Fine Mapping of a Major-Effect QTL Controlling Rhizomatous Growth in Perennial Wildrye Hybrids

Jenny L. Clements
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/gradreports

Recommended Citation
https://digitalcommons.usu.edu/gradreports/809

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Plan B and other Reports by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.
FINE MAPPING OF A MAJOR-EFFECT QTL CONTROLLING RHIZOMATOUS GROWTH IN PERENNIAL WILDRIYE HYBRIDS

by

Jenny L. Clements

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

MASTER OF SCIENCE

in

Statistics

Approved:

Dr. Adele Cutler
Major Professor

Dr. Steve Larson
Committee Member

Dr. John Stevens
Committee Member

UTAH STATE UNIVERSITY
Logan, Utah

2016
I dedicate this project to my father, Russell C. Thompson. Wish you were here.
ABSTRACT

Fine Mapping of a Major-Effect QTL Controlling Rhizomatous Growth in Perennial Wildrye Hybrids

by

Jenny L. Clements

Utah State University, 2016

Quantitative trait loci (QTL) analysis is the science of discovering genes or chromosome regions that influence the expression of quantitative traits. Statistical methods for doing this have evolved from single marker analysis of variance (ANOVA) to more complex approaches such as interval mapping and multiple QTL mapping. The goal of this project was to refine a previously identified QTL chromosome region responsible for rhizome growth in grass hybrids developed from two species of the Triticeae tribe by categorically progeny testing seven recombinant genotypes containing cross overs in a major-effect rhizome QTL. In addition, the analysis addressed physical concerns with the field plot, assessing statistical power, multiple comparison testing and linkage segregation assessment.
ACKNOWLEDGEMENTS

I express deep gratitude to my advisor, Dr. Adele Cutler. The guidance and encouragement that you have given me throughout this entire process has been invaluable. I would also like to thank Dr. Steve Larson for sharing his research and for the many hours spent patiently explaining the science behind it, and Dr. John Stevens for the direction and advice given to improve the analysis done. Thank you to Dr. Jack Staub, USDA Research Leader at the Forage Range and Research Lab, for generously setting up the project and dedicating resources to it.

Special thanks to my husband, Travis Clements, for his patient love and support during this process. To my mother, Janet Thompson, for the offers to help out and randomly bringing over delicious food. To my best friend, Nicole Roundy, for lending her expertise in the written word.

Finally, thanks to all the amazing instructors and friends in the Mathematics and Statistics Department that supported me throughout my time here at Utah State University.

Jenny Clements
# CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION AND LITERATURE REVIEW</td>
<td>1</td>
</tr>
<tr>
<td>2. QTL ANALYSIS BACKGROUND</td>
<td>2</td>
</tr>
<tr>
<td>3. USDA PERENNIAL WILDRYE RESEARCH</td>
<td>7</td>
</tr>
<tr>
<td>4. FINE TUNING OF RHIZOME QTL ANALYSIS</td>
<td>12</td>
</tr>
<tr>
<td>5. CONCLUSION</td>
<td>32</td>
</tr>
<tr>
<td>6. DISCUSSION OF FUTURE WORK</td>
<td>33</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>34</td>
</tr>
<tr>
<td>COMPUTER CODE</td>
<td>36</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. Descriptive Statistics for Rhizome Circumference in cm 2015 ............................21
2. LS Mean Difference Rhizome Circumference in cm 2015  ..................................22
3. Chi-square Goodness of Fit Test of Marker Segregation  ..................................23
LIST OF FIGURES

1. Field Colormap of Rhizome Circumference ...........................................................15
2. Graphical Genotype of Seven BC2 Parents ..........................................................17
3. Power plots ............................................................................................................31
1. Introduction and Literature Review

Phenotypes in an organism are observable characteristics or traits resulting from the interaction between genes and the environment. Quantitative traits are phenotypes that can be measured and vary among individuals in a population. Some examples in plants would be height/width, biomass and tiller number. A quantitative trait locus (QTL) is a region on a chromosome containing one or more genes responsible for affecting a quantitative trait. The goals of QTL mapping techniques are to discover how many genes are responsible for a quantitative trait and where they are located on the chromosome. Identifying locations of genes that affect a quantitative trait have important applications in areas such as plant breeding programs, transgenic technology and epidemiology (Falconer and Mackay, 1996 p. 356). Because the actual genes for QTLs are unobservable, scientists use statistical methods to build genetic maps and link DNA marker information with phenotypic data to identify regions in the genome of the organism that are associated with the phenotype. There are several approaches to QTL mapping, but the three main approaches needed for reference in this project are: single marker analysis, interval mapping, and multiple QTL mapping.

The data for this project was obtained from the USDA Forage and Range Research Lab (FRRL) in Logan, Utah. The mission of FRRL is to develop improved plant materials and management alternatives for sustainable stewardship of rangelands, pastures, and turf in the Western US. Dr. Steve Larson is a plant geneticist with FRRL whose expertise includes gene discovery research in plants (http://www.ars.usda.gov/). One area of his research involves the identification of
genes that control rhizome growth in perennial wildrye plants (Yun et al., 2014). This project is an extension of that research, with the goal of refining the location of a major effect rhizome gene in grasses from the *Leymus* genus of the Triticeae tribe.

This project references several textbooks and papers. In Falconer & Mackay’s textbook *Quantitative Genetics*, the authors describe the basic ideas behind single marker analysis and experimental design considerations (Falconer and Mackay, 1996). In Lynch’s textbook *Genetics and Analysis of Quantitative Traits* the author addresses QTL mapping for inbred lines, which is the common design in plant genetics research for detecting QTLs (Lynch, 1998). A pivotal paper in QTL analysis is Lander and Botstein’s “Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps,” which describes and compares the traditional single marker approach to QTL mapping with interval mapping using LOD scores (Lander and Botstein, 1989). Good reviews of QTL methods, with their strengths and weaknesses, is provided in papers by Broman (2001), Doerge (2001), Kao (1999) and Miles (2008). Background research preceding the data used in this project is in the paper “Genetic Control of Rhizomes and Genomic Localization of a Major-effect Growth Habit QTL in Perennial Wildrye” (Yun et al., 2014).

2. QTL Analysis Background

2.1 Fundamental requirements

The first step in QTL analysis is to identify or develop full-sib families of an organism that show heritable phenotypic variation with regards to the quantitative
trait being studied. The next step is to develop and identify homozygous polymorphic genetic markers that span the genome, distinguish parental lines, and display predictable segregation ratios among full-sib progeny of the parents (Miles, 2008). True breeding, homozygous parental lines can be crossed to produce the heterogeneous F1 progeny. The two most common approaches used for developing QTL mapping populations are to either cross the F1 progeny to themselves (F2) or back-cross (BC) to one of the parents, commonly referred to as the recurrent or tester parent, to produce the next generation of segregating-hybrid progeny. The F2 progeny have a 1:2:1 expected segregation ratio of 1 individual homozygous for parent 1 allele, 2 individuals heterozygous for both parental allele, and 1 individual homozygous for parent 2 allele; whereas the BC progeny have a 1:1 expected segregation ratio of 1 individual heterozygous for both parental alleles to 1 individual homozygous for the recurrent (tester) parent allele. The next generation progeny is scored for each marker genotype and measured for the quantitative response variable. For the QTL to be detected, both the QTL and the marker must be polymorphic and be in linkage disequilibrium (http://passel.unl.edu/).

