robust transcriptional activation of key ABA-associated stress signaling responses in plants, including those associated with both biotic and abiotic signaling mechanisms. Using a combination of physiological, genetic, and whole-transcriptome-based analyses, we further show that a loss of *NDR1* results in a block in stomatal-based mechanisms impacting the response to abiotic and biotic stress signaling. As a mechanism underpinning these responses, we demonstrate that NDR1

associates with the plasma membrane-localized H^+-ATPases AHA1, AHA2, and AHA5, the function of which is to negatively regulate the activity of the guard cell-localized ATPase. We further demonstrate that overexpression of *NDR1* leads to enhanced drought tolerance, suggesting a role for NDR1 in the regulation of processes mediating cellular response, and tolerance, to drought. In fulfillment of a broader role of NDR1, the data presented herein supports a function for NDR1 as a signaling hub required for response to stress.

Results

Previous work demonstrated that a loss of *NDR1* and a concomitant increase in electrolyte leakage is correlated with a loss in focal adhesion between the plasma membrane and cell-wall(Knepper et al., 2011a). Not surprisingly, similar focal adhesion mechanisms have been described in response to abiotic stress elicitation(Feng et al., 2016). To determine if NDR1 functions in abiotic stress perception and signaling, we first evaluated the *ndr1* mutant for response to water withholding. As shown in Figure 3-1a and Figure AA 3-1, WT Col-0, the *ndr1* mutant, the *NDR1*-complementation line (*ndr1/*P_{NDR1}*::T7-NDR1*; (Knepper et al., 2011a)), and the *NDR1* overexpression line (*ndr1/35S::NDR1*; (Coppinger et al., 2004)) showed similar responses at 14 days-post water-withholding (dpw). However, at 16 dpw, the expression of phenotypes associated with water deficit (e.g., leaf wilting) became pronounced in the *ndr1* mutant. Interestingly, under the same conditions, *NDR1* overexpression plants (i.e., *ndr1/35S::NDR1*) did not show outwardly visible signs of drought stress at 21 dpw (Figure 3-1a and Figure AA 3-1), suggesting that overexpression of *NDR1* ameliorates drought-associated phenotype development.

To further evaluate the physiological impact of water withholding and the onset of drought stress across genotypes, we first monitored electrolyte leakage and relative water content (RWC) over the timecourse of water withholding. As shown, from Day 12 until ~16 dpw, all plant genotypes maintained a steady-state level of electrolytes (Figure 3-1b) and RWC (Figure 3-1c). However, from 17-21 dpw, drought phenotypes became more pronounced, concomitant with enhanced electrolyte leakage and a significant loss in RWC. For example, at 18 dpw, the *ndr1* mutant showed an approximate 20% increase in electrolyte leakage compared to WT Col-0 and the *NDR1*-expression lines. Consistent with this, at 21 dpw, electrolyte leakage reached ~60% in the *ndr1* mutant, which represents a 20% increase compared to WT Col-0 and a ~50% increase in leakage compared to the *ndr1/35S::NDR1* overexpression line (Figure 3-1b). Similarly, a decrease in RWC was first observed in the *ndr1* mutant at 16 dpw, and the RWC in WT Col-0 and the *ndr1* mutant dropped ~50% more than in the *NDR1* overexpression line, which maintained RWC of approximately 60% at 21 dpw (Figure 3-1c). In total, these data suggest that NDR1 plays an important role in the response to low water availability and drought stress. As shown in Figure AA 3-2, WT Col-0, the *ndr1* mutant, and the *ndr1/P$_{NDR1}$::T7-NDR1* complementation line did not recover from the effects of drought (25 dpw) following a 2 d post-watering (soil saturation) period. However, in the *ndr1/35S::NDR1* overexpression line, recovery was observed, including ultimate bolting, flowering, and the generation of viable seeds.

To exclude the possibility that T-DNA insertion-site effects underscore the observed response(s) in the *ndr1/35S:NDR1* overexpression line, we assembled a draft genome to determine the location and copy number of the *NDR1* transgene insertion. As

shown in Figure AA 3-3, the right border of the *35S::NDR1* transgene insertion site was located on chromosome 1 at nucleotide (nt) position 5,049,056, and the other end of the plasmid sequence (i.e., left border) was identified at nt 5,049,121. As a result of this insertion event, a 65 nt segment was eliminated. The *35S::NDR1* transgene lies between two genes: AT1G14687 and AT1G14688. The first gene (AT1G14687) encodes a zinc finger homeodomain 14 transcription factor located upstream of the insertion site (Chr1: 5,047,782-5,048,752). The second gene (AT1G14688) is located 404 bp downstream of the insertion site (Ch1: 5,049,526-5,050,983); this gene encodes an E3 ubiquitin ligase. This analysis supports our assertion that a single insertion event of the *35S::NDR1* transgene into an intergenic region is causal to the observed drought tolerance phenotype.

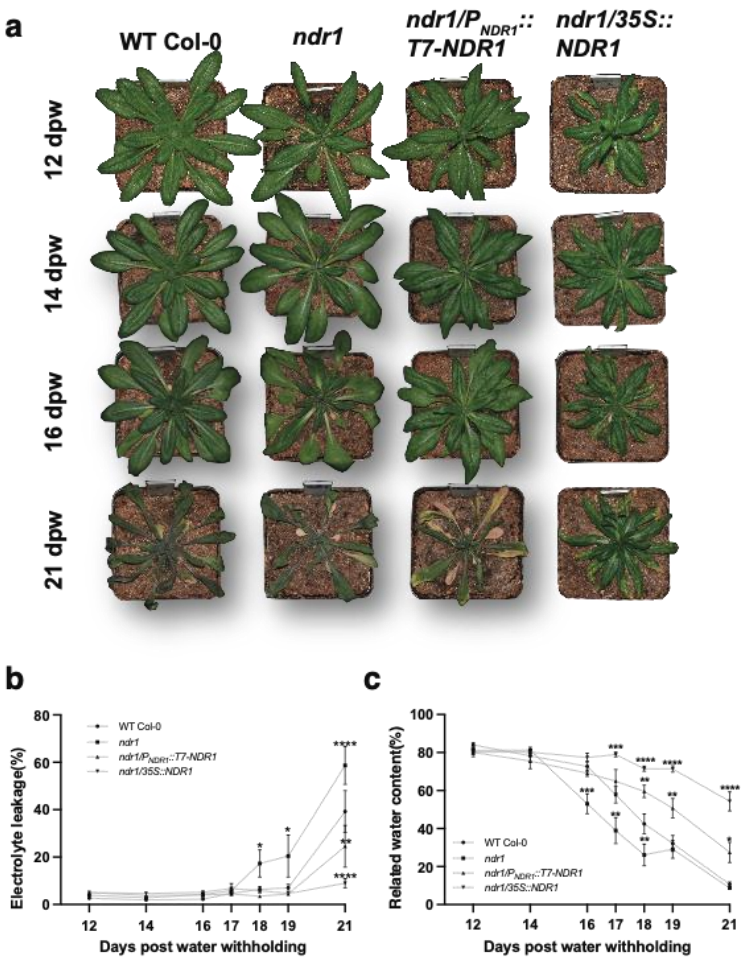

Figure 3-1: The *ndr1* mutant displays enhanced susceptibility to drought. a, Progression of drought response phenotypes from 12 to 21 days post-water withholding (i.e., days post water-withholding, dpw). Representative images from the full timecourse are shown. Additional timepoints are shown in Figure AA 3-1. b, Quantification of electrolyte leakage from drought exposed plants shown in (a). Electrolyte leakage measurements were collected at each indicated time-point, including those shown in Figure AA 3-1. Data points and error bars are the average ± SEM of collections from 5 independent leaves with three independent biological repeats (n ≥ 9). c, Quantification of relative water content (RWC) from drought exposed plants shown in a. Leaves sampled for determining water loss were collected at each indicated timepoint. Data points and error bars are the average ± SEM of collections from 5 independent leaves with three independent biological repeats (n ≥ 10). Statistical analysis was performed using a two-way ANOVA. Asterisks represent significance differences between different plant genotypes at each day. P values ≤ 0.05 were considered significant, where *P < 0.05, **P < 0.01, ***P < 0.005, and ****P < 0.001.

The drought-associated phenotype observed in the *ndr1* mutant suggests the likely involvement of the ABA signaling cascade(Fujii et al., 2007). To investigate this, and to identify possible miscues in ABA biosynthesis and signaling in the *ndr1* mutant, we first employed an ABA-based seed germination assay, previously demonstrated as a high-throughput marker for monitoring ABA sensitivity(Fujii et al., 2007). As shown, WT Col-0 and *ndr1* mutant seeds showed nearly identical germination rates over a 10-day timecourse in the absence of exogenously applied ABA. However, in the presence of 0.5 μM (Figure 3-2a), 1 μM (Figure AA 3-4a), and 2 μM ABA (Figure AA 3-4b), both WT Col-0 and the *ndr1* mutant showed a 1-3 day delay in the germination rate, similar to that previously observed for WT Col-0 (Garciarrubio et al., 1997). However, over the course of the 10d experiment, the *ndr1* mutant consistently showed statistically significant higher levels of germination in the presence of exogenous ABA, as compared to WT Col-0. These results suggest that NDR1 has a role in ABA-mediated regulation of seed germination.

To determine if a reduction in root growth is responsible for the observed drought phenotype in the *ndr1* mutant, we next evaluated root phenotypes during exogenous exposure to increasing concentrations of ABA, polyethylene glycol (PEG 6000; i.e., osmotic stress elicitor;(Joshi et al., 2011)), high salt (NaCl), and mannitol (i.e., osmotic stress elicitor; (Claeys et al., 2014)). As shown in Figure AA 3-5, root elongation in the *ndr1* mutant was similar to WT Col-0 in the presence of increasing concentrations of ABA (Figure AA 3-5a), up to 20% PEG 6000 (Figure AA 3-5b), 150 mM NaCl (Figure AA 3-5c), and 300 mM mannitol (Figure AA 3-5d) and under well-watered conditions (Figure

AA 3-5e). Based on these observations, we surmise that root length is not a primary factor underpinning the observed drought sensitivity in the *ndr1* mutant.

To define the relationship between hormone biosynthesis and stress signaling activation in the *ndr1* mutant, we next evaluated the accumulation of key phytohormones (e.g., abscisic acid, ABA; jasmonic acid, JA; salicylic acid, SA) during conditions of limited water availability. As shown in Figure 3-2b, ABA levels remained the same under well-watered condition across all genotypes, confirming that loss of *NDR1* does not abolish the biosynthesis of ABA. However, at 16 d following the initiation of water withholding, both WT Col-0 and *ndr1* mutant plants showed significant increases in the accumulation of ABA compared to well-watered plants. However, the *ndr1/P$_{NDR1}$::T7-NDR1* and the *ndr1/35S::NDR1* showed a slightly reduced, albeit significant, increase in ABA. These data suggest NDR1 suppresses ABA biosynthesis in response to water stress. As an additional point of reference, we also observed similar *NDR1* mRNA levels under both control and water withholding which indicates drought stress does not impact *NDR1* gene expression (Figure AA 3-6).

As a point of convergence in drought response and plant defense hormone signaling in the *ndr1* mutant, we asked if the accumulation of SA was affected by conditions of limited water availability. Under well-watered conditions, the endogenous levels of SA (Figure 3-2c) and the inactive SA-glycosides (SAG) (Figure 3-2d) correlated with the relative mRNA accumulation of *NDR1* in each of these lines (Figure AA 3-6). SA and SAG levels across all genotypes support our previously published data that the *ndr1* mutant is more susceptible to Pst DC3000, a response that is linked to reductions in the defense hormone SA(Knepper et al., 2011a). Additionally, these data suggest that

overexpression of *NDR1* leads to a de-repression of SA biosynthesis. Interestingly, under drought stress, both SA and SAG levels significantly increased in the *ndr1* mutant at 16 dpw (Figure 3-2 c and d). In the *NDR1* native promoter complementation line, SA levels were enhanced in comparison to WT Col-0 at 16 dpw, and the induction of SA in the complementation was modest. A similar trend was observed in the *35S::NDR1* line, with an approximate 25% increase in SA in response to water withholding (Figure 3-2 c and d).

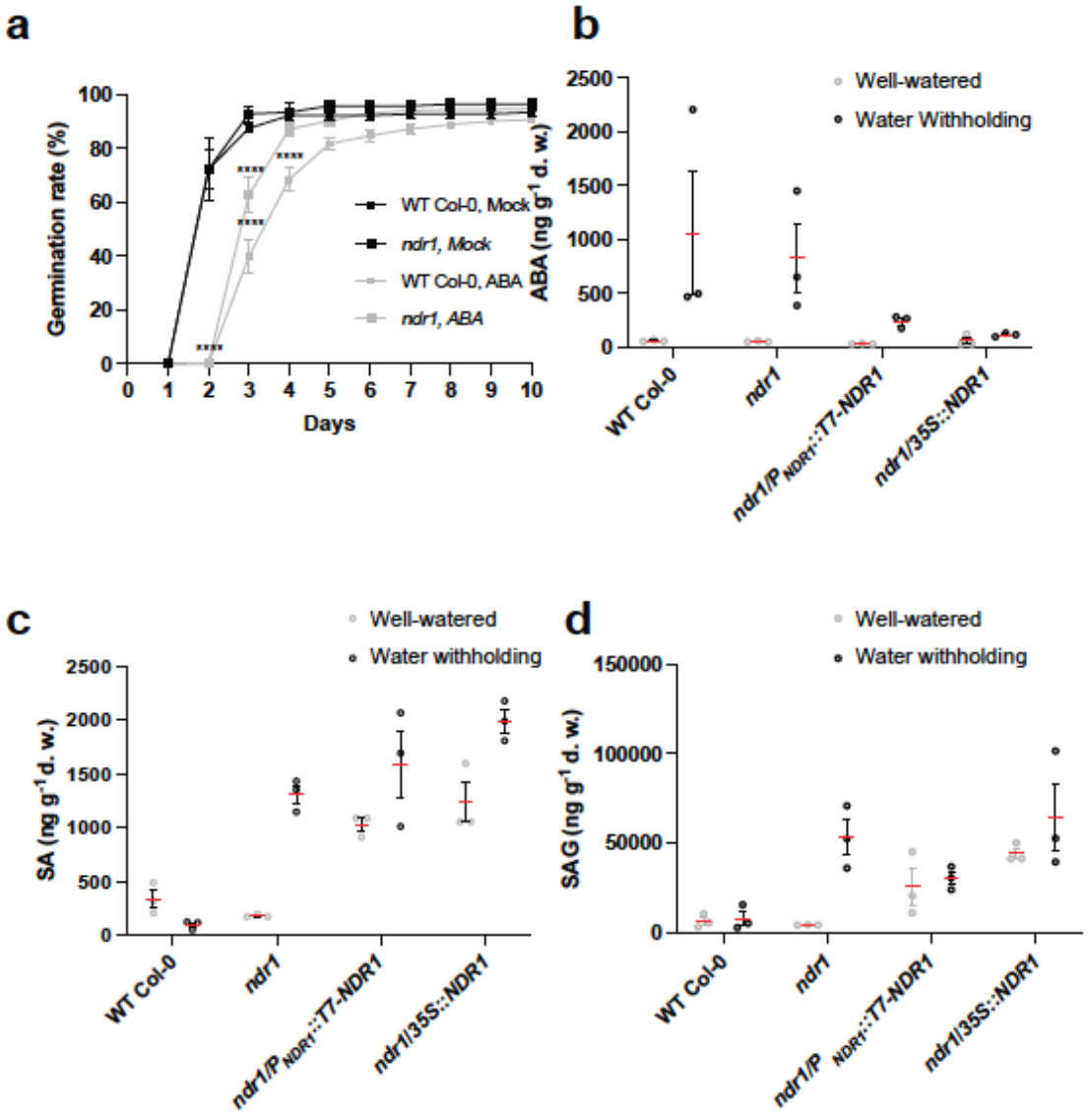

Figure 3-2: *ndr1* mutant plants are less sensitive to ABA treatment. a, Germination assay in the presence of 0.5 µM abscisic acid (ABA), b, Quantification of abscisic acid (ABA), c, salicylic acid (SA) and d, SA-glycosides (SAG) from the leaves of 4-to-5-week-old plants at 16 days post-water withholding (dpw). All data points represent samples collected from 15 individual plants. Statistical analysis was performed using a two-way ANOVA, where $P \leq 0.05$. Bars indicate standard error; red lines represent mean values.

The balance between ABA biosynthesis and catabolism is a critical mechanism that ensures rapid activation, and attenuation, of signaling in response to a battery of external environmental stressors (Umezawa et al., 2006). Based on the data presented in Figure 3-2, we hypothesized that the observed phenotype(s) in the *ndr1* mutant following water withholding are likely not due to a block in the accumulation of key defense and abiotic stress signaling hormone. However, we cannot exclude the possibility that the induction of key ABA biosynthesis genes, as well as ABA responsive marker genes, are not affected, and thus causal to the observed phenotypes. To investigate this, we evaluated the mRNA accumulation of a suite of genes associated ABA biosynthesis and metabolism. As shown in Figure AA 3-7a, mRNA accumulation of the upstream enzyme in ABA biosynthesis, *ZEAXANTHIN EPOXIDASE* (*ABA1*), is elevated to a similar level during drought conditions compared to well-watered controls in WT Col-0, the *ndr1* mutant, the *ndr1/P$_{NDR1}$::T7-NDR1*, and the *ndr1/35S::NDR1* lines. We also examined additional downstream enzymes that are important for ABA biogenesis. For example, a significant increase in *9-CIS-EPOXYCAROTENOID DIOXYGENASE* (*NCED3*) is detected in WT Col-0 and the *ndr1* mutant under drought stress (Figure AA 3-7b). The accumulation of *XANTHOXIN DEHYDROGENASE* (*ABA2*), which encodes a key enzyme in the early conversion steps of xanthoxin to ABA aldehyde, remained to the similar levels and was significantly reduced under water withholding in the *ndr1* mutant (Figure AA 3-7c). The expression level of *ABA2* implies a positive correlation with *NDR1* mRNA levels in all tested lines (Figure AA 3- 6). Interestingly, we found the mRNA expression of ABA biosynthesis aldehyde oxidase (*AAO3*) is significantly induced in WT Col-0 and the *ndr1* mutant but not the *NDR1* over-expressor lines (Figure AA 3-7d).

Moreover, similar mRNA expression patterns were also observed when we measured. *RD29B* and *EARLY RESPONSIVE TO DEHYDRATION* (*ERD4*; (Kiyosue et al., 1994)), both of which are responsive to desiccation and dehydration (Figure AA 3-7e, f).

As shown in Figure 3-1, the *ndr1* mutant exhibited a striking decline in RWC in response to water withholding at 16 dpw. Coincident with this, we detected a slight, yet significant, decrease in the accumulation of *ABA1* mRNA in the *ndr1* mutant compared to WT Col-0. Conversely, by Day 21, and concomitant with the onset of severe drought-associated phenotypes (e.g., Figure 3-1), we observed a significant increase in the mRNA accumulation of *ABA1* in the *ndr1* mutant. At the same time, mRNA accumulation of *ABA2* shifted from being significantly downregulated at 16 dpw in the *ndr1* mutant to significantly upregulated at 21 dpw, as compared to WT Col-0 ($P < 0.05$). Taken together, and in agreement with our quantitative evaluation of RWC over the timecourse of water withholding (Figure 3-1c), we propose that 12 and 16 dpw represents a tipping point for the *ndr1* mutant and WT Col-0 plants, respectively. This hypothesis is consistent with our analysis of ABA biosynthesis, which was first observed to increase in the *ndr1* mutant, and then by 4 days post-water withholding, was similar in both WT Col-0 and *ndr1* plants (Figure 3-2).

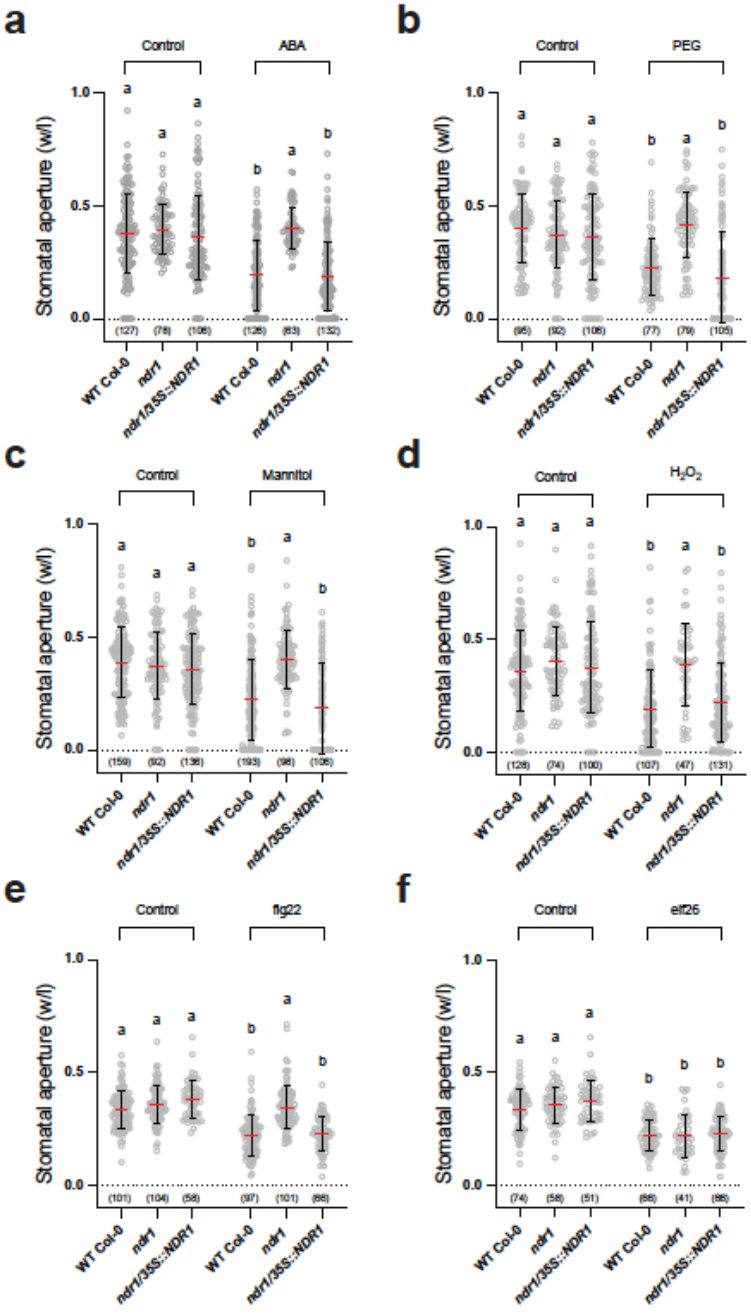

Figure 3-3: Stomatal closure in the *ndr1* mutant is impaired, with differential responses to ABA and PAMPs. Four-to-five-week old plants were subjected to either abiotic or biotic stress elicitation for one hour before imaging and guard cell aperture evaluation. As shown, *ndr1* guard cells do not respond to a, 10 µM ABA, b, 20% PEG, c, 300 mM mannitol, d, H_2O_2, e, 100 nM flg22, f, 100 nM elf26. The *ndr1* guard cells do respond to d, elf26. Statistical analysis was performed using a two way ANOVA, where $P \leq 0.05$. Bars indicate standard error, Red lines represent mean values.

Drought response signaling in plants is regulated by both ABA-dependent and ABA-independent pathways, both of which share converging regulatory mechanisms (Seo & Mas, 2015). For example, many drought-associated genes are regulated by rhythmic changes in the circadian clock, a genetic network that drives plant develop, response to pathogens, and control of hormone biosynthesis (Lu et al., 2017; Seo & Mas, 2015). Stomata are critical components of water deficit signaling, functioning in large part by regulating the diurnal rate of transpiration. As a point of entry for bacterial pathogens, stomata aperture is directly correlated with infection and pathogen colonization (Melotto et al., 2006). To determine if a loss of *NDR1* results in alterations in the clock, we evaluated the diurnal gating of stomatal guard cells as a tractable, predictable, output of clock function, and moreover, diurnal hormone oscillations in hormone accumulation. As shown in Figure AA 3-8a, we observed that the *ndr1* mutant had increased stomatal aperture, compared to WT Col-0, at Zeitgeber time 5 h and 7 h; as indicated, this period corresponds to daylight hours. During Zeitgeber time 15 h – 23 h, the *ndr1* mutant displayed an increase in guard cell aperture (i.e., increased opening) when compared to WT Col-0. As predicted, overexpression of *NDR1* resulted in enhanced guard cell closure over the duration of the 24 h diurnal cycle. In agreement with this, the mRNA expression of key circadian marker genes was also affected in the *ndr1* mutant at timepoints coincident with the hours preceding daylight. Specifically, accumulation of TOC1 (At5G61380) mRNA was significantly increased in the *ndr1* mutant from at 11 and 13 h Zeitgeber time compared to WT Col-0 and the *NDR1* overexpression line (Figure AA 3-8b). In all three genotypes, the periodicity of TOC1 mRNA expression is similar. Further,

the accumulation profiles of ABA and SA over the 24 h diurnal cycle were the same across all 3 genotypes (Figure AA 3-9).

Based on our observations detailed above, we next asked if NDR1 is required for the convergence of signaling between pathogen perception and stomata response. Leaf punches from WT Col-0 and the *ndr1* mutant were treated with 10 μM ABA. As shown in Figure 3-3a, at 1 h post treatment, WT Col-0 stomata exhibited a significant ($P < 0.0001$) reduction in aperture width following ABA treatment, while *ndr1* did not show a change in aperture as compared to mock-treated controls. Similar to the absence of stomatal closure in response to ABA treatment, we also observed that exogenous application of polyethylene glycol (PEG) (Figure 3-3b), mannitol (Figure 3-3c), and hydrogen peroxide (H_2O_2) (Figure 3-3d) did not induce stomatal closure in the *ndr1* mutant.

Immune signaling downstream of flg22 perception is compromised in the *ndr1* mutant (Knepper et al., 2011a, 2011b). To investigate the implication of reduced flg22 signaling and/or general PAMP perception in the *ndr1* mutant, we evaluated stomatal aperture response following flg22 treatment. Whole leaf punches from WT Col-0 exposed to flg22 showed rapid guard cell closure (ca. 1 h; Figure 3-3e), whereas guard cells from the *ndr1* mutant remained open following flg22 treatment. To determine if additional PAMP treatment and perception responses were affected in the *ndr1* mutant, we also evaluated guard cell response following elf26 elicitation (Desikan et al., 2008). Surprisingly, in response to elf26 treatment, stomatal closure in the *ndr1* mutant was similar to that observed in WT Col-0 (Figure 3-3f).

Leaf water content is a major contributor to the restriction of *Pst* DC3000 growth during gene-for-gene mediated resistance in *Arabidopsis* due to the coupling of restricted

vascular activity and water loss through the stomata at the site of infection(Freeman & Beattie, 2009). Indeed, previous work showed that *Pst* DC3000 encounters significantly lower water potentials during incompatible interactions than during compatible interactions (Wright & Beattie, 2004). Based on these observations, we hypothesized that the *ndr1* mutant would be more susceptible to bacterial infection under conditions of drought stress. To test this, we evaluated the impact of drought on the development of disease and enhanced bacterial growth following pathogen infection of 4-week-old WT Col-0 and *ndr1* mutant plants subjected to water withholding. As expected, the difference in *Pst* DC3000 growth in *ndr1* mutant plants under well-watered conditions was approximately 1.5 log CFU cm^{-1} higher than in WT Col-0 (Figure AA 3-10a). Under conditions of water-withholding, a similar differential (ca. 1.5 log CFU cm^{-1}) in *Pst* DC3000 growth was observed in the *ndr1* mutant as compared to WT Col-0. In agreement with the role of NDR1 as a positive regulator of plant defense, the *in planta* growth of *Pst* DC3000 in the *NDR1* native promoter-driven and overexpression lines (Figure AA 3-6) were similar to that observed in WT Col-0. Interestingly, under drought conditions, plants challenged with the avirulent pathogen, *Pst* DC3000-AvrRpt2, showed a response consistent with the hypothesis presented by Wright and Beatie (Wright & Beattie, 2004). Indeed, we observed resistance to Pst DC3000-AvrRpt2 in WT Col-0 and susceptibility in the *ndr1* mutant under well-watered conditions, as previously described (Century et al., 1997; Coppinger et al., 2004; Knepper et al., 2011a) (Figure AA 3-10b). However, under conditions of water-withholding, growth of *Pst* DC3000-AvrRpt2 in the *ndr1* mutant was ~1 log greater than under well-watered conditions. These data are in agreement with previous studies demonstrating that drought stress impacts bacterial growth during

incompatible interactions but not during compatible interactions (Freeman & Beattie, 2009).

Based on the data presented above, we next asked if the requirement for NDR1 is associated with transcriptional activities of immunity and/or ABA-associated signaling processes. To test this, we conducted a comprehensive mRNA-seq analysis to evaluate gene expression changes in WT Col-0, the *ndr1* mutant, and the *NDR1* overexpression line (*ndr1/35S::NDR1*) over a time period of 0-21 day post water-withholding. In addition, control samples (i.e., well-watered) were collected from all genotypes and at all time points. In total, these datasets were comprised of 129 RNA-seq libraries. To explore gene expression difference between treatment groups and WT Col-0 group, samples were grouped using principle component analysis (PCA), the output of which revealed that gene expression in different groups were primarily separated from each other by time, with distinct grouping at Day 0,12,14,16,17,18, 19, and 21 under water withholding treatments, and at Day 0, 12, 16, and 21 under well-watered treatment (Figure AA 3-11).

To establish a baseline of *NDR1*-dependent transcriptional responses (i.e., differentially expressed genes; DEG (P value < 0.01, $|\log_2$-fold change$| \geq 1$)) across genotypes, we first evaluated gene expression profiles from each of the 3 genotypes from 0-21 day post water-withholding. As a control for stress-responsive DEGs, we also included parallel samples (i.e., each genotype, each timepoint) from well-watered plants. As shown in Figure 3-4, we first identified mRNAs that were differentially expressed (compared to well-watered plants) in each group over the time course of water-withholding in *ndr1* and *ndr1/35S::NDR1* compared to WT Col-0 plants, we found more genes are differentially regulated in the *NDR1* overexpression line compared to the *ndr1*

mutant (Figure 3-4a; Data Set AA 3-1 and 3-2). We also identified the same trend as above in the well-watered treatment (Figure 3-4b, Data Set AA 3-3 and 3-4). Next, we examined the pattern of DEGs among each time points to evaluate the temporal pattern of expression of key genes as a function of genotype and treatment. As shown in Figure 3-4b, DEG patterns in the *ndr1/35S::NDR1* overexpression line vs WT Col-0 were similar across the entire water-withholding period, while the pattern varied in the combination of *ndr1* mutant vs WT Col-0. Further analysis of the day-to-day comparison (with Day 0 as a baseline) of DEGs in WT Col-0, the *ndr1* mutant and *ndr1/35S::NDR1* overexpression line revealed more DEGs in WT Col-0 than in the *ndr1* mutant and the *ndr1/35S::NDR1* overexpression line under both water-withholding and well-watered conditions (Figure AA 3-12a, b). The composition of DEGs among the different plant genotypes was found to be most striking at 12 dpw. Indeed, as the pattern of DEGs in *ndr1* mutant was different from WT Col-0, the *ndr1/35S::NDR1* overexpression line was able to restore a complementary set of genes, as compared to WT Col-0 (Figure AA 3-12c, Data Set AA 3-5, 3-6, 3-7, 3-8, 3-9, and 3-10).

To further investigate this, we next identified the DEGs that were constantly up- or downregulated from 12-21 dpw among WT Col-0, *ndr1*, and *ndr1/35S::NDR1* genotypes. From this, we identified 153 genes that are constantly upregulated in the *ndr1* mutant, and 1319 that are constantly upregulated in the 35S-over-expressor line. Additionally, 62 transcripts were identified as constantly downregulated in the *ndr1* mutant, and 831 that are constantly downregulated in the 35S over-expressor line (Data Set AA 3-11, 3-12, 3-13, and 3-14). Among these, 96 transcription factors (TF) belonging to the WRKY, ERF, AP2, and MYB families were identified as either constantly upregulated or downregulated

under water withholding and well-watered conditions (Figure AA 3-13; Data Set AA 3-15). Not surprisingly, and in further support of our hypothesis, we found that WRKY47, a drought-responsive gene, was also among the list of the above DEGs. Additionally, the mRNA expression levels of *WRKY40* and *WRKY70* were also altered significantly under water withholding condition (Figure AA 3-13).

Gene ontology (GO) enrichment analysis resulted in the identification of the above shared transcripts within each group that were continuously up- and/or downregulated in *ndr1* and *ndr1/35S::NDR1* compared to WT Col-0 plants under d water withholding treatment. From this, further analysis identified 995 DEGs that were constantly upregulated in *ndr1/35S::NDR1* line that fell into categories of GO terms that have been reported to regulate reactions drought stress and plant immunity including genes response to hormone (i.e., ABA, JA, and SA), stress, intrinsic components of the plasma membrane, and immune system process. Among these, we identified 324 downregulated genes in GO term groups that response to osmotic and drought stresses in *ndr1/35S::NDR1* under water withholding conditions (Figure 3-4c; Data Set AA 3-16 and 3-17). In support of our assertion that overexpression of *NDR1* is responsible for these

changes in gene expression profiles, we did not observe similar annotated pathways among the 145 up- and 8 downregulated DEGs from the *ndr1* mutant.

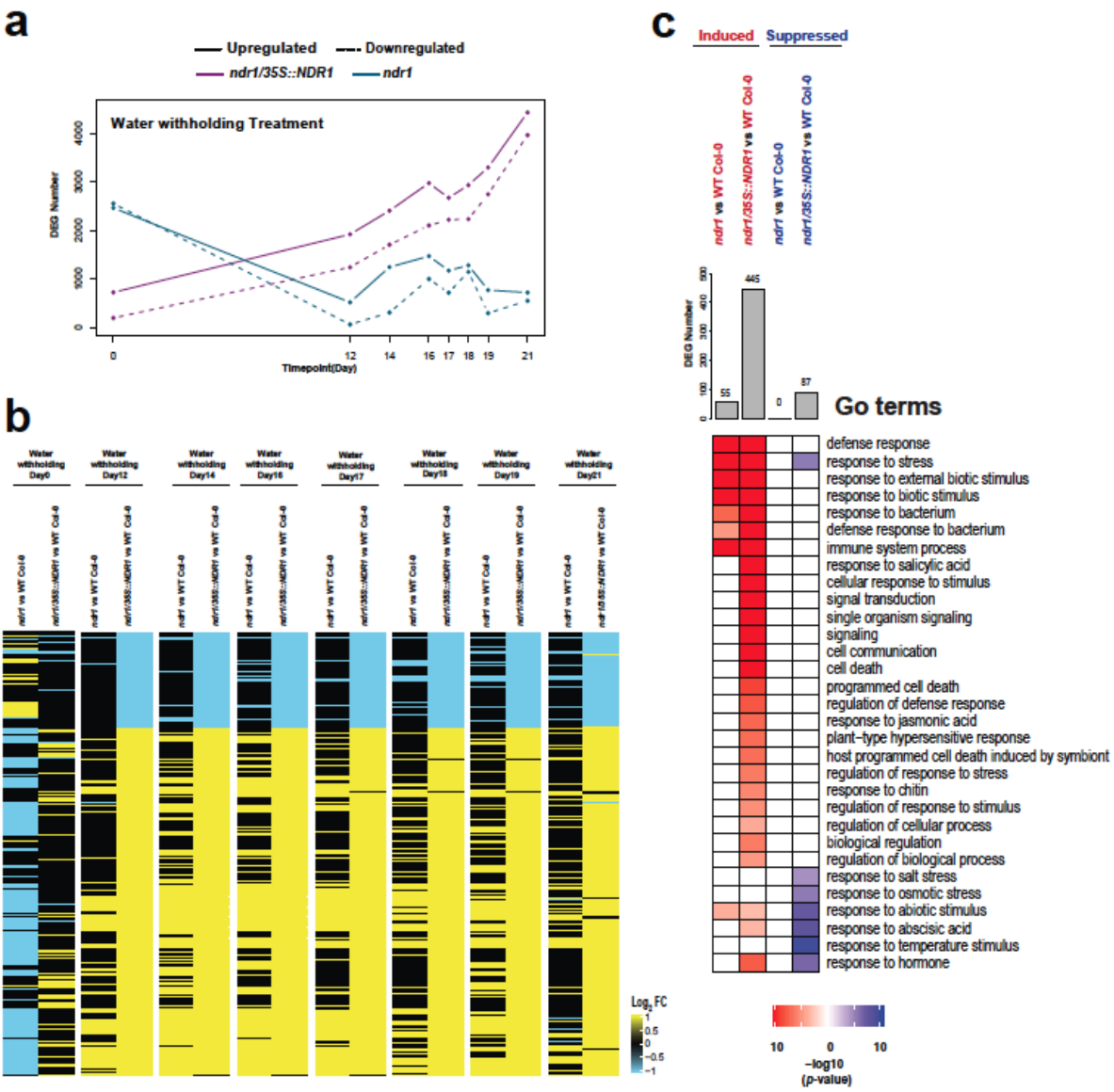

Figure 3-4: Patterns of differential gene expression and the biological process pathways that are regulated by these DEGs during drought stress. a, Expression profile of DEGs under water withholding over the 12-21 day timecourse. b, Expression patterns of genes under water withholding condition. c, GO terms of differential DEGs between *ndr1* vs WT Col-0 and *ndr1/35S::NDR1* vs WT Col-0.

As a final step in our transcriptome analysis, and to associate mRNA expression changes with the known biological activities of the NDR1-dependent immune and abiotic process, we next evaluated the biological processes within each group from 12-21 dpw compared to 0 dpw. Interestingly, we found that changes gene expression occur earlier in *ndr1/35S::NDR1* than in *ndr1* mutant in response to water withholding treatment (Figure 3-4b; Data Set AA 3-18 and 3-19). This is exciting, as it illustrates that the strongest influence on differential gene expression of water-response processes is linked to the overexpression of *NDR1*. Taken together, these data provide compelling evidence that overexpression *NDR1* plays an important role in the regulation of drought stress-responsive genes under conditions of limited water availability.

Table 3-1: NDR1 interacting proteins identified by immunoprecipitation tandem mass spectrometry. Peptide counts are shown for each of 2 independent biological repeats (i.e., BR-1, BR-2). Protein identification required *P* < 0.05 (MOWSE algorithm), minimum of # peptides. None of the proteins listed were identified in the negative control (i.e., WT Col-0) IP-MS/MS experiment.

Gene Name	Gene ID	GO Term	Peptide Count	
			BR-1	*BR-2*
NDR1 (control)	AT3G20600	Non-Race Disease Resistance-1	4	7
AHA1	AT2G18960	H(+)-ATPase-1	16	15
ERD4	AT1G30360	Early-responsive to dehydration stress protein	12	13
PIP2:7	AT4G35100	Plasma membrane intrinsic protein	3	7
HIR2	AT3G01290	PHB domain-containing membrane-associated protein family	8	3
Remorin1	AT3G61260	Remorin family protein	-	4
PIP1:2	AT2G45960	Plasma membrane intrinsic protein	4	4
PIP2:6	AT2G39010	Plasma membrane intrinsic protein	5	3
NHL3	AT5G06320	NDR1/HIN-like 3	2	3
HIR1	AT1G69840	PHB domain-containing membrane-associated protein family	-	2
PIP2:2	AT2G37170	Plasma membrane intrinsic protein	-	2
PIP2:1	AT3G53420	Plasma membrane intrinsic protein	4	2
PEN3	AT1G59870	Plant PDR ABC-type transporter family protein	2	2
PIP1:1	AT3G61430	Plasma membrane intrinsic protein	3	2

Lastly to investigate the relationship between NDR1 function and response to drought as a requirement of ABA signaling, we explored the common and unique DEG in

ndr1 and *ndr1/35S::NDR1* under water withholding and well-watered treatment, respectively (Figure AA 3-14a, b). In addition, we cataloged all GO terms associated with our candidate list of DEGs in the above. We found that unique DEGs in *ndr1/35S::NDR1* under water with-holding treatment were significantly enriched in pathways responding to water and SA. However, DEGs in *ndr1* mutant under water withholding treatment are not enriched in the above biological processes (Figure AA 3-14c; Data Set AA 3-20). Taken together, these data support the hypothesis that immunity and drought process signaling converge as a function of *NDR1* overexpression. Moreover, RNA-seq analysis provides evidence that the expression of *AHA2* is upregulated in *ndr1/35S::NDR1* plant compared to WT Col-0 across all time points (Figure 3-5c). These results indicates that NDR1 functions in signaling processes required for guard cell gating, a mechanism known to control both biotic and abiotic stress signaling responses.

To gain insights into global gene coexpression profiles under water with-holding environments, we performed coexpression analysis in each of the genotypes as a function of water status. In short, this allowed us to categorize genes into coexpression modules and identify the expression levels of module eigengenes (MEs). Using this approach, we identified 15 coexpression modules among WT Col-0, *ndr1*, and the *ndr1/35S::NDR1* overexpression line at all timepoints (Data Set AA 3-21-3-23). As shown in Figure 3-5a, Module 2 contains genes that are associated with processes associated with plant response to water deficit. In this, we found *LEA14* (At1g01470) and *ERD10* (At1g20450), with a reduced eigengene value in the *ndr1/35S::NDR1* overexpression line than in WT Col-0 or the *ndr1* mutant. Interestingly, coexpression profiles from Module 3 also highlight a pattern revealing higher expression in defense-related genes in the

ndr1/35S::NDR1 line, including known stress regulators such as *WRKY70* (At3g45640), *PAD4* (At3g52430), *NPR4* (At4g19660), and *FLS2* (At5g46330). In addition, RNA-seq analysis provided evidence that the expression of *AHA2* was upregulated in *ndr1/35S::NDR1* plant compared to WT Col-0 across all time points (Figure 3-5b). These results indicate that NDR1 likely functions in signaling processes required for guard cell gating, a mechanism known to control both biotic and abiotic stress signaling responses.

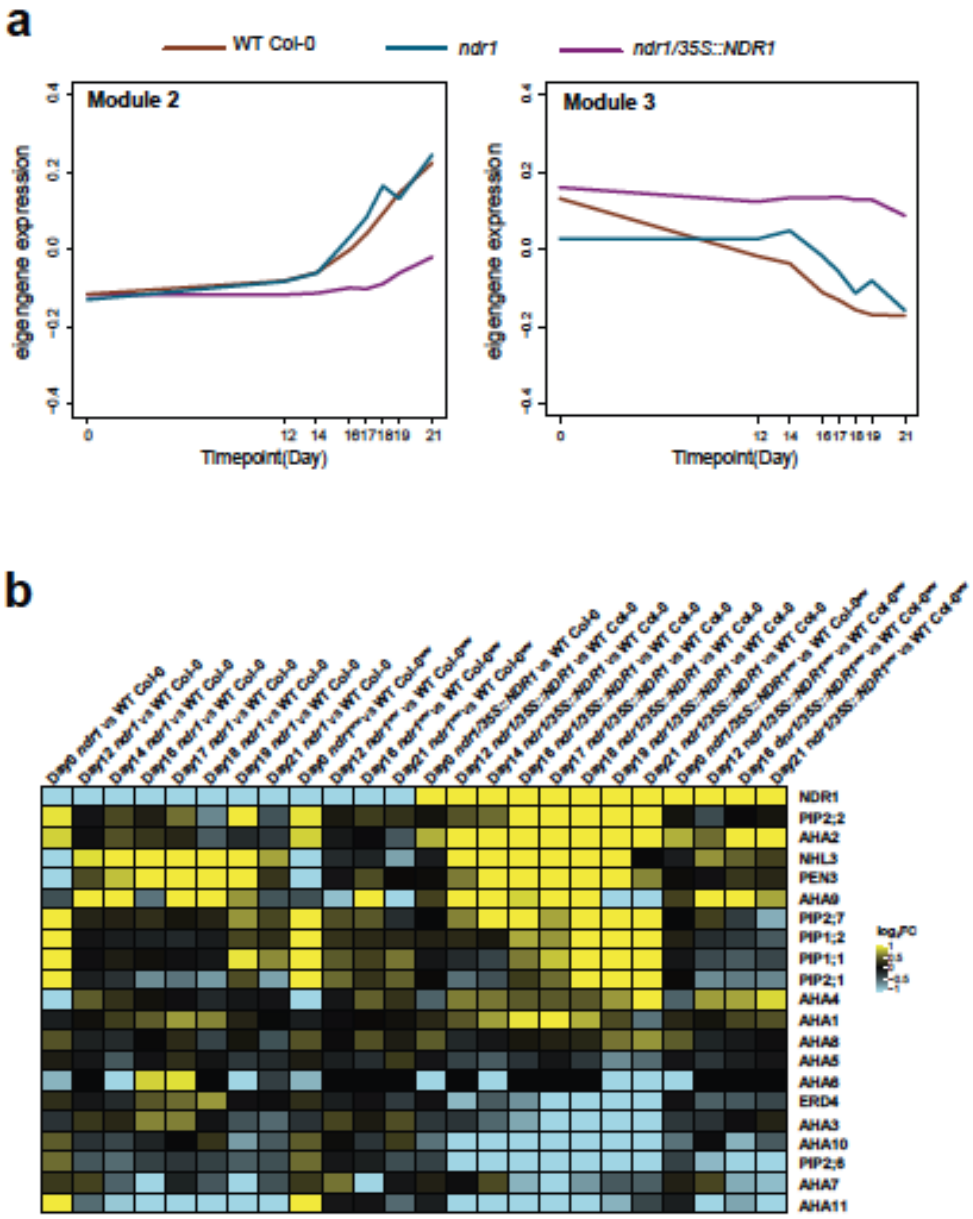

Figure 3-5: Gene co-expression modules and expression patterns of drought-associated genes during water-withholding. a, Eigengene expression modules of WT Col-0, *ndr1*, and *ndr1/35S::NDR1*. b, mRNA accumulation of the mRNAs whose products were identified as NDR1-interacting candidates.

As shown in Figure 3-5b, we observed an increased in the mRNA accumulation of several membrane-associated H^+-ATPases, including *AHA2*, *AHA9*, and *AHA4*. Of

particular note, and in agreement with the MS/MS data, is the identification of an interaction between NDR1 and membrane H^+-ATPase, AHA1 (At2g19860), which in addition to its role in signaling during abiotic stress response, is required for immune signaling following bacterial pathogen infection (Liu et al., 2009). To confirm the specificity of the NDR1-AHA1 interaction, we first performed in planta co-immunoprecipitation (co-IP) assays using a transient *Agrobacterium tumefaciens-Nicotiana benthamiana* heterologous expression system. To do this, we evaluated the pairwise interactions between Flag-NDR1 and HA-epitope-tagged constructs of AHA1, AHA2, and AHA5. As shown in Figure 3-5b, we identified an interaction between Flag-NDR1 and all 3 tested AHA proteins, which supports the NDR1-AHA1 IP-MS/MS interaction data shown in Table 3-1. To further evaluate the interaction between NDR1 and each of the identified AHAs, as well as to provide insight into the cellular address of these interactions, we next used bimolecular fluorescence complementation (BiFC) assays, again confirming the interaction between NDR1 and all tested AHA proteins (Figure AA 3-15).

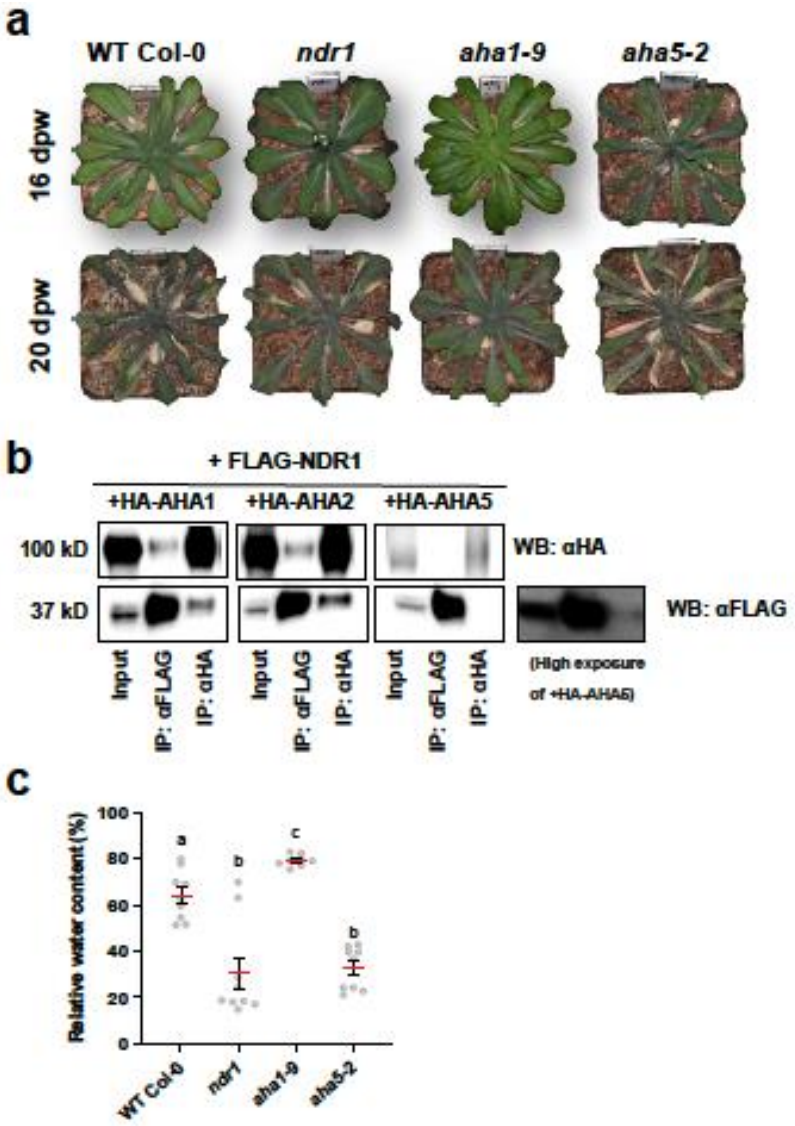

Figure 3-6: The aha1-9 mutant shows increased tolerance to water-withholding conditions compared to WT Col-0 and the *ndr1* and *aha5-2* mutants. a, Representative images of *Arabidopsis* plants undergoing drought stress over a timecourse of 14-to-20-days post water withholding. b, Co-immunoprecipitation of Flag-NDR1 with HA-AHA1, HA-AHA2, and HA-AHA5 c, Relative water content of plants during drought stress at 15 dpw (days post-water-withholding). Statistical analysis was performed using a one-way ANOVA. P values í 0.05 were considered significant. Bars indicate standard error, red lines represent mean values.

Data presented herein support the hypothesis that NDR1 functions in stomatal immunity through its interaction with plasma membrane-associated H^+-ATPases. To

further test this, we evaluated available SALK T-DNA mutant alleles for response to drought to determine if *AHA1* and *AHA5* phenocopy the *ndr1* mutant (Figure AA 3-16). As shown in Figure 3-6a, we observed that the *aha1-9* mutant was in fact more tolerant to the effects of water withholding, as compared to WT Col-0, whereas *aha5-2* was more susceptible – similar to the *ndr1* mutant. These observations were further supported by a quantitative analysis of the RWC over a short timecourse of drought. Indeed, and in agreement with the NDR1-AHA5 protein-protein interaction analysis (Figure 3-5b), we observed that the *aha5-2* mutant had similar levels of RWC as the *ndr1* mutant over the timecourse of water withholding, while the *aha1-9* mutant had RWC levels similar to the *NDR1* overexpression line (Figure 3-6b).

As shown above, NDR1 is essential for stomatal closure in response to a variety of abiotic stressors, including the differential response to at least 2 well-characterized PAMPs. Based on the identification of a physical interaction between NDR1 and AHAs, we next asked if the loss of function of H^+ ATPases phenocopies *ndr1* stomatal response following PAMP perception. As shown in Figure 3-7, we observed differential responses in the *aha1-9* and *aha5-2* mutant lines. Stomata from WT Col-0 and the *aha1-9* mutant rapidly closed following treatment with ABA (Figure 3-7a). However, following ABA treatment, stomata from the *ndr1* and *aha5-2* mutant did not close. A similar result was observed when plants were treated with exogenous flg22 peptide (Figure 3-7b). Conversely, plants treated with the elf26 peptide showed rapid guard cell closure (ca. <1 h; Figure 3-7c). These data are consistent with the observation that *ndr1* mutant stomata

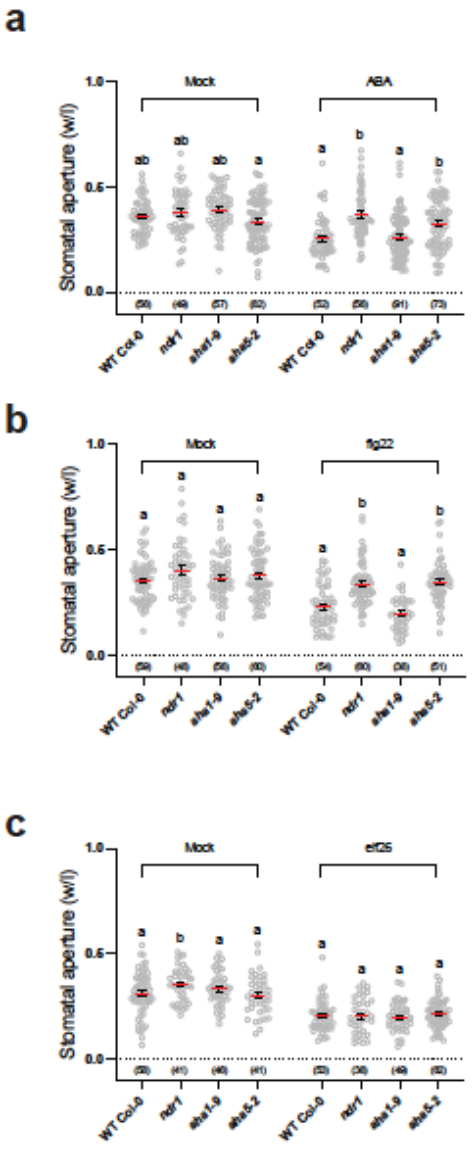

Figure 3-7: The aha5-2 mutant displays impaired stomatal closure in response to both abiotic and biotic stress elicitors. Knockout mutants aha5-2 phenocopies ndr1 in stomatal response to a, ABA, b, flg22, and c, elf26 treatment. Leaf discs were subjected to stress for 1 hour prior to imagining. Each data point represents 3 biological replicates; > 50 stomata were imaged for each replicate. Statistical analysis was performed using a one-way ANOVA. P values í 0.05 were considered significant. Bars indicate standard error, red lines represent mean values.

also do not respond to treatment with ABA or flg22 yet does respond to elf26 (Figure 3-3e, f).

As a mechanism controlling stomatal gating in response to phytopathogen infection, it was recently shown that RIN4 acts as a positive regulator of PM H^+-ATPase activity (Liu et al., 2009). Based on our observation that exogenous application of ABA and flg22 are unable to induce stomata closure in the *ndr1* mutant, as well as NDR1's association with RIN4 (Day et al., 2006), we hypothesized that the PM H^+-ATPase activity in *ndr1* stomata may be affected. To test this, we next monitored PM H^+-ATPase activity in WT Col-0, the *ndr1* mutant, and *ndr1/35S::NDR1* overexpression line, to evaluate NDR1's role in regulating H^+-ATPase activity. In support of our hypothesis, we observed an enhancement in the activity of the H^+-ATPase in the *ndr1* mutant (Figure 3-8a), suggesting a potential role for NDR1 in negatively regulating ATPase activity. Our result showed increasements of H^+-ATPase activity in *aha1-9* and *aha5-2* mutants (Figure 3-8b). Taken together with the genetic and protein interaction data shown above, we assert that NDR1 functions in guard cell gating through its association with stomatal-localized H+-ATPases.

