Fine-mapping Major Domestication QTL on Chromosome Five in *Zea mays*

By

Alessandra M. York

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

Doctor of Philosophy
(Genetics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2018

Date of final oral examination: 07/23/2018

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

John F. Doebley, Professor, Genetics
Richard M. Amasino, Professor, Biochemistry
Natalia de Leon, Professor, Agronomy
Bret A. Payseur, Professor, Genetics
Jerry M. Whitmore Jr., Faculty Associate, CALT/WISCIENCE
Brian S. Yandell, Professor, Statistics

Acknowledgements

First, I want to thank John Doebley for advising me over the past five years. As I started graduate school I could not think of a better fit for my research interests to pursue for my thesis and I'm grateful he gave me that chance. John has shaped me into a better scientist over the years through encouraging me to own my own work and feel confident in myself. I am also very grateful for John letting me pursue my interests in teaching and outreach. He told me very early on that I should own my time in graduate school and make the most out of my time since it is my education. I embraced those words and pushed myself to try anything that interested me, even if it was out of my comfort zone. For this and more, I am very grateful for my experiences.

I have also had the support of wonderful lab members during this time. This dissertation would not be possible without the advice, support or assistance I've received. Everyone has helped me in some sort of way from field and greenhouse work, to assistance in lab experiments, analyses or even by making the lab an enjoyable place to be. I will never forget our Florida trips, the lab March Madness tournament challenges, birthday parties or various lab meals. CJ, Mike, Eric III, Elizabeth, Adam, Brian, David, Wei, Liyan and Qiuyue, thank you for all your support. I also want to thank all of the undergrads that have passed through the lab over the years, especially Lora, Craig, Brandon, Kyle, Jack, Sam, and Bailey. These students had assisted in my thesis project in one form or another and I'm grateful for all their hard work. They also made mentoring a joyful experience that I will never forget.

I feel the saying "it takes a village to raise a child" is very fitting for my experience in graduate school. I have had the wonderful experience of meeting diverse groups of people that have helped shape my graduate career. There has been our collaboration through Panzea that

has exposed me to a broad set of research within the maize genetics field. Through this I've collaborated with others at Cornell University and Iowa State University. I've also had wonderful classmates through the genetics program who have been valuable resources about questions I have about topics in genetics, as a source for edits, or even just for moral support. I would also like to thank my committee for being a wonderful, supportive resource over my career. They have helped shape my thesis into what it is today with their advice and resources, as well as support me in my goals to teach. I was also able to find a community of people that support and foster my love for STEM education and have helped me grow along the way. I am very grateful for Christopher Trimby for giving me the opportunity to be a WISCIENCE Scientific Teaching Fellow and encouraging me to pursue a teaching-as-research project. He's become a valuable mentor to me. He introduced me to Jerry Whitmore, who has become a wonderful advisor for my education research and deserves much of the credit for chapter three of this thesis. Jerry always pushes me to better myself and I am very grateful for all of his time. Cara Theisen has been an incredible source of knowledge and support and challenged me to be a better educator. Devin Wixon also was an amazing support through my time in Delta and has made me a better critical thinker. The teaching fellows and Delta interns provided a diverse, passionate community to learn and grow with. I am forever appreciative of all the people above to help me get where I am today.

Finally, I am grateful for my family and friends for being so supportive through my career. My parents have encouraged me from day one to pursue my dreams and have done nothing but cheer me on since. To my brothers, Dominic and Tristan, thank you for making me laugh and your support. My in-laws have become a second set of parents to me who always ask

about my progress and cheer me on. Many of my friends have been there since childhood and will always be there when I need it. Most of all, thank you to my husband Nathan. Since we were fifteen years old, you have been my friend, lab and study partner, competition at times, and now lifetime companion. You have picked me up when I was down, challenged me when I needed a push, or comforted me when I needed support. None of this would have been possible without you (and snuggles with Lily cat).

Table of Contents

Acknowledgements ..... *i*

Table of Contents ..... *iv*

List of Figures ..... *vii*

List of Tables ..... *viii*

Abstract ..... *x*

Preface ..... *xi*

1 Fine-mapping a major domestication QTL for ear size traits on chromosome five in Zea Mays ... 1

1.1 Abstract 2

1.2 Introduction 3

1.3 Materials and Methods 7

1.3.1 Line development 7

1.3.2 Fine-mapping and Genotype-by-Sequencing 9

1.3.3 Phenotypic analysis 10

1.3.4 Statistical analyses and segregation of phenotypes 10

1.3.5 QTL analysis 11

1.4 Results 12

1.4.1 Phenotypes 12

1.4.2 RC-NIL generation 13

1.4.3 Genotype-by-Sequencing 14

1.4.4 QTL fail to segregate 14

1.4.5 Examining the upstream region 26

1.5 Discussion 28

1.5.1 Complex genetic architecture of ear size traits 28

1.5.2 Pollen Incompatibility 31

1.6 Future Work 31

1.7 Conclusion 33

1.8 Acknowledgements 33

2 Fine-mapping a major domestication QTL for ear shattering trait on chromosome five in Zea Mays 34

2.1 Abstract 35

2.2 Introduction 36

2.3 Materials and Methods 38

2.3.1 Fine-mapping 38

2.3.2 *Phenotypic analysis* 40

2.4 Results 44

2.4.1 Phenotypes 44

2.4.2 Fine-mapping of DIS5.1 46

2.4.3 Main and interaction effects of DIS5.1 and DIS1.1 48

2.4.4 Population genetics 55

2.5 Discussion 57

2.5.1 YABBY gene function and shattering 57

2.5.2 yab6 and role in shattering 58

2.5.3 Selection at yab6 61

2.6 Future Directions 62

2.7 Conclusion 64

2.8 Acknowledgements 65

3 Exploring the Impacts of Peer Mentoring on First-Year STEM Students 66

3.1 Abstract 67

3.2 Literature Review 67

3.3 Program Descriptions 72

3.4 Purpose 76

3.5 Methods 77

3.5.1 Participants 77

3.5.2 Data collection and analysis: surveys 78

3.5.3 Data collection and analysis: focus groups 80

3.6 Limitations 81

3.7 Results 82

3.7.1 Demographics 82

3.7.2 Quantitative findings: survey 83

3.7.3 Qualitative findings: survey 87

3.7.4 Focus groups: students 102

3.7.5 Focus groups: peer leaders 103

3.8 Discussion 106

3.8.1 The Effects for Underrepresented Populations in STEM 107

3.8.2 Peer Leader Support (PLS) Chats 109

3.9 Future Directions 110

3.10 Conclusion 111

3.11 Acknowledgements 112

Appendices 113

1 Fine-mapping a major domestication QTL for ear size traits on chromosome five in *Zea Mays* 114

A.1 Figures 115

A2: Tables 118

2 Fine-mapping a major domestication QTL for ear shattering trait on chromosome five in *Zea Mays* 125

B.1: Figures 126

B.2: Tables 128

3 Exploring the Impacts of Peer Leader Mentoring on Freshman STEM Students 145

C.1: Peer Leader Support (PLS) Chat List 146

C.2: Survey 147
C.3: Student Focus Group Script 148
C.4: Peer Leader Focus Group Script 150

References *152*

List of Figures

Figure P.1: Maize and Teosinte Ear Architecture. xii
Figure P.2: Maize and Teosinte Inflorescence Architecture. xiv

Figure 1.1: Line Development of RC-NILs. 8
Figure 1.2: Segregation of RC-NILs. 15
Figure 1.3: LSMs of B11b RC-NILs for Ear Diameter (ED). 16
Figure 1.4: LSMs of B11b RC-NILs for Kernels Per Rank (KPR). 17
Figure 1.5: LSMs of B11b RC-NILs for Kernel Row Number (KRN). 18
Figure 1.6: LSMs of B46 RC-NILs for Ear Diameter (ED). 19
Figure 1.7: LSMs of B46 RC-NILs for Kernels Per Rank (KPR). 20
Figure 1.8: LSMs of B46 RC-NILs for Kernel Row Number (KRN). 21
Figure 1.9: Ear Diameter QTL Maps. 23
Figure 1.10: Kernels per Rank Trait QTL Maps. 24
Figure 1.11: Kernel Row Number QTL Maps. 25
Figure 1.12: Reassessing Region of Significance. 27

Figure 2.1: Fine-mapping of DIS5.1 on chromosome 5s. 47
Figure 2.2: Interaction Data. 50
Figure 2.3: Teosinte as a dominant allele. 53

Figure 3.1: Comparing average scores of each question at each timepoint of the survey. 84
Figure 3.2: Peer leaders help students gain a better sense of how to become involved at UW-Madison. 85
Figure 3.3: Students feel peer leaders are qualified to be mentors. 86
Figure 3.4: Students report gaining skills from their peer leader. 101

Supplemental Figure A.1: Chromosome Five QTL from F2 Populations. 115
Supplemental Figure A.2: Shannon's QTL mapping of the ear diameter trait (top) and kernel row number (bottom) in the BC2S3 population. 116
Supplemental Figure A.3: Regions of domestication QTL on the short arm of chromosome five. 117

Figure B.1: Sequence Alignment between P6 and P7. 126

List of Tables

Table 1.1: BC6S6 lines used for original ear diameter crosses. 6

Table 2.1: Mixed Linear Models used to produce least squared means for fine-mapping RCNILs. 41
Table 2.2: F2 Population Main Effect Results. 49
Table 2.3: RIL Population Main Effect Results. 52
Table 2.4: F2 interaction data. 54
Table 2.5: Results for different selection tests on different regions around DIS5.1. 56
Table 2.6: Expression data from Lemmon *et al.* (2014). 60

Table 3.1: Mentoring Definitions 69
Table 3.2: Survey statistics examining differences between race/ethnicities from first survey timepoint. 88
Table 3.3: Independent samples test for differences between race/ethnicities at first survey timepoint. 89
Table 3.4: Survey statistics examining differences between race/ethnicities from last survey timepoint. 91
Table 3.5: Independent Samples Test for differences amongst race/ethnicities at last survey timepoint. 92
Table 3.6: Survey statistics examining differences between gender from first survey timepoint. 94
Table 3.7: Independent Samples Test for differences amongst gender at first survey timepoint. 95
Table 3.8: Survey statistics examining differences between gender from last survey timepoint. 97
Table 3.9: Independent Samples Test for differences amongst gender at last survey timepoint. 98

Supplemental Table A.1: Markers for Original RC-NIL Design. 118
Supplemental Table A.2: LSMs developed from B11b for ear size phenotypes. 119
Supplemental Table A.3: LSMs of B46 derived lines for both ear size phenotypes. 122
Supplemental Table A.4: Markers for New RC-NIL Design. 124

Supplemental Table B.1: Primer sets used for fine-mapping. 128
Supplemental Table B.2: Combined shattering LSMs. 129
Supplemental Table B.3: 2014 Fine-mapping data for all traits. 130
Supplemental Table B.4: 2015 Fine-mapping data for Shattering Trait. 132

Supplemental Table B.5: Primer sets used to genotype the F_2 population segregating for both DIS1.1 and DIS5.1. 133

Supplemental Table B.6: LSMs of F2 population shattering interaction data. 134

Supplemental Table B.7: Stocks used for RIL interaction experiment. 135

Supplemental Table B.8: LSMs of RIL interaction data. 137

Supplemental Table B.9: Primers used for Population Genetics Analysis. 138

Supplemental Table B.10: Samples used for diversity of *yab6*. 139

Supplemental Table B.11: Distribution of Specific Lines used for Diversity Sequencing for each marker. 140

Supplemental Table B.12: Results of F2 population analyses to examine epistasis and gene action between DIS1.1 and DIS5.1. 141

Supplemental Table B.13: Results for RIL Population Analyses examining epistasis between DIS1.1 and DIS5.1. 142

Supplemental Table B.14: Structural variants of *yab6*. 143

Supplemental Table B.15: Allele specific expression of *yab6* with five or more reads from Lemmon *et al.* (2014). 144

Abstract

Maize was domesticated from its wild ancestor, teosinte, approximately 9000 years ago and serves as an excellent model to study rapid evolution. This is because maize and teosinte provide an example of extreme morphological divergence in both its plant architecture and structure of its ears. Even with their drastic phenotypic differences, they are able to develop fully fertile offspring that allows us to study the genes responsible for the domestication process. To understand this process better, quantitative trait loci (QTL) mapping has been conducted to identify causal regions of the genome for many domestication traits. One major region of large effect is located on chromosome five. Specifically, many QTL associated with ear morphology are localized to this region. This work has aims to study this and has manifested into two different projects, both examining major domestication QTL identified on chromosome five. One investigates multiple ear size phenotypes co-localizing to the same region, while the other studies a known homolog for domestication of the nonshattering trait in other cereals. Both QTL had been previously identified in different maize-teosinte hybrid mapping populations and were excellent candidates for further study. Each project sheds light on a poorly understood region of the maize genome and the limitations of studying phenotypic traits with a complex genetic architecture. Together, this work demonstrates the complexity of the evolutionary process. In addition to this work, the final chapter highlights my interest in teaching and learning and delves into research that promotes diversity in STEM education by examining the effect of peer mentoring in a first-year STEM classroom.

Preface

The domestication of crop plants provides an excellent model to study rapid evolution of phenotypes. In fact, Charles Darwin highlighted domestication and artificial selection in the first chapter of his landmark book on evolution *On the Origin of Species by Means of Natural Selection* (Darwin 1859). Perhaps inspired by Darwin, plant biologists have invested considerable effort in the study of domestication from an evolutionary perspective. In this process much has been learned. Dozens of genes of large effect on domestication traits have been identified and it is clear that for some traits a simple genetic architecture with genes of large effect was involved in the process of domestication (Purugganan and Fuller 2009; Olsen and Wendel 2013).

Maize (*Zea mays ssp. mays*) and teosinte (*Zea mays ssp. parviglumis*) are an excellent model system to study genetic architecture because of the available genetic resources, drastic phenotypic differences between maize and teosinte, and a long history of study. According to archaeological and genetic evidence, maize was domesticated from teosinte in the Balsas River valley in Mexico around 9000 years ago (Doebley 2004; Piperno *et al.* 2009). The domestication of teosinte into maize led to an increase in yield and ease of harvestability, or the ability to gather ripened crop. To increase yield, there was an increase in ear diameter, kernel row number, cupules per rank, seeds per ear, ear length and fruitcase (kernel) weight (Figure P.1). As to increase harvestability, there was a decrease in tiller number, decrease in lateral branch length and number of branches, decrease in ear shattering, and a change in the inflorescence

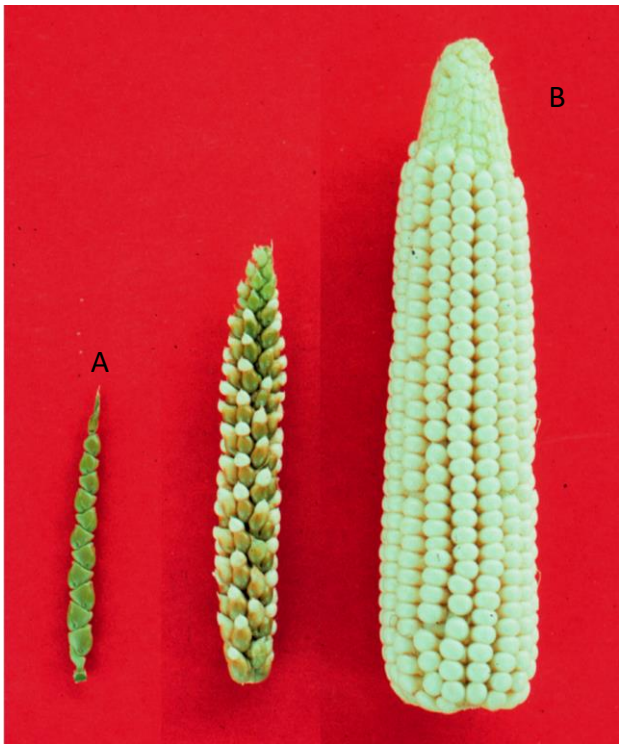


Figure P.1: Maize and Teosinte Ear Architecture.
Teosinte (A) and maize (B) ears differ greatly for almost all ear traits: kernel row number, ear diameter, ear length, kernel weight, seeds per ear, number of cupules per rank and ear internode length. Picture credit: John Doebley.

architecture to become more concentrated to one to two ears and one tassel per plant (Figure P.2). These vast differences make it easy to dissect the genetic architecture of domestication traits because of the wide range of phenotypes seen in hybrid populations. The large library of molecular tools also makes maize-teosinte a great system to study domestication. Teosinte and maize, when crossed, have fertile offspring, allowing us to investigate the genetic differences using maize-teosinte hybrid mapping populations (Beadle 1932; Doebley and Stec 1991, 1993). In addition, many molecular techniques have been optimized in maize (Hake and Ross-Ibarra 2015), it has a reasonable quality reference genome (Schnable *et al.* 2009), and it has a cheap robust genotyping system available (Elshire *et al.* 2011). All of these factors contribute to making teosinte-maize a good system to study the genetic architecture of domestication traits.

Domestication caused a drastic change in the size of the ear of maize as compared its ancestor, teosinte (Figures P.1 and P.2). Teosinte has many (potentially 50 or more) small ears at the nodes of its lateral branches, and each of these ears have few (8-12) kernels. Maize, in contrast, has only a few large ears (typically 1 or 2 in elite lines) that are borne at the end of short lateral branches, and each of these ears have a large number (often 300 or more) of kernels. This change in ear size was brought about by artificial selection during domestication and it allowed ancestral people to harvest the grain more easily since picking two large ears off a plant is faster than picking 50 small ears. Effectively, this change in ear size and number improved the harvestability of the maize relative to teosinte. The fragility of the ear also drastically changed during the process of maize domestication. Teosinte's kernels are formed from an invaginated internode, or cupule, within which the kernel sits, and the glume covers the opening of the cupule, so the kernel is completely hidden. At maturity, disarticulation of the

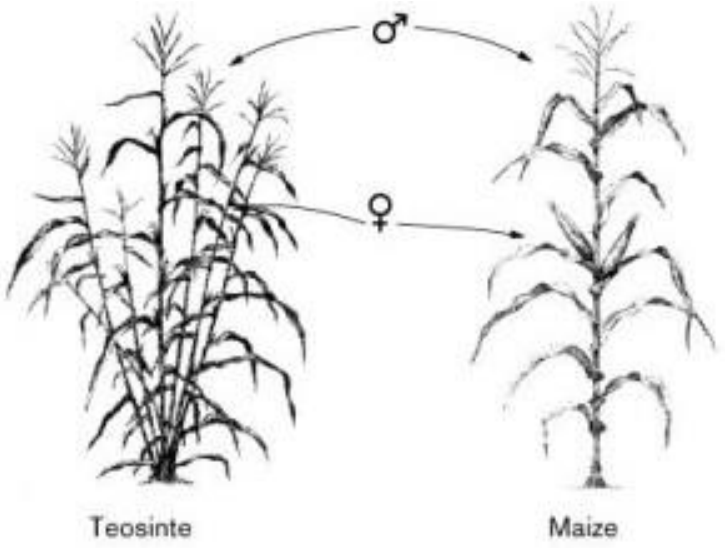


Figure P.2: Maize and Teosinte Inflorescence Architecture.
Maize has fewer inflorescences overall and concentrates in specific regions. Teosinte has numerous longer lateral branches, and many ears along the branches. Modified from Iltis (1983) and Kellogg (1997).

teosinte ear would occur between the fruitcases. The kernel and its protective covering function as a means of biotic seed dispersal. In contrast, kernels of a maize ear have this glume and cupule structure but are reduced in size and develop into the cob structure. This difference appears to be in part due to *teosinte glume architecture (tga1)*, a transcription factor on chromosome four shown to be responsible for exposing the kernel from the hardened protective casing seen on teosinte seed (Wang *et al.* 2005). However, the development of the cob was not the only factor preventing disarticulation. The strength of the cob would have also been selected for by ancestral people to ease harvest. This change resulted in a reduced fragility of the ear since disarticulation no longer occurred, and therefore improved the harvestability of maize relative to teosinte. These changes were essential to make maize the crop it is today. I am interested in investigating the genes underlying the changes in the maize ear during the process of domestication.

It is believed that the genetic architecture of domestication traits is complex, or quantitative, in nature than traditional Mendelian genetics. Quantitative traits are influenced by multiple genes as well as their environment. They typically have a range of phenotypes that echo a normal distribution and tend to have multiple wild-type alleles. Since the early 20th century, geneticists have worked to develop methods for mapping quantitative traits to better understand complex trait variation (Sax 1923). For most of this time geneticists were limited by the lack of available markers. However, since the late 1980's, marker technology has improved dramatically due to development of DNA-based assays, discovery of abundant molecular markers such as microsatellites and indels, a reduction in the cost of DNA sequencing, and the availability of reference genome sequences for many organisms (Mackay *et al.* 2009). With

these advances and the development of better statistical methods for QTL mapping (Lander and Botstein 1989; Broman *et al.* 2003), exploring genetic architecture of complex traits has become easier than ever before. By mapping QTLs in the segregating progeny of crosses between genetically divergent strains, geneticists can answer questions about adaptive evolution. Such studies have answered questions about the genes responsible for the marine to freshwater adaptation of the stickleback fish (Jones *et al.* 2012), tomato fruit size (Alpert *et al.* 1995), seed shattering in rice (Konishi *et al.* 2006), albinism in cave fish (Protas *et al.* 2006) and vernalization in Arabidopsis (Johanson *et al.* 2000) just to name a few.

The work presented in this dissertation sets out to explore the genes responsible for the two phenotypes described above, ear size and disarticulation (shattering) in domesticated maize compared to its ancestral counterparts. This is done in two chapters:

The first chapter describes a fine-mapping project for an ear size QTL discovered on chromosome five that appeared in multiple domestication QTL mapping experiments conducted by the lab but had not yet been successfully fine-mapped due to a complex genetic architecture. I took a subset of lines from a previous lab experiment that had a more isogenic nature to develop homozygous recombinant inbred lines for our QTL. This QTL was not successfully mapped to a single gene, and the QTL ended up shifting to a different region outside our original support interval. While not successful, this study gave us insight on the complex genetic architecture of domestication QTL, as well as how to better design an experiment if one is to examine this QTL in the future.

The second chapter describes a separate fine-mapping project for the shattering trait on chromosome five. A shattering QTL on chromosome five had been identified from a previous

lab study, and within that QTL was a known homolog for a shattering gene in sorghum. Our strategy used lines that segregated for our gene of interest but were isogenic elsewhere in the genome. The causative region mapped directly upstream of the 5'untranslated region (UTR) of the candidate gene, *yab6*. Interaction occurs between this gene, and the other shattering homolog in maize on chromosome one, *ZmYAB2.1*, an enhancer of *teosinte branched* (tb1) found to correlate with differences in ear internode length. This work leads to a better understanding of how different genes altered during maize domestication evolved in concert.

In addition to my work fine-mapping maize domestication traits on chromosome five, I also conducted a teaching-as-research project through the Delta Program in Research, Teaching and Learning's Internship at the University of Wisconsin-Madison and continued the project through my last year in graduate school. This work focuses on the impact that peer mentoring has on first year STEM students in a seminar course. Literature on STEM education has consistently shown that peer mentors in the classroom have contributed to the success of students, especially in minority groups (Chesler and Chesler 2002; Bowling *et al.* 2015; Snyder *et al.* 2016; Zaniewski and Reinholz 2016). However, why peer mentors are beneficial and how they help the students has yet to be examined. Our study is interested in examining what contributes to the beneficial relationship of having peer mentors in the classroom. We specifically examined a case with a large first-year STEM seminar course, Integrated Sciences 100: Exploring Biology, that implements eight peer leaders across its two sections of approximately 100 students each. These peer leaders have been trained through Integrated Mentoring Program and Core Training (IMPaCT), which uses theory from the field of educational leadership to become effective mentors in the classroom. The study implemented

used mixed methods through a sequential explanatory design. First, surveys were conducted across the fall semester to gather quantitative and qualitative data. Students that complete the surveys were invited to participate in the focus groups in the following spring to understand the complexities of this relationship and its benefits. Focus groups were also held with amongst the peer leaders to gain their perspective as well. Our overall goal was to ask the students questions about why they find having a peer mentor in the classroom is beneficial and examine if improvements should be made to the peer mentoring training program. Preliminary data does indeed suggest peer mentors are beneficial, and we plan to use this as a pilot for more extensive studies to solidify our claim in the near future.

Chapter 1:

Fine-mapping a major domestication QTL for ear size traits on chromosome five in *Zea Mays*

1.1 Abstract

Maize ears are much larger in diameter and have more rows of grain than maize's ancestor, teosinte. This significant difference in ear structure makes it a useful trait to study to further understand domestication. The focus of this work is to uncover the causative polymorphism responsible for a major domestication quantitative trait locus (QTL) identified by multiple studies. Specifically, this QTL was recently mapped to a ~2.654 Mbp region in chromosome five. Three BC_6S_6 lines previously identified to segregate for the QTL of interest were selected to create two different families. Known markers were used to identify recombination events in the region and generated a set of 163 recombinant chromosome nearly isogenic lines (RCNILs) from the two families. All 163 RC-NILs were genotyped using genotype-by-sequencing (GBS) to identify the exact recombination breakpoints. GBS also identified additional unexpected regions outside our QTL that were segregating on both chromosomes five and seven. Lines were grown in six replicate blocks and phenotypes were collected for various ear traits including ear diameter, kernels per rank, and kernel row number. A linear mixed model was used to obtain the least squared means (LSMs) for each line and test the observed segregation of the phenotypes. The causal region appeared to map within 20 Mbp upstream of our identified target region. This was most likely due to previously unforeseen segregation in the lines found on a region of chromosome seven since a QTL for ear diameter had already been identified there (Shannon 2013). This data shows that an older, previously described location of ear size QTL by Shannon (2013) was probably more accurate, even though the population used to map the QTL was weaker and not as isogenic as other studies.

1.2 Introduction

The ear underwent dramatic changes during the process of domestication in maize. Teosinte, its ancestor, has many (potentially 50 or more) small ears at the nodes of its lateral branches, and each of these ears have few (8-12) kernels. Maize, in contrast, has only a few large ears (typically 1 or 2 in elite lines) that are borne at the end of short lateral branches, and each of these ears have a large number (often 300 or more) of kernels. This change in ear size was brought about by artificial selection during domestication and it allowed ancestral people to harvest the grain more easily since picking two large ears off a plant is faster than picking 50 small ears. Effectively, this change in ear size and number improved the harvestability of the maize relative to teosinte and one of the largest changes in the domestication process.

Six genomic regions appear to be the largest contributors to the process of domestication in maize (Doebley 2004). Quantitative trait loci (QTL) for a number of domestication traits, including plant and ear architecture traits, have been described (Doebley and Stec 1991, 1993; Briggs *et al.* 2007; Shannon 2013; Lemmon and Doebley 2014). Many of these previously described QTL fall on chromosome five. Some of the most significant domestication QTL found on chromosome five are related to ear size. Doebley and Stec (1991) first reported a large effect QTL for ear size on chromosome five in a maize-teosinte F2 mapping population and later reaffirmed its existence with a different maize-teosinte F2 mapping population (Doebley and Stec 1993) (Supplemental Figure A.1). Subsequently, using a more powerful maize-teosinte mapping population that was backcrossed twice to the maize parent and selfed three times (BC_2S_3), two QTL were identified in this region; one for ear diameter and one for kernel row number. The specific location of these QTL narrowed the interval to a 6.85 Mbp region between

145.98-152.83 Mbp (Shannon 2013). The ear diameter QTL was the largest of the entire domestication study, with a LOD score of 144.3 (Supplemental Figure A.2). Further attempts to fine-map this QTL using a heterogeneous inbred family (HIF) were without success, with both ear diameter and kernel row number not segregating cleanly into two different phenotypic groups (Lemmon 2014). Lemmon conducted a fine-mapping experiment very similar to this one, but the structure of the hybrid population had less crosses that resulted in larger segregating pieces of the genome. One likely reason for the inconclusive results from Lemmon's initial attempt to fine-map this QTL using the BC_2S_3 population is that the genetic background in the fine-mapping lines was too heterogeneous, segregating for multiple small effect QTL distributed throughout the genome. There are also some characteristics of the chromosome that lead to complications in recovering causal polymorphisms for domestication phenotypes. A relevant factor is that chromosome five carries the *gametophyte factor2 (ga2)*, a pollen incompatibility factor that can influence pollination rates of specific genotype combinations (Kermicle and Evans 2010). Also, a reduced rate of recombination around the centromere of chromosome five (102.3-109.2 Mbp) has been observed which makes it more difficult to recover recombinant chromosomes for any fine-mapping experiments in regions near the centromere.

However, using an independent maize-teosinte mapping population comprised of BC_6S_6 lines that were segregating for chromosome 5s, Lemmon and Doebley (2014) confirmed the existence of the QTL and more narrowly mapped them to a 2.654 Mbp region around 166.5-169.2 Mbp. This region included 54 candidate genes (Supplemental Figure A.3). The increased amount of backcrossing and selfing should have eliminated other segregating regions for

teosinte in the genome and developed smaller segregating blocks of teosinte on chromosome five to more narrowly map our QTL. The evidence that the two QTL co-localize in that region yet again suggest a single gene is responsible for influencing both traits. Lemmon and Doebley (2014) found that the major QTL identified originally in Shannon (2013) appear to fractionate and represent two or three linked QTL in their further studies. The linked QTL did not have equal effects; for each trait, one of the linked QTL had a greater effect than the others. The fractionation of QTL could have also been another potential complication that hindered Lemmon (2014)'s initial attempt to find a causal region. The largest QTL Lemmon and Doebley found for both ear diameter and kernel row number were *eard5.3* and *krn5.2* in terms of effect size and LOD score. The QTL co-localized to a physical position around 15-20 Mbp away from the original suspected location Shannon (2013) described at around 166-169 Mbp between molecular markers umc1348 and umc1966 on chromosome five (Table 1.1) (Supplemental Table A.1). This focused my investigation to this region of co-localized QTL.

In this study, I fine-mapped the linked QTL in order to rectify two potential complicating observations that Lemmon encountered in 2014. First, I used BC_6S_6 lines with an increased isogenic background. RFLP markers were used to eliminate teosinte segregating regions of the genome where domestication QTL had been found (Doebley and Stec 1993). Second, I focused on the largest of the linked QTL identified by Lemmon and Doebley (2014) while fixing the other two QTL to avoid the confounding of other segregating QTL for our population. To accomplish this, lines were developed that segregate for only the QTL of interest and were phenotyped in a randomized complete block design (RCBD). A mixed linear model was used to obtain least

Table 1.1: BC6S6 lines used for original ear diameter crosses.
A graphic to show the three lines we crossed to generate recombination in the QTL region of interest. This region is found between markers umc1348 and umc1966. The numbers at the top are the coordinates on chromosome five from AGP version 2. Below are markers used to map the breakpoints of maize and teosinte in all the BC_6S_6 lines. ”T”= teosinte, “M”= maize.

Line	AGPv2	92368559	1.12E+08	1.22E+08	1.42E+08	1.52E+08	1.59E+08	1.67E+08	1.69E+08	1.69E+08
		umc1224	umc1283	bnlg1287	dupssr10	bnlg2323	ZHL0301	umc1348	umc1221	umc1966
Line_B11b		T	T	T	T	T	T	M	M	M
Line_B04		T	T	T	T	T	T	T	T	T
Line_B46		T	T	T	T	M	M	M	M	M

squared means for each line and test their phenotypic segregation. These same methods have been used to detect genes of large effect that appear to be responsible for domestication of maize from its ancestor. Five past successful QTL fine-mapping projects have mapped QTL to single underlying genes in the reference genome (Doebley *et al.* 1995; Wang *et al.* 2005; Hung *et al.* 2012; Wills *et al.* 2013; Yang *et al.* 2016). The more recent experiments (Hung *et al.* 2012; Wills *et al.* 2013; Yang *et al.* 2016) all used specific inbred lines segregating for the QTL of interest, phenotyped the lines and calculated least squared means (LSMs) to find two distinct phenotypic classes that differ for maize versus teosinte haplotypes within a single genomic interval. The genomic interval identified is where the causal factor is found, and sometimes can be very precise. Therefore, I planned to use similar techniques from these previous experiments to investigate this region of interest for an ear size QTL defined by multiple previous studies (Doebley and Stec 1991, 1993; Shannon 2013; Lemmon and Doebley 2014).

1.3 Materials and Methods

1.3.1 Line development

A set of BC_6S_6 lines were developed that were homozygous for the maize inbred W22 throughout their genomes except for the short arm of chromosome five (Lemmon and Doebley 2014). Three of the BC_6S_6 lines (named B11b, B04 and B46) were chosen because they carry maize versus teosinte DNA for the largest effect ear size QTL described above (Table 1.1). In 2014 crosses were made among the three BC_6S_6 lines (Figure 1.1a) and then two F2 populations were generated by selfing eight F_1 plants for each: B11b x B04, and B46 x B04 (Figure 1.1). In 2015 the F_2 plants from these two populations were grown and genotyped using molecular

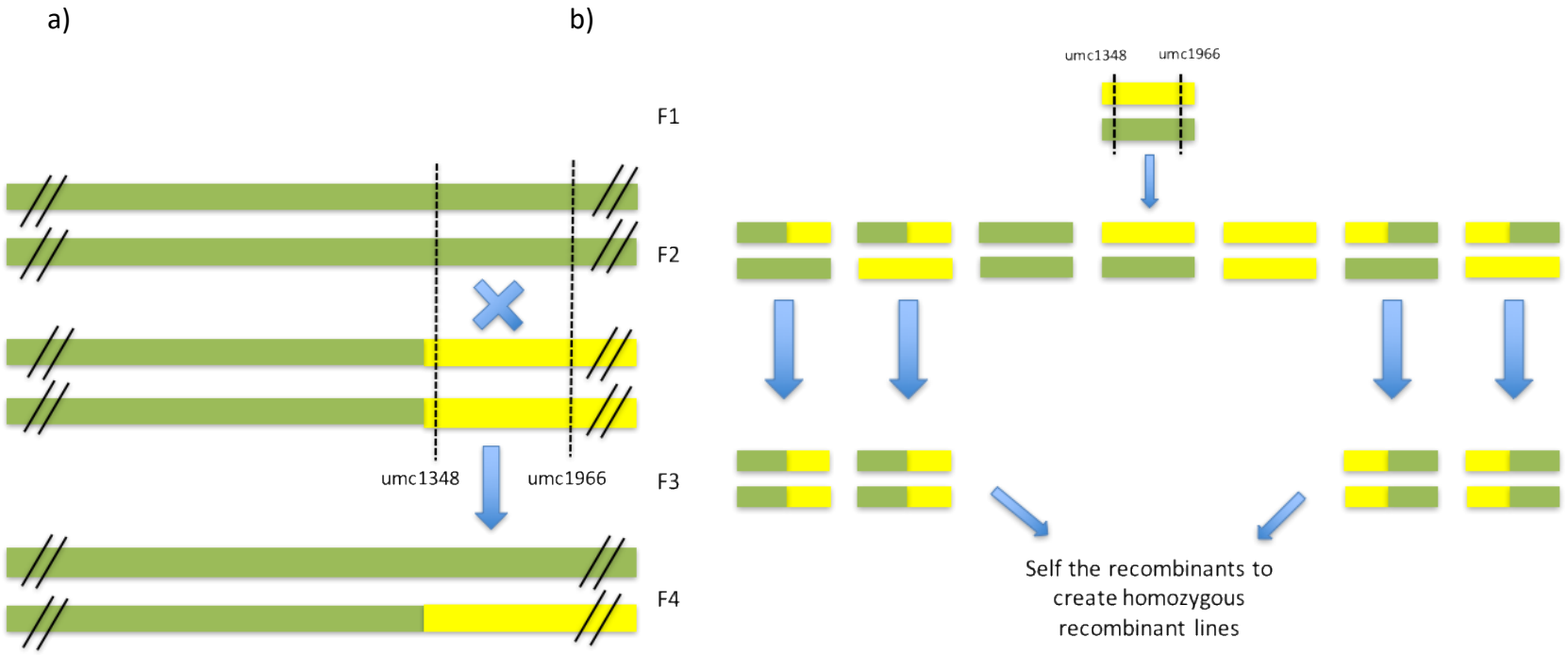


Figure 1.1: Line Development of RC-NILs.
Part a displays the first step of our line development, where individuals were crossed with teosinte over the full region of interest (B04) to lines that segregated for maize where our suspected QTL lie (B46 and B11b). Part b displays the subsequent steps in line development. The bars represent the genotype on each chromosomal arm between the two screening markers, umc1348 and umc1966. The F1 was selfed to generate recombination between our two markers. The F2 was screened to look for individuals that were homozygous maize or teosinte at one marker, and heterozygous for the recombination event at the other. Those identified with the specific genotype of interest were selfed. The F3 was screened for individual plants that are homozygous for the breakpoint, which was expected to be one out of every four. The F3 was selfed to generate the RC-NILs, which are lines that are homozygous for a breakpoint in between the two molecular markers. A total of 163 lines were developed.

markers umc1348 and umc1966 that flank the QTL interval to identify any recombinant chromosomes created by cross-overs (Figure 1.1b). DNA was extracted from seedling leaves in 96-well plates using a CTAB protocol (Doebley and Stec 1991). DNA preparations were used as the substrate for polymerase chain reaction (PCR) with fluorescently tagged primers for both umc1348 and umc1966 in a single reaction. PCR products were assayed on an ABI 3730 DNA Analyzer (Applied Biosystems) and the resulting chromatograms processed using Genescan software (Applied BioSystems). Any plant that had the recombinant genotype was self-pollinated in the field for a total of 170 plants. In fall of 2015, 30 F_3 seeds from each of the F_2 plants with recombinant chromosomes were planted at our winter nursery (27 Farms, Homestead, FL). Plants were genotyped to identify individuals homozygous for a recombination event between the two marker loci in the same manner as the previous field season. One or two plants for each of the 170 F_3 lines that carry the homozygous recombinant chromosome were self-pollinated to generate the set of recombinant chromosome nearly isogenic lines (RC-NILs) (Figure 1.1b).