2.2 Single Marker Analysis

In single marker analysis the progeny is classified by whether the genetic marker is present or absent. Markers close to a QTL will segregate with it through meiosis so significant differences in the means of the two classes infer a QTL is near the marker (Lander and Botstein, 1989). A basic ANOVA model for this is $z_{tk} = \mu + b_t + e_{tk}$.
where \( z_{ik} \) is the phenotypic value for the kth individual with marker genotype i, \( b_i \) is the marker effect, and \( e_{ik} \) is the residual error (Lynch, 1998, p. 442). Some issues with this approach are underestimation of the QTL due to confounding by recombination frequency, false positives due to independent testing of many markers throughout the genome, the exclusion of missing data. Large numbers of progeny are needed to obtain enough statistical power (Lander and Botstein, 1989).

2.3 Interval Mapping (IM)

Interval mapping builds on the single marker approach by incrementally going through the genome defined by ordered pairs of markers. This approach requires evaluation of genetic marker order, distance, and categorical chromosome grouping (genes on different chromosomes are physically independent). This approach uses the maximum likelihood method to assign log of odds (LOD) scores by evaluating the likelihood under the null hypothesis of no QTL at the testing position to the alternative hypothesis of the QTL at the testing position (Doerge, 2002). Likelihood maps are plots of the LOD score versus the map position (Lynch, 1989, p. 448). For a basic understanding of these LOD scores, start by letting \( \mu_A \) and \( \mu_B \) represent the phenotypic means of the QTL depending on the QTL’s genotype. Then assume that the trait variance \( \sigma \) is the same for each genotype. For a given location \( z \), the probability of the observed data can be denoted as \( \Pr(\text{data}|\text{QTL at } z, \mu_A, \mu_B, \sigma) \). These parameters are unknown so maximum likelihood estimates (MLE) of these parameters are obtained using iterative numerical methods to get these estimates by
interpolating the genotypic information between the pairs of markers (Falconer and Mackay, 1989). The MLE solutions are regarded as being a linear regression problem with missing data where none of the genotypes are known, and only their respective probability distributions are known (Lander and Botstein, 1989). A conceptual form of the LOD score is

\[
\text{LOD}(z) = \log_{10} \left[ \frac{\Pr(\text{data} | \text{QTL at } z, \vec{\theta}_A, \vec{\theta}_B)}{\Pr(\text{data} | \text{no QTL})} \right]
\]

with larger values of LOD considered evidence that a QTL is near (Broman, 2001).

Approximate intervals to estimate a QTL location can be constructed by drawing horizontal lines 1, or more commonly 2, LOD scores below the maximum value or peak on the plot (Lynch, 1989, p. 448). An interval is obtained between regions where the plot of the LOD scores intersects with the 2 LOD line. This interval would be considered the most plausible region for the QTL. Resampling methods are employed to improve the construction of these intervals (Lynch, 1989, p. 450). By resampling, researchers avoid issues with multiple hypothesis tests and distribution assumptions (Doerge, 2001).

Interval mapping is more powerful than single marker analysis to detect QTL regions and can be used for detecting multiple QTLs on different chromosomes. However, it is still a one dimensional search through the genome and does not allow for more complex models for handling interactions or epistasis (genes dependent on modifier genes). Another issue is its tendency to detect ghost QTLs which are due to marker interval locations being in the vicinity of the real QTLs (Doerge, 2002). This method is computationally expensive when large numbers of markers are considered.
2.4 Multiple QTL Mapping (MQM)

Multiple QTL mapping techniques model multiple QTLs simultaneously, so they have the ability to account for the presence of other QTLs and interactions between them. Some methods for handling multiple QTLs include single marker analysis using multiple regression, forward selecting interval mapping, composite interval mapping (CIM), and multiple interval mapping (MIM) (Broman 2001). For all of these methods, identifying the best model or subset of models is a daunting task due to the sheer number of possibilities (Doerge, 2002).

The simplest case of handling multiple QTLs extends from the method of single marker ANOVA to employing multiple regression analysis, where a class of models must be identified and then compared to choose the best one. Forward selection, stepwise selection, and backward deletion are some of the methods used for model selection. This approach requires dense markers and a very complete set of genotype data to avoid problems with missing data which eliminates markers from the analysis (Broman 2001).

A MQM technique that allows for missing genotype data is forward selection in the interval mapping context. In this approach, a putative QTL is identified using interval mapping. The residuals from this are used as a new phenotype and interval mapping is repeated. This procedure works well for QTLs that are additive. It is also more powerful for detecting other QTLs and can separate linked QTLs (Broman 2001).
Composite interval mapping (CIM) also employs interval mapping but uses a subset of marker loci as covariates, thus controlling for other QTLs (Kao, 1999). The main issue is choosing suitable markers since the best choice would be ones that are linked (close) to the true QTL, and identifying these is the primary goal of the analysis (Broman 2001).

Multiple interval mapping (MIM) is the multiple regression extension to interval mapping. This approach employs multiple intervals at the same time to build multiple putative QTLs in the model (Kao, 1999). This allows inference on the location of QTLs to positions between markers, makes allowances for missing genotype data, and can allow for interactions between QTLs. There is still the issue of searching through many possible models (Broman, 2001).

2.5 The Role of Sample Size in QTL Analysis

With any inferential method, sample size plays an important role in detecting significant differences in a quantitative trait of interest. Researchers must estimate phenotypic variance and decide what effect sizes are important for detection. QTL effect sizes of (.5 – 1)σ are generally considered large enough effect sizes for applications in basic QTL analysis (Falconer & Mackay, 1996, p. 357). For each marker group, the sample size needed to detect a mean difference of \( \delta \) while controlling for type I error rate at \( \alpha \) and type II error rate at \( \beta \) can be obtained by

\[
n > 2\left[ Z_{\alpha/2} + Z_{\beta} \right] / (\delta / \sigma)^2
\]
The advantage of this method of calculating the sample size is that you don’t need to estimate the trait variance. $\delta / \sigma$ is the standardized effect size, so values of $\delta / \sigma = .5$ in the equation would mean the desired detected effect size is $.5 \sigma$ (Falconer & Mackay, 1996, p. 367).

3 USDA Perennial Grass Research

3.1 Rhizomes

Rhizomes are stems that grow below ground and spread out horizontally to form nodes/shoots. These nodes can then produce roots, leaves, and secondary stem branches. Rhizomes assist in resistance to grazing and over wintering as well as regrowth. As such, rhizomes are important to the survivability and propagation of perennial grasses. Rhizomes are absent in important cereal crops like wheat, and extremely strong in some invasive weeds like quack grass. Understanding the location of genes that determine rhizome growth could be used in breeding programs to improve forage and turf grasses, extend the life of cereal crops, and assist in controlling invasive weeds (Yun et al., 2014).