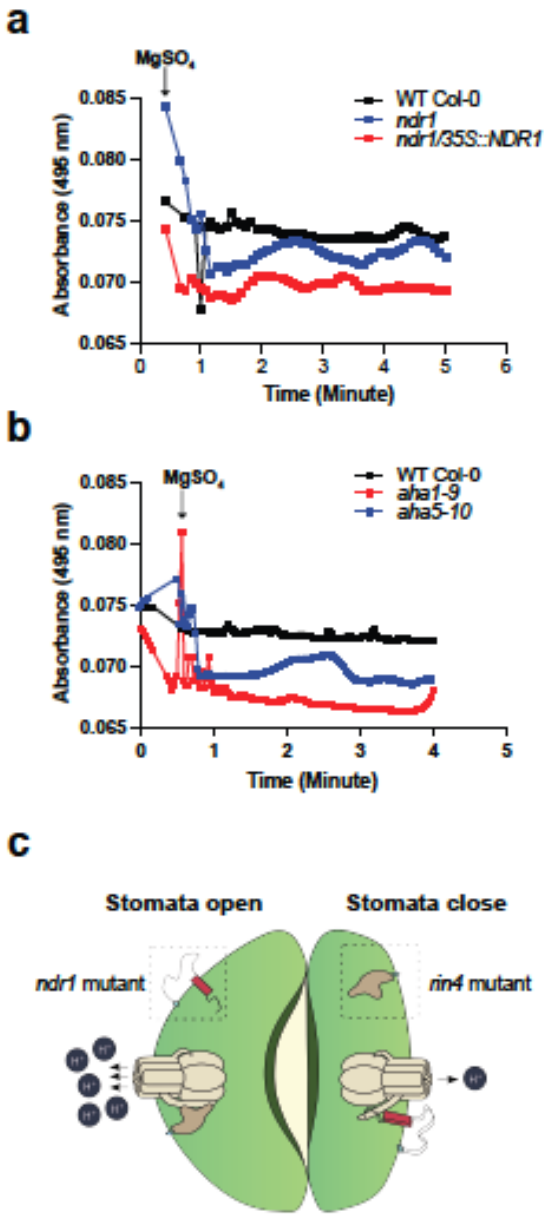

Figure 3-8: H+-ATPase activity is enhanced in the *ndr1* and the aha1-9 H+-ATPase mutant. a, WT Col-0 *ndr1* mutant and *35S::NDR1*, b, WT Col-0, *aha1-9*, and *aha5-2* mutant leaf enriched plasma membrane fractions were purified using an aqueous polymer two phase method, and the assay was conducted on inside-out vesicles. Plasma membrane H+-ATPase activity was measured using an acridine orange efflux assay. Absorbance was measured at 495 nm. Experiments were repeated 2-3 times with independent plasma membrane enriched fractions for each replicate. c, Model illustrating the link between NDR1, RIN4, and the plasma membrane H+- ATPase. NDR1 and RIN4 have opposing functions in both biotic and abiotic stress signaling.

Discussion

NADR is an essential component of the *Arabidopsis* defense signaling network, required for resistance to the bacterial phytopathogen *Pseudomonas syringae pv*. tomato DC3000 (Century et al., 1995; Century et al., 1997; Day et al., 2006; Knepper et al., 2011a, 2011b). As a mediator of the plant PM-CW continuum, NDR1 has been shown to play a role in regulating fluid movement across the membrane in response to pathogen infection, revealing that an increase in electrolyte leakage in the *ndr1* mutant following pathogen inoculation, as compared to WT Col-0 (Knepper et al., 2011a, 2011b). As nutrient acquisition is the primary factor driving pathogen virulence, the observation of a reduction in the PM-CW continuum, resulting in a loss in cellular integrity and the ability to regulate the retention of nutrients, offers a plausible mechanism to explain the enhanced bacterial growth phenotype in the *ndr1* mutant. In total, these findings suggest that NDR1 is required for the maintenance of focal adhesion and the transduction of extracellular stimuli to intracellular signaling.

In the current study, we build upon our previous work to define the mechanism(s) through which NDR1 is required for abiotic and biotic stress signaling. Based on NDR1's role in PM-CW adhesion, we hypothesized that NDR1 may also play a role in abiotic stress signaling, a mechanism which Singh et al. (Singh et al., 2005) previously described for late embryogenesis abundant-14 (LEA14), a structural homolog of NDR1 and a known regulator of desiccation tolerance (Knepper et al., 2011a). The hypothesis that NDR1 is likely associated with processes that converge on plant immunity and abiotic stress (i.e., drought) signaling is based on the described significance of water content during pathogen infections. Indeed, previous studies have shown that pathogens encounter

inhibitory levels of osmotic potential once inside their host, and as part of their virulence response, deploy specific effector molecules which enhance colonization as a function of nutrient and water availability(Aung et al., 2018; Goel et al., 2008). Here, we observed that under conditions of drought, RWC in the *ndr1* mutant decreased much faster than in WT Col-0 plants following water withholding. This decrease in RWC correlated with our previous observation of enhanced electrolyte leakage in the *ndr1* mutant (Knepper et al., 2011a), supporting a role for NDR1 similar to that of the LEA family of proteins as functioning in processes that link biotic and abiotic stress activation through the perception of cell desiccation.

As a common denominator in the avoidance of drought stress and pathogen infection, the regulation of guard cell dynamics (i.e., stomata aperture) through ABA signaling represents a critical process in a plant's response to environmental stress. At a regulatory level, previous data demonstrate that ABA metabolism and catabolism are modulated during host-pathogen interactions, and moreover, that increased levels of ABA drive host susceptibility to pathogen infection (Mine et al., 2017). As a function of the role of NDR1 in this mechanism, data presented herein correlate a decline in RWC in the *ndr1* mutant with an increase in gene expression of the key regulators of ABA metabolism. As a regulatory point in describing the disconnect between increases in ABA biosynthesis (both mRNA and metabolite levels), our data also show that increases in ABA biosynthesis in the *ndr1* mutant did not prevent the *ndr1* mutant from rapidly declining into severe drought stress conditions, and ultimately, death.

We propose that the *ndr1* mutant is unable to perceive changes in ABA levels as an indicator of stress, and as a result, does not activate the appropriate signaling

responses. In the case of abiotic stress, this translates into increased sensitivity to drought. In the case of biotic stress, this results in increased fluid loss during pathogen infection, leading to an enhanced disease susceptibility phenotype. NDR1 has long been associated with disease resistance signaling, including a key role in ETI (Century et al., 1995; Century et al., 1997; Coppinger et al., 2004; Day et al., 2006) and PTI (Knepper et al., 2011a). We next investigated the link between drought, PTI, and NDR1 in stomatal-based signaling. To this end, we examined the role(s) of NDR1 in the regulation of stomata aperture, addressing both the abiotic (i.e., ABA) and biotic (i.e., PAMP) potential for NDR1-mediated signaling. We observed an abrogation in stomata function in the *ndr1* mutant through both an ABA-dependent manner, as well as through PTI-based immune signaling. These data are in agreement with a loss in sensitivity to ABA, as well as offer a possible mechanism supporting the EDS phenotype observed in the *ndr1* mutant (Century et al., 1995; Century et al., 1997; Knepper et al., 2011a).

The activation of PM H^+-ATPase(s) leads to hyperpolarization of the PM, and subsequent induction of inward ion channels, resulting in stomatal opening (Figure 3-8c). In contrast, inhibition of the PM H^+-ATPase and subsequent simulation of anion channel activation initiates membrane depolarization, resulting in the activation of outward rectifying K^+ channels and stomatal closure (Sirichandra, Gu, et al., 2009; Sirichandra, Wasilewska, et al., 2009), less turgor pressure, and the abrogation of stomata closure. This data is further in agreement with recent work by Liu et al. (Liu et al., 2009) which demonstrated that RIN4 modulates the activity of the PM H^+-ATPase to regulate stomatal aperture during pathogen infection.

Based on the observation that a *RIN4* overexpression line exhibits enhanced disease susceptibility and increased PM H^+-ATPase activity, we asked what role NDR1 might play in this process. As a foundation for this question, we exploited previous work by Day et al. (Day et al., 2006) which identified an interaction between NDR1 and RIN4. While the significance of this interaction is not fully defined, we hypothesize that one consequence of NDR1's association with RIN4 is to titrate the negative regulatory function of RIN4. In the case of the mechanism described here, we hypothesize that in the absence of NDR1 (i.e., *ndr1*), there is more unassociated RIN4 protein available to interact with the H^+-ATPase. Such a mechanism is supported by the results described in Liu et al. (Liu et al., 2009) using *RIN4* overexpressing plants. Indeed, the basic principle of RIN4-interacting proteins titrating the negative-regulatory potential has been previously proposed (Axtell & Staskawicz, 2003; Day et al., 2006; Liu et al., 2009), and while direct evidence describing the stoichiometry of RIN4-associated protein complexes is unknown, corollary data support this hypothesis.

In the current study, we demonstrate that NDR1 functions in signaling processes associated with both biotic and abiotic stress response pathways, a function we hypothesize this represents NDR1's role in the maintenance of cellular homeostasis during stress. Similar to the activity of mammalian integrin (Brakebusch & Fassler, 2003; Campbell & Humphries, 2011), we hypothesize that NDR1 links the perception of perturbations to cellular homeostasis to the activation of specific signal transduction processes, including the response to drought and pathogen infection. As a next step in understanding the mechanism of NDR1's function in stress perception, it will be important to determine how NDR1 activity functionalizes various stress perception responses, and

in turn, modulates downstream signaling to respond to these stresses quickly and specifically. We posit that this activity is mediated in large part by the differential association of NDR1 with various key signaling components, such as RIN4 (immunity) and AHA5 (drought).

Materials and Methods

Plant materials and growth conditions

Arabidopsis thaliana and *Nicotiana benthamiana* plants were grown in a BigFoot series growth chamber (BioChambers) at 21°C under a 12 h light/12 h dark cycle with 60% relative humidity and a light intensity of 120 μmol photons m-2s-1. All *Arabidopsis* plants described in this study were in the Columbia-0 (Col-0) background. For seedling-based assays, *Arabidopsis* seeds were surface sterilized in 50% (v/v) bleach and 0.05% (v/v) Tween-20 for 20 min and washed 5 times with sterile distilled water. Seeds were germinated in 6-well plates containing ½-strength Murashige-Skoog (MS; Sigma-Aldrich) and 0.5% (w/v) MES (2-morpholinoethanesulfonic acid; Sigma-Aldrich), pH 5.7, for 4 d at 40 rpm in an orbital shaker at room temperature.

Genome sequencing

Rosette leaves of a single 4-week-old *ndr1/35S::NDR1* plant were collected and ground into a fine powder in liquid nitrogen. Genomic DNA was extracted using the Qiagen plant DNeasy kit and quantified with the Qubit 1x DNA BR assay kit. A DNA-seq library was constructed using 500 nanograms of genomic DNA using the Kapa HyperPlus prep kit according to the manufacturer's protocol with a fragmentation time of 20 min and five PCR cycles. Libraries were size-selected at 300-600 bp using SPRI beads. The

resultant library was sequenced on an Illumina MiSeq using single-end 75 bp mode to achieve ~10X genome coverage.

Raw reads were trimmed and quality-filtered using Trimmomatic v0.38(Bolger et al., 2014). To identify the transgene insertion site(s), we *de novo* assembled the genome of *ndr1/35S::NDR1* plant using SOAPdenovo (Li et al., 2010). A k-mer size of 31 was used, and all other parameters were left at default settings. The *NDR1* overexpression insert was identified using BLASTn with the plasmid sequence as a query against the assembled contigs. Contigs with 95% or less of their contig length covered by insert sequences were extracted and manually inspected. Contigs containing both the plasmid and *Arabidopsis* genome sequences were used to identify the transgene insertion site by mapping the *Arabidopsis* genomic sequences to TAIR10 assembly.

Seed germination assays

Arabidopsis seeds were surface sterilized as described above and plated on ½-strength MS media containing 0.5% (w/v) MES (pH 5.7) and 0.8% (w/v) agar containing ABA (Sigma-Aldrich), NaCl, mannitol (Sigma-Aldrich), and PEG6000 (Sigma-Aldrich). Fifty seeds were plated per plate. For seed stratification, plates were wrapped in aluminum foil and placed at 4°C for 4 d. After 4 d, plates were transferred to a BioChambers model GC8-2H growth chamber set at 20°C with a 12 h light/12 h dark cycle, ~60% relative humidity, and a light intensity of 120 μmol photons $m^{-2}s^{-1}$. For germination assays, radicle emergence was scored as germination. Experiments were repeated 5 times (independent biological replicates), with n = ~250 seeds for each genotype.

Root length assay

Fifteen seedlings of WT Col-0, *ndr1*, and *ndr1/35S::NDR1* plants were grown on ½-strength MS agar plates for 6 d at 22°C in a Percival chamber 36L5 (Geneva Scientific) under a 12 h light/dark cycle with a light intensity of 155 μmol photons $m^{-2}s^{-1}$. After 6 d, germinated seedlings were transferred onto ½-strength MS plates supplemented with increasing concentrations of ABA, mannitol, PEG 6000, and NaCl. Plates were oriented vertically to facilitate linear root growth. Root length was measured 6 d after transfer onto stress-inducing plates (i.e., 12 d post-germination). The average distance of root elongation was calculated from 4 independent biological replicates from 15 seedlings per genotype per experiment.

Drought stress analysis

Prior to sowing seeds, pots containing water and soil were weight-adjusted across all planted genotypes and experiments to standard water content. A single plant was grown in each pot. For watering, the same volume of water was added to the flats containing potted plants every two-to-three days, as needed. Pot weights were measured every day to calibrate water used/lost for each experiment. For initiation of drought conditions at 3-weeks post-germination, flats containing plants, randomized across genotypes, were bottom flooded so that each pot was saturated with water (Day 0). Flats were randomly arranged in the growth chamber to minimize positional effects (e.g., airflow and light intensity). Plants were subjected to two watering regimes: 1) well-watered, wherein plants were watered three times weekly; and 2) drought stress (DS) conditions, wherein plants did not receive water. Leaf samples were harvested at 12, 14, 16, 17, 18,

19, and 21 d post-water withholding (dpw). Leaf samples for RNA extraction and hormone analysis were frozen in liquid nitrogen immediately and kept at -80°C until processing.

Relative water content analysis

Relative water content (RWC) was evaluated using 4-to-5-week old plants (leaves) subjected to control and water-withholding conditions as follows: one fully expanded *Arabidopsis* leaf from each of 15 plants was harvested at 12, 14, 16, 17, 18, 19, and 21 dpw and immediately weighed to obtain the fresh weight (FW). Next, leaves were floated in sterile dH2O in a 15 ml sterile tube for 24 h at 4°C. After 24 h, the turgid weight (TW) of leaves was measured. Sampled leaves were then dried in an oven at 65°C for 24 h, removed to room temperature for 15 min, and the dry weight (DW) was measured. RWC was calculated using the formula:

$$\mathrm{RWC\%} = \left(\frac{[\mathrm{FW} - \mathrm{DW}]}{[\mathrm{TW} - \mathrm{DW}]}\right) \mathrm{X}\ 100$$

Electrolyte leakage analysis

Electrolyte leakage analysis was performed on leaves from 4-to-5-week old plants subjected to control and drought stress conditions as follows: one fully expanded *Arabidopsis* leaf from each of 15 plants was harvested at 12, 14, 16, 17, 18, 19, and 21 dpw and immediately placed into a 15 ml sterile tube. Next, 4 ml of sterile dH2O was added to each tube and the samples were incubated overnight at 4°C. After overnight incubation, the samples were removed from 4°C and were acclimated to room temperature (25°C, ~1 h). Electrolyte readings (μS cm-1) were recorded using a

SevenCompact S230 conductivity meter (Mettler Toledo), and after collection, samples were boiled for 30 min. After boiling, all samples were cooled to room temperature (25°C; ~1 h), and a final reading was collected. Percent electrolyte leakage was calculated using the following formula:

$$\mathrm{EL\%} = \left(\frac{\text{1st reading } (\mu\mathrm{S/cm})}{\text{2nd reading } (\mu\mathrm{S/cm})}\right) \mathrm{X}\ 100$$

Phytohormone analysis and quantification

Quantification of phytohormones was performed as previously described(Velasquez et al., 2017), with minor modifications. In brief, one hundred mg fresh weight (2-3 leaves, 4-to-5 week-old plants) were flash frozen in liquid nitrogen and stored at -80°C until processing. For extraction, frozen tissue was ground using a TissueLyser II (Qiagen) and incubated on a rocking platform at 4°C for 24 h in extraction buffer (80:20 v/v HPLC-grade methanol:water with 0.1% formic acid (v/v), 0.1 g l-1 butylated hydroxytoluene). Samples were centrifuged at 12,000 x g for 10 min at 4°C and resultant supernatants were collected and were filtered through a 0.2 mm PTFE (polytetrafluoroethylene) membrane (Millipore). Abscisic acid (ABA)-d6 (gift from Dr. Daniel Jones, MSU) served as an internal standard. Injections of plant extracts (10 μl per injection) were separated on a fused core Ascentis Express C18 column (2.1 × 100 mm, 2.7 μm; Supelco) installed in the column heater of an Acquity Ultra Performance Liquid Chromatography (UPLC) system (Waters Corporation). A gradient of 0.15% aqueous formic acid (solvent A) and methanol (solvent B) was applied in a 5-min program with a

mobile phase flow rate of 0.4 μl min-1 as follows: 0 to 0.5 min hold at 99% A/1% B, ramp to 70% B at 3 min, ramp to 100% B at 3.01 min, hold at 100% B to 3.5 min, return to 99% A at 3.51 min and hold at 99% A to 5 min. The column was maintained at 40°C and interfaced to an Acquity TQ-D LC-MS/MS system equipped with electrospray ionization and operated in negative ion mode. The capillary voltage and extractor voltage were set at 2.5 kV and 3 V, respectively. The flow rates of cone gas and desolvation gas were 40 and 800 L h-1, respectively. The source temperature was 130°C and the desolvation temperature was 350°C. Collision energies and source cone potentials were optimized for each compound using QuanOptimize software (Waters Corporation). Peak areas were integrated, and the analytes were quantified based on standard curves generated from peak area ratios of analytes. Data acquisition and processing was performed using Masslynx 4.1 software (Waters Corporation). Analytes were quantified by converting peak area to phytohormone concentration (nM) per gram of dry weight of leaf tissue using a standard curve specific to each compound.

Stomata aperture measurements

Leaf punches (0.25 cm2) from 4-to-5-week old plants were floated on opening buffer for 2-3 h at room temperature (25°C) before assaying. After acclimation, control treatments were imaged to establish a baseline for stomatal guard cell aperture. For evaluation of abiotic and biotic treatments, leaf punches were transferred to stress-inducing conditions (e.g., ABA, PEG6000, Mannitol, and H2O2) for 1 h. After 1 h, leaf punches were visualized using an Olympus IX71 microscope at 100X magnification and images were collected using an attached DP71 camera (Olympus). For each treatment, ~25 stomata were imaged for each plant genotype evaluated. Images were analyzed

using ImageJ and the data were expressed as a width by length ratio (W/L). Experiments were repeated 3 independent times. Statistical significance was determined using a one-way ANOVA followed by Tukey's post-test.

Agrobacterium transient expression and co-immunoprecipitation assays

The CDS of AHA1, AHA2, AHA5, and mCherry (as the negative control) are cloned to binary vector pEarlygate201 with an HA tag, and NDR1 is cloned to pEarlygate202 containing a FLAG tag. We optimized AHA5 codon using Invitrogen GeneArt GeneOptimizer to synthesize the CDS of *AHA5*.

Co-immunoprecipitation (co-IP) experiments were performed as previously described (Knepper et al., 2011a), with slight modifications. A. tumefaciens strains expressing epitope-tagged constructs of *NDR1*, *AHA1*, *AHA2*, and *AHA5* fused to a 35S cauliflower mosaic virus promoter (constitutive expression) were used for co-IP assays. For transient expression in *Nicotiana benthamiana, A. tumefaciens* GV3101 strains harboring expression constructs of interest were pre-incubated at room temperature (ca. 22°C) in induction media (10 mM MES pH 5.6, 10 mM MgCl2, 150 mM acetosyringone [Sigma-Aldrich]) for 2 h before hand-infiltration into 5-week-old *N. benthamiana* leaves using a 1-ml needleless syringe at a final concentration of 4 x 108 cells ml-1 for each construct. Inoculated plants were kept at 22°C for 40 h (12 h light/12 h dark). Infiltrated leaves were incubated at room temperature (ca. 25°C) under continuous light for 40 h. After incubation, 15-1 cm2 leaf discs were harvested, flash-frozen in liquid nitrogen, and stored at -80°C until processed. Frozen samples were ground into a fine powder in liquid nitrogen using a mortar and pestle. Once pulverized, the powder was transferred to a chilled mortar and pestle and further homogenized in extraction buffer (50 mM (4-(2-

hydroxyethyl)-1-piperazineethanesulfonic acid) [HEPES] pH 7.5, 50 mM NaCl, 10 mM ethylenediaminetetraacetic acid [EDTA], 0.2% (v/v) triton X-100, and 1 cOmplete, mini, EDTA-free protease inhibitor cocktail tablet [Sigma; 1 tablet/50 ml homogenization buffer]). Following homogenization, samples were transferred using a pipette to sterile 1.5 ml Eppendorf tubes. Samples were centrifuged at 15,000 x g for 20 min at 4°C, and the resultant supernatant (1 ml per tube) was transferred to 1.5 ml tubes. Five μl of the designated monoclonal antibody was added to each tube; an input control containing no antibody was also prepared. Following incubation at 4°C for 1 h (mixing end-over-end), 50 μl of pre-washed protein-G sepharose-4 fast flow (GE Life Sciences) was added to each tube and tumbled end-over-end for 4 h at 4°C. After 4 h, immunocomplexes were pelleted by centrifugation (5000 x g) and were washed 3 times in wash buffer (50 mM HEPES pH 7.5, 50 mM NaCl, 10 mM EDTA, 0.1% (v/v) triton X-100, and 1 cOmplete mini EDTA-free protease inhibitor cocktail tablet [1 tablet/50 ml homogenization buffer]). After the final wash, immunocomplexes were pelleted by centrifugation (5000 x g), the wash buffer was removed by pipetting, and the pelleted samples resuspended in 50 μl of 3x Laemmli buffer [4% (w/v) SDS, 20% (v/v) glycerol, 0.004% (w/v) bromophenol blue, 0.125 M Tris-Cl pH 6.8, and 10% (w/v) DTT (dithiothreitol)]. Samples were boiled for 5 min, cooled on ice (1 min), and 10 μl of the denatured samples was separated by SDS-PAGE. Following SDS-PAGE, samples were transferred to nitrocellulose (GE Healthcare Life Sciences). Signals were detected using the Super Signal West Pico chemiluminescent substrate (ThermoFisher), with anti-HA-HRP (1:1000; Roche), anti-Flag-HRP (1:1000; Sigma), or anti-cMyc-HRP (1:1000; AbCam).

Immunoprecipitation and tandem mass spectroscopy

To identify NDR1-interacting proteins, leaves from 5-week-old transgenic *Arabidopsis* plants expressing native promoter driven T7-NDR1 (i.e., *ndr1/PNDR1::T7-NDR1*;(Knepper et al., 2011a)) were harvested as described above (co-immunoprecipitation assays). Total protein extracts were incubated with 5 μl T7 antibody (EMD Millipore; Cat # 69522) for 4 h, tumbling end-over-end, at 4°C. Following incubation, 50 μL of protein-G sepharose-4 fast flow (GE Life Sciences) (pre-washed and suspended in: pre-washed (50 mM HEPES pH 7.5, 50 mM NaCl, 10 mM EDTA, 0.1% (v/v) triton X-100, and 1 cOmplete mini EDTA-free protease inhibitor cocktail tablet [1 tablet/50 ml homogenization buffer]) was added to each sample and tumbled end-over-end at 4°C for 16 h. Immunoprecipitation reactions were washed 3 times in wash buffer (50 mM HEPES pH 7.5, 50 mM NaCl, 10 mM EDTA, 0.1% (v/v) triton X-100, and 1 cOmplete mini EDTA-free protease inhibitor cocktail tablet [1 tablet/50 ml homogenization buffer]), and suspended in 3x Laemmli buffer. The reaction was boiled for 5 min, and samples were separated by SDS-PAGE (15% Tris-Bis).

Resolved gels were stained with CBB (Coomassie brilliant blue), and the stained protein band corresponding to the expected molecular weight of T7-NDR1 was excised and subjected to in-gel digestion according to previously described methods (Shevchenko et al., 1996), with slight modification. Briefly, excised gel bands were dehydrated in 100% acetonitrile (ACN) and incubated with 10 mM DTT in 100 mM ammonium bicarbonate (pH 8.0) at 56°C for 45 min. After 45 min, samples were dehydrated again and incubated in the dark with 50 mM iodoacetamide in 100 mM ammonium bicarbonate for 20 min. Gel bands were then washed with ammonium bicarbonate and dehydrated again. Sequencing

grade modified trypsin was prepared at a concentration of 0.01 μg μl-1 in 50 mM ammonium bicarbonate, and 50 μl of trypsin was added to each sample such that the gel was completely submerged. Samples were incubated overnight at 37°C. Peptides were extracted from the gel by water bath sonication in a solution of 60% (v/v) ACN/1% (v/v) trichloroacetic acid (TCA) and vacuum dried to ~2 μl. Peptide samples were then re-suspended in 2% (v/v) ACN/0.1% (v/v) trifluoroacetate (TFA) to 20 μl. From this, 5 μl was injected onto a Thermo Acclaim PepMap RSLC 0.075 mm x 20 mm C18 trapping column and washed for 5 min using a Thermo EASYnLC. Peptides were eluted onto a Thermo Acclaim PepMap RSLC 0.075 mm x 500 mm C18 analytical column over 35 min with a gradient of 2% B to 40% B over 24 min, ramping to 100% B at 25 min and held at 100% B for the duration of the run (Buffer A = 99.9% water/0.1% (v/v) formic acid, Buffer B = 80% ACN/0.1% (v/v) formic acid/19.9% (v/v) water).

Eluted peptides were sprayed into a Q-Exactive HF-X mass spectrometer (ThermoFisher) using a FlexSpray spray ion source. Survey scans were taken in the Orbitrap (45,000X resolution, determined at m/z 200) and the top twenty ions in each survey scan were then subjected to automatic higher energy collision induced dissociation (HCD) with fragment spectra acquired at 7,500X resolution. Resultant tandem mass spectrometry (MS/MS) spectra were converted to peak lists using Mascot Distiller (ver. 2.7.0; www.matrixscience.com) and the output was searched against all entries in the TAIR protein sequence database (www.arabidopsis.org; ver. 10). All entries in the UniProt E. coli protein sequence database (downloaded from www.uniprot.org, 2017-11-01), appended with common laboratory contaminants (downloaded from www.thegpm.org, cRAP project), were scanned using Mascot (ver. 2.6). The Mascot

output was analyzed using Scaffold Q+S (ver. 4.8.4; www.proteomesoftware.com) to probabilistically validate protein identifications. Assignments validated using the Scaffold 1% FDR (false discovery rate) confidence filter were considered true.

Biomolecular fluorescence complementation assays

The CDS of *AHA1*, *AHA2*, *AHA5* (codon optimized, described above), and *WRKY40* (as the negative control) were cloned into binary vector pBH1096 (*2X35S::Gateway::HA-cEYFP*) for *cEYFP* fusion, and *NDR1* was cloned into *pBM1089* (*2X35S::myc-nEYFP::Gateway*) for nEYFP fusion. A mixture of Agrobacterium strains expressing nEYFP- and cEYFP-fusion constructs was hand-infiltrated into *N. benthamiana* using the same approach as described above for co-IP assays. Leaf samples were harvested at 2 dpi and were imaged using a confocal microscope system (Olympus FluoView 1000). To accurately compare samples with positive/negative fluorescence, all images were captured using the same laser intensity, dwell time, and PMT voltage. The same linear transformation, abiding by academic criterion of biological imaging, was applied to the LUTs of all images to enhance contrast.

Plasma membrane H+-ATPase activity assay

Plasma membrane isolation was performed according to previously published methods (Jakubowska & Janicka, 2017). In brief, 30 g of 5-week-old *Arabidopsis* leaves were homogenized in 100 ml ice-cold homogenization buffer containing 0.33 M sucrose, 50 mM Hepes-KOH (pH7.5), 10% glycerol (wt/vol), 0.2% BSA (wt/vol), 5 mM DTT (dithiothreitol), 5 mM EDTA, 5 mM ascorbate, 0.6% polyvinylpyrrolidone, 1 mM PMSF (phenylmethylsulfonyl fluoride) and cOmplete EDTA-free protease inhibitor cocktail tablet [1 tablet/100 ml homogenization buffer]). The homogenate was passed through five layers

of Miracloth to remove debris and was then centrifuged at 18,000 x g for 20 min at 4ºC. A microsomal pellet was isolated by ultracentrifugation at 120,000 x g for 30 min at 4ºC and resuspended in a buffer containing 0.33 M sucrose, 3 mM KCl, 0.1 mM EDTA, 1 mM DTT, 1 mM PMSF, 5 mM potassium phosphate (pH7.8) and cOmplete EDTA-free protease inhibitor cocktail tablet [1 tablet/100 mL buffer]). A plasma membrane-enriched fraction was purified from the microsomal fraction by two-phase partitioning(Day et al., 2001). The resuspended microsomal fraction was added to a centrifuge tube containing dextran T500 and polyethylene glycol (PEG) 3350 in 0.33 M sucrose, 3 mM KCl and 5 mM potassium phosphate (pH 7.8). The contents were mixed by inverting 20-25 times and centrifuged in a swinging bucket rotor at 1000 x g for 10 min at 4°C. The PEG upper phase was removed, diluted 5-fold in suspension buffer (0.33M sucrose, 20 mM Hepes-KOH (pH 7.5), 10% glycerol (wt/vol), 0.1% BSA (wt/vol), 2 mM DTT (dithiothreitol), 0.1 mM EDTA, cOmplete EDTA-free protease inhibitor cocktail tablet [1 tablet/100 ml homogenization buffer]) and centrifuged at 120,000 x g for 45 min. The resultant plasma membrane pellet was solubilized in suspending buffer (0.33 M sucrose, 20 mM Hepes-KOH (pH 7.5), 10% glycerol (wt/vol), 0.1% BSA (wt/vol), 2 mM DTT (dithiothreitol), 1 mM EDTA, cOmplete EDTA-free protease inhibitor cocktail tablet [1 tablet/100 ml homogenization buffer]) to a final concentration between 2-4 mg ml-1. H+-pumping activity was detected by a decrease in ACMA absorbance at 495 nm(Palmgren, 1990).The assay buffer contained 20 mM MES-KOH (pH 7.0), 4 mM ATP-Na2, 140 mM KCl, 100 μM ACMA (9-amino-6-chloro-2-methoxy acridine), 0.05% Brij 58, and 50 μg of plasma membrane protein in a total volume of 1 ml. Plasma membrane-enriched samples were pre-incubated at 25ºC for 5 min in assay buffer. The assay was initiated by the

addition of 4 mM $MgSO_4$ and the absorbance (A495nm) was measured over a time interval from 4 to 5 min. Each experiment was repeated 2-3 times with independent plasma membrane preparations.

Quantitative real-time PCR (qPCR)

Total RNA was extracted from the leaves of 4-to-6-week old plants using the RNeasy Plant Mini Kit (Qiagen). RNA concentrations were quantified using a Qubit RNA BR Assay Kit (ThermoFisher) in combination with a Qubit 4 Fluorometer (ThermoFisher). First-strand cDNA was synthesized from 1 μg total RNA using the SuperScript™ III first-strand synthesis system (ThermoFisher). Quantitative real-time PCR (qRT-PCR) was performed using the Applied Biosystems 7500 Real-Time PCR instrument (ThermoFisher) using the SYBR Select qPCR master mix (ThermoFisher). Ubiquitin (*UBQ10*) served as an internal control for amplification. The abundance of target gene transcripts was normalized to *UBQ10* according to the $2^{-\Delta\Delta CT}$ method (Livak & Schmittgen, 2001). All qPCR DNA primers utilized in this study are listed in Table AA 3-1.

TaqMan gene expression assay

Leaves from 3-week-old wild type (WT) Columbia-0 (Col-0), mutant, and complementation lines (i.e., *ndr1*, *ndr1/35S::NDR1*, *aha5-2*, and *aha1-9*), were collected and immediately frozen in liquid nitrogen. Total RNA was extracted using the RNeasy Plant Mini kit (Qiagen). cDNA was synthesized with Maxima H Minus First Strand cDNA Synthesis Kit (Thermo-Fisher). Confirmation of the *aha5-2* mutant line was performed using TaqMan Gene expression Assay kits (Applied Biosystems) for HA5 (Assay ID At02274124_g1) and *UBQ10* (Assay ID At 02353385_S1) according to manufacturer's

protocol in ABI StepOnePlus thermocycler v2.3 (Applied Biosystems). The relative levels of each mRNA transcript were calculated and normalized to the expression of the *UBQ10* gene.

Library preparation and sequencing

mRNA libraries were prepared using the Illumina TruSeq mRNA library preparation kit from three biological repeats of each timepoint of well-watered or water-withholding conditions of WT Col-0, *ndr1*, *ndr1/PNDR1::T7-NDR1* (Knepper et al., 2011a), and *ndr1/35S::NDR1* (Coppinger et al., 2004) plants by the Michigan State Research Technology and Support Facility (RTSF). Multiplexed samples were pooled and sequenced on the Illumina HiSeq 4000 (single end 50 bp mode). Base calling was performed using the Illumina Real Time Analysis (RTA; v2.7.7), and the output of RTA was demultiplexed and converted to FastQ format using Illumina Bcl2fastq (v2.19.1).

Analysis of RNA-sequencing

The adapter sequences and low-quality bases ($q < 10$) were trimmed by Trimmomatic (Bolger et al., 2014). Resultant cleaned reads were mapped to the TAIR10 reference genome using HTseq61. Mapped read counts for each gene were generated using the HTseq command, and differentially expressed genes (DEG) were analyzed using DESeq2 (Love et al., 2014). For the identification of DEGs, a threshold of P-value < 0.01 and at least a 1-fold change was used as a criterion in this study. Gene ontology (GO) term enrichment analysis using Fisher test with a significance Level < 0.05 was performed by using AgriGO v263.

Coexpression network analysis was performed using the R package WGCNA

Normalized read counts of WT Col-0, the *ndr1* mutant, and the NDR1-

overexpression line at 0, 12, 14 16 ,17 18, 19, and 21 day timepoints under water-withholding treatment from DESeq2 (Love et al., 2014) were used for constructing a singed hybrid network using WGCNA (Langfelder & Horvath, 2008). A matrix of Pearson correlation between all pairs of 31319 genes was calculated. The adjacency matrix was then constructed by using a power of 18, a merge CutHeight of 0.25, and module size greater than 30. Average linkage hierarchical clustering was applied to the topological overlap for grouping genes with highly similar coexpression relationships. The expression profiles of each module was summarized by module eigengene resulting in 15 modules. Gene ontology (GO) term enrichment analysis was performed using Fisher test with a significance level < 0.05, and minimum number of mapping entries of 5 was performed by using AgriGO v2 (Tian et al., 2017).

Data availability

Sequence data from this article can be found in the *Arabidopsis* genome initiative or GenBank/EMBL databases under the following accession numbers: At3g20610 (*NDR1*), At2g18960 (*AHA1*), At4g30190 (*AHA2*), At2g24520 (*AHA5*), At5g67030 (*ABA1*), At1g52340 (*ABA2*), At3g14440 (*NCED3*), At5g52300 (*RD29B*), At2g27150 (*AAO3*), At1g30360 (*ERD4*), and At5g61380 (*TOC1*).

The Illumina RNA-seq reads were deposited in BioProject under the BioProject ID: PRJNA595619.

CHAPTER 4: Archaeological Bolivian maize genomes suggest diversity is associated with Inca cultural expansion and environmental variation in South America

The work presented in this chapter:

from Huan Chen et al., 2023,

In Revision

Chapter abstract

Previous archaeological and anthropological studies have demonstrated the myriad ways that cultural and political systems shape access to and preferences for food. However, few studies have carried out biocultural analysis linking specific genotypic/phenotypic traits as evidence of cultural selection in ancient contexts. Here, we address this issue by comparing genomes of maize samples from Bolivia dating to ~500 BP included as an offering with the mummified remains of a young girl to 15 archaeological maize samples spanning at least 5,000 years of evolution, and 87 modern western hemisphere maize samples. Our phylogenetic analysis showed that the archaeological Bolivian maize (aBM) has the closest genetic distance to archaeological maize from ancient Peru, which in turn was derived from central Mexican domesticated varieties. The comparative genomic analysis identified specific modifications in the aBM genome representing traits that were selected by ancient farmers, including shorter growing season, greater stress-resistance, and sweeter taste. The genome of the aBM appears to reflect gene flow from Chile coinciding with the arrival of the Inca empire in the altiplano region surrounding modern-day La Paz, as well as selection for specific traits related to growth in this marginal environment and taste preference. The stronger heat stress response character was selected from American modern maize and was compared to ancient maize. Our study provides insights into the complex biocultural role that Inca culture had in determining the direction of maize diversity in South America.

Introduction

Globally, modern maize (*Zea mays* ssp. *mays*) is one of the most important and productive food crops (Ranum et al., 2014). Prior to European contact in the Americas,

maize had been adapted for a wide range of environments. Its versatility is clear from its extensive geographic distribution today, ranging from 58^{o} north latitude to 40^{o} south latitude, and growing below sea level as well as 3,600 meters above sea level in the Andes (Bonavia, 2013). Studies of archaeological samples and modern molecular data indicate that maize originally evolved in Mexico from the flowering species of grass teosinte (*Zea mays* ssp. *parviglumus*) ~9,000 years before the present (cal. BP) (Doebley, 2004; Matsuoka et al., 2002) and spread into South America by ~7,000 cal. BP (Grobman et al., 2012; Lombardo et al., 2020) and into Eastern North America by ~2500 cal BP (Hart & Lovis, 2013). During this domestication process there was hybridization with teosinte and Mexicana (Hufford et al., 2013) helped maize genetic diversity and adapt to a wider range of climates, soils, and different environmental environment. The dispersal of maize to the rest of the world followed European colonization of the Americas in the 15th and 16th centuries AD (Bonavia, 2013; Rebourg et al., 2003). Its evolution and diversity have been shaped by both cultural and natural selection in a wide range of environmental and social contexts. For example, evidence has shown that cultural preferences were a factor in maintaining otherwise undesirable traits (Bellon, 1996; Knapp, 1997). Maize seeds were selected and retained by farmers for use in the following growing year(s) to maintain or enhance specific morphological traits, such as color and texture, based on cultural preferences of the time, sometimes at the cost of reduced cultivability or nutrition (Louette & Smale, 2000).

Politics and culture had a strong influence on crop planting and food preferences in ancient civilizations. Indeed, in the Andes, previous research showed that under the Inca empire, maize was a sacred crop (Mintz, 1985), a symbol associated with political

power, as well as tightly associated with the life of high-status or elite social groups (Bray, 2003). It was also an important element in numerous cultural events and rituals (Ceruti, 2015). From an agriculture-economic point-of-view, planting a crop on the most suitable land takes priority if a farmer wishes to maximize the output-to-input ratio. However, research has shown that in certain civilizations, such as Andean groups, maize was not a crop for which yield was considered a deciding factor. For example, in Andean highland communities, maize was not the most important food crop during pre-Inca times (Burger & Vandermerwe, 1990), not only because maize was ill-adapted to arid highland conditions and its cultivation was risky during the short growing season, but also because planting maize did not yield a return on investment. In the Andes during the Spanish conquest, the chronicler Cobo observed that, “Indians dropped all that they were doing in order to produce a little maize, even though it cost them more than it was worth”. It seemed to him that work in the fields was "…one of the major forms of recreation and festivals…”, and he was impressed by the "extraordinary fondness for farming" (Cobo & Hamilton, 1990; Doutriaux, 2001). This did not, however, deter Andean farmers from planting an abundance of maize. In fact, while at the very least it was a desirable food, it was more so a symbol of empire (Doutriaux, 2001) and a ritual-feast crop, since its first introduction to the region around 3000 BC (Pearsall, 1978).

Given the availability of ancient maize samples, as well as the ability to associate maize with the cultural rituals of one of the historically most powerful civilizations, we were curious as to how – if at all – maize diversity was associated with Incan cultural influence and modern maize biogeographic patterns in South America. The current study focuses on maize samples from an offering associated with an elite stone tomb burial context from

south of La Paz, Plurinational State of Bolivia, and dated ca. 1470 cal AD. We sequenced six sub samples which were derived from two archaeological Bolivian Maize (aBM) kernels, using short read sequencing. In parallel, we used ^{14}C Accelerator Mass Spectrometer (AMS) radiocarbon dating to determine the approximate age of the samples. Phylogenetic analysis of a total of six aBM,15 archaeological and 87 modern maize samples, 17 modern samples of *Zea mays* ssp. *parviglumis*, 37 modern samples of *Zea mays* ssp. *mexicana* and 1 *Tripsacum* genotype revealed genetic relationships among themselves (Table AA 4-1; Data Set AA 4-1). Finally, comparative genomic analyses identified regions of the genome under selection. Candidate genes associated with ancient Peruvian maize traits contributed to a shorter growing season, more stress-resistance, and a sweeter taste. Based on the data presented herein, we conclude that a combination of Inca culture, the social and political needs of local ethnic groups, and mitigation of certain natural environmental factors may have contributed to the specific preference for and diversification of maize in at least this part of South America. As such, the current study offers insight into the possible pathways that influenced maize genetic diversity, thus providing an important resource to further investigate the factors that drove selection for maize domestication and diversity under the cultural influence of the Inca empire.

Results

Archaeological Bolivian maize sample (aBM)

La Paz is a major city located in the highlands of west-central modern Bolivia, which is 16°30′S, 68°09′W, 3,650 masl, (Figure 4-1a). Dry conditions in this region and the adjacent Andes Mountain range have aided in the preservation of desiccated organic

material. Our maize samples were found within a woven camelid fiber pouch that was one of many offerings associated with the mummified remains of a child reported as being found in a chullpa (stone tower tomb) south of La Paz, Plurinational State of Bolivia around 1890 (Figure 4-1b). The assemblage was acquired by the U.S. Consul to Chile and was sent as a gift to the Michigan State University Museum, where she was curated until 2019. Following a series of recent analyses, she has since been voluntarily repatriated to Bolivia (Lovis, 2019). Given the purported archaeological context of the mummy and her offerings, it is likely that she represents an elite member of society and in the decades following her death, community members periodically visited her chullpa and left offerings. These included the pouch full of well-preserved maize from which we collected and analyzed six subsamples from two kernels (Table AA 4-2 and 4-3). A single Accelerator Mass Spectrometer (AMS) date was obtained on one kernel, yielding a radiocarbon age of 400±26 BP (D-AMS-027148; maize kernel; corrected for fractionation, no $\delta^{13}C$ reported). For the date 400±26 BP the two possible calibrated age ranges are 1440-1520 cal AD (p=0.823) and 1588-1621 cal AD (p=0.177), with a median age of cal AD 1474. (Calibrated at 2σ with Calib 8.2 [IntCal 20]). Two associated maize kernels from the same pouch were subjected to molecular analysis.

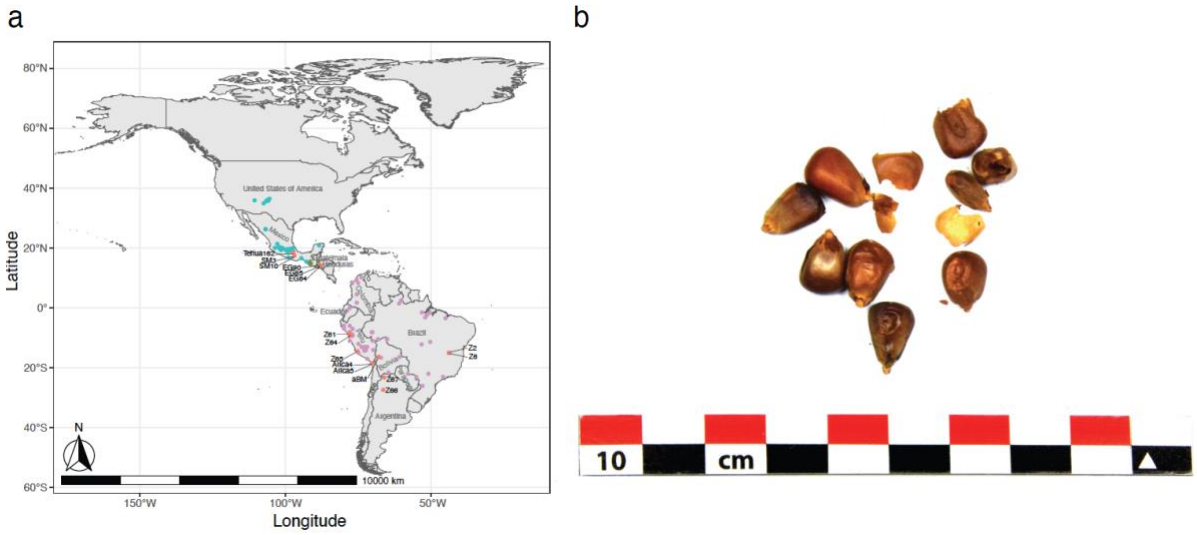

Figure 4-1: Location, appearance, and age of kernel of archaeological Bolivian maize (aBM). (a) Map of maize sample collection and location information. (b) Photographs of kernel showing morphological characteristics.

To provide additional support for this method, we also estimated age using genomic information by identifying base modification patterns associated with ancient DNA (aDNA), including cytosine deamination damage patterns. As predicted, we identified several characteristics indicative of aDNA. First, we observed position-specific substitutions at the ends of the sequence reads (e.g., C→T/G→A(Dabney et al., 2013; Gilbert et al., 2003; Handt et al., 1996; Hansen et al., 2001; Krings et al., 1997) (Figure AA 4-1a), as well as an increase in adenine (A) and guanine (G) residues at the 5' end, concomitant with a reduction of C and T residues (Briggs et al., 2007; Dabney et al., 2013; Stiller et al., 2006) (Figure AA 4-1b). Both of these features are indicative of postmortem deamination patterns. From this, we evaluated the read length and error rate of all substitution types (Figure AA 4-2), including read length range (Paabo, 1989) and high error rate types, both of which are associated with expected aDNA (Prufer et al., 2010).

In total, these data are consistent with the radiocarbon results, supporting that 1) the maize is in fact consistent with an ancient biological sample, and 2) based on the geographic origin of the sample, the kernels potentially descend from the period of the last Inca empire before the Spaniard Francisco Pizarro's conquest of the Inca capital, Cuzco, in A.D. 1533 (Wilson et al., 2007).

Our study focused on identifying and explaining patterns of genetic variability of maize across the Americas, with a specific focus on South and Central America strains that are related to our current aBM samples. We were especially interested in exploring how these patterns of variability found in both ancient and modern samples reflect the influence of cultural systems of behavior on both the spread of variants across the ancient landscape and selection for phenotypic traits by local populations. In particular, it appears that the Inca empire had a large effect on this process of genetic differentiation.

Genetic identity, genetic relatedness, and the spread of maize

After Tupac Inca Yupanqui became the emperor of the Inca Empire in A.D. 1471, the territorial range of the Inca empire was greatly expanded across much of South America, including the lands of the Qulla, who are an indigenous people of western Bolivia, Chile, and Argentina living west of Jujuy and west of Salta Province, and of the Pachamac Aymara south of La Paz, from where the maize samples derive. Because

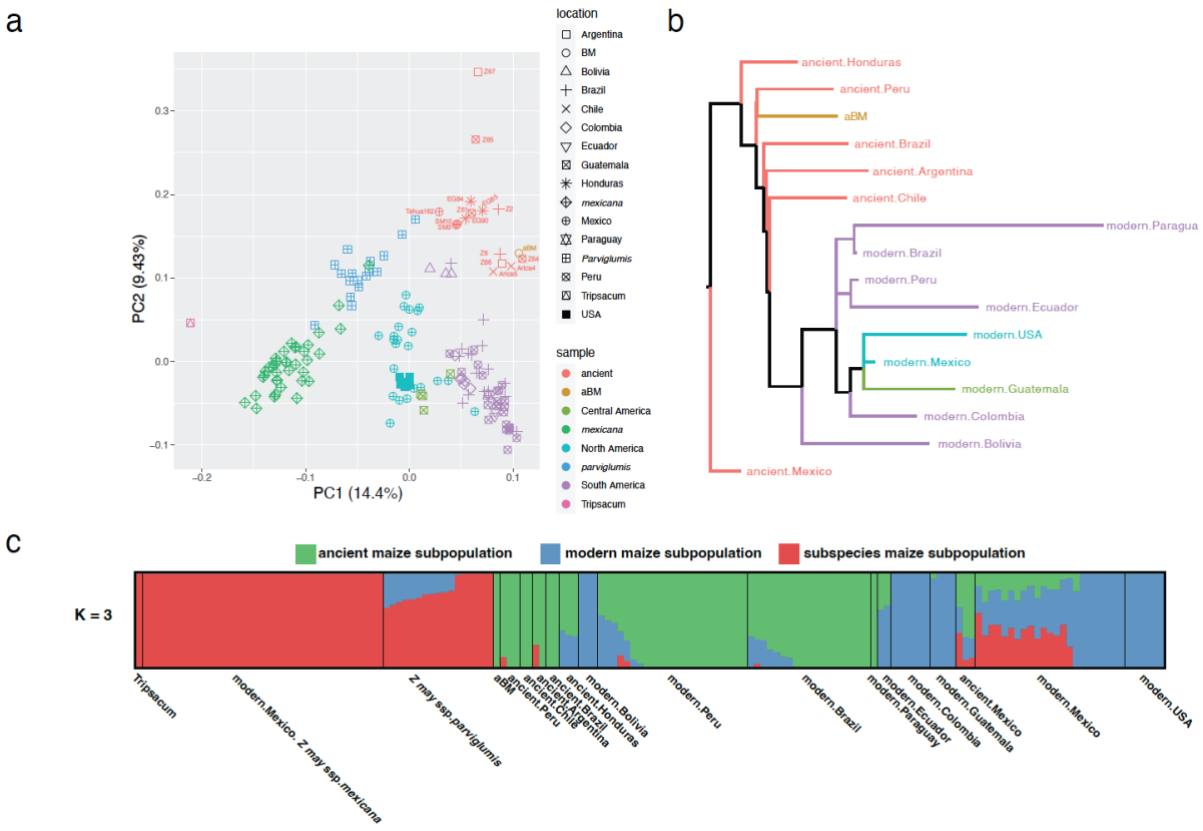

Figure 4-2: Phylogenetic relationships and population structures of diversity of maize. (a) Principal component analysis of ancient maize and modern and maize. (b) Maximum Likelihood (ML) tree of all maize groups. (c) Clustering of the population using ADMIXTURE with K = 3. The colors used in c do not represent the same grouping as panels a and b.

maize was more of a sacred than subsistence crop of the Inca and to a degree symbolic of Inca empire control, the extent of its distribution as a crop may have been related to the extension of social and political hegemony over newly acquired territories. Given the provenience of the aBM and its age, it is very possible it was introduced into western highland Bolivia from the Inca core area – modern Peru. To test this hypothesis, we first evaluated the genetic relatedness of ancient and modern maize samples using a genome sequence approach to assist in identifying the origin(s) of aBM and the relatedness to other maize from different locations. To find the potentially closest relative of aBM, a

principal component analysis (PCA) was used for clustering. As expected, *Tripsacum* is clearly separated from Parviglumis, Mexicana, and modern *Zea mays*, and there is overlap between Parviglumis and Mexicana, but no overlap with maize from South America. This is because Parviglumis is only distributed in Mexico, not in South America, which explains the introgression between maize and Parviglumis in Mexico. Parviglumis is very close to the oldest ancient Mexican maize Tehua162 since Parviglumis is the wild ancestor of maize that was first domesticated. However, Tehua162 is not close to the other two Mexican maize (SM3 and SM10) since it has been shown to be a curious lineage that is largely unrelated to domesticated maize and is probably an extinct lineage of maize domestication. The Central American cluster has overlap with both North and South America, which supports the well-known distribution route of maize(Wang et al., 2021). The results revealed that all the ancient maize was clustered with our aBM in the same group, which also supports the authenticity of our aBM as ancient maize. We were surprised that Z67 did not align very closely with other ancient maize within the cluster, a result due to the relatively young age of Z67 (Argentina; ~100 cal. BP) compared to aBM (~ 500 cal. BP). Among the ancient and modern maize clusters, aBM is most similar to ancient Peruvian maize Z64 (Figure 4-2a). The Maximum Likelihood (ML) tree of all maize

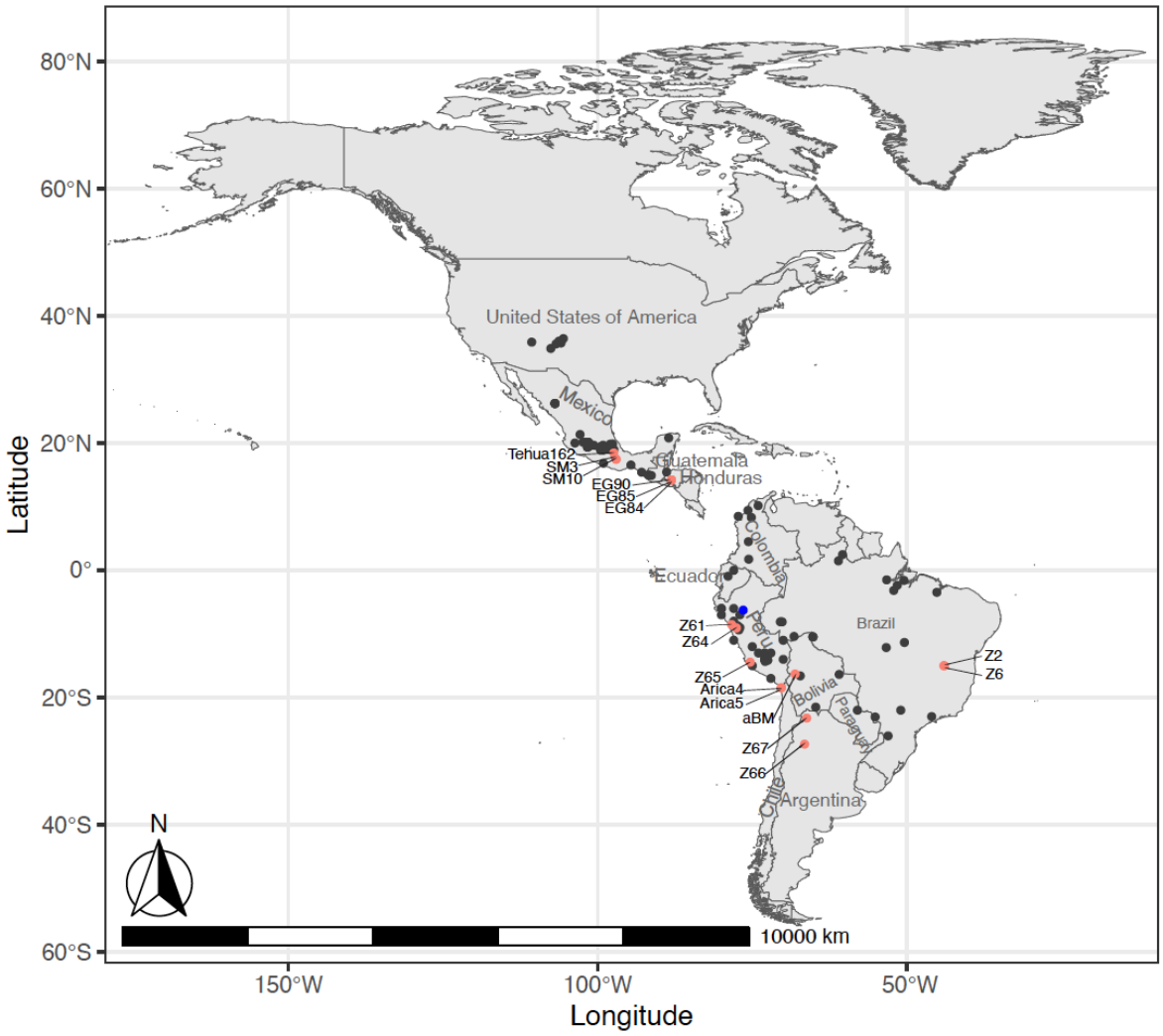

Figure 4-3: Predicted geographic location of archaeological Bolivian maize (aBM). The blue point is the predicted locations of aBM based on the SNPs from all ancient maize. The pink points are the ancient maize samples.

groups also showed the most likely close relative for aBM is ancient Peruvian maize (Figure 4-2b). All the modern Bolivian maize samples, and one Brazilian maize sample which comes from the boundary of Incan territory in Argentina and Brazil, are closest to ancient maize among all modern maize (Figure 4-2a). Both the standard Neighbor-Joining Tree and Maximum Likelihood (ML) Tree also showed topology similar to Figure 4-2a, b (Figure AA 4-3). The cluster consisting of only ancient maize also showed that ancient Peruvian maize Z64 is the closest maize to aBM (Figure AA 4-4). This was also supported

by the maize ecological regions of aBM and Z64. Although Z61 came from a location that is very close to aZ64, we did not see a strong similarity between aBM and Z61, likely due to it being lowland maize and Z64 is a highland maize variety (Figure 4-1a, Table AA 4-1). As the second and third closest maize to aBM, ancient Chilean maize Arica4 and ancient Argentinian Z66 are far away from Z64 by geographic distance, but the characters of coastal maize and highland maize makes them closer to aBM than lowland maize Z61 (Figure 4-2a). This indicates that aBM has strong genetic relatedness with ancient Peruvian highland maize Z64.