1.3.2 Fine-mapping and Genotype-by-Sequencing

The exact recombination breakpoints were determined for the recombinant chromosomes of the RC-NILs through DNA extractions and sequencing. A 300 milligram sample of leaf tissue was collected from each of the homozygous recombinant F3 plants for DNA extraction using a modified CTAB method that removed RNA (Doebley and Stec 1991). This modification is necessary if the DNA is to be used for genotype-by-sequencing (GBS) (Elshire *et al.* 2011), which is a low coverage sequencing method that genotypes thousands of genetic markers across the genome. A combination of GBS markers, PCR-based (indels) markers, and SNPs scored by

Sanger sequencing were used to map the break-points in the recombinant chromosomes. By precisely mapping each breakpoint, association can be made between the line phenotypes and narrowly define regions within the QTL interval.

1.3.3 Phenotypic analysis

Each RC-NIL was grown in a randomized complete block design with four blocks in summer 2016 at the West Madison Agricultural Research Station (UW-WMARS). Individuals were planted with equal spacing in 16-foot (487.68 cm) rows with 1-foot (30.48 cm) in between plants within rows and three-foot (91.44 cm) between rows. Adjacent blocks were separated by 11 feet (335.28 cm). Six ears were harvested from each row of each block for a total of 36 ears phenotyped per recombinant. Ear diameter was measured with calipers in centimeters by measuring the middle of the ear. Kernels per rank was measured by counting the number of kernels vertically per ear. Kernel row number was measured by counting number of kernel rows around the circumference of the ear.

1.3.4 Statistical analyses and segregation of phenotypes

Each phenotypic trait collected was fit to the following linear model using MIXED procedure in SAS (Littell *et al.* 1998) to calculate least square means (LSMs) for each RC-NIL:

$$Y_{ij} = \mu + a_i + b_j + e_{ij}$$

where Y_{ij} is the phenotypic trait value, μ is the overall mean, a_i is the line effect, b_j is the block effect, and e_{ij} is the experimental error. Line effect was considered a fixed effect while the block effect was considered as a random effect.

To fine-map the ear diameter QTL interval correspondence was determined between the introgressed teosinte chromosomal segments and LSMs of the 163 RC-NILs. The

introgressed segments of the 163 RC-NILs were matched to their corresponding LSMs. The 163 RC-NILs were then sorted by their LSMs to test if the lines partition into two distinct phenotypic classes that differ for maize versus teosinte haplotypes within a single genomic interval. Under the assumption that there is a single gene contributing to the QTL, one expects a clean segregation of the lines into two groups based on their LSMs such that there is a small chromosomal interval for which all lines in one group have teosinte genotype and all lines in the other group have maize genotype. The interval thus identified is the causative interval that contains the QTL. Genotype and phenotype data are available in Supplemental Tables A.2 and A.3.

1.3.5 QTL analysis

QTL analyses were conducted using the program R/qtl (Broman *et al.* 2003). LSMs were used as the phenotypic values for each trait for QTL mapping. Genotype-by-sequencing provided the markers, but the est.map command allows us to calculate the genetic distance using the Haldane mapping function. Each trait is independently analyzed using simple interval mapping with the Haley-Knott method (Haley and Knott 1992) using the scanone function. This method considers each point or marker across the genome at a time as the location of a supposed QTL, splits the individuals into groups based on their genotypes to compare the groups' phenotypic averages using a t-test or an ANOVA (Broman and Sen 2009). Therefore, the model for each QTL is the following:

$$P_i=\beta(x_i)+e_i$$

Where P_i is the QTL, $\beta(x)$ is the phenotypic mean of the QTL for genotype x, x_i is the genotype at the QTL being examined for line i, and e_i is the experimental error. Therefore, three possible

means are examined at each QTL: β(MM), β(MT) and β(TT).

Each phenotypic trait's significance threshold at the p=0.05 level was determined by a permutation test (n=10,000). To verify if multiple QTL existed on either chromosome five or seven, the model was refined using a drop-one ANOVA and refineqtl command before running the addqtl command to search for additional QTL. If a new QTL was found, the ANOVA and refineqtl procedures were repeated with the new QTL added to the model. This was repeated until no new significant QTL were located.

1.4 Results

1.4.1 Phenotypes

Three ear size phenotypes were measured in the RC-NILs: ear diameter, kernel row number and kernels per rank. For those traits we did observe differences between the two families that create the RC-NILs. B11b consistently had smaller averages and variation than lines with B46 as its parent. However, modes were quite close, if not the same. Repeatability instead of heritability was calculated for each trait due to the inability to partition the genetic variance from the genetic-by-environmental variance (Fehr *et al.* 1987). It was consistent for each trait between the two parental lines. Ear diameter had the highest repeatability, with 0.713 for B11b offspring and 0.698 for B46 offspring. Kernels per rank had a slightly lower repeatability, with 0.630 for B11b offspring and 0.640 for B46 offspring. Kernel row number had a lower repeatability and was less consistent between lines than other traits, with B11b offspring at 0.367 and B46 offspring at 0.461.

For ear diameter, those with the B11b parent had an average diameter of 29.84 mm and

a variance of 2.48 mm^2. The smallest ear was 20.95 mm and the largest was 36.3 mm. The mode was 30.5 mm. Ears with the B46 parent were slightly larger on average with a value of 32.5 mm (same as the mode). The variance was similar to B11b at 2.94 mm^2. The smallest ear was 24 mm and largest was 38.9 mm.

For kernel row number, B11b offspring's average was 12.28 kernels with a variance of 1.07 $kernels^2$. The smallest value was 8 kernels and the largest was 16 kernels. The mode was 12 kernels per row. For B46 offspring, the average again was slightly larger at 13.19 kernels per row. The variance was also larger at 1.71 $kernels^2$. The smallest value was 10 kernels and the largest value was 18 kernels. The mode was the same at 12 kernels per row.

For kernels per rank, lines derived from B11b had an average of 22.38 kernels and a variance of 9.65 $kernels^2$. The smallest value was 10 kernels and the largest was 31 kernels. The mode was 23 kernels per rank. Lines derived from B46 had a slightly larger average of 24.09 kernels and a slightly increased variance of 11.63 $kernels^2$. The smallest value was 10 kernels and the largest was 33 kernels. The mode was also slightly larger at 25 kernels per rank.

1.4.2 RC-NIL generation

In 2015, a total of 3200 F2 plants were grown from the two separate parent crosses and each genotyped to identify any recombinant chromosomes created by cross-overs between the molecular markers that flank the QTL interval (Table 1.1). Overall, 170 plants were identified to have a recombination event in the QTL interval of interest and were kept for further studies. After screening the F3 generation for homozygous recombinant individuals, a total of 163 RC-NILs were recovered. Seven were lost due to mold or poor pollinations.

1.4.3 Genotype-by-Sequencing

The expectation was that segregation should only be observed from approximately 145,215,779 bp to 181,195,080 bp on chromosome five, which surrounds our region of interest. However, depending on the parent line, additional segregation was observed in other regions of the genome (Figure 1.2). The lines that were derived from the B46 parent also segregated on chromosome seven from 143,050,224 bp to 172,779,320 bp. The lines that were derived from the B11b parent segregated on chromosome five from 0 to around 5,903,953 bp and segregated in the same region on chromosome seven that B46 lines did. However, B11b lines were fixed for teosinte from 155,340,573 bp to 172,779,320 bp. This suggested that B04, the shared parent between the lines, was responsible for the segregation on chromosome seven. Consequently, the lines were separated into two parental groups based on their differing genotypes.

1.4.4 QTL fail to segregate

When the LSMs of each phenotype were organized from smallest to largest, the corresponding genotypes appeared to have little to no correlation with the phenotype. The results were not the clean segregation between the maize and teosinte phenotypes that were expected (Figures 1.3-1.8). The cleanest segregating phenotype was kernel row number in the B46 lines, where there was some correlation between phenotype and genotype but there were still obvious outliers (Figure 1.8). Therefore, QTL mapping was performed using R/qtl to understand where the causal regions were predicted to be (Broman *et al.* 2003). For the lines with the B11b parent, an ear diameter QTL was found to pass the significance threshold of p=0.05 on chromosome seven from approximately 146.3 Mbp – 155.4 Mbp with a LOD score of 7.1, where

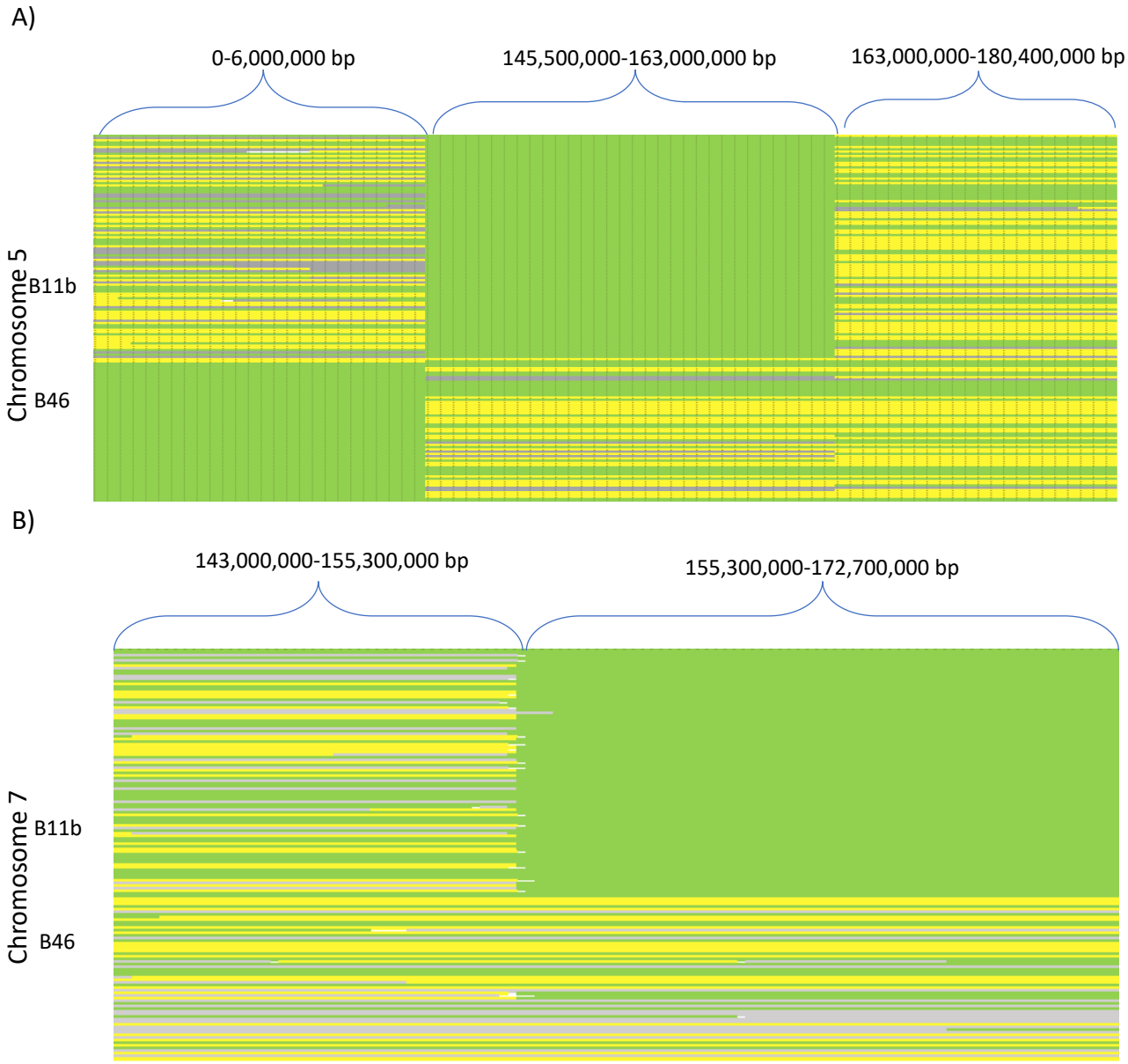


Figure 1.2: Segregation of RC-NILs.
A graphic to show the segregation of genotypes amongst the different RC-NILs. Figure 1.2A demonstrates the segregation pattern on chromosome five while figure 1.2B shows the segregation pattern on chromosome seven. Most of this genotype segregation was unexpected, with the exception of chromosome five past 145 Mbp. Segregation appears to be line dependent, with B11b fixed for teosinte on chromosome seven from 155.3 Mbp to 172.7 Mbp and B46 fixed on chromosome five from the beginning of the chromosome to 6 Mbp. Y-axis is each individual RC-NIL, and X-axis is the genotype at each of the markers. Yellow=maize, green=teosinte and grey=heterozygous

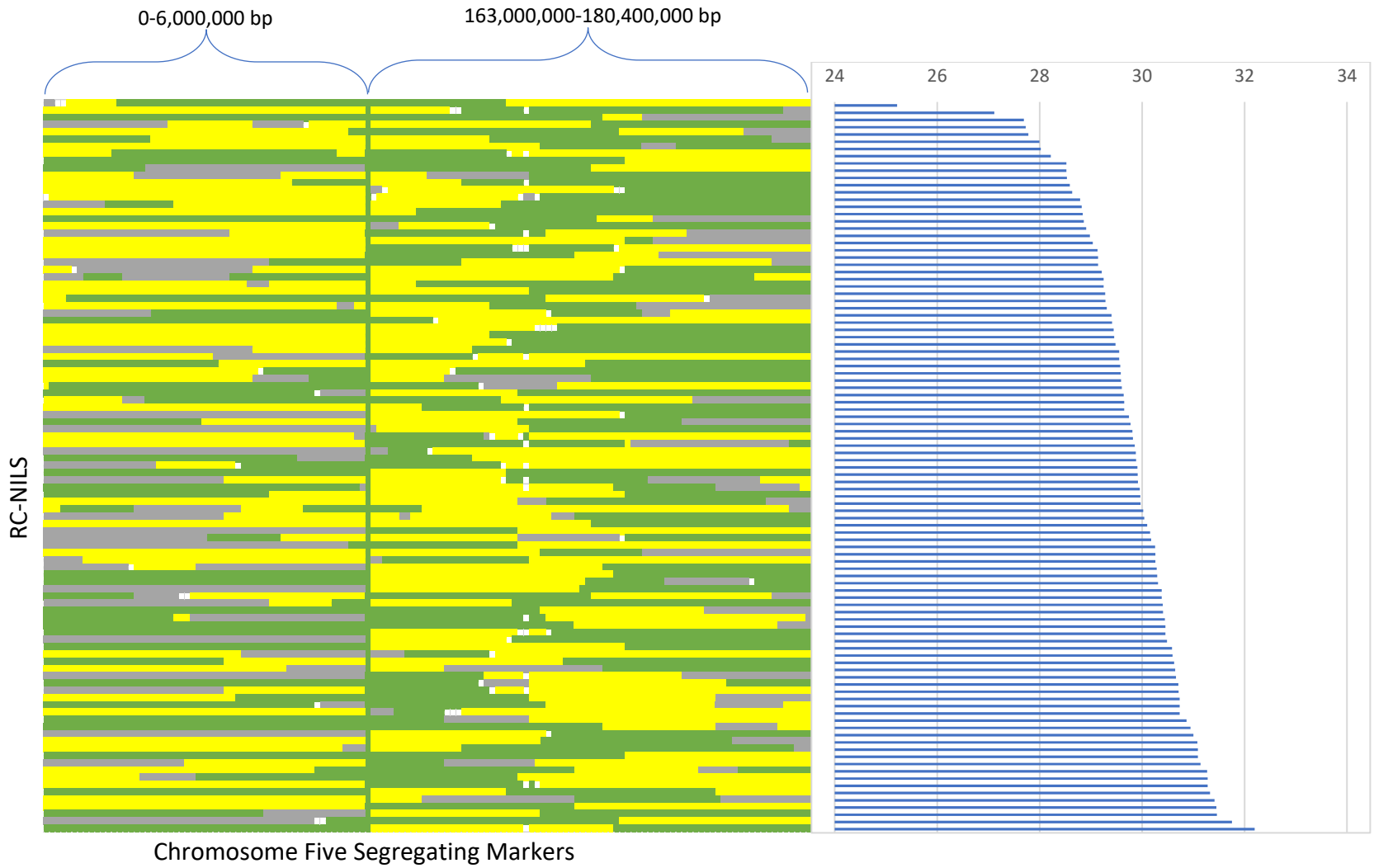


Figure 1.3: LSMs of B11b RC-NILs for Ear Diameter (ED).

A graphic representing the phenotyping of ear diameter in the B11b RC–NILs. RC-NILs are organized on the y-axis from smallest ED to largest ED, as demonstrated by the bar graph seen on the right. X-axis is the genotype at the marker in our specific region of interest on chromosome five. Everything left of the green line is the region segregating from approximately 0-6 Mbp, whereas everything to the right of the green line is approximately 163-180.4 Mbp. Yellow=maize, green=teosinte, grey=heterozygous and white=missing data.

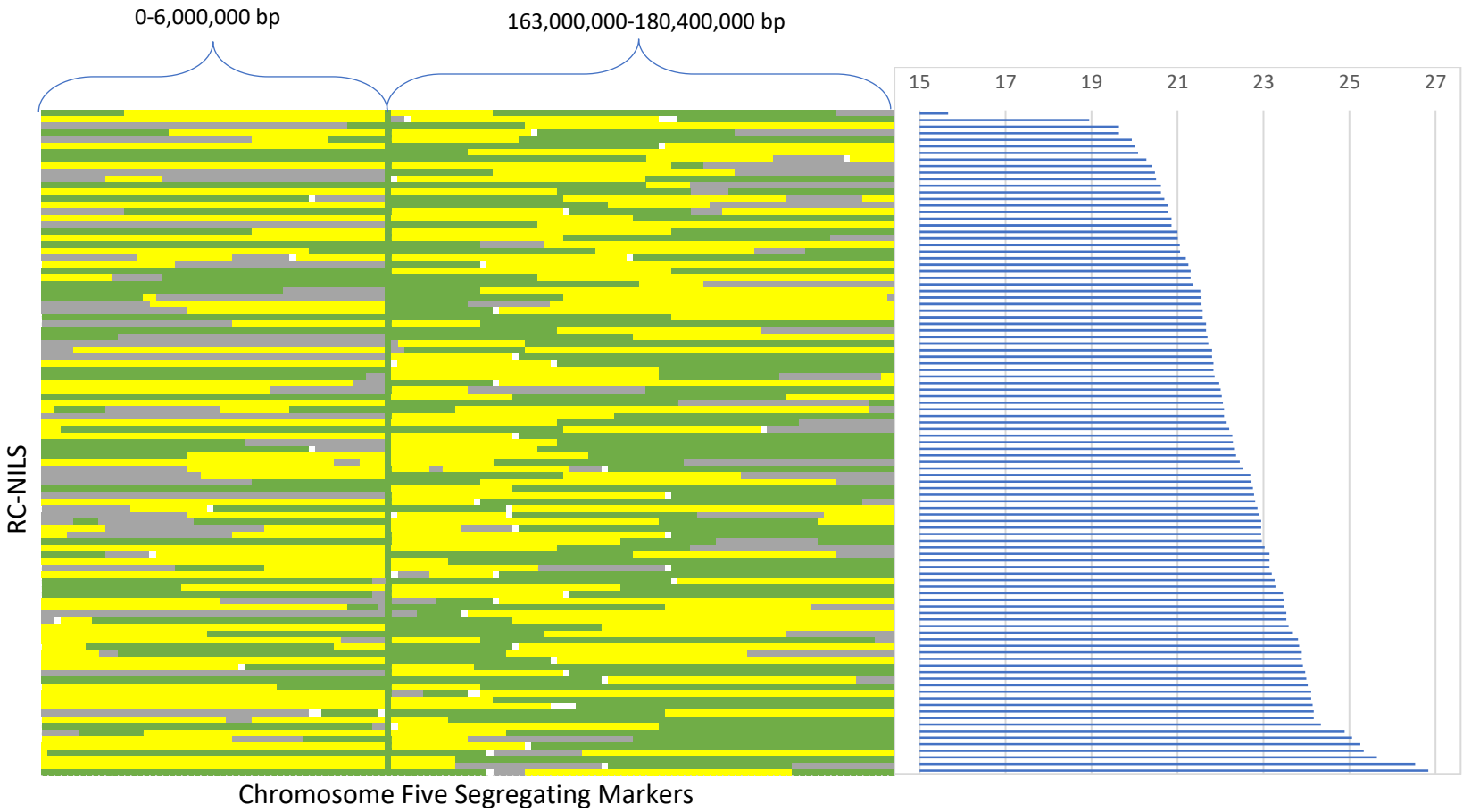


Figure 1.4: LSMs of B11b RC-NILs for Kernels Per Rank (KPR).

A graphic representing the phenotyping of kernels per rank in the B11b RC–NILs. RC-NILs are organized on the y-axis from smallest KRN to largest KPR, as demonstrated by the bar graph seen on the right. X-axis is the genotype at the marker in our specific region of interest on chromosome five. Everything left of the green line is the region segregating from approximately 0-6 Mbp, whereas everything to the right of the green line is approximately 163-180.4 Mbp. Yellow=maize, green=teosinte, grey=heterozygous and white=missing data.

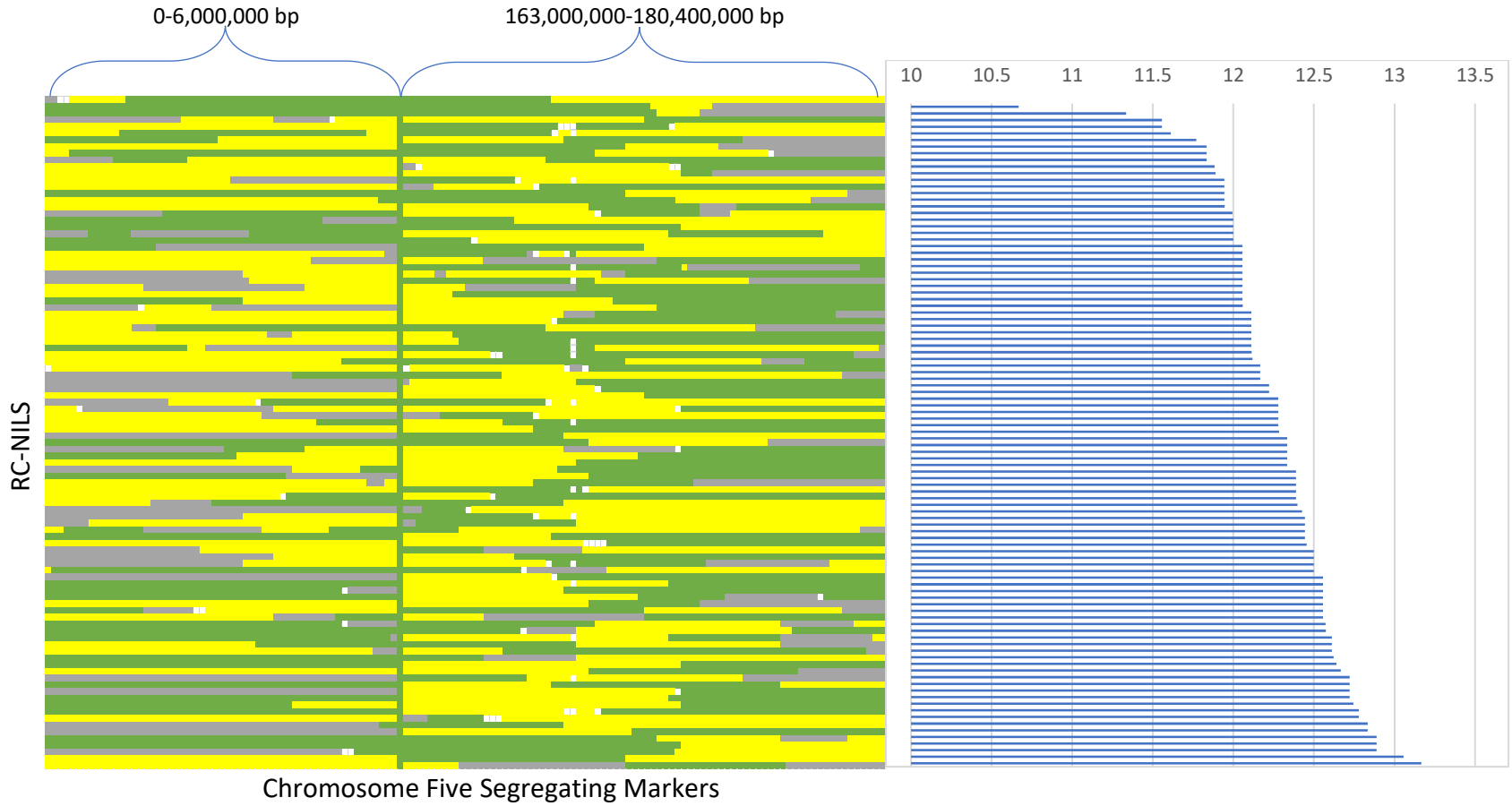


Figure 1.5: LSMs of B11b RC-NILs for Kernel Row Number (KRN).

A graphic representing the phenotyping of kernel row number in the B11b RC–NILs. RC-NILs are organized on the y-axis from smallest KRN to largest KRN, as demonstrated by the bar graph seen on the right. X-axis is the genotype at the marker in our specific region of interest on chromosome five. Everything left of the green line is the region segregating from approximately 0-6 Mbp, whereas everything to the right of the green line is approximately 163-180.4 Mbp. Yellow=maize, green=teosinte, grey=heterozygous and white=missing data.

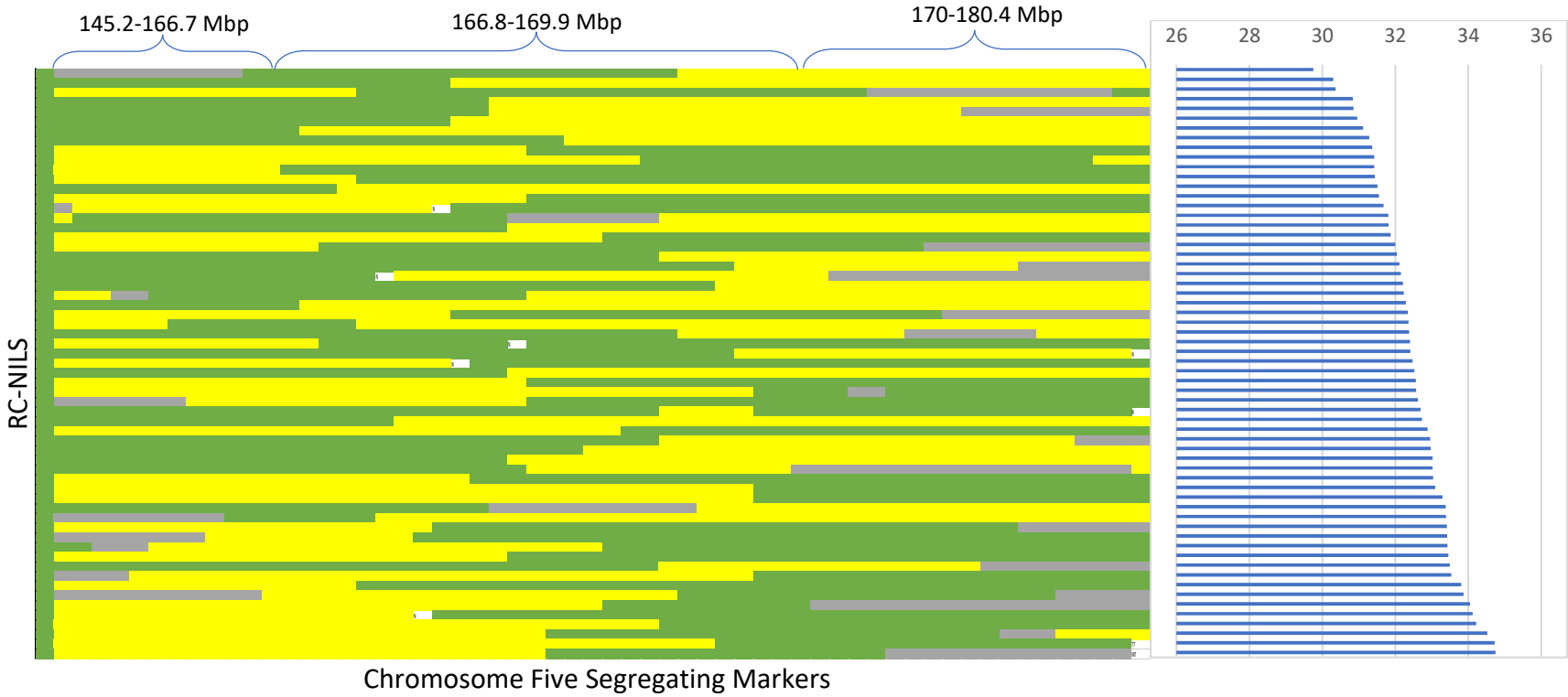


Figure 1.6: LSMs of B46 RC-NILs for Ear Diameter (ED).

A graphic representing the phenotyping of ear diameter in the B11b RC–NILs. RC-NILs are organized on the y-axis from smallest ED to largest ED, as demonstrated by the bar graph seen on the right. X-axis is the genotype at the markers on chromosome five from 145.5 to 180.4 Mbp. Yellow=maize, green=teosinte, grey=heterozygous and white=missing data.

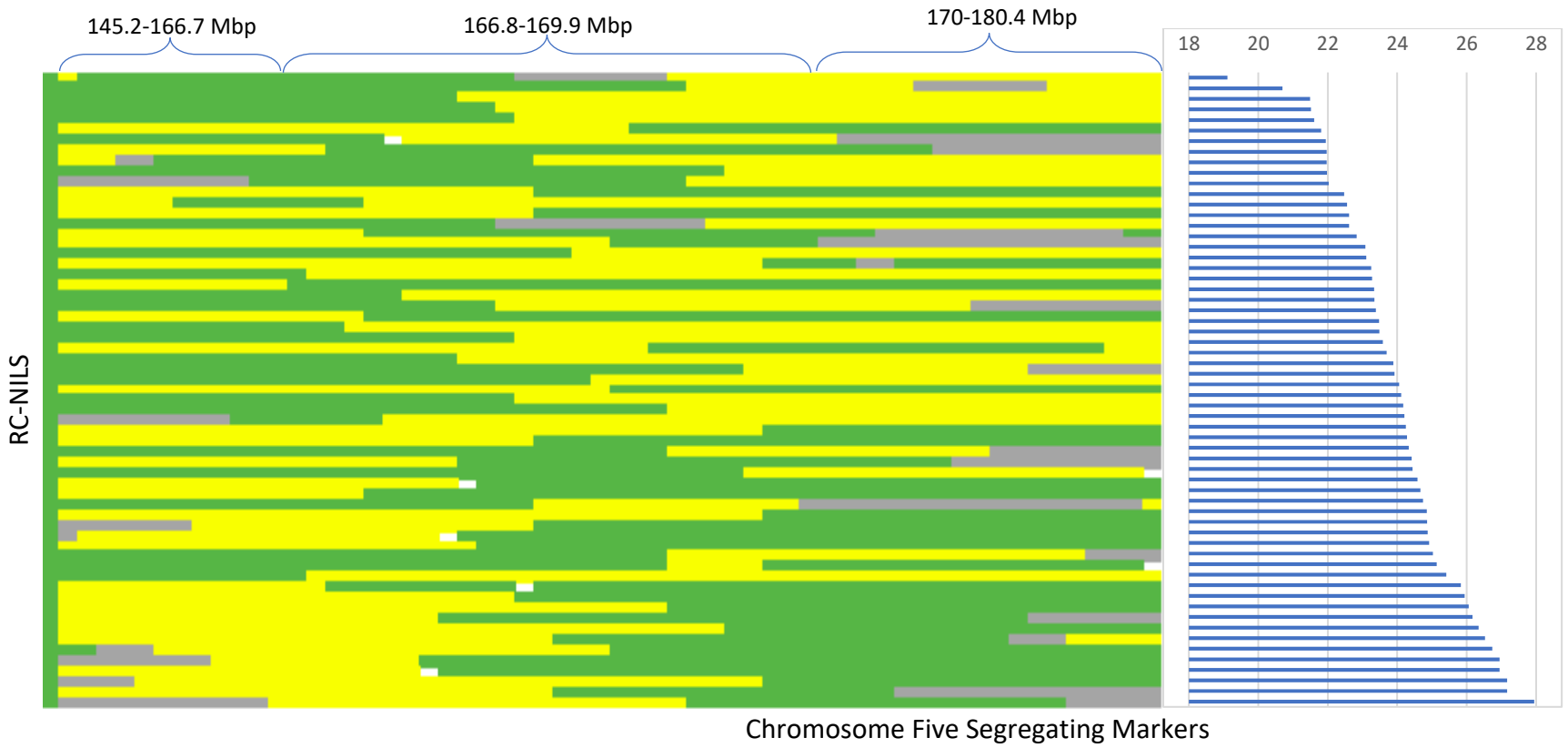


Figure 1.7: LSMs of B46 RC-NILs for Kernels Per Rank (KPR).

A graphic representing the phenotyping of kernels perk rank in the B46 RC–NILs. RC-NILs are organized on the y-axis from smallest KPR to largest KPR, as demonstrated by the bar graph seen on the right. X-axis is the genotype at the markers on chromosome five from 145.5 to 180.4 Mbp. Yellow=maize, green=teosinte, grey=heterozygous and white=missing data.

Figure 1.8: LSMs of B46 RC-NILs for Kernel Row Number (KRN).

A graphic representing the best results from our phenotyping of the RC -NILs. RC-NILs are organized on the y-axis from smallest KRN to largest KRN, as demonstrated by the bar graph seen on the right. X-axis is the genotype at the markers on chromosome five from 145.5 to 180.4 Mbp. Yellow=maize, green=teosinte, grey=heterozygous and white=missing data.

there was a previously identified domestication QTL for ear diameter (Figure 1.9) (Shannon 2013). Multiple QTL were found for the kernels per rank trait. A kernels per rank QTL was found near the ear diameter QTL on chromosome seven with a LOD score of 7.855, and it was the most significant for the trait (Figure 1.10). There was also a significant QTL on chromosome five, with a LOD score of 5.956. The QTL on chromosome five and seven were found to have a significant interaction ($p<.0001$), indicating that the causal factors underneath the underlying QTL work together to produce the phenotype. No kernel row number QTL was found to be significant in this population (Figure 1.11).