3.2 Plant Materials

Perennial grass species *Leymus cinereus* (C) and *Leymus triticoides* (T), commonly known as basin wildrye and creeping wildrye, are native to western North America and display many divergent traits. The genus *Leymus* belongs to the Triticeae grass tribe contains 400-500 species of grasses that including important
cereal crop species wheat, barley, and rye. *Leymus cinereus*, or basin wildrye, has an upright, compact growth habit described as caespitose. *Leymus triticoides* (T), or creeping wildrye, is distinctly different in its growth habit as it spreads out and forms sod and therefore is considered rhizomatous (Yun et al., 2014). These plants are allotetraploid (2n=4x=28), meaning the progeny inherits two sets of seven chromosomes from each parent, and show disomic inheritance (Larson et al., 2012). That these two plants can be crossbred and have distinctly different growth habits makes them good candidates for identifying major effect QTLs for rhizome growth.

Two hybrids, TC1 and TC2, were developed by crossing one *triticoides* (T) plant with two *L. cinereus* (C) plants. The TTC1 and TTC2 backcross mapping families were derived by taking the two hybrid plants, TC1 and TC2, and backcrossing them with the same *L. triticoides* plant (T-tester). This crossing allows for the detection of dominant *L. cinereus* effects. The TCC family was derived from one *L. cinereus* (C-tester) backcrossed with the TC1 hybrid for the purpose of detecting dominant *L. triticoides* gene effects. It is from this latter backcross progeny that plant materials were obtained for this project. All backcrosses in this study are considered pseudo-backcross populations because the original parents used to make the hybrid TC1 and TC2 populations were not available, so plants of the same species and geographic location were used instead (Yun et al., 2014).
3.3 Study Design and Data Description

Two hundred and fifty of the TCC hybrid backcross progeny were planted in the Spring of 2009 in a randomized complete block design with five clones of each progeny genotype (genet) per plot and two replicate sets (blocks) of the 250 TCC progeny, for a total of 10 clones each genet. Perimeters of rhizome spreading from each clone were measured in mid-June of 2011. Rhizome meristems from TC1 and TC2 were used to create an expressed gene sequence tag (EST) library. The EST sequences were aligned with known genome sequences of *Brachypodium*, rice and sorghum. Additional PCR markers were designed with predicted locations on the long arm of chromosome 6. Additional EST primers were also designed from an ortholog of the rice chromosome 4 Ra gene with predicted map location on chromosome group 2. The TCC population was also genotyped using 16 of the most informative primer combinations from the TTC1 and TTC2 populations (Yun et al., 2014).
3.4 Genetic Map Construction

The TCC population was genotyped for 411 amplified fragment length polymorphism (AFLP) loci and 122 EST markers that were present in the hybrids, absent in C-tester, and segregated among the 250 TCC progeny (Larson et al. 2012). The GLM procedure of SAS was used to obtain the least square means of each of the ten measurements by plant ID to create the phenotype file needed for building linkage groups and QTL maps (SAS Institute, 1999). The linkage group map was built using JoinMap Version 4.0 (Van Ooijen, 2006) and the double haploid (DH) model. The DH model was used to analyze meiosis in the TC hybrids because the actual parental plants of the hybrids were not available for determining linkage phases (Larson et al., 2012). Linkage groups (LG) were selected using a minimum LOD linkage threshold of ten using the TTC consensus map and perennial wildrye EST alignments to other grass genome reference sequences (wheat and barley) in order to identify two sets of seven homoeologous groups and provide a reference gene map (Larson et al. 2012). Thus, two sets of seven linkage groups (LG) were identified and numbered LG1a-LG7a and LG1b-LG7b, where a and b arbitrarily distinguish different sets or subgenomes of allotetraploid *Leymus*. Linear map orders and distances among markers within these 14 groups were determined by regression mapping using those linkages with recombination frequencies smaller than 0.4. MapQTL 6.0 (Van Ooijen, 2009) was used to build QTL maps using the interval mapping approach (Yun et al., 2014).
3.5 QTL Findings

Maps from TTC1 and TTC2 populations detected QTLs in linkage groups LG-3a, 3b and 6a. The interval QTL mapping (IM) procedure from the TCC population detected significant QTL’s on LG2a and LG6a with LOD values of 8.9 and 26.8 with 15.2% and 29% of the variation explained, respectively. The LG6a QTL was the only QTL detected in backcrosses to both species. The LOD value of the TCC LG6a QTL was 31.6 with 37.1% variation explained using a marker from the LG2a QTL as cofactor in a multiple QTL model (MQM). The LOD value of the TCC LG2a QTL was 11.1 with 11.3% variation explained using a marker from the LG6a QTL as a MQM cofactor. This increased the observed LOD scores for linkage group 6 but the general shape was maintained. Because LG6a presented significant LOD scores in both the reciprocal TTC and TCC pseudo-backcross experiments, the QTL in that region has additive or incomplete-dominant gene effects on rhizome growth (Yun et al., 2014).

From this point forward the focus will be on the QTL region in LG-6a. The two LOD drop off bounds from the IM approach for this region were located within a 3.8 centimorgan (cM) region between marker Ltc0171 (≈ 137.8 cM) and E37M63.202 (≈141.6 cM) with the latter marker’s location at the distal end of LG6a (Larson et al., 2012; Yun et al., 2015). Jackknife samples were obtained by randomly deleting 50% of the observations and then reanalyzing the remaining data using the IM approach. The markers producing the maximum LOD peak were recorded. This was repeated for a total of 1,000 jackknife samples. The 95% and 99% jackknife bounds were where 95% and 99% of the QTL peaks occurred, respectively. This procedure was
repeated for 10%-delete jackknife samples and then again using the MQM approach for both the 10% and 50%-delete jackknife samples (Yun et al., 2014). It should be noted that the MQM bounds were virtually identical to those which were obtained using the IM approach. In the 10%-delete jackknife samples the 95% bounds were between LG6L026 (≈139.9 cM) and LG6L129 (≈141.1 cM). In the 50%-delete jackknife samples, the 99% bounds were located between Ltc171 (≈137.8 cM) and the end of the chromosome at E37M62 (≈141.5 cM), which is similar to the two-LOD drop-off interval. It is within these bounds of 3.8 cM that the location of a major effect additive gene was postulated (Yun et al., 2015).