In the current study, spatial auto-correlations in genetic data, which allowed us to predict the geographic location of aBM by comparing it to a set of maize samples of known geographic origin (Battey et al., 2020). Using this prediction method, we conclude that aBM came from a location nearby Z64 (Figure 4-3). This is interesting, as historically, evidence from recordings of capacocha rituals showed that they were able to source their food from the connection stations between high-altitude ritual sites and other regions of the empire (Wilson et al., 2007). Indeed, data generated herein show that our aBM has a strong relationship with ancient Peruvian highland maize in terms of both genetic identity and location. To further this line of investigation, PCA clustering based on different countries of origin showed that maize within Mexico subsequently diverged into two dominant groups, with one group retaining similarity with current local Mexican maize, and another group being similar to maize from neighboring countries, including the United States and Guatemala (Figure 4-2a). Among the second dominant group, a small part of Mexican maize being similar to maize from South America conforms to the maize spread route after it was introduced from Mexico and indicates that maize diverged ever since it

was introduced to South American. This was also supported by the subsequently diverged patterns among South America maize being very similar to each other - especially maize in Brazil and Peru (Figure 4-2a). However, modern maize varieties in Brazil and Peru are not closely related to ancient varieties from the same region (Figure 4-1a; Figure 4-2b; Data Set AA 4-1). The establishment of this new pattern of diversity appears to coincide with the Spanish conquest of South America and the dissolution of the Inca empire. It is possible that European farmers and merchants may have favored different phenotypic traits, and the subsequent spread of specific varieties followed the new global geopolitical maps of the Colonial era.

The analyses also provide interesting insights into the initial spread of ancient maize in the more distant past. The ancient maize cluster showed that all the ancient maize varieties that came from South America are genetically similar to each other. Combining this with the age of these ancient maize varieties and their distance on the PCA figure, the earliest maize in South America appears to be derived from ancient Honduran maize (Figure 4-2b; Table AA 4-1). This finding is consistent with previous research revealing that the original maize came from teosinte in Mexico, then dispersed northward into North America and southward into South America (Doebley, 2004; Lombardo et al., 2020; Matsuoka et al., 2002; Mercer et al., 2008; van Heerwaarden et al., 2011). After maize was introduced into and dispersed across Central and South America, it diverged into different groups in regions that would eventually become the Inca empire.

Following the establishment of an important maize culture center in Peru, maize was domesticated and diverged into different groups, which then spread along with the Inca culture (COOK, 1925). This can be verified in previous research showing that there were 24 different maize groups in Peru during the peak Inca times (Bonavia, 2013). After this (and prior to the Colonial period), one of these groups was spread into newly conquered territories in Chile and Argentina, and from there it slowly diverged. At the same time, maize groups from the Inca territory also dispersed into areas outside of the empire. As we can see, there is one modern Brazilian maize species (Brazil Parana) from the location of near Brazil's border with Argentina that are close to ancient Chile maize Arica4 and Arica5, ancient Argentina maize Z66, ancient Brazilian maize Z6, Ancient Peru maize Z64 and aBM (Figure 4-2a; Data Set AA 4-1). All of the modern Bolivian maize also showed the similar information as the above Brazilian maize (Figure 4-2a). The cluster of ancient maize showed the above ancient maize (Arica4, Arica5, Z66, Z6, Z64, and aBM) are very close to each other. Furthermore, clustering the samples using ADMIXTURE (K=3) showed the aBM genome structure is more similar to all the ancient maize in South America and all the ancient maize from South America are similar to each other (Figure 4-2c, Figure AA 4-5).

The mixture history of aBM from f3 statistic did not show it is a mixture of any two ancient maize populations, but the f3 outgroup with *Tripsacum* showed maize Z67 shared the longest evolutionary time with aBM, followed by maize from ancient Peru, ancient Chile, ancient Argentina, ancient Brazil, ancient Mexico, and ancient Honduras (Figure AA 4-6a). The qpAdm method results with small stderr, and showed aBM has very high proportions of ancestry coming from ancient Peruvian maize and ancient Argentinian

maize (Table AA 4-3). Therefore, it indicates aBM is a group that came from ancient Inca territory maize. In addition, we found gene flow from ancient Chilean maize to modern Paraguay, Brazil, Peru and Ecuador (Figure AA 4-6b and c). This finding is also consistent with the above previous research revealing that one of the maize groups from Peru during Inca peak times spread into newly conquered territories, then slowly diverged to fit the local environment following Incan culture influence. After this, diverged groups were introduced outside of Inca conquered territories.

Our results also contradict the previous hypothesis that maize in South America mixed with *Zea mays* ssp. *parviglumis* to form a new maize variety (Bonavia, 2013), revealed by the fact that PCA clustering did not include any ancient maize in the same cluster as *Zea mays* ssp. *parviglumis* (Figure 4-2a), and the genome structure also showed that *Zea mays* ssp. *parviglumis* is more similar to ancient Mexican maize, modern Mexican maize and *Zea mays* ssp. *mexicana* (Figure 4-2c).

Taken together, these results support our hypothesis that the ancestor of all South American maize dispersed from Mexico and southward through the continent, after which it was cultivated, its characteristics managed, and then dispersed southward and diverged into different groups following the cultural influence of the Inca empire (Cerda-Hurtado et al., 2018; Mercer et al., 2008). After this, maize, both as a primary food and as a symbol of power, was dispersed from Peru throughout the Inca empire, in the end driving maize diversity in South America.

Inca culture, social organization and maize diversity inferences

After establishing the identity of aBM and its relatedness to other regional samples, we explored how aBM was shaped by human behavior, and specifically how cultural

selection by populations within the Inca empire influenced maize diversity in South America. We compare the previous selection of ancient Peruvian maize without Incan

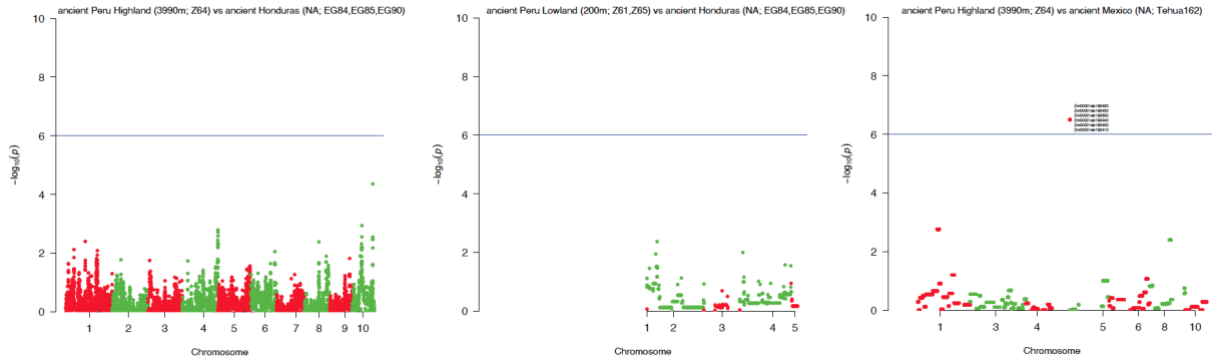

Figure 4-4: Signal of selection of ancient Peruvian maize for genome-wide SNP through genome comparison. Each point indicates a SNP. The blue horizontal line shows the genome-wide significance level ($p = 1 \times 10-6$). Genes associated with significant SNPs are labeled. SNP=single-nucleotide polymorphism.

cultural influence with two significantly older ancient maize groups, which are from ancient Mexico and ancient Honduras (Table AA 4-1). We discovered more common SNPs on the genome region between ancient Peruvian maize and ancient Honduras maize compared to ancient Mexican maize. However, it is unsurprising, considering the closer relationship in terms of geographical proximity, chronological age, and domestication processes between the ancient maize of Honduras and ancient maize of Peru (Table AA 4-5). To identify the signatures of selection by scanning the genome, we utilized the genome comparison method XP-EHH, with cutoff metrics for significant strong signatures of selection being set at value 6. The significant selection was found only between ancient Peruvian maize and ancient Mexican maize - Tehua162, and the six target candidate genes are related to transcription of genes and the energy-related aspects of plant survival (Figure 4-4; Figure AA 4-7; Data Set AA 4-2). Considering there are more

genomic regions shared between evolutionarily related groups, there should be more regions on the genome that are shared between aBM and highland maize since aBM is closely related to ancient highland Peruvian maize. As we expected, aBM exhibits a higher degree of SNP commonality with ancient highland maize in contrast to coastal and lowland maize (Table AA 4-6). The level of SNP commonality between aBM and ancient Honduran maize, as well as ancient Mexican maize, is observed to be lower when compared to that with ancient Peruvian maize (Table AA 4-6). This indicates that there should be more selection signature differences between aBM and the ancient Peruvian highlands maize resulting from the close relationship in both genetic identity and location.

The same XP-EHH method that was used above was also used here. Comparisons were evaluated against other ancient maize groups (Figure 4-5). While not surprising, more significant signatures were identified between aBM and ancient Peruvian highland maize and ancient Argentina highland maize (Figure 4-5; Figure AA 4-8). We hypothesize that this may be due to aBM being the cultivated group from ancient Peruvian highland maize because maize from the ancient Peruvian highlands likely underwent more harsh natural environment tolerance selection to adapt to the higher elevation under Incan cultural influence in the new territory. Followed by ancient Chilean Coastal maize and ancient Brazilian lowland maize, fewer signatures of selection were detected.

Our above results, suggesting that aBM has a very strong genetic identity and location relationship with ancient Peruvian highland maize, also support this. As expected, significant signatures of selection were found when samples from aBM and ancient Peruvian highland maize were compared (Figure 4-5; Table AA 4-6). We identified windows with all significant signatures by setting window size with 5kb, as previous research(Brodie et al., 2016) used to find possible candidate genes (Data Set AA 4-3). The gene ontology enrichment result showed that only one Go term “glycine

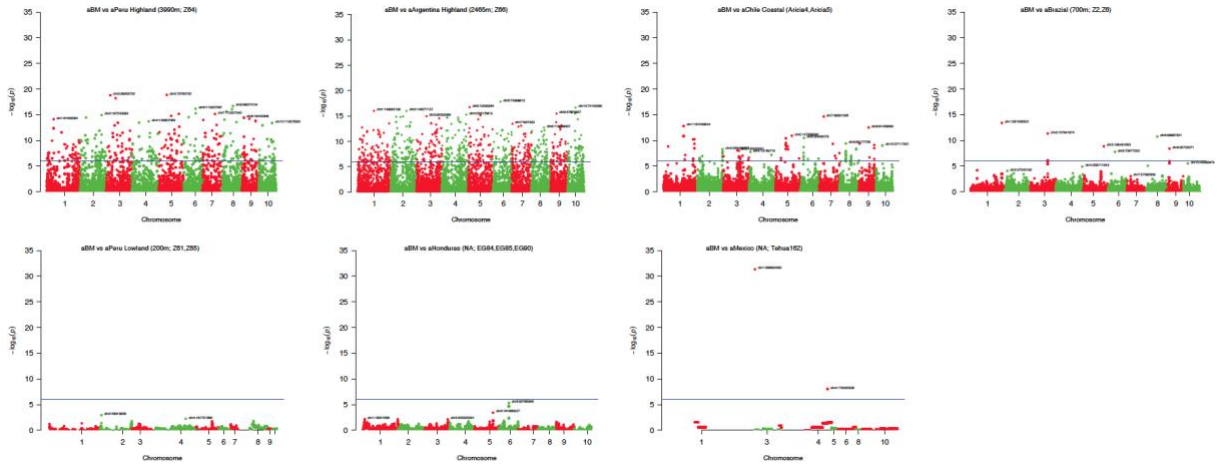

Figure 4-5: Signal of selection of archaeological Bolivian maize (aBM) for genome-wide SNP through genome comparison. Each point indicates a SNP. The blue horizontal line shows the genome-wide significance level (p = 1 × 10−6). Genes associated with significant SNPs are labeled. SNP=single-nucleotide polymorphism.

dehydrogenase (decarboxylating) activity” was enriched by using the target genes from the ancient Peruvian highland and ancient Argentinian highland maize, respectively, and no GO term was found from the target genes in other ancient groups (Data Set AA 4-4). To check which kind of biological process these target genes are involved in, GO Term category showed key variants that only exist between aBM and ancient Peruvian highland maize are involved in root development (Zm00001eb052290), disease (Zm00001eb388620, Zm00001eb050460, Zm00001eb052290), response to stress

(Zm00001eb017930, Zm00001eb050460, Zm00001eb060520, Zm00001eb165400), starch synthase (Zm00001eb194550), faster-growing (Zm00001eb050460, Zm00001eb052290) and better adaptation to daily environmental changes (Zm00001eb017930) (Data Set AA 4-4). In addition, genetic differentiation on flowering, starch and stress response relative biological processes were detected using fixation index (Fst) as in previous research(Swarts et al., 2017) between aBM and other maize groups. It showed aBM has the most Fst similarity on flowering, starch, and cold resistance with ancient Argent highland maize Z67, on response to drought with ancient Argent highland maize Z66, and on light relative biology processes with ancient Peru lowland maize (Z61, Z65) (Data Set AA 4-5; Figure AA 4-9). This indicates traits that are adapted for highland environments have been selected in aBM.

The results suggest that after the aBM groups of maize were introduced into Peru, phenotypic selection by local groups resulted in sweeter and faster-growing variants. This was also consistent with evidence from previous research that showed maize and chicha-maize are associated with elite lifeways(Bray, 2003). This also supported the conclusion that aBM has a strong relationship with ancient maize in Peru based on both genetic identity and location. Aside from this, it also showed that aBM grew stronger, faster and matured more quickly. These traits seem essential for cultivars in the relatively harsh conditions of the highlands, aiding in resisting drought and frost, and potentially providing flexibility in the timing of planting. Moreover, aBM may have had altered polysaccharide structure, which in addition to adding to the taste and texture could also have facilitated fermentation for the creation of chicha.

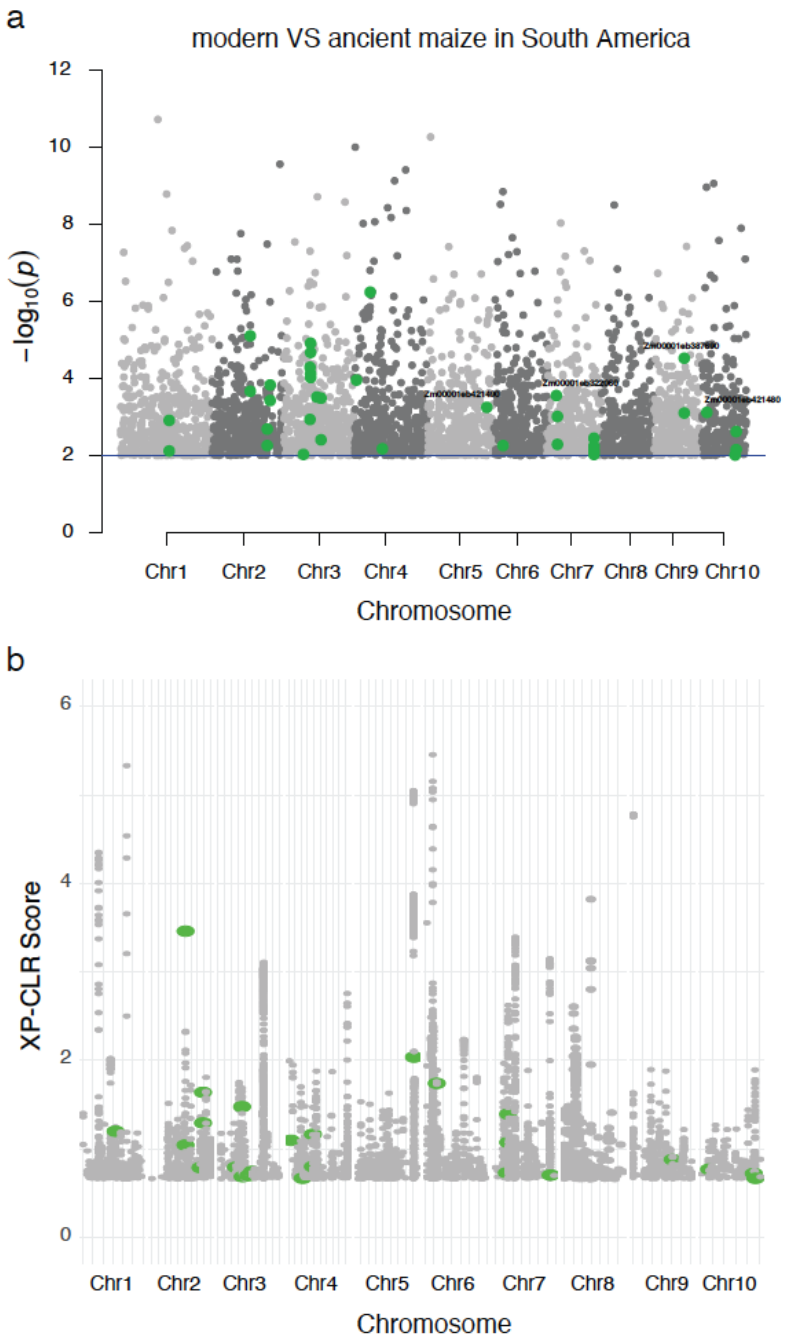

Figure 4-6: Genome-wide distribution of ancient maize specific selective sweeps in South America. a Signal of selection of modern maize compared to ancient maize population in South America. Each point indicates a SNP. The blue horizontal line shows the genome-wide significance level ($p=1 \times 10^{-2}$). b XP-CLR estimated by comparing the modern maize against the ancient maize population in South America. Only top 1% XP-CLR are showed (K = 5,000). Each point indicates a xpclr score for each 5,000 bp region. The green points indicate the common significant signals of selection that were found in both methods.

Finally, we investigated the kind of selective sweep modern maize experienced compared to ancient maize in South America by calculating XP-CLR for 5kb windows across the genome and the genome comparison method XP-EHH. Using this approach, we identified the overlapping regions of highest 1% XP-CLR peaks and the significant selection in XP-EHH (Figure 4-6a and b; Figure AA 4-10). In total, four genes were found, the gene product of Zm00001eb387690 protein is the heat shock protein, which play a crucial role in cellular protection and response to high temperature stress. The function of Zm00001eb421480 as glycolytic process benefits a fundamental part of plant survival and growth and helps convert excess glucose produced through photosynthesis into starch and stored within plant cells (Data Set AA 4-7). This suggests that the preference for selection of maize was in part for stronger stress response and altered grinding or cooking texture. In addition, we were also curious about the gene sweep between ancient and modern maize in the whole of the Western Hemisphere. The same analysis methods used for the comparison of the selection between aBM and ancient maize in South America were applied here (Figure 4-7a and b; Figure AA 4-11). Five target genes were found, and gene Zm00001eb387690 that was found in aBM and ancient maize in South America was also found here (Data Set AA 4-6). All this showed that in the stages of maize diversification, the characteristics of stronger stress response were selected. The genetic differentiation also showed the difference between ancient maize and modern maize (Figure AA 4-12; Data Set AA 4-5).

In addition, both target genes of the top 1% XP-CLR peaks in ancient and modern maize across South America and the entire Western Hemisphere also exhibited enhanced stress and disease response. The target genes, which are related to benefits

in fundamental parts of plants survival and growth traits, were selected over time (Table AA 4-7; Data Set AA 4-7 and 4-8). It is very interesting that the stronger heat stress response character was selected in the whole America modern maize compared to ancient maize. This may be related to global warming resulting from industrialization in the 18th and 19th centuries.

In conclusion, the ancient Bolivian maize included in the funerary accoutrements of a young elite Andean girl belongs to the family of ancient Peruvian maize groups, and likely was introduced to the Aymara region following Inca expansion in the late 15th century. Following its dispersal as a crop across the entire Inca empire, maize was further modified by regional populations. Some changes to the genome reflect selection to mitigate local environmental conditions, including a shorter growing season/maturation rate and greater stress-resistance. The presence of genes associated with polysaccharide structure likely reflect cultural preferences for texture and taste and/or to facilitate the fermentation process during the production of chicha. Thus, patterns of regional diversity in pre-Contact South America maize reflect complex biocultural processes. These data may also inform subsequent studies about Colonial dispersals and the origin of global maize varieties (Rebourg et al., 2003).

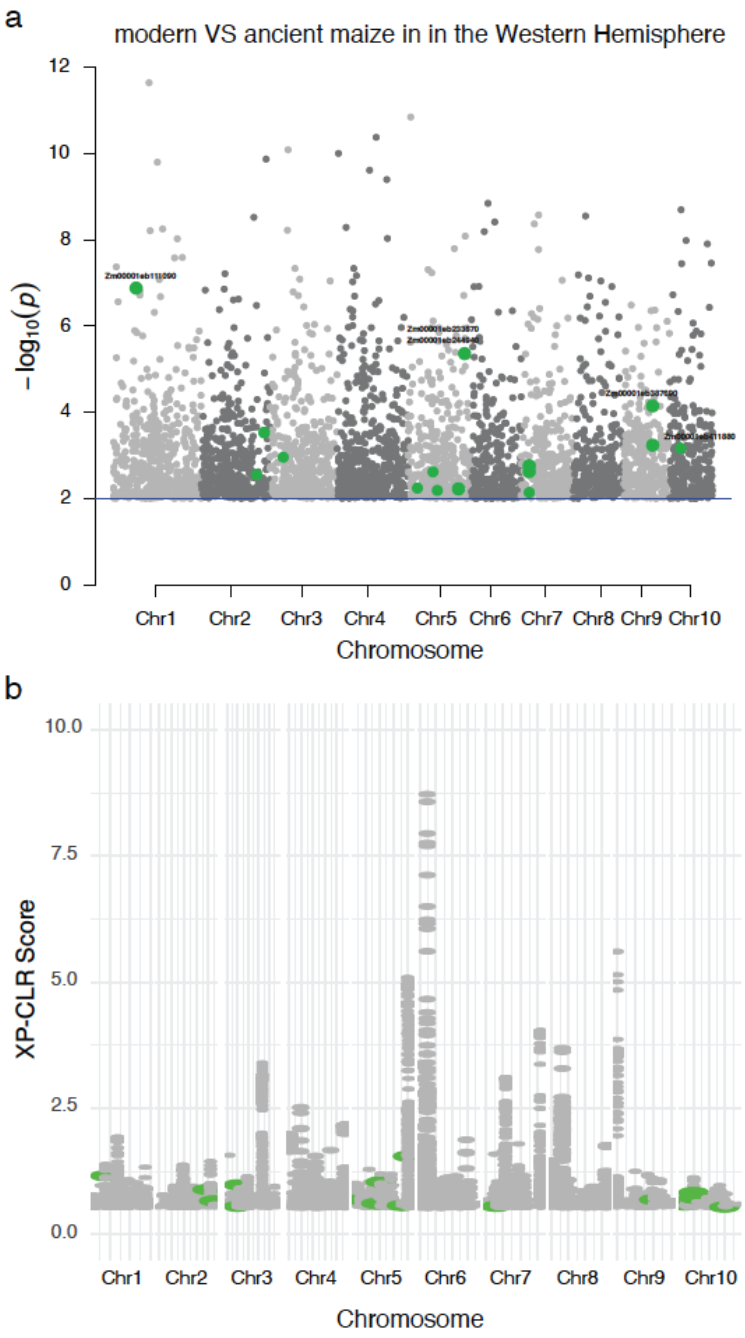

Figure 4-7: Genome-wide distribution of ancient maize specific selective sweeps in the Western Hemisphere. a Signal of selection of modern maize compared to ancient maize population in the western hemisphere. Each point indicates a SNP. The blue horizontal line shows the genome-wide significance level (p=1 × 10-2). b XP-CLR estimated by comparing the modern maize against the ancient maize population. Only top 1% XP-CLR are showed (K = 5,000). Each point indicates a xpclr score for each 5,000 bp region. The green points indicate the common significant signals of selection that were found in both methods.

Materials and Methods

Plant material

The two archaeological Bolivian maize kernel samples were collected prior to 1890 AD from the area south of modern La Paz, Bolivia, curated at the MSU Museum, Michigan State University, and based on AMS assay dates from 401 to 582 cal. BP, with a median age of 1474 cal AD (Table 4-1). In addition, data from 13 archaeological samples collected from previous research were also used, including two from Chile dating from 750 to 1,100 BP, two from Brazil dating from 510 to 690 BP, three from Peru dating from 600 to 1,000 BP, two from Argentina, dating from 980 to 1040 BP and from 70 to 130 BP, three from Honduras dating from 1,740 to 2,300 BP, and one from Mexico dating to 5,310 BP (Table AA 4-1)(Kistler et al., 2018; Kistler et al., 2020).

aDNA extraction

Archaeological maize remains were processed in a clean laboratory environment at Michigan State University. Whole kernels were crushed with a mortar and pestle. Six hundred uL of fresh PTB (N-phenacylthiazolium bromide) buffer (0.4 mg/mL proteinase K, T2.5 mM PTB, 50 nM dithiothreitol) was added. The maize in the buffer was added to 1260 uL of lysis buffer. An additional 600 uL of PTB buffer was used to rinse the mortar and added to the lysis buffer. The individually ground maize kernels were rotated overnight at 37C. The remaining DNA extraction followed Swarts et al. (Swarts et al., 2017).

Library construction, amplification, and sequence

DNA extracts were converted to Illumina-compatible paired-end DNA libraries in a clean setting using by SMARTer ThurPLEX DNA-seq from Takara Bio and amplified with

7 PCR cycles. The generated libraries were cleaned to remove remaining adaptors with AxyPrep Mag beads from Fisher Scientific and ran on a 2100 Bioanalyzer (Agilent Technologies) to check for quality. The libraries were indexed to enable pooling on the NovaSeq SP at University of Illinois.

Geographic location map

The geographical coordinates of each sample were obtained from previously published studies, as indicated (Table AA 4-1). Utilizing latitude and longitude information, the samples were plotted on a geographic map, resulting in the successful mapping of 140 samples, excluding 17 *Zea mays* ssp. *parviglumis* and 1 *Tripsacum* samples that lacked location information. The geographic data visualization was carried out in R using the ggplot2, sf, leaflet, rnaturalearth, rnaturalearthdata, ggspatial, and rnaturalearthhires packages.

Genome analysis

Data for ancient and modern samples (maize landraces, *Zea mays* ssp. *parviglumis*, *Zea mays* ssp. *mexicana* and *Tripsacum*) were obtained from previously published work (Kistler et al., 2018; Kistler et al., 2020; Vallebueno-Estrada et al., 2016) and HapMap2 by downloading sequencing from Sequence Read Archive (SRA) (https://www.ncbi.nlm.nih.gov/sra/) (Chia et al., 2012). To have concentrated and accurate results, we downloaded the raw sequences (SRA formal) for each sample and transformed them into fastq raw read format using fastq-dump. For the derived aDNA maize samples, raw reads were removed adapters and paired reads were merged by AdapterRemoval (Lindgreen, 2012), the first 5nt of both ends of reads were cut off, and mapped to B73 v5 reference genome by BWA v0.7.17 with disabled seed -l 1024 and q

20 based on the recommended parameter (Schubert et al., 2012) to improve ancient DNA mapping. Duplicated reads were removed by GATK tools Picard (https://gatk.broadinstitute.org/hc/en-us). We combined the six aBM sequence samples into one sample called combbam_aBM. This was achieved by mapping each aBM sequence sample to the genome and subsequently removing the duplicate reads. Finally, the resulting duplicate-removed files were recombined into a single read alignment (Table AA 4-2). The reason we merged them into one sample is that PCA does not show clear evidence of two distinct clusters, and we needed higher coverage for the analysis (Figure AA 4-1c; Table AA 4-3). This observation suggests that the two corn kernels originated from the same individual plant due to their identical genotypes. MapDamage v.2.2.1 (Jonsson et al., 2013) was used to verify cytosine deamination profiles consistent with authentically ancient DNA. The coverage of the assembled genome sequence was measured using Qualimap v2.2.1 (Okonechnikov et al., 2016).

For modern maize, raw reads were removed adapters and were trimmed using Trimmomatic(Bolger et al., 2014); parameters were set as follows: LEADING = 20, TRAILING = 20 SLIDINGWINDOW:4:20 MINLEN =25. Trimmed reads were mapped to the B73 v5 reference genome using BWA. Duplicated reads were removed by GATK tools Picard (https://gatk.broadinstitute.org/hc/en-us).

Population genetic and phylogenetic analyses

We created a genotype file using publicly available sequences for a total of 15 - archaeological and 87 - modern maize, 17 - modern samples of *Zea mays* ssp. *parviglumis*, 37 - modern samples of *Zea mays* ssp. *mexicana*, and 1-*Tripsacum* genotypes. The GATK pipeline was used to call single-nucleotide polymorphisms (SNPs)

from all sequencing data using the maize reference genome B73 (v5). HaplotypeCaller option was used to call SNP and only SNP VCF files contain genotypes that required a minimal depth-of coverage of 2. These were used to merge all identified SNPs from comparisons of all sequences as a raw final VCF file (https://gatk.broadinstitute.org/hc/en-us) using CombineGVCFs. Indels were removed and the remaining SNP was filtered by MQ < 30, QD<2.0, FS > 60.0, SOR > 3.0, MQRankSum < -12.5, ReadPosRankSum < -8.0 by following the GATK pipeline due to its higher positive calling rate and sensitivity for low-coverage sequencing data(Yu & Sun, 2013). SNPs for chromosomes 1-10 were used for downstream analysis.

Next, the SNP VCF file was filtered using plink v1.9(Purcell et al., 2007) with settings established (--indep-pairwise 50 5 0.2 --maf 0.02 --geno 0.5) to only keep SNPs with a minor allele frequency ≥0.02 and at least half of all samples called according the parameter used in previous research (Kistler et al., 2020). This approach led to the identification of 16,081,009 SNPs, out of a total 169,926,437 SNPs (DP >/= 2). A Maximum Likelihood (ML) tree was constructed with TreeMix (Pickrell & Pritchard, 2012) between different maize groups. ML phylogenies tree was inferred and visualized by FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). The final genotype classifications were assigned and colored according to location information for each sample. The same SNP also was used to calculate principal component analysis (PCA) with plink v1.9 (Purcell et al., 2007). The PCA data was plotted in R using the ggplot2 package (Wickham, 2009).

7,603 SNPs existing in both aBM and all modern maize, out of a total of 16,081,009 SNPs, were used to build the phylogenetic tree. A Maximum Likelihood (ML) was built by

SNPhylo (Lee et al., 2014) with parameter: -l 0.5 -m 0.02 -M 0.5 -o *Tripsacum* -b 1000; A Neighbor-Joining Tree was built by VCF2PopTree (Subramanian et al., 2019) using a Number of Difference model.

Population structure analyses

To determine the level of genetic differentiation among different populations of maize, ADMIXTURE (Alexander & Lange, 2011) graph was used for population structure analysis by implementing a block relaxation approach (Alexander & Lange, 2011; Alexander et al., 2009). The independent replications were shown for each of K (K=1-10, K = number of genetic clusters), and echo K was visualized by pong software (Behr et al., 2016). The selected optimal of K = 3 was determined by the minimum CV error. To determine the patterns of population splits and mixtures in the history of all populations, gene flow was determined by Treemix (Pickrell & Pritchard, 2012) on a total of 69,721 nucleotides, which were not missing more than 10% of all maize samples, the optimal edge was determined by OptM (Fitak, 2021) with K=1000, and edge = 1 was shown as the optimal option.

The f3 statistic (Durand et al., 2011; Green et al., 2010) with the admixr package (Petr et al., 2019) as part of the software package of ADMIXTURE (Alexander & Lange, 2011) was used to detect which maize population shows evidence of ancient BM admixture. The qpAdm(Harney et al., 2021; Patterson et al., 2012) method was used to detect the proportions of ancestry coming from a set of source populations for aBM.

Geographical location prediction

For prediction the possible geographical location of aBM, a deep neural network method (Battey et al., 2020) was applied using only ancient maize samples because all

the ancient maize were clustered in the same group. In total, 9,946,141 SNPs out of a total of 16,081,009 SNPs that exclude SNPs that are 100% missing in all ancient samples, were used to predict the possible geographical location of aBM.

Identification of selection sweeps

For detecting selection on aBM compared to other different ancient maize groups, to detect genomic signatures of recent selective sweeps, the cross-population extended haplotype homozygosity (XP-EHH) (Sabeti et al., 2007) method based on population comparison was applied using 16,081,009 SNPs. We compared the aBM to the other different groups, which are ancient Mexico Tehua162, ancient Mexico SM310 (SM3, SM10), ancient Peru highland (Z64), ancient Argentina highland (Z66), ancient Chile coastal (Arica4, Arica 5), ancient Brazil Lowland (Z2, Z6), ancient Peru Lowland (Z61, Z65), and ancient Honduras (EG84, EG85, and EG 90), and only SNP existing in at least two samples with --indep-pairwise 50 5 0.5 in ancient maize or modern maize as queries to identify the selection sweeps. The reason that we treated Tehua162 and SM3,10 as two different groups is because their locations were found to be different. We did not compare aBM with Z67 because Z67 is only ~100 years old, which is much younger than the age of aBM. A commonly used threshold - $\log_{10}(p) = 6$ was set as the cutoff threshold for identifying the signal of significant selection. The regions with -log10(p) >= 6 were designated as candidate sweeps and the candidate genes were identified within 5 Kb of the selected candidate genomic regions as suggested in previous research(Brodie et al., 2016). For detecting genomic signatures of recent selective sweeps on modern maize compared to ancient maize groups, SNPs that exist in at least two ancient samples and modern samples were used, we applied the same XP-EHH method as above. We also

applied a cross-population composite likelihood ratio (XP-CLR)(Chen et al., 2010) calculated for 5KB windows using the romberg integral calculation method across the genome between ancient maize and modern maize. The selection sweeps were scanned for selection on the genotype data from modern maize with ancient maize as a reference with a sliding size of the window of 0.005 units of Morgan with the maximum number of SNPs in each window set to 500 and the correlation levels set to 0.95. The regions with XP-CLR values in the top 1% of distributions were designated as candidate sweeps and the candidate genes of the top 1% peak were identified within 5kb of selected candidate genomic regions as used in previous research (Pavlidis et al., 2012). MaizeMine from MaizeGDB (Schnable et al., 2009) was used to conduct gene enrichment with P-value < 0.05, with P calculated by using a hypergeometric distribution (Berkopec, 2007).

Fixation index (Fst) analysis

Flowering, starch, cold, drought, light and heat related genes were found using Go Term (Data Set AA 4-5). Target SNPs on the above genes were obtained from 16,081,009 and Fst between different groups was calculate by using vcfR (Knaus & Grunwald, 2017), adegenet (Jombart & Ahmed, 2011) and hierfstat (de Meeus & Goudet, 2007) packages in R. The classification of maize as highland or lowland was based on the elevation above sea level. Lowland maize refers to elevations lower than 1,000 meters, while highland maize refers to elevations higher than 1,000 meters. This classification was only applied to maize populations that did not have a clearly defined classification from previous research.

Data availability

The authors declare that all data supporting the findings of this study are included in the manuscript and its supplementary files or are available from the corresponding author upon request. Source data are provided with this paper. The genome sequence data of ancient Bolivian Maize has been deposited in the NCBI database under BioProject accession PRJNA886637. Custom scripts and single nucleotide polymorphism (SNP) calls are available on Dryad (DOI: https://doi.org/10.5061/dryad.w6m905qtd).

CHAPTER 5: Conclusion and future directions

How do plants respond to biotic and abiotic stress? What mechanism are used to respond to stress, as well as what kind of key traits were selected in ancient plants to aid plant evolution into the modern plants that we see today? All of these are important, and largely unanswered, questions that remain in the broad field of plant biology area. This dissertation used computational biology approaches to help uncover answers to these questions.

Chapter 2: Contrasting transcriptional responses to *Fusarium virguliforme* colonization in symptomatic and asymptomatic hosts

Conclusion and contribution: In Chapter 2, I identified the reason "why *Fusarium virguliforme* is asymptomatic on monocot host and a symptomatic eudicot host". In this Chapter, we revealed that 1) maize has earlier defense response to fungal in both orthologous and non-orthologous gene levels; 2) maize root defense marker induced in SDS-resistant soybean cultivars defend the *Fusarium virguliforme*; 3) Constitutive expression of maize genes within orthologous defense has rapid upregulation at Day 0; 4) The defense expression patterns of orthologous transcription factors in maize which control senescence has down-regulated in maize compared to up-regulated in soybean; 5) Biological processes control the senescence were not changed at earlier time points compared to soybean. In summary, the trick to surviving the disease might be a matter of timing when to mount a defense response between maize and soybean. This Chapter contributed that the computational biology and big data approaches helped yield a big picture view of disease progression and prevention and fill the gap of studying between broad hosts.

Future directions: Root senescence was observed in the experiment, but I am not sure it is the reason or the result of fungal infection. I would explore the molecular-genetic and biochemical networks associated with pathogen-induced root senescence.

Chapter 3: NDR1 and the *Arabidopsis* plasma membrane ATPase AHA5 are required for processes that converge on drought tolerance and immunity

Conclusion and contribution: In Chapter 3, I identified that NDR1 plays an important role in the response to low water availability and drought stress by suppressing ABA biosynthesis in response to drought stress, requiring for the convergence of signaling between pathogen perception and stomata response, working in signaling processes required for guard cell gating, withing in stomatal osmoregulation immunity through its interaction with *AHA5*, and control the stomata aperture to respond to drought stress. Identified the new role of NDR1 in responding to drought stress. This chapter contributed to the indentation of the new role of NDR1 in responding to drought stress.

Future directions: Since NDR1 is a critical component of the plant immune system, first discovered in 1995, increasing evidence has shown that NDR1 may have multiple functions in plant immunity and response to stress. I would explore the multiple roles of NDR1 in biotic and abiotic stress.

Chapter 4: Archaeological Bolivian maize genomes suggest diversity is associated with Inca cultural expansion and environmental variation in South America

Conclusion and contribution: In Chapter 4, I identified the closest relative maize to archaeological Bolivian maize (aBM) is ancient Peruvian maize and the geographic location of aBM may come from also nearby ancient Peruvian maize. Besides, I also identified that key traits, including shorter growing seasons, greater stress-resistance, and sweeter taste, are selected in ancient maize by the Inca culture. Finally, I also identified

that the stronger heat stress response character was selected from modern maize and was compared to ancient maize in South America. This chapter contributed to using a combination of biology and culture to identify the key trait selection under the Incan cultural influence in South America.

Future directions: The Incan culture has an influence on maize diversity in South America, did it have an influence on the certain group of people who lived in the territory that the Inca empire conducted? If it did, what kind of influence it have and what's the potential contribution to population diversity in South America? To answer these questions, I will explore the story behind the mummy and the culture she presented.

Chapter 6: The evolution and function prediction of the *ADF* gene family in *Arabidopsis thaliana*

Conclusion and contribution: In the Supplement Chapter, I identified that ADF most likely originated in the last common ancestor of animals, fungi, and plants, and the plant ADF family is unique, and likely emerged in the last common ancestor of eukaryotes by doing the evolutionary analysis of ADF among whole representative eukaryotes species. Five pairs of segmental duplication genes of the *AtADF* gene were detected. AtADF5 and AtADF9 are from the same clade, but their expression levels are very different indicating that duplication genes can lose the function of ancestral genes. This chapter contributed to using the *ADF* gene family as a case to explore the study of immunity genes.

Unfinished questions and future directions: To answer the question of the evolution of the *AtADF* gene expression pattern, a total of 24,123 RNA-Seq samples of *Arabidopsis* with the Col-0 genotype collected from GEO will be used. To answer the novel or function loss of ADF, more single-cell, protein, and metadata will be used to find

the narrow down the specific function of each gene. For future direction, only less than 50% of protein function remains are known in *Arabidopsis*, therefore, machine learning (ML) will be used to predict protentional immunity genes from the half unknown function genes.

Other aspects

Besides research, during graduate school, I served as the representative of 5+ graduate students in the Genetics and Genome Sciences program in Fall 2023. I participated in the Genomes to Fields (G2F) Genotype by Environment Prediction Competition in 2023 with team members from the AIMazie group, leading the writing of genomics and computational sections, and generated preliminary data for the proposal “The Temperature-Immunity Nexus: Activation of Immunity by Low Temperature”, co-creating an Apple App called BU (Be You) as part of The Apple Five group at the Apple Developer Academy MSU Foundation Program 2023. Additionally, I volunteered for outreach activities on both Girls Math and Science Day 2023 and outreach in Donley Elementary School and mentored an undergraduate student in the Graduate Women in Science (GWIS) undergraduate mentoring program in Fall 2023.

CHAPTER 6: The evolution and function prediction of the *ADF* gene family in *Arabidopsis thaliana*

The chapter is the supplementary chapter

Chapter Abstract

The immune system functions as a danger perception and signaling transducer for plants, and the actin cytoskeleton is one of the major nodes of immune signaling cascades. The ACTIN DEPOLYMERIZING FACTORS (ADFs) are among the most widely conserved family of actin-binding proteins in all eukaryotes, where they are mainly involved in the recycling of G-actin by filament depolymerization during rapid dynamic actin cytoskeleton turnover. ADF is essential for plants as it is one of the actin-binding proteins (ABP) among approximately 100 ABPs in the plant cytoskeleton. ADF regulates the plant's immune system by playing a fundamental role in the transition between monomeric (G-actin) and filamentous (F-actin) to speed up the movement of the cytoskeleton, which contributes to the growth, development, and responses to various stresses of the organisms. The actin dynamic which is controlled by ADFs has the essential function of keeping cell motility, development, differentiation, and signal transduction cytokinesis functioning properly, while maintaining the structure of the eukaryotic cell surface. Over millions of years of evolution, the *ADF* gene family has expanded in plants. The significance of the expanded number of *ADF* members in plants is hypothesized to underpin functional specificity based on temporal, tissue, and developmental patterns of expression. Recent research in our lab showed that ADF4 interacts with Actin 2/7 to enter the nucleus and interact with WRKY transcription factors 29 (WRKY29), resulting in improved immunity in *Arabidopsis* by regulating the expression of downstream genes. In addition, the functionally analogous cofilins (COFs) in humans play a key role as a moonlight gene (a gene with multiple functions genes) in Alzheimer's and Parkinson's diseases, and DNA damage repair in cancer. In plants, the known

functions of ADFs are mainly related to filament depolymerization, with only ADF4 having been reported to have a function in improving innate immune response in *Arabidopsis*. However, the functions of other ADFs in the immune system remain largely unknown. Indeed, there is an overall lack of knowledge concerning ADFs' multiple functions, despite their value as an agricultural species model. Since evolution is the primary power of developing new functions for genes by using different duplication mechanisms, determining the *ADF* gene family evolution process and duplication mechanism will enhance the understanding of the *ADF* gene family. At the same time, we believe that characterizing the function of individual *ADF* genes will enhance the understanding of the plant immune system. This is motivated by a situation in humans, where knowledge of functions of individual COF genes has aided in the discovery of new ways to improve the immune system. This chapter outlines a research plan to determine the evolution process of the *ADF* gene family, predict potential functions of individual *ADF* genes, verify other biological functions of ADFs in different tissues of *Arabidopsis* through experimental approaches. This study may provide some foundational knowledge for improving the plant immune system in response to abiotic and biotic stresses.

Introduction

Actin cytoskeleton plays a central role in all eukaryotes and regulates various cellular functions in multiple cellular processes, including plant growth, response to environmental changes, cell shape, division and differentiation, facilitating cytoplasmic streaming, organelle movement, regulation of polarity and stomatal movement (Day et al., 2011; Henty-Ridilla et al., 2013; Hussey et al., 2006; Pollard & Cooper, 2009; Staiger, 2000; Staiger & Blanchoin, 2006; Szymanski & Cosgrove, 2009; Zhao et al., 2011). The

actin cytoskeleton is not only highly organized, but also highly dynamic within plant cells, rapid reorganization and turnover are achieved through the change between filamentous (F) and globular (G), which is precisely regulated by several actin-binding proteins (ABPs) (Staiger & Blanchoin, 2006). Numerous actin-binding proteins regulate actin organization and dynamics, and thus various biological functions (Henty-Ridilla et al., 2013; Higaki et al., 2007; Huang et al., 2015; Sun et al., 2013; van Gisbergen & Bezanilla, 2013). ACTIN DEPOLYMERIZING FACTORS (ADFs) are an ancient protein that is conserved. They are an important ABP in eukaryotes regardless of their ability to depolymerize F-actin, yielding products with ends that can serve as sites for new filament initiation. ADF functions to keep the cellular balance of F– and G-actin and increases the dynamics of actin polymerization and depolymerization (Andrianantoandro & Pollard, 2006; Pavlov et al., 2007).

Beyond their major function as a depolymerization of actin microfilaments, plant ADFs are involved in various biological processes such as hypocotyl elongation, innate immunity, nematode infection, polar growth, stomatal movement and response to biotic and abiotic stresses (Clement et al., 2009; Dong et al., 2001; Henty-Ridilla et al., 2014; Inada et al., 2016; Qian et al., 2019; Tian et al., 2009; Zhao et al., 2016; Zheng et al., 2013; Zhu et al., 2017). The *ADF* gene family in the *Arabidopsis thaliana* genome consists of 11 functional genes, which can be divided into four subclasses (subclasses I–IV)(MacRobbie & Kurup, 2007). The differentiation of tissue has been reported for ADFs. Subclass I includes *AtADF1*, *AtADF2, AtAD1*, *and AtADF4*, which are expressed throughout plants at relatively high levels. However, *AtADF7* and *AtADF10* in subclass II are only expressed in flowers, particularly in pollens. *AtADF8* and *AtADF11* in subclass

IIb are only specifically expressed in root epidermal cells. Subclass III and subclass IV are expressed in a wide variety of tissues, but the expression of *AtADF5* and *AtADF9* are not seen in root apical meristem (Ruzicka et al., 2007). Results from Dong et al. also confirmed the distinct expression pattern of *AtADF1*, *AtADF5,* and *AtADF9*, and the expression of *AtADF5* is clearly restricted.

Physiological function difference has also been reported for ADFs. Transient expression of *AtADF1*, *AtADF5*, *AtADF6*, *AtADF7*, and *AtADF10* inhibits the resistance of barley against powdery mildew fungus (Miklis et al., 2007). *AtADF1* plays an important role in plant thermal adaptation by blocking the high-temperature-induced stability of actin filaments (L. Wang et al., 2023). *AtADF3* has an important function in the phloem (Mondal et al., 2018). *AtADF4* plays a role in controlling gene-for-gene resistance activation and MAPK-signaling via the coordinated regulation of actin cytoskeletal dynamics, *R*-gene transcription, and controlling the morphological change of plants (Peng & Huang, 2006; Porter et al., 2012). What is more, *AtADF4* also contributes to the stochastic dynamic turnover of actin filaments in the cortical array of epidermal cells (Henty et al., 2011). *AtADF5* plays a role in drought stress by participating in regulating stomatal closure and acting as a downstream target gene of *CBFs* in response to low-temperature stress (Qian et al., 2019; Zhang et al., 2021). Overexpression of *AtADF6* inhibits the transport resistant protein RPW8.2-YFP for powdery mildew fungus (Wang et al., 2009). *AtADF7* promotes the turnover of longitudinal actin cables by severing actin filaments in pollen tubes to promote normal pollen tube growth and inhibits VLN1 to regulate F-actin dynamics in root hair formation in response to osmotic stress (Bi et al., 2022; Zheng et al., 2013). Purified AtADF9 shows strong F-actin bundling activity, whereas the loss of *AtADF9* expression

causes early flowering phenotypes (Burgos-Rivera et al., 2008; Tholl et al., 2011). *AtADF10* regulates vesicle trafficking and pollen tube growth by shaping the organization of apical actin filaments (Jiang et al., 2017). Subclass I ADFs play a role in the regulation of nuclear organization and gene expression and are a novel regulator of endoreplication (Inada et al., 2021; Matsumoto et al., 2023). Although both *AtADF7* and *AtADF10* are expressed in pollen, which belongs to Subclass II, their roles are distinct due to their differences in expression and localization (Bou Daher et al., 2011; Daher & Geitmann, 2012).

Recent experimental research has been contributing to the study of ADF thanks to the development and application of next-generation sequencing technology. There are however no reports regarding the *ADF* gene expression profiles in various tissues, in response to various stress, or their potential biological function by using large public data. In this study, we analyzed the *ADF* gene structure and the encoded conserved motifs and determined the genomic locations and duplication events of the *ADF* genes. Furthermore, ADFs in selective reprehensive modern species were identified and used to construct a phylogenetic tree. To further investigate the protentional function of ADF in *Arabidopsis thaliana*, we collected the expression profiles of 28,633 RNA-Seq samples and one leaf single-cell samples from NCBI, analyzed the expression profiles of these 11 functional genes in different tissues, as well as in response to different stress. What is more, we also analyzed the protein evolution of ADFs to explore the protentional function. The results of this study will promote our understanding of the *ADF* gene family and provide the basis for future study of ADFs in plants.

Results

Evolutionary relationships among ADF members in representative eukaryote species

To investigate the evolutionary relationships and characteristics of the *ADF* genes with other homologous genes, 38 eukaryotes species and 1 bacteria species as an outgroup were used to construct a species tree in Figure 6-1a, and the longest amino acid sequences of ADF protein from all species were used to construct an ADF gene family phylogenetic tree in Figure 6-1b.

The phylogenetic tree showed *ADF* most likely originated in the last common ancestor of animals, fungi, and plants. Two main ADF clades were classified as plant-specific and not plant-specific. The clade of not plant-specific integrates yeast *ADF* genes with invertebrates, protists, metazoan, fungi, and vertebrates, which indicates that this is most likely the ancestral clade of the *ADF* gene family. Plant-specific clade only integrates plant-specific. Taken together, these findings suggest that plant-specific clade emerged in the last common ancestor of eukaryotes. The ADF in plants and other kingdoms are completely separated.

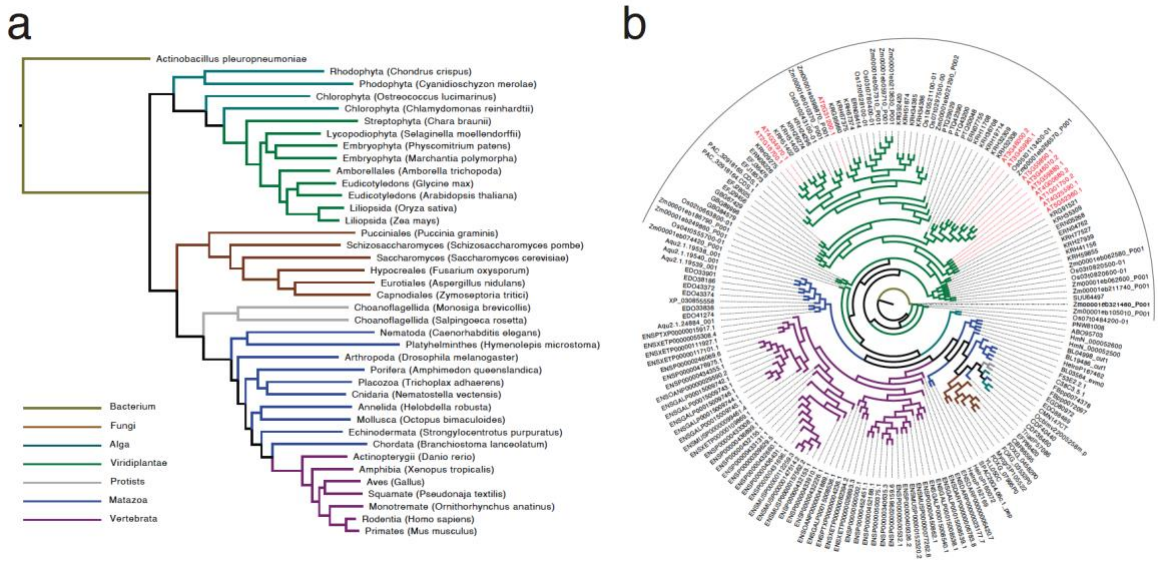

Figure 6-1: Species evolution tree and phylogenetic tree of ADFs. (a) Evolution relationship among 38 eukaryote with 1 bacterium as outgroup. (b) Phylogenetic tree of ADF family members in the 38 eukaryotes species and most similar ADF protein in 1 bacterium by using the longest isoform. ADF proteins in *Actinobacillus pleuropneumoniae, Amborella trichopoda, Amphimedon queenslandica, Arabidopsis thaliana, Aspergillus nidulans, Branchiostoma lanceolatum, Caenorhabditis elegans, Chara braunii, Chlamydomonas reinhardtii, Chondrus crispus, Cyanidioschyzon merolae, Danio rerio, Drosophila melanogaster, Fusarium oxysporum, Gallus, Glycine max, Helobdella robusta, Homo sapiens, Hymenolepis microstoma, Marchantia polymorpha, Monosiga brevicollis, Mus musculus, Nematostella vectensis, Octopus bimaculoides, Ornithorhynchus anatinus, Oryza sativa, Ostreococcus lucimarinus, Physcomitrium patens, Pseudonaja textilis, Puccinia graminis, Saccharomyces cerevisiae, Salpingoeca rosetta, Schizosaccharomyces pombe, Selaginella moellendorffii, Strongylocentrotus purpuratus, Trichoplax adhaerens, Xenopus tropicalis, Zea mays* and *Zymoseptoria tritici* are prefixed by SU, ER, Aq, AT, CB, BL, C3 or F5, GB, PN, CD, CM, ENSD, FB, FO, ENSG, KR, He, ENSP0, Hm, PT, EDQ, ENSM, EDO, Oc, ENSO, Os, AB, PA, ENSPT, EFP, YL, EG, SP, EFJ, XP, Tr, ENSX, Zm, My. ADF in *Arabidopsis thaliana* are marked in red. The grey half-circle line indicates the plant-specific clade. ADF proteins from *Arabidopsis* thaliana are marked as red. The different color lines represent different kingdoms.

Conserved motifs analysis of ADF family members in eight representative species

A species tree of eight representative species was contracted by OrthoFinder in Figure 6-2, and based on this tree, we performed conserved motifs of *ADF* gene family

members. It showed all plant ADF has the unique motif 7 and the non-functional *ADF* gene lacks motif 4 in *Arabidopsis thaliana*. This may partly explain the reason why all the ADF in plants were in the plant-specific clade and why putative ADF in *Arabidopsis thaliana* is without any function. To investigate the *ADF* gene family members in plants, ADFs in *Arabidopsis thaliana* will be used to study evolution history and function prediction.