For the lines with the B46 parent, no QTL were identified for ear diameter (Figure 1.9). Kernels per rank had a significant QTL on chromosome five, but the target region shifted to 163.42-176.91 Mbp (Figure 1.10). Kernel row number also had a QTL on chromosome five with a LOD score of 7.94, but this QTL was found to be upstream of our previous target region from around 155.03-166 Mbp (Figure 1.11). Thus, both kernel row number and kernels per rank appear to have a "shadow peak". It is possible this is due to the flipping of genotypes on each side of our recombinant region (166.58-169.23 Mbp). The R/qtl program finds a similar correlation of genotype to phenotype on both sides of our recombinant regions making it appear that there are causal factors on either side of our recombinant region. The existence of a shadow peak is confirmed by searching for additional QTL, which is done by dropping the most significant QTL from each chromosome from the model. We found that the peak disappears when additional QTL are added to the model with the addqtl function. However, the "true" QTL for kernel row number and kernels per rank appear on opposite sides of the recombinant region. The QTL for kernel row number appears to be upstream of our causal region, while for kernels per rank the "true" QTL appears to be downstream, meaning the two

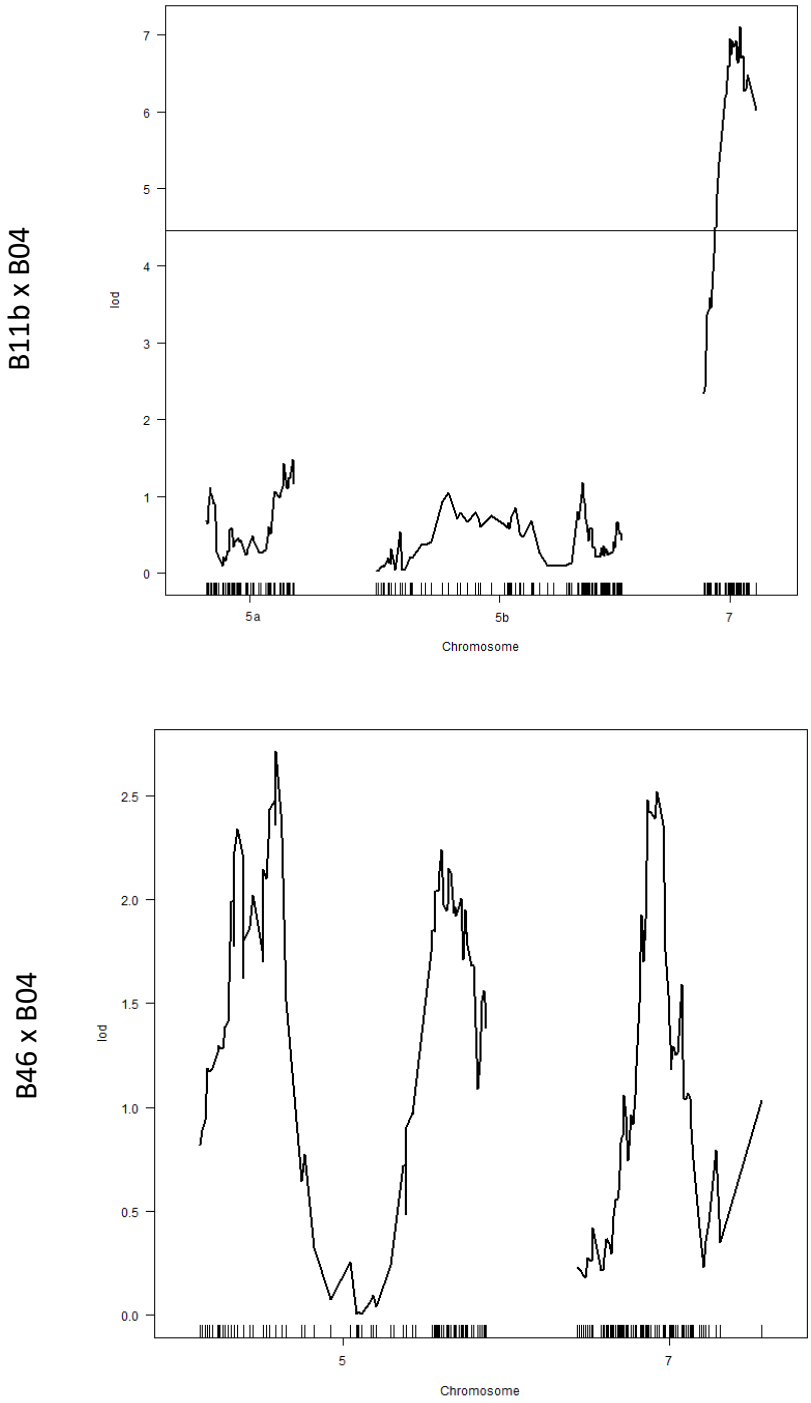


Figure 1.9: Ear Diameter QTL Maps.
The figure on top is the QTL graph for B11b, where there is one significant QTL on chromosome seven with a LOD score of 7.1. This peak explains 27.66% of the variance of the phenotype. Chromosome 5a in top figure on the x axis is the segregating region on chromosome five from approximately 0-6,000,000 bp. Chromosome 5b in the top figure on the x axis is the segregating region on chromosome five from approximately 163,000,000-180,000,000 bp. The figure on bottom is the QTL graph for the ear B46, with no significant peaks.

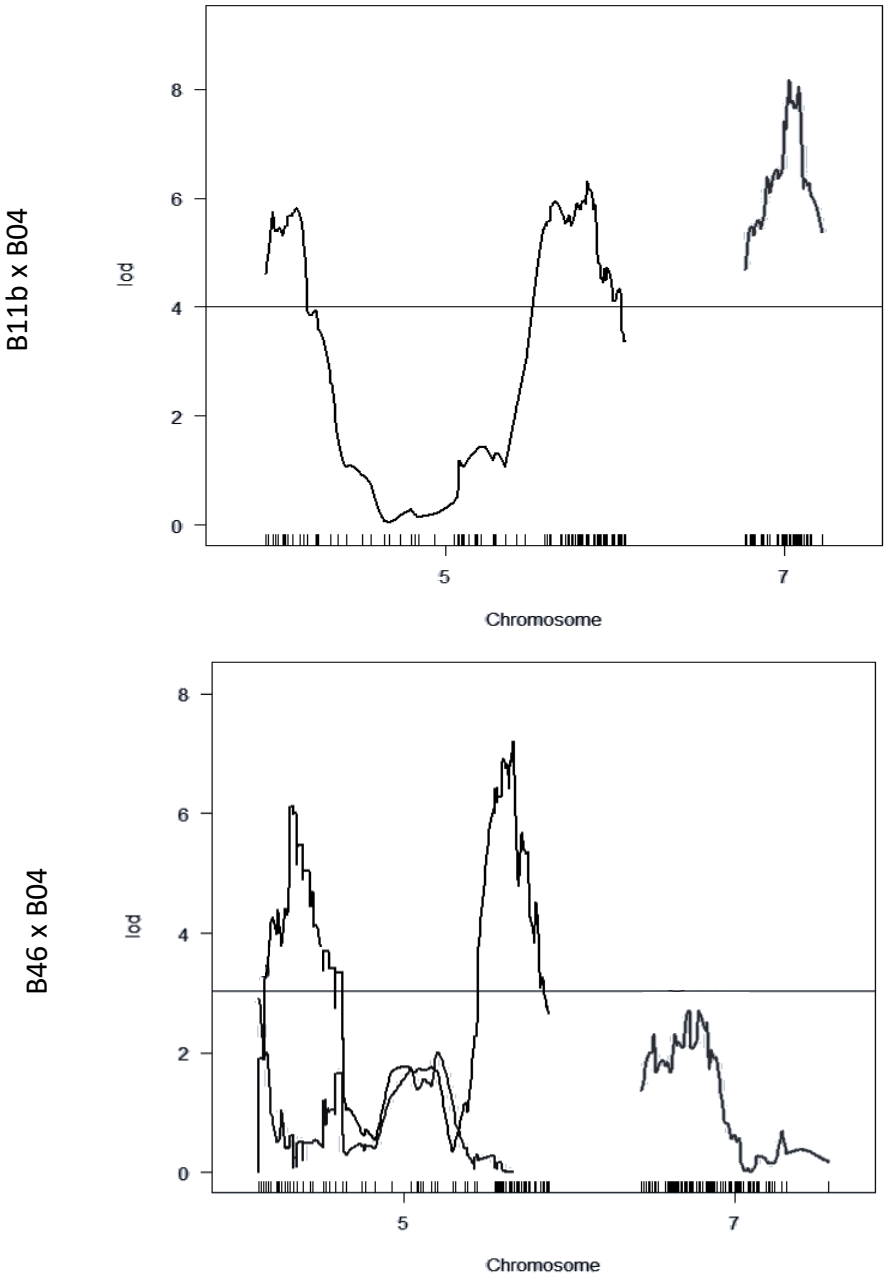


Figure 1.10: Kernels per Rank Trait QTL Maps.
Figure on the top is the QTL graph for B11b, where there are two significant peaks, one on chromosome five from 163.42-176.91 Mbp with a LOD score of 5.956 and the other on chromosome seven from 148.326-154.608 Mbp with a LOD score of 7.855. These QTL represent 18.9% and 26.1% of the variance of the phenotype, respectively. There is significant interaction between these two QTL (p=<0.0001). Figure on bottom is the B46 lines, where there's only one significant QTL on chromosome five from 147.541-172.951 Mbp with a LOD score of 6.873. This QTL accounts for 33.493% of the variance of the phenotype. The "peaks" on the right of our QTL is a shadow peak due to the flipping of genotypes on the other side of our recombinant region.

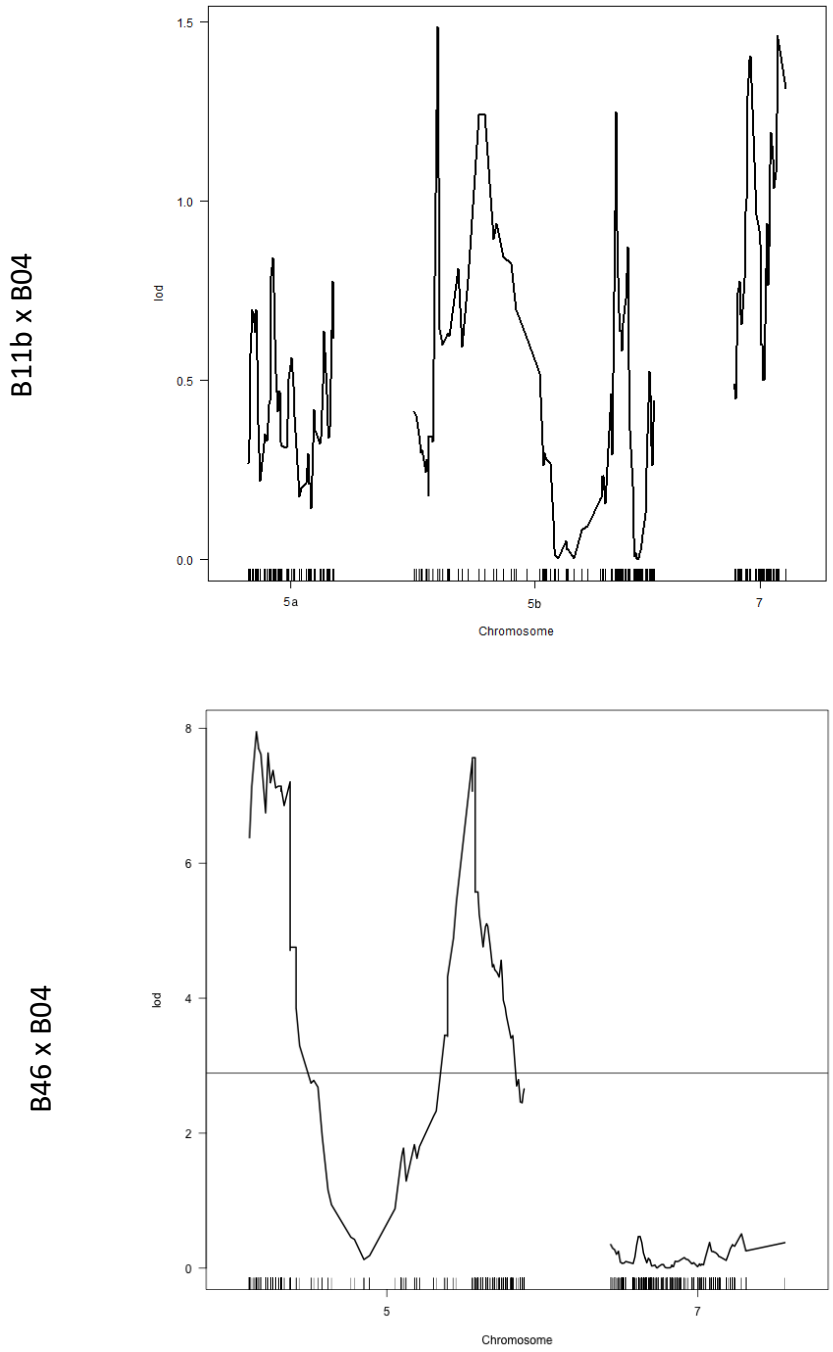


Figure 1.11: Kernel Row Number QTL Maps.
The figure on top is for the B11b lines, with no significant QTL peaks. Chromosome 5a in top figure on the x axis is the segregating region on chromosome five from approximately 0-6,000,000 bp. Chromosome 5b in the top figure on the x axis is the segregating region on chromosome five from approximately 163,000,000-180,000,000 bp. The figure on the bottom is the QTL graph for the B46 lines, where the QTL on chromosome five has a LOD score of 7.94. This peak explains 44.6% of the variance of the phenotype. The “peak” on the right of our QTL is a shadow peak due to the flipping of genotypes on the other side of our recombinant region.

phenotypes are probably not caused by the same factor.

1.4.5 Examining the upstream region

Fine-mapping efforts then became focused on examining the region upstream of our original target region (166-169 Mbp), from approximately 145 Mbp-166 Mbp. Approximately 3000 kernels were planted in 32-kernel flats made up of the four lines (A115-3, A285-1, 2221 and 1594), all heterozygous in our new region of interest, were screened in the greenhouse in early 2017 (Figure 1.12). DNA was extracted and samples were genotyped with their respective fluorescent-tagged or PCR molecular markers (Supplemental Table A.4). Plants that were identified as heterozygous for a recombinant event were selfed and kept for further study.

Although 200+ recombinants were originally identified, many failed to flower appropriately. The majority of plants (over 70%) had either ear growth stunted or failed to shed pollen. In total only 53 recombinants produced seed with 16 plants producing less than ten kernels.

Seed from the recombinants were planted in the field in summer 2017 in a completely randomized design (CRD) to be phenotyped for ear diameter, kernel row number and kernels per rank. Approximately 1350 individual kernels were planted in the field. The plants were grown with rows that were 30 inches (76.2 cm) apart and plants within rows 1-foot (30.48 cm) apart. Only 275 individuals germinated, likely due to inadequate pollination conduction in the previous season. DNA was extracted from the available plants as previously described and genotyped using fluorescent-tagged or PCR molecular markers. Plants that were identified to be homozygous for the recombinant event were subsequently selfed. If only heterozygous recombinants were produced for a specific genotype, they were also selfed to screen in future studies. In addition, if few individuals were homozygous for a recombinant event for a specific

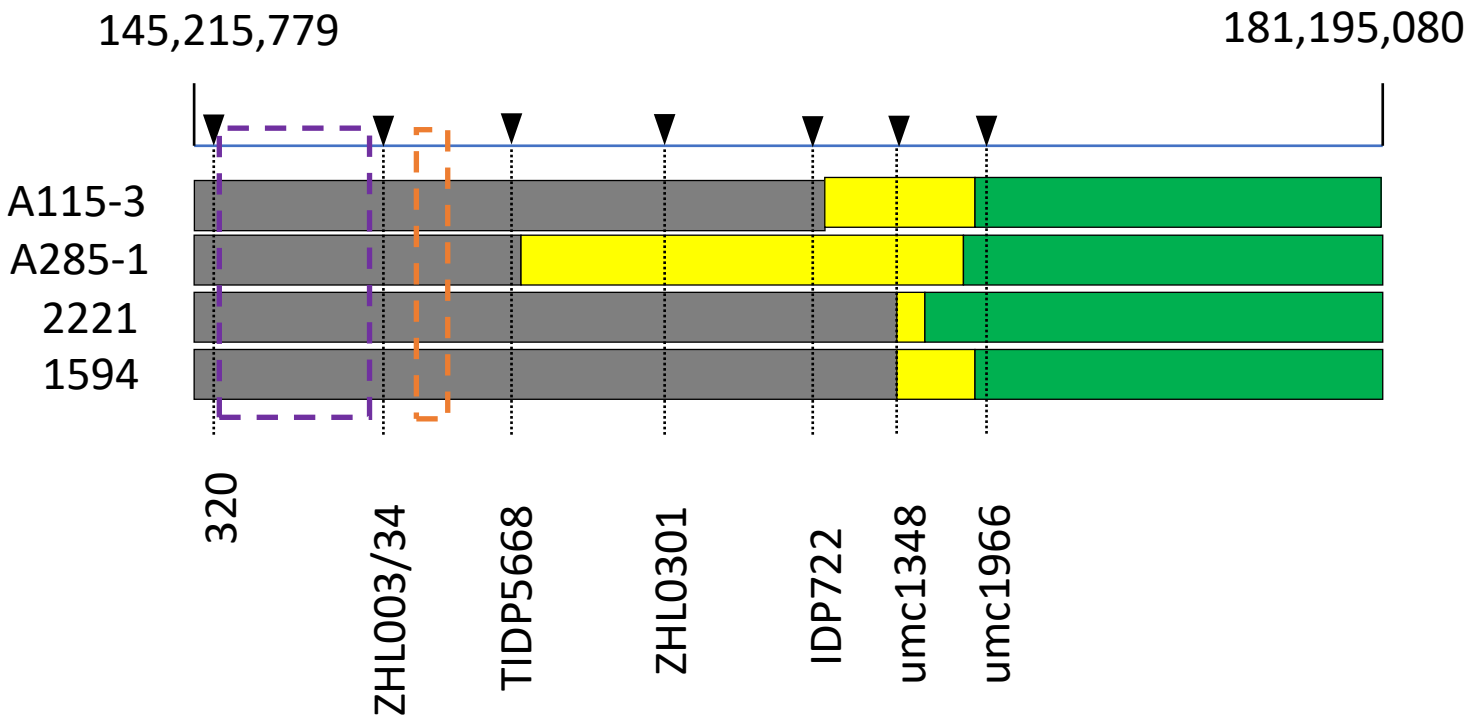


Figure 1.12: Reassessing Region of Significance.
A subset of the original B46 lines were selected that segregated upstream of our original target region seen here. The genotypes at each marker in that region are displayed. Teosinte=green, maize=yellow and heterozygous=grey. The purple dashed box signifies where Shannon 2013 found the major ear diameter QTL, while the orange dashed box signifies where Shannon 2013 found the major kernel row number QTL. The approximate location of each of the new primers (Supplemental Table A.4) is defined in the lower portion of the figure.

genotype, heterozygous individuals for that genotype were selfed in case it failed to produce kernels for further study. All ears that were selfed were harvested.

We identified nineteen types of recombinants from the entire region. Many of these are the individuals that produced more than 50 kernels in the greenhouse. For seven individuals, only heterozygous recombinants were found, so they were selfed for the potential to be screened in further studies. By the end of the 2017 field season, a total of eleven recombinants could be used in further phenotyping analyses and six heterozygous ears could be screened for recombination events again. Two potential recombinants, one homozygous and heterozygous, were lost due to poor ear development.

1.5 Discussion

1.5.1 Complex genetic architecture of ear size traits

Numerous studies were successful in identifying causative factors underlying QTL in maize. Similar methods have been previously used to examine underlying causal genes in identified domestication QTL (Hung *et al.* 2012; Wills *et al.* 2013; Yang *et al.* 2016). Instead of finding an underlying causative factor in our region on chromosome five for ear size traits our results suggest a new region of interest to investigate for a controlling factor responsible for kernel row number.

Prior analyses had predicted a similar QTL location for ear size traits such as ear diameter and kernel row number. In Shannon 2013 it was predicted that the ear diameter QTL, *diam5.2*, was located at approximately 145.98-150.51 Mbp (AGPv2) and the kernel row number QTL, *krn5.2*, was located at approximately 151.88-152.83 Mbp (AGPv2) on chromosome five.

Lemmon and Doebley (2014) had predicted a shift to these QTL to co-localize around 167.4 to 169.5 Mbp (AGPv2). In their study, their QTLs *eard5.3* and *krn5.2* contributed 25.6% and 53.4% of the variation, respectively. However, the current mapping study appears to indicate that the region is found upstream of the original location (Shannon 2013). This is likely related to the presence of an ear size QTL on chromosome seven for the lines used in this study, which was also observed in Shannon (2013). This is also be supported by the evidence that we find a significant QTL for chromosome five in B46, a population that segregates starting at 152 Mbp, while B11b, a population that starts segregating around 166 Mbp, does not present the same segregation pattern.

QTL in other segregating regions beyond the 6 Mbp-170 Mbp region on chromosome five may have influenced the predicted location of the QTL found by Lemmon and Doebley (2014). The original development of the lines in their study used restriction fragment length polymorphism (RFLP) markers to eliminate other teosinte segments on other chromosomes and not GBS data, which may have led to bias in the QTL results. There were no RFLP markers on chromosome seven, so therefore segregation was not examined on chromosome seven. However, there were previously identified domestication QTL on chromosome seven, especially for ear size (Doebley and Stec 1991, 1993; Shannon 2013). This explains how we found the unexpected segregation on chromosome seven.

Other potential reasons that we are unable to identify a causal gene could be due to a phenotype's genetic architecture and limitations of the statistical model. Shannon (2013) highlighted that she found 35 QTL for ear diameter. While ear diameter did have the largest

heritability of all of her traits, this was still 11 QTL more than the trait with the second largest number of QTL. The concern with that many linked QTL was that the accuracy may not be there due to limitations of the statistical model. According to her simulation work examining the number of QTL and population size, this is near the upper limit of expected QTL for a highly heritable trait for a population of its size. This suggested to her that there are probably more than 35 causal loci of small effect that control ear diameter, where many of those 35 "causal loci" may actually be multiple-linked loci. Shannon ran additional simulated data sets with a differing number of underlying causative loci and found that no matter the distribution or the number of loci, the number of QTL identified by software was consistently around 30-35 even if up to 300 underlying loci contributed to the phenotype tested. This suggests that the statistical model cannot accurately predict causal loci in traits with many loci of small effect.

In addition to her simulation data there was also biological evidence pointing to a similar conclusion. Shannon (2013) found three ear diameter QTL across all of chromosome five, while Lemmon and Doebley (2014) found three QTL for ear diameter just from 7 to 169 Mbp. These results may explain why we are not finding significance for ear diameter on chromosome five in this study's QTL mapping experiments. When examining a region that supposedly had a factor controlling ear diameter, there were no significant factors found on chromosome five and the only significant factor appeared on chromosome seven. Both other phenotypes examined in this experiment, kernels per rank and kernel row number, had a significant QTL for their respective trait on chromosome five. It seems the more crossovers that were introduced, the more factors with reducing significance appear for ear diameter. This leads us to conclude that ear diameter may be a trait that is controlled by many loci of small

effect.

1.5.2 Pollen Incompatibility

The pollen incompatibility factor *ga2* mentioned in the introduction could also have influenced our results when examining the region upstream of our original target interval. Pollen grains carrying a *Ga* allele have a significant advantage in effecting fertilization over pollen carrying the *ga* allele on the silks of plants that are *Ga/Ga* or *Ga/ga* but not of plants that are *ga*/*ga* (Kermicle and Evans 2010). This is one of many pollen incompatibility factors around the genome and studies have shown that the effect at one factor can override the effect at other factors (Burnham and Clark 1954; Kermicle and Allen 1990; Kermicle and Evans 2010). It is suggested that these incompatibility factors exist to reinforce reproductive isolation due to maize-teosinte hybrids having low fitness. While the exact location of *ga2* has yet to be identified, the maize genome maps it between markers IDP8236 (AGPv2 161,083,179-161,084,055) and IDP722 (AGPv2 163,897,310 - 163,897,997). This falls in the region we later examined and could explain why we had difficulty recovering seed from recombinants when studying the upstream region.

1.6 Future Work

This study reaffirmed the importance of cleanly segregating lines seen in Lemmon and Doebley (2014). Even though we had isogenic lines, we still were not able to locate the causal factor for any of the three ear size traits examined but instead provided more evidence to how crucial it is to only use lines that segregate in the region of interest. Therefore, for future work on this project, GBS data from all eleven BC_6S_6 parent lines should be collected to find which pair(s)

would result in the cleanest segregation. By identifying any outlying segregating regions individuals can be selected for the most isogenic background to create a maximum level of success. The three parents, B04, B11b and B46, were originally chosen because a teosinte allele in a theoretically all-maize background would make the phenotype stand out more. However, since little difference was seen in germination amongst the eleven parents in the original stages of the project all would be good candidates to sequence. The results provided by GBS would dictate the future path of the experiment. If all the parents have regions that segregate elsewhere in the genome another backcross to the W22 inbred maize parent would be needed and then screened for segregation of the specified problematic region in the offspring using PCR markers. If parents are found that would result in a lack of segregation elsewhere in the genome the same crossing design that was used to generate the RC-NILs could be repeated with our new parents and screened for recombinants in the expanded 20Mbp region. This would involve taking the F1 seed already generated in 2014, screening F2 offspring for segregation in the 20 Mbp upstream of umc1348 and umc1966 using fluorescently tagged molecular markers, and selfing individual plants heterozygous for a recombination event. The F3 would then be screened for a homozygous recombination event to generate RC-NILs to phenotype. This would at least alleviate the problem of segregation affecting the phenotype outside of our region on chromosome five because little could be done about any problems resulting from *ga2* complications. This design would allow us to better answer what factors contribute to the kernel row number phenotype and the structure of the genetic architecture of the ear diameter trait.

1.7 Conclusion

Through multiple studies we are consistently finding significance for this region on chromosome five for ear size (Doebley and Stec 1991, 1993; Shannon 2013; Lemmon and Doebley 2014), however the region of interest shifts slightly between populations due to the difference in genetic backgrounds. This study identifies a larger region where we believe a kernel row number QTL lies that has approximately 180 annotated genes in the 10 Mbp upstream of our original region of interest between molecular markers umc1348 and umc1966. No particular candidate gene of interest lies within that region which makes this a significant challenge to pursue in the future with such few recombinants.

1.8 Acknowledgements

I'd like to thank Zachary Lemmon for helping get this project off the ground and supplying the original parent lines. I would also thank Eric Rentmeester, Adam Mittermaier and Wei Xue for assistance with planting and maintenance of my plants. I would like to thank Craig DeValk for all of his hard work and enthusiasm to assist in the kernels per rank trait associated with this chapter. It would also not be possible without the help from all the numerous undergraduates that helped assist on this project including Lora Daskalska, Brandon Kim, Samuel Lawton, Jack Schnell and Bailey Spiegelberg.

Chapter 2:

Fine-mapping a major domestication QTL for ear shattering trait on chromosome five in *Zea Mays*

2.1 Abstract

A major change that consistently occurred during crop domestication was the disappearance of seed shattering in the domesticated counterpart. In the case of maize domestication, the wild ancestor's (teosinte) ears shatter at maturity for increased dispersal; therefore, intact ears would have been essential to increase yield during harvest. In this study we conducted a fine-mapping experiment on a previously identified QTL for nonshattering in maize that demonstrates the trait is controlled by *yab6*, which encodes a YABBY transcription factor. We generated 14 recombinant chromosome nearly isogenic lines (RC-NILs) and conducted Sanger sequencing to narrow down the causal region, which appears to map upstream of the gene's 5' untranslated region (UTR). We also demonstrated that *yab6* does not act alone, but interacts with *ZmYAB2.1*, another YABBY transcription factor that has been previously identified as a domestication gene responsible for the length of ear internodes and the density of kernels in an ear. Similar to *ZmYAB2.1*, weak to no evidence for selection is found near the causal region in exon 1. Overall our results have confirmed the proposed role of *yab6* as the main factor, working in concert with a known domestication gene *ZmYAB2.1*, to regulate nonshattering in maize.

2.2 Introduction

During the process of domestication of various cereal crops a major trait selected against was seed shattering. In wild crops seeds shed naturally at maturity in order to continue propagating the species. However, this would have made harvesting crops rather inefficient and would result in a large loss of yield. Ancient farmers would have had to strike a balance between harvesting mature seed versus increasing their yield. A mutation in a plant causing a nonshattering phenotype would have been advantageous for ancient farmers and propagated for generations to come. Selecting against the shattering trait would have resulted in these crops becoming increasingly reliant on humans, which is a hallmark of distinguishing wild crops from their domesticated counterparts.

While there have not been any causative genes identified in the evolution of nonshattering in maize (*Zea mays* ssp. *mays*) from its wild progenitor (*Zea mays* ssp. *parviglumis*), several have been identified in other cereal crops such as wheat (Simons *et al.* 2006), rice (Li *et al.* 2006; Konishi *et al.* 2006), and sorghum (Lin *et al.* 2012).The first gene identified was *Q*, a major gene involved in the domestication of wheat affecting shattering, glume shape, glume tenacity, spike length, plant height and time to spike emergence (Simons *et al.* 2006). *Q* is a member of the AP2 class of transcription factors that is best known to regulate the development of inflorescence structure and flowering. The ortholog has been demonstrated to determine the number of floral meristems produced (Chuck *et al.* 2008) and is a possible candidate for an increase in kernel row number (Calderón *et al.* 2016). In rice multiple genes responsible for shattering have been identified including *sh4* and *qSH1* (Li *et al.* 2006; Konishi *et al.* 2006). *sh4* is a Myb3 transcription factor that was found when one amino

acid change was made in the binding domain that could change a plant's phenotype from shattering to nonshattering and affect the formation of an abscission layer. *qSH1* is a BEL1-type homeobox gene and a single nucleotide mutation in the 5' regulatory region would lead to the absence of formation of an abscission layer. While both genes result in the lack of formation of an abscission layer they were identified through different crosses, with *sh4* found by crossing *O. sativa* ssp. *indica* with either of two wild species *O. rufipogon* or *O. nivara* and *qSH1* by crossing *O. sativa* ssp. *indica* cultivar Kasalath, a nonshattering to *O. sativa* ssp. *japonica* cultivar Nipponbare. None of the orthologs to these genes have been found to be under selection in other crop species.

However, this does not appear to be true for *Shattering1* (*Sh1)*. *Sh1* is a gene in sorghum that encodes for a YABBY transcription factor, with ten variants in three haplotypes that contribute to the nonshattering phenotype (Lin *et al.* 2012). In addition to their work in sorghum, Lin *et al.* (2012) also reaffirmed that there are orthologs to *Sh1* in maize, rice and foxtail millet, which had been hypothesized to underlie multiple QTL found in previous work (Paterson *et al.* 1995; Devos and Gale 2000). In maize specifically there are two orthologs of *Sh1* that are believed to be responsible for nonshattering, one on chromosome 1L and the other on chromosome 5S. Shannon (2013) identified QTL for this phenotype at both these locations in a domestication QTL mapping experiment and they were the two largest of the four QTL found for the trait. The QTL on chromosome five, DIS5.1, was the largest of the study with a LOD score of 59.1, explaining 23.1% of the phenotypic variance. It also had only 11 genes in the interval—with the ortholog of *Sh1* being the only transcription factor. The QTL on chromosome 1, DIS1.1, was the second largest with a significantly reduced LOD score of 10.32 and explained 3.5% of

the variance. This QTL had 59 annotated genes, but the *Sh1* ortholog is one of two transcription factors. This YABBY transcription factor, *ZmYAB2.1*, was later described to work epistatically with the teosinte branched (*tb1*) gene to control the length of internodes within the maize ear and a plausible candidate to also control nonshattering (Yang *et al.* 2016).

This study builds off the work found in Lin *et al.* (2012) and Shannon (2013) to confirm that indeed the maize ortholog of the sorghum shattering gene *Sh1* on chromosome five, *yab6*, is main factor for controlling the shattering domestication trait in maize. The causal factor(s) appears to align to a region just upstream of the 5' UTR of *yab6*. The teosinte allele in either *yab6* or *ZmYAB2.1* has a dominant effect where the phenotype of the heterozygote is similar to an individual homozygous for teosinte. Evidence is found for an interaction effect at the phenotype level between both genes. This report provides data that is consistent with previous work in the other *Sh1* homologs including the other identified homolog in maize, *ZmYAB2.1* (Yang *et al.* 2016).

2.3 Materials and Methods

2.3.1 Fine-mapping

A set of recombinant chromosome nearly isogenic lines (RC-NILs) were developed for fine-mapping the DIS5.1 QTL starting with a single family that was backcrossed twice to the maize parent and selfed three times (BC_2S_3). This family (MR0079B) was chosen because it segregated for maize and teosinte genotype where the DIS5.1 QTL is located and is homozygous for maize genotype at another major shattering QTL (DIS1.1) on chromosome one (Shannon 2013). Heterozygous MR0079B individuals were selfed to create a large segregating family that was

screened for crossovers between the molecular markers P1 and P9 in the genes that flank *yab6* (GRMZM2G157310 and GRMZM5G883855) (Supplemental Table B.1). DNA was extracted from seedling leaves in 96-well plates using a modified CTAB protocol (Doebley and Stec 1991). The DNA preparations were used as the substrate for PCR with fluorescently tagged primers for both P1 and P9 in a single reaction. PCR products were assayed on an ABI 3730 DNA Analyzer (Applied Biosystems) and the resulting chromatograms processed using Genescan software (Applied BioSystems). Fourteen individuals with crossovers in the interval between P1 and P9 in one of their two chromosomes, were identified and self-pollinated. Progeny from these fourteen plants were then screened to identify those plants that were homozygous for the recombinant event between the two molecular markers in the same manner as the previous field season. Individuals that were homozygous for the breakpoint were selfed to generate the set of RC-NILs.

Each of the 14 RC-NILs along with two control lines (homozygous non-recombinant maize and teosinte in the *yab6* interval) were grown in a randomized block design of eight blocks in both 2014 and 2015 at the West Madison Agricultural Experiment Station. For 2014 ten kernels were sown in each plot at a spacing of 1 foot (30.48 cm) apart within the plots. In 2015 sixteen kernels were sown in each plot with the same spacing as the previous field season. The goal was to phenotype five plants per plot in 2014 and fourteen plants per plot in 2015. Due to incomplete germination for some plots between 69 and 120 ears per each RC-NIL were obtained.

2.3.2 *Phenotypic analysis*

The shattering phenotype is measured as a weakened or more fragile cob in maize. Therefore, to measure this phenotype an ear is dropped from six feet above a hard floor and the number of segments into which the ear shattered was recorded as the shattering score. If the ear did not shatter the shattering score would have been recorded as one. For the fine-mapping analysis three other phenotypes were measured to examine possible pleiotropic effects of DIS5.1: ear length, ear diameter and cupule length. Ear length was measured with a ruler in centimeters, while ear diameter was measured with calipers in millimeters by measuring the middle of the ear. Cupule length was measured with a ruler by the length of 4 to 10 cupules in millimeters and then dividing by the number of cupules. This trait is also known as 10-kernel length (10KL) and ear internode length in other publications (Studer and Doebley 2011; Yang *et al*. 2016). The three additional ear phenotypes were collected only on the ears harvested in the 2014 field season.

The phenotypic data collected was fit to a linear model using MIXED procedure in SAS (Littell *et al.* 1998) to calculate least square means (LSMs) for each RC-NIL (Table 2.1). To fine-map the interval within which the shattering QTL is located correspondence was determined between the introgressed teosinte chromosomal segments and LSMs of the 14 RC-NILs. The introgressed segments of the 14 RC-NILs were matched to their corresponding LSMs. Then the 14 RC-NILs were sorted by their LSMs to test if the lines partition into two distinct phenotypic classes that differ for maize versus teosinte haplotypes within a single genomic interval. Under the assumption that there is a single gene contributing to the QTL one expects a clean segregation of the lines into two groups based on their LSMs such that there is a small

Table 2.1: Mixed Linear Models used to produce least squared means for fine-mapping RCNILs.

Y_{ij} or Y_{ijk} is the phenotype for a line, μ is the population mean, a_i is the line effect, b_j is the block effect, c_k is the year effect, and e_{ij} or e_{ijk} is the experimental error. Line and year were included as fixed effects while the block effect was random.