4 Fine Tuning of Rhizome QTL Analysis

4.1 Plant Selection

Seven of the 250 TCC recombinant progeny were identified as having chromosome crossovers in the 3.8-cM TCC LG-6a rhizome QTL interval where the gene was postulated. A total of 13 different female *L. cinereus* plants from two different varieties (TH and ACC) were selected as testers for cross pollination by the seven TCC progeny containing chromosome crossovers in the LG6a rhizome QTL. When possible, two female plants from different *L. cinereus* varieties (TH and ACC) were used to test gene effects of each of the seven TCC recombinant progeny. However, the TH x TCC133 cross was unsuccessful so all of the TCC133 BC2 progeny were obtained from the crossing with one ACC female. Conversely, all of the TCC125 BC2 progeny derived from crosses with two different TH plants because
there were no receptive ACC plants available for test crossing. Thus, a total of 13 subfamilies resulted from crosses of the seven recombinant TCC progeny, used as male pollen parents, with 13 different plants from two *L. cinereus* varieties, used as female seed parents. A total of 100 seeds for each population were planted in the greenhouse in the Fall of 2012 and three clones from each plant were then planted in a Utah State University field near Richmond, Utah in the Spring of 2013 (Larson, personal communication, June 2015). The resulting hybrids were labeled according to parental genotypes and then by plant number.

4.2 Study Design and Data Description

The cloned hybrids were planted in the field using a randomized block design with a serpentine order. The perimeters of rhizome spreading were measured in both 2014 and 2015 with the 2015 measurements used for the analysis. A number of physical events were experienced in the field over these two years which included damage by voles, weed infestation, shearing of plants by wind-driven soil (sand) particles (at least two events), and soil redistribution caused by two other events of fast-moving flood water that went through the middle part of the field (Larson, personal communication, October 2015). To assess whether there should be any blocking done by areas of the field, a color map of the 2015 rhizome perimeter measurements was compared to a random assignment of the same measurements (Figure 1). There did not appear to be an obvious field location effect.
Figure 1: Color maps of rhizome perimeters, actual versus randomized.
Primers for PCR amplification were chosen for detecting markers in the LG-6a regions based on the QTL map and were specific to areas where the male plant was heterozygous in the 50%-delete jackknife 99% interval region. Parents and hybrids were genotyped for each marker. In some cases, multiple markers resulted from the same primer with band size indicated by the right most number. Different amplicon (sizes) from the same primer can arise from different genome locations with independent inheritance patterns or from different parts of the same gene resulting in identical inheritance patterns (Larson, personal communication, October 2015). Figure 2 lists the markers in order of approximate location on LG6a with respect to the IM jackknife bounds, the LOD plot and graphical genotypes of the seven TCC hybrid male plants.
Figure 2: Graphical genotypes of seven recombinant TCC progeny used as BC2 test-cross parents with the LOD plot from the TCC QTL analysis. The solid lines on the LOD plot indicate the IM approach and the dashed lines indicate the MQM approach. Solid and dashed vertical lines on the QTL LOD graph also indicate the 2 LOD drop off thresholds for the 31.6-LOD MQM QTL peak and the 26.8-LOD IM QTL peak. For the 10% and 50% delete jackknife samples, the box indicates the 95% bounds and the whiskers indicates the 99% bounds for the 10% delete (shorter interval) and 50% delete jackknife samples (longer interval).
4.3 Analysis Description

For each population analyzed, the ANOVA fixed factor was the presence or absence of the genetic marker (A – present, B – absent). When possible, random blocking factors were female subfamily (ACC, TH) with plant ID nested in the subfamily. For this approach there were six populations analyzed. In addition, a one-way ANOVA was performed on each subfamily data to test the significant marker effect with ID as a random effect, resulting in 13 subpopulations analyzed. For the analysis to be informative, the marker tested should be present and heterozygous in the male plant (A) and homozygous absent in the female plant (B). Markers were eliminated if the male and female plant shared the same genotype or if the marker was absent in the male parent. If the marker was present in both females, it was completely removed from consideration. If the marker was present in just one female subfamily, the other subfamily analysis was used for assessment. The SAS GLIMMIX procedure was used for obtaining ANOVA p-values for each population as a measure of linkage (SAS Institute, version 9.3). In some of the analysis, the residuals showed right fanning and a log transform was performed to get those p-values. The log transformed p-values did not change the significance of the results so they were not reported.

The means procedure of SAS was used to get basic descriptive statistics of means, standard deviations, and sample sizes. The GLIMMIX procedure of SAS was used to obtain estimates of the Least Squares (LS) mean difference by marker, taking into account the other factors (subfamily and ID), and provided 95% confidence interval estimates of the LS mean difference (SAS Institute, version 9.3).
Microsoft Excel (version 2010) was used to obtain p-values for a Chi-square goodness of fit test to test if the progeny follow the expected 1:1 segregation ratio for backcross population and to calculate correlation matrices on the marker information. The latter was used to assess marker similarity in inheritance patterns within each of the plant populations. Markers close together on a chromosome should exhibit similar inheritance patterns among the progeny and would be highly correlated (Larson, personal communication, October 2015).

Multiple testing considerations are that the Chi-square tests and the ANOVA tests resulted in 20 p-values each. However, the markers were all taken from the same region on Chromosome 6 and within a four cM region, thus they should be highly correlated. A Bonferroni correction on the p-values would be too conservative. Holm’s method allowed for dependency by doing a Bonferroni correction in a stepdown fashion (Shaffer, 1986). There was a different approach by Benjamini and Hochberg that controlled the false discovery rate (FDR) that was valid for p-values under any kind of dependency and was more liberal than controlling for family wide error (FEW) rates (Benjamini and Yekateuli, 2001). It was important to consider how to define a family of hypothesis tests. One option was to address multiple comparison adjustments within a population and the other was to consider them all as one family. The Holm’s method for multiple testing was first calculated within each population. Holm’s was also applied to all 20 p-values, using the MULTTEST procedure in SAS (SAS Institute, version 9.3). In addition, dependent FDR adjustments were applied. Significance was assessed at the 0.05 level. None of the methods for adjusting the p-
values changed the significance of the ANOVA results but did have impact on some of the Chi-square test results.

4.4 Results

4.4.1 Descriptive Statistics with 95% Confidence Bounds

Differences in means ranged from 1.5 to 40 centimeters (cm). Negative effect sizes showed up consistently in the IAA10-5utr.182 markers across all populations. This can occur if a marker has a reversed phase relative to other markers (i.e. the parental source of marker allele is different from other markers). The TCC-083 population resulted in the smallest mean differences. TCC-254 resulted in the largest mean differences.
Table 1: Descriptive Statistics for Rhizome Circumference in cm 2015