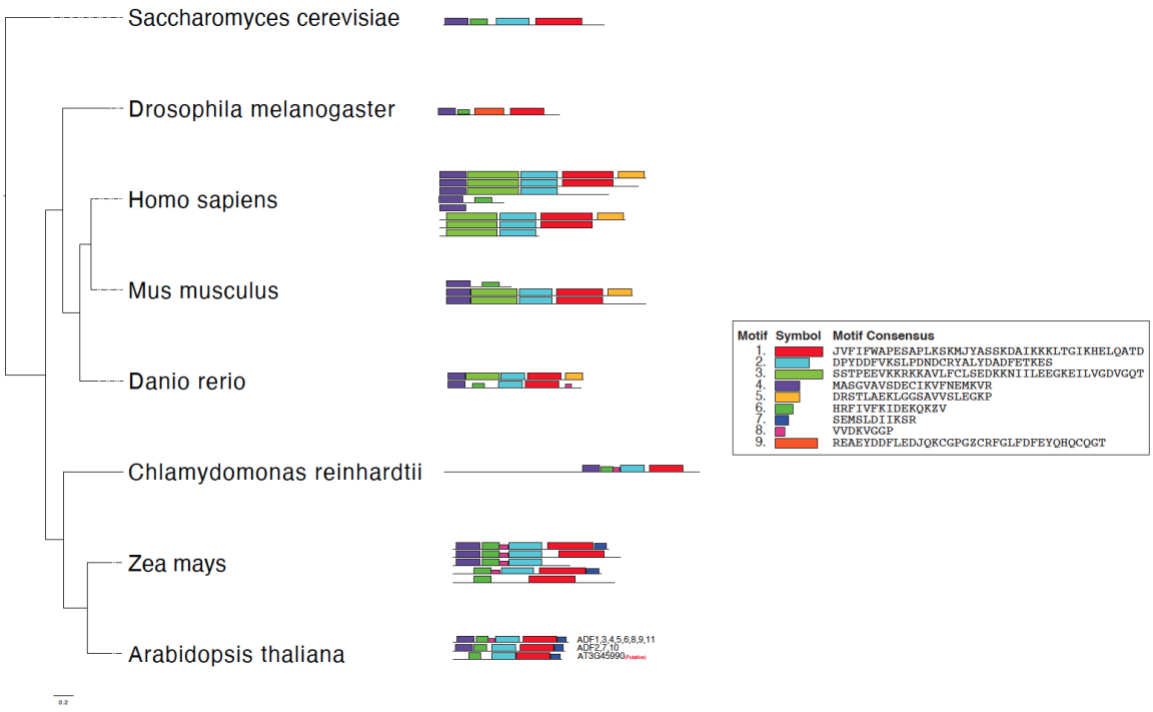

Figure 6-2: Phylogenetic relationship and motif composition of ADFs among eight representative species. (A) Phylogenetic relationship of ADFs. (B) Distribution of predicted conserved motifs type of ADFs.

Gene structure of *AtADF* genes in *Arabidopsis thaliana*

Phylogenetic analysis of all ADF classifications indicates these 12 genes could be divided into four clusters in Figure 6-3a. The gene structure of *ADF* genes in *Arabidopsis* was analyzed by aligning the CDS region and the genome sequence from the TAIR

(https://www.arabidopsis.org/). As shown in previous research (Inada, 2018), all the *ADF* genes have three exons, indicating that all of these *ADF* genes' structures are relatively simple.

Cis-acting elements are important for gene transcription. In our study, a 1000 bp sequence upstream of the transcription start site of transcripts was extracted as promoters to predict cis-acting elements in 11 *AtADF* (*AtADF1* - *AtADF11*) in Figure 6-3b. A total number of 22 kinds of cis-acting elements were detected and classified into five categories, which are cycle (2), development (1), hormone (1), stress (16), and transcription (3) for a total of 693 predicted cis-acting elements in *Arabidopsis thaliana*. All *AtADF* genes contain TATA-box and MYB. The number of the TATA-box elements (293) was the highest. The number of MYB was the highest among stress relative elements kinds. The number of development elements was the lowest (Table AA 6-1).

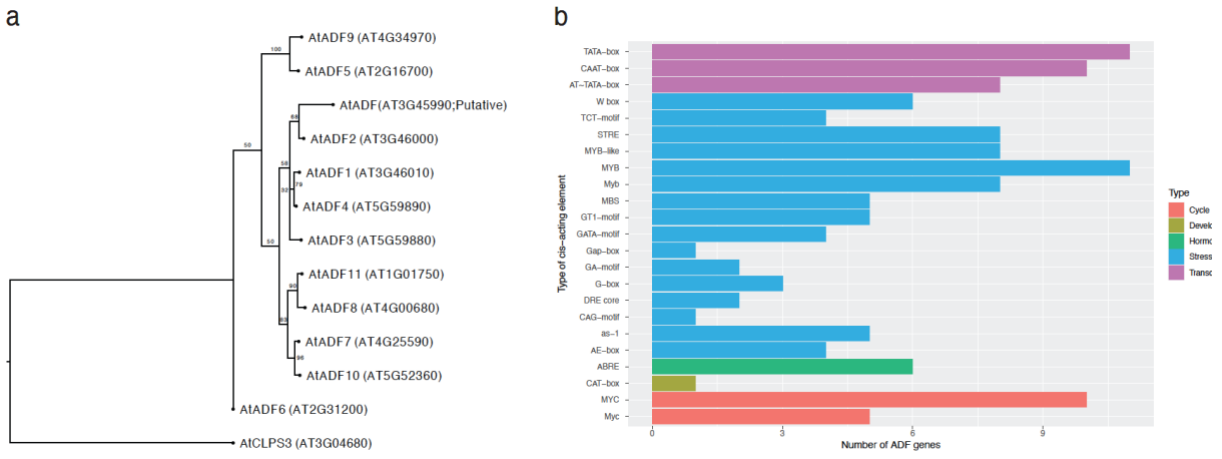

Figure 6-3: Phylogenetic relationship and prediction and analysis of cis-acting elements in the AtADF gene family promoter region. (A) The Maximum Likelihood phylogenetic tree of ADFs was constructed using the longest amino acid sequence from the longest isoform with CLPS3 as outgroup. (B) Classification cis-acting elements. Twenty-two kinds of cis-acting elements were divided into five categories. Different colors represent different kinds of categories.

Analysis of the physicochemical properties of proteins of the *AtADF* gene family

The basic information about the *AtADF* gene family and the proteins is shown in Table 6-1. The longest genome sequence is *ADF* (AT3G45990; Putative) at 3359bp and the shortest is 945bp (*ADF11*). The average length is 1490 bp for all *AtADFs*.

We also analyzed the physicochemical properties of the proteins, which are amino acid length, molecular weight (MW), theoretical pI, Aliphatic index, and grand average of hydropathy (GRAVY) (Table 6-1). The amino acid length of each AtADF is very similar from 133 to 150. The molecular weight varied from 15.820 kDa (AtADF7) to 17.942 kDa (ADF2) and the GRAVY of all studied AtADF were below zero. The maximum aliphatic index value was 83.53 (AtADF4), and the minimum value was 70.07 (AtADF6). The results of hydrophilicity and hydrophobicity analysis indicated that all AtADF family proteins are hydrophilic proteins. The results of subcellular localization of the proteins analyzed using online software and information obtained from TAIR showed that all AtADF family proteins are cytoplasmic proteins.

Table 6-1: *AtADF* gene family protein properties table.

Name	Chr	Position	Longe-st isofor-m	Protein Characteristics					Subcellular Location
				Amino Acids (aa)	MW (KDa)	Theor-etical pI	Aliphatic index	GR-AVY	
AtADF1 (AT3G46010)	Chr3	16909217-16910845	AT3G46010.2	150	17.339	7.62	80.67	-0.289	Cytoplasm; Chloroplast; vascular tissues of all organs
AtADF2 (AT3G46000)	Chr3	16907334-16909119	AT3G46000.2	156	17.942	6.59	74.42	-0.362	Cytoplasm;
AtADF3 (AT5G59880)	Chr5	24120164-24121949	AT5G59880.1	139	15.922	5.94	75.83	-0.415	Cytoplasm
AtADF4 (AT5G59890)	Chr5	24122348-24123912	AT5G59890.1	139	16.034	6.14	83.53	-0.313	Cytoplasm
AtADF5 (AT2G16700)	Chr2	7244500-7246069	AT2G16700.1	143	16.467	6.83	70.91	-0.348	Cytoplasm; Expressed exclusively in root tip meristem.
AtADF6 (AT2G31200)	Chr2	13293953-13295410	AT2G31200.1	146	16.707	5.46	70.07	-0.574	Cytoplasm; Expressed in vascular tissues of all organs.
AtADF7 (AT4G25590)	Chr4	13058936-13060250	AT4G25590.1	137	15.820	5.08	73.36	-0.389	Nucleus; Specifically expressed in pollen.
AtADF8 (AT4G00680)	Chr4	279596-281072	AT4G00680.2	150	17.353	5.12	71.47	-0.461	Cytoplasm; Expressed in the root trichoblast cells and developed root hairs.
AtADF9 (AT4G34970)	Chr4	16653704-16654955	AT4G34970.1	141	16.369	7.66	60.85	-0.551	Cytoplasm
AtADF10 (AT5G52360)	Chr5	21257934-21259408	AT5G52360.1	137	15.884	5.57	76.20	-0.445	Nucleus; Specifically expressed in pollen.
AtADF11 (AT1G01750)	Chr1	275188-276310	AT1G01750.2	145	16.886	5.54	72.62	-0.561	Nucleus
ADF (AT3G45990; Putative)	Chr3	16900778-16904136	AT3G45990.1	133	15.904	7.74	79.92	-0.599	Unknown

Chromosome localization and collinearity analysis of *AtADF* genes family

The relationship between the location of the *ADF* gene family on the chromosomes and the collinearity within the *Arabidopsis thaliana* is shown in Figure 6-3. A total of 12 *AtADF* genes, which are 11 functional genes and 1 putative gene, are partially distributed on the chromosomes of *Arabidopsis thaliana* and are scattered on whole five chromosomes. Only one *ADF* gene was located on chromosome 1 (*AtADF11*), two genes on chromosome 2 (*AtADF5* and *AtADF6*), three genes on chromosome 3 (*AtADF1*, *AtADF2* and *AtADF*-putative), chromosome 4 (*AtADF7*, *AtADF8* and *AtADF9*) and chromosome 5 (*AtADF3*, *AtADF4* and *AtADF10*), respectively.

The replication relationship between genes within a species could be demonstrated by collinearity analysis. One tandem duplication between *ADF2* and ADF putative gene was detected among whole *ADF* gene family members. ADF gene family members have five collinear relationships within species, which are *AtADF3*:*AtADF5*, *AtADF3*:*AtADF9*, *AtADF3*:*AtADF*-putative gene, *AtADF5*:*AtADF9*, and *AtADF7*:*AtADF10*, as shown in Figure 6-2b. The selection pressure on the *ADF* gene duplication events in *Arabidopsis thaliana* was evaluated by calculating the rates of non-synonymous (Ka) and synonymous (Ks) nucleotide substitutions (Table AA 6-2). The average Ka/Ks ratios for the five duplicated pairs were less than 1.00, implying the *ADF* genes evolved under strong purifying selection.

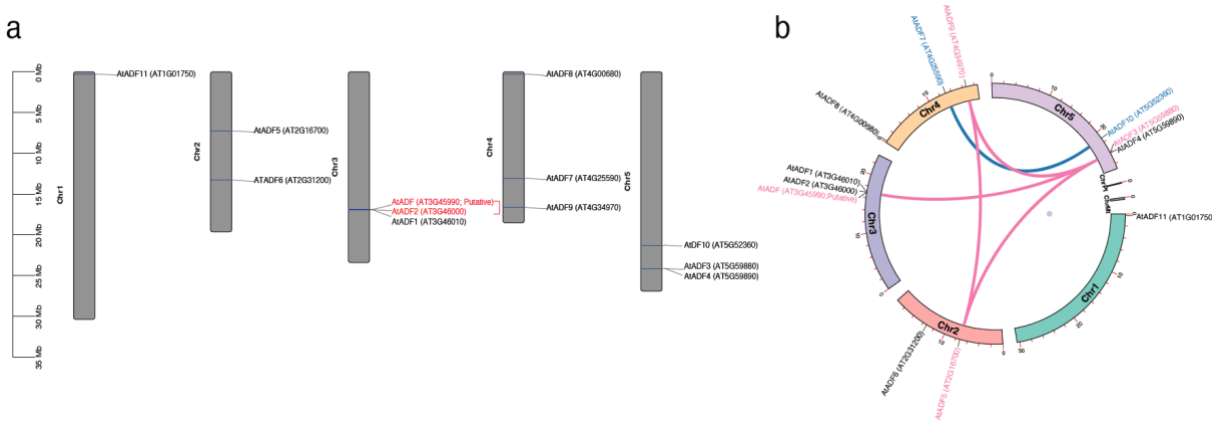

Figure 6-4: Chromosomal location and duplication events of *ADF* in *Arabidopsis thaliana*. (A) Chromosomal location of *ADF* genes in *Arabidopsis* thaliana. Twelve *ADF* genes were distributed on chromosomes 1, 2, 3, 4 and 5. The chromosome numbers are indicated on the left side of each vertical bar. The red line between genes indicates tandem duplication of *ADF* genes. (B) Gene duplication events of *AtADFs*. The different colors represent seven chromosomes. The segmentally duplicated genes are linked by different color lines.

Systematic evolutionary relationships among ADF gene members in monocots and dicots

To determine the collinear relationship of the 12 *ADF* family genes among other species, collinearity analysis was performed using *Glycine max* (dicot), *Zea mays* (monocot), and *Oryza sativa* (monocot) with *ADF* family members from *Arabidopsis thaliana* (dicot). The results of the analysis are shown in Figure 6-3. *Arabidopsis thaliana* (dicot) had 15 collinearities with *Glycine max*, 1 collinearity with *Zea mays*, and 4 collinearities with *Oryza sativa*, respectively. *ADF1*, *ADF2*, *ADF4,* and *ADF8* in *Arabidopsis thaliana* had no collinearity with genes from other species (Table AA 6-3). Among these three species, *ADF* gene family members in *Arabidopsis thaliana* are most closely related to the *ADF* gene family in *Glycine max*. This is because they both belong to the dicot lineage. The collinear relationship analysis of the *ADF* gene family with other

plant species can help provide the foundation of the genetic relationship and gene function of species.

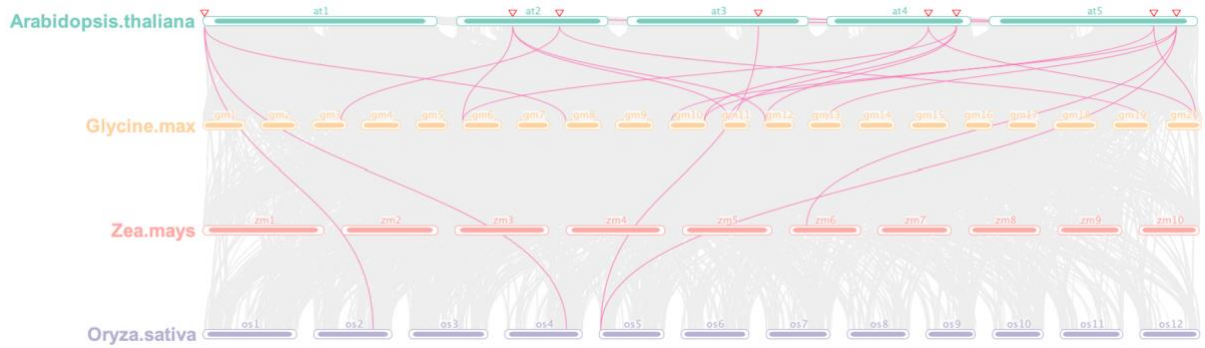

Figure 6-5: Collinearity relationship between *ADF* gene family members in *Arabidopsis thaliana* (dicto), *Glycine max* (dicot), *Zea mays* (monocot) and *Oryza sativa* (monocot). The chromosome number is marked above the chromosome. The chromosomes of *Arabidopsis thaliana*, *Glycine max*, *Zea mays* and *Oryza sativa* are marked with different colors. The collinear relationship between the ADF gene family members among four different species is connected by purple-colored lines. The red triangle represents the location of *ADF* genes on *Arabidopsis* thaliana.

Single-cell expression of *AtADF* genes in leaf tissues of *Arabidopsis thaliana*

The results of single-cell RNA-Seq (scRNA-Seq) are shown in Figure 6-6. Totally nice *AtADF* genes were detected in all leaf tissue except *AtADF7* and *AtADF10*, which is consistent with results from previous research. Both *AtADF7* and *AtADF10* are specifically expressed in the pollen (Bou Daher et al., 2011; Daher & Geitmann, 2012). *AtADF1-6* were significantly expressed in the companion cell, *AtADF1-4* were significantly expressed in the guard cell, *AtADF1-3* were significantly expressed in the epidermis cell, *AtADF1* and *AtADF3* were significantly expressed in the bundle sheath, and the xylem parenchyma cell.

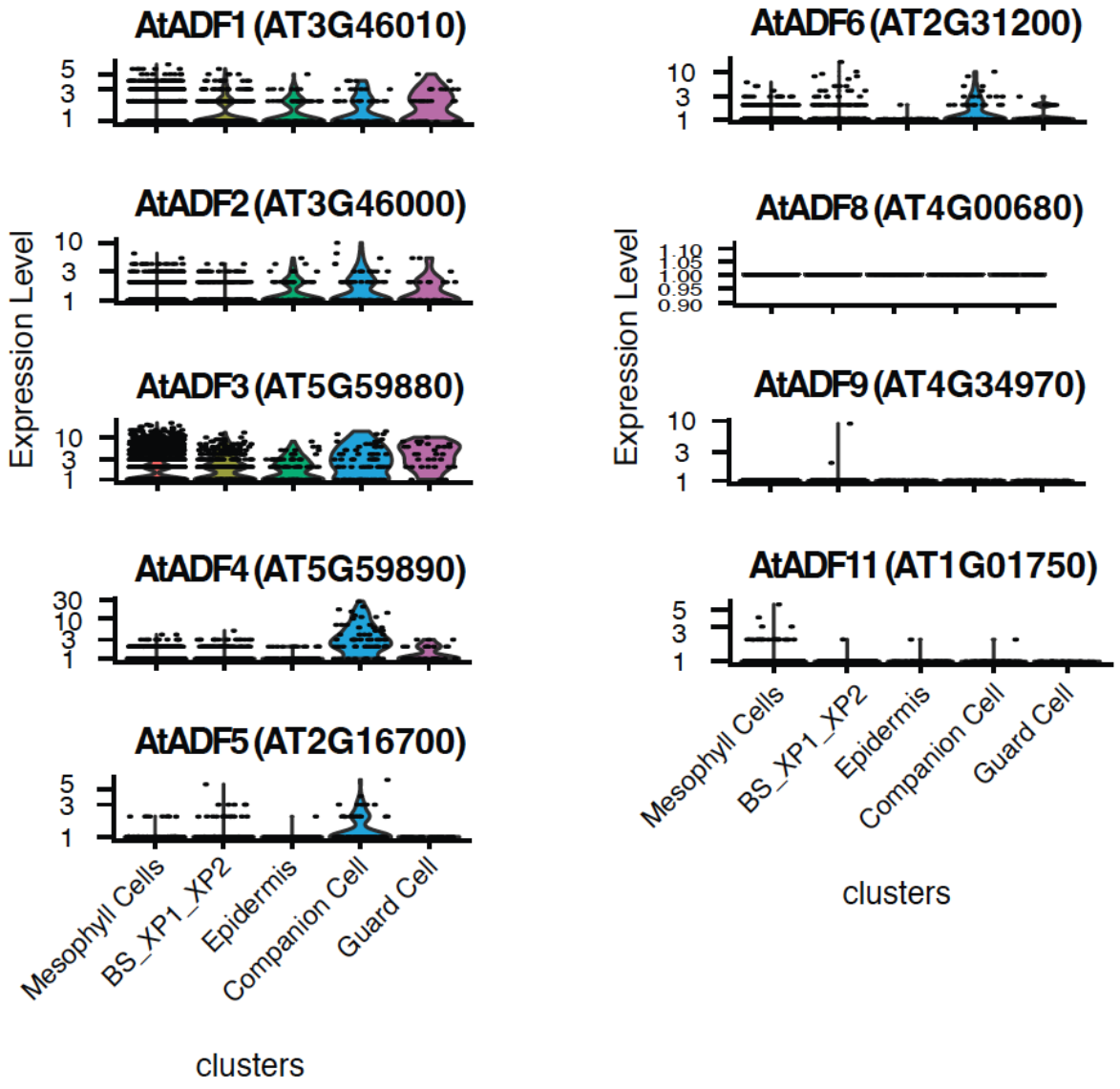

Figure 6-6: Single-cell expression of *AtADF* genes in leaf tissue. Violin plot showing transcript expression. The Y axis is the log scale of the count. The X axis is the cell-type cluster.

Discussion

The *ADF* gene family is conserved and relatively small in higher plants. Our evolutionary analysis of *ADF* among whole representative eukaryote species suggests that *ADF* most likely originated in the last common ancestor of animals, fungi, and plants.

Interestingly, the research described herein supports the hypothesis that the plant ADF family is unique and likely emerged in the last common ancestor of eukaryotes. Indeed, if we track whole genome duplication events, genes in the same gene family come from a common ancestor through duplication. Some of them could become pseudogenes, but some will keep the same function as the common ancestor or evolve various functions (Nan et al., 2017). Five pairs of segmental duplication genes of the *AtADF* gene were detected (Figure 6-4). Among these, *AtADF5* and *AtADF9* are from the same clade (Figure 6-3a), *AtADF9* has lower expression levels in companion cells compared to *AtADF5*, indicating that duplication genes can lose the function of ancestral genes. *Arabidopsis thaliana* is the most important model plant for crop reproductive development. The *ADF* gene family has been reported to play a crucial role in plants' development and response to stresses. Although a few evolutionary function prediction studies have been done for the *ADF* gene family in *Arabidopsis thaliana*, no study of evolution to function prediction of the *AtADF* gene family has been reported. Concurrent with an increased number of computational and data-rich resources with *Arabidopsis*, coupled with advances in computational biology, an informatics-driven prediction of *AtADF* has become less difficult and more powerful for guiding experimental studies, such as the work described herein.

Materials and Methods

Species tree construction

Species tree construction among 39 species. The protein sequence data of selected species, which presents each phylum of eukaryotes, were downloaded from Ensembl and Ensembl Plants database (https://useast.ensembl.org/index.html;

http://plants.ensembl.org/index.html). The species were built using all orthogroups by OrthoFinder (Emms & Kelly, 2015, 2019). The same method as above was used to build the species tree in eight representative species. The representative species were selected based on the genome having been fully annotated and representing significant temporal spacing from common ancestry.

Phylogenetic tree construction ADF

The phylogenetic tree was performed using the maximum likelihood (ML) method in RAxML (Randomized Axelerated Maximum Likelihood) (Stamatakis, 2014), with 1,000 bootstrap replicates. Sequences of ADF proteins from selected species were identified based on the orthogroup genes results using OrthoFinder (Emms & Kelly, 2015, 2019). Only the longest protein sequence from the longest isoform was used to build the phylogenetic tree. The *ADF* gene from Bacterium was found by using *ADF* genes from *Arabidopsis thaliana* to do the Blastp (Altschul et al., 1997) with the E value < 0.05. The above methods were also used in the phylogenetic tree construction of ADFs in eight representative modern species and ADFs in *Arabidopsis thaliana*.

Motif analysis of the ADF gene family in eight representative species

Conserved motifs of ADFs among eight species were predicted using the Multiple Em for Motif Elicitation (MEME) with the following parameters: classic mode, the number of motifs eques 10, the size of motif between 6 wide and 50 wide (inclusive), A 0-order background model generated from the supplied sequences and zero or one occurrence (of a contributing motif site) per sequence.

Gene structure of *AtADF* in *Arabidopsis thaliana*

The genome sequence data of *Arabidopsis* were downloaded from The *Arabidopsis Information* Resource (TAIR) database. Gene structure is displayed by the Gene Structure Display Server (GSDS) (Hu et al., 2015) program using the CDS and corresponding genomic sequences. The final gene structure was displayed using the figure from the previous research (Inada, 2018).

Cis-acting elements of the AtADF gene family in *Arabidopsis thaliana*

A promoter is typically located directly upstream or at the 5′ end of the transcription initiation site, which is a short region of DNA (100–1,000 bp)(Le et al., 2019; Lin et al., 2019). Genes that contain common motifs in their promoter regions have similar expression patterns in low-oxygen response in *Arabidopsis* root (Klok et al., 2002; Vilo et al., 2000). Thus, the 1,000 bp upstream of the transcription start site of all ADF transcripts was extracted as promoters to predict cis-acting elements using PlantCARE (Lescot et al., 2002).

Chromosomal localization, gene duplication and calculating Ka/ks values of *AtADF*

All the *AtADF* genes were mapped to chromosomes based on physical location information from the TAIR database of the genome. Gene duplication in the *Arabidopsis thaliana* genome was analyzed with the Multiple Collinearity Scan toolkit (MCScanX) (Wang et al., 2012), and if an *AtADF* has more than one transcript, only the longest transcript in the annotation was used. The BLASTP (Altschul et al., 1990) was used with -evalue 1e-5 and MCScanX was used with MATCH_SIZE as 2. The calculation of ka and ks substitution of each duplicated AtADF gene was performed using add_kaks_to_synteny.pl from MCScanX.

Physicochemical properties and protein location prediction

The physicochemical properties of the *AtADF*, which are MW (kDa), theoretical pI, instability index, and GRAVY, was predicted by Expasy (https://web.expasy.org/protparam/). The subcellular localization of the AtADF proteins were collected from TAIR, ThaleMine (https://bar.utoronto.ca/thalemine/begin.do), and also predicted by WoLF PSORT(Horton et al., 2007).

Single-cell RNA-sequencing (scRNA-seq) expression pattern of AtADF genes from *Arabidopsis thaliana*

A total of twenty-two single-cell sequence samples with wild type *Arabidopsis thaliana* were collected from previous research, which are two biological replicates in leaf (Kim et al., 2021), five biological replicates from root nuclear (Farmer et al., 2021), seven samples from 5 days root tips tissue (Shahan et al., 2022), three sample from root protoplasts (Ryu et al., 2019). The raw single-cell data was filtered by Seurat (Butler et al., 2018) with cells that have unique feature counts over 2,500 or less than 200 counts. Plotting the transcriptomes in two dimensions using Uniform Manifold Approximation and Projection (McInnes L, 2018). Cluster specific marker genes were identified with Parameters Log FC ≥ 0.25, p_val_adj <0.01, PCT.1 > 0.1, PCT.2 < 0.1 using previous research as a reference (Denyer et al., 2019).

Expression pattern of *AtADF* genes from *Arabidopsis* in different tissues and in response to different stress at different development stages

A total of 24,123 RNA-Seq sample of *Arabidopsis thaliana* with Col-0 genotype were collected from GEO (Gene Expression Omnibus), Trimmomatic (Bolger et al., 2014) was used to remove the adapter and quality control with LEADING:20 TRAILING:20 SLIDINGWINDOW:4:20 MINLEN:20, salmon (Patro et al., 2017) was used to call TPM

(Transcripts Per Million) value from RNA-Seq data. The TPM of the isoform of all *AtADF* genes is used to do PCA (Principal Component Analysis). The R package pheatmap was used to generate the heatmap for each cluster.

BIBLIOGRAPHY

Abdelsamad, N. A., MacIntosh, G. C., & Leandro, L. F. S. (2019). Induction of ethylene inhibits development of soybean sudden death syndrome by inducing defense-related genes and reducing Fusarium virguliforme growth. PLoS One, 14(5), e0215653. https://doi.org/10.1371/journal.pone.0215653

Alexa A., a. R. J. (2018). topGO: Enrichment analysis for gene ontology., R package version 2.36.0.

Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. BMC Bioinformatics, 12, 246. https://doi.org/10.1186/1471-2105-12-246

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res, 19(9), 1655-1664. https://doi.org/10.1101/gr.094052.109

Allen, T. W., Bradley, C.A., Sisson, A.J., Byamukama, E., Chilvers, M.I., Coker, C.M., Collins, A.A., Damicone, J.P., Dorrance, A.E., Dufault, N.S., Esker, P.D., Faske, T.R., Giesler, L.J., Grybauskas, A.P., Hershman, D.E., Hollier, C.A., Isakeit, T., Jardine, D.J., Kelly, H.M., Kemerait, R.C., Kleczewski, N.M., Koenning, S.R., Kurle, J.E., Malvick, D.K., Markell, S.G., Mehl, H.L., Mueller, D.S., Mueller, J.D., Mulrooney, R.P., Nelson, B.D., Newman, M.A., Osborne, L., Overstreet, C., Padgett, G.B., Phipps, P.M., Price, P.P., Sikora, E.J., Smith, D.L., Spurlock, T.N., Tande, C.A., Tenuta, A.U., Wise, K.A., and Wrather, J.A. . (2017). Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2010 to 2014. Plant Health Prog. , 18, 19-27.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. J Mol Biol, 215(3), 403-410. https://doi.org/10.1016/S0022-2836(05)80360-2

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res, 25(17), 3389-3402. https://doi.org/10.1093/nar/25.17.3389

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics, 31(2), 166-169. https://doi.org/10.1093/bioinformatics/btu638

Andrianantoandro, E., & Pollard, T. D. (2006). Mechanism of actin filament turnover by severing and nucleation at different concentrations of ADF/cofilin. Mol Cell, 24(1), 13-23. https://doi.org/10.1016/j.molcel.2006.08.006

Askew, C., Sellam, A., Epp, E., Hogues, H., Mullick, A., Nantel, A., & Whiteway, M. (2009). Transcriptional regulation of carbohydrate metabolism in the human pathogen Candida albicans. PLoS Pathog, 5(10), e1000612. https://doi.org/10.1371/journal.ppat.1000612

Atkinson, N. J., & Urwin, P. E. (2012). The interaction of plant biotic and abiotic stresses: from genes to the field. J Exp Bot, 63(10), 3523-3543. https://doi.org/10.1093/jxb/ers100

Aung, K., Jiang, Y., & He, S. Y. (2018). The role of water in plant-microbe interactions. Plant J, 93(4), 771-780. https://doi.org/10.1111/tpj.13795

Axtell, M. J., & Staskawicz, B. J. (2003). Initiation of RPS2-specified disease resistance in Arabidopsis is coupled to the AvrRpt2-directed elimination of RIN4. Cell, 112(3), 369-377. https://doi.org/10.1016/s0092-8674(03)00036-9

Baetsen-Young, A., Man Wai, C., VanBuren, R., & Day, B. (2020). Fusarium virguliform e Transcriptional Plasticity Is Revealed by Host Colonization of Maize versus Soybean. Plant Cell, 32(2), 336-351. https://doi.org/10.1105/tpc.19.00697

Bagnaresi, P., Biselli, C., Orru, L., Urso, S., Crispino, L., Abbruscato, P., Piffanelli, P., Lupotto, E., Cattivelli, L., & Vale, G. (2012). Comparative transcriptome profiling of the early response to Magnaporthe oryzae in durable resistant vs susceptible rice (Oryza sativa L.) genotypes. PLoS One, 7(12), e51609. https://doi.org/10.1371/journal.pone.0051609

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

Balmer, D., de Papajewski, D. V., Planchamp, C., Glauser, G., & Mauch-Mani, B. (2013b). Induced resistance in maize is based on organ-specific defence responses. Plant J, 74(2), 213-225. https://doi.org/10.1111/tpj.12114

Balmer, D., Planchamp, C., & Mauch-Mani, B. (2013a). On the move: induced resistance in monocots. J Exp Bot, 64(5), 1249-1261. https://doi.org/10.1093/jxb/ers248

Bandara, A. Y., Weerasooriya, D. K., Conley, S. P., Allen, T. W., & Esker, P. D. (2020). Modeling the relationship between estimated fungicide use and disease-associated yield losses of soybean in the United States II: Seed-applied fungicides vs seedling diseases. PLoS One, 15(12), e0244424. https://doi.org/10.1371/journal.pone.0244424

Banerjee, S., Walder, F., Buchi, L., Meyer, M., Held, A. Y., Gattinger, A., Keller, T., Charles, R., & van der Heijden, M. G. A. (2019). Agricultural intensification reduces

microbial network complexity and the abundance of keystone taxa in roots. ISME J, 13(7), 1722-1736. https://doi.org/10.1038/s41396-019-0383-2

Bashir, K., Matsui, A., Rasheed, S., & Seki, M. (2019). Recent advances in the characterization of plant transcriptomes in response to drought, salinity, heat, and cold stress. F1000Res, 8. https://doi.org/10.12688/f1000research.18424.1

Bates, D., Machler, M., Bolker, B., and Walker, S. . (2015). Fitting linear mixed-effect models using lme4. J. Stat. Soft 67, 1-48.

Battey, C. J., Ralph, P. L., & Kern, A. D. (2020). Predicting geographic location from genetic variation with deep neural networks. Elife, 9. https://doi.org/10.7554/eLife.54507

Behr, A. A., Liu, K. Z., Liu-Fang, G., Nakka, P., & Ramachandran, S. (2016). pong: fast analysis and visualization of latent clusters in population genetic data. Bioinformatics, 32(18), 2817-2823. https://doi.org/10.1093/bioinformatics/btw327

Bellon, M. R. (1996). The dynamics of crop infraspecific diversity: A conceptual framework at the farmer level. Economic Botany, 50(1), 26-39. https://doi.org/Doi10.1007/Bf02862110

Berkopec, A. (2007). HyperQuick algorithm for discrete hypergeometric distribution. Journal of Discrete Algorithms, 5(2), 341-347. https://doi.org/10.1016/j.jda.2006.01.001

Bi, S., Li, M., Liu, C., Liu, X., Cheng, J., Wang, L., Wang, J., Lv, Y., He, M., Cheng, X., Gao, Y., & Wang, C. (2022). Actin depolymerizing factor ADF7 inhibits actin bundling protein VILLIN1 to regulate root hair formation in response to osmotic stress in Arabidopsis. PLoS Genet, 18(9), e1010338. https://doi.org/10.1371/journal.pgen.1010338

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114-2120. https://doi.org/10.1093/bioinformatics/btu170

Bolt, S., Zuther, E., Zintl, S., Hincha, D. K., & Schmulling, T. (2017). ERF105 is a transcription factor gene of Arabidopsis thaliana required for freezing tolerance and cold acclimation. Plant Cell Environ, 40(1), 108-120. https://doi.org/10.1111/pce.12838

Bonavia, D. (2013). Maize : origin, domestication, and its role in the development of culture. Cambridge University Press.

Boretti, A. R., L. (2019). Reassessing the Projections of the World Water Development Report. NPJ Clean. Water, 2019, 2, 15.

Bou Daher, F., van Oostende, C., & Geitmann, A. (2011). Spatial and temporal expression of actin depolymerizing factors ADF7 and ADF10 during male gametophyte development in Arabidopsis thaliana. Plant Cell Physiol, 52(7), 1177-1192. https://doi.org/10.1093/pcp/pcr068

Boyer, J. S. (1982). Plant productivity and environment. Science, 218(4571), 443-448. https://doi.org/10.1126/science.218.4571.443

Brakebusch, C., & Fassler, R. (2003). The integrin-actin connection, an eternal love affair. EMBO J, 22(10), 2324-2333. https://doi.org/10.1093/emboj/cdg245

Bray, T. L. (2003). Inka pottery as culinary equipment: Food, fasting, and gender in imperial state design. Latin American Antiquity, 14(1), 3-28. https://doi.org/Doi 10.2307/972232

Briggs, A. W., Stenzel, U., Johnson, P. L., Green, R. E., Kelso, J., Prufer, K., Meyer, M., Krause, J., Ronan, M. T., Lachmann, M., & Paabo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. Proc Natl Acad Sci U S A, 104(37), 14616-14621. https://doi.org/10.1073/pnas.0704665104

Brodie, A., Azaria, J. R., & Ofran, Y. (2016). How far from the SNP may the causative genes be? Nucleic Acids Res, 44(13), 6046-6054. https://doi.org/10.1093/nar/gkw500

Bu, Q., Jiang, H., Li, C. B., Zhai, Q., Zhang, J., Wu, X., Sun, J., Xie, Q., & Li, C. (2008). Role of the Arabidopsis thaliana NAC transcription factors ANAC019 and ANAC055 in regulating jasmonic acid-signaled defense responses. Cell Res, 18(7), 756-767. https://doi.org/10.1038/cr.2008.53

Burger, R. L., & Vandermerwe, N. J. (1990). Maize and the Origin of Highland Chavin Civilization - an Isotopic Perspective. American Anthropologist, 92(1), 85-95. https://doi.org/DOI 10.1525/aa.1990.92.1.02a00060

Burgos-Rivera, B., Ruzicka, D. R., Deal, R. B., McKinney, E. C., King-Reid, L., & Meagher, R. B. (2008). ACTIN DEPOLYMERIZING FACTOR9 controls development and gene expression in Arabidopsis. Plant Mol Biol, 68(6), 619-632. https://doi.org/10.1007/s11103-008-9398-1

Burkhardt, A., & Day, B. (2016). Transcriptome and Small RNAome Dynamics during a Resistant and Susceptible Interaction between Cucumber and Downy Mildew. Plant Genome, 9(1). https://doi.org/10.3835/plantgenome2015.08.0069

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol, 36(5), 411-420. https://doi.org/10.1038/nbt.4096

Campbell, I. D., & Humphries, M. J. (2011). Integrin structure, activation, and interactions. Cold Spring Harb Perspect Biol, 3(3). https://doi.org/10.1101/cshperspect.a004994

Canto-Pastor, A., Mason, G. A., Brady, S. M., & Provart, N. J. (2021). Arabidopsis bioinformatics: tools and strategies. Plant J, 108(6), 1585-1596. https://doi.org/10.1111/tpj.15547

Cao, F. Y., Yoshioka, K., & Desveaux, D. (2011). The roles of ABA in plant-pathogen interactions. J Plant Res, 124(4), 489-499. https://doi.org/10.1007/s10265-011-0409-y

Century, K. S., Holub, E. B., & Staskawicz, B. J. (1995). NDR1, a locus of Arabidopsis thaliana that is required for disease resistance to both a bacterial and a fungal pathogen. Proc Natl Acad Sci U S A, 92(14), 6597-6601. https://doi.org/10.1073/pnas.92.14.6597

Century, K. S., Shapiro, A. D., Repetti, P. P., Dahlbeck, D., Holub, E., & Staskawicz, B. J. (1997). NDR1, a pathogen-induced component required for Arabidopsis disease resistance. Science, 278(5345), 1963-1965. https://doi.org/10.1126/science.278.5345.1963

Cerda-Hurtado, I. M., Mayek-Perez, N., Hernandez-Delgado, S., Muruaga-Martinez, J. S., Reyes-Lara, M. A., Reyes-Valdes, M. H., & Gonzalez-Prieto, J. M. (2018). Climatic adaptation and ecological descriptors of wild beans from Mexico. Ecol Evol, 8(13), 6492-6504. https://doi.org/10.1002/ece3.4106

Ceruti, M. C. (2015). Frozen Mummies from Andean Mountaintop Shrines: Bioarchaeology and Ethnohistory of Inca Human Sacrifice. Biomed Res Int, 2015, 439428. https://doi.org/10.1155/2015/439428

Chang, H. X., Domier, L. L., Radwan, O., Yendrek, C. R., Hudson, M. E., & Hartman, G. L. (2016). Identification of Multiple Phytotoxins Produced by Fusarium virguliforme Including a Phytotoxic Effector (FvNIS1) Associated With Sudden Death Syndrome Foliar Symptoms. Mol Plant Microbe Interact, 29(2), 96-108. https://doi.org/10.1094/MPMI-09-15-0219-R

Chen, H., Patterson, N., & Reich, D. (2010). Population differentiation as a test for selective sweeps. Genome Research, 20(3), 393-402. https://doi.org/10.1101/gr.100545.109

Chen, Y., Yin, H., Gao, M., Zhu, H., Zhang, Q., & Wang, Y. (2016). Comparative Transcriptomics Atlases Reveals Different Gene Expression Pattern Related to Fusarium Wilt Disease Resistance and Susceptibility in Two Vernicia Species. Front Plant Sci, 7, 1974. https://doi.org/10.3389/fpls.2016.01974

Chia, J. M., Song, C., Bradbury, P. J., Costich, D., de Leon, N., Doebley, J., Elshire, R. J., Gaut, B., Geller, L., Glaubitz, J. C., Gore, M., Guill, K. E., Holland, J., Hufford, M. B., Lai, J., Li, M., Liu, X., Lu, Y., McCombie, R., . . . Ware, D. (2012). Maize HapMap2 identifies extant variation from a genome in flux. Nat Genet, 44(7), 803-807. https://doi.org/10.1038/ng.2313

Chisholm, S. T., Coaker, G., Day, B., & Staskawicz, B. J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. Cell, 124(4), 803-814. https://doi.org/10.1016/j.cell.2006.02.008

Chowdhury, S., Basu, A., & Kundu, S. (2017). Biotrophy-necrotrophy switch in pathogen evoke differential response in resistant and susceptible sesame involving multiple signaling pathways at different phases. Sci Rep, 7(1), 17251. https://doi.org/10.1038/s41598-017-17248-7

Chuberre, C., Plancot, B., Driouich, A., Moore, J. P., Bardor, M., Gugi, B., & Vicre, M. (2018). Plant Immunity Is Compartmentalized and Specialized in Roots. Front Plant Sci, 9, 1692. https://doi.org/10.3389/fpls.2018.01692

Claeys, H., Van Landeghem, S., Dubois, M., Maleux, K., & Inze, D. (2014). What Is Stress? Dose-Response Effects in Commonly Used in Vitro Stress Assays. Plant Physiol, 165(2), 519-527. https://doi.org/10.1104/pp.113.234641

Clement, M., Ketelaar, T., Rodiuc, N., Banora, M. Y., Smertenko, A., Engler, G., Abad, P., Hussey, P. J., & de Almeida Engler, J. (2009). Actin-depolymerizing factor2-mediated actin dynamics are essential for root-knot nematode infection of Arabidopsis. Plant Cell, 21(9), 2963-2979. https://doi.org/10.1105/tpc.109.069104

Cobo, B., & Hamilton, R. (1990). Inca religion and customs (1st ed.). University of Texas Press.

COOK, O. F. (1925). PERU AS A CENTER OF DOMESTICATION: Tracing the Origin of Civilization Through the Domesticated Plants. Journal of Heredity, 16(3), 95-110. https://doi.org/10.1093/oxfordjournals.jhered.a102570

Coppinger, P., Repetti, P. P., Day, B., Dahlbeck, D., Mehlert, A., & Staskawicz, B. J. (2004). Overexpression of the plasma membrane-localized NDR1 protein results in enhanced bacterial disease resistance in Arabidopsis thaliana. Plant J, 40(2), 225-237. https://doi.org/10.1111/j.1365-313X.2004.02203.x

Crick, F. H. (1968). The origin of the genetic code. J Mol Biol, 38(3), 367-379. https://doi.org/10.1016/0022-2836(68)90392-6

Dabney, J., Meyer, M., & Paabo, S. (2013). Ancient DNA damage. Cold Spring Harb Perspect Biol, 5(7). https://doi.org/10.1101/cshperspect.a012567

Daher, F. B., & Geitmann, A. (2012). Actin depolymerizing factors ADF7 and ADF10 play distinct roles during pollen development and pollen tube growth. Plant Signal Behav, 7(7), 879-881. https://doi.org/10.4161/psb.20436

Davila Olivas, N. H., Coolen, S., Huang, P., Severing, E., van Verk, M. C., Hickman, R., Wittenberg, A. H., de Vos, M., Prins, M., van Loon, J. J., Aarts, M. G., van Wees, S. C., Pieterse, C. M., & Dicke, M. (2016). Effect of prior drought and pathogen stress on Arabidopsis transcriptome changes to caterpillar herbivory. New Phytol, 210(4), 1344-1356. https://doi.org/10.1111/nph.13847

Day, B., Dahlbeck, D., & Staskawicz, B. J. (2006). NDR1 interaction with RIN4 mediates the differential activation of multiple disease resistance pathways in Arabidopsis. Plant Cell, 18(10), 2782-2791. https://doi.org/10.1105/tpc.106.044693

Day, B., & Graham, T. (2007). The plant host pathogen interface: cell wall and membrane dynamics of pathogen-induced responses. Ann N Y Acad Sci, 1113, 123-134. https://doi.org/10.1196/annals.1391.029

Day, B., Henty, J. L., Porter, K. J., & Staiger, C. J. (2011). The pathogen-actin connection: a platform for defense signaling in plants. Annu Rev Phytopathol, 49, 483-506. https://doi.org/10.1146/annurev-phyto-072910-095426

Day, R. B., Okada, M., Ito, Y., Tsukada, K., Zaghouani, H., Shibuya, N., & Stacey, G. (2001). Binding site for chitin oligosaccharides in the soybean plasma membrane. Plant Physiol, 126(3), 1162-1173. https://doi.org/10.1104/pp.126.3.1162

De Coninck, B., Timmermans, P., Vos, C., Cammue, B. P., & Kazan, K. (2015). What lies beneath: belowground defense strategies in plants. Trends Plant Sci, 20(2), 91-101. https://doi.org/10.1016/j.tplants.2014.09.007

de Meeus, T., & Goudet, J. (2007). A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels. Infect Genet Evol, 7(6), 731-735. https://doi.org/10.1016/j.meegid.2007.07.005

Denyer, T., Ma, X., Klesen, S., Scacchi, E., Nieselt, K., & Timmermans, M. C. P. (2019). Spatiotemporal Developmental Trajectories in the Arabidopsis Root Revealed Using High-Throughput Single-Cell RNA Sequencing. Dev Cell, 48(6), 840-852 e845. https://doi.org/10.1016/j.devcel.2019.02.022

Des Marais, D. L., Hernandez, K. M., & Juenger, T. E. (2013). Genotype-by-environment interaction and plasticity: exploring genomic rsponses of plants to the abiotic environment. Ann Rev Ecol, Evol, System, 44, 5-29.

Deshmukh, S., Huckelhoven, R., Schafer, P., Imani, J., Sharma, M., Weiss, M., Waller, F., & Kogel, K. H. (2006). The root endophytic fungus Piriformospora indica

requires host cell death for proliferation during mutualistic symbiosis with barley. Proc Natl Acad Sci U S A, 103(49), 18450-18457. https://doi.org/10.1073/pnas.0605697103

Desikan, R., Horak, J., Chaban, C., Mira-Rodado, V., Witthoft, J., Elgass, K., Grefen, C., Cheung, M. K., Meixner, A. J., Hooley, R., Neill, S. J., Hancock, J. T., & Harter, K. (2008). The histidine kinase AHK5 integrates endogenous and environmental signals in Arabidopsis guard cells. PLoS One, 3(6), e2491. https://doi.org/10.1371/journal.pone.0002491

Deveau, A., Kohler, A., Frey-Klett, P., & Martin, F. (2008). The major pathways of carbohydrate metabolism in the ectomycorrhizal basidiomycete Laccaria bicolor S238N. New Phytol, 180(2), 379-390. https://doi.org/10.1111/j.1469-8137.2008.02581.x

Diamond, M., Reape, T. J., Rocha, O., Doyle, S. M., Kacprzyk, J., Doohan, F. M., & McCabe, P. F. (2013). The fusarium mycotoxin deoxynivalenol can inhibit plant apoptosis-like programmed cell death. PLoS One, 8(7), e69542. https://doi.org/10.1371/journal.pone.0069542

Divon, H. H., & Fluhr, R. (2007). Nutrition acquisition strategies during fungal infection of plants. FEMS Microbiol Lett, 266(1), 65-74. https://doi.org/10.1111/j.1574-6968.2006.00504.x

Doebley, J. (2004). The genetics of maize evolution. Annu Rev Genet, 38, 37-59. https://doi.org/10.1146/annurev.genet.38.072902.092425

Doehlemann, G., Okmen, B., Zhu, W., & Sharon, A. (2017). Plant Pathogenic Fungi. Microbiol Spectr, 5(1). https://doi.org/10.1128/microbiolspec.FUNK-0023-2016

Dong, C. H., Xia, G. X., Hong, Y., Ramachandran, S., Kost, B., & Chua, N. H. (2001). ADF proteins are involved in the control of flowering and regulate F-actin organization, cell expansion, and organ growth in Arabidopsis. Plant Cell, 13(6), 1333-1346. https://doi.org/10.1105/tpc.13.6.1333

Doutriaux, M. (2001). Power, ideology and ritual: the practice of agriculture in the Inca Empire. Pap Kroeber Anthropol Soc(85), 91-108. https://www.ncbi.nlm.nih.gov/pubmed/19035027

Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for Ancient Admixture between Closely Related Populations. Molecular Biology and Evolution, 28(8), 2239-2252. https://doi.org/10.1093/molbev/msr048

Edman, P., & Begg, G. (1967). A protein sequenator. Eur J Biochem, 1(1), 80-91. https://doi.org/10.1007/978-3-662-25813-2_14

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol, 16(1), 157. https://doi.org/10.1186/s13059-015-0721-2

Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol, 20(1), 238. https://doi.org/10.1186/s13059-019-1832-y

Etalo, D. W., Stulemeijer, I. J., van Esse, H. P., de Vos, R. C., Bouwmeester, H. J., & Joosten, M. H. (2013). System-wide hypersensitive response-associated transcriptome and metabolome reprogramming in tomato. Plant Physiol, 162(3), 1599-1617. https://doi.org/10.1104/pp.113.217471

Fahad, S., Bajwa, A. A., Nazir, U., Anjum, S. A., Farooq, A., Zohaib, A., Sadia, S., Nasim, W., Adkins, S., Saud, S., Ihsan, M. Z., Alharby, H., Wu, C., Wang, D., & Huang, J. (2017). Crop Production under Drought and Heat Stress: Plant Responses and Management Options. Front Plant Sci, 8, 1147. https://doi.org/10.3389/fpls.2017.01147

Fan, J., Hill, L., Crooks, C., Doerner, P., & Lamb, C. (2009). Abscisic acid has a key role in modulating diverse plant-pathogen interactions. Plant Physiol, 150(4), 1750-1761. https://doi.org/10.1104/pp.109.137943

Farmer, A., Thibivilliers, S., Ryu, K. H., Schiefelbein, J., & Libault, M. (2021). Single-nucleus RNA and ATAC sequencing reveals the impact of chromatin accessibility on gene expression in Arabidopsis roots at the single-cell level. Mol Plant, 14(3), 372-383. https://doi.org/10.1016/j.molp.2021.01.001

Fehr WR, C. C. (1977). Stages of soybean development. Iowa

State University Special Report, 87. http://lib.dr.iastate.edu/specialreports/87.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol, 17(6), 368-376. https://doi.org/10.1007/BF01734359

Feng, W., Lindner, H., Robbins, N. E., 2nd, & Dinneny, J. R. (2016). Growing Out of Stress: The Role of Cell- and Organ-Scale Growth Control in Plant Water-Stress Responses. Plant Cell, 28(8), 1769-1782. https://doi.org/10.1105/tpc.16.00182

Fitak, R. R. (2021). OptM: estimating the optimal number of migration edges on population trees using Treemix. Biol Methods Protoc, 6(1), bpab017. https://doi.org/10.1093/biomethods/bpab017

Fox, J., and Weisberg, S. . (2011). An R companion to applied regression. . Thousand Oaks {CA}: Sage.

Freeman, B. C., & Beattie, G. A. (2009). Bacterial growth restriction during host resistance to Pseudomonas syringae is associated with leaf water loss and localized cessation of vascular activity in Arabidopsis thaliana. Mol Plant Microbe Interact, 22(7), 857-867. https://doi.org/10.1094/MPMI-22-7-0857

Fujii, H., Verslues, P. E., & Zhu, J. K. (2007). Identification of two protein kinases required for abscisic acid regulation of seed germination, root growth, and gene expression in Arabidopsis. Plant Cell, 19(2), 485-494. https://doi.org/10.1105/tpc.106.048538

Garcia-Gomez, B. I., Campos, F., Hernandez, M., & Covarrubias, A. A. (2000). Two bean cell wall proteins more abundant during water deficit are high in proline and interact with a plasma membrane protein. Plant J, 22(4), 277-288. https://doi.org/10.1046/j.1365-313x.2000.00739.x

Garciarrubio, A., Legaria, J. P., & Covarrubias, A. A. (1997). Abscisic acid inhibits germination of mature Arabidopsis seeds by limiting the availability of energy and nutrients. Planta, 203(2), 182-187. https://doi.org/10.1007/s004250050180

Gauthier, J., Vincent, A. T., Charette, S. J., & Derome, N. (2019). A brief history of bioinformatics. Brief Bioinform, 20(6), 1981-1996. https://doi.org/10.1093/bib/bby063

Gdanetz, K., and Trail, F. . (2017). The wheat microbiome under four management strategies, and potential for endophytes in disease protection. Phytobiomes J., 1, 158-168.