Trait	Mixed Linear Model
Cupule Length	$Y_{ij} = \mu + a_i + b_j + e_{ij}$
Ear diameter	$Y_{ij} = \mu + a_i + b_j + e_{ij}$
Ear Length	$Y_{ij} = \mu + a_i + b_j + e_{ij}$
Shattering	$Y_{ijk} = \mu + a_i + b_j + c_k + e_{ijk}$

chromosomal interval for which all lines in one group have the teosinte haplotype and all lines in the other group have the maize haplotype. This interval is the causative interval that contains the QTL. Genotype and phenotype data are available in Supplemental Tables B.2, B.3 & B.4.

2.3.3 Analysis of epistasis and gene action

In addition to examining the main effect of DIS5.1, the individual main effect of DIS1.1 and the interaction between the two QTL were estimated. To assay this an F2 population was created to segregate for both QTL. Lines were generated by taking a stock that was homozygous for the teosinte allele at DIS1.1 but homozygous for the maize allele at DIS5.1 (W22-ETB1_IN1207 F4, Stock ID: 49842) and crossing it to a stock that was homozygous for the maize allele at DIS1.1 but homozygous for the teosinte allele at DIS5.1 (W22-QTLShatter5c_TCL, Stock ID: 56730). The F1s were then selfed to generate the F2 population. The F2 population was grown in the Summer of 2015 in a fully randomized design and all plants were genotyped at DIS1.1 and DIS5.1 with markers P080-223 and P028-031, respectively (Supplemental Table B.5). All individuals were only scored for the shattering phenotype.

For the F2 population, the main and interaction effects were tested using the following linear model implemented in SAS to calculate LSMs:

$$Y_{ij} = \mu + a_i + b_j + (ab)_{ij} + e_{ij}$$

where Y_{ij} is the trait value for the shattering phenotype, μ is the population mean, a_i is the DIS1.1 effect, b_j is the DIS5.1 effect, $(ab)_{ij}$ is the interaction effect between DIS1.1 and DIS5.1, and e_{ij} is the experimental error.

Among the F2 population, plants were identified that were homozygous at both QTL and self-pollinated to produce a set of true-breeding recombinant inbred lines (RILs). This process

resulted in DIS1.1::DIS5.1 homozygous RILs with the following genotypes: M:M, T:M, M:T, and T:T. The RILs were grown in a randomized complete block design in Summer 2016 with 8 blocks and 16 plants per plot spaced 1 foot (30.48 cm) apart. In addition to the RILs three plots of W22 (control) were sown in each block. The shattering phenotype was measured on all ears in addition to cupule length, ear diameter, ear length and kernel row number to see if there is any evidence for pleiotropic effects. Cupule length, ear diameter and ear length were measured in the same manner as in the fine-mapping experiment above. Kernel row number was measured by counting the number of kernel rows around the circumference of the ear.

For the RILs the main and interaction effects were tested using MIXED procedure in SAS with the following linear model:

$$Y_{ijkln} = \mu + \beta(a_i) + \beta(b_i) + \delta(ab)_i + c_j(d_k) + c_j(f_l) + e_{ijkl}$$

where Y is the trait value for the specified phenotype (cupule length, ear diameter, ear length, kernel row number and shattering), μ is the overall mean, $\beta(a_i)$ is the genotype for DIS5.1 for line i, $\beta(b_i)$ is the genotype for DIS1.1 for line i, $\delta(ab)_l$ is the interaction effect between the genotypes at DIS1.1 and DIS5.1, c_j is the block effect, d_k and f_l are the respective X and Y coordinates within the block, and e is the experimental error. The DIS1.1 and DIS5.1 genotype effect, DIS1.1::DIS5.1 interaction effect, and position in the block were included as fixed effects while the block effect was fit as a random effect. Genotype and phenotype data are available in Supplemental Tables B.6, B.7 & B.8.

2.3.4 Population genetics

Evidence of selection during domestication was investigated around DIS5.1 through the Hudson-Kreitman-Aguade (HKA) test (Hudson *et al.* 1987), Tajima's D (Tajima 1989), and

coalescent simulation (Innan and Kim 2004) to test if our region of interest deviates from the theory of neutral evolution. The ratio for nucleotide diversity (π) and nucleotide polymorphism (θ) were both additionally calculated in maize and teosinte. In order to do this, seven regions in and around DIS5.1 were sequenced including in the two flanking genes GRMZM2G157310 and GRMZM5G883855, 2kb upstream of *yab6*, three regions within *yab6* and 6kb downstream of *yab6* (Supplemental Table B.9). All seven regions were sequenced in a variety accessions of maize and teosinte, as well as an outgroup, *Zea Diploperennis* (Supplemental Tables B.10 and B.11).

Sequences were also included from six previously identified neutral genes (AY104395, AY106816, AY107192, AY107248, AY111546, AY111711) (Zhao *et al.* 2011) as controls for the HKA test. All sequence alignments were done using Sequencher 5.1 (Gene Codes Corporation). Nucleotide diversity, nucleotide polymorphism, HKA test (Hudson *et al.* 1987), and Tajima's D using DnaSP v5.10.01 (Librado and Rozas 2009). Percent loss in diversity was calculated as $(1-\pi_M/\pi_T)*100$. Coalescent simulation was performed using software from Innan and Kim (2004) with the same parameters used by Zhao *et al.* (2011).

2.4 Results

2.4.1 Phenotypes

For the fine-mapping population there was quite a range in the shattering trait. The highest shattering score identified was 5, although few individuals have that high of a score. The average shattering score for the population was 1.593 with a variance of 0.705. Repeatability instead of heritability was calculated due to the inability to partition the genetic variance from

the genetic-by-environmental variance (Fehr *et al.* 1987). The repeatability of the lines for the first year was 0.8133 and increased to 0.8747 in the second year.

The possibility of pleiotropy for the QTL was also examined in the first year of fine-mapping data by measuring phenotypes for other ear traits. In the RC-NILs cupule length ranged from 3.29 mm to 6.43 mm and averaged around 4.7 mm with a variance of only 0.13 mm^2. This seemed to echo the variation seen by Yang *et al.* (2016) in *ZmYAB2.1*. The repeatability was 0.5859. For ear diameter the RC-NILs had an average diameter of 31.24 mm and a variance of 4.51 mm^2, with a wide range from 22.98 to 35.61 mm and a repeatability of 0.5194. The ear length phenotype also had significant variation, ranging from 8.90 cm to 20.50 cm with an average of 16.29 cm and a variance of 1.95 cm^2, with a repeatability of 0.4800.

For the F2 interaction data the average shattering score was 1.891 and a variance of 0.725. While this was higher than the average for the fine-mapping data this population still had the same maximum shattering score of 5 and a similar variance. These individuals did not have replicates to examine the heritability or repeatability of the trait.

In the RIL population we saw an average shattering score closer to the fine-mapping population of 1.56 and a decreased variance of 0.52. We still found that the maximum shattering score was 5 in this population as well. Repeatability instead of heritability was calculated for each trait not only due to the inability to partition the genetic variance from the genetic-by-environmental variance, but also due to the selection on the population (Fehr *et al.* 1987). Repeatability was significantly lower with a value of 0.6333. Pleiotropy was also measured in the RIL population. Cupule length had an average of 5.34 mm, which is larger than the fine-mapping score, and a variance of 0.57 mm^2. These results were not surprising since

now a known gene that controls the phenotype was segregating in the population in addition to DIS5.1. Repeatability of cupule length was slightly higher than shattering at 0.7349. The average for the ear diameter phenotype was 28.25 mm and a large variance of 16.21 mm^2, with a range from 12.95 mm to 55.55 mm. Repeatability was around 0.6963. Ear length had a range of 5.8 cm to 19.2 cm and an average of 12.36 cm and a variance of 4.86 cm^2. Its repeatability was 0.6450. Kernel row number was the most evenly distributed trait with a range from 8-18 and an average of 12.73 and a variance of 1.89 cm^2. The repeatability of this trait was 0.5798.

2.4.2 Fine-mapping of DIS5.1

To map the genetic factor(s) controlling shattering the RC-NILs were sorted by their LSMs for shattering and values were compared to the haplotypes of the lines in the QTL interval (Figure 2.1). The upper portion of Figure 2.1 shows that the lines segregate into two groups – a maize-like group above the green line and a teosinte-like group below. The two groups are largely differentiated by the chromosomal interval between markers P6 and P8 with the exception of two RC-NILs (R02 and R09), which have cross-overs within this interval. This segment falls upstream of the duplicated *yab6* exons in a presumptive regulatory region. While there is significant variation for shattering within each of these two clusters, this grouping shows the greatest differentiation between groups relative to differences within the maize and teosinte classes. These data suggest that DIS5.1 is a regulatory element for *yab6*.

As noted above, RC-NILs R02 and R09 have cross-overs within the P6-P8 causative interval and have intermediate LSMs for shattering. They show greater shattering scores than other lines in the maize class but far smaller shattering scores than lines in the teosinte class. To precisely map the cross-overs for these two lines DNA sequences of the two RC-NILs R02 and

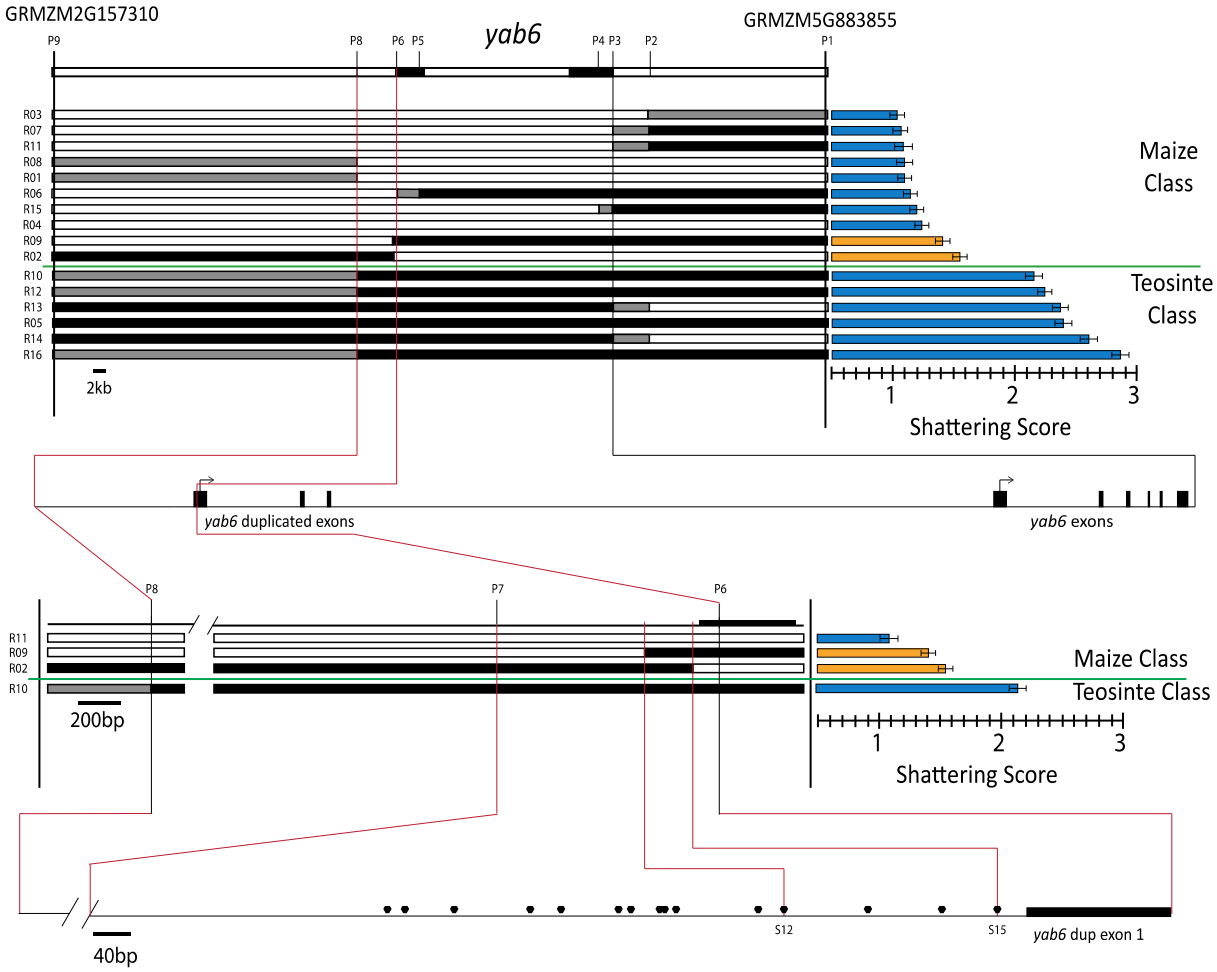


Figure 2.1: Fine-mapping of DIS5.1 on chromosome 5s.
Genotypes of the 14 RC-NILs and two controls carrying different introgressions within *yab6* region are shown to the left. White bars represent maize genotype, black bars represent teosinte genotype and grey bars represent unknown genotype. Least squared means and standard errors of shattering score are shown for each RC-NIL on the right. The two bars in green represent phenotypically intermediate RC-NILs. Markers P6 and P8 flank the narrowed interval, which is expanded at the bottom of the figure. Triangles show polymorphisms that differ between the maize and teosinte haplotypes. The two phenotypically intermediate RC-NILS (2223 and 2632) have cross-overs between markers S12 and S15. The green line is the threshold line for the “maize” shattering class and "teosinte” shattering class.

R09 as well as the maize and teosinte control lines were determined for the interval from markers P6 to P7 (Figure 2.1). These two lines complicate the simple interpretation that a single causative polymorphism between P6-P8 controls shattering as the two lines do not share maize genotype in common with all the other maize-class RC-NILs throughout this region. A causative factor producing a more teosinte-like phenotype than the other maize-class lines could even fall between P5 and P6. Given these observations we hypothesize that multiple causal polymorphisms between P5-P8 are involved in controlling the trait including at least one factor in the P8-S12 interval and one in the S15-P5 interval. The intermediate phenotype of these two lines can then be explained since they each have the teosinte haplotype in one of these intervals and maize in the other. Interaction effects between the polymorphisms could also be involved.

Looking closely at the sequence alignment for RC-NILs R02 and R09, maize and teosinte controls, a total of 15 polymorphisms were identified in the consensus region between P6 and P7 (Supplemental Figure B.1). These include 9 SNPs and 6 indels that range in size from 1 to 4 bp in length. These are all potential causal polymorphisms.

2.4.3 Main and interaction effects of DIS5.1 and DIS1.1

To estimate the relative main effects of DIS1.1 and DIS5.1 QTLs an F2 population was assayed in which both QTL segregated for maize and teosinte alleles. A total of 594 F2 plants were scored for shattering. Both additive (a) and dominance effects (d) were estimated for the shattering trait (Table 2.2). DIS1.1 and DIS5.1 were both statistically significant ($p < 0.05$) for the phenotype with p-values less than p=0.0001. However, DIS1.1 (a=0.0427) was not found to have as strong of an additive effect as DIS5.1 (a= 0.2563) but does have a much higher

Table 2.2: F2 Population Main Effect Results.
Below are the additive (a) and dominance (d) effects for each phenotype at both QTLs, in addition to the measure of dominance (d/a). P-values were calculated in SAS.

QTL	Background at other QTL	a	d	d/a	P-value
1	**M**	0.314102564	0.406056587	1.292751583	
1	**H**	-0.023431595	0.460260948	-19.64274947	
1	**T**	-0.078431373	0.321465428	-4.098684211	
1	**Overall**	**0.042727787**	**0.439562845**	**10.28751718**	**<0.0001**
5	**M**	0.451357466	0.250921736	0.451357466	
5	**H**	0.212794918	0.206154485	0.212794918	
5	**T**	0.058823529	-0.031612645	0.058823529	
5	**Overall**	**0.256270147**	**0.202439919**	**0.256270147**	**<0.0001**

Figure 2.2: Interaction Data.
As mentioned above, since F2 population showed an interaction but saw large amounts of variation due to the number of ears recovered for each phenotype (top), the interaction experiment was repeated using a RIL population with 8 blocks of 30 plots (bottom). Estimate is the average shattering score for each phenotype. The general pattern is the same in both seasons, where the teosinte allele has a less-than-additive effect. Interaction tests for each population are found in Supplemental Table B.12 (2015) and B.13 (2016).

dominance to additive ratio (10.2875 vs 0.2563, respectively). This dominance effect is further seen when examining LSMs of each genotype (Figure 2.2). The average shattering score of an individual that is heterozygous at one of the two QTL has a similar average shattering score to an individual that is homozygous for teosinte at the same QTL. This demonstrates we do not see an additive effect between the loci.

The experiment was repeated with the set of true-breeding recombinant inbred lines (RILs). This process resulted in 27 DIS1.1::DIS5.1 homozygous RILs with the genotypes: M:M (8 ines) T:M (5 lines), M:T (6 lines), and T:T (8 lines). Due to the lack of heterozygotes only additive effects were estimated (Table 2.3). Results were similar to the F2 population where DIS1.1 is less significant for the shattering trait than DIS5.1, but the difference in the significance of each QTL has widened (0.0144 vs 0.1712, respectively). However, in the RIL population DIS1.1 no longer appears statistically significant (p=0.0690) for the shattering trait. Overall, the observations confirmed previous results where DIS5.1 is found to be have a stronger effect on the phenotype than DIS1.1 for the shattering trait. When examining other phenotypes for pleiotropy, statistical significance ($p < 0.05$) is seen in at least one of the QTL examined, if not both. Cupule length, ear diameter, and kernel row number are found to all be significant for both chromosome one and five. Ear length was only significant for chromosome one, but chromosome five was just above the significance threshold at 0.0594.

In addition to the individual effects of DIS1.1 and DIS5.1 the interaction effects between these two QTL were also examined (Figure 2.3, Table 2.4). When heterozygotes were included when analyzing the F2 population the interaction between DIS1.1 and DIS5.1 accounts for 1.71% of the shattering phenotype (Supplemental Table B.12). If all heterozygotes are dropped

Table 2.3: RIL Population Main Effect Results.
Below are the additive effects for each phenotype at both QTLs. P-values were calculated in SAS.

	QTL	a	Background at other QTL	P-value
Shattering	**1**	0.13095238	**M**	
	1	-0.118693	**T**	
	1	-0.002040	**Overall**	**0.0690**
	5	0.29257466	**M**	
	5	0.04292929	**T**	
	5	0.18500435	**Overall**	**<0.0001**
	QTL	**a**	**Background at other QTL**	**P-value**
Cupule Length	**1**	0.24685717	**M**	
	1	0.21811	**T**	
	1	0.2499635	**Overall**	**<0.0001**
	5	0.20331344	**M**	
	5	0.17456628	**T**	
	5	0.21681965	**Overall**	**<0.0001**
	QTL	**a**	**Background at other QTL**	**P-value**
Ear Diameter	**1**	-0.8002578	**M**	
	1	-1.1029736	**T**	
	1	-1.0756499	**Overall**	**<0.0001**
	5	-0.6961419	**M**	
	5	-0.9988577	**T**	
	5	-0.9337204	**Overall**	**<0.0001**
	QTL	**a**	**Background at other QTL**	**P-value**
Ear Length	**1**	-0.235885	**M**	
	1	-0.8159696	**T**	
	1	-0.6022425	**Overall**	**<0.0001**
	5	0.15192065	**M**	
	5	-0.428164	**T**	
	5	-0.1583788	**Overall**	**0.9106**
	QTL	**a**	**Background at other QTL**	**P-value**
Kernel Row Number	**1**	-0.1013655	**M**	
	1	-0.3247242	**T**	
	1	-0.2593057	**Overall**	**<0.0001**
	5	-0.0969027	**M**	
	5	-0.3202614	**T**	
	5	-0.2175594	**Overall**	**<0.0001**

Figure 2.3: Teosinte as a dominant allele.
In 2015, individuals were selected within our F2 population that were heterozygous at either DIS1.1 or DIS5.1 to analyze the effect of main effects of the maize and teosinte allele at each locus. Estimate is the average shattering score for each phenotype. At both homologs, it was found that heterozygous individuals tend to show a similar shattering phenotype to teosinte individuals, and not an intermediate phenotype.

Table 2.4: F2 interaction data.
This data was generated in SAS by fitting our interaction model to our data to estimate the main and interaction effects of the QTL. The 2015 data was fit to the model with the heterozygotes (row one) and without heterozygotes (row 2). The 2016 RIL experiment had no heterozygous samples. Contrasts were only calculated for the data that examined heterozygotes.

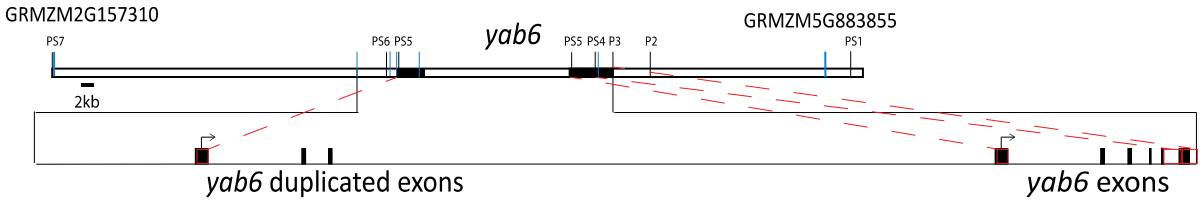
	Main Effects		Interaction	Contrasts			
Year	**DIS1.1**	**DIS5.1**	**DIS1.1 x DIS5.1**	**A x A**	**D x D**	**D x A**	**A x D**
2015	**4.6626**	**3.7858**	**1.7136**	**0.7963**	**1.5035**	**3.260***	**1.660**
2015	**2.0636**	**9.6710**	**5.7250**	**-**	**-**	**-**	**-**
2016	**0.1792**	**7.6312**	**2.3525**	**-**	**-**	**-**	**-**

from the F2 population analyses the interaction appears to explain more of the phenotype, at about 5.7%. When examining the RIL population, the interaction between DIS1.1 and DIS5.1 is statistically significant (p=<0.0001) and explains approximately 2.35% of the phenotype (Supplemental Table B.4, Supplemental Table B.13). For the other ear traits, ear length, and kernel row number were both found to have a statistically significant interaction between the two QTL. The interaction value for ear diameter was near our statistical significance threshold at 0.0716. Cupule length had no significant interaction between the two QTL.

2.4.4 Population genetics

After conducting different statistical tests to examine any departure from expectations, evidence of some form of selection near exon 1 of *yab6* is found, although it is not strong. A significant departure ($p < 0.05$) is only found in the coalescent simulation test but not either in Tajima's D nor the HKA test (Table 2.5). This is not surprising considering Tajima's D has been demonstrated to have low power for detecting selection in maize (Tenaillon *et al.* 2004). It is possible that the HKA test was not significant due to the different parameters of the test. While both the HKA test and coalescent simulation have similar goals of testing the neutral theory of evolution, they achieve it in slightly different ways. Demographic history of an organism influences the HKA test, and crops tend to violate many of the assumptions of the neutral equilibrium model used to examine whether selection occurred (Wright and Gaut 2005). On the other hand, coalescent simulations estimate the duration and severity of bottleneck by using summary statistics of the population, including recombination. Since maize is known to undergo a drastic bottleneck during the process of domestication this might influence the results and possibly explain why significance was seen with coalescent simulation and not with

Table 2.5: Results for different selection tests on different regions around DIS5.1. P-values are shown for pair-wise HKA test and coalescent simulation, where a significant p-value is marked with an asterisk. Tajima's D statistics are shown for Tajima's test, but no regions were found to be significant.



Description	Pair-wise HKA Test		Tajima's D		Coalescent Simulation	Diversity Loss (%)
	Maize	Teosinte	Maize	Teosinte		
GRMZM2G157310 (primers 55-56) (PS7)	0.997	0.998	-0.212	-1.319	0.2526	36.37
2kb upstream yab6 duplicated exons (primers 42-44) (PS6)	0.267	0.853	-0.294	-0.136	0.1508	61.38
Within exon 1 yab6 (primers 19-145) (PS5)	0.998	0.988	1.261	-1.652	**0.0493***	43.21
In between exon 5 and 6 yab6 (primers 73-97) (PS4)	N/A	N/A	-0.974	-0.562	0.4996	38.84
Around exon 6 yab6 (primers 28-30) (PS3)	0.607	1	-1.803	0.291	0.4325	67.55
AC233885.1_FG002 (primers 158-160) (PS2)	0.224	0.815	-0.571	-0.336	0.2364	66.63
GRMZM5G883855 (primers 148-163) (PS1)	0.458	1	-1.205	0.548	0.3718	62.14

the HKA test. Nonetheless, the significant coalescent simulation result was directly downstream of the causal region our fine-mapping experiment identified, within exon 1, and provides weak evidence of selection at best for the phenotype during the domestication process.

2.5 Discussion

2.5.1 YABBY gene function and shattering

The YABBY gene family is comprised of a class of transcription factors that has been previously described in *Arabidopsis*. This work has shown that YABBY gene family has been involved in the initiation of outgrowth of the lamina, maintenance of polarity, and establishment of the leaf margin (Finet *et al.* 2016). YABBY genes are characterized by their two domains: a N-terminal zinc finger domain and a C-terminal helix-loop-helix motif (Bowman and Smyth 1999). There are six YABBY genes found in *Arabidopsis*: CRABS CLAW(CRC), FILAMENTOUS FLOWER (FIL), YABBY2 (YAB2), YABBY3 (YAB3), INNER NO OUTER (INO or YABBY4), and YABBY5 (YAB5) (Bowman and Smyth 1999; Sawa *et al.* 1999; Siegfried *et al.* 1999; Bowman 2000). Four of the six (FIL, YAB2, YAB3 and YAB5) are expressed in the leaf and many leaf-derived organs. CRC and INO are expressed in tissue evolutionarily derived from leaves such as carpels and nectaries for CRC and the abaxial domain of the outer integument for INO (Villanueva *et al.* 1999).

YABBY homologs have been identified in rice, sorghum, and maize, where they appear to have a different role than in *Arabidopsis* (Jang *et al.* 2004; Toriba *et al.* 2007; Lin *et al.* 2012). In rice there are eight classified YABBY genes that fall into four classes (Toriba *et al.* 2007). The two major classes comprise of six of the eight identified YABBY genes and have differing patterns of expression and sequence characteristics. Both classes have different expression

patterns than in *Arabidopsis* where they are either seen expressed in parts of the developing plant containing meristems and primordia of organs such as leaf and floral organs or as a general expression pattern of almost all organs suggesting a broader function. This is confirmed with the homologs of *OsYAB2* found in maize (*ZmYAB2.1* and *yab6*) and sorghum (*Sh1*) where they appear to function in the maintenance of abscission layer and its formation (Lin *et al.* 2012; Yang *et al.* 2016). In *ZmYAB2.1 in situ* RNA hybridization shows that expression is localized to a narrow band where the abscission layer would form (Yang *et al.* 2016). Using this information, in addition to what is presented in this work, we have multiple pieces of evidence suggesting that these two YABBY transcription factors, *ZmYAB2.1* and *yab6,* appear to regulate the formation of an abscission layer. Significance was found in our chromosome five QTL for the ear internode (cupule length) phenotype as well as shattering providing evidence for *yab6* to have a similar role to *ZmYAB2.1*.

2.5.2 yab6 and role in shattering

Many previous studies have hypothesized that an important factor for the shattering trait is in this region on the short arm of chromosome five (Doebley and Stec 1991, 1993; Paterson *et al.* 1995; Devos and Gale 2000; Lin *et al.* 2012; Shannon 2013). The work presented here was able to confirm that polymorphisms in the regulatory region of *yab6* are a major cause of nonshattering in maize.

We also show that *ZmYAB2.1* plays a role in shattering as well and that the interaction between the alleles can affect the ability to shatter. We found there is a less-than-additive effect between the two alleles with the factor on chromosome five having a larger effect. This was expected based the size of the QTLs found in Shannon's mapping experiment (Shannon

2013). However, unexpectedly we consistently find that the phenotype with the largest shattering score are those that have the genotype M::T (Figure 2.2). Our hypothesis is that this is most likely due to our lines not selectively segregating for just our QTLs of interest. The line chosen to cross for teosinte at the chromosome one marker was IN1207 from the *ZmYAB2.1* experiment (Yang *et al.* 2016), and is almost completely teosinte from 257-275 Mbp, with a small exception between markers SBM04 and umc1411 (approximately 460 kbp) where it is maize. Something within that 18 Mbp beyond *ZmYAB2.1* must contribute to the phenotype, with the most likely candidate being *tb1*(Doebley *et al.* 1995).

We also find that the teosinte allele has a dominance effect at both DIS1.1 and DIS5.1 (Figure 2.3) where the phenotype of the heterozygous allele is hard to distinguish from the phenotype of the teosinte allele. This does not seem to be especially surprising given the results we see examining Lemmon *et al.* (2014)'s expression data. Maize samples show reduced levels of expression, especially in the W22 maize inbred sample, and teosinte on average has a much higher level of expression (Table 2.6). Therefore, one teosinte allele would result in an increase of expression compared to a homozygous maize individual and lead to the shattering phenotype observed.

In the B73 genome, the inbred line the maize genome is built upon, there is a partial duplication event of *yab6* within the syntenic block (Lin *et al.* 2012). This change is present in the majority of maize inbreds and absent in all teosinte inbreds (Supplemental Table B.14). In B73 the duplicated exons are found to be identical to their original counterparts. The introns between exons are found to be almost identical with only a 45 bp insertion and an A to G substitution, all approximately 120 bp upstream of exon 2. This insertion is in a repetitive

Table 2.6: Expression data from Lemmon *et al.* (2014).
We selected for reads that mapped to the original exons and calculated the reads per kilobase per million (RPKM) for each line. For this calculation, we established a threshold of up to two mismatches per read in order for it to count as mapping to our region. This provides further evidence of selection during domestication and is consistent with other genes responsible for shattering.

Strain	Reads Per Kilobase Per Million (RPKM)
B73	**2.948**
CML103	**2.222**
Ki3	**8.665**
Mo17	**2.016**
Oh23	**0.923**
W22	**0.274**
TIL01	**6.756**
TIL03	**5.399**
TIL05	**4.282**
TIL09	**13.535**
TIL10	**12.393**
TIL11	**8.141**
TIL14	**8.561**
TIL15	**16.209**
TIL25	**6.154**

region, specifically a string of TATA repeats. The duplicated region extends upstream of exon 1 by approximately 250bp, which made it hard to distinguish causal polymorphisms of the shattering phenotype. However, in the newly released W22 maize inbred genome sequence (Springer *et al.* 2018) this duplication does not appear to exist. Therefore, more work would need to be pursued to better understand the structure of the gene.

2.5.3 Selection at yab6

Strong evidence is not seen for selection occurring in the region of *yab6*. We sampled seven regions around *yab6* including the neighboring genes, upstream and downstream of *yab6,* and in and around exons 1, 5 and 6 (Table 2.5, Supplemental Table B.9). We observe evidence for selection only in a single coalescent simulation test within exon 1. While similar patterns of selection are observed for our shattering homolog on chromosome one (Yang *et al.* 2016), there are much stronger patterns of selection for other maize domestication genes (Wang *et al.* 2005; Studer *et al.* 2011). Like the suggestions that Yang *et al.* (2016) provides, it is possible that (1) we are seeing a false positive result for the coalescent simulation test, (2) since our only close sample to the causative region is solely within an exon it isn't detecting the sweep or (3) that *yab6* is a selection of a soft sweep at or near its 5' region (Innan and Kim 2004; Hermisson and Pennings 2005). Further evidence for weak or no selection is that Hufford *et al.* (2012) as well as subsequent follow-up analyses with XP-CLR by the Hufford lab did not see evidence for selection by genomic selection scans. These results suggest that no selection occurred or it was too weak to leave a signature of selection. This was also observed by Yang *et al.* (2016). However, it is possible that this is because there is a 27-kb gap in the genome between the duplicated and original exons, and therefore could be hiding evidence of selection (Jiao *et al.*

2017). In summary, we find that the selection at genes controlling the shattering trait appear to be weak at best, or have no selection.

While we find weak evidence to no evidence of selection through population genetic analyses, there appears to be evidence of selection observed when examining expression data. We found that various teosinte inbred lines consistently have more reads than maize inbreds that map to the region where *yab6* is located, with one exception (Table 2.6). Using Lemmon *et al.* (2014)'s data we gathered data for reads that mapped within *yab6*'s original exons. When quantified the reads per kilobase per million (RPKM) are on average much higher in teosinte inbreds than maize inbreds. Lemmon *et al.* also reported allele-specific expression in *yab6* for four maize lines and two teosinte lines, where teosinte alleles were consistently expressed at higher levels than the maize allele (Supplemental Table B.15). This provides further evidence for selection on an upstream cis-regulatory element, which are thought to be targets during the evolutionary process. Similarly, Yang *et al.* (2016) reported that *ZmYAB2.1* had increased levels of expression when the teosinte allele was present and found that Lemmon *et al.* also reported evidence of allele specific expression where teosinte alleles are expressed at higher levels than maize alleles. In summary, the expression data provides evidence that maize was selected for reduced levels of expression in the genes responsible for shattering.

2.6 Future Directions

While we have identified that *yab6* is a major contributor in the nonshattering phenotype in maize there is still some characterization of the gene that needs to be pursued. We must conduct more extensive work to verify the gene structure and the possible duplication using

molecular techniques. The maize reference genome uses an inbred called B73 in which our causal region appeared to have undergone a duplication event. However, this study used the maize inbred W22 and when the genome was recently published it does not appear to have the duplication of our causal region. Therefore, we need to sequence this region to better to understand if there truly is a duplication event in B73 and, if there is, verify if the duplication event also exists in W22. To do this we will develop primers that sequence all the way across the whole gene as well as a control gene and conduct Sanger sequencing on both B73 and W22 samples. Then, using Sequencher 5.1 (Gene Codes Corporation) we can align all the reads to examine if the number of reads varies between *yab6* and our control gene in B73 and W22 samples. This will help verify if both reference genomes are correct and help guide further projects to pursue.

I also would want to explore the expression patterns of *yab6*. I would conduct qPCR to quantify the possible differences in expression between the maize and teosinte copy of *yab6*. Using RNA extracted from immature ears I would amplify cDNA with reverse transcriptase and regular oligo primers that span across at least two exons of *yab6* and repeat the same for a control gene (actin) to normalize the expression levels. I would also want to examine the interaction between *ZmYAB2.1* and *yab6* to see if the expression at one gene could have an effect on the other. While preliminary qPCR data had been performed previously the results were inconsistent. More biological replicates would be needed to verify the validity of the results.

It would be interesting to investigate the spatial expression domain of *yab6* and see if it differs from its paralog *ZmYAB2.1*. *ZmYAB2.1*'s expression is observed in a narrow band

subtending to the spikelet pair (Yang *et al.* 2016). Since the genes are about 87% identical and seem to have a similar purpose there is a high chance that their physical expression patterns are similar as well. To verify its expression pattern, I would collect maize and teosinte immature ears and fix them in a para-formaldehyde solution and create RNA probes that are specific for *yab6*. I then would work with our collaborator to conduct *in situ* hybridization to examine both the maize and teosinte alleles in maize and teosinte genetic backgrounds.

I could also investigate the difference in the mRNA abundance between maize and teosinte using an allele specific expression assay. This would be done by extracting RNA from immature ears of F1 hybrid plants that are heterozygous to create cDNA. Primers would be created to examine a size difference between maize and teosinte alleles. PCR is then run by using fluorescently-tagged primers we created on the cDNA and assaying the products, similar to our fine-mapping methods. We would then obtain the area under the peaks corresponding to maize and teosinte using a software called Gene Marker (Softgenetics, State College, PA). This will allow me to calculate the relative expression of the alleles by taking the ratio of the area under the peak for maize versus teosinte alleles, and compare to Lemmon *et al.* (2014) results seen in Supplemental Table B.15. This experiment would only yield meaningful results if it can be verified that there is no duplication of the causal region.