<table>
<thead>
<tr>
<th>Male Plant</th>
<th>Marker</th>
<th>Mean\textsubscript{A}</th>
<th>SD\textsubscript{A}</th>
<th>N\textsubscript{A}</th>
<th>Mean\textsubscript{B}</th>
<th>SD\textsubscript{B}</th>
<th>N\textsubscript{B}</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCC-004</td>
<td>IAA10-5utr.223.6</td>
<td>85.51</td>
<td>28.93</td>
<td>51</td>
<td>67.49</td>
<td>20.42</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Ltch177.088</td>
<td>82.64</td>
<td>25.96</td>
<td>119</td>
<td>58.16</td>
<td>19.60</td>
<td>116</td>
</tr>
<tr>
<td>TCC-083</td>
<td>LG6aL.030.215</td>
<td>57.24</td>
<td>15.82</td>
<td>105</td>
<td>60.70</td>
<td>18.66</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>Ltch171.160</td>
<td>64.23</td>
<td>15.37</td>
<td>47</td>
<td>66.57</td>
<td>18.51</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Ltch173.119</td>
<td>57.68</td>
<td>15.27</td>
<td>105</td>
<td>59.89</td>
<td>19.11</td>
<td>116</td>
</tr>
<tr>
<td>TCC-125</td>
<td>IAA10-5utr.182</td>
<td>63.33</td>
<td>18.21</td>
<td>98</td>
<td>85.39</td>
<td>32.14</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>LG6aL.024</td>
<td>63.44</td>
<td>17.97</td>
<td>101</td>
<td>85.61</td>
<td>31.92</td>
<td>116</td>
</tr>
<tr>
<td></td>
<td>Ltch177.088</td>
<td>63.44</td>
<td>17.97</td>
<td>101</td>
<td>85.61</td>
<td>31.92</td>
<td>116</td>
</tr>
<tr>
<td>TCC-133</td>
<td>IAA10_5utr_224</td>
<td>70.29</td>
<td>24.76</td>
<td>107</td>
<td>57.04</td>
<td>15.60</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Ltch177.088</td>
<td>71.38</td>
<td>24.60</td>
<td>105</td>
<td>56.55</td>
<td>16.88</td>
<td>91</td>
</tr>
<tr>
<td>TCC-162</td>
<td>LG6aL.024.201</td>
<td>68.14</td>
<td>24.20</td>
<td>118</td>
<td>57.88</td>
<td>17.34</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>Ltch182.142</td>
<td>68.14</td>
<td>24.20</td>
<td>118</td>
<td>57.61</td>
<td>17.27</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>Ltch182.152</td>
<td>58.80</td>
<td>17.26</td>
<td>105</td>
<td>65.59</td>
<td>23.86</td>
<td>140</td>
</tr>
<tr>
<td>TCC-254</td>
<td>IAA10-5utr.182</td>
<td>77.14</td>
<td>22.32</td>
<td>72</td>
<td>117.02</td>
<td>48.74</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>IAA10-5utr.223</td>
<td>117.02</td>
<td>48.74</td>
<td>58</td>
<td>77.14</td>
<td>22.32</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Ltch177.88</td>
<td>117.02</td>
<td>48.74</td>
<td>58</td>
<td>77.14</td>
<td>22.32</td>
<td>72</td>
</tr>
<tr>
<td>TCC-265</td>
<td>IAA10-5utr.182</td>
<td>58.53</td>
<td>22.84</td>
<td>53</td>
<td>88.52</td>
<td>44.88</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>IAA10-5utr.223</td>
<td>88.52</td>
<td>44.88</td>
<td>29</td>
<td>58.53</td>
<td>22.84</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Ltch177.88</td>
<td>74.48</td>
<td>35.11</td>
<td>64</td>
<td>58.85</td>
<td>20.05</td>
<td>97</td>
</tr>
</tbody>
</table>
Table 2: LS Mean Difference Rhizome Circumference in cm 2015

<table>
<thead>
<tr>
<th>Male Plant</th>
<th>Marker</th>
<th>Diff (A-B)</th>
<th>S.E.</th>
<th>DF</th>
<th>CI Lower</th>
<th>CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCC-004</td>
<td>IAA10-5utr.223.6</td>
<td>17.95</td>
<td>4.79</td>
<td>71</td>
<td>8.41</td>
<td>27.50</td>
</tr>
<tr>
<td></td>
<td>Ltc177.088</td>
<td>24.62</td>
<td>3.35</td>
<td>145</td>
<td>18.00</td>
<td>31.23</td>
</tr>
<tr>
<td>TCC-083</td>
<td>LG6aL.030.215</td>
<td>-2.00</td>
<td>2.66</td>
<td>127</td>
<td>-7.26</td>
<td>3.27</td>
</tr>
<tr>
<td></td>
<td>Ltc171.160</td>
<td>-2.33</td>
<td>3.87</td>
<td>67</td>
<td>-10.05</td>
<td>5.40</td>
</tr>
<tr>
<td></td>
<td>Ltc173.119</td>
<td>-1.46</td>
<td>2.68</td>
<td>129</td>
<td>-6.77</td>
<td>3.84</td>
</tr>
<tr>
<td></td>
<td>IAA10-5utr.223</td>
<td>21.88</td>
<td>4.59</td>
<td>127</td>
<td>12.79</td>
<td>30.97</td>
</tr>
<tr>
<td></td>
<td>LG6aL.024</td>
<td>22.68</td>
<td>5.08</td>
<td>113</td>
<td>12.62</td>
<td>32.74</td>
</tr>
<tr>
<td></td>
<td>Ltc177.088</td>
<td>20.71</td>
<td>4.63</td>
<td>130</td>
<td>11.56</td>
<td>29.87</td>
</tr>
<tr>
<td>TCC133</td>
<td>IAA10_5utr_224</td>
<td>13.92</td>
<td>3.35</td>
<td>113</td>
<td>7.29</td>
<td>20.55</td>
</tr>
<tr>
<td></td>
<td>Ltc177.088</td>
<td>15.70</td>
<td>3.45</td>
<td>109</td>
<td>8.86</td>
<td>22.54</td>
</tr>
<tr>
<td>TCC-162</td>
<td>LG6aL.024.201</td>
<td>9.69</td>
<td>3.36</td>
<td>150</td>
<td>3.06</td>
<td>16.33</td>
</tr>
<tr>
<td></td>
<td>Ltc182.142</td>
<td>9.95</td>
<td>3.33</td>
<td>152</td>
<td>3.37</td>
<td>16.53</td>
</tr>
<tr>
<td></td>
<td>Ltc182.152</td>
<td>-6.64</td>
<td>3.47</td>
<td>152</td>
<td>-13.49</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>IAA10-5utr.223</td>
<td>39.89</td>
<td>7.08</td>
<td>83</td>
<td>25.80</td>
<td>53.98</td>
</tr>
<tr>
<td></td>
<td>Ltc177.88</td>
<td>39.89</td>
<td>7.08</td>
<td>83</td>
<td>25.80</td>
<td>53.98</td>
</tr>
<tr>
<td>TCC-265</td>
<td>IAA10-5utr.182</td>
<td>-29.97</td>
<td>7.71</td>
<td>45</td>
<td>-45.50</td>
<td>-14.44</td>
</tr>
<tr>
<td></td>
<td>IAA10-5utr.223</td>
<td>29.97</td>
<td>7.71</td>
<td>45</td>
<td>14.44</td>
<td>45.50</td>
</tr>
<tr>
<td></td>
<td>Ltc177.88</td>
<td>15.34</td>
<td>4.94</td>
<td>89</td>
<td>5.52</td>
<td>25.16</td>
</tr>
</tbody>
</table>

4.4.2 Chi-square Test of Segregation

The observed segregation ratios for most markers was not significantly different from the expected ratio of 1:1 or frequency 0.5 (see Table 3). The TCC-265 BC2 population showed significant evidence of segregation distortion for all markers examined even after adjusting the p-values using Holm’s method within the population. One possible explanation for segregation distortion is that plants are not progeny of the intended parents. Sometimes the source of pollen is
difficult to control or perhaps the wrong plant was used as a source of pollen.