Gilbert, M. T., Hansen, A. J., Willerslev, E., Rudbeck, L., Barnes, I., Lynnerup, N., & Cooper, A. (2003). Characterization of genetic miscoding lesions caused by postmortem damage. Am J Hum Genet, 72(1), 48-61. https://doi.org/10.1086/345379

Goel, A. K., Lundberg, D., Torres, M. A., Matthews, R., Akimoto-Tomiyama, C., Farmer, L., Dangl, J. L., & Grant, S. R. (2008). The Pseudomonas syringae type III effector HopAM1 enhances virulence on water-stressed plants. Mol Plant Microbe Interact, 21(3), 361-370. https://doi.org/10.1094/MPMI-21-3-0361

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H., Hansen, N. F., Durand, E. Y., Malaspinas, A. S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prufer, K., Meyer, M., Burbano, H. A., . . . Paabo, S. (2010). A draft sequence of the Neandertal genome. Science, 328(5979), 710-722. https://doi.org/10.1126/science.1188021

Grobman, A., Bonavia, D., Dillehay, T. D., Piperno, D. R., Iriarte, J., & Holst, I. (2012). Preceramic maize from Paredones and Huaca Prieta, Peru. Proc Natl Acad Sci U S A, 109(5), 1755-1759. https://doi.org/10.1073/pnas.1120270109

Guo, Q., Major, I. T., & Howe, G. A. (2018). Resolution of growth-defense conflict: mechanistic insights from jasmonate signaling. Curr Opin Plant Biol, 44, 72-81. https://doi.org/10.1016/j.pbi.2018.02.009

Gupta, A., Hisano, H., Hojo, Y., Matsuura, T., Ikeda, Y., Mori, I. C., & Senthil-Kumar, M. (2017). Global profiling of phytohormone dynamics during combined drought and pathogen stress in Arabidopsis thaliana reveals ABA and JA as major regulators. Sci Rep, 7(1), 4017. https://doi.org/10.1038/s41598-017-03907-2

Gupta, A., Rico-Medina, A., & Cano-Delgado, A. I. (2020). The physiology of plant responses to drought. Science, 368(6488), 266-269. https://doi.org/10.1126/science.aaz7614

Haffner, E., Karlovsky, P., & Diederichsen, E. (2010). Genetic and environmental control of the Verticillium syndrome in Arabidopsis thaliana. BMC Plant Biol, 10, 235. https://doi.org/10.1186/1471-2229-10-235

Haffner, E., Konietzki, S., & Diederichsen, E. (2015). Keeping Control: The Role of Senescence and Development in Plant Pathogenesis and Defense. Plants (Basel), 4(3), 449-488. https://doi.org/10.3390/plants4030449

Handt, O., Krings, M., Ward, R. H., & Paabo, S. (1996). The retrieval of ancient human DNA sequences. Am J Hum Genet, 59(2), 368-376. https://www.ncbi.nlm.nih.gov/pubmed/8755923

Hansen, A., Willerslev, E., Wiuf, C., Mourier, T., & Arctander, P. (2001). Statistical evidence for miscoding lesions in ancient DNA templates. Mol Biol Evol, 18(2), 262-265. https://doi.org/10.1093/oxfordjournals.molbev.a003800

Harney, E., Patterson, N., Reich, D., & Wakeley, J. (2021). Assessing the performance of qpAdm: a statistical tool for studying population admixture. Genetics, 217(4). https://doi.org/10.1093/genetics/iyaa045

Hart, J. P., & Lovis, W. A. (2013). Reevaluating What We Know About the Histories of Maize in Northeastern North America: A Review of Current Evidence. Journal of Archaeological Research, 21(2), 175-216. https://doi.org/10.1007/s10814-012-9062-9

Haudenshield, J. S., & Hartman, G. L. (2011). Exogenous Controls Increase Negative Call Veracity in Multiplexed, Quantitative PCR Assays for Phakopsora pachyrhizi. Plant Dis, 95(3), 343-352. https://doi.org/10.1094/PDIS-01-10-0023

Henty, J. L., Bledsoe, S. W., Khurana, P., Meagher, R. B., Day, B., Blanchoin, L., & Staiger, C. J. (2011). Arabidopsis actin depolymerizing factor4 modulates the stochastic dynamic behavior of actin filaments in the cortical array of epidermal cells. Plant Cell, 23(10), 3711-3726. https://doi.org/10.1105/tpc.111.090670

Henty-Ridilla, J. L., Li, J., Blanchoin, L., & Staiger, C. J. (2013). Actin dynamics in the cortical array of plant cells. Curr Opin Plant Biol, 16(6), 678-687. https://doi.org/10.1016/j.pbi.2013.10.012

Henty-Ridilla, J. L., Li, J., Day, B., & Staiger, C. J. (2014). ACTIN DEPOLYMERIZING FACTOR4 regulates actin dynamics during innate immune signaling in Arabidopsis. Plant Cell, 26(1), 340-352. https://doi.org/10.1105/tpc.113.122499

Higaki, T., Sano, T., & Hasezawa, S. (2007). Actin microfilament dynamics and actin side-binding proteins in plants. Curr Opin Plant Biol, 10(6), 549-556. https://doi.org/10.1016/j.pbi.2007.08.012

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., & Nakai, K. (2007). WoLF PSORT: protein localization predictor. Nucleic Acids Res, 35(Web Server issue), W585-587. https://doi.org/10.1093/nar/gkm259

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. Biom J, 50(3), 346-363. https://doi.org/10.1002/bimj.200810425

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., & Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics, 31(8), 1296-1297. https://doi.org/10.1093/bioinformatics/btu817

Huang, S., Qu, X., & Zhang, R. (2015). Plant villins: versatile actin regulatory proteins. J Integr Plant Biol, 57(1), 40-49. https://doi.org/10.1111/jipb.12293

Hufford, M. B., Lubinksy, P., Pyhajarvi, T., Devengenzo, M. T., Ellstrand, N. C., & Ross-Ibarra, J. (2013). The genomic signature of crop-wild introgression in maize. PLoS Genet, 9(5), e1003477. https://doi.org/10.1371/journal.pgen.1003477

Humphry, M., Bednarek, P., Kemmerling, B., Koh, S., Stein, M., Gobel, U., Stuber, K., Pislewska-Bednarek, M., Loraine, A., Schulze-Lefert, P., Somerville, S., & Panstruga, R. (2010). A regulon conserved in monocot and dicot plants defines a functional module in antifungal plant immunity. Proc Natl Acad Sci U S A, 107(50), 21896-21901. https://doi.org/10.1073/pnas.1003619107

Hussey, P. J., Ketelaar, T., & Deeks, M. J. (2006). Control of the actin cytoskeleton in plant cell growth. Annu Rev Plant Biol, 57, 109-125. https://doi.org/10.1146/annurev.arplant.57.032905.105206

Huysmans, M., Buono, R. A., Skorzinski, N., Radio, M. C., De Winter, F., Parizot, B., Mertens, J., Karimi, M., Fendrych, M., & Nowack, M. K. (2018). NAC Transcription Factors ANAC087 and ANAC046 Control Distinct Aspects of Programmed Cell Death in the Arabidopsis Columella and Lateral Root Cap. Plant Cell, 30(9), 2197-2213. https://doi.org/10.1105/tpc.18.00293

Inada, N. (2018). Correction to: Plant actin depolymerizing factor: actin microfilament disassembly and more. J Plant Res, 131(3), 567. https://doi.org/10.1007/s10265-018-1013-1

Inada, N., Higaki, T., & Hasezawa, S. (2016). Nuclear Function of Subclass I Actin-Depolymerizing Factor Contributes to Susceptibility in Arabidopsis to an Adapted Powdery Mildew Fungus. Plant Physiol, 170(3), 1420-1434. https://doi.org/10.1104/pp.15.01265

Inada, N., Takahashi, N., & Umeda, M. (2021). Arabidopsis thaliana subclass I ACTIN DEPOLYMERIZING FACTORs and vegetative ACTIN2/8 are novel regulators of endoreplication. J Plant Res, 134(6), 1291-1300. https://doi.org/10.1007/s10265-021-01333-0

Jacob, P., Hirt, H., & Bendahmane, A. (2017). The heat-shock protein/chaperone network and multiple stress resistance. Plant Biotechnol J, 15(4), 405-414. https://doi.org/10.1111/pbi.12659

Jakubowska, D., & Janicka, M. (2017). The role of brassinosteroids in the regulation of the plasma membrane H(+)-ATPase and NADPH oxidase under cadmium stress. Plant Sci, 264, 37-47. https://doi.org/10.1016/j.plantsci.2017.08.007

Jarzyniak, K. M., & Jasinski, M. (2014). Membrane transporters and drought resistance - a complex issue. Front Plant Sci, 5, 687. https://doi.org/10.3389/fpls.2014.00687

Jiang, Y., Wang, J., Xie, Y., Chen, N., & Huang, S. (2017). ADF10 shapes the overall organization of apical actin filaments by promoting their turnover and ordering in pollen tubes. J Cell Sci, 130(23), 3988-4001. https://doi.org/10.1242/jcs.207738

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics, 27(21), 3070-3071. https://doi.org/10.1093/bioinformatics/btr521

Jones, A. M. (2001). Programmed cell death in development and defense. Plant Physiol, 125(1), 94-97. https://doi.org/10.1104/pp.125.1.94

Jonsson, H., Ginolhac, A., Schubert, M., Johnson, P. L., & Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics, 29(13), 1682-1684. https://doi.org/10.1093/bioinformatics/btt193

Joshi, R., Shukla, A., & Sairam, R. J. (2011). In vitro screening of rice genotypes for drought tolerance using polyethylene glycol. Acta Physiol. Plantarum, 33, 2209.

Katan, J. (2017). Diseases caused by soilborne pathogens: Biology, management and challenges. Journal of Plant Pathology, Vol. 99, No. 2 (July 2017), pp. 305-315.

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., & Phillips, D. C. (1958). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature, 181(4610), 662-666. https://doi.org/10.1038/181662a0

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Methods, 12(4), 357-360. https://doi.org/10.1038/nmeth.3317

Kim, J. Y., Symeonidi, E., Pang, T. Y., Denyer, T., Weidauer, D., Bezrutczyk, M., Miras, M., Zollner, N., Hartwig, T., Wudick, M. M., Lercher, M., Chen, L. Q., Timmermans, M. C. P., & Frommer, W. B. (2021). Distinct identities of leaf phloem cells revealed by single cell transcriptomics. Plant Cell, 33(3), 511-530. https://doi.org/10.1093/plcell/koaa060

Kistler, L., Maezumi, S. Y., Gregorio de Souza, J., Przelomska, N. A. S., Malaquias Costa, F., Smith, O., Loiselle, H., Ramos-Madrigal, J., Wales, N., Ribeiro, E. R., Morrison, R. R., Grimaldo, C., Prous, A. P., Arriaza, B., Gilbert, M. T. P., de Oliveira Freitas, F., & Allaby, R. G. (2018). Multiproxy evidence highlights a complex evolutionary legacy of maize in South America. Science, 362(6420), 1309-1313. https://doi.org/10.1126/science.aav0207

Kistler, L., Thakar, H. B., VanDerwarker, A. M., Domic, A., Bergstrom, A., George, R. J., Harper, T. K., Allaby, R. G., Hirth, K., & Kennett, D. J. (2020). Archaeological Central American maize genomes suggest ancient gene flow from South America. Proc Natl Acad Sci U S A, 117(52), 33124-33129. https://doi.org/10.1073/pnas.2015560117

Kiyosue, T., Yamaguchi-Shinozaki, K., & Shinozaki, K. (1994). Cloning of cDNAs for genes that are early-responsive to dehydration stress (ERDs) in Arabidopsis thaliana L.: identification of three ERDs as HSP cognate genes. Plant Mol Biol, 25(5), 791-798. https://doi.org/10.1007/BF00028874

Klok, E. J., Wilson, I. W., Wilson, D., Chapman, S. C., Ewing, R. M., Somerville, S. C., Peacock, W. J., Dolferus, R., & Dennis, E. S. (2002). Expression profile analysis of the low-oxygen response in Arabidopsis root cultures. Plant Cell, 14(10), 2481-2494. https://doi.org/10.1105/tpc.004747

Knapp, G. (1997). Changing fortunes: Biodiversity and peasant livelihood in the Peruvian Andes. Geographical Review, 87(4), 568-570. https://doi.org/Doi 10.2307/215240

Knaus, B. J., & Grunwald, N. J. (2017). vcfr: a package to manipulate and visualize variant call format data in R. Mol Ecol Resour, 17(1), 44-53. https://doi.org/10.1111/1755-0998.12549

Knepper, C., Savory, E. A., & Day, B. (2011a). Arabidopsis NDR1 is an integrin-like protein with a role in fluid loss and plasma membrane-cell wall adhesion. Plant Physiol, 156(1), 286-300. https://doi.org/10.1104/pp.110.169656

Knepper, C., Savory, E. A., & Day, B. (2011b). The role of NDR1 in pathogen perception and plant defense signaling. Plant Signal Behav, 6(8), 1114-1116. https://doi.org/10.4161/psb.6.8.15843

Kolander, T. M., Bienapfl, J. C., Kurle, J. E., & Malvick, D. K. (2012). Symptomatic and Asymptomatic Host Range of Fusarium virguliforme, the Causal Agent of Soybean Sudden Death Syndrome. Plant Dis, 96(8), 1148-1153. https://doi.org/10.1094/PDIS-08-11-0685-RE

Kong, W., Chen, N., Liu, T., Zhu, J., Wang, J., He, X., & Jin, Y. (2015). Large-Scale Transcriptome Analysis of Cucumber and Botrytis cinerea during Infection. PLoS One, 10(11), e0142221. https://doi.org/10.1371/journal.pone.0142221

Kopecka, R., Kameniarova, M., Cerny, M., Brzobohaty, B., & Novak, J. (2023). Abiotic Stress in Crop Production. Int J Mol Sci, 24(7). https://doi.org/10.3390/ijms24076603

Koyama, T. (2014). The roles of ethylene and transcription factors in the regulation of onset of leaf senescence. Front Plant Sci, 5, 650. https://doi.org/10.3389/fpls.2014.00650

Koyama, T., Sato, F., & Ohme-Takagi, M. (2017). Roles of miR319 and TCP Transcription Factors in Leaf Development. Plant Physiol, 175(2), 874-885. https://doi.org/10.1104/pp.17.00732

Kretschmer, M., Damoo, D., Djamei, A., & Kronstad, J. (2019). Chloroplasts and Plant Immunity: Where Are the Fungal Effectors? Pathogens, 9(1). https://doi.org/10.3390/pathogens9010019

Krings, M., Stone, A., Schmitz, R. W., Krainitzki, H., Stoneking, M., & Paabo, S. (1997). Neandertal DNA sequences and the origin of modern humans. Cell, 90(1), 19-30. https://doi.org/10.1016/s0092-8674(00)80310-4

Laluk, K., & Mengiste, T. (2010). Necrotroph attacks on plants: wanton destruction or covert extortion? Arabidopsis Book, 8, e0136. https://doi.org/10.1199/tab.0136

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., . . . International Human Genome Sequencing, C. (2001). Initial sequencing and analysis of the human genome. Nature, 409(6822), 860-921. https://doi.org/10.1038/35057062

Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics, 9, 559. https://doi.org/10.1186/1471-2105-9-559

Lanubile, A., Muppirala, U. K., Severin, A. J., Marocco, A., & Munkvold, G. P. (2015). Transcriptome profiling of soybean (Glycine max) roots challenged with pathogenic and non-pathogenic isolates of Fusarium oxysporum. BMC Genomics, 16, 1089. https://doi.org/10.1186/s12864-015-2318-2

Le, N. Q. K., Yapp, E. K. Y., Nagasundaram, N., & Yeh, H. Y. (2019). Classifying Promoters by Interpreting the Hidden Information of DNA Sequences via Deep Learning and Combination of Continuous FastText N-Grams. Front Bioeng Biotechnol, 7, 305. https://doi.org/10.3389/fbioe.2019.00305

Leandro, L. F. S., Eggenberger, S., Chen, C., Williams, J., Beattie, G. A., & Liebman, M. (2018). Cropping System Diversification Reduces Severity and Incidence of Soybean Sudden Death Syndrome Caused by Fusarium virguliforme. Plant Dis, 102(9), 1748-1758. https://doi.org/10.1094/PDIS-11-16-1660-RE

Lee, T. H., Guo, H., Wang, X., Kim, C., & Paterson, A. H. (2014). SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. BMC Genomics, 15, 162. https://doi.org/10.1186/1471-2164-15-162

Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouze, P., & Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res, 30(1), 325-327. https://doi.org/10.1093/nar/30.1.325

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G., Kristiansen, K., Li, S., Yang, H., Wang, J., & Wang, J. (2010). De novo assembly of human genomes with massively parallel short read sequencing. Genome Res, 20(2), 265-272. https://doi.org/10.1101/gr.097261.109

Li, Z., Pan, X., Guo, X., Fan, K., & Lin, W. (2019). Physiological and Transcriptome Analyses of Early Leaf Senescence for ospls1 Mutant Rice (Oryza sativa L.) during the Grain-Filling Stage. Int J Mol Sci, 20(5). https://doi.org/10.3390/ijms20051098

Lin, H., Liang, Z. Y., Tang, H., & Chen, W. (2019). Identifying Sigma70 Promoters with Novel Pseudo Nucleotide Composition. IEEE/ACM Trans Comput Biol Bioinform, 16(4), 1316-1321. https://doi.org/10.1109/TCBB.2017.2666141

Lindgreen, S. (2012). AdapterRemoval: easy cleaning of next-generation sequencing reads. BMC Res Notes, 5, 337. https://doi.org/10.1186/1756-0500-5-337

Liu, J., Elmore, J. M., Fuglsang, A. T., Palmgren, M. G., Staskawicz, B. J., & Coaker, G. (2009). RIN4 functions with plasma membrane H+-ATPases to regulate stomatal apertures during pathogen attack. PLoS Biol, 7(6), e1000139. https://doi.org/10.1371/journal.pbio.1000139

Liu, Z., Marella, C. B. N., Hartmann, A., Hajirezaei, M. R., & von Wiren, N. (2019). An Age-Dependent Sequence of Physiological Processes Defines Developmental Root Senescence. Plant Physiol, 181(3), 993-1007. https://doi.org/10.1104/pp.19.00809

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. Methods, 25(4), 402-408. https://doi.org/10.1006/meth.2001.1262

Lofgren, L. A., LeBlanc, N. R., Certano, A. K., Nachtigall, J., LaBine, K. M., Riddle, J., Broz, K., Dong, Y., Bethan, B., Kafer, C. W., & Kistler, H. C. (2018). Fusarium graminearum: pathogen or endophyte of North American grasses? New Phytol, 217(3), 1203-1212. https://doi.org/10.1111/nph.14894

Lombardo, U., Iriarte, J., Hilbert, L., Ruiz-Perez, J., Capriles, J. M., & Veit, H. (2020). Early Holocene crop cultivation and landscape modification in Amazonia. Nature, 581(7807), 190-193. https://doi.org/10.1038/s41586-020-2162-7

Lorrain, C., Marchal, C., Hacquard, S., Delaruelle, C., Petrowski, J., Petre, B., Hecker, A., Frey, P., & Duplessis, S. (2018). The Rust Fungus Melampsora larici-populina Expresses a Conserved Genetic Program and Distinct Sets of Secreted Protein Genes During Infection of Its Two Host Plants, Larch and Poplar. Mol Plant Microbe Interact, 31(7), 695-706. https://doi.org/10.1094/MPMI-12-17-0319-R

Louette, D., & Smale, M. (2000). Farmers' seed selection practices and traditional maize varieties in Cuzalapa, Mexico. Euphytica, 113(1), 25-41. https://doi.org/Doi 10.1023/A:1003941615886

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

Lovis, W. A. (2019). A Mummy Returns Home. MSU Today, MSU 360, Faculty Voice.

Lu, H., McClung, C. R., & Zhang, C. (2017). Tick Tock: Circadian Regulation of Plant Innate Immunity. Annu Rev Phytopathol, 55, 287-311. https://doi.org/10.1146/annurev-phyto-080516-035451

Lu, S., Friesen, T. L., & Faris, J. D. (2011). Molecular characterization and genomic mapping of the pathogenesis-related protein 1 (PR-1) gene family in hexaploid

wheat (Triticum aestivum L.). Mol Genet Genomics, 285(6), 485-503. https://doi.org/10.1007/s00438-011-0618-z

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

MacRobbie, E. A. C., & Kurup, S. (2007). Signalling mechanisms in the regulation of vacuolar ion release in guard cells. New Phytol, 175(4), 630-640. https://doi.org/10.1111/j.1469-8137.2007.02131.x

Mahajan, S., & Tuteja, N. (2005). Cold, salinity and drought stresses: an overview. Arch Biochem Biophys, 444(2), 139-158. https://doi.org/10.1016/j.abb.2005.10.018

Majid I, A. N. (2019). Signal transduction in leaf senescence: an overview. In M Sarwat, M Tuteja (eds), Senescence Signalling and Control in Plants. Academic Press, Cambridge, MA,, pp 41–59.

Majumdar, R., Rajasekaran, K., Sickler, C., Lebar, M., Musungu, B. M., Fakhoury, A. M., Payne, G. A., Geisler, M., Carter-Wientjes, C., Wei, Q., Bhatnagar, D., & Cary, J. W. (2017). The Pathogenesis-Related Maize Seed (PRms) Gene Plays a Role in Resistance to Aspergillus flavus Infection and Aflatoxin Contamination. Front Plant Sci, 8, 1758. https://doi.org/10.3389/fpls.2017.01758

Malcolm, G. M., Kuldau, G. A., Gugino, B. K., & Jimenez-Gasco Mdel, M. (2013). Hidden host plant associations of soilborne fungal pathogens: an ecological perspective. Phytopathology, 103(6), 538-544. https://doi.org/10.1094/PHYTO-08-12-0192-LE

Mang, H. G., Qian, W., Zhu, Y., Qian, J., Kang, H. G., Klessig, D. F., & Hua, J. (2012). Abscisic acid deficiency antagonizes high-temperature inhibition of disease resistance through enhancing nuclear accumulation of resistance proteins SNC1 and RPS4 in Arabidopsis. Plant Cell, 24(3), 1271-1284. https://doi.org/10.1105/tpc.112.096198

Matsumoto, T., Higaki, T., Takatsuka, H., Kutsuna, N., Ogata, Y., Hasezawa, S., Umeda, M., & Inada, N. (2023). Arabidopsis thaliana Subclass I ACTIN DEPOLYMERIZING FACTORs Regulate Nuclear Organization and Gene Expression. Plant Cell Physiol. https://doi.org/10.1093/pcp/pcad092

Matsuoka, Y., Vigouroux, Y., Goodman, M. M., Sanchez, G. J., Buckler, E., & Doebley, J. (2002). A single domestication for maize shown by multilocus microsatellite genotyping. Proc Natl Acad Sci U S A, 99(9), 6080-6084. https://doi.org/10.1073/pnas.052125199

McInnes L, H. J., Melville J. (2018). UMAP: uniform manifold approximation and projection for dimension reduction. arXiv eprint arXiv, 1802.03426.

McKenzie, A. T., Katsyv, I., Song, W. M., Wang, M., & Zhang, B. (2016). DGCA: A comprehensive R package for Differential Gene Correlation Analysis. BMC Syst Biol, 10(1), 106. https://doi.org/10.1186/s12918-016-0349-1

Melotto, M., Underwood, W., Koczan, J., Nomura, K., & He, S. Y. (2006). Plant stomata function in innate immunity against bacterial invasion. Cell, 126(5), 969-980. https://doi.org/10.1016/j.cell.2006.06.054

Mercer, K., Martinez-Vasquez, A., & Perales, H. R. (2008). Asymmetrical local adaptation of maize landraces along an altitudinal gradient. Evol Appl, 1(3), 489-500. https://doi.org/10.1111/j.1752-4571.2008.00038.x

Michielse, C. B., & Rep, M. (2009). Pathogen profile update: Fusarium oxysporum. Mol Plant Pathol, 10(3), 311-324. https://doi.org/10.1111/j.1364-3703.2009.00538.x

Miklis, M., Consonni, C., Bhat, R. A., Lipka, V., Schulze-Lefert, P., & Panstruga, R. (2007). Barley MLO modulates actin-dependent and actin-independent antifungal defense pathways at the cell periphery. Plant Physiol, 144(2), 1132-1143. https://doi.org/10.1104/pp.107.098897

Mine, A., Berens, M. L., Nobori, T., Anver, S., Fukumoto, K., Winkelmuller, T. M., Takeda, A., Becker, D., & Tsuda, K. (2017). Pathogen exploitation of an abscisic acid- and jasmonate-inducible MAPK phosphatase and its interception by Arabidopsis immunity. Proc Natl Acad Sci U S A, 114(28), 7456-7461. https://doi.org/10.1073/pnas.1702613114

Mintz, S. W. (1985). Sweetness and power : the place of sugar in modern history. Viking.

Mittler, R., Kim, Y., Song, L., Coutu, J., Coutu, A., Ciftci-Yilmaz, S., Lee, H., Stevenson, B., & Zhu, J. K. (2006). Gain- and loss-of-function mutations in Zat10 enhance the tolerance of plants to abiotic stress. FEBS Lett, 580(28-29), 6537-6542. https://doi.org/10.1016/j.febslet.2006.11.002

Mondal, H. A., Louis, J., Archer, L., Patel, M., Nalam, V. J., Sarowar, S., Sivapalan, V., Root, D. D., & Shah, J. (2018). Arabidopsis ACTIN-DEPOLYMERIZING FACTOR3 Is Required for Controlling Aphid Feeding from the Phloem. Plant Physiol, 176(1), 879-890. https://doi.org/10.1104/pp.17.01438

Murata, M., Richardson, J. S., & Sussman, J. L. (1985). Simultaneous comparison of three protein sequences. Proc Natl Acad Sci U S A, 82(10), 3073-3077. https://doi.org/10.1073/pnas.82.10.3073

Nakashima, K., Yamaguchi-Shinozaki, K., & Shinozaki, K. (2014). The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress

responses including drought, cold, and heat. Front Plant Sci, 5, 170. https://doi.org/10.3389/fpls.2014.00170

Nan, Q., Qian, D., Niu, Y., He, Y., Tong, S., Niu, Z., Ma, J., Yang, Y., An, L., Wan, D., & Xiang, Y. (2017). Plant Actin-Depolymerizing Factors Possess Opposing Biochemical Properties Arising from Key Amino Acid Changes throughout Evolution. Plant Cell, 29(2), 395-408. https://doi.org/10.1105/tpc.16.00690

Nations., U. (2023). Available online: https://www.un.org.

Ngaki, M. N., Wang, B., Sahu, B. B., Srivastava, S. K., Farooqi, M. S., Kambakam, S., Swaminathan, S., & Bhattacharyya, M. K. (2016). Tanscriptomic Study of the Soybean-Fusarium virguliforme Interaction Revealed a Novel Ankyrin-Repeat Containing Defense Gene, Expression of Whose during Infection Led to Enhanced Resistance to the Fungal Pathogen in Transgenic Soybean Plants. PLoS One, 11(10), e0163106. https://doi.org/10.1371/journal.pone.0163106

Noctor, G., Mhamdi, A., & Foyer, C. H. (2014). The roles of reactive oxygen metabolism in drought: not so cut and dried. Plant Physiol, 164(4), 1636-1648. https://doi.org/10.1104/pp.113.233478

O'Connell, R. J., Thon, M. R., Hacquard, S., Amyotte, S. G., Kleemann, J., Torres, M. F., Damm, U., Buiate, E. A., Epstein, L., Alkan, N., Altmuller, J., Alvarado-Balderrama, L., Bauser, C. A., Becker, C., Birren, B. W., Chen, Z., Choi, J., Crouch, J. A., Duvick, J. P., . . . Vaillancourt, L. J. (2012). Lifestyle transitions in plant pathogenic Colletotrichum fungi deciphered by genome and transcriptome analyses. Nat Genet, 44(9), 1060-1065. https://doi.org/10.1038/ng.2372

Okonechnikov, K., Conesa, A., & Garcia-Alcalde, F. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics, 32(2), 292-294. https://doi.org/10.1093/bioinformatics/btv566

Paabo, S. (1989). Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci U S A, 86(6), 1939-1943. https://doi.org/10.1073/pnas.86.6.1939

Palmgren, M. G. (1990). An H-ATPase Assay: Proton Pumping and ATPase Activity Determined Simultaneously in the Same Sample. Plant Physiol, 94(3), 882-886. https://doi.org/10.1104/pp.94.3.882

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods, 14(4), 417-419. https://doi.org/10.1038/nmeth.4197

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., & Reich, D. (2012). Ancient admixture in human history. Genetics, 192(3), 1065-1093. https://doi.org/10.1534/genetics.112.145037

Pavlidis, P., Jensen, J. D., Stephan, W., & Stamatakis, A. (2012). A critical assessment of storytelling: gene ontology categories and the importance of validating genomic scans. Molecular Biology and Evolution, 29(10), 3237-3248. https://doi.org/10.1093/molbev/mss136

Pavlov, D., Muhlrad, A., Cooper, J., Wear, M., & Reisler, E. (2007). Actin filament severing by cofilin. J Mol Biol, 365(5), 1350-1358. https://doi.org/10.1016/j.jmb.2006.10.102

Pearsall, D. M. (1978). Phytolith analysis of archeological soils: evidence for maize cultivation in formative ecuador. Science, 199(4325), 177-178. https://doi.org/10.1126/science.199.4325.177

Pell, E. J., Miller, J.D., and Bielenberg, D.G. . (2004). 20 effects of airborne pollutants. In Plant Cell Death Processes, L.D. Noodén, ed (San Diego: Academic Press), pp. , 295-306.

Peng, S. Q., & Huang, D. F. (2006). [Expression of an Arabidopsis actin-depolymerizing factor 4 gene (AtADF4) in tobacco causes morphological change of plants]. Zhi Wu Sheng Li Yu Fen Zi Sheng Wu Xue Xue Bao, 32(1), 52-56. https://www.ncbi.nlm.nih.gov/pubmed/16477131

Petr, M., Vernot, B., & Kelso, J. (2019). admixr-R package for reproducible analyses using ADMIXTOOLS. Bioinformatics, 35(17), 3194-3195. https://doi.org/10.1093/bioinformatics/btz030

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet, 8(11), e1002967. https://doi.org/10.1371/journal.pgen.1002967

Podzimska-Sroka, D., O'Shea, C., Gregersen, P. L., & Skriver, K. (2015). NAC Transcription Factors in Senescence: From Molecular Structure to Function in Crops. Plants (Basel), 4(3), 412-448. https://doi.org/10.3390/plants4030412

Pollard, T. D., & Cooper, J. A. (2009). Actin, a central player in cell shape and movement. Science, 326(5957), 1208-1212. https://doi.org/10.1126/science.1175862

Porter, K., Shimono, M., Tian, M., & Day, B. (2012). Arabidopsis Actin-Depolymerizing Factor-4 links pathogen perception, defense activation and transcription to cytoskeletal dynamics. PLoS Pathog, 8(11), e1003006. https://doi.org/10.1371/journal.ppat.1003006

Prufer, K., Stenzel, U., Hofreiter, M., Paabo, S., Kelso, J., & Green, R. E. (2010). Computational challenges in the analysis of ancient DNA. Genome Biol, 11(5), R47. https://doi.org/10.1186/gb-2010-11-5-r47

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet, 81(3), 559-575. https://doi.org/10.1086/519795

Pusztahelyi, T., Holb, I.J., and Pócsi, I. (2016). Plant-fungal interactions: special secondary metabolites of the biotrophic, necrotrophic, and other specific interactions. In: Mérillon, J.M., Ramawat, K. (eds) Fungal Metabolites. Reference Series in Phytochemistry. Springer, Cham. pp. , 1-58.

Qi, J., Song, C. P., Wang, B., Zhou, J., Kangasjarvi, J., Zhu, J. K., & Gong, Z. (2018). Reactive oxygen species signaling and stomatal movement in plant responses to drought stress and pathogen attack. J Integr Plant Biol, 60(9), 805-826. https://doi.org/10.1111/jipb.12654

Qian, D., Zhang, Z., He, J., Zhang, P., Ou, X., Li, T., Niu, L., Nan, Q., Niu, Y., He, W., An, L., Jiang, K., & Xiang, Y. (2019). Arabidopsis ADF5 promotes stomatal closure by regulating actin cytoskeleton remodeling in response to ABA and drought stress. J Exp Bot, 70(2), 435-446. https://doi.org/10.1093/jxb/ery385

Ranum, P., Pena-Rosas, J. P., & Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. Ann N Y Acad Sci, 1312, 105-112. https://doi.org/10.1111/nyas.12396

Rebourg, C., Chastanet, M., Gouesnard, B., Welcker, C., Dubreuil, P., & Charcosset, A. (2003). Maize introduction into Europe: the history reviewed in the light of molecular data. Theoretical and Applied Genetics, 106(5), 895-903. https://doi.org/10.1007/s00122-002-1140-9

Rejeb, I. B., Pastor, V., & Mauch-Mani, B. (2014). Plant Responses to Simultaneous Biotic and Abiotic Stress: Molecular Mechanisms. Plants (Basel), 3(4), 458-475. https://doi.org/10.3390/plants3040458

Ruzicka, D. R., Kandasamy, M. K., McKinney, E. C., Burgos-Rivera, B., & Meagher, R. B. (2007). The ancient subclasses of Arabidopsis Actin Depolymerizing Factor genes exhibit novel and differential expression. Plant J, 52(3), 460-472. https://doi.org/10.1111/j.1365-313X.2007.03257.x

Ryu, K. H., Huang, L., Kang, H. M., & Schiefelbein, J. (2019). Single-Cell RNA Sequencing Resolves Molecular Relationships Among Individual Plant Cells. Plant Physiol, 179(4), 1444-1456. https://doi.org/10.1104/pp.18.01482

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E. H., McCarroll, S. A., Gaudet, R., Schaffner, S. F., Lander, E. S., International HapMap, C., Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A., . . . Stewart, J. (2007). Genome-wide detection and characterization of positive selection in human populations. Nature, 449(7164), 913-918. https://doi.org/10.1038/nature06250

Saga, H., Ogawa, T., Kai, K., Suzuki, H., Ogata, Y., Sakurai, N., Shibata, D., & Ohta, D. (2012). Identification and characterization of ANAC042, a transcription factor family gene involved in the regulation of camalexin biosynthesis in Arabidopsis. Mol Plant Microbe Interact, 25(5), 684-696. https://doi.org/10.1094/MPMI-09-11-0244

Saijo, Y., & Loo, E. P. (2020). Plant immunity in signal integration between biotic and abiotic stress responses. New Phytol, 225(1), 87-104. https://doi.org/10.1111/nph.15989

Salleh, F. M., Evans, K., Goodall, B., Machin, H., Mowla, S. B., Mur, L. A., Runions, J., Theodoulou, F. L., Foyer, C. H., & Rogers, H. J. (2012). A novel function for a redox-related LEA protein (SAG21/AtLEA5) in root development and biotic stress responses. Plant Cell Environ, 35(2), 418-429. https://doi.org/10.1111/j.1365-3040.2011.02394.x

Savory, E. A., Adhikari, B. N., Hamilton, J. P., Vaillancourt, B., Buell, C. R., & Day, B. (2012). mRNA-Seq analysis of the Pseudoperonospora cubensis transcriptome during cucumber (Cucumis sativus L.) infection. PLoS One, 7(4), e35796. https://doi.org/10.1371/journal.pone.0035796

Schafer, P., Khatabi, B., & Kogel, K. H. (2007). Root cell death and systemic effects of Piriformospora indica: a study on mutualism. FEMS Microbiol Lett, 275(1), 1-7. https://doi.org/10.1111/j.1574-6968.2007.00848.x

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T. A., Minx, P., Reily, A. D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., . . . Wilson, R. K. (2009). The B73 maize genome: complexity, diversity, and dynamics. Science, 326(5956), 1112-1115. https://doi.org/10.1126/science.1178534

Schommer, C., Palatnik, J. F., Aggarwal, P., Chetelat, A., Cubas, P., Farmer, E. E., Nath, U., & Weigel, D. (2008). Control of jasmonate biosynthesis and senescence by miR319 targets. PLoS Biol, 6(9), e230. https://doi.org/10.1371/journal.pbio.0060230

Schubert, M., Ginolhac, A., Lindgreen, S., Thompson, J. F., Al-Rasheid, K. A., Willerslev, E., Krogh, A., & Orlando, L. (2012). Improving ancient DNA read mapping against

modern reference genomes. BMC Genomics, 13, 178. https://doi.org/10.1186/1471-2164-13-178

Seo, P. J., & Mas, P. (2015). STRESSing the role of the plant circadian clock. Trends Plant Sci, 20(4), 230-237. https://doi.org/10.1016/j.tplants.2015.01.001

Shahan, R., Hsu, C. W., Nolan, T. M., Cole, B. J., Taylor, I. W., Greenstreet, L., Zhang, S., Afanassiev, A., Vlot, A. H. C., Schiebinger, G., Benfey, P. N., & Ohler, U. (2022). A single-cell Arabidopsis root atlas reveals developmental trajectories in wild-type and cell identity mutants. Dev Cell, 57(4), 543-560 e549. https://doi.org/10.1016/j.devcel.2022.01.008

Shevchenko, A., Wilm, M., Vorm, O., & Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem, 68(5), 850-858. https://doi.org/10.1021/ac950914h

Singh, S., Cornilescu, C. C., Tyler, R. C., Cornilescu, G., Tonelli, M., Lee, M. S., & Markley, J. L. (2005). Solution structure of a late embryogenesis abundant protein (LEA14) from Arabidopsis thaliana, a cellular stress-related protein. Protein Sci, 14(10), 2601-2609. https://doi.org/10.1110/ps.051579205

Sirichandra, C., Gu, D., Hu, H. C., Davanture, M., Lee, S., Djaoui, M., Valot, B., Zivy, M., Leung, J., Merlot, S., & Kwak, J. M. (2009). Phosphorylation of the Arabidopsis AtrbohF NADPH oxidase by OST1 protein kinase. FEBS Lett, 583(18), 2982-2986. https://doi.org/10.1016/j.febslet.2009.08.033

Sirichandra, C., Wasilewska, A., Vlad, F., Valon, C., & Leung, J. (2009). The guard cell as a single-cell model towards understanding drought tolerance and abscisic acid action. J Exp Bot, 60(5), 1439-1463. https://doi.org/10.1093/jxb/ern340

Sisic, A., Bacanovic-Sisic, J., Al-Hatmi, A. M. S., Karlovsky, P., Ahmed, S. A., Maier, W., de Hoog, G. S., & Finckh, M. R. (2018). The 'forma specialis' issue in Fusarium: A case study in Fusarium solani f. sp. pisi. Sci Rep, 8(1), 1252. https://doi.org/10.1038/s41598-018-19779-z

Song, W. M., & Zhang, B. (2015). Multiscale Embedded Gene Co-expression Network Analysis. PLoS Comput Biol, 11(11), e1004574. https://doi.org/10.1371/journal.pcbi.1004574

Staden, R. (1979). A strategy of DNA sequencing employing computer programs. Nucleic Acids Res, 6(7), 2601-2610. https://doi.org/10.1093/nar/6.7.2601

Staiger, C. J. (2000). Signaling to the Actin Cytoskeleton in Plants. Annu Rev Plant Physiol Plant Mol Biol, 51, 257-288. https://doi.org/10.1146/annurev.arplant.51.1.257

Staiger, C. J., & Blanchoin, L. (2006). Actin dynamics: old friends with new stories. Curr Opin Plant Biol, 9(6), 554-562. https://doi.org/10.1016/j.pbi.2006.09.013

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30(9), 1312-1313. https://doi.org/10.1093/bioinformatics/btu033

Stiller, M., Green, R. E., Ronan, M., Simons, J. F., Du, L., He, W., Egholm, M., Rothberg, J. M., Keates, S. G., Ovodov, N. D., Antipina, E. E., Baryshnikov, G. F., Kuzmin, Y. V., Vasilevski, A. A., Wuenschell, G. E., Termini, J., Hofreiter, M., Jaenicke-Despres, V., & Paabo, S. (2006). Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA. Proc Natl Acad Sci U S A, 103(37), 13578-13584. https://doi.org/10.1073/pnas.0605327103

Su, J., Yang, L., Zhu, Q., Wu, H., He, Y., Liu, Y., Xu, J., Jiang, D., & Zhang, S. (2018). Active photosynthetic inhibition mediated by MPK3/MPK6 is critical to effector-triggered immunity. PLoS Biol, 16(5), e2004122. https://doi.org/10.1371/journal.pbio.2004122

Subramanian, S., Ramasamy, U., & Chen, D. (2019). VCF2PopTree: a client-side software to construct population phylogeny from genome-wide SNPs. PeerJ, 7, e8213. https://doi.org/10.7717/peerj.8213

Sun, C., Liang, W., Yan, K., Xu, D., Qin, T., Fiaz, S., Kear, P., Bi, Z., Liu, Y., Liu, Z., Zhang, J., & Bai, J. (2022). Expression of Potato StDRO1 in Arabidopsis Alters Root Architecture and Drought Tolerance. Front Plant Sci, 13, 836063. https://doi.org/10.3389/fpls.2022.836063

Sun, T., Li, S., & Ren, H. (2013). Profilin as a regulator of the membrane-actin cytoskeleton interface in plant cells. Front Plant Sci, 4, 512. https://doi.org/10.3389/fpls.2013.00512

Sussmilch, F. C., Atallah, N. M., Brodribb, T. J., Banks, J. A., & McAdam, S. A. M. (2017). Abscisic acid (ABA) and key proteins in its perception and signaling pathways are ancient, but their roles have changed through time. Plant Signal Behav, 12(9), e1365210. https://doi.org/10.1080/15592324.2017.1365210

Sussmilch, F. C., & McAdam, S. A. M. (2017). Surviving a Dry Future: Abscisic Acid (ABA)-Mediated Plant Mechanisms for Conserving Water under Low Humidity. Plants (Basel), 6(4). https://doi.org/10.3390/plants6040054

Swarts, K., Gutaker, R. M., Benz, B., Blake, M., Bukowski, R., Holland, J., Kruse-Peeples, M., Lepak, N., Prim, L., Romay, M. C., Ross-Ibarra, J., Sanchez-Gonzalez, J. J., Schmidt, C., Schuenemann, V. J., Krause, J., Matson, R. G., Weigel, D., Buckler, E. S., & Burbano, H. A. (2017). Genomic estimation of complex traits reveals

ancient maize adaptation to temperate North America. Science, 357(6350), 512-515. https://doi.org/10.1126/science.aam9425

Szymanski, D. B., & Cosgrove, D. J. (2009). Dynamic coordination of cytoskeletal and cell wall systems during plant cell morphogenesis. Curr Biol, 19(17), R800-811. https://doi.org/10.1016/j.cub.2009.07.056

Team., R. D. C. (2010). R: a language and environment for statistical computing. . In Vienna, Austria: R Foundation for Statistical Computing.

Thirumalaikumar, V. P., Devkar, V., Mehterov, N., Ali, S., Ozgur, R., Turkan, I., Mueller-Roeber, B., & Balazadeh, S. (2018). NAC transcription factor JUNGBRUNNEN1 enhances drought tolerance in tomato. Plant Biotechnol J, 16(2), 354-366. https://doi.org/10.1111/pbi.12776

Tholl, S., Moreau, F., Hoffmann, C., Arumugam, K., Dieterle, M., Moes, D., Neumann, K., Steinmetz, A., & Thomas, C. (2011). Arabidopsis actin-depolymerizing factors (ADFs) 1 and 9 display antagonist activities. FEBS Lett, 585(12), 1821-1827. https://doi.org/10.1016/j.febslet.2011.05.019

Tian, M., Chaudhry, F., Ruzicka, D. R., Meagher, R. B., Staiger, C. J., & Day, B. (2009). Arabidopsis actin-depolymerizing factor AtADF4 mediates defense signal transduction triggered by the Pseudomonas syringae effector AvrPphB. Plant Physiol, 150(2), 815-824. https://doi.org/10.1104/pp.109.137604

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W., & Su, Z. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res, 45(W1), W122-W129. https://doi.org/10.1093/nar/gkx382

Uga, Y., Sugimoto, K., Ogawa, S., Rane, J., Ishitani, M., Hara, N., Kitomi, Y., Inukai, Y., Ono, K., Kanno, N., Inoue, H., Takehisa, H., Motoyama, R., Nagamura, Y., Wu, J., Matsumoto, T., Takai, T., Okuno, K., & Yano, M. (2013). Control of root system architecture by DEEPER ROOTING 1 increases rice yield under drought conditions. Nat Genet, 45(9), 1097-1102. https://doi.org/10.1038/ng.2725

Umezawa, T., Fujita, M., Fujita, Y., Yamaguchi-Shinozaki, K., & Shinozaki, K. (2006). Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. Curr Opin Biotechnol, 17(2), 113-122. https://doi.org/10.1016/j.copbio.2006.02.002

V. N. Njiti, M. A. S., R. J. Suttner, M. E. Schmidt, P. T. Gibson. (1996). Soybean Response to Sudden Death Syndrome: Inheritance Influenced by Cyst Nematode Resistance in Pyramid × Douglas Progenies. Crop Sci., 36, 1165-1170.

Vallebueno-Estrada, M., Rodriguez-Arevalo, I., Rougon-Cardoso, A., Martinez Gonzalez, J., Garcia Cook, A., Montiel, R., & Vielle-Calzada, J. P. (2016). The earliest maize

from San Marcos Tehuacan is a partial domesticate with genomic evidence of inbreeding. Proc Natl Acad Sci U S A, 113(49), 14151-14156. https://doi.org/10.1073/pnas.1609701113

van der Does, H. C., & Rep, M. (2017). Adaptation to the Host Environment by Plant-Pathogenic Fungi. Annu Rev Phytopathol, 55, 427-450. https://doi.org/10.1146/annurev-phyto-080516-035551

van Gisbergen, P. A., & Bezanilla, M. (2013). Plant formins: membrane anchors for actin polymerization. Trends Cell Biol, 23(5), 227-233. https://doi.org/10.1016/j.tcb.2012.12.001

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. Proc Natl Acad Sci U S A, 108(3), 1088-1092. https://doi.org/10.1073/pnas.1013011108

Vargas, W. A., Martin, J. M., Rech, G. E., Rivera, L. P., Benito, E. P., Diaz-Minguez, J. M., Thon, M. R., & Sukno, S. A. (2012). Plant defense mechanisms are activated during biotrophic and necrotrophic development of Colletotricum graminicola in maize. Plant Physiol, 158(3), 1342-1358. https://doi.org/10.1104/pp.111.190397

Velasquez, A. C., Oney, M., Huot, B., Xu, S., & He, S. Y. (2017). Diverse mechanisms of resistance to Pseudomonas syringae in a thousand natural accessions of Arabidopsis thaliana. New Phytol, 214(4), 1673-1687. https://doi.org/10.1111/nph.14517

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., . . . Zhu, X. (2001). The sequence of the human genome. Science, 291(5507), 1304-1351. https://doi.org/10.1126/science.1058040

Vilo, J., Brazma, A., Jonassen, I., Robinson, A., & Ukkonen, E. (2000). Mining for putative regulatory elements in the yeast genome using gene expression data. Proc Int Conf Intell Syst Mol Biol, 8, 384-394. https://www.ncbi.nlm.nih.gov/pubmed/10977099

Wang, J., Jacobs, J. L., Byrne, J. M., & Chilvers, M. I. (2015). Improved Diagnoses and Quantification of Fusarium virguliforme, Causal Agent of Soybean Sudden Death Syndrome. Phytopathology, 105(3), 378-387. https://doi.org/10.1094/PHYTO-06-14-0177-R

Wang, J., Jacobs, J. L., Roth, M. G., & Chilvers, M. I. (2019). Temporal Dynamics of Fusarium virguliforme Colonization of Soybean Roots. Plant Dis, 103(1), 19-27. https://doi.org/10.1094/PDIS-03-18-0384-RE

Wang, L., Cheng, J., Bi, S., Wang, J., Cheng, X., Liu, S., Gao, Y., Lan, Q., Shi, X., Wang, Y., Zhao, X., Qi, X., Xu, S., & Wang, C. (2023). Actin Depolymerization Factor ADF1 Regulated by MYB30 Plays an Important Role in Plant Thermal Adaptation. Int J Mol Sci, 24(6). https://doi.org/10.3390/ijms24065675

Wang, L., Josephs, E. B., Lee, K. M., Roberts, L. M., Rellan-Alvarez, R., Ross-Ibarra, J., & Hufford, M. B. (2021). Molecular Parallelism Underlies Convergent Highland Adaptation of Maize Landraces. Molecular Biology and Evolution, 38(9), 3567-3580. https://doi.org/10.1093/molbev/msab119

Wang, P., Qi, S., Wang, X., Dou, L., Jia, M. A., Mao, T., Guo, Y., & Wang, X. (2023). The OPEN STOMATA1-SPIRAL1 module regulates microtubule stability during abscisic acid-induced stomatal closure in Arabidopsis. Plant Cell, 35(1), 260-278. https://doi.org/10.1093/plcell/koac307

Wang, W., Wen, Y., Berkey, R., & Xiao, S. (2009). Specific targeting of the Arabidopsis resistance protein RPW8.2 to the interfacial membrane encasing the fungal Haustorium renders broad-spectrum resistance to powdery mildew. Plant Cell, 21(9), 2898-2913. https://doi.org/10.1105/tpc.109.067587

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., Lee, T. H., Jin, H., Marler, B., Guo, H., Kissinger, J. C., & Paterson, A. H. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res, 40(7), e49. https://doi.org/10.1093/nar/gkr1293

Wickham, H. (2009). Ggplot2 : elegant graphics for data analysis. Springer. http://bvbr.bibbvb.de:8991/F?func=service&doc_library=BVB01&doc_number=017387312&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. , https://ggplot2.tidyverse.org.

Williams, B., Kabbage, M., Kim, H. J., Britt, R., & Dickman, M. B. (2011). Tipping the balance: Sclerotinia sclerotiorum secreted oxalic acid suppresses host defenses by manipulating the host redox environment. PLoS Pathog, 7(6), e1002107. https://doi.org/10.1371/journal.ppat.1002107

Wilson, A. S., Taylor, T., Ceruti, M. C., Chavez, J. A., Reinhard, J., Grimes, V., Meier-Augenstein, W., Cartmell, L., Stern, B., Richards, M. P., Worobey, M., Barnes, I., & Gilbert, M. T. (2007). Stable isotope and DNA evidence for ritual sequences in Inca child sacrifice. Proc Natl Acad Sci U S A, 104(42), 16456-16461. https://doi.org/10.1073/pnas.0704276104

Wimalanathan, K., Friedberg, I., Andorf, C.M., and Lawrence-Dill, C.J. (2018). Corn GO annotation—methods, evaluation, and review (Corn-GAMER). . Plant Direct 2, e00052.

Wojciechowska, N., Sobieszczuk-Nowicka, E., & Bagniewska-Zadworna, A. (2018). Plant organ senescence - regulation by manifold pathways. Plant Biol (Stuttg), 20(2), 167-181. https://doi.org/10.1111/plb.12672

Wong, K. H., Tan, W. L., Kini, S. G., Xiao, T., Serra, A., Sze, S. K., & Tam, J. P. (2017). Vaccatides: Antifungal Glutamine-Rich Hevein-Like Peptides from Vaccaria hispanica. Front Plant Sci, 8, 1100. https://doi.org/10.3389/fpls.2017.01100

Woo, H. R., Kim, H. J., Nam, H. G., & Lim, P. O. (2013). Plant leaf senescence and death - regulation by multiple layers of control and implications for aging in general. J Cell Sci, 126(Pt 21), 4823-4833. https://doi.org/10.1242/jcs.109116

Wrather, S. R. K. a. J. A. (2010). Suppression of Soybean Yield Potential in the Continental United States by Plant Diseases from 2006 to 2009. Plant Health Prog., 10.1094/php-2010-1122-1001-rs.

Wright, C. A., & Beattie, G. A. (2004). Pseudomonas syringae pv. tomato cells encounter inhibitory levels of water stress during the hypersensitive response of Arabidopsis thaliana. Proc Natl Acad Sci U S A, 101(9), 3269-3274. https://doi.org/10.1073/pnas.0400461101

Wu, A., Allu, A. D., Garapati, P., Siddiqui, H., Dortay, H., Zanor, M. I., Asensi-Fabado, M. A., Munne-Bosch, S., Antonio, C., Tohge, T., Fernie, A. R., Kaufmann, K., Xue, G. P., Mueller-Roeber, B., & Balazadeh, S. (2012). JUNGBRUNNEN1, a reactive oxygen species-responsive NAC transcription factor, regulates longevity in Arabidopsis. Plant Cell, 24(2), 482-506. https://doi.org/10.1105/tpc.111.090894

Xi Yuan, H. W., Jiating Cai, Dayong Li & Fengming Song. (2019). NAC transcription factors in plant immunity. Phytopathology Research, 1, Article number: 3.

Xia, J., Zhao, Y., Burks, P., Pauly, M., and Brown, P.J. (2018). A sorghum NAC gene is associated with variation in biomass properties and yield potential. Plant Direct, 2, 10.1002/pld3.70.

Xie, Q., Frugis, G., Colgan, D., & Chua, N. H. (2000). Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. Genes Dev, 14(23), 3024-3036. https://doi.org/10.1101/gad.852200

Xiong, L., & Zhu, J. K. (2003). Regulation of abscisic acid biosynthesis. Plant Physiol, 133(1), 29-36. https://doi.org/10.1104/pp.103.025395

Xu, J., Audenaert, K., Hofte, M., & De Vleesschauwer, D. (2013). Abscisic Acid Promotes Susceptibility to the Rice Leaf Blight Pathogen Xanthomonas oryzae pv oryzae by Suppressing Salicylic Acid-Mediated Defenses. PLoS One, 8(6), e67413. https://doi.org/10.1371/journal.pone.0067413

Yamaguchi, M., Ohtani, M., Mitsuda, N., Kubo, M., Ohme-Takagi, M., Fukuda, H., & Demura, T. (2010). VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem vessel formation in Arabidopsis. Plant Cell, 22(4), 1249-1263. https://doi.org/10.1105/tpc.108.064048

Yamaguchi-Shinozaki, K., & Shinozaki, K. (2006). Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. Annu Rev Plant Biol, 57, 781-803. https://doi.org/10.1146/annurev.arplant.57.032905.105444

Yang, Z., & Ohlrogge, J. B. (2009). Turnover of fatty acids during natural senescence of Arabidopsis, Brachypodium, and switchgrass and in Arabidopsis beta-oxidation mutants. Plant Physiol, 150(4), 1981-1989. https://doi.org/10.1104/pp.109.140491

Yasuda, M., Ishikawa, A., Jikumaru, Y., Seki, M., Umezawa, T., Asami, T., Maruyama-Nakashita, A., Kudo, T., Shinozaki, K., Yoshida, S., & Nakashita, H. (2008). Antagonistic interaction between systemic acquired resistance and the abscisic acid-mediated abiotic stress response in Arabidopsis. Plant Cell, 20(6), 1678-1692. https://doi.org/10.1105/tpc.107.054296

York, L. M., Nord, E. A., & Lynch, J. P. (2013). Integration of root phenes for soil resource acquisition. Front Plant Sci, 4, 355. https://doi.org/10.3389/fpls.2013.00355

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

Yu, X., & Sun, S. (2013). Comparing a few SNP calling algorithms using low-coverage sequencing data. BMC Bioinformatics, 14, 274. https://doi.org/10.1186/1471-2105-14-274

Zhang, H., Zhu, J., Gong, Z., & Zhu, J. K. (2022). Abiotic stress responses in plants. Nat Rev Genet, 23(2), 104-119. https://doi.org/10.1038/s41576-021-00413-0

Zhang, P., Qian, D., Luo, C., Niu, Y., Li, T., Li, C., Xiang, Y., Wang, X., & Niu, Y. (2021). Arabidopsis ADF5 Acts as a Downstream Target Gene of CBFs in Response to Low-Temperature Stress. Front Cell Dev Biol, 9, 635533. https://doi.org/10.3389/fcell.2021.635533

Zhang, W., Zhao, F., Jiang, L., Chen, C., Wu, L., & Liu, Z. (2018a). Different Pathogen Defense Strategies in Arabidopsis: More than Pathogen Recognition. Cells, 7(12). https://doi.org/10.3390/cells7120252

Zhang, X., Valdes-Lopez, O., Arellano, C., Stacey, G., & Balint-Kurti, P. (2017). Genetic dissection of the maize (Zea mays L.) MAMP response. Theor Appl Genet, 130(6), 1155-1168. https://doi.org/10.1007/s00122-017-2876-6

Zhang, Y., Tian, L., Yan, D. H., & He, W. (2018b). Genome-Wide Transcriptome Analysis Reveals the Comprehensive Response of Two Susceptible Poplar Sections to Marssonina brunnea Infection. Genes (Basel), 9(3). https://doi.org/10.3390/genes9030154

Zhao, S., Jiang, Y., Zhao, Y., Huang, S., Yuan, M., Zhao, Y., & Guo, Y. (2016). CASEIN KINASE1-LIKE PROTEIN2 Regulates Actin Filament Stability and Stomatal Closure via Phosphorylation of Actin Depolymerizing Factor. Plant Cell, 28(6), 1422-1439. https://doi.org/10.1105/tpc.16.00078

Zhao, Y., Zhao, S., Mao, T., Qu, X., Cao, W., Zhang, L., Zhang, W., He, L., Li, S., Ren, S., Zhao, J., Zhu, G., Huang, S., Ye, K., Yuan, M., & Guo, Y. (2011). The plant-specific actin binding protein SCAB1 stabilizes actin filaments and regulates stomatal movement in Arabidopsis. Plant Cell, 23(6), 2314-2330. https://doi.org/10.1105/tpc.111.086546

Zhao, Z., Li, Y., Zhao, S., Zhang, J., Zhang, H., Fu, B., He, F., Zhao, M., & Liu, P. (2018). Transcriptome Analysis of Gene Expression Patterns Potentially Associated with Premature Senescence in Nicotiana tabacum L. Molecules, 23(11). https://doi.org/10.3390/molecules23112856

Zheng, Y., Xie, Y., Jiang, Y., Qu, X., & Huang, S. (2013). Arabidopsis actin-depolymerizing factor7 severs actin filaments and regulates actin cable turnover to promote normal pollen tube growth. Plant Cell, 25(9), 3405-3423. https://doi.org/10.1105/tpc.113.117820

Zhu, J., Nan, Q., Qin, T., Qian, D., Mao, T., Yuan, S., Wu, X., Niu, Y., Bai, Q., An, L., & Xiang, Y. (2017). Higher-Ordered Actin Structures Remodeled by Arabidopsis ACTIN-DEPOLYMERIZING FACTOR5 Are Important for Pollen Germination and Pollen Tube Growth. Mol Plant, 10(8), 1065-1081. https://doi.org/10.1016/j.molp.2017.06.001

APPENDIX

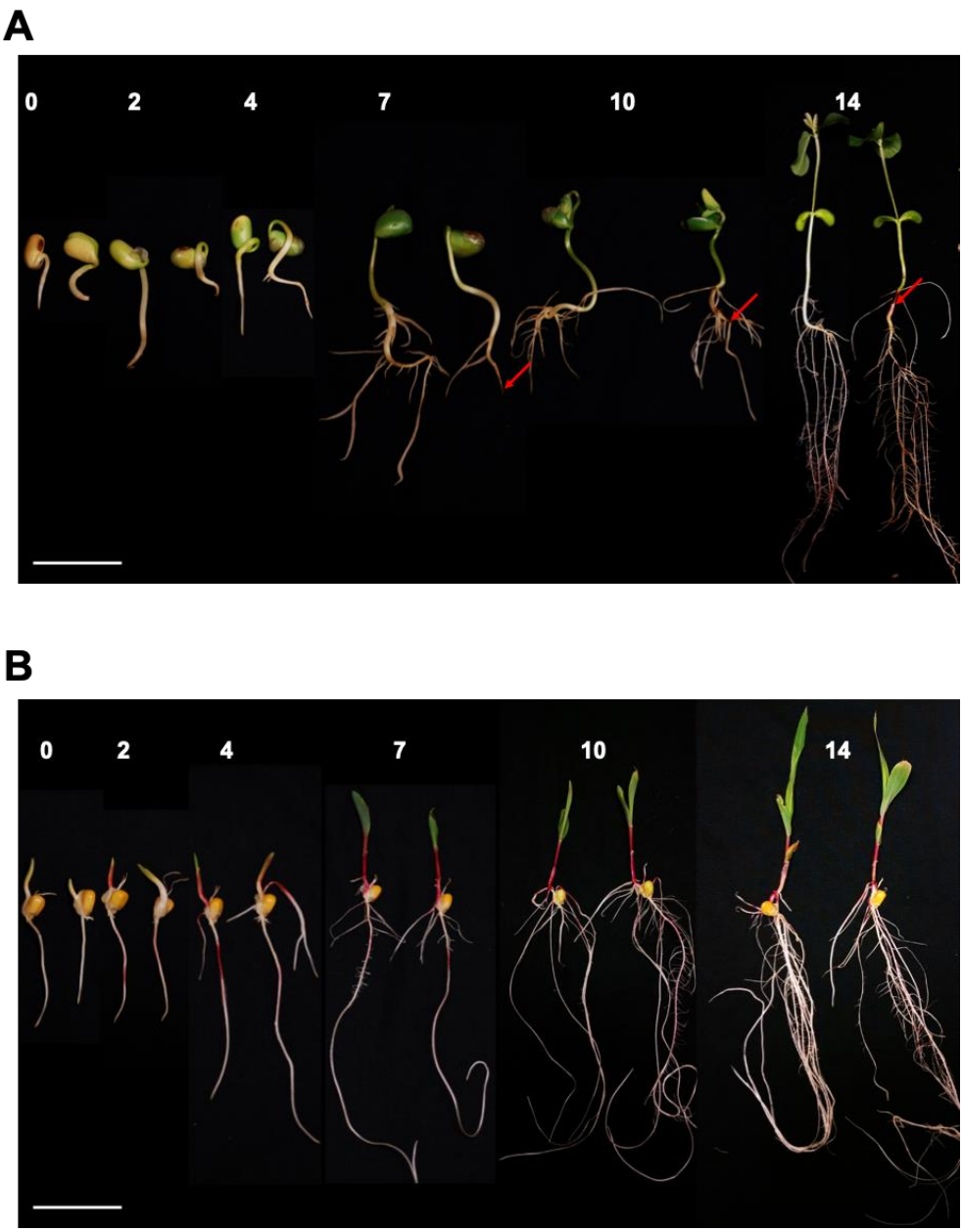

Figure AA 2-1: Plant growth and development across the colonization timecourse. (A) *Fusarium virguliforme* inoculated soybean. (B) *Fusarium virguliforme* inoculated maize. Numbers indicate days after inoculation (DAI). (Supports Figure 2-1). Arrows are shown to highlight pathogen-induced necrosis phenotypes. Bar = 4 cm.