2.7 Conclusion

In this study we have taken the approach of phenotype-to-gene mapping to better understand whether a maize homolog for a known shattering gene in sorghum, *yab6,* is also responsible for the same phenotype in maize. Our results indicate that *yab6* does indeed appear to play a

significant role in domestication. This was not too surprising given that the other shattering homolog found in maize, *ZmYAB2.1*, has also been associated with maize domestication for playing a role in ear internode length (cupule length) (Yang *et al.* 2016). The underlying genetic factor(s) of the shattering phenotype appear to lie in a regulatory region upstream of *yab6* where it is believed two or more variants lead to the shattering phenotype. Evidence shows that both *ZmYAB2.1* and *yab6* independently contribute to the phenotype but also appear to have an interaction effect. Pleiotropic effects appear to be seen as well in both genes. We also see that not all genes that contribute to the domestication of maize have strong patterns of selection. Both our genes contributing to the shattering trait show softer sweeps or little evidence for selection at all. Overall, this work provides more evidence that parallel selection during domestication amongst multiple cereals was indeed observed in the seed shattering phenotype.

2.8 Acknowledgements

I'd like to thank the many people that contributed to the work on this project: David Wills for creating the original crosses, Chin Jian Yang and Zhihong Lang for doing the original screening for the RC-NILs, Weidong Wang for much of the original fine-mapping work and population genetic analyses, and Michael Tuholski for the follow up fine-mapping and interaction work. I would also like to thank Wei Xue for help analyzing Lemmon's RNAseq data. Thanks to Jesse Rucker, Elizabeth Buschert, Eric Rentmeester and Adam Mittermaier for all assistance on the project.

Chapter 3:

Exploring the Impacts of Peer Mentoring on First-Year STEM Students

3.1 Abstract

This study uses a mixed-methods approach to examine the relationship between peer mentors and first-year STEM undergraduates. The results show that over the semester students increase their appreciation and perception of the value the peer mentors have to offer. It also suggests that many students do benefit from their relationship with the peer mentors and learn various skills from them. Students consistently report that their leaders are well-qualified to be mentors and help them get involved at the university. Although the literature suggests that underrepresented students benefit more from having peer mentors in the classroom we do not have data that supports that claim. However, we plan to repeat the experiment in the upcoming year to see if the results can be replicated before conclusions can be drawn. The goal of this work is to provide a more qualitative approach to the literature to understand what about the peer mentoring relationship first-year STEM students find beneficial in order to help initiatives on college campuses support underrepresented students in STEM fields.

3.2 Literature Review

The definition of mentoring or mentors varies across different disciplines which can lead to confusion over what is actually being measured (Jacobi 1991), and has made little progress since addressed by Jacobi (Crisp and Cruz 2009). This creates a problem because there is a lack of shared characteristics to quantify a mentor-mentee relationship. The shared consensus for mentoring relationships is focused on the following characteristics (Jacobi 1991):

1) A helping relationship focused on achievement and support

2) Mentors showing a greater level of expertise or achievement compared to the mentee or protégé
3) It is a reciprocal relationship between mentors and mentees
4) The relationship is personal due to direct interaction between the mentor and mentee or protégé, and
5) Mentors provide one, if not all three, of these potential components: emotional and psychological support, direct assistance with career and professional development, and role modeling.

In Table 3.1 we outline some of the various definitions of mentoring or mentoring roles. As mentoring is a prominent form of education in graduate studies, it is not unreasonable that mentoring could be used in undergraduate education as well. Colleges and universities have implemented the technique in some shape or form with much success (Campbell and Campbell 1997; Tinto 1998) and further implementation has become a national priority (Girves *et al.* 2005). Peer mentoring however is unique in its approach. Research has also shown that peers play an important role in transitioning to college (Brissette *et al.* 2002; Ishler and Schreiber 2002), learning and academic performance (Astin 1993; Terenzini *et al.* 1996), satisfaction (Astin 1993), persistence, and retention (Tinto 1993; Thomas 2000). With peers being such an influential piece in a student's life, peer mentoring would have advantages compared to other mentoring relationships. Most traditional mentoring focuses on a particular structure that can be time consuming for faculty, but peer mentoring allows the networking of a community while de-emphasizing seniority or hierarchy and allowing for flexibility (Chesler and Chesler 2002).

Table 3.1: Mentoring Definitions

Name	Definition
Peer leader	'students who have been selected and trained to offer educational services to their peers [that] are intentionally designed to assist in the adjustment, satisfaction and persistence of students toward attainment of their educational goals' (Ender and Kay 2001)(Shook and Keup 2012).
Peer leaders in PLTL	undergraduates who have successfully completed the course that work collaboratively with classroom support staff to facilitate small group problem-solving outside of regular class hours (Woodward *et al.* 1993; Wamser 2006; Quitadamo *et al.* 2009; Snyder *et al.* 2015, 2016). This can also be known as a supplemental instruction (SI) leader, where supplemental instruction targets high-risk students (Stout and McDaniel 2006).
STEM ambassador	facilitates peer-led study sessions and planning and coordinating social events, as well as informally mentoring students through these activities (Bowling *et al.* 2015).
Near-peer mentoring	dyadic platonic relationship between a more experienced student (mentor) and a less experienced student (mentee) at the same institution with frequent direct face-to-face contact. Mentors range in age from second year undergraduates to graduate students (Zaniewski and Reinholz 2016).
Peer mentoring	an institutionally supported program facilitates a relationship amongst peers where one student is providing support for at least one peer (Newton and Ender 2010).

Peers also can apply a positive influence on their peers in a way that is less intimidating than staff or faculty (Cuseo 1991).

Specifically in STEM peer mentoring has been shown to be successful in a variety of outcomes, whether it is an improvement of grade performance of students (Wamser 2006; Snyder *et al.* 2015, 2016; Mayer *et al.* 2017), an increase in the retention of students in STEM fields (Chesler and Chesler 2002; Wamser 2006; Bowling *et al.* 2015; Snyder *et al.* 2015; Zaniewski and Reinholz 2016) or social integration (Allen *et al.* 1999). It is also well documented that measures like these are to be more beneficial for underrepresented minority groups in STEM (Chesler and Chesler 2002; Bowling *et al.* 2015; Snyder *et al.* 2016; Zaniewski and Reinholz 2016). While many of these studies examine smaller groups, the effect seems to resonate across all the different applications of peer mentoring. However, why it leads to success is not quite so well understood. Literature has shown that it can be tied to a sense of belonging and more positive identities in STEM (Carlone and Johnson 2007; Zaniewski and Reinholz 2016). It has also been shown that the more students interact with others students and faculty, the more likely they will persist in college (Nora 1987; Astin 1999).

As mentioned before, peer mentoring implementation styles widely differ across various colleges/universities in the literature. Some peer mentors lead discussion sections (Wamser 2006; Quitadamo *et al.* 2009; Snyder *et al.* 2015, 2016) while some also have social components (Bowling *et al.* 2015; Zaniewski and Reinholz 2016). Others provide more of a mentoring role by having regular meetings to discuss how things are going (Zaniewski and Reinholz 2016). Students can also vary in their training as well. Many times the qualification to be a peer mentor is to have successfully completed a course or have shown to be successful in

their major (Mayer *et al.* 2017). Some have also been trained in group dynamics and learning theory (Quitadamo *et al.* 2009) or in leadership training and development (Bowling *et al.* 2015). Others had mentoring learning communities throughout the mentoring process to provide support (Zaniewski and Reinholz 2016).

Beyond the benefits mentioned above for the students it is relatively easy to implement. It can require as little as one faculty member's time to supervise the implementation of this method (Woodward *et al.* 1993). Peer mentoring can also be low cost, as you can recruit peer leaders to participate for credit (Zaniewski and Reinholz 2016) on a volunteer basis (Mayer *et al.* 2017) or pay minimal stipends (Quitadamo *et al.* 2009). As discussed above, training can be quite minimal for students but still show successful outcomes. All of these reasons allow even universities with limited resources to pursue some form of peer mentoring.

On the University of Wisconsin-Madison's campus peer mentors (here after referred to as peer leaders) have been implemented in many STEM courses but their impacts have not been well studied. As a part of this study we are interested in learning students' perceptions of peer leaders in a freshman seminar course, whether or not students feel that they have benefitted from having peer leaders, and if so, how they benefitted from that relationship. We want to know if there are specific activities or interactions first-year students most benefit from and ways to continue it in future classes. If we find students are not benefitting from this relationship we are interested in knowing what we can improve to better the first-year transition experience. This will help us evaluate the peer leaders' training program on campus called the Integrated Mentoring Program and Core Training (IMPaCT) and see if any improvements can be made for future classes of peer leaders.

We also hope to contribute to the literature because while many quantitative measures of student success such as grades or percent retention have been collected to examine the effects of peer mentorship, the literature is limited when describing how peer mentors are valuable to the students. We are curious about factors such as access, the relatability of a mentor, certain personality traits, or ways they help students feel more connected to the campus. In our study, we hope to go beyond typical measures of student success outlined in other literature by using our focus groups and surveys to examine more behavioral questions instead.

3.3 Program Descriptions

University of Wisconsin-Madison is a large state university with approximately 30,000 undergraduates. According to the 2017-2018 University of Wisconsin-Madison Data Digest in fall 2017 the university comprised of approximately 51.2% women and 48.8% men. 73.18% of students identified as Caucasian, 5.21% identified as Hispanic, 2.11% identified as African-American, 0.20% identified as American Indian, 5.99% identified as Asian, 0.10% identified as Native Hawaiian, 3.28% identified as two or more of the above categories, 9.13% identified as International students, and 0.80% identified as unknown. Approximately 10.61% of students are what the university classifies as a “targeted minority” student, which is defined as domestic students who are Hispanic/Latino, African American, American Indian, or Southeast Asian (Hmong, Vietnamese, Laotian, Cambodian). That number is close to what was seen in the 2017 freshman class, where 715 of the 6,610 students (10.8%) were targeted minority students. At the university Biology is consistently the largest major by degrees conferred on campus, accounting for 7.4% of all degrees awarded in 2016-2017.

The IMPaCT program was developed as part of a response to address the issue of student retention amongst underrepresented populations at the University of Wisconsin-Madison. It was developed as part of a grant from the Howard Hughes Medical Institute (HHMI) specifically aiming to help first year, first generation and underrepresented minority undergraduates succeed. This was a part of larger overarching project to get students more involved in the university's community. The program itself was developed using core principles from educational leadership theory to prepare students for mentoring roles.

Since much of our project relied on the peer leaders in the course it was important that they were exposed to proper classroom training. To become a peer leader interested students applied to the IMPaCT program where they indicated what program(s) they were interested in participating in as a peer leader. The only requirement of the interested peer leader is that their GPA is higher than 3.0. Each course or program director then interviews candidates and selects for the best fit. Once selected, individuals are required to be trained by the IMPaCT program. This program is designed to prepare students to be an effective leader. The training is done over a semester long course where they explore and discuss issues of diversity, ethics, social justice, community, and civic responsibility in relation to leadership skills. This course creates safe spaces to discuss some of the more difficult topics mentioned above and fosters self-reflection to encourage students to better explore them. These skills the peer leaders develop in the course are then applied when they interact with the students during classroom activities or their office hours. If a student had a conflict that would not allow them to take the course, they could participate in the IMPaCT retreat, which was a weekend-long course that includes much of the material from the semester-long course.

We specifically examined a freshman seminar, sections 1 and 2 of Integrated Sciences 100: Exploring Biology. Exploring Biology is a course that was developed by the Institute for Biology Education (IBE), which is now called the Wisconsin Institute for Science Education and Community Engagement (WISCIENCE), to introduce freshman students to the five core concepts of biology and to understand the relatedness of the field. Exploring Biology was developed under the guidance of the American Association for the Advancement of Science (AAAS) Vision and Change report. It also contains the foundation of a first-year seminar course to provide students with support during the transition from high school to college. Each section contained about 100 students in a WisCEL classroom that fosters student-centered learning with 19 tables of six students. There were two sets of four instructors comprised of graduate students and post-doctoral fellows who were a part of the WISCIENCE Teaching Fellows Program that are supervised by a course director. While the course had implemented peer leaders previously, this was the first time that there was a peer leader for each instructor.

We evaluated each Exploring Biology section separately because they functioned independently of each other. Though the structure for the course was the same with a focus is on varying topics in the life sciences the specific topics taught in each section of the Exploring Biology course was different between sections due to the different set of instructors. The major assignments of the course were very similar and have similar point distributions. The implementation of peer leaders in the course was overall the same. Peer leaders were meant to be support staff in the classroom and help students with the transition process. To achieve this, the peer leaders were assigned to manage and work with different tables of students throughout class, address an array of topics on student success and campus involvement

through Peer Leader Support (PLS) Chats (Appendix C.1), which were approximately 10-15 minutes at the beginning of each class, and hold office hours/be available to students via email. Through these different activities both sets of students were able to interact and develop relationships with their respective peer leaders.

Our hope is that by networking with other students they would develop a feeling of belonging and inclusivity. We also wanted to strengthen this relationship by assigning tables for our peer leaders to visit and make regular conversation with students about the class and their adjustment to college life. Our project also focused on shared discovery because we were trying to get students to broaden their scope of knowledge and seek to learn more about what they found interesting within class assignments as well as on campus. Through all of our different activities the goal was to encourage students to get excited about learning and take advantage of all the University of Wisconsin-Madison campus has to offer through academics, activities, supporting staff and more.

Our overall learning goals (LO) for the project were as follows:

1) Students will describe how they value their peer leaders
2) Students will report a beneficial relationship with the peer leaders
3) Students will reflect on the value of peer leaders and how they would be valuable to others
4) Students will describe skills peer leaders help with
5) Students will assess the strengths of having peer leaders in the classroom
6) Peer leaders will reflect and evaluate their training with the IMPaCT program

LOs 1-4 were targeted with our survey evaluating their relationship with their peer leaders. Survey results were compared to self-reported demographic data, course grade, and the students' pre/post-tests from the course to examine any correlation in success in the course or overall growth over the course. LOs 5 and were evaluated by the focus groups. They provide more qualitative data than our surveys but should provide necessary information from the students about the contributions peer leaders made to any behavioral or psychological changes. These questions were was coded using the content analysis method (Taylor-Powell and Renner 2003), in addition to giving us direct quotes. Much of this data was representative of individual opinions rather than of the whole of the course but summarizing statements could be synthesized about the effectiveness of the program.

3.4 Purpose

Literature has consistently shown that peer mentorship in STEM has contributed to an increase in retention, GPA, and social integration in students, especially in underrepresented groups (Chesler and Chesler 2002; Bowling *et al.* 2015; Snyder *et al.* 2016; Zaniewski and Reinholz 2016). This is believed to be due to students developing a sense of belonging and more positive science identities (Carlone and Johnson 2007; Zaniewski and Reinholz 2016). Yet how peer mentors achieve those outcomes has not been examined. Our study is interested in exploring how first-year STEM undergraduates benefit from having peer mentors in the classroom using a sequential explanatory design. We specifically examined a case with a large freshmen STEM seminar course, Integrated Sciences 100: Exploring Biology, that implements eight peer leaders across its two sections. This study was conducted using three iterations of a

survey across the fall semester to gather quantitative and qualitative data. Students that complete the surveys were invited to participate in the focus groups in the following spring to understand the complexities of this relationship and its benefits. Focus groups were also held with amongst the peer leaders to gain their perspective as well. Our overall goal was to go beyond the standard measures of success the literature presents, such as grades and retention rate, and gather data directly from the students about why they find having a peer mentor in the classroom is beneficial for them. Preliminary data does indeed suggest peer mentors are beneficial, and we plan to use this as a pilot for more extensive studies to solidify our claim in the near future. This study sought to answer the following questions:

- How do first-year STEM undergraduates describe their relationship with their peer leader in a first-year seminar course?
- How do students feel peer mentorship can be improved in a first-year STEM seminar course?

3.5 Methods

3.5.1 Participants

The students that were invited to be part of the study were undergraduates enrolled in the Exploring Biology course. Each class was recruited through a short introduction at the beginning of one of the classes to explain the purpose of the research project. Although students were required to take the surveys in order to receive credit in the course, participation in the study was voluntary. The only requirement other than being a student in the course was that the student had to be 18 to consent to participate. Participation information and consent were at

the beginning of survey timepoints 1 and 2. If students did not want to participate in the study any further, they were able to withdraw by contacting the study team leader by email.

3.5.2 Data collection and analysis: surveys

We used a mixed-methods design, specifically sequential explanatory design, to approach our research study. The quantitative data of the experiment came from our survey questions. The qualitative data is from the three open ended questions on the survey and the focus groups, which was informed by our survey data (see 3.5.3).

Over the course of the semester students received links from Qualtrics (Qualtrics, Provo UT) with our survey evaluating the peer leader experience (Appendix C.2). It is a subsampling of a survey found in the literature to analyze relationships between others: Basic Psychological Needs Scales- focused under "relationships" with three additional free response questions to have students further explain their answers. In addition, students are also asked about how frequently they interact with their mentor and in what capacity. The survey was sent at three timepoints throughout this semester: at the beginning of October, November and December. In Exploring Biology, students were required to take all surveys for class credit, where they received 1 point for each survey completed. Reminders were sent through Qualtrics if the student had not completed it. If a student forgot to complete the survey they were offered the chance to use a reusable link to access the survey and complete it for credit.

For our data students were to complete all three of the surveys in the appropriate order. We had 102 students from both Exploring Biology sections take all the surveys on time and in the appropriate order. We had an additional 44 students who still completed the surveys in order, but they had to use the reusable link to have their survey results connected to their

name. Therefore, a total of 146 students completed in order, consented to all the surveys, and were used to conduct the following analyses.

Quantitative questions were formatted using a likert scale between 1 to 7. A score of "0" represents no answer. Data was analyzed for each question by calculating an average score. Questions were analyzed generally to compare classes as well as separate by other demographic data (parent's education level, gender, class [Monday/Wednesday], interest in biology and race/ethnicity). Two power analyses procedures were conducted to determine further statistical tests. If we were interested in having a medium effect size of 0.15 and a power analysis of 0.95, we would need each of our demographic classes to have a sample size of 699 students. This could only be possible if we were to conduct the experiment over multiple years. Based solely on this year's data, the power analysis was 0.123 when examining race/ethnicity differences and 0.127 when examining gender differences. Due to the lack of power more general statistical analyses were chosen.

Each demographic class was analyzed using general statistical descriptors. To examine for statistical significance Levene's Test was used to examine the variance of each population. A simple t-test was used to examine if there are statistically significant different means ($p < 0.05$) between groups. A 95% confidence interval was also determined. For qualitative/short answer questions, questions were coded using the content analysis method, where emergent codes were identified after combing through the data (Taylor-Powell and Renner 2003). Sometimes multiple codes fit multiple answers.

3.5.3 Data collection and analysis: focus groups

We planned to hold multiple focus groups in the following spring semester: one for each section's peer leaders and student focus groups for each of the Exploring Biology sections. Students received an email before the start of the spring semester if they had completed all the surveys to be recruited for the focus groups. For their participation students were offered pizza for their time and two participants were selected at random to receive $20 Amazon gift cards. Depending on the turn out we received for participating in the focus groups we tried to get a representative sampling of those in the course based on the surveys.

Focus groups were held in the WISCIENCE Department spaces. The interviewer for the focus groups was consistent for all of the Exploring Biology students and was not directly responsible for or affiliated with the course. This was done to try to obtain information that had no biases. The same was true for the peer mentor focus groups but was a separate individual. Depending on the size of the focus group each session lasted between 20 minutes to one hour. All students who participate had to sign a waiver of consent, administered by the interviewer, and were allowed time for any questions or concerns. The scripts for each of the focus groups were prepared in advance (Appendices C.3 and C.4). The student focus group was asked about how they define their relationship with their peer leader in addition to the positive and negatives of having peer leaders in the classroom. The peer leader focus group asked questions to understand how they became peer leaders, their point of view how things went over the semester and what could be done to improve the role in the future. The sessions were audio recorded and transcribed at a later date. After they were transcribed the tapes were destroyed.

Each question was coded and themes were attributed. This allows the ability to pull direct quotes for any future publication or poster.

3.6 Limitations

There are many cofounding factors that may skew our results. As mentioned before, each Exploring Biology course functioned independently from each other. Each unit of the course differed between the two courses and were graded by their own set of instructors. Therefore, we expected there to be an instructor effect between the courses. This could also affect the interaction with the peer leaders since two of the research team members were instructors in the same section of Exploring Biology. The Monday section of Exploring Biology also participated in the First-Year Interest Group (FIG) Program at University of Wisconsin-Madison. In the FIG program students take a freshman seminar course (in this case, Exploring Biology) and two other courses related to the topic to develop an academic learning community. In the case of Exploring Biology this was either math and a humanities course, chemistry and a humanities course, or math and chemistry. This can also lead to a more diverse student population in "FIG'ed" courses, where participation of underrepresented students makes up 25% of the FIG program, but 11% of the freshmen class (*UW-Madison Strategic Diversity Update* 2013).

We also had limitations within our focus groups. While 87.5% of the peer leaders were able to participate in the focus groups we only had a sample size of n=3 for the Exploring Biology focus groups, which in ended up being one-on-one interviews. More information could have been obtained if we had more students we could interview or if we could have had

multiple students in a room at once to interact with one another. We also recognize there were prior unexpected relationships. Of the students that participated in the focus groups, one student had participated in a program or a course led by Dr. Whitmore, the faculty mentor of this project, and therefore could have led to some bias in results.

3.7 Results

3.7.1 Demographics

In total, 147 students completed all three surveys in the appropriate order. 79 of those students were from the Monday section of Exploring Biology while 68 of the students were from the Wednesday section. 27.2% of students were male while 72.8% were female. 74.8% of students self-identified as Caucasian, 0.7% identified as Black or African American, 14.3% identified as Asian, 2.7% identified as Hispanic, 1.4% identified as Other, 1.4% preferred not to disclose their race/ethnicity and 4.8% identified as 2 or more of the above categories. Therefore, for the following analyses we grouped all of the non-Caucasian groups into a category identified as “other”. 12.6% of students reported that their parents were high school graduates or had less than a high school degree. 10.9% reported that their parents attended some college, 7.5% reported that their parents held a 2-year degree, 32.7% said one or both of their parents held a 4-year degree, 28.6% reported that one or both parents held a professional degree and 9.5% reported one or both parents held a doctorate degree. Our demographic data is close to what is representative as a whole at the university, with the exception of gender.

We also asked for students’ interest in biology across the course. At the end of the course, 31.1% of the students reported a “very high” interest in biology, 29.5% reported a

"high" interest in biology, 24.2% reported a "moderate" interest in biology, 6.1% reported a "low" interest in biology and 9.1% reported a "very low" interest in biology.

3.7.2 Quantitative findings: survey

The general trends seen are that the student/peer leader relationship was reported to be more favorable as the semester went on, as expected (Figure 3.1). Over time students reported that they felt more confident, comfortable, and cared about in their relationship with their peer leader. They also consistently reported an increase in the quality of their relationship, more effectively using their peer leader, benefitting from their relationship, belonging on campus (Figure 3.2) and learning valuable skills. As the semester went on they also reported that their peer leader appeared more qualified for their position (Figure 3.3). Interestingly, the only non-positive correlation as time went on was the amount of time spent interacting with peer leaders. There was an increase in the second survey timepoint but it actually went down slightly in the final survey timepoint. However, this does not appear to be statistically significant ($p=0.885$). Either the stress of wrapping up the semester with finals or that students would have already enrolled for classes the following semester were hypothesized to be the causes.

Data was also collected about the ways students and peer leaders interacted outside of class. We found if they did at all it was always 10 minutes or less with two exceptions. Almost all communication was done in person (in class) or by email. Less than five students reported using student office hours. We had less than five instances where students reported having communication with their peer leader through social media or by phone (including texting).

Subgroups were analyzed based on the demographic questions collected during the study. While we collected many demographic questions from students in this chapter we chose

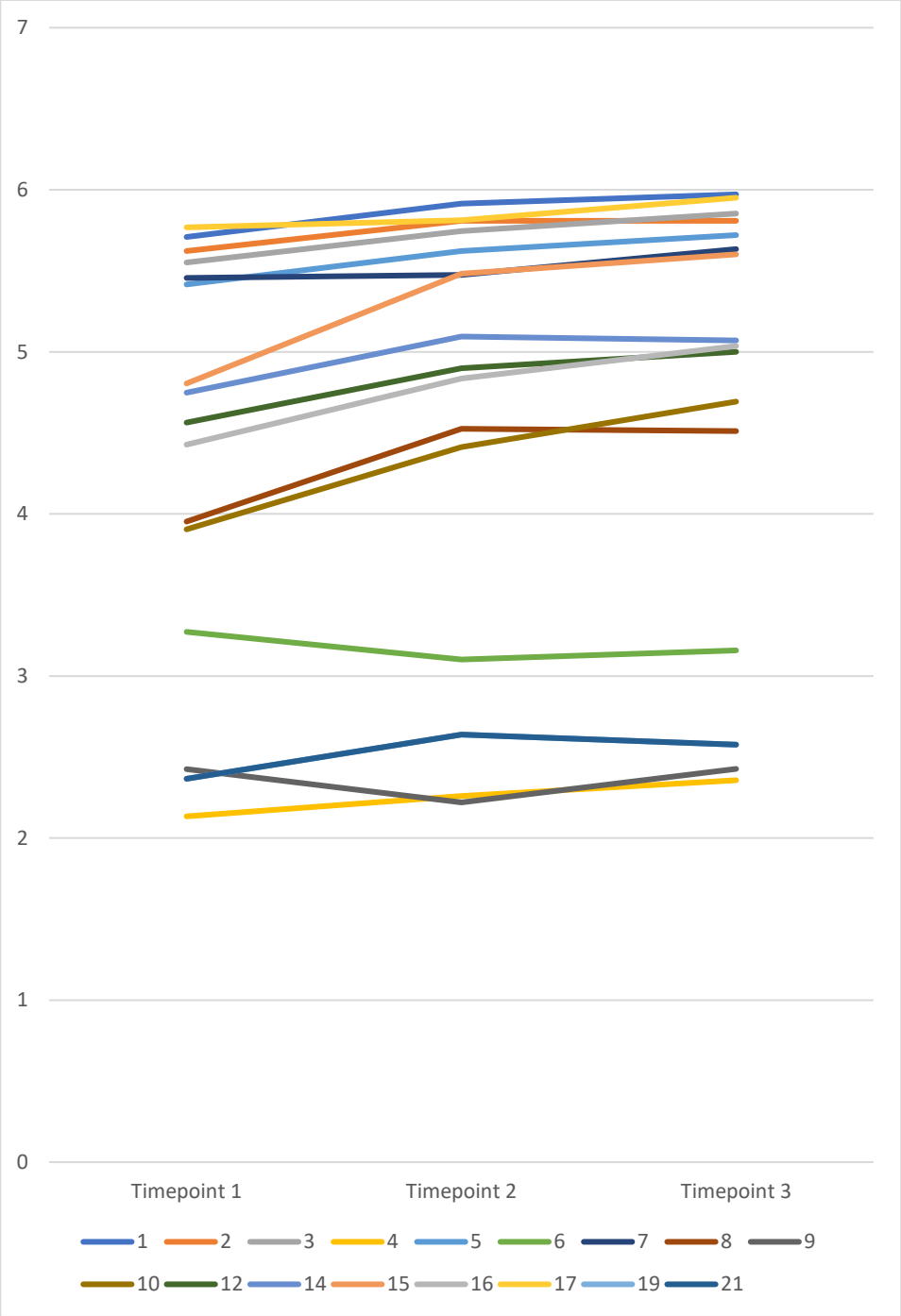


Figure 3.1: Comparing average scores of each question at each timepoint of the survey.

Only question 4, 6, 9 and 17 had a negative downturn over time. Question 4 addressed feelings of inadequacy or incompetency with their peer leader, question 6 was about distance in relationship with peer leader, and question 9 addressed feelings of pressure to be different around peer leaders, all we hoped to stay low or decrease. Question 21 addresses average time spent with peer leaders in class. Full list of questions can be found in Appendix C.2.

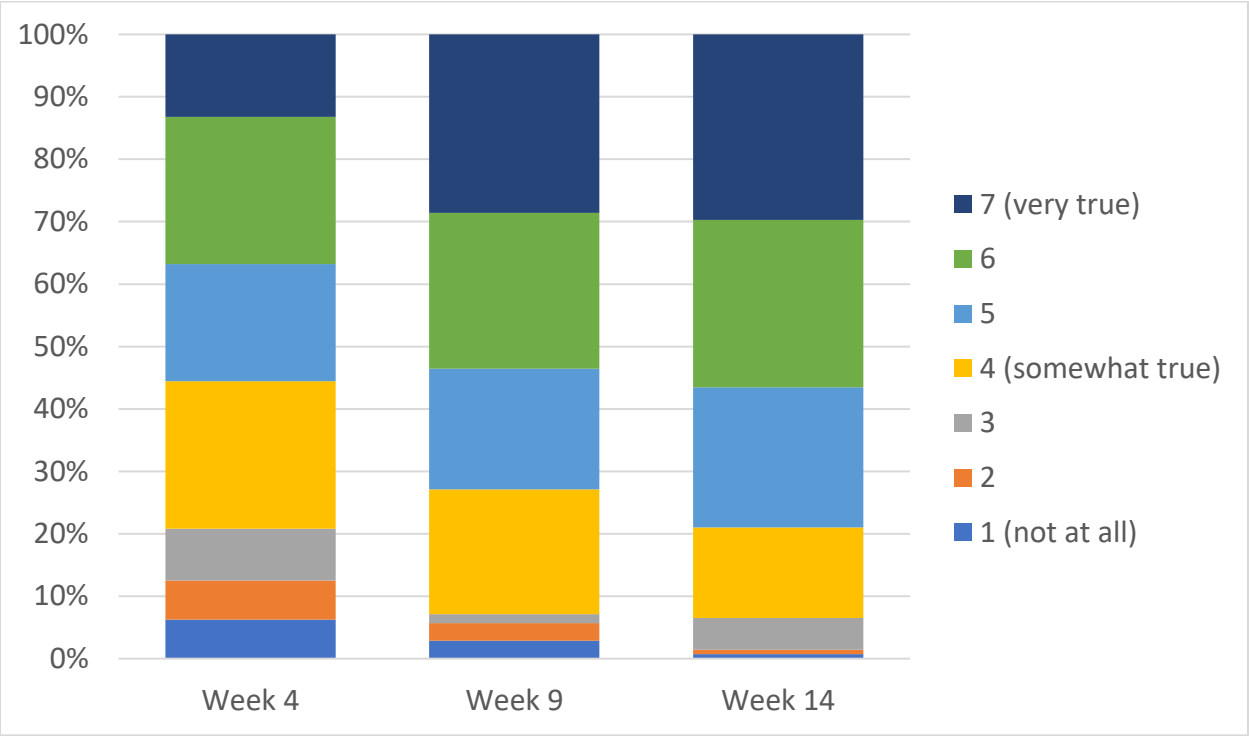


Figure 3.2: Peer leaders help students gain a better sense of how to become involved at UW-Madison.
As the semester progressed, students increasing reported that their peer leader helped them understand how to become involved on UW-Madison's campus. This was most likely from their Peer Leader Support Chats held at the beginning of class. Week 4= timepoint 1, week 9= timepoint 2 and week 14= timepoint 3.

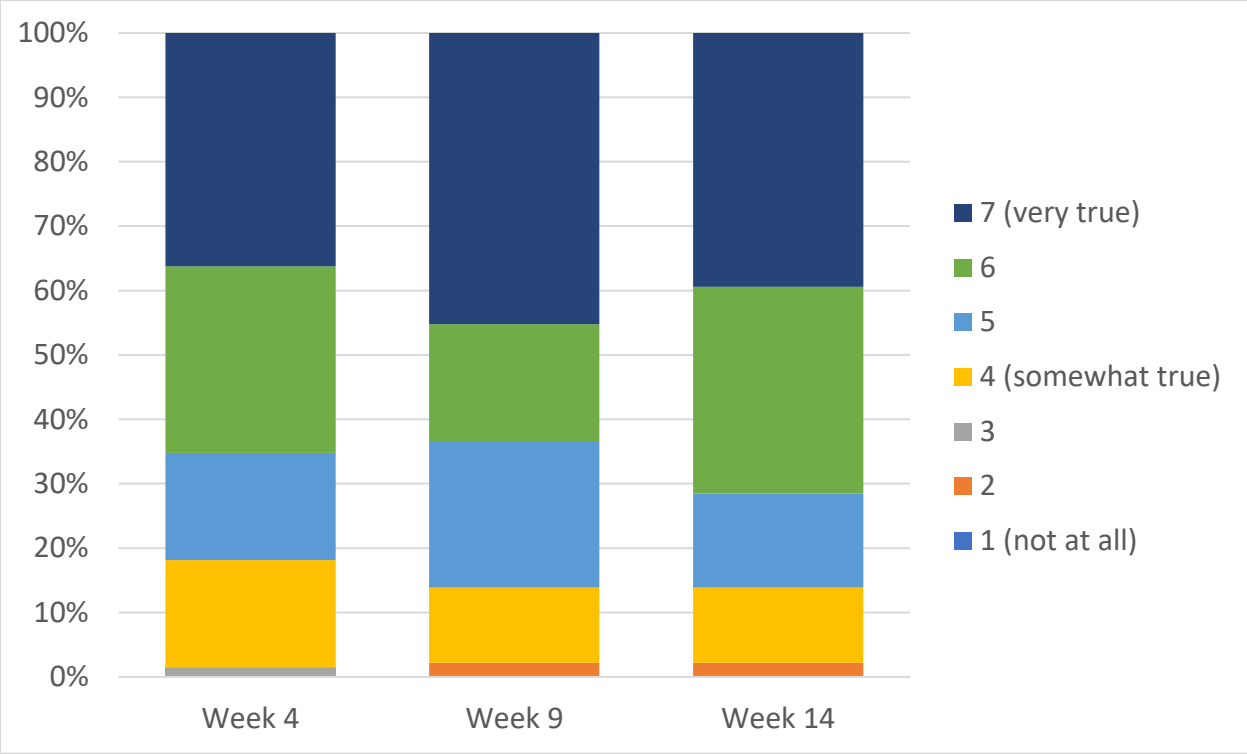


Figure 3.3: Students feel peer leaders are qualified to be mentors.
Many students reported from the beginning of the course that their Peer Leader to be well-qualified. However, as the semester progressed, the average score increased. By the end of the semester, over 70% of students gave a score of either 6 or 7. Week 4= timepoint 1, week 9= timepoint 2 and week 14= timepoint 3.

to focus on race/ethnicity and gender. It was expected that individuals who identified with certain subgroups, such as underrepresented students in the field, may find more of a benefit to having a mentor than others. When comparing groups based on race/ethnicity we did not see that there were statistically significant differences for any of the six questions between groups (Tables 3.2-3.5). The most divergent question was examining the question about benefitting from the peer leader relationship by Levene's Test in the last survey timepoint (p=0.193) (Table 3.5). We also examined differences between genders for the same six questions (Tables 3.6-3.9). In the case of gender, we found that women tend to report gaining more skills than men at the end of the semester (p=0.005) (Table 3.9). This was also near the significance threshold for Levene's Test (p=0.090), but our t-test still concluded significance even if the variances were not equal (p= 0.012) (Table 3.9). However, no other questions appeared to be statistically significant (p > 0.05). There were no statistically significant differences found between responses from students in either the Monday or Wednesday sections.

3.7.3 Qualitative findings: survey

Preliminary coding analyses were conducted on the three short answer questions in the surveys. Each question was analyzed as a whole and not examined individually. The three short answer questions focused on what the relationship was like between the student and their peer leader, how or how are they not using their peer leader effectively, and what skills they were gaining from their peer leader.

Short answer #1 examined the relationship between the peer leader and whether or not it was high quality. Of 433 responses, 262 times students reported a positive relationship with

Table 3.2: Survey statistics examining differences between race/ethnicities from first survey timepoint.

Question	Category	N	Mean	Std. Deviation	Std. Error Mean
High quality relationship	White	109	3.73	1.869	0.179
	Other	37	3.97	1.907	0.314
Effectively utilizing PL	White	109	4.18	1.791	0.172
	Other	37	4.14	1.946	0.320
Recommend program	White	109	4.30	2.012	0.193
	Other	37	4.41	2.006	0.330
Benefitting from PL	White	109	4.35	1.792	0.172
	Other	37	4.00	2.095	0.344
Succeeding in college	White	109	4.59	1.832	0.175
	Other	37	4.57	1.676	0.275
Gaining skills	White	109	4.22	1.941	0.186
	Other	37	3.89	2.025	0.333

Table 3.3: Independent samples test for differences between race/ethnicities at first survey timepoint.