Segregation distortion can also be caused by deleterious genes that may reduce the survival of gametes or developing embryos or genes controlling mate compatibility including self-compatibility or species compatibility (Larson et al. 2012).

Table 3: Chi-Square Goodness of Fit Test of Marker Segregation

<table>
<thead>
<tr>
<th>Male Pollinator</th>
<th>Marker</th>
<th>P_A</th>
<th>N_A</th>
<th>N_B</th>
<th>N</th>
<th>P-value</th>
<th>W/in Holm's</th>
<th>Global Holm's</th>
<th>Adj FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCC-004</td>
<td>IAA10-5utr.223.6</td>
<td>0.43</td>
<td>51</td>
<td>67</td>
<td>118</td>
<td>0.1408</td>
<td>0.2815</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>Ltc177.088</td>
<td>0.51</td>
<td>119</td>
<td>116</td>
<td>235</td>
<td>0.8448</td>
<td>0.8449</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>TCC-083</td>
<td>LG6aL.030.215</td>
<td>0.48</td>
<td>105</td>
<td>113</td>
<td>218</td>
<td>0.5879</td>
<td>0.9187</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>Ltc171.160</td>
<td>0.41</td>
<td>47</td>
<td>68</td>
<td>115</td>
<td>0.0502</td>
<td>0.1506</td>
<td>0.8032</td>
<td>0.7224</td>
</tr>
<tr>
<td></td>
<td>Ltc173.119</td>
<td>0.48</td>
<td>105</td>
<td>116</td>
<td>221</td>
<td>0.4594</td>
<td>0.9187</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>TCC-125</td>
<td>IAA10-5utr.182</td>
<td>0.45</td>
<td>98</td>
<td>120</td>
<td>218</td>
<td>0.1362</td>
<td>0.5449</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>IAA10-5utr.223</td>
<td>0.51</td>
<td>111</td>
<td>107</td>
<td>218</td>
<td>0.7865</td>
<td>0.9257</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>LG6aL.024</td>
<td>0.47</td>
<td>101</td>
<td>116</td>
<td>217</td>
<td>0.3086</td>
<td>0.9257</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>Ltc177.088</td>
<td>0.47</td>
<td>101</td>
<td>116</td>
<td>217</td>
<td>0.3086</td>
<td>0.9257</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>TCC133</td>
<td>IAA10-5utr.224</td>
<td>0.53</td>
<td>107</td>
<td>94</td>
<td>201</td>
<td>0.3592</td>
<td>0.6346</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>Ltc177.088</td>
<td>0.54</td>
<td>105</td>
<td>91</td>
<td>196</td>
<td>0.3173</td>
<td>0.6346</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>TCC-162</td>
<td>LG6aL.024.201</td>
<td>0.49</td>
<td>118</td>
<td>124</td>
<td>242</td>
<td>0.6997</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>Ltc182.142</td>
<td>0.48</td>
<td>118</td>
<td>127</td>
<td>245</td>
<td>0.5653</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>Ltc182.152</td>
<td>0.43</td>
<td>105</td>
<td>140</td>
<td>245</td>
<td>0.0254</td>
<td>0.0760</td>
<td>0.4309</td>
<td>0.4560</td>
</tr>
<tr>
<td>TCC-254</td>
<td>IAA10-5utr.182</td>
<td>0.55</td>
<td>72</td>
<td>58</td>
<td>130</td>
<td>0.2195</td>
<td>0.6585</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>IAA10-5utr.223</td>
<td>0.45</td>
<td>58</td>
<td>72</td>
<td>130</td>
<td>0.2195</td>
<td>0.6585</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>Ltc177.88</td>
<td>0.45</td>
<td>58</td>
<td>72</td>
<td>130</td>
<td>0.2195</td>
<td>0.6585</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>TCC-265</td>
<td>IAA10-5utr.182</td>
<td>0.65</td>
<td>53</td>
<td>29</td>
<td>82</td>
<td>0.0080</td>
<td>0.0241</td>
<td>0.1608</td>
<td>0.2231</td>
</tr>
<tr>
<td></td>
<td>IAA10-5utr.223</td>
<td>0.35</td>
<td>29</td>
<td>53</td>
<td>82</td>
<td>0.0080</td>
<td>0.0241</td>
<td>0.1608</td>
<td>0.2231</td>
</tr>
<tr>
<td></td>
<td>Ltc177.88</td>
<td>0.40</td>
<td>64</td>
<td>97</td>
<td>161</td>
<td>0.0093</td>
<td>0.0241</td>
<td>0.1674</td>
<td>0.2231</td>
</tr>
</tbody>
</table>
4.4.3 ANOVA Results by Population

Recall that the seven male (pollen) parents of the BC2 families were chosen to be progeny tested by selecting those TCC plants with crossovers in the putative TCC LG6a rhizome QTL region. See Figure 2. Markers were then selected for genotyping the progeny of each male based on where the male parent was heterozygous in the TCC LG6a rhizome QTL region. If the ANOVA test for that marker was significant, this indicates that the male parent carries at least one gene affecting rhizome growth within the region of the TCC LG6a rhizome QTL region where that male was heterozygous for *L. cinereus* and *L. triticoides* alleles. Assuming that the TCC LG6a rhizome QTL was caused by only one gene, segregation of significant marker effects on rhizome development eliminate regions of the TCC LG6a rhizome QTL that were homozygous for *L. cinereus* marker alleles in the male parent. Conversely, assuming that the power of our test is adequate, failure to detect significant marker effects on rhizome development in a BC2 family indicates that the male parent did not carry a gene affecting rhizome growth within the region TCC LG6a rhizome QTL region where that male was heterozygous for *L. cinereus* and *L. triticoides* alleles. Likewise, failure to detect significant marker effects on rhizome development in a BC2 family indicates that the TCC LG6a rhizome QTL was caused by a gene that was present in a region of the TCC LG6a rhizome QTL that was homozygous for *L. cinereus* marker alleles in the male parent. Results will be presented for each of the seven recombinant TCC genotypes that were progeny tested for segregation of the LG6a rhizome
gene(s) as male pollen parents of 13 BC2 families in crosses with one or two different *L. cinereus* female tester plants.

TCC-004 progeny test

The crossover in the TCC-004 recombinant genotype occurred between homozygous LG6al.082 and heterozygous LG6aL.026 (Fig. 2). The TCC-005 genotype is heterozygous for *L. cinereus* and *L. triticoides* marker alleles from LG6aL.026 to the most distal marker below LG6aL.026 (Fig. 2). Two markers were selected below LG6al.026 and the results from both produced significant p-values. Ltc177.088 had the smallest p-value. Both markers inheritance patterns were strongly correlated with r = 0.88. The ANOVA results excluded the QTL region above LG6aL.082 and validated the region below LG6aL.026. See Figure 2.