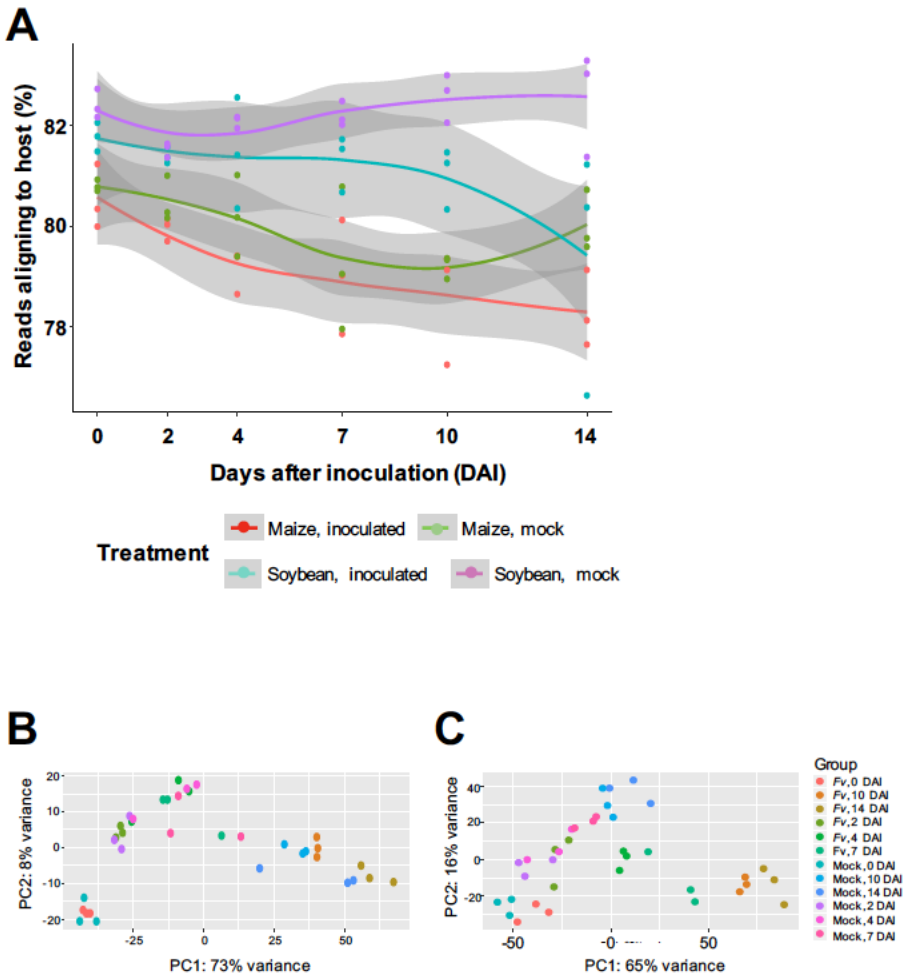

Figure AA 2-2: Read alignment and principal component analysis. (A) Percent of unique trimmed reads aligning to each respective genome within corresponding treatments of maize or soybean inoculated with *F. virguliforme* or mock (water). Each sample is indicated by a colored dot, and lines are the average of three independent biological replicates. Grey shading indicates SEM. Reads were mapped to either to Glycine_max_v2.0 or *Zea_mays*.AGPv4 genomes with HISAT2/2.1.0. Mapped reads were counted with HTSeq v. 0.6.1 and were normalized by dividing by the library size for each sample calculated by DESeq2/1.2.0. The data was then log2-transformed. The PCA was calculated in R (ver. 3.5.1) in DESeq2 (ver. 1.2.0). (B) and (C) principal component analysis (PCA) of gene expression values of either soybean (B) or maize (C) inoculated with *Fusarium virguliforme*- or mock-inoculated with water. (Supports Figure 2-2).

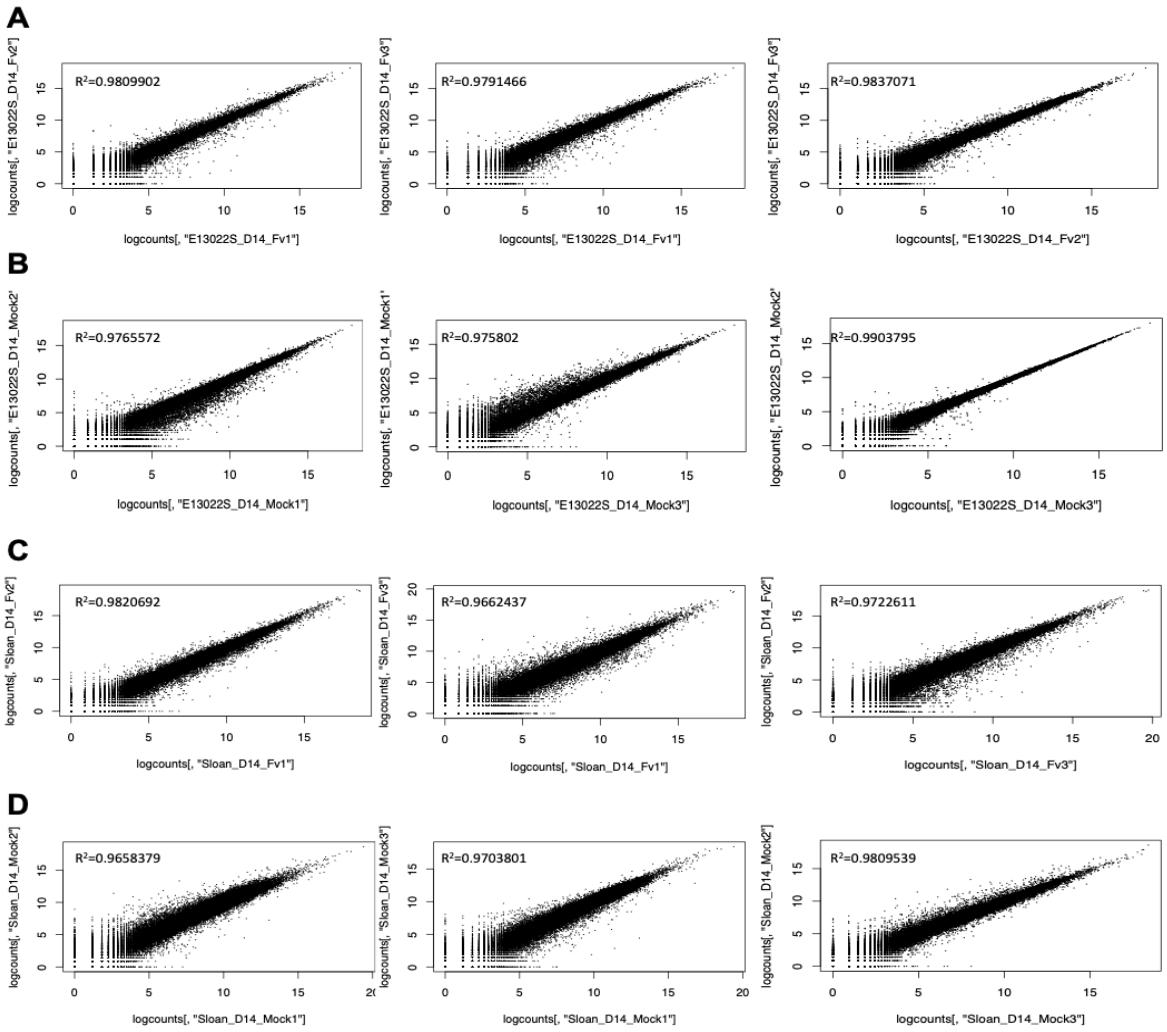

Figure AA 2-3: Correlation of gene expression values in three biological replicates from 14 DPI samples of maize inoculated with *Fusarium virguliforme* (A) or mock (B) inoculated with water, and soybean inoculated with *Fusarium virguliforme* (C) or mock (D) inoculated with water. Reads were mapped to the B73 genome with HISAT2/2.1.0. Mapped reads were counted using HTSeq v. 0.6.1 and were normalized by dividing by the library size for each sample calculated by DESeq2/1.2.0. The data were then log2-transformed with the addition of one. R2 is the correlation coefficient, calculated in R (v.3.5.1). (Supports Figure 2-2).

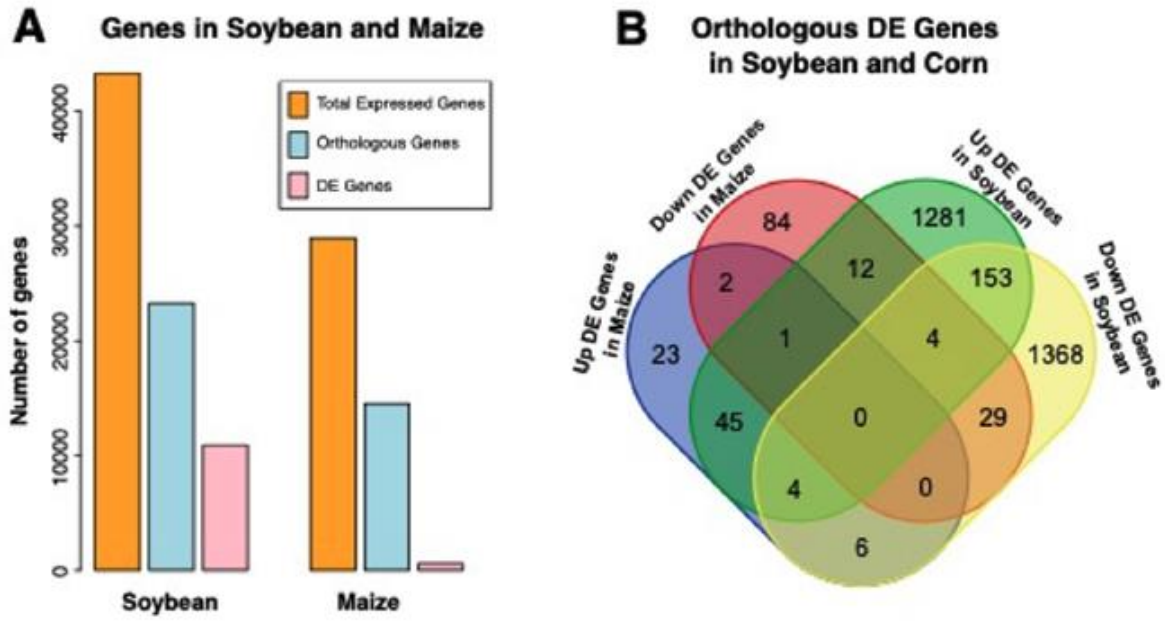

Figure AA 2-4: Graphical representation of expressed and differentially expressed genes in soybean and maize following *F. virguliforme* inoculation. (A) Number of differentially-expressed genes in soybean and maize. (B) Venn graph illustration of differentially-expressed gene (DEG) overlap between treatments in soybean and maize. (Supports Table 2-1).

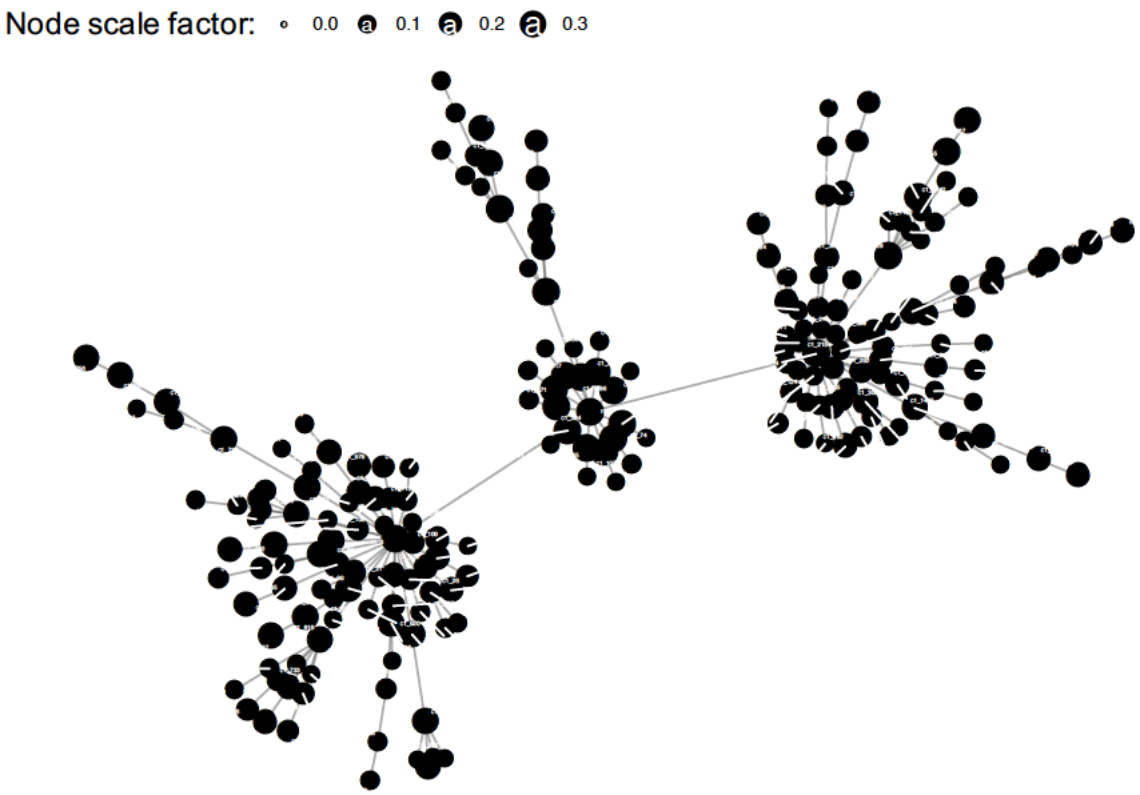

Figure AA 2-5: Module hierarchy of hubs containing defense-related genes. Module hierarchy of hubs was generated using MEGENA by setting module.pval less than 0.05. Modules that included defense-related genes are shown. Different node scale factor size means scale number to adjust node sizes. (Supports Figure 2-2).

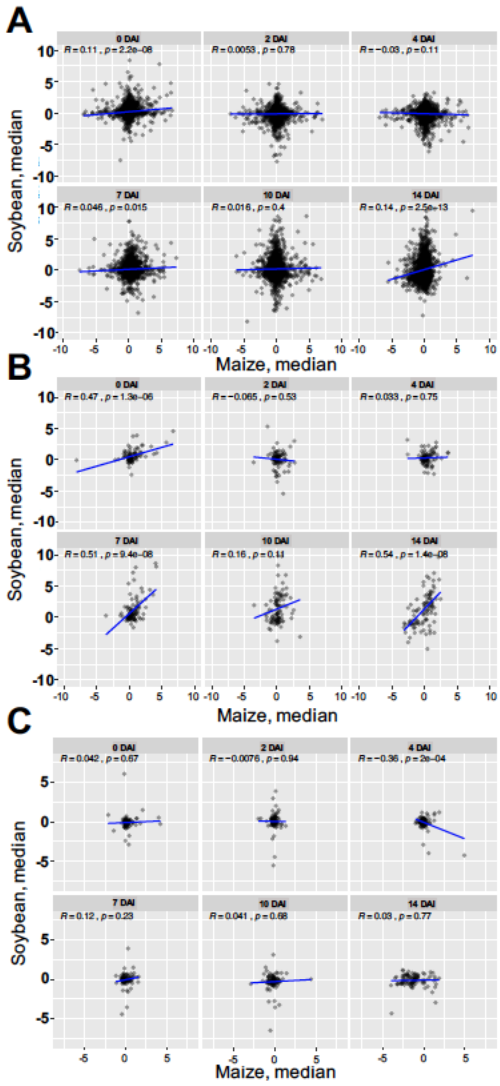

Figure AA 2-6: Correlation of median transformed log2-fold changes between mock- and *F. virguliforme*-inoculated orthogroups within soybean compared to maize at each timepoint over the colonization timecourse. (A) Orthogroups uniquely differentially regulated in soybean, (B) orthogroups differentially regulated in maize and soybean, (C) orthogroups uniquely differentially regulated only in maize. (Supports Figure 2-3).

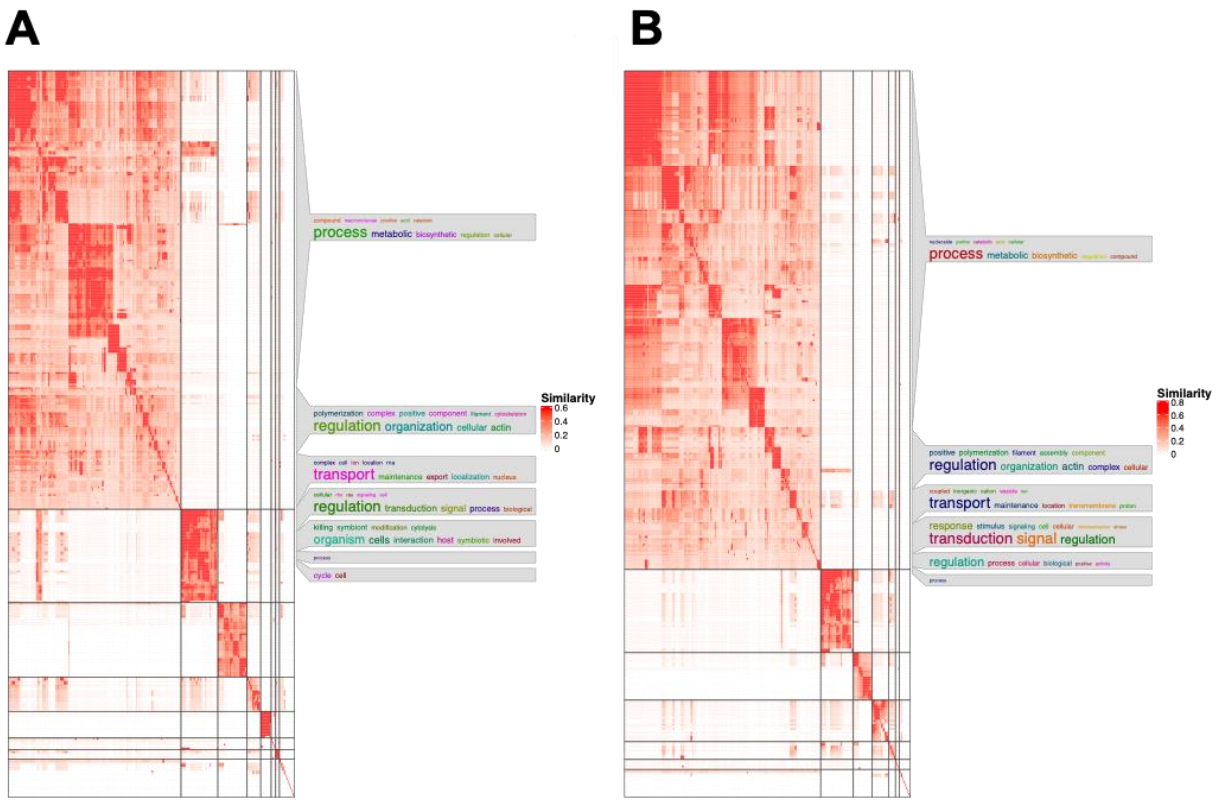

Figure AA 2-7: Gene ontology enrichment of differentially-expressed genes from *Fusarium virguliforme* in maize and soybean. The gene ontology enrichment of biological processes of *Fusarium virguliforme* involved in growth on (A) maize and (B) soybean across 0– 14 days post inoculation. The GO biological process (BP) was split into groups using cluster method binary_cut() based on the similarity matrix. Blocks located in the diagonal where GO terms are present in the same block indicate that these share similar functions. All clusters with a size less than 5 were merged into one single cluster for representation in the heatmap. The word summarizes the functions with keywords in every GO cluster, top words with the highest frequency and larger font size indicate increased frequency/representation. Font colors were generated randomly. (Supports Figure 2-2).

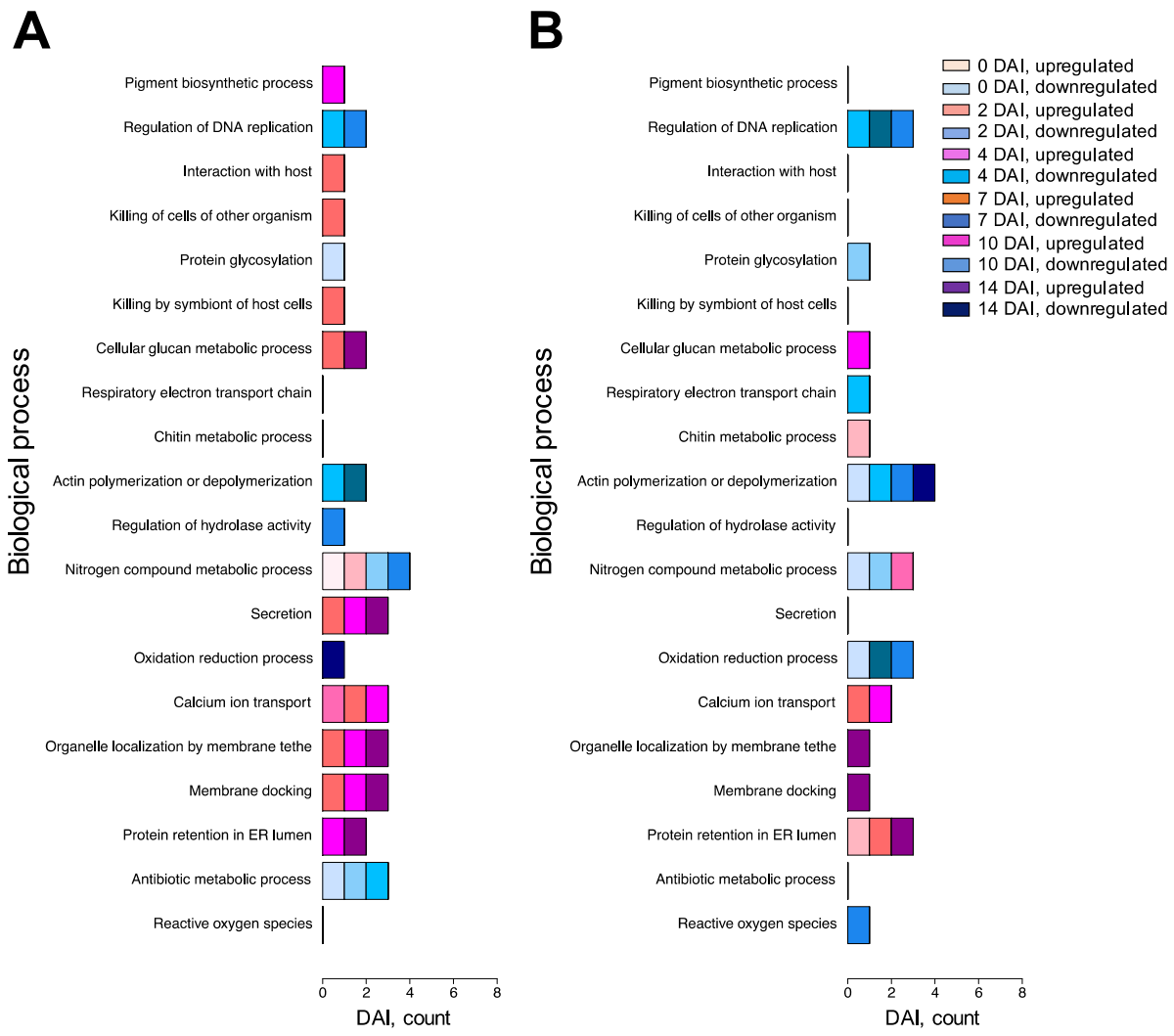

Figure AA 2-8: Biological processes of differentially expressed genes from *Fusarium virguliforme* at each day after inoculation (DAI) during growth on maize and soybean. (A) Enrichment of biological processes from *F. virguliforme* at each DAI across the timecourse on (A) maize and (B) soybean. (Supports Figure 2-2).

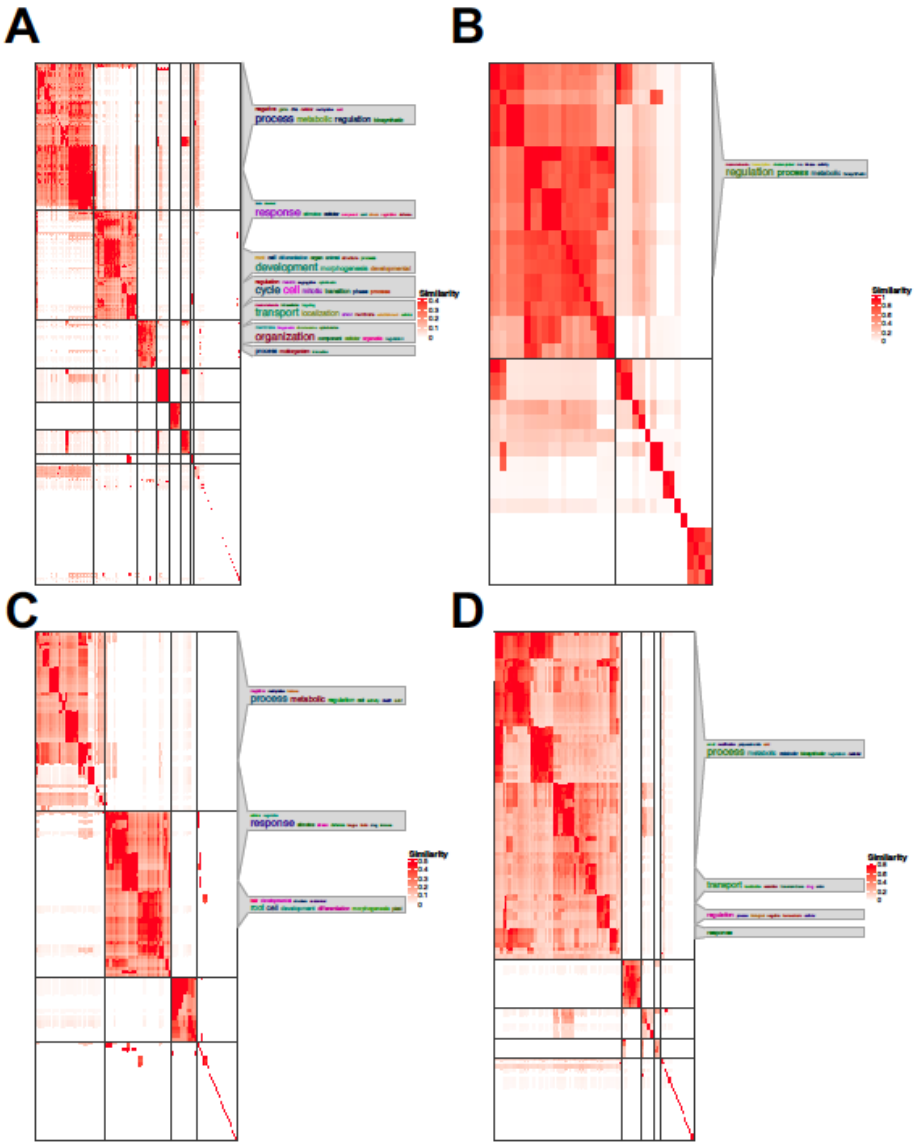

Figure AA 2-9: Gene ontology (GO) enrichment of differentially-expressed putatively orthologous and non-orthologous genes between maize and soybean. The gene ontology enrichment of orthologous genes on (A) maize and (B) soybean across 0–14 days postinoculation. The GO enrichment of non-orthologous genes on (C) maize and (D) soybean across 0–14 days post-inoculation. The GO biological process (BP) was split into groups using cluster method binary_cut() based on the similarity matrix. Blocks located in the diagonal illustrate GO terms in the same block which share similar functions. All clusters with size less than 5 were merged into one single cluster for illustration in the heatmap. GO terms summarize the functions with keywords in every GO cluster. Font size (i.e., larger) illustrates GO terms with the highest frequency. Word colors were randomly assigned. (Supports Figures 2-3 and 2-4).

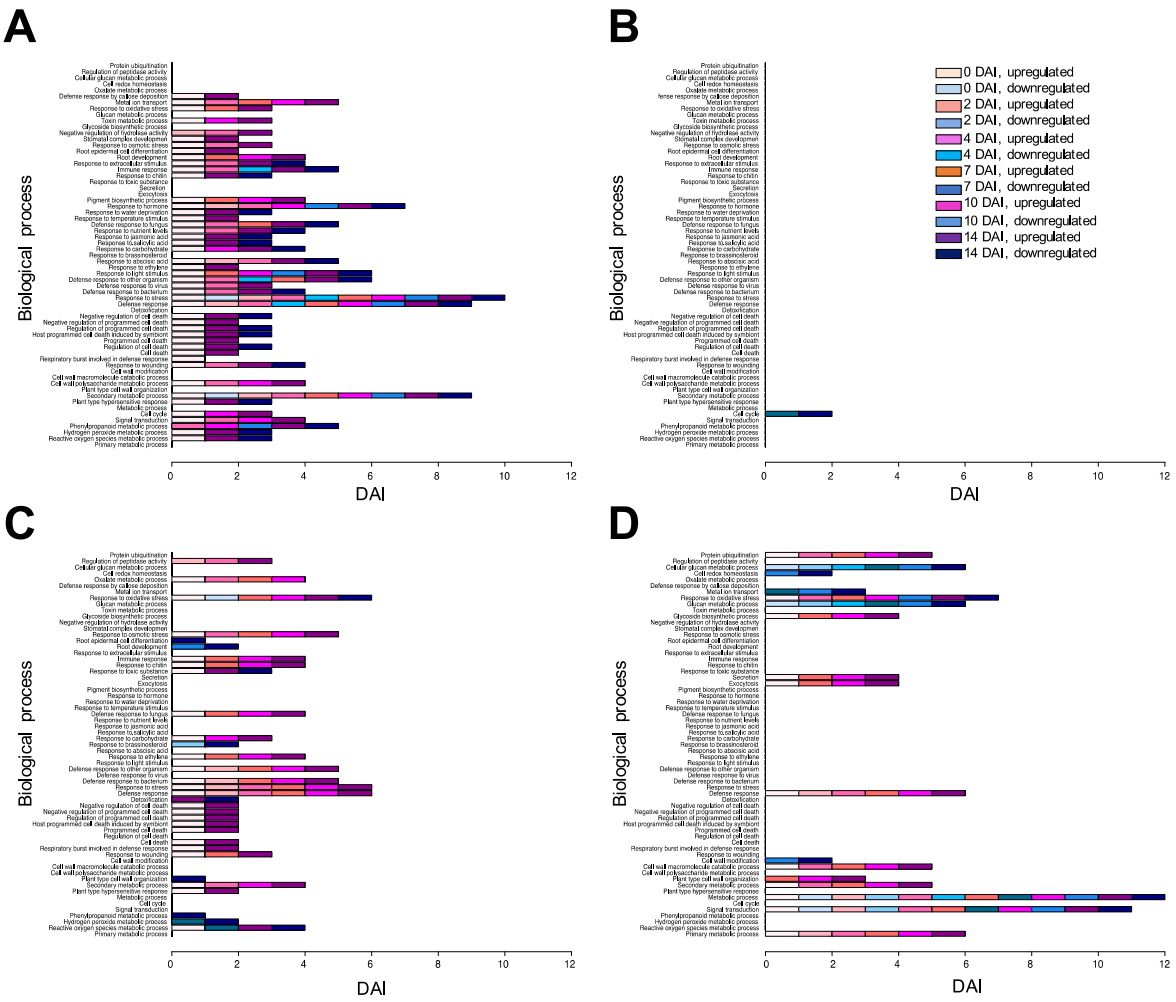

Figure AA 2-10: Biological processes of differentially-expressed genes from maize and soybean at each day after inoculation (DAI). Biological processes of orthologous genes from (A) maize and (B) soybean at each day post-inoculation time course. The biological processes of non-orthologous gene on (C) maize and (D) soybean at each day post-inoculation over the 2-week timecourse. (Supports Figure 2-3 and 2-4).

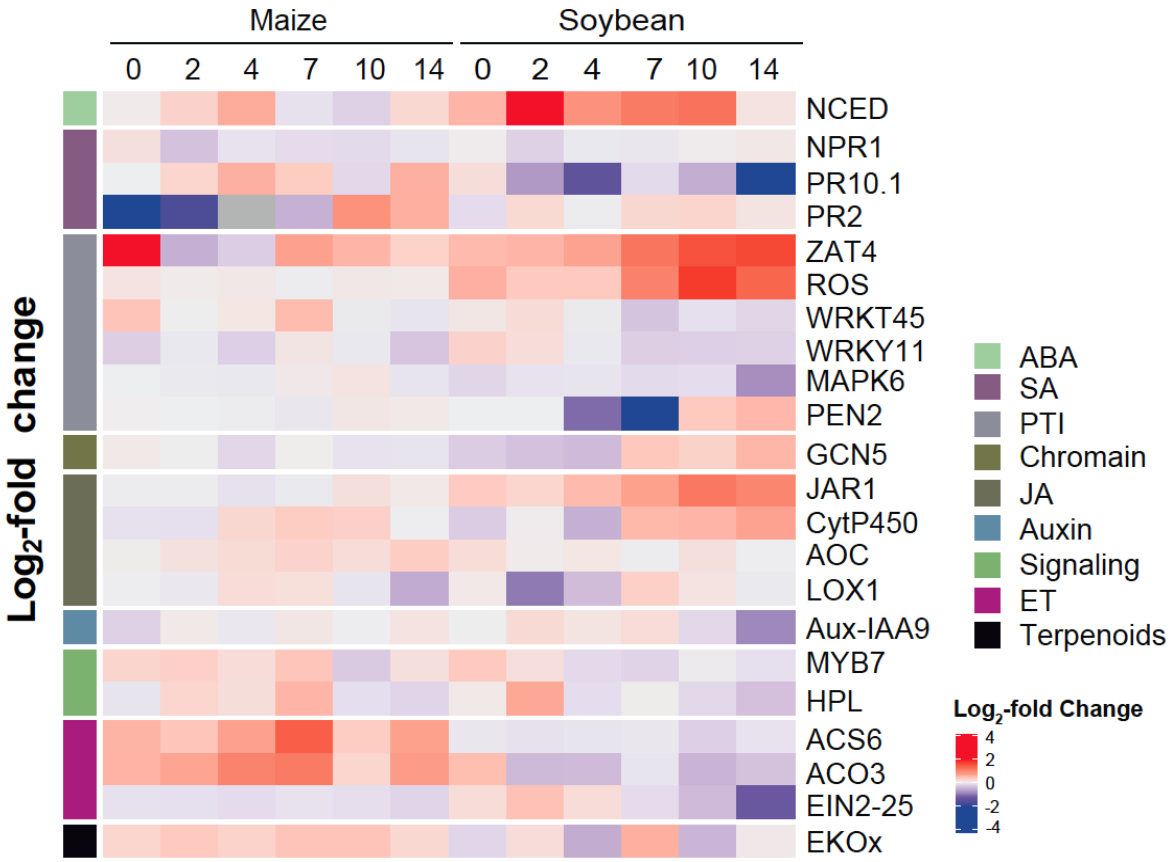

Figure AA 2-11: Expression of orthologous genes in defense-relevant processes. Each gene heatmap illustrates temporal changes from 0–14 DAI in maize and soybean. Different color sidebar indicates different defense-relevant processes. (Supports Figure 2-5).

Table AA 2-1: Significantly induced genes in maize. To be considered significantly induced, a 1-fold change or greater in expression of *F. virguliforme* treated samples compared to mock-inoculated. An adjusted P < 0.05 was required. DAI, days after inoculation.

		Maize, inoculated					
	DAI	0	2	4	7	10	14
Maize, mock	0	92					
	2		13				
	4			11			
	7				24		
	10					52	
	14						457

		Maize, inoculated; upregulated					
	DAI	0	2	4	7	10	14
Maize, mock	0	86					
	2		7				
	4			7			
	7				21		
	10					35	
	14						151

		Maize, inoculated; downregulated					
	DAI	0	2	4	7	10	14
Maize, mock	0	6					
	2		5				
	4			4			
	7				3		
	10					17	
	14						306

Table AA 2-2: Significantly induced genes in soybean. To be considered significantly induced, a 1-fold change or greater in expression between *F. virguliforme*- and mock-inoculated was required. An adjusted P < 0.05 (Benjamini & Hochberg (BH) Test adjusted p-values) was required. DAI, days after inoculation.

	Soybean, inoculated					
Soybean, mock	DAI	0	2	4	7	10	14
	0	771					
	2		156				
	4			266			
	7				1,737		
	10					4,564	
	14						9,542

	Soybean, inoculated; upregulated					
Soybean, mock	DAI	0	2	4	7	10	14
	0	724					
	2		121				
	4			232			
	7				1,538		
	10					3,256	
	14						4,957

	Soybean, inoculated; downregulated					
Soybean, mock	DAI	0	2	4	7	10	14
	0	47					
	2		35				
	4			34			
	7				199		
	10					1,308	
	14						4,585

Table AA 2-3: Gene ontology enrichment of biological processes from genes of maize significantly upregulated between *Fusarium virguliforme*-inoculated and mock-inoculated treatments across a timecourse of two weeks. P-values were generated using a Fischer Statistical Test.

GO term	Ontology	Description	Adjusted P-value
GO:0055114	P	oxidation-reduction process	0.0012
GO:0016114	P	terpenoid biosynthetic process	0.0012
GO:0006721	P	terpenoid metabolic process	0.002
GO:0006952	P	defense response	0.0026
GO:0051346	P	negative regulation of hydrolase activity	0.0098
GO:0044710	P	single-organism metabolic process	0.0098
GO:0008299	P	isoprenoid biosynthetic process	0.0098
GO:0051248	P	negative regulation of protein metabolic process	0.016
GO:0032269	P	negative regulation of cellular protein metabolic process	0.016
GO:0050832	P	defense response to fungus	0.021
GO:0006720	P	isoprenoid metabolic process	0.022
GO:0009620	P	response to fungus	0.032

Table AA 2-4: Gene ontology enrichment of biological processes in soybean significantly upregulated between *Fusarium virguliforme*-inoculated and mock-inoculated treatments over a timecourse of two weeks. P-values were generated using a Fischer Statistical Test.

GO term	Ontology	Description	Adjusted P-value
GO:0055114	P	oxidation reduction	2.40E-25
GO:0008152	P	metabolic process	1.40E-21
GO:0043687	P	post-translational protein modification	4.20E-19
GO:0006464	P	protein modification process	1.50E-18
GO:0006468	P	protein amino acid phosphorylation	1.10E-17
GO:0043412	P	macromolecule modification	7.80E-17
GO:0016310	P	phosphorylation	2.60E-16
GO:0006796	P	phosphate metabolic process	2.60E-16
GO:0006793	P	phosphorus metabolic process	2.60E-16
GO:0051704	P	multi-organism process	3.70E-11
GO:0009875	P	pollen-pistil interaction	4.40E-11
GO:0008037	P	cell recognition	4.40E-11
GO:0048544	P	recognition of pollen	4.40E-11
GO:0009856	P	pollination	4.40E-11
GO:0000003	P	reproduction	8.00E-11
GO:0022414	P	reproductive process	8.00E-11
GO:0032501	P	multicellular organismal process	4.90E-10
GO:0006979	P	response to oxidative stress	1.80E-05
GO:0016567	P	protein ubiquitination	9.60E-05
GO:0070647	P	protein modification by small protein conjugation or removal	0.00012
GO:0032446	P	protein modification by small protein conjugation	0.00012
GO:0006952	P	defense response	0.00031
GO:0007154	P	cell communication	0.00048
GO:0009607	P	response to biotic stimulus	0.0014
GO:0016052	P	carbohydrate catabolic process	0.0015
GO:0044283	P	small molecule biosynthetic process	0.0028
GO:0042221	P	response to chemical stimulus	0.0028
GO:0043436	P	oxoacid metabolic process	0.0031
GO:0044282	P	small molecule catabolic process	0.0031
GO:0006082	P	organic acid metabolic process	0.0031

Table AA 2-4 (cont'd)

GO:0019752	P	carboxylic acid metabolic process	0.0031

GO term	Ontology	Description	Adjusted P-value
GO:0044267	P	cellular protein metabolic process	0.005
GO:0016053	P	organic acid biosynthetic process	0.0072
GO:0046394	P	carboxylic acid biosynthetic process	0.0072
GO:0008610	P	lipid biosynthetic process	0.0096
GO:0044281	P	small molecule metabolic process	0.011
GO:0046164	P	alcohol catabolic process	0.012
GO:0044275	P	cellular carbohydrate catabolic process	0.017
GO:0006066	P	alcohol metabolic process	0.019
GO:0032787	P	monocarboxylic acid metabolic process	0.029
GO:0006006	P	glucose metabolic process	0.029
GO:0006631	P	fatty acid metabolic process	0.031
GO:0006720	P	isoprenoid metabolic process	0.038
GO:0008299	P	isoprenoid biosynthetic process	0.038
GO:0005996	P	monosaccharide metabolic process	0.041
GO:0019318	P	hexose metabolic process	0.045

Table AA 2-5: Significantly induced non-orthologous gene numbers in maize. To be considered significantly induced, gene expression values were required to be 1-fold change or greater in *F. virguliforme* treated compared to mock-inoculated with an adjusted P < 0.05 (Benjamini & Hochberg (BH) Test adjusted p-values) was required. DAI, days after inoculation.

		Maize, Inoculated					
	DAI	0	2	4	7	10	14
Maize, mock	0	57					
	2		9				
	4			6			
	7				18		
	10					37	
	14						291

		Maize, inoculated; upregulated					
	DAI	0	2	4	7	10	14
Maize, mock	0	54					
	2		6				
	4			4			
	7				17		
	10					28	
	14						107

		Maize, inoculated; downregulated					
	DAI	0	2	4	7	10	14
Maize, mock	0	3					
	2		3				
	4			2			
	7				1		
	10					9	
	14						184

Table AA 2-6: Significantly induced non-orthologous gene numbers in soybean. To be considered significantly induced, genes were required to be 1-fold change or greater in *Fusarium virguliforme* treated compared to mock-inoculated with an adjusted P < 0.05 (Benjamini & Hochberg (BH) Test adjusted p-values) was required. DAI, days after inoculation.

		Soybean, inoculated					
	DAI	0	2	4	7	10	14
Soybean, mock	0	450					
	2		110				
	4			180			
	7				1099		
	10					2707	
	14						5530

		Soybean, inoculated; upregulated					
	DAI	0	2	4	7	10	14
Soybean, mock	0	418					
	2		86				
	4			158			
	7				971		
	10					1899	
	14						2944

		Soybean, inoculated; downregulated					
	DAI	0	2	4	7	10	14
Soybean, mock	0	32					
	2		24				
	4			22			
	7				128		
	10					808	
	14						2586

Table AA 2-7: Gene ontology enrichment of biological processes in non-orthologous gens of maize significantly upregulated between *Fusarium virguliforme*-inoculated and mock-inoculated treatments across a timecourse of two weeks.

GO term	Ontology	Description	Adjusted P-value
GO:0050896	P	response to stimulus	5.33E-14
GO:0002679	P	respiratory burst involved in defense re...	1.78E-05
GO:0042221	P	response to chemical	1.78E-05
GO:0045730	P	respiratory burst	1.78E-05
GO:0009628	P	response to abiotic stimulus	1.78E-05
GO:0006952	P	defense response	3.82E-05
GO:0051704	P	multi-organism process	3.91E-05
GO:0006811	P	ion transport	0.00023134
GO:0015849	P	organic acid transport	0.00032902
GO:0006082	P	organic acid metabolic process	0.00061075
GO:0006950	P	response to stress	0.00061075
GO:0043436	P	oxoacid metabolic process	0.00061075
GO:0046942	P	carboxylic acid transport	0.00062641
GO:0019752	P	carboxylic acid metabolic process	0.00079318
GO:0015711	P	organic anion transport	0.00103642
GO:0032787	P	monocarboxylic acid metabolic process	0.00115672
GO:0017144	P	drug metabolic process	0.00155853
GO:1901700	P	response to oxygen-containing compound	0.00155853
GO:0002252	P	immune effector process	0.00155853
GO:0010033	P	response to organic substance	0.00203583
GO:0044283	P	small molecule biosynthetic process	0.00253817
GO:0080167	P	response to karrikin	0.0025574
GO:0006820	P	anion transport	0.00257496
GO:0001101	P	response to acid chemical	0.00431842
GO:0009607	P	response to biotic stimulus	0.00540989
GO:0044281	P	small molecule metabolic process	0.00540989
GO:0046394	P	carboxylic acid biosynthetic process	0.00603207
GO:0016053	P	organic acid biosynthetic process	0.00663717
GO:0071702	P	organic substance transport	0.00663717
GO:0055085	P	transmembrane transport	0.00764977
GO:0009698	P	phenylpropanoid metabolic process	0.00883584
GO:0007154	P	cell communication	0.01226122
GO:0051716	P	cellular response to stimulus	0.01226783

Table AA 2-7 (cont'd)

GO term	Ontology	Description	Adjusted P-value
GO:1903825	P	organic acid transmembrane transport	0.01226783
GO:0051707	P	response to other organism	0.01439472
GO:0009605	P	response to external stimulus	0.01460592
GO:0043207	P	response to external biotic stimulus	0.0185075
GO:0010038	P	response to metal ion	0.0185075
GO:0002376	P	immune system process	0.0185075
GO:1905039	P	carboxylic acid transmembrane transport	0.02166732
GO:0010035	P	response to inorganic substance	0.02291405
GO:0000165	P	MAPK cascade	0.030285
GO:0019748	P	secondary metabolic process	0.030285
GO:0023014	P	signal transduction by protein phosphory...	0.03454733
GO:0098656	P	anion transmembrane transport	0.03540565
GO:0006810	P	transport	0.037015
GO:0006721	P	terpenoid metabolic process	0.037015
GO:0034220	P	ion transmembrane transport	0.04079204
GO:0009687	P	abscisic acid metabolic process	0.0424987
GO:0043288	P	apocarotenoid metabolic process	0.0424987
GO:1902644	P	tertiary alcohol metabolic process	0.0424987
GO:0006865	P	amino acid transport	0.0424987
GO:1900424	P	regulation of defense response to bacter...	0.0424987

Table AA 2-8: Gene ontology enrichment of biological processes among non-orthologous genes of soybean significantly upregulated between *Fusarium virguliforme*- inoculated and mock-inoculated treatments across a timecourse of two weeks.

GO term	Ontology	Description	Adjusted P-value
GO:0051704	P	multi-organism process	3.70E-05
GO:0009719	P	response to endogenous stimulus	3.70E-05
GO:0009725	P	response to hormone	8.39E-05
GO:0042221	P	response to chemical	0.000192478
GO:0098542	P	defense response to other organism	0.000192478
GO:0010033	P	response to organic substance	0.000285544
GO:0071495	P	cellular response to endogenous stimulus	0.000285544
GO:0009605	P	response to external stimulus	0.000481195
GO:0042493	P	response to drug	0.000501759
GO:0032870	P	cellular response to hormone stimulus	0.000636658
GO:1901700	P	response to oxygen-containing compound	0.0007403
GO:0050896	P	response to stimulus	0.000925375
GO:0071310	P	cellular response to organic substance	0.001216207
GO:0001101	P	response to acid chemical	0.001216207
GO:0014070	P	response to organic cyclic compound	0.001233833
GO:0051707	P	response to other organism	0.002844311
GO:0050776	P	regulation of immune response	0.002844311
GO:0070887	P	cellular response to chemical stimulus	0.002844311
GO:0043207	P	response to external biotic stimulus	0.002844311
GO:0009607	P	response to biotic stimulus	0.0037015
GO:0007154	P	cell communication	0.0037015
GO:0071229	P	cellular response to acid chemical	0.0037015
GO:0006952	P	defense response	0.0037015
GO:0007165	P	signal transduction	0.0037015
GO:0009751	P	response to salicylic acid	0.00384956
GO:0044550	P	secondary metabolite biosynthetic processes	0.003986231
GO:0023052	P	signaling	0.005483704
GO:0009755	P	hormone-mediated signaling pathway	0.007931786
GO:0009966	P	regulation of signal transduction	0.008636833
GO:0019748	P	secondary metabolic process	0.008636833
	P	cellular response to organic cyclic comp...	0.01146271

Table AA 2-8 (cont'd)

GO:0071407

GO term	Ontology	Description	Adjusted P-value
GO:0023051	P	regulation of signaling	0.012193176
GO:0035556	P	intracellular signal transduction	0.012193176
GO:0006955	P	immune response	0.0133254
GO:0010646	P	regulation of cell communication	0.014806
GO:0002682	P	regulation of immune system process	0.014806
GO:0032501	P	multicellular organismal process	0.016948974
GO:0010374	P	stomatal complex development	0.018982051
GO:1901701	P	cellular response to oxygen-containing c...	0.022209
GO:1901615	P	organic hydroxy compound metabolic proce...	0.024173061
GO:0010941	P	regulation of cell death	0.024173061
GO:0032101	P	regulation of response to external stimulus	0.024173061
GO:0009595	P	detection of biotic stimulus	0.024173061
GO:0009697	P	salicylic acid biosynthetic process	0.024173061
GO:0031640	P	killing of cells of other organism	0.024173061
GO:0044364	P	disruption of cells of other organism	0.024173061
GO:0048583	P	regulation of response to stimulus	0.024173061
GO:0000165	P	MAPK cascade	0.024173061
GO:0023014	P	signal transduction by protein phosphorylation	0.0251702
GO:0009617	P	response to bacterium	0.029031373
GO:0046677	P	response to antibiotic	0.031320385
GO:0006468	P	protein phosphorylation	0.036316604
GO:0043900	P	regulation of multi-organism process	0.037015
GO:0032502	P	developmental process	0.037688
GO:0042742	P	defense response to bacterium	0.038336964
GO:0001906	P	cell killing	0.041560702
GO:0007584	P	response to nutrient	0.043396897
GO:0031347	P	regulation of defense response	0.046885667
GO:0042445	P	hormone metabolic process	0.046885667

Table AA 2-9: DNA primer sequences used for the quantification of *Fusarium virguliforme* by quantitative PCR.

Primer/Probe	Sequence (5′ - 3′)	Length (bp)
F6-3	GTAAGTGAGATTTAGTCTAGGGTAGGTGAC	30
R6	GGGACCACCTACCCTACACCTACT	25
Fv-Prb3	6FAM-TTTGGTCTAGGGTAGGCCG-MGBNFQ	19
HHIC-F	CTAGGACGAGAACTCCCACAT	21
HHIC-R	CAATCAGCGGGTGTTTCA	18
HHIC-Prb	5HEX-TCGGTGTTGATGTTTGCCATGGT-3IABkFQ	23

Data Set AAs

All the Data Set AA are deposited at koaa021_Supplementary_Data from https://doi.org/10.1093/plcell/koaa021.

Data Set AA 2-1: Read mappings of maize and soybean libraries to either reference genomes of soybean (Wm82.a2.v1) and maize (B73 RefGen_v4, AGPv4).

Data Set AA 2-2: DEGs at each timepoint of maize colonization between *F. virguliforme*- and mock-inoculated, with mock as a comparison base.

Data Set AA 2-3: DEGs at each timepoint of soybean colonization between *F. virguliforme*- and mock-inoculated, with mock as a baseline for comparison.

Data Set AA 2-4: Significant differential gene pairs between *F. virguliforme*-inoculated (corA) and mock (corB) treatments in maize.

Data Set AA 2-5: Gene ontology enrichment of genes in modules for soybean.

Data Set AA 2-6: Gene identities assigned to corresponding orthogroups for maize and soybean.

Data Set AA 2-7: DE (log2-fold change) soybean genes have orthologs in maize.

Data Set AA 2-8: DE (log2-fold change) maize genes that have orthologs in soybean.

Data Set AA 2-9: DE orthogroup median of log2- fold changes between mock and inoculated treatments within each host.

Data Set AA 2-10: Gene ontology annotation of orthogroups between maize and soybean.

Data Set AA 2-11: Gene Ontology enrichment of DE orthogroups.

Data Set AA 2-12: DE maize genes log2-fold changes without orthologues in soybean.

Data Set AA 2-13: DE soybean genes log2-fold changes that are not orthologous with maize.

Data Set AA 2-14: Gene ontology enrichment of DE upregulated nonorthologous genes in maize.

Data Set AA 2-15: Gene ontology enrichment of DE upregulated nonorthologous soybean genes.

Data Set AA 2-16: Gene ontology enrichment of DE downregulated nonorthologous maize genes.

Data Set AA 2-17: Gene ontology enrichment of DE downregulated nonorthologous soybean genes.

Data Set AA 2-18: Gene list of 182 orthologous gens that were significantly upregulated in maize.

Data Set AA 2-19: DE transcription orthogroup median log2-fold changes between inoculated and mock in each host. Differential expression is grouped by shared or unique in each host.

Data Set AA 2-20: Orthologous transcription factors.

Data Set AA 2-21: Orthologous and nonorthologous genes are involved in a variety of regulatory pathways that control senescence.

Data Set AA 2-22: TFs that positively and negatively regulate senescence-associated gene processes.

Data Set AA 2-23: Gene expression of senescence-associated genes and NAC transcription factors.