		Levene's Test for Equality of Variances		t-test for equality of the means					95% Confidence interval of the difference	
		F	Sig.	t	df	Sig (2 tailed)	Mean Difference	Std. Error Difference	Lower	Upper
High quality relationship	Equal variances assumed	0.104	.904	-.669	144	.505	-.239	.357	-.946	.468
	Equal variances not assumed			-.662	61.128	.510	-.239	.361	-.961	.483
Effectively utilizing PL	Equal variances assumed	.480	.490	.139	144	.890	.048	.348	-.640	.737
	Equal variances not assumed			.133	58.072	.894	.048	.363	-.678	.775
Recomme	Equal variances assumed	.061	.806	-.268	144	.789	-.103	.382	-.859	.653

	Equal variances not assumed			-.269	62.339	.789	-.103	.382	-.866	.661
Benefitting from PL	**Equal variances assumed**	307	.581	.979	144	.329	.349	.356	-.355	1.053
	Equal variances not assumed			.906	54.969	.369	.349	.385	-.423	1.120
Succeeding in college	**Equal variances assumed**	1.031	.312	.057	144	.954	.020	.341	-.655	.694
	Equal variances not assumed			.060	67.435	.952	.020	.327	-.632	.671
Gaining skills	**Equal variances assumed**	.021	.885	.879	144	.381	.328	.373	-.410	1.066
	Equal variances not assumed			.861	60.009	.393	.328	.381	-.434	1.091

Table 3.4: Survey statistics examining differences between race/ethnicities from last survey timepoint.

Question	Category	N	Mean	Std. Deviation	Std. Error Mean
High quality relationship	White	109	4.43	1.838	0.176
	Other	37	4.38	2.126	0.349
Effectively utilizing PL	White	109	4.35	1.823	0.175
	Other	37	4.35	1.814	0.298
Recommend program	White	109	4.75	1.722	0.165
	Other	37	4.89	2.105	0.346
Benefitting from PL	White	109	4.72	1.851	0.177
	Other	37	4.92	1.831	0.301
Succeeding in college	White	109	5.41	1.486	0.142
	Other	37	5.42	1.658	0.273
Gaining skills	White	109	4.77	1.767	0.169
	Other	37	4.62	2.060	0.339

Table 3.5: Independent Samples Test for differences amongst race/ethnicities at last survey timepoint.

		Levene's Test for Equality of Variances		t-test for equality of the means					95% Confidence interval of the difference	
		F	Sig.	t	df	Sig (2 tailed)	Mean Difference	Std. Error Difference	Lower	Upper
High quality relationship	Equal variances assumed	.454	.501	.145	144	.855	.053	.364	-.667	.733
	Equal variances not assumed			.135	55.391	.893	.053	.391	-.731	.837
Effectively utilizing PL	Equal variances assumed	.359	.550	-.008	144	.994	-.003	.346	-.687	.682
	Equal variances not assumed			-.008	62.461	.994	-.003	.346	-.693	.688
Recomme	Equal variances assumed	1.022	.314	-.402	144	.688	-.140	.347	-.826	.547

	Equal variances not assumed			-.364	53.294	.717	-.140	.383	-.909	.629
Benefitting from PL	**Equal variances assumed**	1.713	.193	-.579	144	.688	-.140	.347	-.826	.547
	Equal variances not assumed			-.364	53.294	.717	-.140	.383	-.909	.629
Succeeding in college	**Equal variances assumed**	.058	.809	.057	144	.955	.017	.291	-.559	.592
	Equal variances not assumed			.054	56.918	.957	.017	.307	-.599	.632
Gaining skills	**Equal variances assumed**	.655	.420	.425	144	.672	.149	.351	-.545	.843
	Equal variances not assumed			.394	55.095	.695	.149	.379	-.610	.908

Table 3.6: Survey statistics examining differences between gender from first survey timepoint.

Question	Category	N	Mean	Std. Deviation	Std. Error Mean
High quality relationship	Male	40	3.85	1.929	0.305
	Female	106	3.77	1.863	0.181
Effectively utilizing PL	Male	40	4.30	1.814	0.287
	Female	106	4.12	1.835	0.178
Recommend program	Male	40	4.45	2.230	0.353
	Female	106	4.28	1.921	0.187
Benefitting from PL	Male	40	4.35	1.902	0.301
	Female	106	4.23	1.868	0.181
Succeeding in college	Male	40	4.28	1.921	0.304
	Female	106	4.70	1.730	0.168
Gaining skills	Male	40	3.90	2.098	0.332
	Female	106	4.23	1.909	0.185

Table 3.7: Independent Samples Test for differences amongst gender at first survey timepoint.

		Levene's Test for Equality of Variances		t-test for equality of the means					95% Confidence interval of the difference	
		F	Sig.	t	df	Sig (2 tailed)	Mean Difference	Std. Error Difference	Lower	Upper
High quality relationship	Equal variances assumed	.006	.941	.219	144	.827	.076	.349	-.614	.766
	Equal variances not assumed			.215	68.165	.830	.076	.355	-.631	.784
Effectively utilizing PL	Equal variances assumed	.053	.819	.523	144	.602	.177	.339	-.494	.848
	Equal variances not assumed			.525	70.972	.601	.177	.338	-.496	.851
Recomme	Equal variances assumed	1.101	.296	.448	144	.655	.167	.373	-.570	.904

	Equal variances not assumed			.419	62.099	.677	.167	.399	-.630	.964
Benefitting from PL	**Equal variances assumed**	.013	.909	.355	144	.723	.124	.348	-.565	.812
	Equal variances not assumed			.352	69.167	.726	.124	.351	-.577	.824
Succeeding in college	**Equal variances assumed**	.210	.647	-1.278	144	.203	-.423	.331	-1.077	.231
	Equal variances not assumed			-1.219	64.285	.227	-.423	.347	-1.117	.270
Gaining skills	**Equal variances assumed**	.483	.488	-.897	144	.371	-.326	.364	-1.046	.393
	Equal variances not assumed			-.859	64.830	.393	-.326	.380	-1.085	.432

Table 3.8: Survey statistics examining differences between gender from last survey timepoint.

Question	Category	N	Mean	Std. Deviation	Std. Error Mean
High quality relationship	Male	40	4.35	2.070	0.327
	Female	106	4.44	1.852	0.180
Effectively utilizing PL	Male	40	4.40	1.878	0.297
	Female	106	4.33	1.798	0.175
Recommend program	Male	40	4.60	1.630	0.258
	Female	106	4.86	1.890	0.184
Benefitting from PL	Male	40	4.83	1.781	0.282
	Female	106	4.75	1.872	0.182
Succeeding in college	Male	40	5.08	1.607	0.254
	Female	106	5.55	1.481	0.144
Gaining skills	Male	40	4.05	2.062	0.326
	Female	106	4.99	1.687	0.164

Table 3.9: Independent Samples Test for differences amongst gender at last survey timepoint.

		Levene's Test for Equality of Variances		t-test for equality of the means					95% Confidence interval of the difference	
		F	Sig.	t	df	Sig (2 tailed)	Mean Difference	Std. Error Difference	Lower	Upper
High quality relationship	Equal variances assumed	.003	.957	-.263	144	.793	-.093	.355	-.795	.608
	Equal variances not assumed			-.250	63.955	.803	-.093	.373	-.839	.653
Effectively utilizing PL	Equal variances assumed	.005	.943	.207	144	.837	.070	.338	-.598	.737
	Equal variances not assumed			.203	67.629	.840	.070	.345	-.618	.757
Recomme	Equal variances assumed	1.058	.305	-.764	144	.446	-.258	.338	-.927	.410

	Equal variances not assumed			-.817	80.866	.416	-.258	.316	-.888	.371
Benefitting from PL	**Equal variances assumed**	1.001	.319	.232	144	.817	.080	.343	-.598	.758
	Equal variances not assumed			.238	73.545	.813	.080	.335	-.588	.748
Succeeding in college	**Equal variances assumed**	.104	.747	-1.678	144	.095	-.472	.281	-1.028	.084
	Equal variances not assumed			-1.617	65.486	.111	-.472	.292	-1.055	.111
Gaining skills	**Equal variances assumed**	2.907	.090	-2.821	144	**.005***	-.941	.333	-1.600	-.282
	Equal variances not assumed			-2.577	59.775	**.012***	-.941	.365	-1.671	-.210

their peer leader, whereas 154 times students reported a neutral relationship and 2 times students reported a negative relationship (although it was reported to be neutral or positive in following survey timepoints). When students were asked to describe their relationship the highest reported relationship was describing a mentor-student relationship (140), followed by a teacher-student relationship (106) and then acquaintances (94). Some students reported a friendship between them and their peer leader (28), classmates (3) or even familial relationship (2). There were some students who reported no relationship (60). Many who reported no relationship consistently reported no relationship throughout the survey timepoints.

Short answer #2 asked students to explain how they are or are not using their peer leader effectively. The most reported answer by far was that students use their peer leaders effectively by asking them questions (219). Many times, students did not clarify what their questions were about so it was hard to differentiate if these questions were about Exploring Biology or more broadly about college and UW/campus-related issues. Some students did say they specifically sought advice from their peer leaders (39) or used peer leaders to specifically ask about class (36) or for campus resources (26). There were students who recognized that they did not use peer leaders to their full potential. Occasionally some students clarified they should ask more questions (51) or seek more advice from their peer leaders (11). There were others that said they could be using more of their resources (34) or that they specifically do not utilize a resource, which was many times noted to be their office hours (52). The second most common answer was that students did not interact with their peer leader (68).

The last short answer question probed students about the skills peer leaders give them (Figure 3.4). The most reported answer was that the peer leaders help them find campus

Figure 3.4: Students report gaining skills from their peer leader.
Students were asked three times throughout the semester what skills they gained from their peer leader, and these are a tally of all their answers. We found that students would list different skills gained each survey, unless they reported no skills gained.

resources or how to get involved (153), followed by soft skills such as time management and study skills (89), problem solving/critical thinking/other academic skills (79), communication/social skills (39), helped them personally (21), and professional skills (12). Some students reported the skills they learned were from the topics covered in their peer leader Support (PLS) Chats (10), but the actual skills may fall into any of the categories above. There were also students who reported that peer leaders did not help them gain any skills (56).

3.7.4 Focus groups: students

In reality, our focus groups turned out to be interviews with students due to lack of response from students. Two students from the Monday Exploring Biology course and one from the Wednesday Exploring Biology course were interviewed. The interviews elicited a wide variety of responses.

Two students cited that they had a positive experience with the peer mentors while the other felt it was a neutral experience. This appeared to be due to the expectations of the student and their preconceived notions to what a peer leader was. The student that had a neutral experience conveyed that they felt the peer leader's role was to help more with the class material itself, even though peer leaders are explicitly not there to teach content. The other students described the peer leader's role as more of an advising role, where they could ask about majors or resources on campus. These two students also appear to take advantage of the resources that were given to them by the peer leaders, while the student who had a neutral experience seemed to use the peer leaders just to ask questions about the content in class, although they did mention that their advice was useful.

All students cited that they felt the peer leaders had adequate training but sometimes needed clarification on some of the activities in class. This echoes some of what the peer leaders said—that sometimes communication was not always clear between the instructors and the peer leaders. Some students appeared to be more forgiving if they found value in some other aspect of the peer leader.

The two students that reported a positive experience said they themselves would be interested in pursuing peer mentoring in the future, whether for this course or for other programs offered through IMPaCT. Both felt that they would like to give back to freshmen like themselves to learn about resources on a campus as large as UW-Madison, and also each mentioned how it would help them develop new skills they were interested in gaining, such as communication or presentation skills. Another factor that could relate to their interest in being a peer leader is that they knew people who mentored others. Each reported that they knew individuals who participated in peer mentoring programs, whether it was their friends or acquaintances they had met on campus.

3.7.5 Focus groups: peer leaders

The peer leaders reported a positive overall experience during their time as a peer mentor. Many of the peer leaders said they pursued the IMPaCT program either because they themselves had peer leaders in a course or program and wanted to replicate that for others, or they had learned about the opportunity through others who had participated in a course or program with peer leaders. They all felt that they either benefitted or could have benefitted from having peer mentors as freshmen and they wanted to share their experiences. Some had mentioned participating in IMPaCT would help gain communication skills and experience

working with others. Peer leaders felt that their training did prepare them as well as it could have for the class and that the skills taught in the training were useful to them. However, they all seemed nervous about implementing them in the classroom. They felt they were most valuable by giving their weekly presentations in class or actually seeing students implement what was discussed. The peer leaders discussed that they enjoyed the connections they formed with students and helping advise students on topics beyond the class. Some reported feeling least valuable when they held office hours and they found few students actually took advantage of the resource. Others felt that all experiences, whether good or bad, were learning experiences and therefore positive. Peer leaders reported still staying in contact with a few former students, whether it is by seeing them on campus or through social media or email. Throughout their participation in the program peer leaders felt that they grew and gained communication and interpersonal skills along the way and many are planning on continuing some form of mentoring in the upcoming future through IMPaCT. One of the peer mentors even goes as far as saying “being a part of [IMPact] is one of the best choices I made once I came to campus”. Another stated that “it has been a part of my undergrad experience” here at UW-Madison and “it has shaped who she is today”. All report a “positive” or “very positive” experience.

The few suggestions of ways to improve the program were focused on the theme of better communication, whether that be with the instructional staff or the students. Here are some examples of what the peer leaders suggested for improvement:

- Clarification to the students what the role of the peer leader is in the course—they felt as if students did not warm up to them as much because they feel they were presented

more like instructors. This was something that was echoed in the student surveys, where sometimes students felt they were not colleagues but rather instructor-like. The peer leaders felt this could have been remedied at the beginning of the course.

- Clarification on the course to the students: One peer leader suggested clarifying the idea of the course or at least better communicating the purpose of the course to the students, which may also aid the suggestion above. If students better understand the purpose of the course and the purpose of having peer leaders in such a course, it would make the course more effective as a whole.
- Possibly more peer leaders: Some peer leaders also felt it could have been better if there were more peer leaders because of the limit on time they could spend with the students. They felt that if they were required to meet with less students then they could have more of an effect.
- Adding incentives for students to come to office hours: As noted above, some peer leaders felt their office hours were the time they felt least valuable because many times no students attended. A suggestion was to offer points (whether assigned to their final grade or as extra credit) to attend office hours at some point during the semester.
- More time to meet and chat with students in class: Sometimes due to the lesson plan peer leaders had limited time to chat with students within the class time. A peer leader suggested possibly have ice breakers, increase the amount of break time or have mini discussions led by the peer leaders to discuss topics more at length. They suggested this is another way to encourage more interaction with peer leaders.

- Strengthen the relationship between the instructional team and peer leaders earlier, develop clearer communication between the two groups and learning how to better adjust to the physical classroom set-up: Those that suggested clear communication between the instructors and peer leaders did have a different meeting arrangement than the others— with the Monday peer leaders the weekly instructor met with all peer leaders to discuss the plan for class the following week while for Wednesday peer leaders only the most senior peer leader attended the weekly instructional meeting. Therefore, it sounds like it may be advantageous to pursue the Monday peer leader arrangement in future class iterations. It was also mentioned that when the outline of the activities for a lesson were presented clearer and had additional materials for them, that helped the peer leaders be more effective in the classroom.
- Meeting former peer leaders before the fall semester: It was also discussed that having former course peer leaders meet with the new peer leaders and sharing their experiences would be beneficial. They may provide tips or stories that could help them become better peer leaders.

3.8 Discussion

We found, as predicted, that students would report benefitting from having peer leaders in the classroom. We also found many students felt that peer leaders have given them valuable skills they can use during their time at UW and beyond. Some of these factors did appear to have larger increases in scores over the course of the semester in underrepresented populations. This was seen in both the surveys and the focus groups/student interviews.

3.8.1 The Effects for Underrepresented Populations in STEM

Although we had a small population (n=3) for our focus group, our two students who did have positive experiences and discussed interest in being a peer leader themselves were underrepresented students, with one being a first-generation college student. She explicitly stated how helpful it was to have an older student to ask questions to because she didn't have someone to advise her on college being the first in her family to go to a 4-year school. This was somewhat echoed in the survey, but more data would need to be collected to draw conclusions.

This pilot study shows some results that may support our conclusion but one of the biggest problems is our lack of diversity in our populations. White women alone make up 58.5% of our population. 27.4% of our study was male, which is drastically lower than the population of males at the university (48.8%). By repeating this study in the upcoming fall, hopefully we can increase our numbers to have a representative population that we can draw conclusions from. We also hope to increase our focus group numbers to get a better representation of students and their experiences as well.

Our limited findings suggesting value in peer mentors echo the literature that has shown evidence that programs like this significantly can help grades and retention rates, a problem UW currently needs to address. At UW-Madison, for all courses in 2016-2017 academic year underrepresented minority (URM) students have a D/F/Drop rate of 11%, compared to non-URM student D/F/Drop rate of 6% (*Grade Gap Report for Undergraduate Courses in 2016-2017* 2017). First-generation college students tend to also have an increase in D/F/Drop rate than non-first-generation college students by 50% in courses overall. With introductory STEM

courses this is very apparent. In Biology 151, which is the first course in the 2-semester intro Biology course sequence here on campus, URM students had a 20% D/F/Drop rate over the course of the 2016-2017 year, compared to 7% D/F/Drop rate for non-URM students. The two introductory chemistry courses, Chemistry 103 and 104 see similar disparities (18% versus 10%, 14% versus 7%, respectively). This trend is also repeated for introductory math courses as well (for 101: Intermediate Algebra it is 33% versus 16%, and for College Algebra 112 it is 31% versus 13%). This trend sometimes continues for other higher-level courses as well, such as Genetics 466, Botany 260, Math 221 & 222, and Physics 103.

When it comes to retention, URM students are 71.6% likely to return for the second year while the general average for all students in the UW system is 80.7% (*Information Memorandum on Retention and Graduation*). The six-year graduation rate continuing at the same UW institution is also lower amongst URM students, averaging at 41.5%, whereas the general average is almost 50% greater at 59.3% graduating from the institution (*Information Memorandum on Retention and Graduation*). By creating support networks early on with classmates and upperclassmen mentors, underrepresented students can learn about resources that are on campus or create study groups to lower that drop rate. It would be interesting to possibly follow up with these students in the future to see if the exposure to all these resources affected their general outcomes in college. The other possibility would be trying to extend the IMPaCT program to more introductory courses to extend those benefits to others. The IMPaCT program has recently begun training peer leaders for the introductory Biochemistry course; however, these students function more in the PLTL format discussed in the introduction.

3.8.2 Peer Leader Support (PLS) Chats

We believe that the activity with the most major impact in the peer mentoring relationship was the Peer Leader Support (PLS) chats. While students did highlight that they received personal advice from their peer leaders or other more intimate or personalize resources, we consistently found that one of the largest influences the peer mentors had was on student involvement on campus (Figures 3.2 and 3.4). It was not surprising, given the fact that the PLS chats has the most participants and sometimes the only interaction students had with their peer leader. Therefore, it is possible to conclude that our results, such as the skills gained, are influenced by the topics chosen in those chats (Appendix C.1), where many of the topics are centered around UW's resources and campus involvement. Peer leaders may choose to emphasize these topics for two different reasons: 1) peer leaders themselves may have chosen these select topics due to their own struggles and challenges during their first year in college, or 2) could be influenced by the mentoring training curriculum.

it is possible that the next set of peer leaders may cover entirely different topics during their chats that could yield different results if the experiment were repeated. Hopefully, we can work with the peer mentors about working with the students more on the topics of the chats presented in the course. By integrating students' opinions on what they are interested in hearing or gathering feedback on the topics peer leaders discussed in the chats, we can improve the peer leader experience by giving students more power over their learning environment. While this may lead to different results from this study, we can better address challenges the first-year students have.

3.9 Future Directions

With our current results we effectively answered our learning goals with our assessments. With that being said, our team plans to make small changes to the phrasing of some questions before we repeat the experiment in the upcoming year. While the short answer questions did supply us with data that fit our learning goals it would have been better to provide a multiple-choice format for some of our free response questions. For example, something that was unexpected was that students sometimes had a hard time addressing who was a peer leader and who was an instructor. Some data had to be thrown out because it was clear the data was not about a peer leader but an instructor. In the future we plan to re-word that question to provide students the names of the peer leaders to verify the answers they chose were describing the peer leaders. This could have skewed our results in question 27 when students were to describe skills they have gained, and that may be why we have seen such high results for critical thinking/problem solving. However, in our student interviews one of our students did describe peer leaders encouraging them to think critically about the material in the course.

We also hope to spread the survey out a little more across the semester instead of at the beginning of each month or reduce the data down to two survey collection points. This will most likely reduce the survey fatigue we could see with the students. We are also interested in conducting similar surveys or focus groups with other programs IMPaCT peer leaders participate in. Since peer leaders have slightly different roles in different programs it would be interesting to see if the benefits of peer leaders are seen to be similar amongst all the different groups of students or if there are areas of improvement for these other peer leaders that are in other courses or programs.

The adjustments suggested by the peer leaders have already started to be implemented. The two sections of Exploring Biology will deliver the same content but just have different instructional staff and peer leaders. This will help both the instructional staff as well as the peer leaders form more of a community and foster a support system for one another. The suggested changes of clarifying the role of the peer leader in the course has also been added to the agenda for the course. The peer leaders will all be expected to meet with a member of the instructional staff once a week to prepare for the upcoming lesson. The IMPaCT student association and their monthly meetings helps develop connections between new and old peer mentors. All of these changes will hopefully strengthen the role of peer leaders in the course.

3.10 Conclusion

Thus far, our data supports our hypotheses of peer leaders benefitting first-year STEM students and their transition to college. They have reported gaining from this relationship and gained skills from their peer leaders over the semester. We have evidence that students feel their peer leaders are trained well, and most suggestions for improvement fall into the application within the Exploring Biology course itself and not the actual training of the peer leaders. After sharing this data at the 2018 UW-Madison Teaching and Learning Symposium we have found more programs and instructors interested in implementing the IMPaCT in their own classrooms as a way to improve their learning communities.

3.11 Acknowledgements

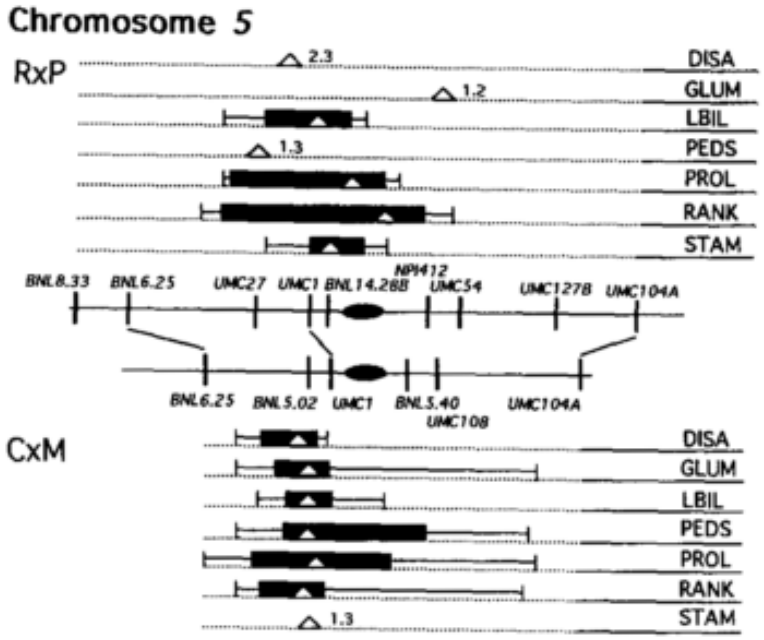
Special thanks are given to our research team, Kassi Crocker, Clay Smith, Christopher Trimby and Jerry Whitmore Jr. They have been a fantastic team to work with, even though this is a side project for all of us. I would also like to thank Amanda Butz for assisting with running focus groups and advice on our IRB. Devin Wixon gave us great assistance on the development and feedback on the project. The Delta Internship Fall 2017 cohort also provided feedback and advice on the project. Cara Theisen and the WISCIENCE Teaching Fellows were kind enough to let us use the course as a pilot for our study and were grateful for their cooperation, even with our study's hiccups. We would also like to thank the peer leaders of Exploring Biology for cooperating with us and allowing us to conduct such work.

Appendices

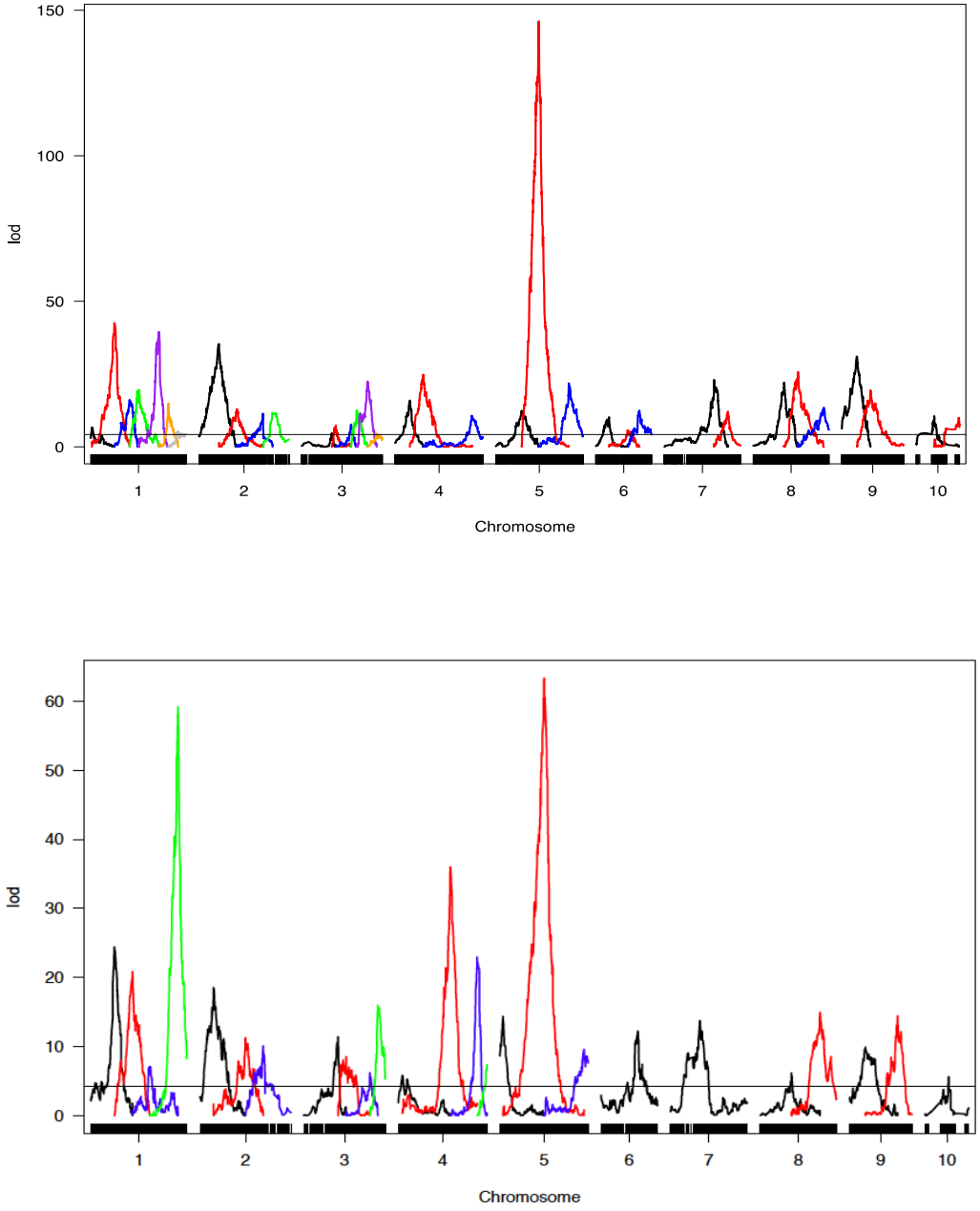
Appendix A:

Fine-mapping a major domestication QTL for ear size traits on chromosome five in *Zea Mays*

A.1 Figures

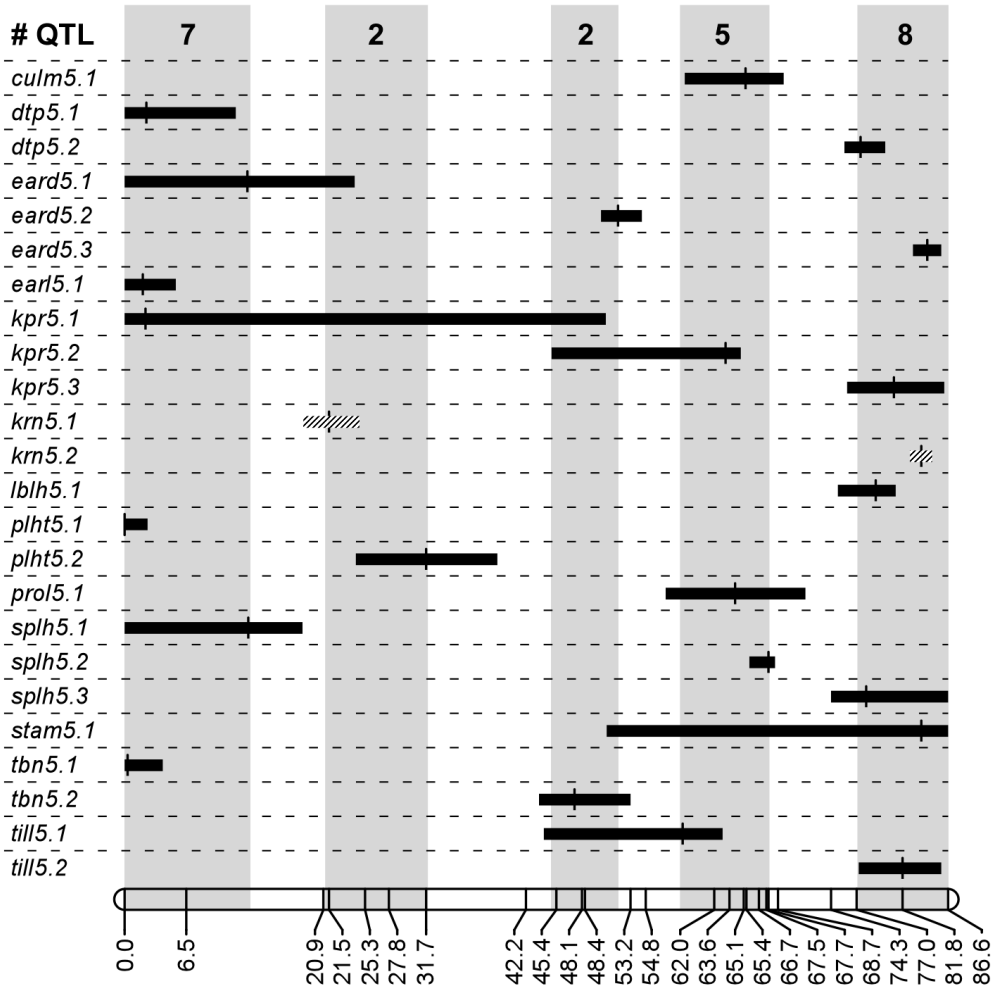


Supplemental Figure A.1: Chromosome Five QTL from F2 Populations.
Doebley and Stec (1991, 1993) published domestication QTL studies examining results from two different F2 crosses of maize and teosinte. In the middle of the image, there is a graphical representation of two chromosome maps, with the black oval being the centromere. Each tick mark on the line labels PCR markers that were used to genotype the F2 population. CxM is the cross between Chapalote (a maize landrace) and *Zea mays ssp. mexicana* (a teosinte) (Doebley and Stec 1991). RxP is the cross between Reventador (a maize landrace) and *Zea mays ssp. parviglumis* (a teosinte) (Doebley and Stec 1993). Both "rank" (number of rows of cupules) and "LBIL" (average length of internodes in the primary lateral inflorescence) are ear size-related domestication traits and show significant QTL on chromosome five designated by the triangles. Black solid bars represent 1-LOD support interval and whiskers represent 2-LOD support interval.



Supplemental Figure A.2: Shannon's QTL mapping of the ear diameter trait (top) and kernel row number (bottom) in the BC2S3 population.

X-axis is the position of markers along the genome, chromosome by chromosome. Y-axis is the LOD score. In the top figure for ear diameter, there appears to be a very large peak on chromosome five with a LOD score of 144.43. This was the largest QTL found in her entire domestication QTL study. In the bottom figure, you will see the QTL map for the kernel row number trait. The largest peak is found on chromosome five with a LOD score of 65.83 and is located directly downstream from the large peak for ear diameter.



Supplemental Figure A.3: Regions of domestication QTL on the short arm of chromosome five.

X-axis is the chromosomal map of a portion of chromosome five in map units. The QTL named "eard.5.3" is the major effect QTL seen in Figure 3, and it overlaps with the QTL labeled "krn5.2" in a region that is 2.654 Mbp in length. Modified from Lemmon and Doebley (2014).

A2: Tables

Supplemental Table A.1: Markers for Original RC-NIL Design.
The primers used are outlined below, as well as where they map to within the maize genome (AGPv2 and AGPv4). The size of the B73 product is listed.

Marker (Primer set)	Left Primer	Right Primer	Location	B73 Amplicon Size (bp)	Polymorphism
umc1348	CTCACTGACACTTGAACACACACG	TTACTGGTCTCCTGATCCTTAGCG	166,576,570-166,576,708 **(AGPv2)** 170,513,623-170,513,761 **(AGPv4)**	138	Teo 8bp deletion
umc1221	GCAACAGCAACTGGCAACAG	AAACAGGCACAAAGCATGGATAG	168,671,955-168,672,050 **(AGPv2)** 172,643,917-172,644,012 **(AGPv4)**	95	Maize 10bp deletion
umc1966	GTTTTCGACGAGGGGACTACATTT	CACGGTTGAGAACTTCGCTTGTAG	169,230,959-169,231,114 **(AGPv2)** 173,180,225-173,180,380 **(AGPv4)**	155	Teo 20bp deletion

Supplemental Table A.2: LSMs developed from B11b for ear size phenotypes.
List of each RC-NIL, their LSMs and standard errors for all phenotypes scored.