<table>
<thead>
<tr>
<th>TCC-004 Markers</th>
<th>F Value</th>
<th>DF</th>
<th>Pr &gt; F</th>
<th>W/in Holm's</th>
<th>Global Holm's</th>
<th>Adj FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAA10-5utr.223.6</td>
<td>14.06</td>
<td>71</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.0033</td>
<td>0.0020</td>
</tr>
<tr>
<td>Ltc177.088</td>
<td>54.09</td>
<td>145</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

TCC-083 progeny test

The crossover in the TCC-083 recombinant genotype occurred between heterozygous LG6al.082 and homozygous LG6aL.026. The TCC-083 genotype is heterozygous for *L. cinereus* and *L. triticoides* marker alleles from LG6al.082 to Ltc0171 and well above Ltc0171 (Fig. 2). Three markers were selected above LG6aL.082. None of the three markers resulted in significant p-values.
Correlation in the markers inheritance patterns ranged from 0.57 to 0.98. These results were consistent with the exclusion of the region above LG6aL.082 from the TCC-004 results. See Figure 2.

<table>
<thead>
<tr>
<th>TCC-083 Markers</th>
<th>F-Value</th>
<th>DF</th>
<th>Pr &gt; F</th>
<th>W/in Holm’s</th>
<th>Global Holm’s</th>
<th>Adj FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>LG6aL.030.215</td>
<td>0.56</td>
<td>127</td>
<td>0.4556</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Ltc171.160</td>
<td>0.36</td>
<td>67</td>
<td>0.5505</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Ltc173.119</td>
<td>0.30</td>
<td>129</td>
<td>0.5848</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

TCC-125 progeny test

The crossover in the TCC-125 recombinant genotype occurred between homozygous LG6aL.026 and heterozygous IAA10. The TCC-125 genotype is heterozygous for *L. cinereus* and *L. triticoides* marker alleles from IAA10 and Ltc177 to the most distal marker below IAA10 (Fig. 2). Four markers were selected at IAA10 and below. All resulted in significant p-values. Ltc177.088 produced the smallest p-value. These markers inheritance patterns were strongly correlated (between 0.90 and 0.95). Based on the ANOVA results, the region above Lg6aL.026 was excluded and below IAA10 was validated. See Figure 2.

<table>
<thead>
<tr>
<th>TCC-125 Marker</th>
<th>F Value</th>
<th>DF</th>
<th>Pr &gt; F</th>
<th>W/in Holm’s</th>
<th>Global Holm’s</th>
<th>Adj FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAA10-5utr.182</td>
<td>23.57</td>
<td>127</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>IAA10-5utr.223</td>
<td>22.68</td>
<td>127</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>LG6aL.024</td>
<td>19.94</td>
<td>113</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Ltc177.088</td>
<td>20.02</td>
<td>130</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.0002</td>
<td>0.0002</td>
</tr>
</tbody>
</table>
TCC-133 progeny test

The crossover in the TCC-133 recombinant genotype occurred between heterozygous IAA10 and homozygous Ltc182. The TCC-083 genotype is heterozygous for *L. cinereus* and *L. triticoides* marker alleles from IAA10 and Ltc177 to Ltc0171 and well above Ltc0171 (Fig. 2). Two closely-linked markers were selected including IAA10 and Lt177. Both markers in this population were significant and their inheritance patterns were strongly correlated (r = 0.93). The Ltc177.088 marker produced the smallest p-value. The ANOVA results excluded the region below Ltc182 and validated above IAA10. See Figure 2.

<table>
<thead>
<tr>
<th>TCC133 Marker</th>
<th>F Value</th>
<th>DF</th>
<th>Pr &gt; F</th>
<th>W/in Holm’s</th>
<th>Global Holm’s</th>
<th>Adj FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAA10-5utr.224</td>
<td>17.29</td>
<td>113</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.0007</td>
<td>0.0005</td>
</tr>
<tr>
<td>Ltc177.088</td>
<td>20.71</td>
<td>109</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.0002</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

TCC-162 progeny test

The crossover in the TCC-162 recombinant genotype occurred between homozygous IAA10 and heterozygous Ltc182. The TCC-162 genotype is heterozygous for *L. cinereus* and *L. triticoides* marker alleles from Ltc182 to the most distal marker below Ltc182 (Fig. 2). Markers were selected at and below Ltc182. LG6aL.024.201 and Ltc182.142 markers were significant in this population and their inheritance pattern had a correlation of 1. Ltc182.152 raw p-values showed some evidence (< 0.10) but this marker was different from the other by band size. Its inheritance pattern was also not strongly correlated with the
other markers (0.015 and 0.024). This could be due to an unintentional amplification of a different region on the chromosome. The ANOVA results excluded the region above the IAA10-5utr marker and validate below Ltc182. This exclusion combined with the TCC-133 exclusion resulted in a reduction of the QTL region to between IAA10 and Ltc182. See Figure 2.

<table>
<thead>
<tr>
<th>TCC-162 Marker</th>
<th>F Value</th>
<th>DF</th>
<th>Pr &gt; F</th>
<th>W/in Holm’s</th>
<th>Global Holm’s</th>
<th>Adj FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>LG6aL.024.201</td>
<td>8.33</td>
<td>150</td>
<td>0.0045</td>
<td>0.0090</td>
<td>0.0224</td>
<td>0.0201</td>
</tr>
<tr>
<td>Ltc182.142</td>
<td>8.92</td>
<td>152</td>
<td>0.0033</td>
<td>0.0099</td>
<td>0.0197</td>
<td>0.0158</td>
</tr>
<tr>
<td>Ltc182.152</td>
<td>3.66</td>
<td>152</td>
<td>0.0576</td>
<td>0.0576</td>
<td>0.2305</td>
<td>0.2439</td>
</tr>
</tbody>
</table>

TCC-254 progeny test

The crossover in the TCC-254 recombinant genotype occurred between heterozygous LG6aL.030 and homozygous Ltc173. The TCC-254 genotype is heterozygous for *L. cinereus* and *L. triticoides* marker alleles from LG6aL.030 to the most distal marker below LG6aL.030 (Fig. 2). Markers were selected around IAA10. These markers were identical in all ANOVA results and the inheritance pattern perfectly correlated in the TH family. The AC family was excluded from this analysis due to segregation distortion but the analysis that includes them both yields similar results with p-values <0.0001. These results excluded the region above Ltc173 and validated below LG6aL.030. See Figure 2.