Additional Supplementary Data Modules are deposited at Dryad: https://datadryad.org/stash/share/bl9xHgsTaIvsAZjYeInMc78Y1fNz8-Px3uGBsw7C99Q

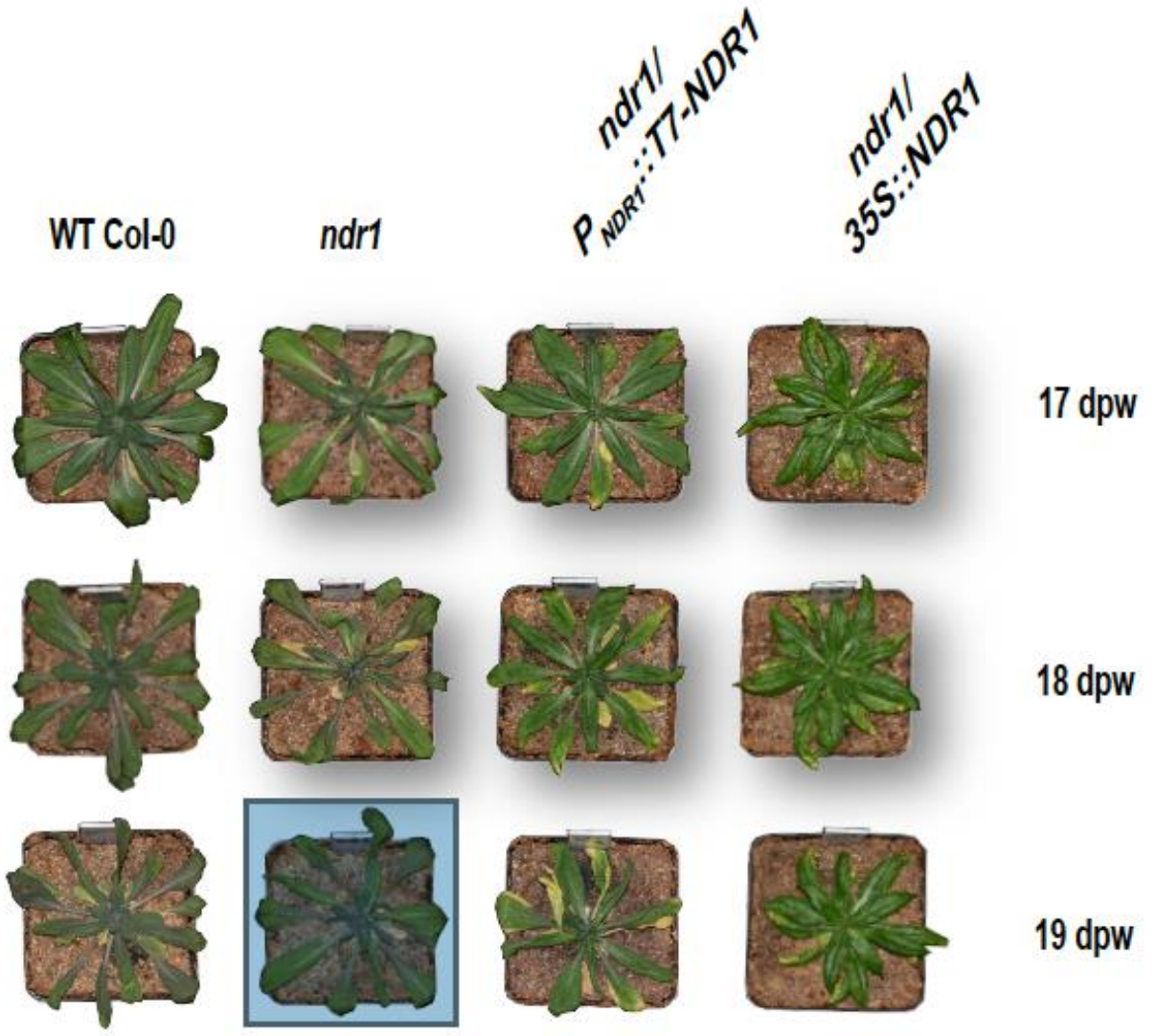

Figure AA 3-1: Extended timecourse phenotypes of drought-induced plants shown in Figure 3-1. Each photo is representative of at least 15 individual plants per timepoint.

Figure AA 3-2: Drought phenotype recovery. Photos were taken at 2 d post-watering following 25 d of water withholding.

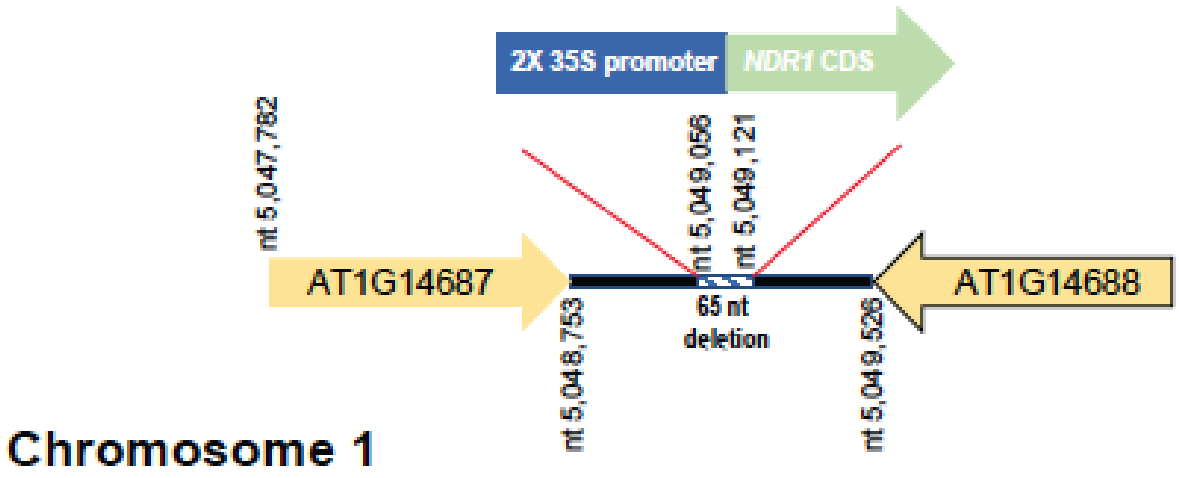

Figure AA 3-3: Genome sequence analysis of *ndr1/35S::NDR1*. Illustration of the *35S::NDR1* insertion site on chromosome 1.

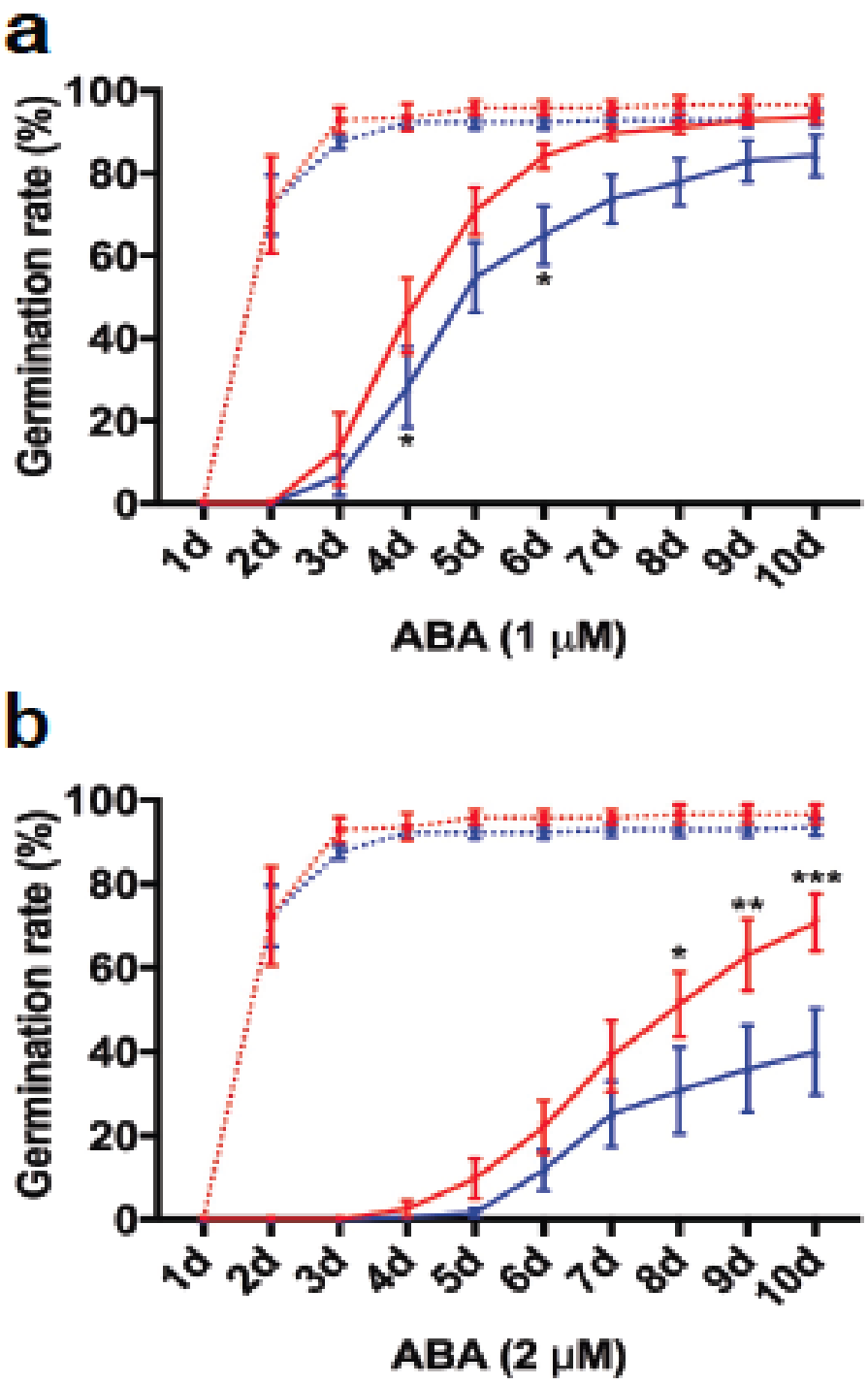

Figure AA 3-4: ABA-induced inhibition of seed germination. Fifty seeds were plated on a single Petri dish, per experiment. Radicle emergence was scored as germination. Varying concentrations of ABA were dissolved in MS media. a, 1 μM and b, 2 μM. Five independent biological replicates were carried out with n = ~250 seeds for each genotype.

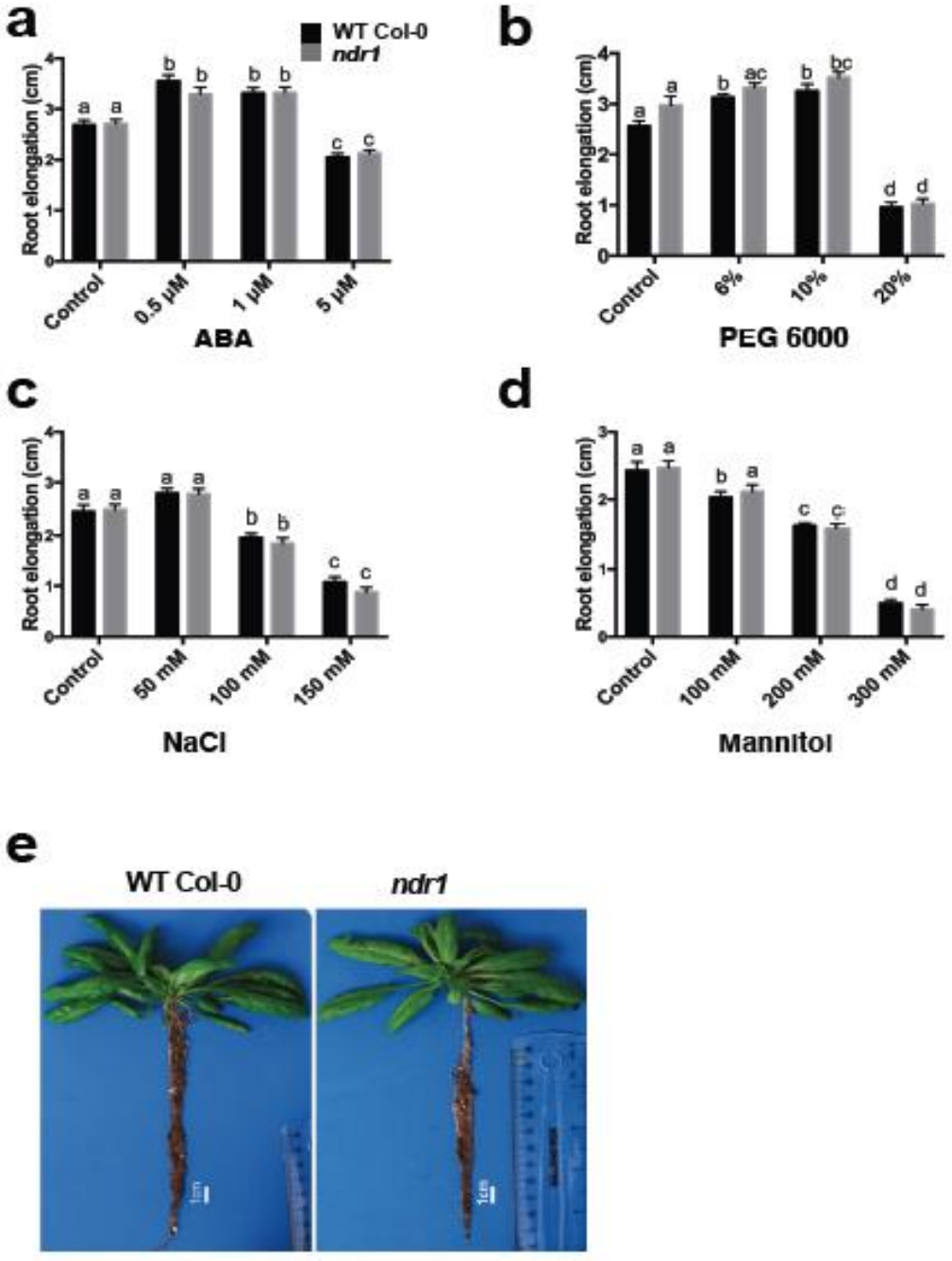

Figure AA 3-5: Root elongation inhibition assays in the *ndr1* mutant in response to abiotic chemical stressors. a, ABA, b, PEG6000, c, NaCl, and d, Mannitol. e, Root length of plants grown under normal conditions for 4 weeks in soil. Root length was measured at 6 days after transferring seedlings onto plates containing salt described above. The average of root elongation was measured from 4 independent biological replicates from 15 seedlings per genotype per experiment.

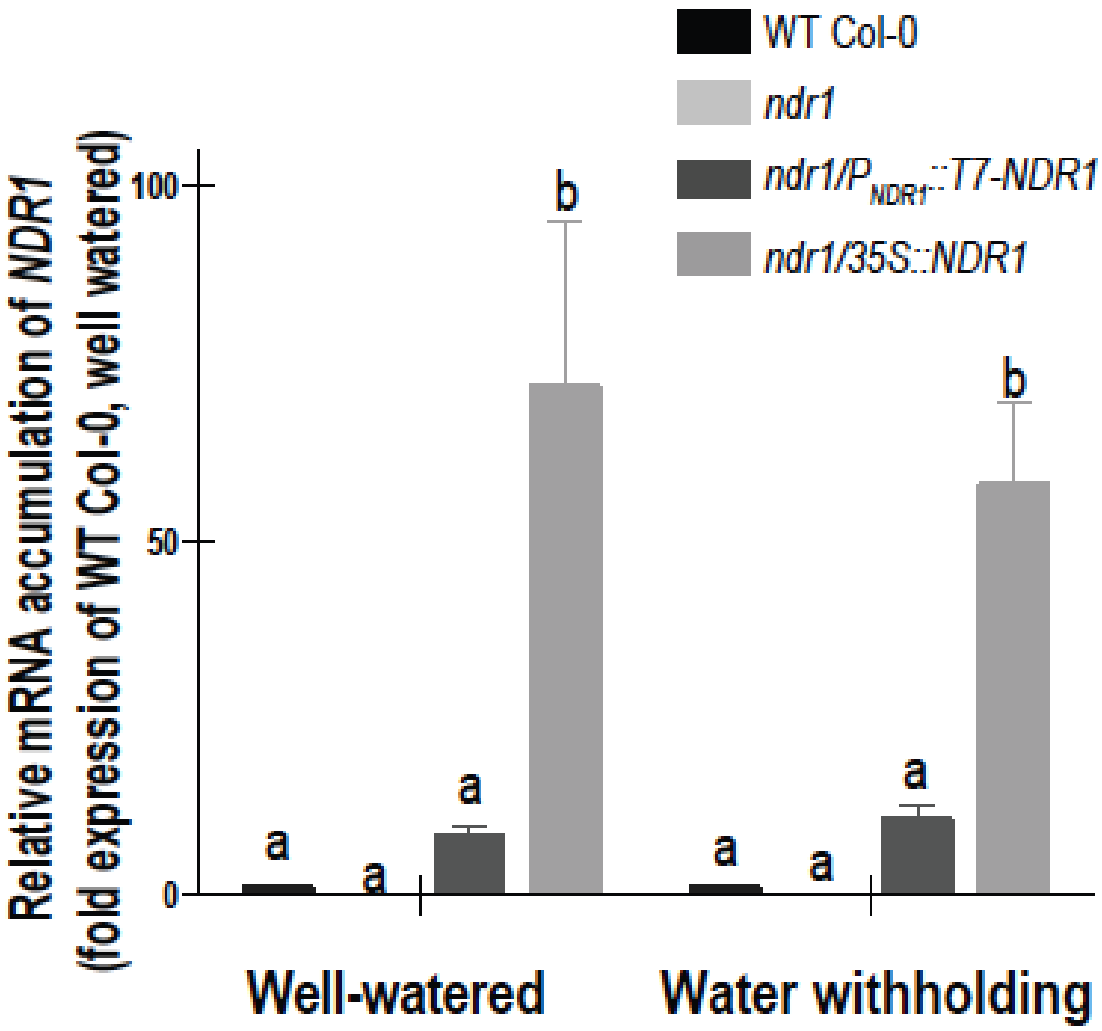

Figure AA 3-6: Relative mRNA level of *NDR1*. *NDR1* mRNA accumulation in WT Col-0, the ndr1 mutant, the *ndr1/P$_{NDR1}$::T7-NDR1* complementation line, and the *ndr1/35S::NDR1* overexpression line under well-watered and drought conditions. Three biological repeats were performed (n=9), $p<0.05$.

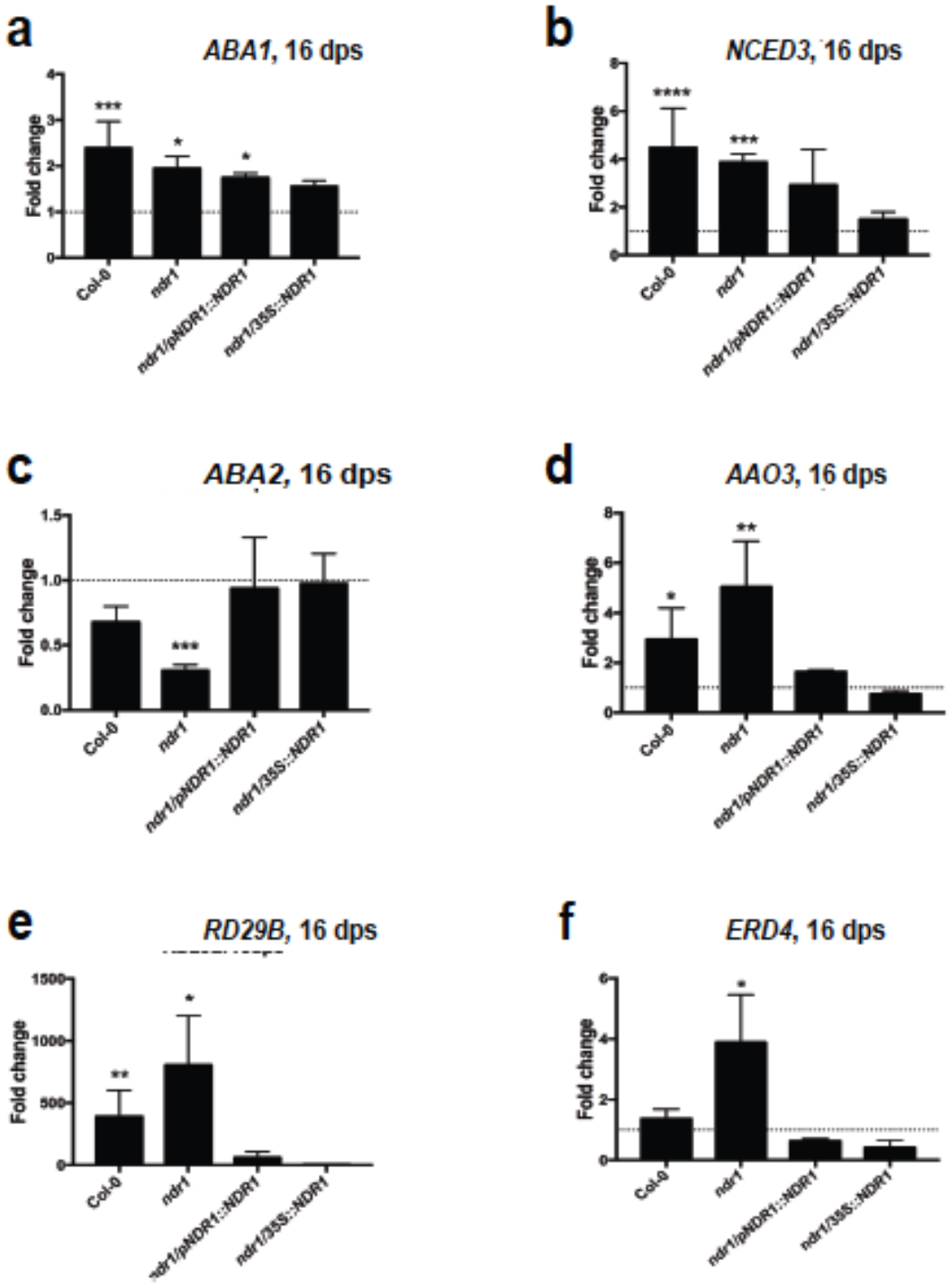

Figure AA 3-7: mRNA expression levels of ABA related genes. mRNA were measured at a, 16 day and b, 21 day post drought stress. Three biological repeats were performed (n=9), $p<0.05$.

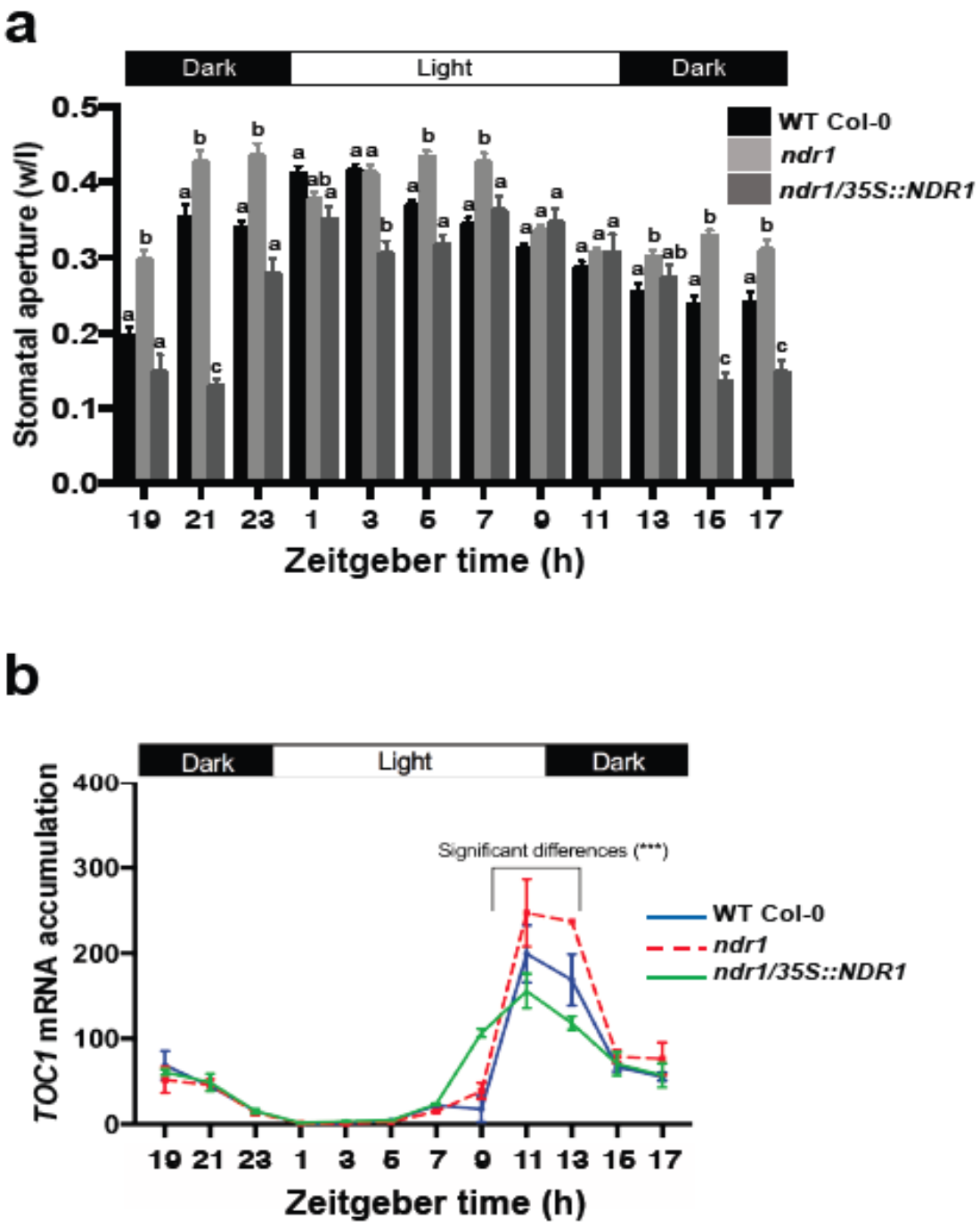

Figure AA 3-8: Stomatal aperture during the diurnal cycle. Plants were assayed every 2 h over a 24 h time period to analyze a, stomatal aperture, and b, TOC1 mRNA accumulation over a 24 h diurnal cycle. Three biological repeats were performed (n=9), $p<0.05$.

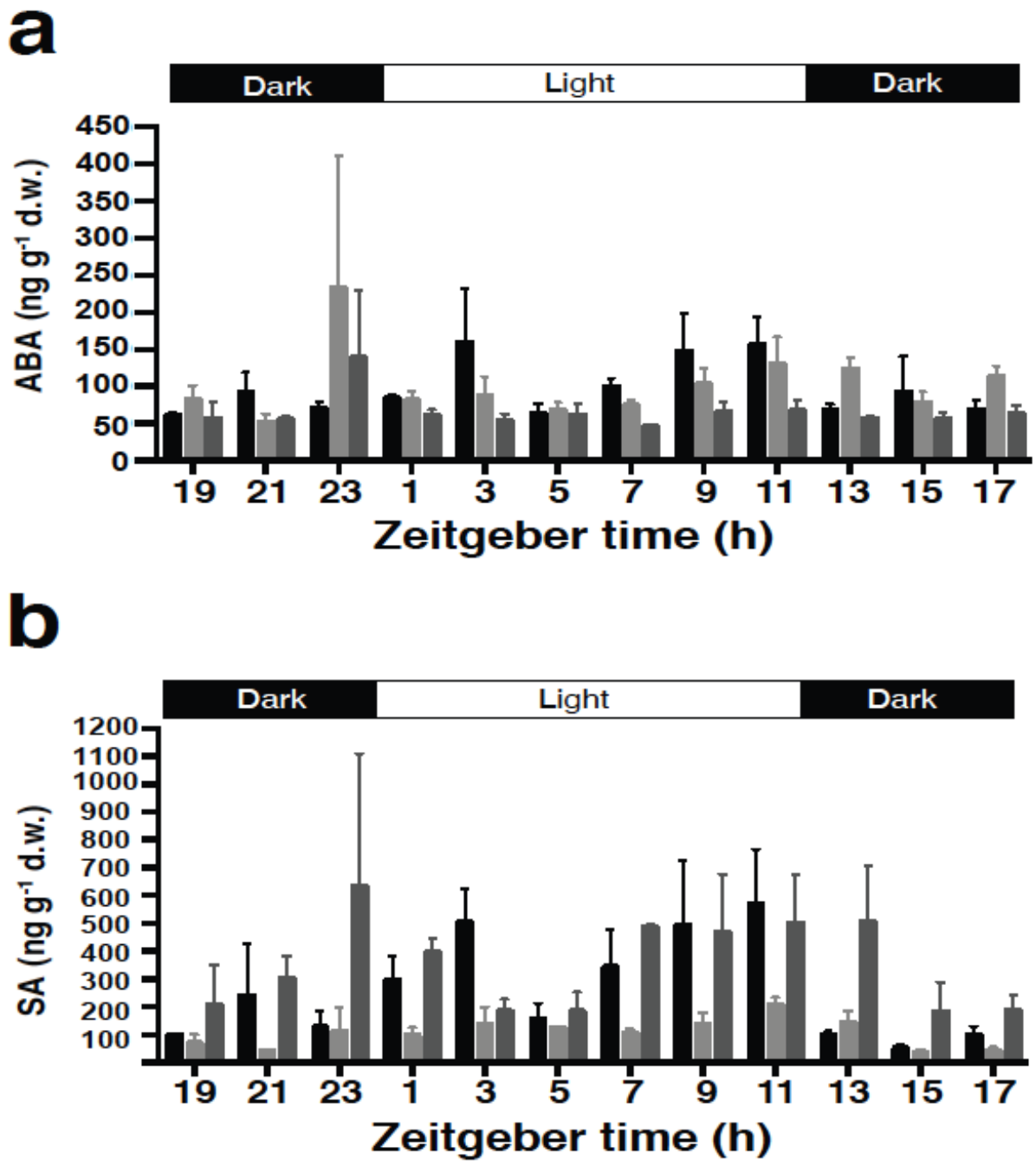

Figure AA 3-9: Hormone accumulation during the 24 hours diurnal cycle. a, ABA accumulation and b, SA accumulation. Three biological repeats were performed (n=9), $p<0.05$.

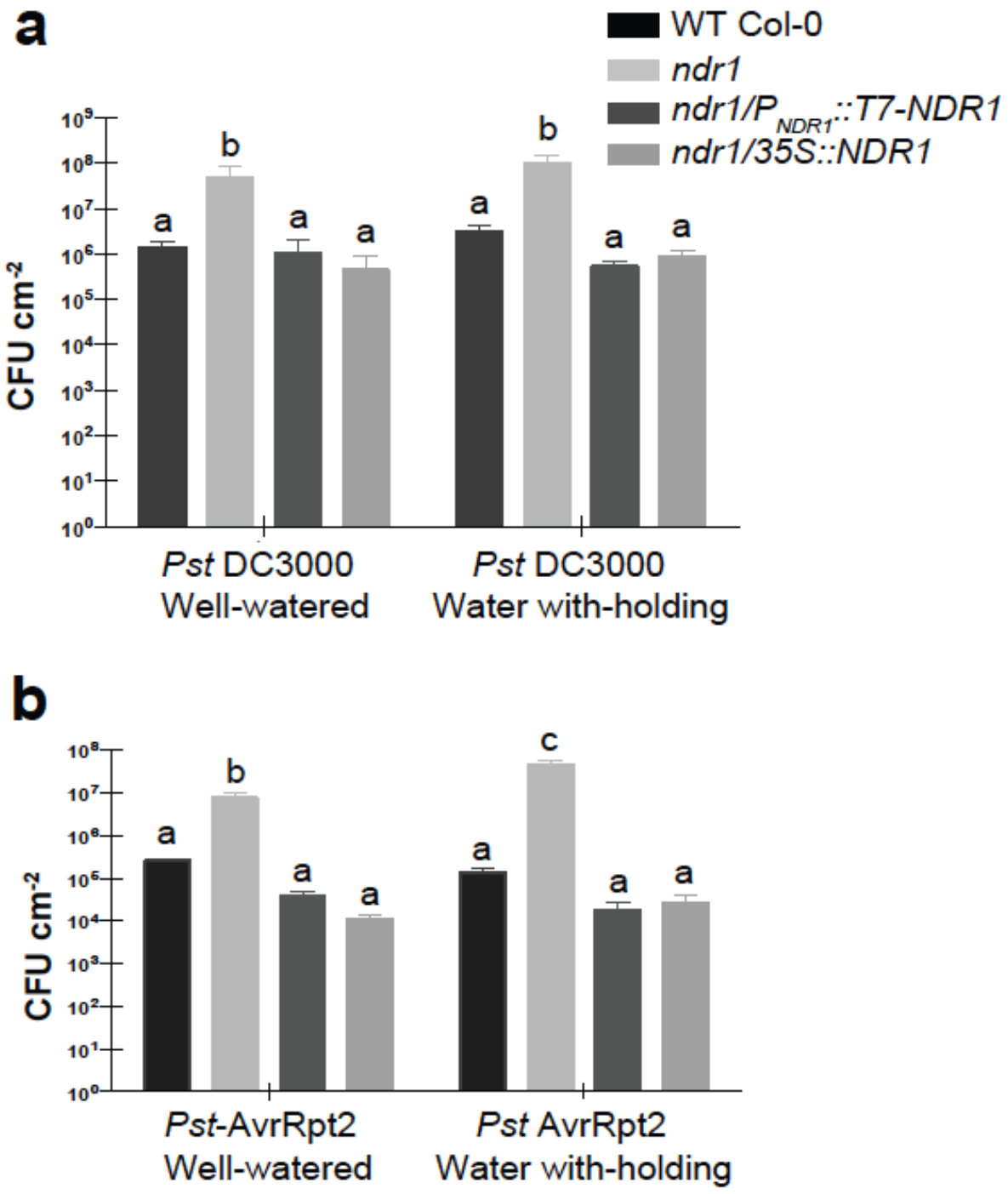

Figure AA 3-10: *Pst* DC3000 and *Pst* DC3000-AvrRpt2 *in planta* growth curves during drought stress. Plants were subjected to drought stress and infected with a, the virulent pathogen Pst. DC3000 as well as the avirulent pathogen b, Pst-AvrRpt2. Plant were infiltrated with *Pst* DC3000 at 12 dpw. Three biological repeats were performed (n=9), $p<0.05$.

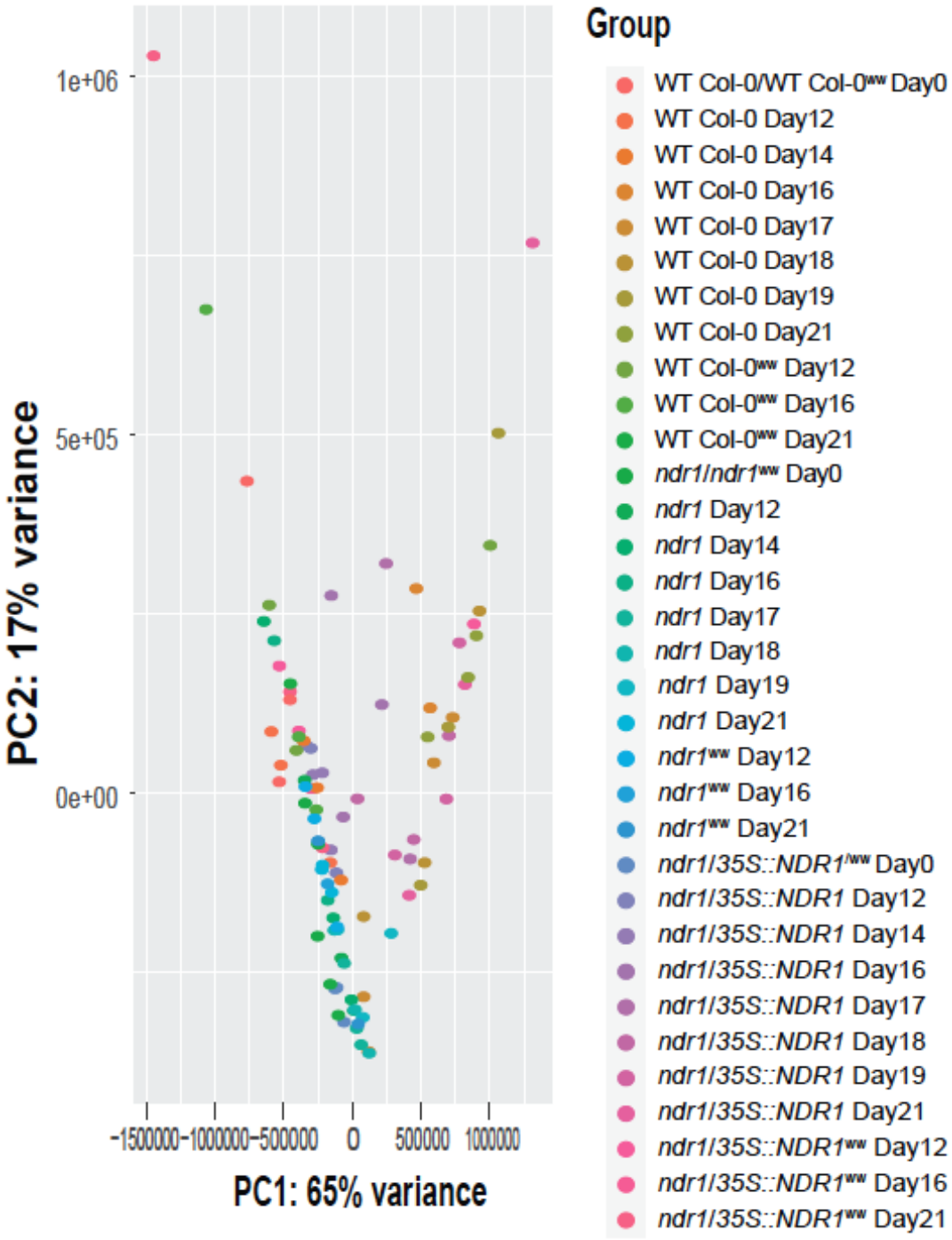

Figure AA 3-11: Principal component analysis (PCA) of gene expression values. Gene expression values of both water withholding condition and well-watered condition during time course.

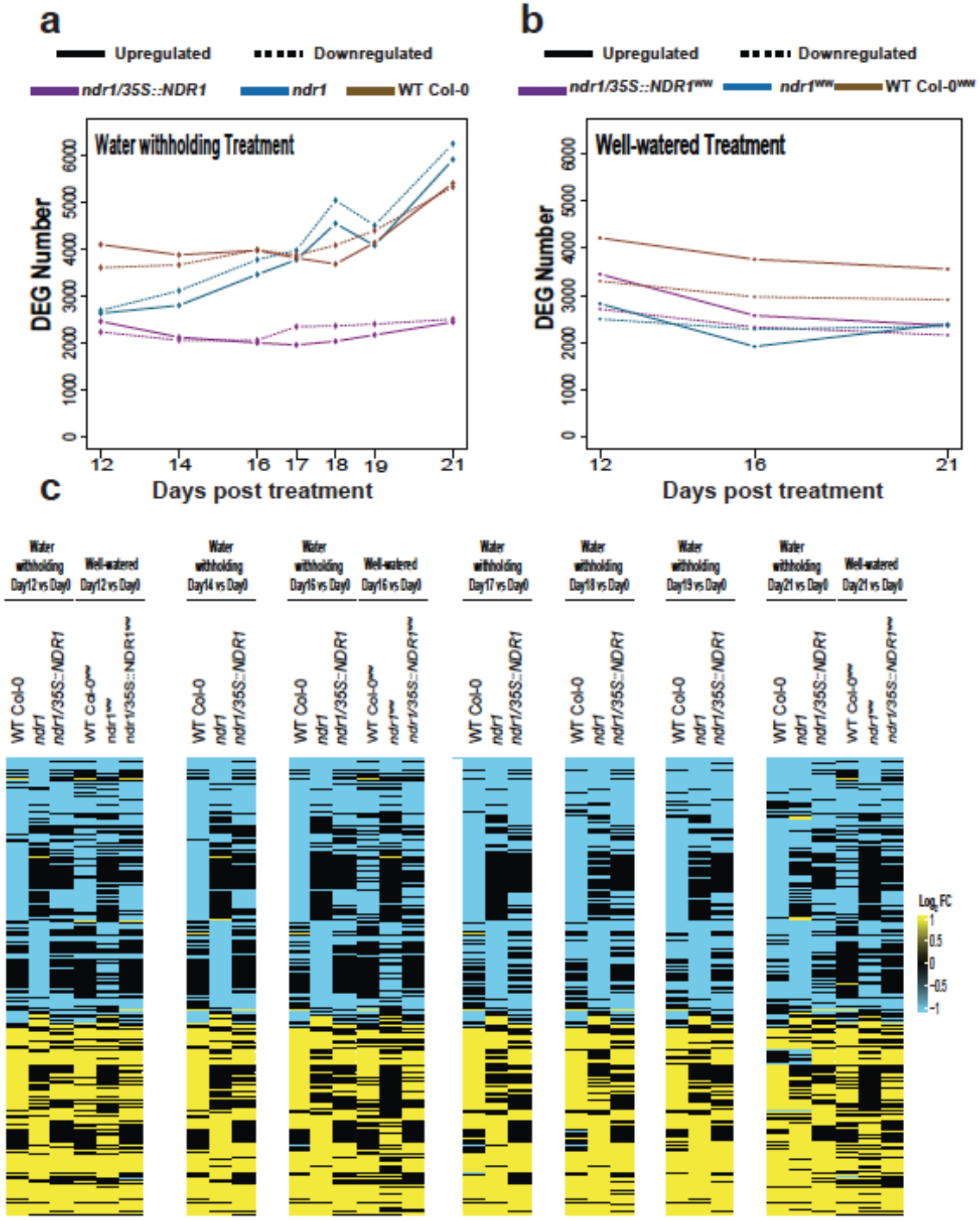

Figure AA 3-12: Total number of DEGs for each treatment compared to Day 0. The total number of DEGs was plotted for each timepoint, for each comparison in both a, water withholding condition and b, well-watered plants. As shown, more DEGs were identified during the timecourse of water withholding as compared to the well-watered plants. c, Heat map of DEGs in WT Col-0, the *ndr1*, mutant and *ndr1/35S::NDR1* lines.

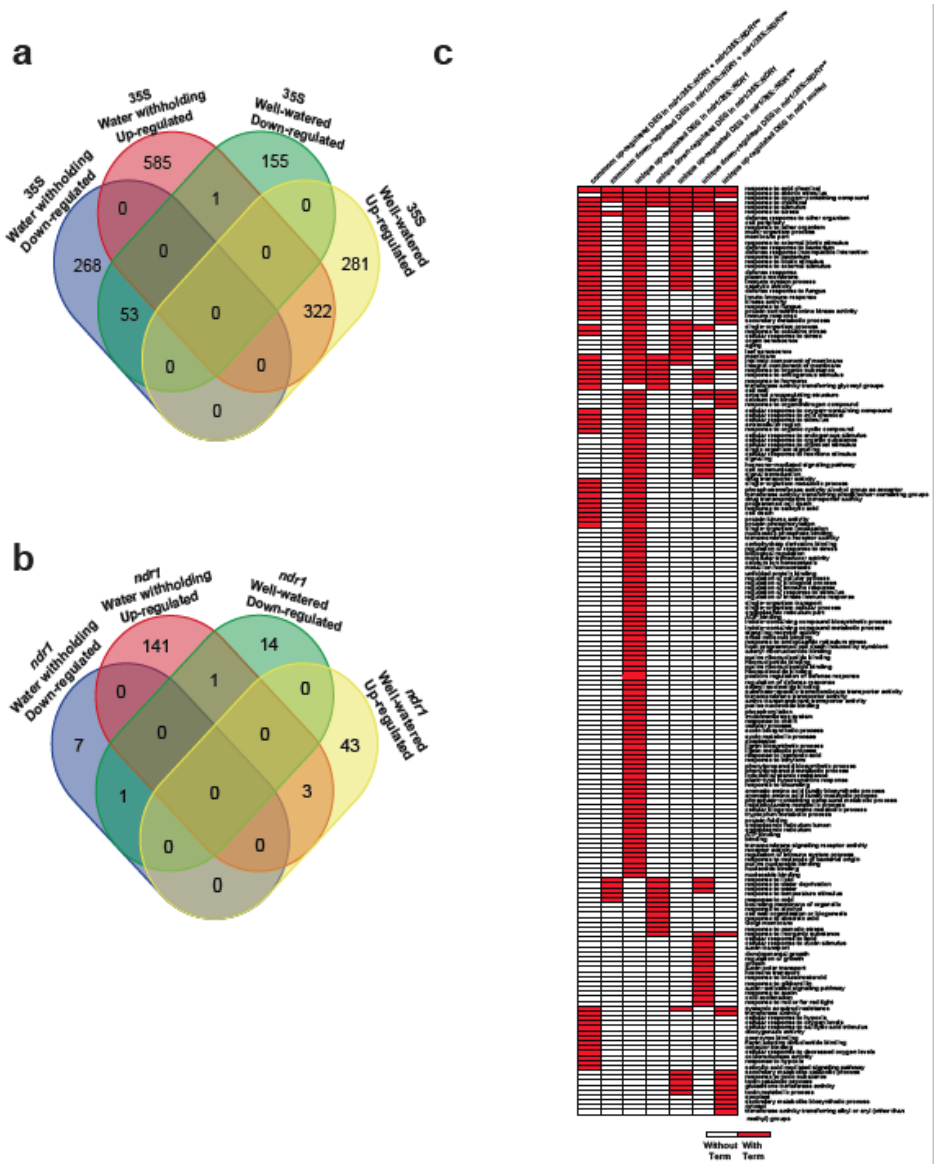

Figure AA 3-13: Expression patterns of transcription factors that response to drought stress. Transcription factor mRNAs identified as up- or downregulated across all timepoints of the treatment.

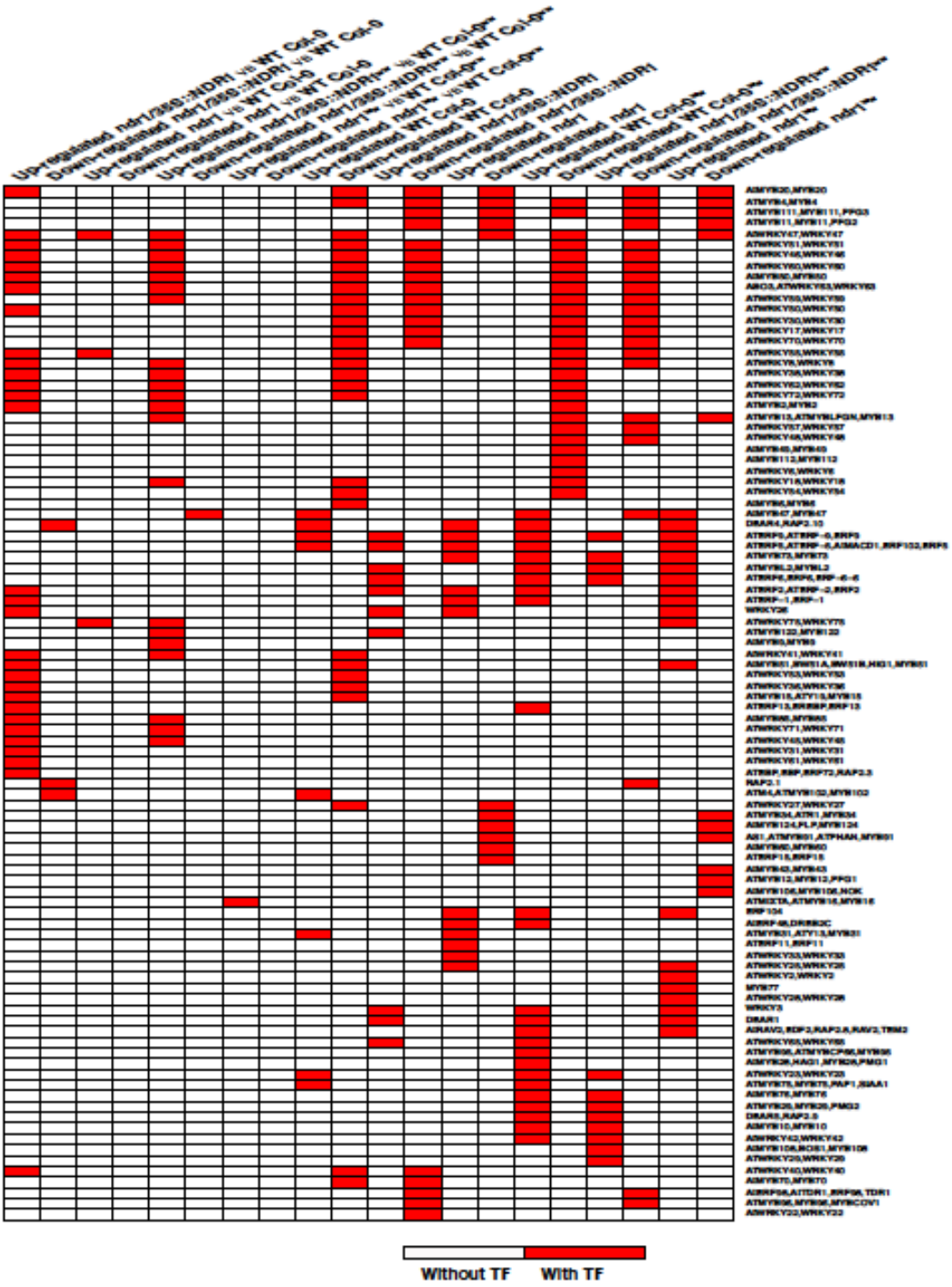

Figure AA 3-14: Common up- or downregulated genes under well water and water withholding conditions. a, The number of genes in *ndr1/35s::ndr1* lines. b, the number of genes in ndr1 mutant. c, GO terms that are associated with genes identified from a and b.

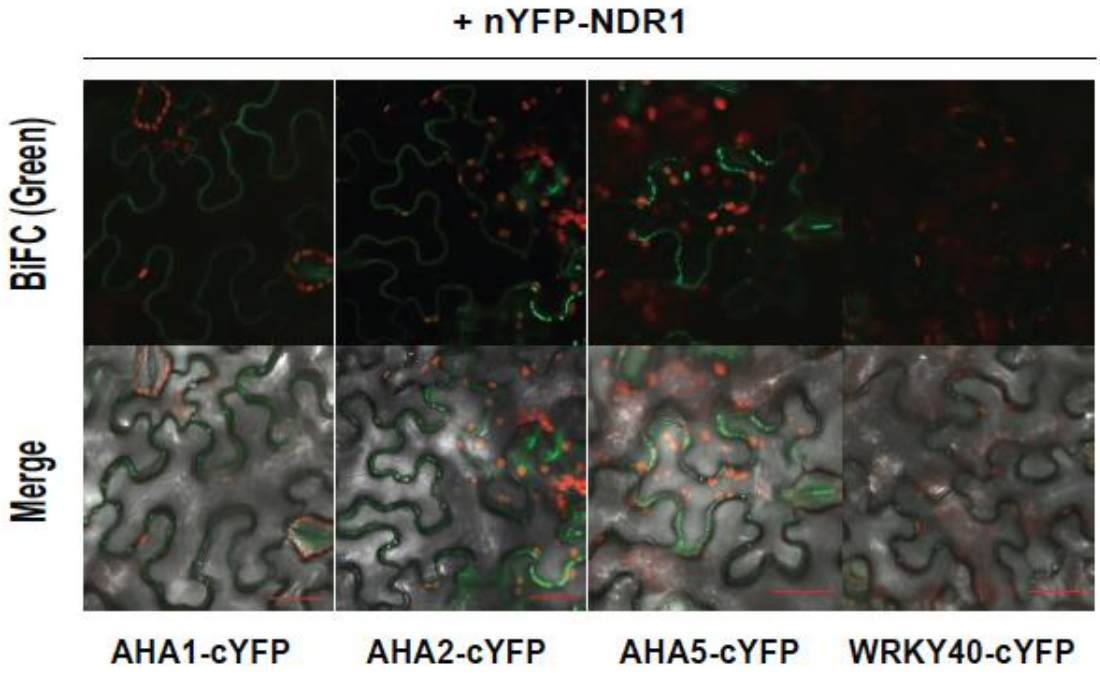

Figure AA 3-15: BiFC confirmed NDR1 associates with AHA1, AHA2 and AHA5. Co-expression of n-YFP-NDR1 with H^+-ATPase-cYFP proteins in N. benthamiana leaves. WRKY40-c-YFP was employed as a negative control.

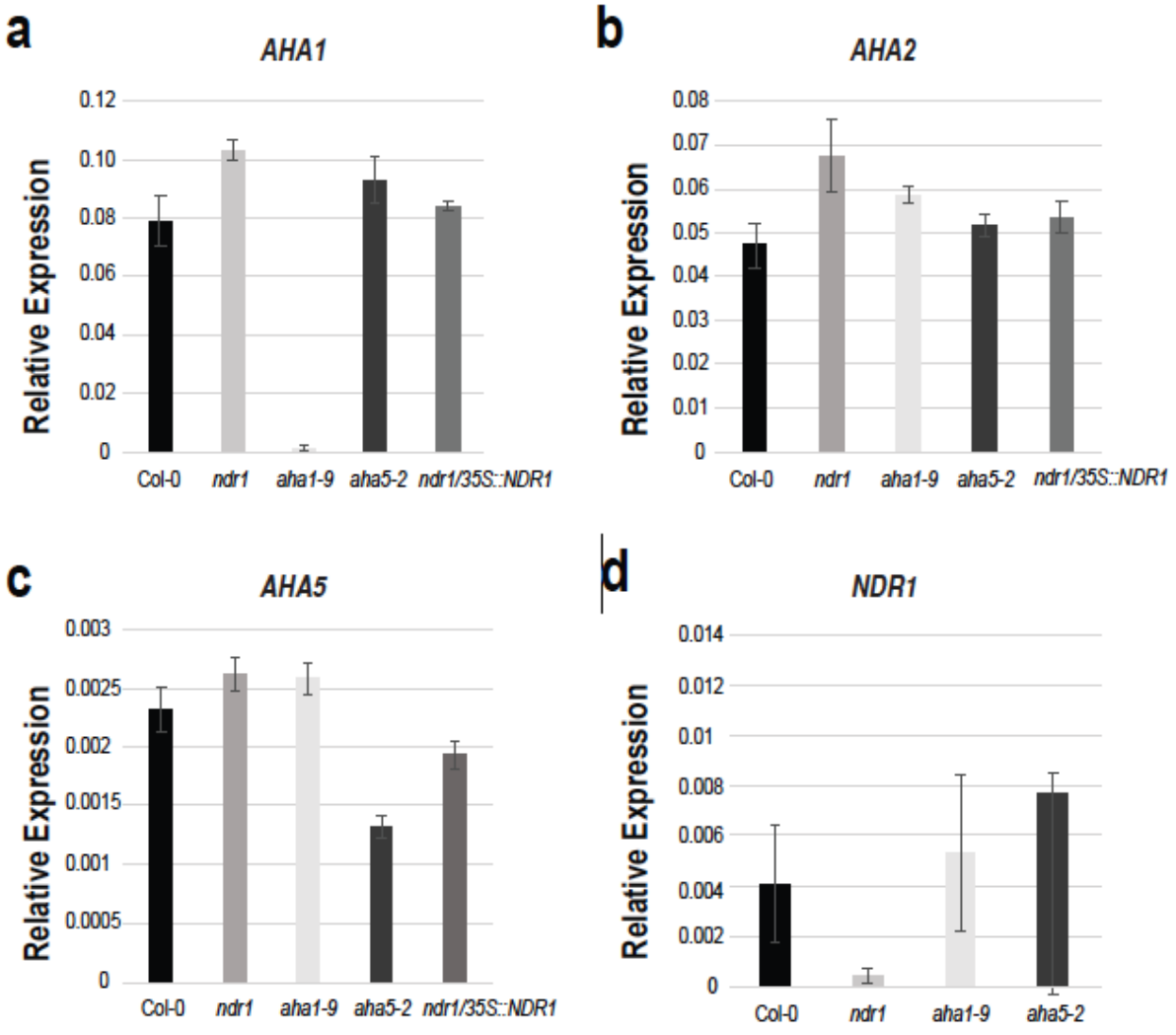

Figure AA 3-16: mRNA expression levels of *AHA* in 4 *Arabidopsis* genotypes. a, *AHA1*, b, *AHA2*, c, AHA5, and d, *NDR1*.

Data Set AAs

All the Data Set AA are deposited at

https://www.biorxiv.org/content/10.1101/2021.06.10.445978v1.supplementary-material

Data Set AA 3-1: Locus list of DEG genes among *ndr1* mutant vs WT Col-0 lines at Day 0, Day 12, Day 14, Day 16, Day 17, Day 18, Day 19 and Day 21 under water withholding treatment.

Data Set AA 3-2: Locus list of DEG genes among *ndr1/35S::NDR1* Overexpression vs WT Col-0 lines at Day 0, Day 12, Day 14, Day 16, Day 17, Day 18, Day 19 and Day 21 under water withholding treatment.