Recombinant	LSM: Ear Diameter	Standard Error	LSM: Kernel Row Number	Standard Error	LSM: Kernels per Rank	Standard Error
A001-5	27.9982	0.2433	11.8339	0.1893	22.1986	0.4701
A002-2	28.6347	0.2412	11.8889	0.1753	24.1667	0.4636
A003-2	30.2569	0.2412	12.8333	0.1753	22.8611	0.4636
A004-4	29.775	0.2412	11.8333	0.1753	23.9722	0.4636
A005-1	30.4056	0.2412	12.3889	0.1753	23.6667	0.4636
A006-1	29.1319	0.2412	11.6111	0.1753	21.5278	0.4636
A007-4	29.4167	0.2412	12.0556	0.1753	19.6389	0.4636
A008-2	30.9472	0.2412	12.8889	0.1753	24.1667	0.4636
A009-5	29.0403	0.2412	11.9444	0.1753	22.0833	0.4636
A010-4	30.6639	0.2412	12.7222	0.1753	25.6389	0.4636
A013-3	30.2875	0.2412	12.1111	0.1753	24.0278	0.4636
A014-4	27.6903	0.2412	11.5556	0.1753	20.8611	0.4636
A015-1	28.0206	0.2478	11.9932	0.1821	22.0269	0.4839
A017-3	30.7351	0.2433	12.5726	0.1775	23.2587	0.4701
A018-1	29.1389	0.2412	11.8333	0.1753	22.75	0.4636
A019-5	29.4056	0.2412	12	0.1753	23.5833	0.4636
A021-5	30.1014	0.2412	12.2778	0.1753	24.1111	0.4636
A023-1	30.1569	0.2412	12.3333	0.1753	21.0556	0.4636
A024-4	29.9651	0.2433	12.7441	0.1775	22.9444	0.4636
A026-6	29.9708	0.2412	12.1111	0.1753	22.2873	0.4701
A027-4	30.8706	0.2433	12.6224	0.1775	23.8277	0.4701
A028-5	31.2681	0.2412	12.1667	0.1753	23.5278	0.4636
A032-5	27.7306	0.2412	11.5556	0.1753	20.7778	0.4636
A033-8	29.5528	0.2412	11.9444	0.1753	20.4167	0.4636
A034-5	30.583	0.2528	12.6399	0.1873	22.9571	0.4991
A035-5	31.2803	0.2433	12.3982	0.1775	23.9945	0.4701
A037-5	28.8639	0.2412	11.3333	0.1753	21.5556	0.4636
A041-5	29.8847	0.2412	12	0.1753	21.1944	0.4636
A042-5	30.4439	0.2433	12.1126	0.1775	22.9444	0.4636
A043-3	31.1403	0.2412	12.5	0.1753	23.8889	0.4636
A044-4	30.7139	0.2412	12.4444	0.1753	21.6667	0.4636
A045-3	28.5222	0.2412	12	0.1753	18.9444	0.4636
A046-2	29.5514	0.2412	12.5	0.1753	23.1944	0.4636
A047-5	30.4111	0.2412	12.3333	0.1753	23.1389	0.4636
A048-3	28.5236	0.2412	12.0556	0.1753	21.5833	0.4636
A050-1	29.8111	0.2412	12.2222	0.1753	23.9193	0.4701
A051-5	30.2611	0.2412	12.4444	0.1753	21.8333	0.4636
A052-1	30.4889	0.2412	12.5556	0.1753	24.3333	0.4636
A053-5	28.7931	0.2412	12.1667	0.1753	21.5556	0.4636
A054-1	32.2	0.2412	12.5556	0.1753	26.5278	0.4636
A057-1	29.9583	0.2412	12.6111	0.1753	22.6944	0.4636
A059-2	29.8194	0.2412	12.0556	0.1753	22.4444	0.4636

A060-5	30.6486	0.2412	12.0556	0.1753	25.25	0.4636
A061-4	31.325	0.2412	12.7222	0.1753	24.1111	0.4636
A062-4	29.8556	0.2412	12.0556	0.1753	23.2778	0.4636
A063-2	30.0236	0.2412	12.4444	0.1753	21	0.4636
A064-3	30.3833	0.2412	12.8333	0.1753	23.8889	0.4636
A065-7	31.0806	0.2412	12.6667	0.1753	25.0556	0.4636
A066-1	29.2833	0.2412	11.8333	0.1753	20.0833	0.4636
A067-2	29.4833	0.2412	12.1111	0.1753	22.5278	0.4636
A069-5	31.4625	0.2412	12.3889	0.1753	25.3333	0.4636
A070-3	29.6417	0.2412	12.5556	0.1753	22.0556	0.4636
A071-4	30.6251	0.2433	12.0584	0.1775	22.0016	0.4701
A072-1	29.3125	0.2412	12.3889	0.1753	20.4722	0.4636
A073-8	30.0431	0.2412	12.0556	0.1753	23.4722	0.4636
A074-3	28.9806	0.2412	12.0556	0.1753	20.6944	0.4636
A075-3	29.1431	0.2412	12.1667	0.1753	20	0.4636
A076-5	29.9139	0.2412	12.4444	0.1753	22.3333	0.4636
A078-4	29.7431	0.2412	12.7222	0.1753	22.7778	0.4636
A081-2	27.1167	0.276	12.1193	0.2105	20.2724	0.5667
A082-8	29.9097	0.2412	12.2778	0.1753	20.5	0.4636
A083-3	29.9222	0.2412	12.5	0.1753	22.0833	0.4636
A084-1	29.25	0.2412	12	0.1753	21.9722	0.4636
A085-2	28.5333	0.2412	12.0556	0.1753	21.3611	0.4636
A086-5	29.2153	0.2412	12.2778	0.1753	21.8056	0.4636
A087-2	30.3111	0.2412	12.5556	0.1753	23.1389	0.4636
A088-5	30.2583	0.2412	12.5556	0.1753	20.8611	0.4636
A089-3	30.3847	0.2412	12.5556	0.1753	22.3611	0.4636
A090-2	28.8444	0.2412	12.0556	0.1753	21.25	0.4636
A091-3	30.1764	0.2412	12.3333	0.1753	22.8889	0.4636
A092-3	28.9097	0.2412	11.9444	0.1753	21.6667	0.4636
A094-4	31.0903	0.2412	12.8889	0.1753	21.8611	0.4636
A095-1	29.5764	0.2412	12.3333	0.1753	21.6944	0.4636
A097-11	30.4556	0.2412	12.7778	0.1753	23.0278	0.4636
A098-10	30.5944	0.2412	12.2778	0.1753	22.1389	0.4636
A101-12	27.7809	0.2455	11.9454	0.1798	19.934	0.4768
A102-3	29.8791	0.2478	12.4267	0.1821	19.6365	0.4913
A104-3	25.2208	0.2412	10.6667	0.1753	15.6667	0.4636
A169-4	31.4514	0.2412	13.0556	0.1753	23.8056	0.4636
A170-2	30.7319	0.2412	12.6111	0.1753	21.5833	0.4636
A182-1	31.0889	0.2412	12.6111	0.1753	26.8333	0.4636
A186-3	28.2196	0.2433	11.7697	0.1775	21.005	0.4701
A190-10	29.65	0.2412	12.1111	0.1753	20.7778	0.4636
A192-4	31.2819	0.2412	12.3889	0.1753	24.1389	0.4636
A199-1	29.5833	0.2412	12.3889	0.1753	22.8056	0.4636
A200-4	31.0042	0.2412	12.2222	0.1753	24.8889	0.4636
A201-3	30.4542	0.2412	11.9444	0.1753	22.7222	0.4636
A208-1	28.5917	0.2412	12.2778	0.1753	21.8333	0.4636

A210-11	30.7361	0.2412	12.7778	0.1753	21.3056	0.4636
A219-4	29.4567	0.2433	12.284	0.1775	21.7193	0.4701
A225-2	29.4467	0.2433	12.4554	0.1775	23.4476	0.4768
A226-1	31.7528	0.2412	12.8889	0.1753	22.2778	0.4636
A228-5	29.2514	0.2412	12.1111	0.1753	21.3056	0.4636
A238-1	30.2972	0.2412	12.7222	0.1753	23.4722	0.4636
A247-1	28.8203	0.2433	11.8839	0.1775	20.6111	0.4636
A249-5	29.5903	0.2412	12.5556	0.1753	23.5278	0.4636
A250-5	29.2806	0.2412	12.3333	0.1753	21.8056	0.4636
A263-5	29.6028	0.2412	12.5	0.1753	20.6111	0.4636
A265-2	29.6528	0.2412	12.1111	0.1753	22.9444	0.4636
A268-9	31.4125	0.2412	13.1667	0.1753	23.1389	0.4636
A272-2	30.7061	0.2433	12.574	0.1775	21.0584	0.4701

Supplemental Table A.3: LSMs of B46 derived lines for both ear size phenotypes.
List of each RC-NIL, their LSMs and standard errors for all phenotypes scored.

Recombinant	LSM: Ear Diameter	Standard Error	LSM: Kernel Row Number	Standard Error	LSM: Kernels per Rank	Standard Error
A105-1	32.3778	0.3033	13.0556	0.2073	22.1986	0.4701
A108-2	31.4194	0.3033	13.5	0.2073	24.1667	0.4636
A109-5	34.1222	0.3033	14	0.2073	22.8611	0.4636
A110-5	33.4958	0.3033	12.9444	0.2073	23.9722	0.4636
A112-4	33.017	0.3075	12.8217	0.2131	23.6667	0.4636
A114-1	31.6806	0.3033	13.1111	0.2073	21.5278	0.4636
A115-3	33.8667	0.3033	13.2778	0.2073	19.6389	0.4636
A118-5	31.8139	0.3033	12.3889	0.2073	24.1667	0.4636
A119-5	31.1139	0.3033	12.6667	0.2073	22.0833	0.4636
A121-1	32.1542	0.3033	12.6667	0.2073	25.6389	0.4636
A122-9	32.2903	0.3033	12.3333	0.2073	24.0278	0.4636
A125-5	30.8317	0.3054	12.6279	0.2101	20.8611	0.4636
A129-4	31.2819	0.3033	12.8333	0.2073	22.0269	0.4839
A131-3	30.2949	0.3075	12.1767	0.2131	23.2587	0.4701
A132-1	33.2903	0.3033	13.6111	0.2073	22.75	0.4636
A133-4	33.0148	0.3054	12.6846	0.2101	23.5833	0.4636
A134-5	31.3667	0.3033	13.3333	0.2073	24.1111	0.4636
A137-1	33.0361	0.3033	13.8889	0.2073	21.0556	0.4636
A141-6	32.5681	0.3033	12.9444	0.2073	22.9444	0.4636
A142-1	33.412	0.3054	13.5989	0.2101	22.2873	0.4701
A143-8	34.2167	0.3033	13.7778	0.2073	23.8277	0.4701
A144-2	34.7181	0.3033	14.3889	0.2073	23.5278	0.4636
A145-3	32.0358	0.3054	12.9709	0.2101	20.7778	0.4636
A146-2	31.8708	0.3033	13.4444	0.2073	20.4167	0.4636
A147-3	34.7417	0.3033	14.0556	0.2073	22.9571	0.4991
A148-2	31.4208	0.3033	13.3333	0.2073	23.9945	0.4701
A149-5	30.8528	0.3033	12.5556	0.2073	21.5556	0.4636
A150-5	30.9528	0.3033	12.6111	0.2073	21.1944	0.4636
A153-2	30.3639	0.3033	13.2778	0.2073	22.9444	0.4636
A154-10	31.5472	0.3033	13.4444	0.2073	23.8889	0.4636
A155-5	32.9721	0.3054	14.1749	0.2101	21.6667	0.4636
A156-2	31.8097	0.3033	12.4444	0.2073	18.9444	0.4636
A157-3	34.0472	0.3033	13.8889	0.2073	23.1944	0.4636
A158-1	32.9514	0.3033	13.2222	0.2073	23.1389	0.4636
A159-5	33.3917	0.3033	13.2778	0.2073	21.5833	0.4636
A162-2	32.5611	0.3033	13.4444	0.2073	23.9193	0.4701
A165-1	33.3833	0.3033	12.8333	0.2073	21.8333	0.4636
A168-4	33.0934	0.3054	13.7132	0.2101	24.3333	0.4636
A275-1	32.6139	0.3033	13.3333	0.2073	21.5556	0.4636
A276-3	32.2319	0.3033	12.8333	0.2073	26.5278	0.4636
A280-1	33.5301	0.3054	13.7137	0.2101	22.6944	0.4636
A282-2	31.9958	0.3033	13.2778	0.2073	22.4444	0.4636

A285-1	33.4264	0.3033	13.5	0.2073	25.25	0.4636
A289-5	32.3417	0.3033	13.3333	0.2073	24.1111	0.4636
A292-1	32.8847	0.3033	13.7778	0.2073	23.2778	0.4636
A293-1	32.5181	0.3033	12.8333	0.2073	21	0.4636
A295-3	32.7267	0.3075	13.111	0.2131	23.8889	0.4636
A296-3	32.1129	0.3054	12.5709	0.2101	25.0556	0.4636
A297-3	32.2115	0.3054	12.8566	0.2101	20.0833	0.4636
A299-10	32.4014	0.3033	13.1667	0.2073	22.5278	0.4636
A304-5	32.5292	0.3033	13.6667	0.2073	25.3333	0.4636
A305-1	31.5125	0.3033	12.6111	0.2073	22.0556	0.4636
A307-3	34.5236	0.3033	14	0.2073	22.0016	0.4701
A308-4	32.4764	0.3033	13.0556	0.2073	20.4722	0.4636
A309-4	32.3597	0.3033	12.7222	0.2073	23.4722	0.4636
A321-1	29.7542	0.3033	13	0.2073	20.6944	0.4636
A329-1	33.4208	0.3033	13.6111	0.2073	20	0.4636
A330-3	31.4344	0.3054	13.1396	0.2101	22.3333	0.4636
A332-1	33.45	0.3033	13.8333	0.2073	22.7778	0.4636
A333-5	33.8062	0.3054	13.1989	0.2101	20.2724	0.5667
A335-1	32.4056	0.3033	12.9444	0.2073	20.5	0.4636
A336-2	32.6875	0.3033	12.2778	0.2073	22.0833	0.4636

Supplemental Table A.4: Markers for New RC-NIL Design.
Since this involves looking at a larger region of interest on chromosome five spanning about 20 Mbp, we have multiple markers within that region to genotype our different lines.

Primer Set Name	Primers	Location	B73 Amplicon Size (bp)	Polymorphism
320	TCAAAAACACAATTGATCAACAAA GCTGGATGGTTGAGTGCAG	145,738,660-145,738,909 **(AGPv2)** 149,178,313-149,178,562 **(AGPv4)**	249	Teo 3bp deletion
ZHL0033-ZHL0034	AGCAAAGCATGGGCTAGTGT GCCATGCTGCTTATGGATCT	151,446,508-151,446,736 **(AGPv2)** 155,134,027-155,134,255 **(AGPv4)**	228	Maize 1bp deletion
TIDP5668	AACTCACTCACATCCTCACCG TGACGAGCTCCGTTAGTTCC	154,808,141-154,808,557 **(AGPv2)** 158,594,015-158,594,431 **(AGPv4)**	416	Teo 4bp deletion
ZHL0301	AACAGCTTTGCTTCCCTGAA CCCAGAGGATCCAGAGTCAG	159,447,674-159,447,786 **(AGPv2)** 163,216,947-163,217,059 **(AGPv4)**	112	Teo 4bp deletion
IDP722	ACCAAGCGTTAAGACAACGG ATCACCATCTTCCTCATCGC	163,897,431-163,897,721 **(AGPv2)** 167,774,918-167,775,208 **(AGPv4)**	290	Significant deletion in maize, PCR marker
umc1348	CTCACTGACACTTGAACACACACG TTACTGGTCTCCTGATCCTTAGCG	166,576,570-166,576,708 **(AGPv2)** 170,513,623-170,513,761 **(AGPv4)**	138	Teo 8bp deletion

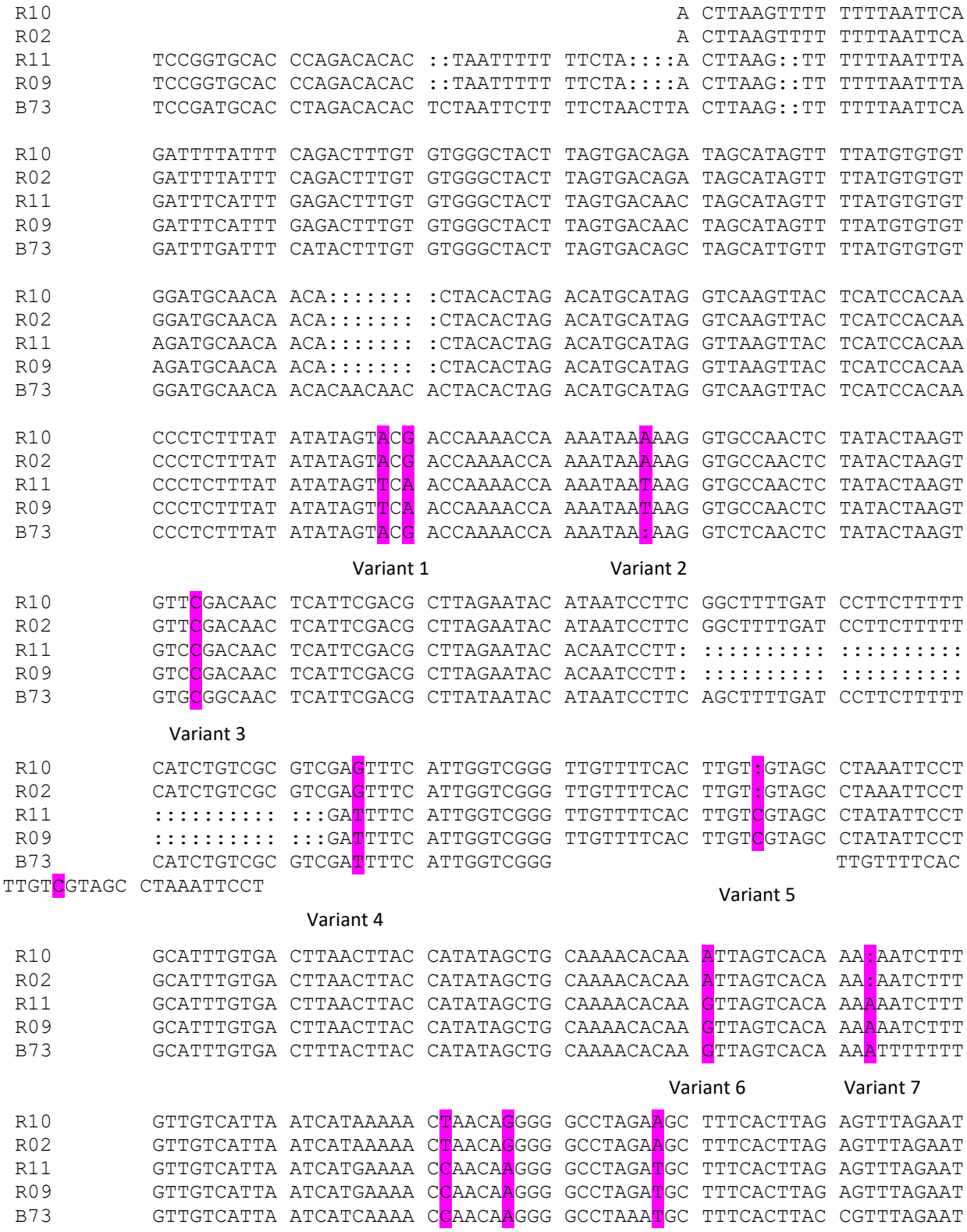
Appendix B:

Fine-mapping a major domestication QTL for ear shattering trait on chromosome five in *Zea Mays*

B.1: Figures

Supplemental Figure B.1: Sequence Alignment between P6 and P7.

Variants between maize and teosinte are highlighted.



Variant 8 and 9 Variant 10

```
R10  GAAGGGATGG GTACTAAAGT ATTGTAGTAG AATATGAATA GATTAGACAG ATACTTA::G
R02  GAAGGGATGG GTACTAAAGT ATTGTAGTAG AATATGAATA GATTAGACAG ATACTTA::G
R11  GAAGGGATGG GTACTAAAGT ATTGTAGTAG AATATGAATA GATTAGACAG ATACTTATAG
R09  GAAGGGATGG GTACTAAAGT ATTGTAGTAG AATATGAATA GATTAGACAG ATACTTATAG
B73  GGAGGGATGG GTACTAAAGT ATTGTAGTAG AATATGAATA GATTGGACAT ATG:GT::C
```

Variant 11

```
R10  CTATATGGTC CTATATTATT TTACCCATGG ATTTCGTTCT AACACTTGTC TATTTTATTA
R02  CTATATGGTC CTATATTATT TTACCCATGG ATTTCGTTCT AACACTTGTC TATTTTATTA
R11  CTATATGGTC CTATATTATT TTATCCATGG ATTTCGTTCT AACACTTGTC TATTTTATTA
R09  CTATATGGTC CTATATTATT TTATCCATGG ATTTCGTTCT AACACTTGTC TATTTTATTA
B73  CTAT::G::: :T:::T:ATT TTATCCATGG ATTTCGTTCT AACACTTGTC ::TGTT:TTA
```

Variant 12

```
R10  CTTAAAAACA CTCGTGTATG CGGAAGAGTG CTGTTTAATG GACTAGGGGG GGTA::ACAA
R02  CTTAAAAACA CTCGTGTATG CGGAAGAGTG CTGTTTAATG GACTAGGGGG GGTA::ACAA
R11  CTTAAAAACA CTCGTGTATG CGGAAGAGTG CTGTTTAATG GACTAGGG:: :GTA::ACAA
R09  CTTAAAAACA CTCGTGTATG CGGAAGAGTG CTGTTTAATG GACTAGGGGG GGTA::ACAA
B73  CTTAAAAACA CTCGTGTATG CGGAAGAGTG CTGTTTAATA TAAT:GGACT AGTAGGACAA
```

Variant 13

```
R10  AATAATACAG AAATTATTTG GACCAGCGCC GAGGAAGACA ATCCGTTTTG AATGAATCTG
R02  AATAATACAG AAATTATTTG GACCAGCGCC GAGGAAGACA ATCCGTTTTG AATGAATCTG
R11  AATAATACAG AAATTATTTG GACCAGCGCC GAGGAAGACA ATCCGTTTTG AATGAATCTG
R09  AATAATACAG AAATTATTTG GACCAGCGCC GAGGAAGACA ATCCGTTTTG AATGAATCTG
B73  AATAATACAG AAATTATTTG GACCAGCGCC GAGGAAGACA ATTCGTTTTG AATGAATCTG

R10  CTGCTGGCTG CTGAGTAGTC AGACATAGAG AATCCTTCTA GACACAGCGA TGTGCCGGCC
R02  CTGCTGGCTG CTGAGTAGTC AGACATAGAG AATCCTTCTA GACACAGCGA TGTGCCGG::
R11  C:GCTGGCTG CTGAGTAGTC AGACATAGAG AATCCTTCTA GACACAGCGA TGTGCCGG::
R09  CTGCTGGCTG CTGAGTAGTC AGACATAGAG AATCCTTCTA GACACAGCGA TGTGCCGGCC
B73  CTGCTGGCTG CTGAGTAGTC GGACATAGAG AATCCTTCTA GACACAGCGA CGTGCCGG::
```

Variant 14 Variant 15

```
R10  GGTGCCCCAG CCATTNCTTG T:TTCTCCAA CGACC
R02  ::TGCCCCAG CNATTCATGG GCTTCCCCCA C
R11  ::TGCCCCAG CAATTCATGG GCTTCGCCC: CNCCC
R09  GGTGCCCCAG CCATTCATGG GCNTCCCCN: C
B73  ::TGCCCCAG CAATTCATTG TGGCATTTCG CAGCCCACAT CAACCCCTTC ACAC::GAAC
```

Variant 15

```
B73  CAAGAAAAGC CACTACTGCT TTCTCTCTCT CTCTCTCACA CACACACACA GACACAAATA

B73  AAAAGAAATC AGTAGTTCGA TT
```

B.2: Tables

Supplemental Table B.1: Primer sets used for fine-mapping.

The primers used are outlined below, as well as where they map to within the maize genome (AGPv4). The size of the B73 product as well as what type of polymorphism that occurs within the amplicon is listed. Primers sets 1 and 9 (known as P1 and P9) were used as flanking markers, while primer sets 2-8 (or P2-P8) were used for fine-mapping.

Primer Set Name	Primer Creator	Primers	Position (AGPv4)	Position of Duplicate (AGPv4)	B73 Amplicon Size (bp)	Polymorphism
Primer Set 1-GRMZM5G883855	DW/CY	CAAAGCTGGGTRTGCTTCCA TGCTGAACCGTTGCTTGATT	16,558,691-16,558,456	N/A	273	Maize 5bp deletion
Primer Set 2-158-160	WW	AGGTCTCGAAAGGGCATCA GGCATCGTAGAAGAGAATGG	16,642,491-16,643,210	N/A	719	Teo 3bp deletion
Primer Set 3-28-30	WW	GGCCCTGTGAAAGTAGCAAC AATTTAGGAGGGACTGAACTAAAC	16,649,364 - 16,648,828	N/A	536	Teo 2bp deletion
Primer Set 4-55-120	55-CY 120-WW	ACGTGCAGCATATGATCCAG CTCGCCATGCATGACATGAAC	16,651,886-16,652,386	16,709,143–16,709,643	500	PCR?
Primer Set 5-80-126	WW	GTTCTGGACAGGAAGTGATTG CATATAGCTTTGTTTGGTGC	16,652,848-16,653,231	16,710,105–16,710,533	371	PCR?
Primer Set 6-123-124	WW	GAGGTCAGATAGGATTGGACG AATCAGTAGTTCGATTTCTC	16,656,490-16,656,621	16,713,792–16,713,923	131	PCR?
Primer Set 7-17-91	WW	TGATGCCTACTATGTAAACTAGTCG ATATAAAGAGGGTTGTGGATGAG	16,714,948-16,714,639	N/A	355	Maize 4bp deletion
Primer Set 8-209-210	WW	TGCATGTGCCAGCTACCCT CGCGTACTATGCTATGCTAC	16,720,077-16,720,433	N/A	356	Teo 2bp deletion
Primer Set 9-GRMZM2G157310	DW/CY	ACACTGTGAGAGCACACTTC GACTTCGATGGCACGCTCAC	16,770,388-16,770,768	N/A	380	Maize 5bp deletion

Supplemental Table B.2: Combined shattering LSMs.
List of each RC-NIL, their LSMs and standard errors for the shattering phenotype. This was using both years (2014, 2015) of data.

Stock	RC-NIL Name	LSM Estimate	Standard Error (SE)
R01	IN2209	1.0969	0.05741
R02	IN2223	1.5512	0.05951
R03	IN2258	1.0364	0.06060
R04	IN2301	1.2388	0.05785
R05	IN2405	2.3994	0.06975
R06	IN2509	1.1443	0.05695
R07	IN2540	1.0686	0.05880
R08	IN2587	1.0968	0.06641
R09	IN2632	1.4098	0.05979
R10	IN2691	2.1549	0.06976
R11	IN2863	1.0867	0.07347
R12	IN2927	2.2437	0.05832
R13	IN3069	2.3718	0.06498
R14	IN3071	2.6055	0.07076
R15	IN3131	1.1969	0.05716
R16	IN3168	2.8646	0.07065

Supplemental Table B.3: 2014 Fine-mapping data for all traits.
List of each RC-NIL, their LSMs and standard errors for all phenotypes scored.

Fine-mapping Trait	Years Scored	RC-NIL	LSM	Standard Error (SE)
Cupule Length	2014	IN2209	4.7046	0.06405
Cupule Length	2014	IN2223	4.8136	0.06167
Cupule Length	2014	IN2258	4.7006	0.06120
Cupule Length	2014	IN2301	4.6498	0.06068
Cupule Length	2014	IN2405	5.0885	0.06233
Cupule Length	2014	IN2509	4.6088	0.06409
Cupule Length	2014	IN2540	4.4621	0.05989
Cupule Length	2014	IN2587	4.7110	0.08291
Cupule Length	2014	IN2632	4.5902	0.05953
Cupule Length	2014	IN2691	4.5420	0.1198
Cupule Length	2014	IN2863	4.7030	0.1585
Cupule Length	2014	IN2927	4.8254	0.05971
Cupule Length	2014	IN3069	4.7349	0.1203
Cupule Length	2014	IN3071	4.8950	0.1569
Cupule Length	2014	IN3131	4.5059	0.06284
Cupule Length	2014	IN3168	4.9434	0.1110
Ear Diameter	2014	IN2223	31.6509	0.3249
Ear Diameter	2014	IN2258	32.0250	0.3227
Ear Diameter	2014	IN2301	32.5789	0.3231
Ear Diameter	2014	IN2405	29.4604	0.3346
Ear Diameter	2014	IN2509	30.9583	0.3427
Ear Diameter	2014	IN2540	32.6288	0.3188
Ear Diameter	2014	IN2587	30.2629	0.4529
Ear Diameter	2014	IN2632	29.9540	0.3126
Ear Diameter	2014	IN2691	31.0640	0.6836
Ear Diameter	2014	IN2863	31.8426	0.9258
Ear Diameter	2014	IN2927	30.7081	0.3175
Ear Diameter	2014	IN3069	30.2456	0.6984
Ear Diameter	2014	IN3071	29.3575	0.8609
Ear Diameter	2014	IN3131	32.2106	0.3309
Ear Length	2014	IN3168	31.5226	0.6333
Ear Length	2014	IN2209	15.6575	0.2981
Ear Length	2014	IN2223	16.7933	0.2924
Ear Length	2014	IN2258	15.3471	0.2908
Ear Length	2014	IN2301	16.2195	0.2877
Ear Length	2014	IN2405	17.4772	0.2925
Ear Length	2014	IN2509	15.6794	0.2995
Ear Length	2014	IN2540	16.0311	0.2848

Ear Length	2014	IN2587	16.3085	0.3654
Ear Length	2014	IN2632	16.1499	0.2853
Ear Length	2014	IN2691	15.9667	0.4933
Ear Length	2014	IN2863	16.6675	0.6294
Ear Length	2014	IN2927	16.7215	0.2844
Ear Length	2014	IN3069	17.3205	0.4882
Ear Length	2014	IN3071	17.4551	0.6523
Ear Length	2014	IN3131	16.2292	0.2971
Ear Length	2014	IN3168	18.1928	0.4597
Shattering	2014	IN2209	1.0750	0.08944
Shattering	2014	IN2223	1.3835	0.09047
Shattering	2014	IN2258	1.0250	0.08944
Shattering	2014	IN2301	1.2500	0.08944
Shattering	2014	IN2405	2.4074	0.09268
Shattering	2014	IN2509	1.2115	0.09154
Shattering	2014	IN2540	1.0750	0.08944
Shattering	2014	IN2587	1.0928	0.1248
Shattering	2014	IN2632	1.4250	0.08944
Shattering	2014	IN2691	1.5673	0.1751
Shattering	2014	IN2863	1.0629	0.2462
Shattering	2014	IN2927	2.0500	0.08944
Shattering	2014	IN3069	1.9673	0.1751
Shattering	2014	IN3071	2.2629	0.2462
Shattering	2014	IN3131	1.2328	0.09507
Shattering	2014	IN3168	2.6673	0.1751

Supplemental Table B.4: 2015 Fine-mapping data for Shattering Trait.
List of each RC-NIL, their LSMs and standard errors for the shattering phenotype.

RC-NIL	LSM	Standard Error (SE)
IN2209	1.1045	0.06822
IN2223	1.6376	0.07212
IN2258	1.0307	0.07489
IN2301	1.2267	0.06914
IN2405	2.3774	0.09486
IN2509	1.1114	0.06650
IN2540	1.0564	0.07107
IN2587	1.1089	0.07431
IN2632	1.3874	0.07319
IN2691	2.2530	0.07318
IN2863	1.1098	0.07490
IN2927	2.3424	0.07009
IN3069	2.4373	0.06692
IN3071	2.6418	0.07159
IN3131	1.1805	0.06569
IN3168	2.9072	0.07430

Supplemental Table B.5: Primer sets used to genotype the F_2 population segregating for both DIS1.1 and DIS5.1.

The positions of the primers in the maize genome (AGPv4) and the B73 amplicon size are shown.

Primer Set Name	Primers	Position (AGPv4)	Amplicon Size (bp)	Polymorphism
Chromosome 1: CJ080-CJ223	GACGACATTCCTTCGCCTTG CTGGATGGCCCTCTGATTAG	264,942,508-264,942,672	164	Teo 14bp deletion
Chromosome 5: Mt028-Mt03	AGTGAAAACAACCCGACCAA TGTRGATGCAACAACACTACA	16,714,469-16,714,683	214	Teo 32bp deletion

Supplemental Table B.6: LSMs of F2 population shattering interaction data.
List of each genotype combination, their LSMs and standard errors for the shattering phenotype.

Genotypes at DIS1.1	Genotypes at DIS5.1	LSM	Standard Error (SE)
maize	maize	1.0385	0.1579
het	maize	1.7586	0.1460
teosinte	maize	1.6667	0.1644
maize	het	1.7407	0.0996
het	het	2.1776	0.0778
teosinte	het	1.6939	0.1150
maize	teosinte	1.9412	0.1381
het	teosinte	2.1842	0.0934
teosinte	teosinte	1.7843	0.1128

Supplemental Table B.7: Stocks used for RIL interaction experiment.
The table lists the 28 different stocks that represent the four genotypic classes plus the W22 control. Each stock was grown in one plot per block with the exception of W22 our control that had three plots per block.

Stock	Stock Description	Genotypes at DIS1.1 : DIS5.1	Stock ID
1	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; M)	M:M	61751
2	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; M)	M:M	61786
3	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; M)	M:M	61793
4	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; M)	M:M	61809
5	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; M)	M:M	61818
6	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; M)	M:M	61834
7	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; M)	M:M	61840
8	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; M)	M:M	61847
9	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; T)	M:T	61768
10	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; T)	M:T	61800
11	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; T)	M:T	61808
12	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; T)	M:T	61839
13	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; T)	M:T	61856
14	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (M; T)	M:T	61876
15	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; M)	T:M	61752
16	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; M)	T:M	61766
17	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; M)	T:M	61774
18	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; M)	T:M	61825
19	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; M)	T:M	61850

20	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; T)	T:T	61747
21	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; T)	T:T	61772
22	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; T)	T:T	61776
23	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; T)	T:T	61778
24	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; T)	T:T	61798
25	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; T)	T:T	61802
26	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; T)	T:T	61806
27	W22-ETB1_IN1207 x QTLShatter5c_TCL F3 (T; T)	T:T	61844
28	W22 control	M:M	59329

Supplemental Table B.8: LSMs of RIL interaction data.

List of each genotype combination, their LSMs and standard errors for the all phenotypes measured.

Fine-mapping Trait	Genotypes at DIS1.1 : DIS5.1	LSM	Standard Error (SE)
Cupule Length	M;M	4.8148	0.0426
Cupule Length	M;T	5.5379	0.0390
Cupule Length	T;M	5.2479	0.2115
Cupule Length	T;T	6.3461	0.1782
Ear Diameter	M;M	30.1115	0.4674
Ear Diameter	M;T	28.5741	0.4513
Ear Diameter	T;M	28.4031	0.5020
Ear Diameter	T;T	25.9720	0.4534
Ear Length	M;M	12.6260	0.2899
Ear Length	M;T	13.1281	0.2829
Ear Length	T;M	12.2019	0.3057
Ear Length	T;T	11.2087	0.2837
Kernel Row Number	M;M	13.0908	0.1366
Kernel Row Number	M;T	12.8959	0.1295
Kernel Row Number	T;M	12.9391	0.1508
Kernel Row Number	T;T	12.1498	0.1306
Shattering	M;M	1.1366	0.0846
Shattering	M;T	1.9094	0.0817
Shattering	T;M	1.3583	0.0908
Shattering	T;T	1.5629	0.0821

Supplemental Table B.9: Primers used for Population Genetics Analysis.
The positions of the primers in the maize genome (AGPv4) and the B73 amplicon size are shown.

Primer Set Name	Primer Creator	Primers	Position (AGPv4)	Position of Duplicate (AGPv4)	B73 Amplicon Size (bp)
Primer Set 1-148-163	CY	TGTGGTTCAGATGAATGGTGA AGAAGTACCAGCGTGCTGTG	16,554,546-16,554,916	N/A	370
Primer Set 2-158-160	DW/CY	AGGTCTCGAAAGGGCATCA GGCATCGTAGAAGAGAATGG	16,642,491-16,643,210	N/A	719
Primer Set 3-28-30	WW	GGCCCTGTGAAAGTAGCAAC AATTTAGGAGGGACTGAACTAAAC	16,648,805-16,649,383	N/A	578
Primer Set 4-73-97	WW	TCTTGGACTTTCTGGGGATG TAACCCGGACATAAGCCATAG	16,649,314-16,649,978	N/A	664
Primer Set 5-19-145	WW	CAAGTTGCAGTGCACGTAGC AAGCCACTACTGCTTTCTC	16,656,285-16,656,681	16,713,587-16,713,983	396
Primer Set 6-42-44	WW	TATAGTCAAGTTTTAAAATGTGGC ATGGACCCTTCTATAGAGTCTCCTA	16,715,169-16,716,162	N/A	993
Primer Set 7-55-56	WW	TCACCGTCCCCGTTCCCTC GGTAGTGCTGATACAACGCTTCTC	16,770,548-16,771,115	N/A	567

Supplemental Table B.10: Samples used for diversity of *yab6*.

Line	Type	Race/Species	Country
MR01	Landrace Inbred	Araguito	Venezuela
MR02	Landrace Inbred	Assiniboine	USA
MR03	Landrace Inbred	Bolita	Mexico
MR04	Landrace Inbred	Canilla	Venezuela
MR05	Landrace Inbred	Cateto	Bolivia
MR06	Landrace Inbred	Chapalote	Mexico
MR07	Landrace Inbred	Comiteco	N/A
MR08	Landrace Inbred	Costeno	Venezuela
MR09	Landrace Inbred	Cravo Riogranense	Brazil
MR10	Landrace Inbred	Cristalino Norteno	Chile
MR11	Landrace Inbred	Cuban Flint	Cuba
MR12	Landrace Inbred	Havasupai	USA
MR13	Landrace Inbred	Hickory King	USA
MR14	Landrace Inbred	Longfellow Flint	Canada
MR15	Landrace Inbred	Palomero De Jalisco	Mexico
MR16	Landrace Inbred	Pepetilla	Mexico
MR17	Landrace Inbred	Pisankalla	Bolivia
MR18	Landrace Inbred	Reventador	Mexico
MR19	Landrace Inbred	Santa Domingo	USA
MR20	Landrace Inbred	Shoe Peg	USA
MR21	Landrace Inbred	Tabloncillo	Mexico
MR22	Landrace Inbred	Tuxpeno	Mexico
MR23	Landrace Inbred	Zapalote Chico	Mexico
MR24	Landrace Inbred	Chullpi	Peru
MR25	Landrace Inbred	Pororo	Bolivia
MR26	Landrace Inbred	Pollo	Colombia
MR28	Landrace Inbred	N/A	N/A
TIL01	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL03	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL04	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL05	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL06	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL07	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL08	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL09	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL10	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL11	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL12	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL14	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL15	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL16	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL17	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL18	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
TIL25	Teosinte Inbred	*Zea mays* ssp. *parvigulmis*	Mexico
outgroup	Open pollinated	*Zea diploperennis*	Mexico

Supplemental Table B.11: Distribution of Specific Lines used for Diversity Sequencing for each marker.

Each box with "X" means that we were able to sequence this region flanked by the markers with the specific line of interest.

Lines	55-56	42-44	19-145	73-97	28-30	158-160	148-163
MR01	X		X	X	X	X	X
MR02	X	X	X	X		X	
MR03	X	X	X	X		X	
MR04					X	X	X
MR05				X	X	X	X
MR06	X	X		X	X	X	X
MR07	X	X	X	X		X	X
MR08	X	X		X	X	X	
MR09	X	X		X		X	X
MR10	X	X		X		X	X
MR11	X	X	X	X	X	X	X
MR12	X	X		X	X	X	
MR13	X	X	X	X	X		X
MR14	X		X	X	X	X	
MR15					X	X	
MR16	X	X		X	X		X
MR17		X	X		X	X	X
MR18		X		X	X	X	
MR19	X	X	X	X	X	X	X
MR20	X	X		X	X	X	X
MR21		X	X	X	X	X	X
MR22	X	X		X	X	X	X
MR23	X	X	X	X	X	X	X
MR24	X			X	X	X	X
MR25		X	X			X	X
MR26	X	X	X	X	X		X
MR28	X				X		X
TIL01	X		X	X	X	X	X
TIL03	X	X		X	X	X	X
TIL04	X		X	X	X	X	X
TIL05	X		X		X		X
TIL06	X	X	X	X	X		X
TIL07	X	X	X	X	X		X
TIL08	X	X	X	X	X	X	
TIL09	X	X	X	X	X		
TIL10	X		X	X	X	X	X
TIL11	X		X	X	X	X	
TIL12	X		X	X	X	X	X
TIL14	X	X	X	X	X		
TIL15	X		X	X	X		
TIL16	X	X	X	X	X	X	
TIL17	X	X	X	X	X		X
TIL18	X		X	X		X	
TIL25	X		X	X		X	X
outgroup	X	X	X		X	X	X

Supplemental Table B.12: Results of F2 population analyses to examine epistasis and gene action between DIS1.1 and DIS5.1.

The p-values were calculated by in SAS.

Summary	p-value	a	d
DIS1.1	$8.318*10^{-7}$	0.042728	0.439563
DIS5.1	$9.476*10^{-6}$	0.25627	0.20244
DIS1.1 * DIS5.1	0.07359		

R^2= 0.122

Adjusted R^2 model is 0.106 ($p=1.284*10^{-9}$)

Supplemental Table B.13: Results for RIL Population Analyses examining epistasis between DIS1.1 and DIS5.1.

The p-values were calculated in SAS.

Cupule Length	p-value	a
DIS1.1	<.0001	-0.002040
DIS5.1	<.0001	0.197499
DIS1.1 * DIS5.1	0.1997	
X(block)	0.0067	
Y(block)	0.0005	

R^2= 0.14694445

Ear Diameter	p-value	a
DIS1.1	<.0001	0.2499635
DIS5.1	<.0001	0.2168197
DIS1.1*DIS5.1	0.0716	
X(block)	0.0938	
Y(block)	<.0001	

R^2= 0.16865706

Ear Length	p-value	a
DIS1.1	<.0001	-0.6022425
DIS5.1	0.0594	-0.1583788
DIS1.1 * DIS5.1	<.0001	
X(block)	0.0045	
Y(block)	0.0010	

R^2= 0.206824

Kernel Row Number	p-value	a
DIS1.1	<.0001	-0.2593057
DIS5.1	<.0001	-0.2175594
DIS1.1 * DIS5.1	0.0007	
X(block)	0.2378	
Y(block)	0.2897	

R^2= 0.0975329

Shattering	p-value	a
DIS1.1	0.1564	-0.002040
DIS5.1	<.0001	0.18500435
DIS1.1 * DIS5.1	<.0001	
X(block)	<.0001	
Y(block)	0.1584	

R^2= 0.16364599

Supplemental Table B.14: Structural variants of *yab6*.

^ represents individuals with a B73-type insertion on chromosome one's *ZmYAB2.1*. * represents individuals with a Mo-17 insertion on chromosome one's *ZmYAB2.1*. Modified from Lin *et al.* (2012).

Strain	B73-type insertion	Phenotype (% Nondisarticulating fruitcases)
TIL01	no	0
TIL03^	no	0
TIL04	no	0
TIL05	no	0
TIL06	no	0
TIL07^	no	0
TIL08	no	0
TIL09^	no	22
TIL10^	no	0
TIL11	no	0
TIL12	no	0
TIL14^	no	18
TIL15	no	0
TIL16^	no	0
TIL17^	no	6
TIL25	no	0
Mo17*	no	100
P39	yes	100
Ky21*	yes	100
M37W	yes	100
CML247^	no	100
Ki11^	no	100
B73^	yes	100
CML103^	yes	100
CML277^	yes	100
HP301^	yes	100
Oh7B^	yes	100
Tx303^	yes	100
CML322^	?	100
MO18W^	yes	100
Tzi8^	yes	100
MS71^	yes	100
Oh43^	yes	100
NC350^	no	100
CML69^	no	100
CML333^	no	100
B97^	no	100
CML52^	yes	100
Il14H^	?	100
CML228^	yes	100
M162W^*	no	100
KI3^	yes	100
NC358^	yes	100
W22	no	100

Supplemental Table B.15: Allele specific expression of *yab6* with five or more reads from Lemmon *et al.* (2014).

F1 Hybrid	Percentage of Reads Belonging to Maize (%)	Total Number of Reads
B73 x TIL01	14	51
B73 x TIL11	20	71
CML103 x TIL11	22	41
Ki3 x TIL11	32	44
W22 x TIL01	21	78
W22 x TIL11	16	61

Appendix C:

Exploring the Impacts of Peer Leader Mentoring on Freshman STEM Students

C.1: Peer Leader Support (PLS) Chat List

Week	Topics
1	Note Taking Strategies
2	No talk– Librarians visit
3	Campus Involvement
4	DARS Report (Academic Scheduling Tool)
5	Stress Management
6	No talk-Library Resources Lesson
7	Research at UW-Madison
8	Study Techniques and University Health Resources
9	No talk- visiting talk on cover letters/resumes
10	Certificates (minors)
11	How to take advantage of your summer
12	Food
13	Finals study tips
14	No talk- Poster Fair

C.2: Survey

1. When I am with my Peer Leader, I feel free to be who I am.
2. When I am with my Peer Leader, I feel like a competent person.
3. When I am with my Peer Leader, I feel cared about.
4. When I am with my Peer Leader, I often feel inadequate or incompetent.
5. When I am with my Peer Leader, I have a say in what happens and I can voice my own opinion.
6. When I am with my Peer Leader, I often feel a lot of distance in our relationship.
7. When I am with my Peer Leader, I feel very capable and effective.
8. When I am with my Peer Leader, I feel a lot of closeness and friendship.
9. When I am with my Peer Leader, I feel controlled and pressured to be certain ways.
10. My Peer Leader and I enjoy a high quality relationship.
11. Please explain your answer to the question above. What is your relationship with your Peer Leader like? What makes or doesn't make it high quality?
12. I am effectively using my Peer Leader.
13. Please explain your answer to the question above. How are you using or not using your Peer Leader effectively?
14. I am benefitting from the Peer Leader relationship.
15. I would recommend the Peer Leader mentoring program to others.
16. From working with my Peer Leader, I am gaining a better sense of how to become more involved at UW-Madison.
17. From working with my Peer Leader, I am gaining new skills.
18. Please explain your answer to the question above. What are the new skills your Peer Leader is helping you gain?
19. My Peer leader is well qualified to be a mentor.
20. What Peer Leader have you interacted with the most?
21. How often do you interact with a Peer Leader?
22. Have you communicated with a Peer Leader outside of class?
    a. If you answered "Yes" to the question above, on average how often do you communicate with a Peer Leader outside of class?
    b. If you answered "Yes" to the question above, how do you communicate?
23. How old are you?
24. What is your gender? Select the one you most identify with.
25. What is your race/ethnicity? Select all that apply.
26. Currently, what is your level of interest in biology?
27. What are the educational level(s) of your guardian(s)/parent(s)?
28. I am enrolled in…

C.3: Student Focus Group Script
Student Focus Group about Peer Leaders

Introduction: Purpose of the Focus Group

My name is Jerry Whitmore Jr and I am the Director of First Year and Retention Programs in WISCIENCE. You were invited to this focus group because you participated in the Exploring Biology course this past fall. Our discussion should take about an hour. We are interested in knowing more about your perspectives and ideas about the success of Peer Leaders in the course. WISCIENCE staff, like me, train the peer mentors and then work with WISCIENCE Teaching Fellows to facilitate discussions in the Exploring Biology course within WISCIENCE (the Wisconsin Institute for Science Education and Community Engagement).

The information you give us today will be shared with other study team members to help improve the peer leader experience. We will NOT use your name. We hope that you can feel comfortable to speak freely about your experiences and we ask that everyone be respectful of what everyone else has to say.

Please go around, introduce yourselves, and think about your experience with peer leaders in the course and come up with one word to describe peer leaders.

1. Would you ever consider being a peer leader yourself?
    - Did you know other people involved in peer mentoring/interested in peer mentoring?
    - What are the reasons would you consider being a peer leader?
    - Would you be interested in doing it for this course (Exploring Biology) or another course? Why?

2. A lot of students come to UW-Madison and choose to participate in many Peer Mentor programs.
    - Did you have any preconceived notions about a peer leader?
        - Were they met? Were you disappointed?

3. Do you feel like your peer leader had adequate training?
    - Do you feel that they had good communication between them and the instructors?
    - Do you feel that they had good communication with the students?

4. Can you describe the first time you interacted with a peer leader?
    - Describe what you got out of the mentoring experience? Was is a positive experience or not so much?
    - What could have made that first experience better?

- What are the most valuable experiences you have had with a peer leader? Explain.
- What are the least valuable experiences you have had with a peer leader? Explain.

5. Please describe what you got out of the Mentor experience?
    - What skills do you feel you have developed/grown within the Mentor experience?
    - What skills do you think peer leaders could have helped with, but didn't?

6. Would you say peer mentoring was a positive experience, negative experience or neutral experience?
    - For those of you have had a positive experience, what helped you come to this conclusion? Did someone help you in deciding your experience
    - For those of you who had a negative experience, what helped you come to this conclusion?
    - How do you think other students felt about peer leaders in the course?

Digging Deeper

How were you supported in your experience with peer leaders?

Overall, how do you feel you have changed within your first semester in college?

What do you like most about having a Peer leader?

What could be done to improve the peer leader program?

Do you feel that peer leaders made a difference?

Were you given opportunities to provide feedback to the fellows or to faculty about your experience with peer mentors?

Closing

Ok, to close here, let's go around once and hear from everyone on a "final thoughts" question: If there was one (or two) thing (s) that the peer leaders could have done differently or better to help students have a great experience, what would that be?

Thank you very for your time. We learned so much from this conversation and will pass on the information to people who can make changes to improve students' experiences.

C.4: Peer Leader Focus Group Script
Peer Mentor Focus Group

Introduction: Purpose of the Focus Group

My name is Jerry Whitmore Jr and I am the Director of First Year and Retention Programs in WISCIENCE. You were invited to this focus group because you participated in the Exploring Biology course as a peer mentor. Today I would like to learn about your experience as a peer mentor. Our discussion should take about an hour. We are interested in knowing more about your perspectives and ideas. WISCIENCE staff, like me, train the peer mentors and then work with WISCIENCE Teaching Fellows to facilitate discussions in the Exploring Biology course within WISCIENCE (the Wisconsin Institute for Science Education and Community Engagement).

The information you give us today will be shared with only study team members to help improve the peer leader experience. We will NOT use your name. We hope that you can feel comfortable to speak freely about your experiences and we ask that everyone be respectful of what everyone else has to say.

Please go around, introduce yourselves, and think about your experience as a Peer Mentor or Leader and come up with one word to describe why you chose to be a part of the program.

7. How did you get involved with peer mentoring with WISCIENCE?
    - Did you know other people involved?
    - Did you know people who chose to do peer mentoring in another department or program area? Why?
    - Did you think about doing peer mentoring somewhere else?
    - Returners- what made you decide to return for a second year?

8. A lot of students come to UW-Madison and choose to participate in many Peer Mentor programs.
    - Why did you choose to become a peer mentor for Exploring Biology? Experience/resume/develop skills? Credit? Like helping others?

9. After you completed the peer mentor training, do you feel like you were ready to succeed as a leader? Why or why not?
    - What was your training experience like? Did you receive enough training? Too little training? Quality of training? Did you feel prepared?
    - What did you find to be the most useful and least useful with training?
    - Did you do any peer mentoring outside of Exploring Biology? Did students continue to reach out to you after being introduced in Exploring Biology?

10. Can you describe the first time you acted in your role as Peer Mentor?

- Describe what you got out of the Mentor experience? Was is a positive experience or not so much?
- What could have made that first experience better?
- What are the most valuable experiences you have had within the Mentor role?
- What are the least valuable experiences you have had within the Mentor role?

11. Please describe what you got out of the Mentor experience?
    - Did you know what skills you wanted to learn before the experience? If so, how did that impact your decision to be a peer mentor?
    - What skills do you feel you have developed/grown within the Mentor experience?

12. Would you say peer mentoring was a positive experience, negative experience or neutral experience?
    - For those of you have had a positive experience, what helped you come to this conclusion? Did someone help you in deciding your experience
    - For those of you who had a negative experience, what helped

Digging Deeper

How were you supported in your experience as a Mentor? collaborated

Overall, how do you feel you have changed within your Peer Mentor role?

What do you love the most about being a Peer Mentor?

What could be done to improve the peer mentor program?

Do you feel that you made a difference? How does it feel when you're able to help an underclassman?

Were you given opportunities to provide feedback to the fellows or to faculty about your experience as a peer mentor?

Closing

Ok, to close here, let's go around once and hear from everyone on a "final thoughts" question: If there was one (or two) thing (s) that the peer training or WISCIENCE staff could have done differently or better to help students become better mentors and/or have a great experience, what would that be?

Thank you very for your time. We learned so much from this conversation and will pass on the information to people who can make changes to improve students' experiences.

References

Allen T. D., McManus S. E., Russell J. E. A., 1999 Newcomer Socialization and Stress: Formal Peer Relationships as a Source of Support. J. Vocat. Behav. **54**: 453–470.

Alpert K. B., Grandillo S., Tanksley S. D., 1995 fw 2.2:a major QTL controlling fruit weight is common to both red- and green-fruited tomato species. Theor. Appl. Genet. **91**: 994–1000.

Astin A. W., 1993 *What Matters in College? Four Critical Years Revisited*. Jossey-Bass.

Astin A. W., 1999 Student involvement: A developmental theory for higher education. J. Coll. Stud. Dev. **40**: 518–529.

Beadle G. W., 1932 Studies of Euchlaena and its Hybrids with Zea. Z. Indukt. Abstamm. Vererbungsl. **62**: 291–304.

Bowling B., Doyle M., Taylor J., Antes A., 2015 Professionalizing the Role of Peer Leaders in STEM. J. STEM Educ. Innov. Res. **16**: 30–39.

Bowman J., Smyth D., 1999 CRABS CLAW, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel portein with zinc finger and helix-loop-helix domains. Development **126**: 2387–2396.

Bowman J. L., 2000 The YABBY gene family and abaxial cell fate. Curr. Opin. Plant Biol. **3**: 17–22.

Briggs W. H., McMullen M. D., Gaut B. S., Doebley J., 2007 Linkage mapping of domestication loci in a large maize teosinte backcross resource. Genetics **177**: 1915–28.

Brissette I., Scheier M. F., Carver C. S., 2002 The Role of Optimism in Social Network Development, Coping, and Psychological Adjustment During a Life Transition. J. Pers. Soc. Psychol. **82**: 102–111.

Broman K. W., Wu H., Sen S., Churchill G. A., 2003 R/qtl: QTL mapping in experimental crosses. Bioinformatics **19**: 889–90.

Broman K. W., Sen S., 2009 *A guide to QTL mapping with R/qtl*. Springer.

Burnham C., Clark E., 1954 Gametophyte factor studies-- specificity tests. Maize Genet. Coop News Lett. **28**: 60.

Campbell T. A., Campbell D. E., 1997 Faculty/Student Mentor Program: Effects on Academic Performance and Retention. Res. High. Educ. **38**: 727–742.

Calderón C. I., Yandell B. S., Doebley J. F., 2016 Fine Mapping of a QTL Associated with Kernel Row Number on Chromosome 1 of Maize. PLoS One **11**: e0150276.

Carlone H. B., Johnson A., 2007 Understanding the Science Experiences of Successful Women of Color: Science Identity as an Analytic Lens. J. Res. Sci. Teach. J Res Sci Teach **44**: 1187–1218.

Chesler N. C., Chesler M. A., 2002 Gender-Informed Mentoring Strategies for Women Engineering Scholars: On Establishing a Caring Community. J. Eng. Educ. **91**: 49–55.

Chuck G., Meeley R., Hake S., 2008 Floral meristem initiation and meristem cell fate are regulated by the maize AP2 genes ids1 and sid1. Development **135**: 3013–9.

Crisp G., Cruz I., 2009 Mentoring College Students: A Critical Review of the Literature Between 1990 and 2007. Res. High. Educ. **50**: 525–545.

Cuseo J. B., 1991 *The freshman orientation seminar : a research-based rationale for its value, delivery, and conten*. University of South Carolina, Columbia.

Darwin C., 1859 *On the Origin of Species by Means of Natural Selection*. John Murray, London.

Devos K. M., Gale M. D., 2000 Genome Relationships: The Grass Model in Current Research. Plant Cell **12**: 637–646.

Doebley J., Stec a., 1991 Genetic-Analysis of the Morphological Differences between Maize and Teosinte. Genetics **129**: 285–295.

Doebley J., Stec A., 1993 Inheritance of the morphological differences between maize and teosinte: comparison of results for two F2 populations. Genetics **134**: 559–70.

Doebley J., Stec A., Gustus C., 1995 teosinte branched1 and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics **141**: 333–46.

Doebley J., 2004 The genetics of maize evolution. Annu. Rev. Genet. **38**: 37–59.

Elshire R. J., Glaubitz J. C., Sun Q., Poland J. A., Kawamoto K., Buckler E. S., Mitchell S. E., 2011 A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One **6**: e19379.

Ender S., Kay K., 2001 *Peer Leadersip Programs: A Rationale and Review of the Literature.* (S Hamid, Ed.). National Resource Center for the First-Year Experience and Students in Transition, University of South Carolina, Columbia.

Fehr W. R., Fehr E. L., Jessen H. J., 1987 *Principles of Cultivar Development: Theory and Technique*. Macmilan Publishing Company, New York.

Finet C., Floyd S. K., Conway S. J., Zhong B., Scutt C. P., Bowman J. L., 2016 Evolution of the YABBY gene family in seed plants. Evol. Dev. **18**: 116–126.

Girves J. E., Zepeda Y., Gwathmey J. K., 2005 Mentoring in a Post-Affirmative Action World. J. Soc. Issues **61**: 449–479.

Grade Gap Report for Undergraduate Courses in 2016-2017, 2017

Hake S., Ross-Ibarra J., 2015 Genetic, evolutionary and plant breeding insights from the domestication of maize. Elife **4**: e05861.

Haley C. S., Knott S. A., 1992 A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity (Edinb). **69**: 315–324.

Hermisson J., Pennings P. S., 2005 Soft Sweeps: Molecular Population Genetics of Adaptation From Standing Genetic Variation. Genetics **169**: 2335–2352.

Hudson R. R., Kreitman M., Aguadé M., 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116**: 153–9.

Hufford M. B., Xu X., Heerwaarden J. van, Pyhäjärvi T., Chia J.-M., Cartwright R. A., Elshire R. J., Glaubitz J. C., Guill K. E., Kaeppler S. M., Lai J., Morrell P. L., Shannon L. M., Song C., Springer N. M., Swanson-Wagner R. A., Tiffin P., Wang J., Zhang G., Doebley J., McMullen M. D., Ware D., Buckler E. S., Yang S., Ross-Ibarra J., 2012 Comparative population genomics of maize domestication and improvement. Nat. Genet. **44**: 808–811.

Hung H.-Y., Shannon L. M., Tian F., Bradbury P. J., Chen C., Flint-Garcia S. A., McMullen M. D., Ware D., Buckler E. S., Doebley J. F., Holland J. B., 2012 ZmCCT and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. Proc. Natl. Acad. Sci. U. S. A. **109**: E1913-21.

Iltis H. H., 1983 From Teosinte to Maize: The Catastrophic Sexual Transmutation. Science (80-.). **222**: 886–894

Information Memorandum on Retention and Graduation, Madison.

Innan H., Kim Y., 2004 Pattern of polymorphism after strong artificial selection in a domestication event. Proc. Natl. Acad. Sci. U. S. A. **101**: 10667–72.

Ishler J., Schreiber S., 2002 First-Year Female Students: Perceptions of Friendship. J. First-Year Exp. Students Transit. **14**: 89–104.

Jacobi M., 1991 Mentoring and Undergraduate Academic Success: A Literature Review. Rev. Educ. Res. **61**: 505–532.

Jang S., Hur J., Kim S.-J., Han M.-J., Kim S.-R., An G., 2004 Ecotpic expression of OSYAB1 causes extra stamens and carpels in rice. Plant Mol. Biol. **56**: 133–143.

Jiao Y., Peluso P., Shi J., Liang T., Stitzer M. C., Wang B., Campbell M. S., Stein J. C., Wei X., Chin

C.-S., Guill K., Regulski M., Kumari S., Olson A., Gent J., Schneider K. L., Wolfgruber T. K., May M. R., Springer N. M., Antoniou E., McCombie W. R., Presting G. G., McMullen M., Ross-Ibarra J., Dawe R. K., Hastie A., Rank D. R., Ware D., 2017 Improved maize reference genome with single-molecule technologies. Nature **546**: 524.

Johanson U., West J., Lister C., Michaels S., Amasino R., Dean C., 2000 Molecular Analysis of FRIGIDA, a Major Determinant of Natural Variation in Arabidopsis Flowering Time. Science (80-.). **290**: 344–347.

Jones F. C., Grabherr M. G., Chan Y. F., Russell P., Mauceli E., Johnson J., Swofford R., Pirun M., Zody M. C., White S., Birney E., Searle S., Schmutz J., Grimwood J., Dickson M. C., Myers R. M., Miller C. T., Summers B. R., Knecht A. K., Brady S. D., Zhang H., Pollen A. A., Howes T., Amemiya C., Baldwin J., Bloom T., Jaffe D. B., Nicol R., Wilkinson J., Lander E. S., Palma F. Di, Lindblad-Toh K., Kingsley D. M., 2012 The genomic basis of adaptive evolution in threespine sticklebacks. Nature **484**: 55–61.

Kellogg E. A., 1997 Plant evolution: The dominance of maize. Curr. Biol. **7**: R411–R413.

Kermicle J. L., Allen J., 1990 Cross Incompatibility between maize and teosinte. Maydica **35**: 399–408.

Kermicle J. L., Evans M. M. S., 2010 The Zea mays Sexual Compatibility Gene ga2: Naturally Occurring Alleles, Their Distribution, and Role in Reproductive Isolation. J. Hered. **101**: 737–749.

Konishi S., Izawa T., Lin S. Y., Ebana K., Fukuta Y., Sasaki T., Yano M., 2006 An SNP caused loss of seed shattering during rice domestication. Science **312**: 1392–6.

Lander E. S., Botstein D., 1989 Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121**: 185–199.

Lemmon Z. H., Doebley J. F., 2014 Genetic Dissection of a Genomic Region with Pleiotropic Effects on Domestication Traits in Maize Reveals Multiple Linked QTL. Genetics **198**: 345–53.

Lemmon Z. H., 2014 The complex inheritance of maize domestication traits and gene expression.

Lemmon Z. H., Bukowski R., Sun Q., Doebley J. F., 2014 The Role of cis Regulatory Evolution in Maize Domestication. PLoS Genet. **10**: e1004745.

Li C., Zhou A., Sang T., 2006 Rice Domestication by Reducing Shattering. Science (80-.). **311**: 1936–1939.

Librado P., Rozas J., 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25**: 1451–2.

Lin Z., Li X., Shannon L. M., Yeh C.-T., Wang M. L., Bai G., Peng Z., Li J., Trick H. N., Clemente T. E., Doebley J. F., Schnable P. S., Tuinstra M. R., Tesso T. T., White F., Yu J., 2012 Parallel domestication of the Shattering1 gene in cereals. Nat. Genet. **44**: 720–725.

Littell R. C., Henry P. R., Ammerman C. B., 1998 Statistical analysis of repeated measures data using SAS procedures. J. Anim. Sci. **76**: 1216–31.

Mackay T. F. C., Stone E. A., Ayroles J. F., 2009 The genetics of quantitative traits: challenges and prospects. Nat. Rev. Genet. **10**: 565–77.

Mayer R. E., Christoffersen R., Fiorella L., Mayer R. E., Fiorella L., 2017 Enhancing Undergraduate Success in Biology Through the Biomentors Program Enhancing Undergraduate Success in Biology through the Biomentors Program The Biomentors program is a relatively low-cost program because all participants are volunteers , and the p. **79**: 23–27.

Newton F., Ender S., 2010 *Students helping students: A guide for peer educators on college campuses*. Jossey-Bass, San Francisco.

Nora A., 1987 Determinants of Retention among Chicano College Students: A Structural Model. Source Res. High. Educ. **2610446**: 31–59.

Olsen K. M., Wendel J. F., 2013 Crop plants as models for understanding plant adaptation and diversification. Front. Plant Sci. **4**: 290.

Paterson A. H., Lin Y.-R., Li Z., Schertz K. F., Doebley J. F., Pinson S. R., Liu S.-C., Stansel J. W., Irvine J. E., 1995 Convergent Domestication of Cereal Crops by Independent Mutations at Corresponding Genetic Loci. Science (80-.). **269**: 1714–1718.

Piperno D. R., Ranere A. J., Holst I., Iriarte J., Dickau R., 2009 Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. Proc. Natl. Acad. Sci. U. S. A. **106**: 5019–24.

Protas M. E., Hersey C., Kochanek D., Zhou Y., Wilkens H., Jeffery W. R., Zon L. I., Borowsky R., Tabin C. J., 2006 Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. Nat. Genet. **38**: 107–11.

Purugganan M. D., Fuller D. Q., 2009 The nature of selection during plant domestication. Nature **457**: 843–8.

Quitadamo I. J., Brahler C. J., Crouch G. J., 2009 Peer-led Team Learning: A prospective method for increasing critical thinking in undergraduate science courses. Sci. Educ. **18**: 29–39.

Sawa S., Ito T., Shimura Y., Okada K., 1999 FILAMENTOUS FLOWER Controls the Formations and Development of Arabidopsis Inflorescences and Floral Meristems. Plant Cell **11**: 69–86.

Sax K., 1923 The Association of Size Differences with Seed-Coat Pattern and Pigmentation in PHASEOLUS VULGARIS. Genetics **8**: 552–60.

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., Rock S. M., Belter E., Du F., Kim K., Abbott R. M., Cotton M., Levy A., Marchetto P., Ochoa K., Jackson S. M., Gillam B., Chen W., Yan L., Higginbotham J., Cardenas M., Waligorski J., Applebaum E., Phelps L., Falcone J., Kanchi K., Thane T., Scimone A., Thane N., Henke J., Wang T., Ruppert J., Shah N., Rotter K., Hodges J., Ingenthron E., Cordes M., Kohlberg S., Sgro J., Delgado B., Mead K., Chinwalla A., Leonard S., Crouse K., Collura K., Kudrna D., Currie J., He R., Angelova A., Rajasekar S., Mueller T., Lomeli R., Scara G., Ko A., Delaney K., Wissotski M., Lopez G., Campos D., Braidotti M., Ashley E., Golser W., Kim H., Lee S., Lin J., Dujmic Z., Kim W., Talag J., Zuccolo A., Fan C., Sebastian A., Kramer M., Spiegel L., Nascimento L., Zutavern T., Miller B., Ambroise C., Muller S., Spooner W., Narechania A., Ren L., Wei S., Kumari S., Faga B., Levy M. J., McMahan L., Buren P. Van, Vaughn M. W., Ying K., Yeh C.-T., Emrich S. J., Jia Y., Kalyanaraman A., Hsia A.-P., Barbazuk W. B., Baucom R. S., Brutnell T. P., Carpita N. C., Chaparro C., Chia J.-M., Deragon J.-M., Estill J. C., Fu Y., Jeddeloh J. A., Han Y., Lee H., Li P., Lisch D. R., Liu S., Liu Z., Nagel D. H., McCann M. C., SanMiguel P., Myers A. M., Nettleton D., Nguyen J., Penning B. W., Ponnala L., Schneider K. L., Schwartz D. C., Sharma A., Soderlund C., Springer N. M., Sun Q., Wang H., Waterman M., Westerman R., Wolfgruber T. K., Yang L., Yu Y., Zhang L., Zhou S., Zhu Q., Bennetzen J. L., Dawe R. K., Jiang J., Jiang N., Presting G. G., Wessler S. R., Aluru S., Martienssen R. A., Clifton S. W., McCombie W. R., Wing R. A., Wilson R. K., 2009 The B73 maize genome: complexity, diversity, and dynamics. Science **326**: 1112–5.

Shannon L. M., 2013 The Genetic Architecture of Maize Domestication and Range Expansion.

Shook J. L., Keup J. R., 2012 The benefits of peer leader programs: An overview from the literature. New Dir. High. Educ. **2012**: 5–16.

Siegfried K. R., Eshed Y., Baum S. F., Otsuga D., Drews G. N., Bowman J. L., 1999 Members of the YABBY gene family specify abaxial cell fate in Arabidopsis. Development **126**: 4117–4128.

Simons K. J., Fellers J. P., Trick H. N., Zhang Z., Tai Y.-S., Gill B. S., Faris J. D., 2006 Molecular characterization of the major wheat domestication gene Q. Genetics **172**: 547–55.

Snyder J. J., Elijah Carter B., Wiles J. R., 2015 Implementation of the peer-led team-learning instructional model as a stopgap measure improves student achievement for students opting out of laboratory. CBE Life Sci. Educ. **14**: 1–6.

Snyder J. J., Sloane J. D., Dunk R. D. P., Wiles J. R., 2016 Peer-Led Team Learning Helps Minority Students Succeed. PLoS Biol. **14**: 1–7.

Springer N. M., Anderson S. N., Andorf C. M., Ahern K. R., Bai F., Barad O., Barbazuk W. B., Bass

H. W., Baruch K., Ben-Zvi G., Buckler E. S., Bukowski R., Campbell M. S., Cannon E. K. S., Chomet P., Dawe R. K., Davenport R., Dooner H. K., Du L. H., Du C., Easterling K. A., Gault C., Guan J.-C., Hunter C. T., Jander G., Jiao Y., Koch K. E., Kol G., Köllner T. G., Kudo T., Li Q., Lu F., Mayfield-Jones D., Mei W., McCarty D. R., Noshay J. M., Portwood J. L., Ronen G., Settles A. M., Shem-Tov D., Shi J., Soifer I., Stein J. C., Stitzer M. C., Suzuki M., Vera D. L., Vollbrecht E., Vrebalov J. T., Ware D., Wei S., Wimalanathan K., Woodhouse M. R., Xiong W., Brutnell T. P., 2018 The maize W22 genome provides a foundation for functional genomics and transposon biology. Nat. Genet.: 1.

Stout M., McDaniel A., 2006 Benefits to supplemental instruction leaders. In: Stone M, Jacobs G (Eds.), *Supplemental Instruction: new visions for empowering student learning*, San Francisco, pp. 55–62.

Studer A., Zhao Q., Ross-Ibarra J., Doebley J., 2011 Identification of a functional transposon insertion in the maize domestication gene tb1. Nat. Genet. **43**: 1160–3.

Tajima F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123**: 585–95.

Taylor-Powell E., Renner M., 2003 *Analyzing Qualitative Data*. Madison.

Tenaillon M. I., U'Ren J., Tenaillon O., Gaut B. S., 2004 Selection Versus Demography: A Multilocus Investigation of the Domestication Process in Maize. Mol. Biol. Evol. **21**: 1214–1225.

Terenzini P. T., Pascarella E. T., Blimling G. S., 1996 Students' out-of-class experiences and their influence on learning and cognitive development: A literature review. J. Coll. Stud. Dev. **37**: 149–162.

Thomas S. L., 2000 Ties That Bind: A Social Network Approach to Understanding Student Integration and Persistence. J. Higher Educ. **71**: 591.

Tinto V., 1993 *Leaving College: Rethinking the Causes and Cures of Student Attrition*. University of Chicago Press, Chicago.

Tinto V., 1998 Colleges as communities: Exploring the educational character of student persistence. J. Higher Educ. **68**.

Toriba T., Harada K., Takamura A., Nakamura H., Ichikawa H., Suzaki T., Hirano H.-Y., 2007 Molecular characterization the YABBY gene family in Oryza sativa and expression analysis of OsYABBY1. Mol Genet Genomics **277**: 457–468.

UW-Madison Strategic Diversity Update, 2013

Villanueva J. M., Broadhvest J., Hauser B. A., Meister R. J., Schneitz K., Gasser C. S., 1999 INNER NO OUTER regulates abaxial-adaxial patterning in Arabidopsis ovules. Genes Dev. **13**:

3160–3169.

Wamser C. C., 2006 Peer-Led Team Learning in Organic Chemistry : Effects on Student Performance , Success , and Persistence in the Course. J. Chem. Educ. **83**: 1562–1566.

Wang H., Nussbaum-Wagler T., Li B., Zhao Q., Vigouroux Y., Faller M., Bomblies K., Lukens L., Doebley J. F., 2005 The origin of the naked grains of maize. Nature **436**: 714–719.

Wills D. M., Whipple C. J., Takuno S., Kursel L. E., Shannon L. M., Ross-Ibarra J., Doebley J. F., 2013 From many, one: genetic control of prolificacy during maize domestication. PLoS Genet. **9**: e1003604.

Woodward A. E., Weiner M., Gosser D., 1993 Projects supported by the NSF Division of Undergraduate Education Problem Solving Workshops in General Chemistry. J. Chem. Educ. **70**: 651.

Wright S. I., Gaut B. S., 2005 Molecular Population Genetics and the Search for Adaptive Evolution in Plants. Mol. Biol. Evol. **22**: 506–519.

Yang C. J., Kursel L. E., Studer A. J., Bartlett M. E., Whipple C., Doebley J. F., 2016 A gene for genetic background in Zea mays: fine-mapping enhancer of teosinte branched1.2 (etb1.2) to a YABBY class transcription factor. bioRxiv.

Zaniewski A. M., Reinholz D., 2016 Increasing STEM success: a near-peer mentoring program in the physical sciences. Int. J. STEM Educ. **3**: 14.

Zhao Q., Weber A. L., McMullen M. D., Guill K., Doebley J., 2011 MADS-box genes of maize: frequent targets of selection during domestication. Genet. Res. (Camb). **93**: 65–75.