Data Set AA 3-3: Locus list of DEG genes among *ndr1* mutant vs WT Col-0 lines at Day 0, Day 12, Day 16 and Day 21 under well-watered treatment.

Data Set AA 3-4: Locus list of DEG genes among *ndr1* mutant vs WT Col-0 at Day 0, Day 12, Day 16 and Day 21 under well-watered treatment.

Data Set AA 3-5: Locus list of DEG genes within WT Col-0 group at Day 12, Day 14, Day 16, Day 17, Day 18, Day19 and Day 21 vs Day 0 under water withholding treatment.

Data Set AA 3-6: Locus list of DEG genes within *ndr1* mutant group at Day 12, Day 14, Day 16, Day 17, Day 18, Day 19 and Day 21 vs Day 0 under water withholding treatment.

Data Set AA 3-7: Locus list of DEG genes within ndr1/35S::NDR1 overexpression group at Day 12, Day 14, Day 16, Day 17, Day 18, Day 19 and Day 21 vs Day 0 under water withholding treatment.

Data Set AA 3-8: Locus list of DEG genes within WT Col-0 group at Day 12, Day 16, and Day 21 vs Day 0 under well-watered treatment.

Data Set AA 3-9: Locus list of DEG genes within the *ndr1* mutant group at Day 12, Day 16, and Day 21 vs Day 0 under well-watered treatment.

Data Set AA 3-10: Locus list of DEG genes within *ndr1/35S::NDR1* overexpression group at Day 12, Day 14, Day 16, Day 17, Day 18, Day 19, and Day 21 vs Day 0 under well-watered treatment.

Data Set AA 3-11: Locus list of consistent upregulated and downregulated DEG genes among *ndr1* mutant vs WT Col-0, and *ndr1/35S::NDR1* overexpression vs WT Col-0 lines at Day 12 to Day 21 under water withholding treatment, respectively.

Data Set AA 3-12: Locus list of consistently upregulated and downregulated DEG genes among *ndr1* mutant vs WT Col-0, and *ndr1/35S::NDR1* overexpression vs WT Col-0 lines

at Day 12 to Day 21 under well-watered treatment, respectively.

Data Set AA 3-13: Locus list of consistent upregulated and downregulated DEG genes among WT Col-0, *ndr1* mutant, and *ndr1/35S::NDR1* overexpression lines at Day 12 to Day 21 vs Day 0 under water withholding treatment, respectively.

Data Set AA 3-14: Locus list of consistent upregulated and downregulated DEG genes among WT Col-0, *ndr1* mutant, and *ndr1/35S::NDR1* overexpression at Day 12 to Day 21 vs Day 0 under well-watered treatment, respectively.

Data Set AA 3-15: 96 transcript factors among the constantly differentially expressed genes.

Data Set AA 3-16: Gene ontology enrichment of consistently upregulated and downregulated DEG genes among *ndr1* mutant vs WT Col-0, and *ndr1/35S::NDR1* overexpression vs WT Col-0 lines at Day 12 to Day 21 under water withholding and well-watered treatment, respectively.

Data Set AA 3-17: DEGs in selected gene ontology enrichment among *ndr1* mutant vs WT Col-0, and *ndr1/35S::NDR1* overexpression vs WT Col-0 lines at Day 12 to Day 21 under water withholding and well-watered treatment, respectively.

Data Set AA 3-18: Gene ontology enrichment of consistently upregulated and downregulated DEG genes among WT Col-0, *ndr1* mutant and *ndr1/35S::NDR1* overexpression lines at Day 12 to Day 21 vs Day 0 under water withholding and well-watered treatment, respectively.

Data Set AA 3-19: DEGs in selected gene ontology enrichment among WT Col-0, *ndr1* mutant, and *ndr1/35S::NDR1* overexpression lines at Day 12 to Day 21 vs Day 0 under water withholding and well-watered treatment, respectively.

Data Set AA 3-20: The common and unique upregulated and downregulated DEGs among *ndr1* mutant vs WT Col-0, and *ndr1/35S::NDR1* overexpression vs WT Col-0 lines at Day 12 to Day 21 under water withholding and well-watered treatment, respectively.

Data Set AA 3-21: Gene modules list among WT Col-0, *ndr1* mutant and *ndr1/35S::NDR1* overexpression lines under water withholding treatment.

Data Set AA 3-22: Eigengene expression value of each module among WT Col-0, *ndr1* mutant and *ndr1/35S::NDR1* overexpression lines under water withholding treatment.

Data Set AA 3-23: Gene Ontology Enrichment of each gene module among WT Col-0, *ndr1* and *ndr1/35S::NDR1* overexpression lines under water withholding treatment.

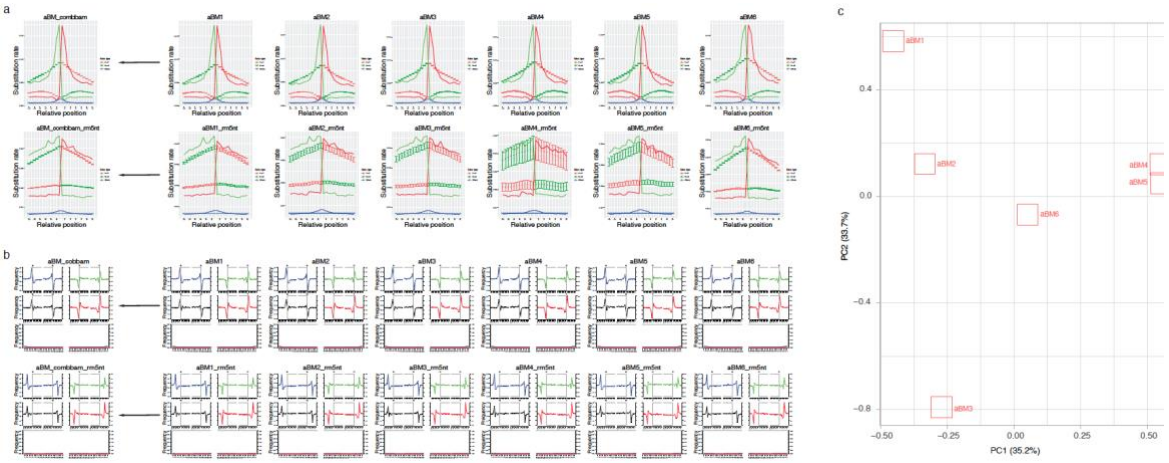

Figure AA 4-1: Ancient DNA (aDNA) damage pattern of archaeological Bolivian maize (aBM). (a) Cytosine deamination damage patterns for the combined bam file of six aBM sequence samples and individual six aBM sequence samples. The position specific substitutions from the 5'end(red) and the 3'end(green) of a read. The red line corresponds to C to T substitutions, the green line corresponds to G to A substitutions, and the blue line represents other types of substitutions. (b) Ancient DNA damage profile. The four upper plots show the base frequencies inside and outside of a read, where the open grey box corresponds to a read. The two lower plots show the position specific substitutions from the 5'end(left) and the 3'end(right) of a read. The red line corresponds to C to T substitutions, the blue line corresponds to G to A substitutions, and the fade line represent other types of substitutions. (c) Principal component analysis of six aBM sequence samples.

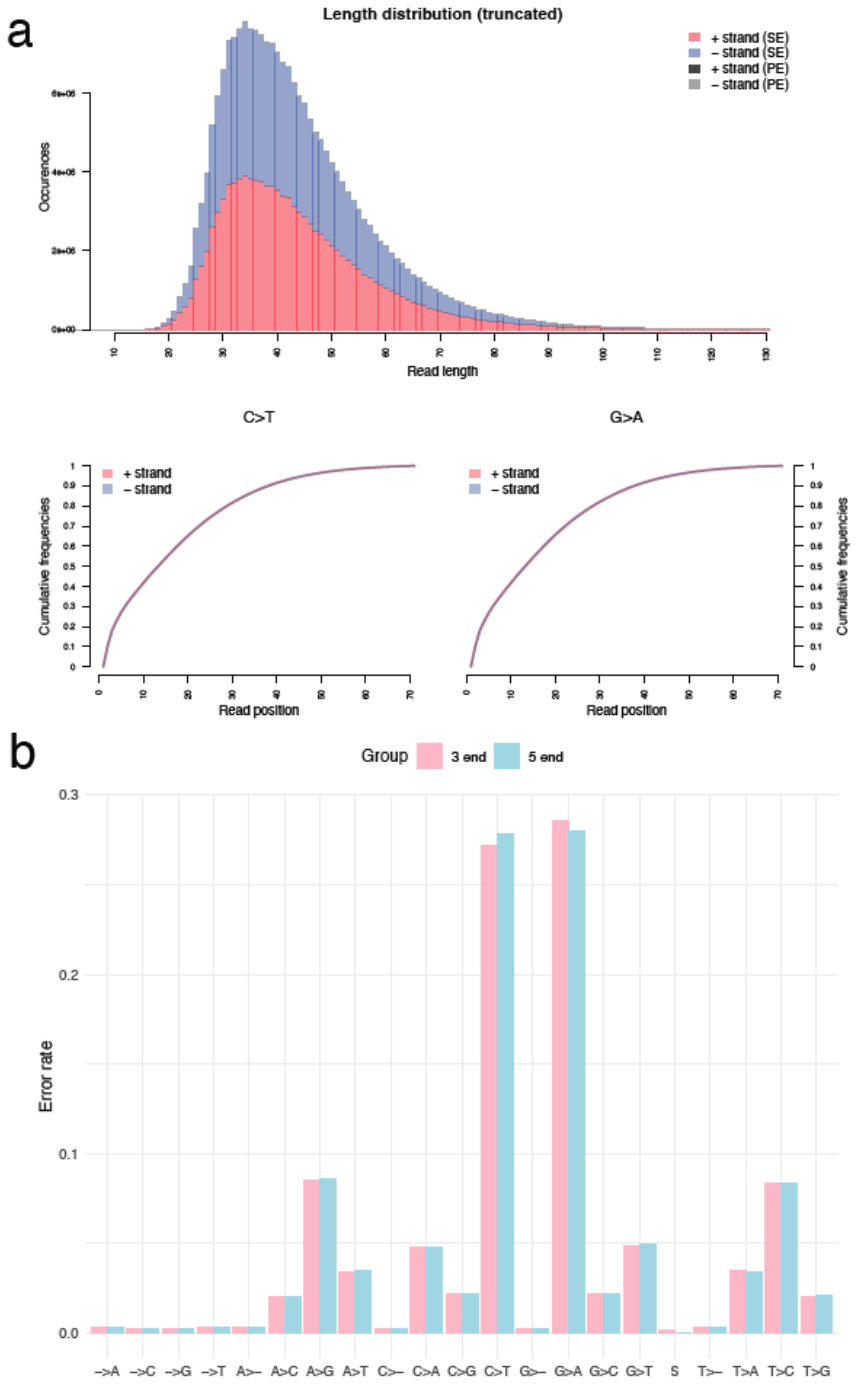

Figure AA 4-2: Archaeological Bolivian maize (aBM) DNA read length and error rate. (a) Read length distribution of combined bam files of all six aBM sequence samples prior 5 nt trimmed. (b) Error rate (%) per substitution type among all misincorporation in the combined bam aBM sample. Pink indicates the error rate in 3 end and blue indicate the error rate in 5 end.

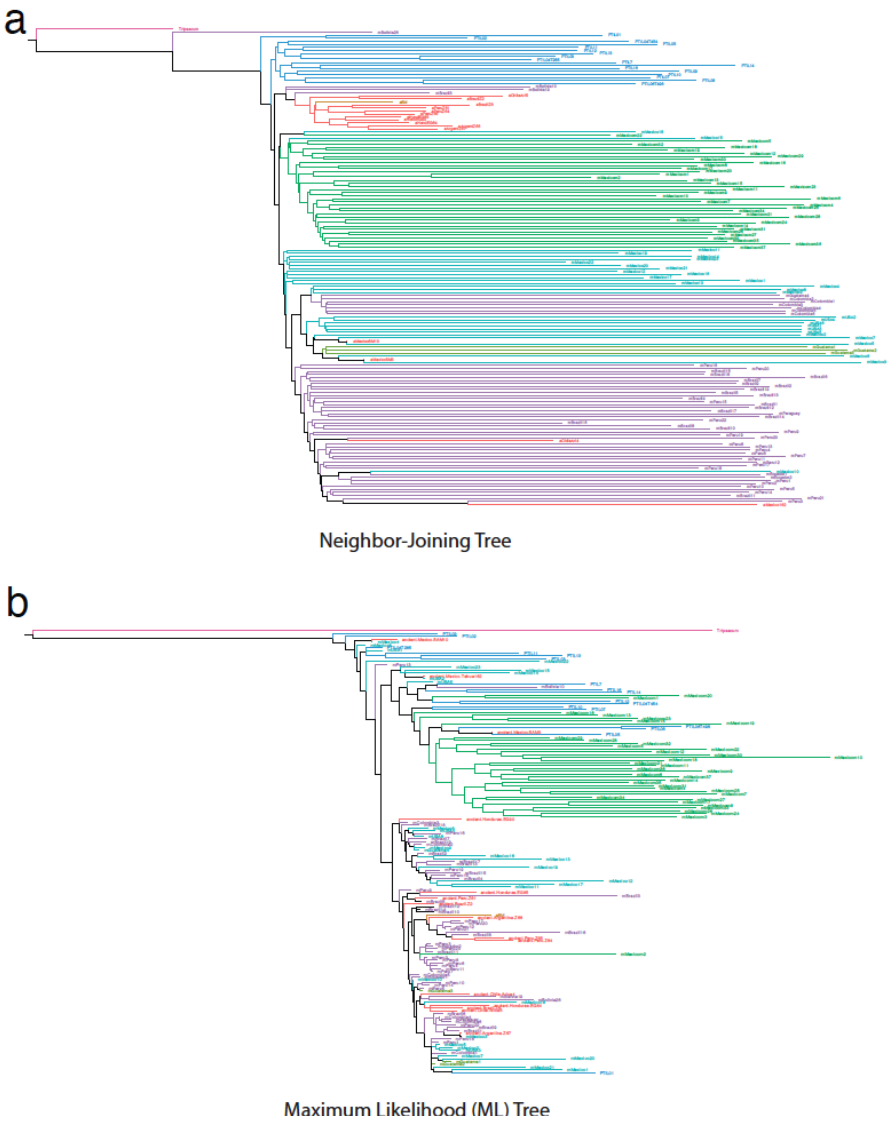

Figure AA 4-3: Phylogenetic tree of archaeological Bolivian maize (aBM). (a) Neighbor-Joining tree with VCF2Pop. (b) Maximum Likelihood (ML) tree with SNPhylo. Colors match the color in Figure 4- 2a and b. PTIL (*Zea mays* ssp. *parviglumis*), mMexicom (modern Mexico; *Zea mays* ssp. *mexicana*), mMexicom (modern Mexico), mParaguay (modern Paraguay), mBrazil (modern Brazil), mPeru (modern Peru), mEcuador (modern Ecuador), mGuatemala (modern Guatemala), mColombia (modern Colombia), and mBolivia (modern Bolivia).

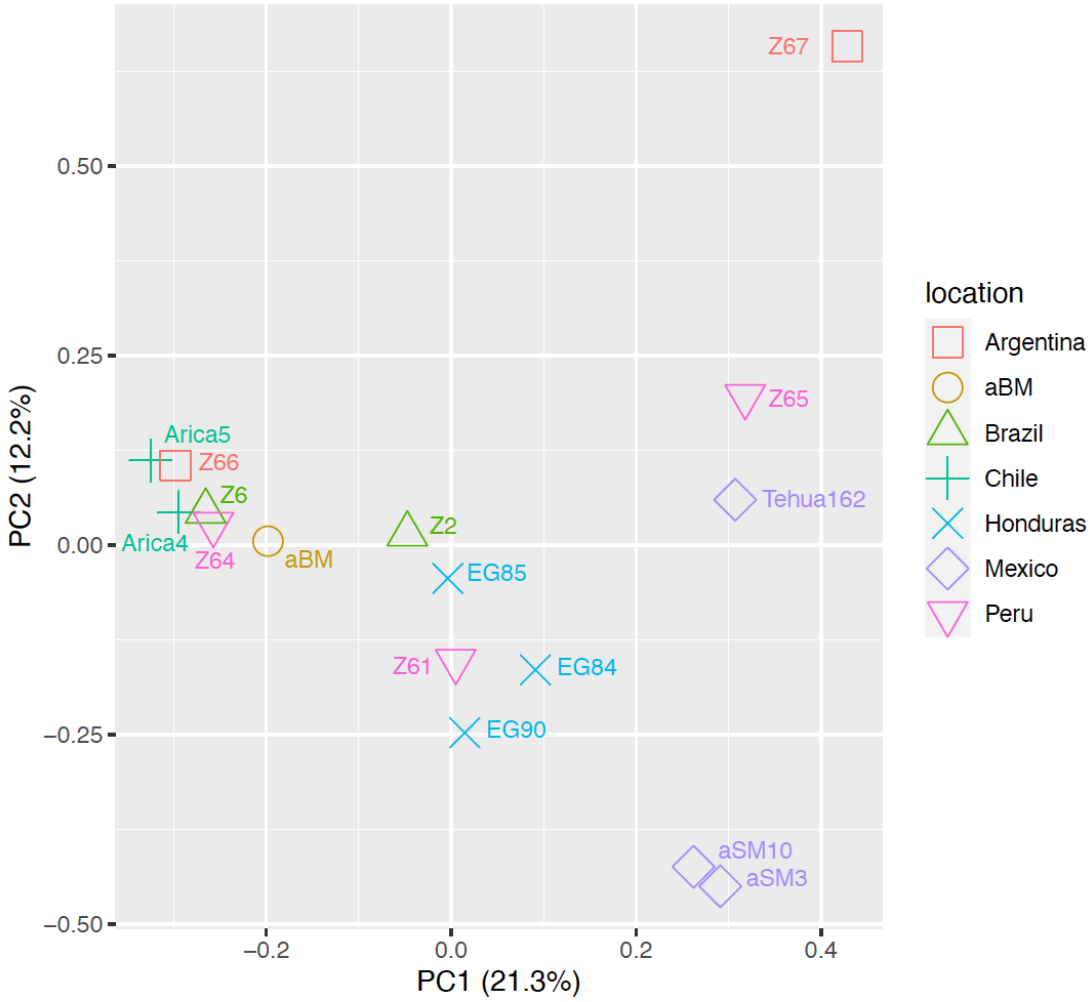

Figure AA 4-4: Principal component analysis of ancient maize samples. Samples with the same color and shape indicate they came from the same country.

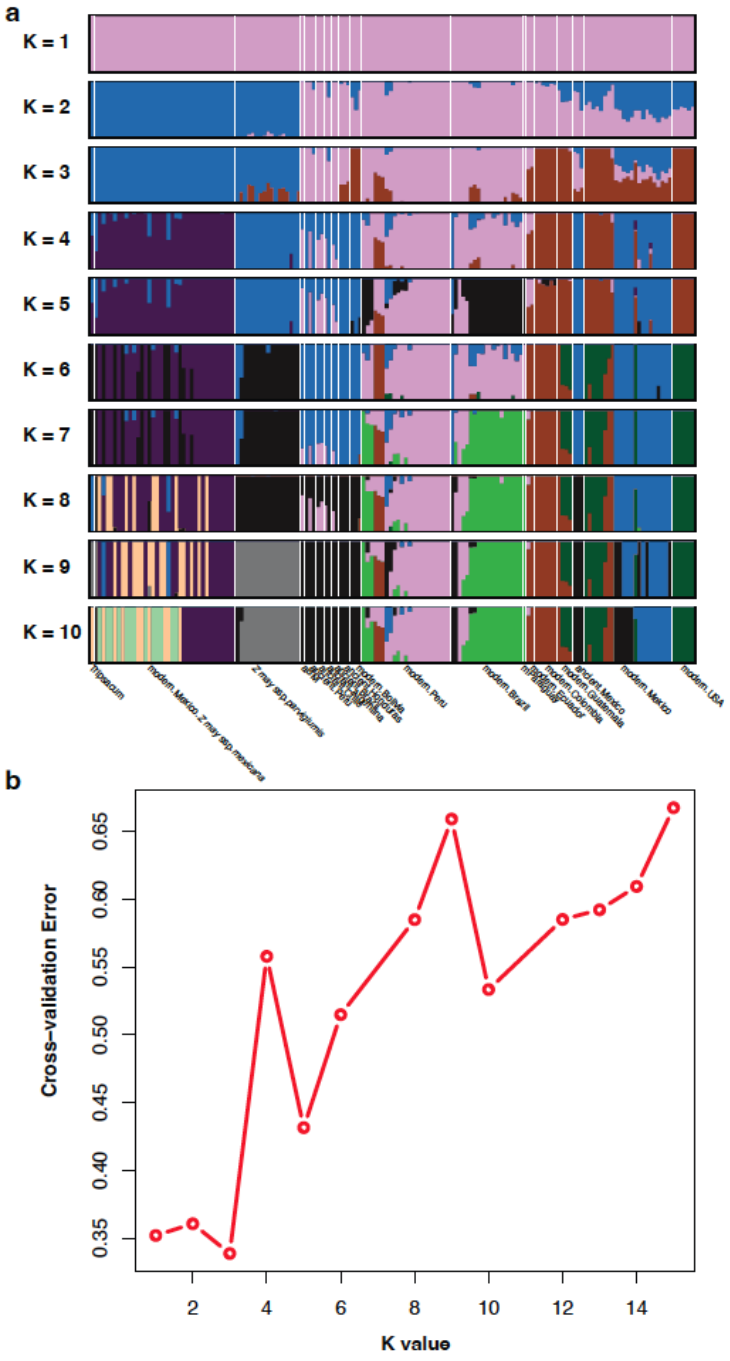

Figure AA 4-5: Population structures analysis. (a) Clustering of the population using ADMIXTURE with K = 1 to 10 with filtered 16,081,009 SNPs. The colors used do not represent the same as Figure 2a and b. (b) Cross validation error at different K using ADMIXTURE. The optimal K was determined by the minimum CV error (K=3).

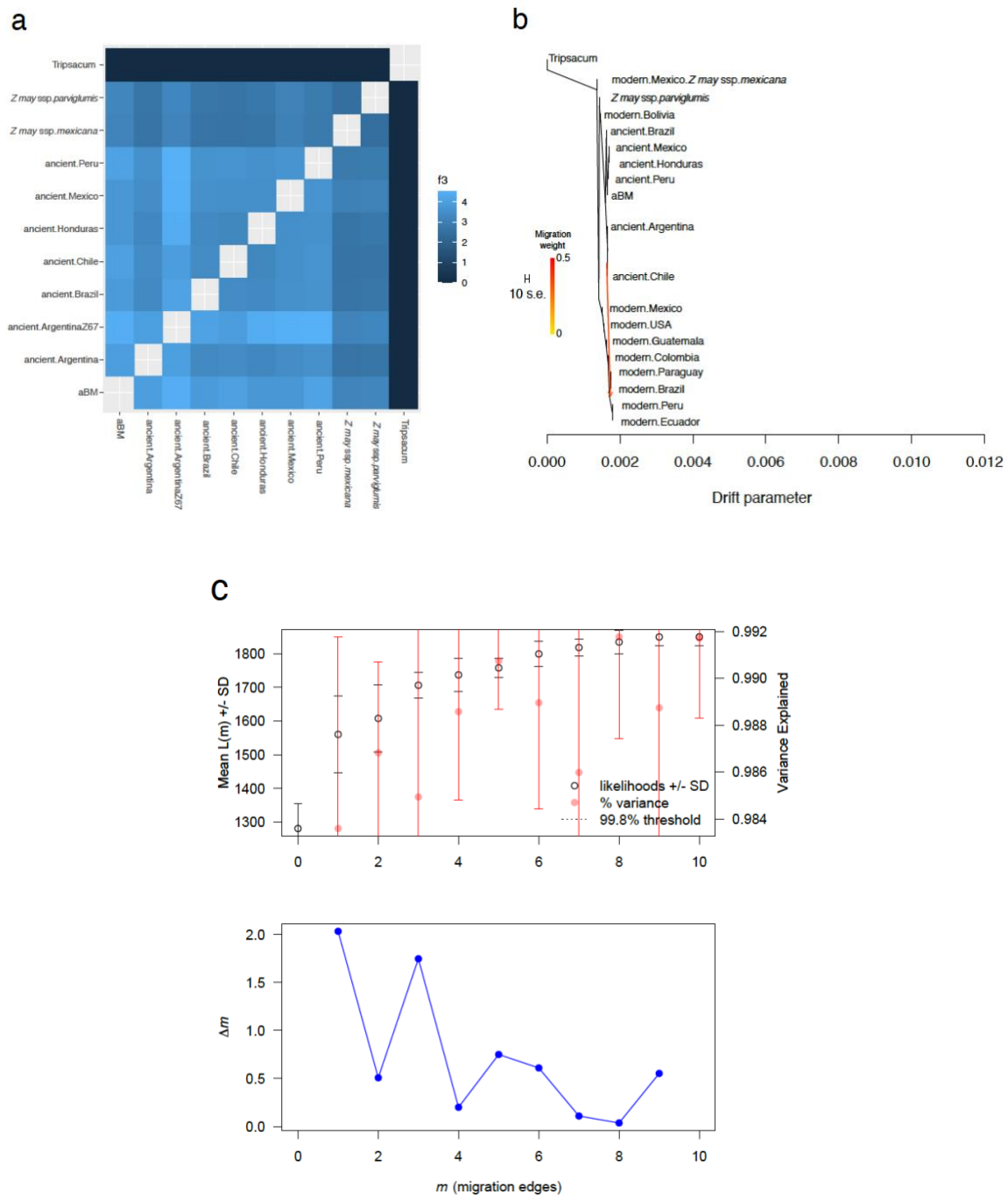

Figure AA 4-6: Estimated branch length between a pair of populations with respect to a common outgroup *Tripsacum* and gene flow. (a) Heatmap of the values of pairwise $f3$ statistics. (b) The tree structure of the graph inferred by Treemix for the maize. (c) The output produced by OptM for maize data.

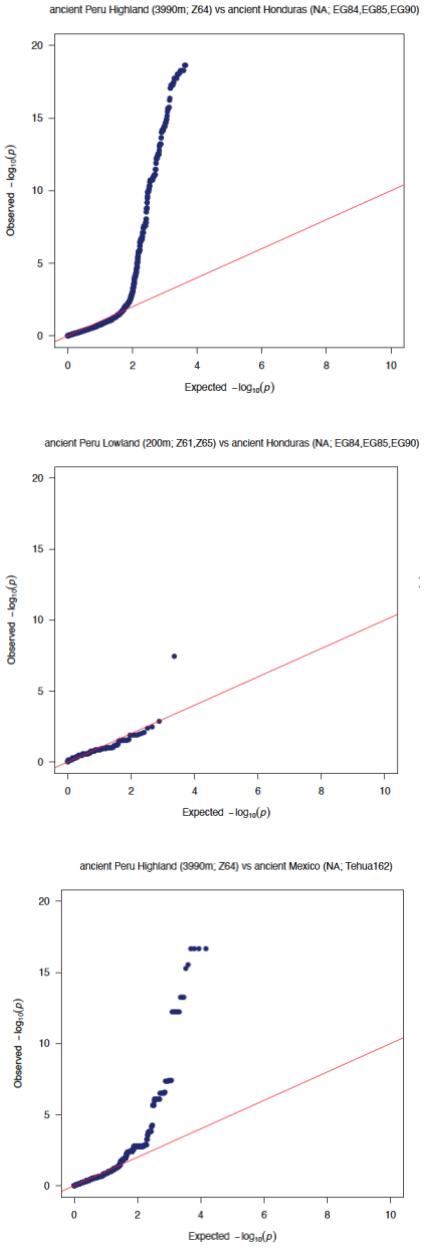

Figure AA 4-7: Q-Q plot for genome-wide selection in ancient Peruvian maize compared to ancient Honduras and ancient Mexico maize. The Q-Q plot between maize from ancient Peru highland vs ancient Honduras, ancient Peru lowland vs ancient Honduras, ancient Peru highland vs ancient Mexico Tehua162, respectively. Each point indicates a value from expected and observed. Red line represents the expected distribution of the p-value, while the blue trend represents the observed distribution. No Q-Q plot of ancient Peru highland and lowland vs ancient Mexico SM3,10 is shown because no common data exist among them.

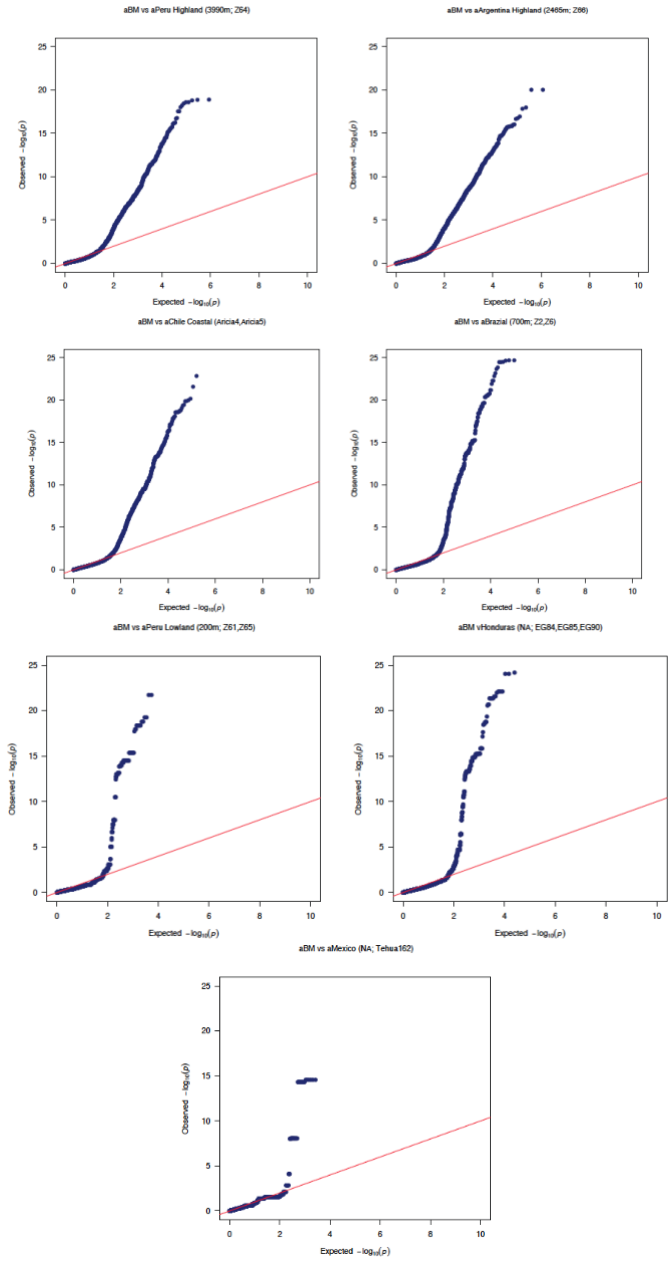

Figure AA 4-8: Q-Q plot for genome-wide selection in archaeological Bolivian maize (aBM) genome-wide selection compared to other ancient maize groups. The Q-Q plot between aBM vs ancient Peru Highland, ancient Argentina Highland, ancient Chile Coastal, ancient Brazil, ancient Peru Lowland, ancient Honduras, ancient Peru Highland, and ancient Mexico Tehua162, respectively. Each point indicates a value from expected and observed. Red line represents the expected distribution of the p-value, while blue trend represents the observed distribution. No Q-Q plot of aBM vs ancient MexicoSM3,10 is shown because no common data exist between them.

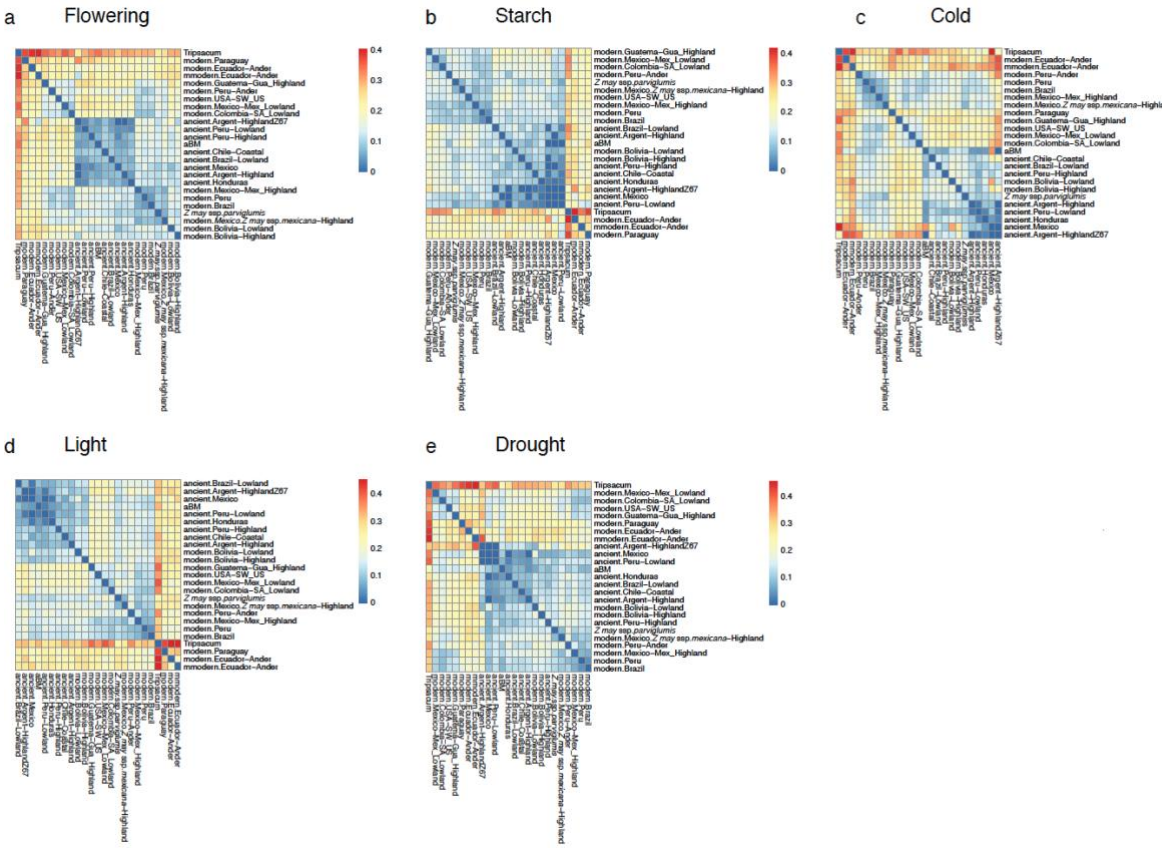

Figure AA 4-9: Fixation index (Fst) of selection on aBM. (a-e) Heatmap of Fst of flowering, starch, cold, light, drought and heat relative biological process.

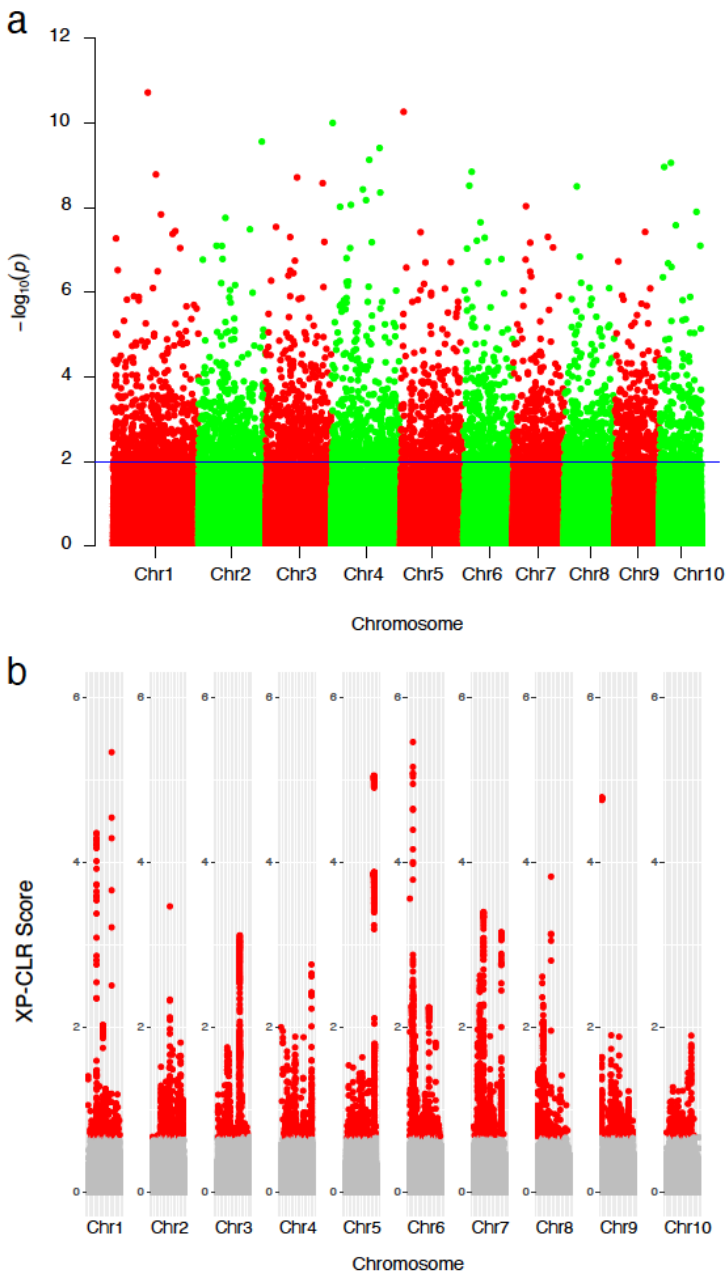

Figure AA 4-10: XP-CLR score for genome-wide selection and selective sweeps distribution between modern maize and ancient maize in South America. (a) XP-CLR estimated by comparing the modern maize against the ancient maize population in South America. Red points indicate the top 1% XP-CLR score. (b) Signal of selection of modern maize compared to ancient maize population in South America. Each point indicates a SNP. The blue horizontal line shows the genome-wide significance level ($p=1 \times 10^{-2}$).

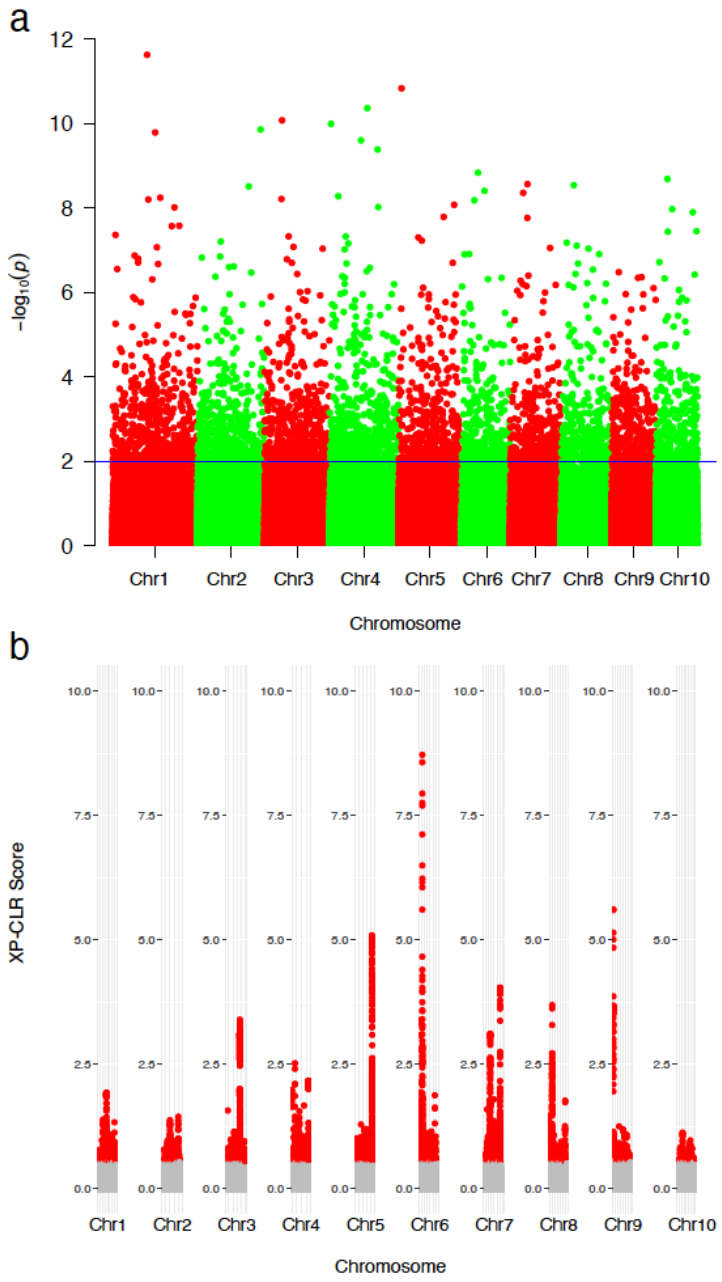

Figure AA 4-11: XP-CLR score for genome-wide selection and selective sweeps distribution between modern maize and ancient maize in the Western Hemisphere. (a) XP-CLR estimated by comparing the modern maize against the ancient maize population in the western hemisphere. Red points indicate the top 1% XP-CLR score. (b) Signal of selection of modern maize compared to ancient maize population in the western hemisphere. Each point indicates a SNP. The blue horizontal line shows the genome-wide significance level (p=1 × 10-2).

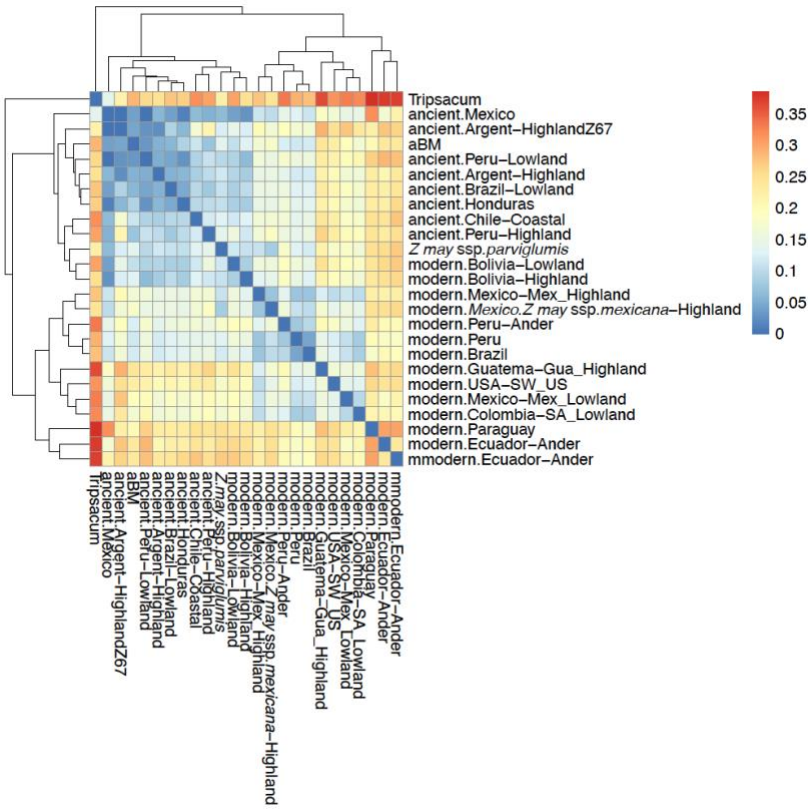

Figure AA 4-12: Fixation index (Fst) of heat-resistance selection on modern maize compared to ancient maize. Heatmap of Fst of heat relative biological process between modern and ancient maize.

Table AA 4-1: Geographical information about the ancient maize samples used in this study.

ID	Country	Age (BP)	Location	Latitude	Longitude	Altitude (m above sea level)
Arica-4	Chile	990 +/- 30	Arica, Chile, coastal	18.47S	70.32W	coastal
Arica-5	Chile	780 +/- 30	Arica, Chile, coastal	18.47S	70.32W	coastal
Z2	Brazil	570 +- 60	Brazil: Peruacu Valley, Boquete cave	15S	44W	700
Z6	Brazil	630+- 60	Brazil: Peruacu Valley, Lapa de Hora	15S	44W	700
Z61	Peru	800 +/- 30	Peru: Ancash	8.531S	78.341W	200
Z64	Peru	630 +/- 30	Peru: Ancash	9.064S	77.564W	3990
Z65	Peru	970 +/- 30	Peru: Inca	14.433S	75.342W	220
Z66	Argentina	1010 +/- 30	Argentina: Catamerca	27.342S	66.550W	2465
Z67	Argentina	100 +/- 30	Argentina: Jujuy	23.240S	66.211W	3700
EG84	Honduras	1870 – 1740	El Gigante rock shelter	14.22N	88.06W	NA
EG85	Honduras	2300 – 2070	El Gigante rock shelter	14.22N	88.06W	NA
EG90	Honduras	2300 – 2120	El Gigante rock shelter	14.22N	88.06W	NA
Tehu-a162	Mexico	5310	Mexico: Tehuacan, Puebla	18.46N	97.39W	NA
SM3	Mexico	4190 +/- 30	Mexico: San Marcos cave, Tehuacan Valley	17.48N	97.03W	NA
SM10	Mexico	4240 +/- 30	Mexico: San Marcos cave, Tehuacan Valley	17.48N	97.03W	NA

* The latitude and longitude information for SM3 and SM10 were determined based on the Tehuacan Valley location as provided in 'Myxomycetes associated with dryland ecosystems of the Tehuacán-Cuicatlán Valley Biosphere Reserve, Mexico."

Table AA 4-2: Paleogenetic characterization of six archaeological Bolivian maize sequence samples.

Sample	QC-passed reads	Total number of read mapped	Read mapped (%)	Mean Coverage	Std Coverage
aBM_combbam_rm5nt.srt.bam	334,380,500	311,508,504	93.16%	5.0472X	24.0833X
Bolivian_Maize_1_ATCACG_ L001_rm5nt.srt.bam	62,275,438	58,228,120	93.50%	0.9291X	4.5697X
Bolivian_Maize_2_CGATGT_ L001_rm5nt.srt.bam	75,746,744	70,631,420	93.25%	1.1524X	5.631X
Bolivian_Maize_3_TTAGGC_ L001_rm5nt.srt.bam	57,680,293	52,864,311	91.65%	0.9049X	4.9229X
Bolivian_Maize_4_TGACCA_ L001_rm5nt.srt.bam	11,737,197	10,683,928	91.03%	0.1852X	1.1184X
Bolivian_Maize_5_ACAGTG_ L001_rm5nt.srt.bam	11,871,062	11,121,964	93.69%	0.1716X	0.9095X
Bolivian_Maize_6_GCCAAT_ L001_rm5nt.srt.bam	115,069,766	107,978,761	93.84%	1.7039X	8.1491X
aBM_combbam_rm5nt.DR.bam	247,125,512	224,253,516	90.74%	3.487X	9.3418X
Bolivian_Maize_1_ATCACG_ L001_rm5nt.DR.bam	46,786,487	42,739,169	91.35%	0.6574X	1.921X
Bolivian_Maize_2_CGATGT_ L001_rm5nt.DR.bam	55,157,767	50,042,443	90.73%	0.778X	2.1457X
Bolivian_Maize_3_TTAGGC_ L001_rm5nt.DR.bam	42,799,698	37,983,716	88.75%	0.6169X	1.9418X
Bolivian_Maize_4_TGACCA_ L001_rm5nt.DR.bam	9,908,815	8,855,546	89.37%	0.1475X	0.7011X
Bolivian_Maize_5_ACAGTG_ L001_rm5nt.DR.bam	10,471,348	9,722,250	92.85%	0.1469X	0.6374X
Bolivian_Maize_6_GCCAAT_ L001_rm5nt.DR.bam	82,001,397	7,4910,392	91.35%	1.1404X	2.7747X

Table AA 4-3: Total number of genomic sites covered at variable depths in the archaeological Bolivian maize (aBM) sample (aBM_combbam_rm5nt.DR).

Depth	aBM_Combbam
>1X	66.36%
>2X	47.80%
>3X	36.02%
>4X	27.89%
>5X	21.99%
>6X	17.57%
>7X	14.19%
>8X	11.59%
>9X	9.55%
>10X	7.94%

Table AA 4-4: Estimated proportions of ancestry coming from source populations for archaeological Bolivian maize (aBM).

Target	source1	source2	source 1_ proportion	source 2_ proportion	stderr_ source 1	stderr_ source 2	nsnps_ used	nsnps_ target	pvalue
aBM	aMexico	aPeru	0.201	0.799	0.71	0.71	44721	829374	0.989846
aBM	aMexico	aArgentina	0.122	0.878	0.62	0.62	48319	829374	0.679729
aBM	aMexico	aBrazil	9.17	-8.17	43.5	43.5	49309	829374	0.671428
aBM	aMexico	aChile	0.052	0.948	1.27	1.27	54282	829374	0.0836292
aBM	aMexico	aHonduras	-0.017	1.07	0.464	0.464	46376	829374	0.611175

Table AA 4-5: Shared SNP numbers between maize from ancient Peru and ancient Mexico, and from ancient Peru and ancient Honduras.

Group	Shared SNP	Target gene
aPeru Highland (3990m; Z64) vs aMexico SM310	1,328	0
aPeru Highland (3990m; Z64) vs aMexico Tehua162	39,383	6
aPeru Lowland (200m; Z61, Z65) vs aMexico SM310	124	0
aPeru Lowland (200m; Z61, Z65) vs aMexico Tehua162	2,624	0
aPeru Highland (3990m; Z64) vs aHonduras (EG84, EG85, EG90)	64,234	0
aPeru Highland (200m; Z61, Z65) vs aHonduras (EG84, EG85, EG90)	10,829	0

Table AA 4-6: Shared SNP number and target gene number within 5kb between archaeological Bolivian maize (aBM) and ancient maize group (log^P=1X10^{-6}).

Group	Shared SNP	Target gene
aBM vs aPeru Highland (3990m; Z64)	443,211	206
aBM vs aArgentina Highland (2465m; Z66)	591,129	259
aBM vs aChile (Coastal; Arica4, Arica5)	407,921	26
aBM vs aBrazil (700m; Z2, Z6)	152,521	3
aBM vs aPeru Lowland (200m; Z64)	31,002	0
aBM vs aHonduras (EG84, EG85, EG90)	41,302	0
aBM vs aMexico (Tehua162)	24,580	0

Table AA 4-7: Target gene number within 5kb of top 1% XP-CLR.

Group	Target gene
Modern vs ancient	191
Modern vs ancient in South America	229
aBM vs aChile (Coastal; Arica4, Arica5)	26
aBM vs aBrazil (700m; Z2, Z6)	3
aBM vs aPeru Lowland (200m; Z64)	0
aBM vs aHonduras (EG84, EG85, EG90)	0

Data Set AAs

Data Set AA 4-1: Modern Sample information.

Data Set AA 4-2: Target gene of significant selection on maize between ancient Peru vs ancient Honduras, and ancient Peru vs ancient Mexico.

Data Set AA 4-3: Target gene of selection on maize between aBM vs ancient Brazil (700m; Z2, Z6), aBM vs ancient Chile Coastal (Aricia4, Aricia5), aBM vs ancient Argentina Highland (2465m; Z66), and aBM vs ancient Peru Highland(3990m; Z64).

Data Set AA 4-4: Gene Ontology Enrichment and target genes in Gene Ontology Enrichment and UniProt.GO of aBM vs ancient Brazil (700m; Z2, Z6), aBM vs ancient Chile Coastal (Aricia4, Aricia5), aBM vs ancient Argentina Highland (2465m; Z66), and aBM vs ancient Peru Highland(3990m; Z64).

Data Set AA 4-5: Go Terms and fixation index (Fst) are related to flowering, starch, cold, light, drought and heat.

Data Set AA 4-6: Target genes of modern vs ancient, and target genes of modern vs ancient in South America with top 1% XPCLR & XPEHH with $p<0.01$.

Data Set AA 4-7: Target genes of top 1% XPCLR in. modern vs ancient and modern vs ancient in South.

Data Set AA 4-8: Gene Ontology Enrichment of target genes of top 1% XPCLR in modern vs ancient in the western hemisphere, and modern vs ancient in South America.

Table AA 6-1: Type of cis-acting element.

Cycle	Myc	0	1	0	2	0	1	0	0	1	1	0
	MYC	0	0	4	3	11	2	2	8	1	0	10
Development	CAT-box	2	0	0	0	0	0	0	0	0	0	0
Hormone	ABRE	2	0	4	0	0	1	3	4	0	0	6
Stress	AE-box	0	4	0	0	0	1	0	0	1	1	0
	as-1	0	0	0	4	9	0	3	2	0	0	2
	DRE core	0	3	0	0	6	0	0	0	0	0	0
	GATA-motif	2	0	2	0	3	0	0	1	0	0	0
	G-box	0	0	6	0	0	0	1	0	0	0	2
	GT1-motif	2	0	0	0	0	1	0	0	1	1	2
	MBS	2	3	2	6	0	0	0	1	0	0	0
	STRE	0	1	0	1	2	1	2	0	1	1	2
	TCT-motif	0	0	0	0	5	0	1	0	1	2	0
	W box	6	4	2	3	3	0	0	0	0	0	2
	CAG-motif	0	0	0	0	0	0	0	0	0	0	2
	GA-motif	0	0	0	0	3	0	0	0	1	0	0
	Gap-box	0	0	0	0	3	0	0	0	0	0	0
	Myb	2	3	4	6	3	0	0	2	1	0	3
	MYB	4	4	4	11	3	4	2	6	1	7	8
	MYB-like	2	0	2	0	0	2	2	1	1	3	2
Transcription	CAAT-box	0	5	2	9	6	3	2	10	4	6	14
	TATA-box	38	41	20	27	62	2	8	35	32	13	15
	AT~TATA-box	4	0	0	3	21	0	2	13	9	1	1
	Motif name	ADF1	ADF2	ADF3	ADF4	ADF5	ADF6	ADF7	ADF8	ADF9	ADF10	ADF11

Table AA 6-2: Ka/Ks ratios of collinearity *ADF* genes in *Arabidopsis thaliana*.

No.	Gene Name	Chr	Gene Name	Chr	Ka	Ks	Ka/Ks	Duplication Event
1	*ADF5* (AT2G16700)	2	*ADF3* (AT5G59880)	5	-2	-2	1	segmental duplication
2	*ADF9* (AT4G34970)	4	*ADF3* (AT5G59880)	5	-2	-2	1	segmental duplication
3	*ADF7* (AT4G25590)	4	*ADF10* (AT5G52360)	5	-2	-2	1	segmental duplication
4	*ADF5* (AT2G16700)	2	*ADF9* (AT4G34970)	4	0.10	0.88	0.11	segmental duplication
5	*ADF* (AT3G45990; Putative)	3	*ADF3* (AT5G59880)	5	0.23	0.63	0.37	segmental duplication

Table AA 6-3: Collinearity relationship of *ADF* gene in *Arabidopsis thaliana* with *Glycine max*, *Zea mays*, and *Oryza sativa*.

No.	Gene Name	Sequence ID
1	AT3G45990.1	AT5G59880.1
2	AT5G59880.1	KRH32308
3	AT5G59880.1	KRH19714
4	AT5G59880.1	Os05t0113400-01
5	AT5G59880.1	Zm00001eb266570_P001
6	AT2G16700.1	AT4G34970.1
7	AT2G16700.1	KRH29274
8	AT2G16700.1	KRH24296
9	AT2G16700.1	KRH51402
10	AT2G31200.1	KRG95660
11	AT2G31200.1	KRH67374
12	AT4G25590.1	AT5G52360.1
13	AT4G25590.1	KRH35309
14	AT4G25590.1	KRG91521
15	AT2G16700.1	AT4G34970.1
16	AT4G34970.1	KRH29274
17	AT4G34970.1	KRH24296
18	AT4G34970.1	KRH51402
19	AT1G01750.1	KRH41156
20	AT1G01750.1	Os02t0663800-01
21	AT1G01750.1	Os04t0555700-01
22	AT3G45990.1	AT5G59880.1
23	AT3G45990.1	Os05t0113400-01
24	AT4G25590.1	AT5G52360.1
25	AT5G52360.1	KRH35309
26	AT5G52360.1	KRG91521