<table>
<thead>
<tr>
<th>TCC-254 Markers</th>
<th>F Value</th>
<th>DF</th>
<th>Pr &gt; F</th>
<th>W/in Holm’s</th>
<th>Global Holm’s</th>
<th>Adj FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAA10-5utr.182</td>
<td>31.71</td>
<td>83</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>IAA10-5utr.223</td>
<td>31.71</td>
<td>83</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Ltc177.088</td>
<td>31.71</td>
<td>83</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
TCC-265 progeny test

The crossover in the TCC-265 recombinant genotype tested as a male pollen parent also occurred between heterozygous LG6aL.030 and homozygous Ltc173. The TCC-265 genotype is heterozygous for *L. cinereus* and *L. triticoides* marker alleles from LG6aL.030 to the most distal marker below LG6aL.030 (Fig. 2). Markers were selected around IAA10 and all produced significant p-values. There was also strong correlation in the markers inheritance patterns (.95 to 1). The IAA10 markers produced the smallest p-values in this population. These results also exclude the region above Ltc173. See Figure 2.

<table>
<thead>
<tr>
<th>TCC-265 Markers</th>
<th>F Value</th>
<th>DF</th>
<th>Pr &gt; F</th>
<th>W/in Holm’s</th>
<th>Holm’s</th>
<th>Adj FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAA10-5utr.182</td>
<td>15.10</td>
<td>45</td>
<td>0.0003</td>
<td>0.0009</td>
<td>0.0033</td>
<td>0.0020</td>
</tr>
<tr>
<td>IAA10-5utr.223</td>
<td>15.10</td>
<td>45</td>
<td>0.0003</td>
<td>0.0009</td>
<td>0.0033</td>
<td>0.0020</td>
</tr>
<tr>
<td>Ltc177.088</td>
<td>9.64</td>
<td>89</td>
<td>0.0026</td>
<td>0.0047</td>
<td>0.0179</td>
<td>0.0131</td>
</tr>
</tbody>
</table>

4.4.4 Power Analysis

The TCC083 population did not show significant marker effects for any of the markers selected. Power calculations were done to assess whether the sample size was sufficient to detect marker effects in that population. The sample size needed for 90% power and an effect size of .5σ for α = 0.05 is 49 per group. The sample size needed for 90% power and a standardized effect size of 1σ for α = 0.05 is 21 per group. As such, the sample sizes obtained were sufficient.
In the TCC083 results, standard deviations between 15 and 20 were observed. An effect size of 14 was seen in the original 250 TCC mapping for that QTL region (Yun et al., 2015). Figure 3 shows power plots obtained using the POWER procedure for one-way ANOVA significance tests in SAS (SAS Institute, version 9.3)
Figure 3: Power plots for each effect size of 15, 10 and 5 with three different levels of standard deviation, 15, 20, and 25 as indicated in the lower panel of the plots.
Conclusion

From the IM results, the most conservative jackknife bounds of the TCC LG6a rhizome QTL were between Ltc.171 and E37M62.202 on the distal end of LG6aL, which defined an ostensible 3.8-cM interval, corresponding to 106 gene models in the barley genome reference sequence, and major-effect gene controlling rhizome proliferation in closely related perennial Triticeae grasses of genus *Leymus* (Yun et al., 2014). Results of my analysis demonstrate the gene(s) responsible for the TCC LG6a rhizome QTL were present in a heterozygous configuration in six recombinant TCC progeny (TCC004, TCC135, TCC133, TCC162, TCC254, and TCC265) and homozygous absent in a one recombinant TCC progeny genotype (TCC083). Based on the reported genotypes of these seven TCC recombinant progeny (Figure 2), the only TCC LG6a rhizome QTL region that could be heterozygous among TCC004, TCC135, TCC133, TCC162, TCC254, and TCC265 and homozygous in TCC083 is located in a presumed region of overlap between chromosome cross-overs in the formation of the recombinant TCC-133 and TCC-125 progeny genotypes. Because TCC-133 and TCC-162 had crossovers in the region between IAA10 and Ltc182 and TCC-133 segregated above IAA10 and TCC-125 segregated below Ltc182, their combined results rigorously confirm presence of a rhizome gene within the 3.8 cM TCC LG6a rhizome QTL interval described by Yun et al. (2015) and further reduced the QTL region to a sharply defined region between IAA10 (≈140.3 cM) and Ltc182 (≈141.1 cM), which is less than 0.8 cM.
6 Discussion of Future Work

Currently 100 more markers are being developed within the 0.8-cM region between IAA10 and LGL.024 for genotyping the TCC-133 and TCC-162 genotypes and their derived BC2 populations. I anticipate that results of my analysis, combined with additional genotyping of the TCC-133 and TCC-162 plants, will further reduce the 0.8-cM LG6a rhizome gene interval. Moreover, I anticipate that comparisons of gene expression in rhizomes of selected TCC progeny and refinements of the corresponding gene interval in the barley genome reference sequence will help identify a relatively small subset of candidate genes that may be responsible for this biologically important trait.
References


15. The ANOVA results, descriptive statistics, power plots, multi testing adjustments for this paper were generated using SAS software, Version 9.3 of SAS System for Windows. Copyright © 2011.


19. URL [http://passel.unl.edu/](http://passel.unl.edu/)
Sample SAS Single Marker Analysis Code

/*Analysis for TC004*/
/*Read and prepare data from open excel file*/
FILENAME xlsdat2 DDE "excel|TCC004!r3c1:r290c10" lrecl=10240;  * m1 $ m2 $ m3 $ m4 $ m5 $ m6 $ m7 $ ;
data TCC004; /*name of data file*/
infile xlsdat2; /*this references open excel file*/
INPUT ID $ family $ subfamily $ rep $ m1 $ m2 $ m3 $ m4 $ CIRC_14 CIRC_15;
label m1 = "IAA10-5utr.223.6"
    m2 = "IAA10-5utr.222.6"
    m3 = "IAA10-5utr.182"
    m4 = "Ltc177.088";
if subfamily='Parent' then delete;
run;
/*subset by TH-13.5 subfamily*/
data TCC004_1;
set TCC004;
where subfamily = 'TH-13.5';
run;
/*subset by AC17-6 subfamily*/
data TCC004_2;
set TCC004;
where subfamily = 'Acc636-1';
run;
/*data transform when fanning is present*/
data Tdata;
set TCC004;
lny    = log(CIRC_15);
run;
/*Code for Marker ANOVA ID nested in subfamily, One-Way ANOVA ID is random effect, descriptive statistics*/
ods html image_dpi=300 style=journal;
ods graphics on;
title 'Circ_15 ANOVA m4 with ID nested in subfamily';
PROC glimmix data = TCC004 plots = residualpanel;
class m4 ID subfamily;
model CIRC_15 = m4;
random subfamily ID(subfamily);
lsmeans m4/ alpha = .05 e cl tdiff pdiff plot;
RUN;

/*Analysis for subfamily TH-13.5*/
ods graphics on;
title 'Circ_15 ANOVA m4 for TH-13.5 with ID as random effect';
PROC glimmix data = TCC004_1 plots = residualpanel;
class m4 ID;
model CIRC_15 = m4;
random ID;
lsmeans m4/ alpha = .05 e cl tdiff pdiff plot;
RUN;
ods graphics off;

title "Descriptive Statistics m4 for TH-13.5";
proc means data = TCC004_1 N MEAN STD;
class m4;
var CIRC_15;
run;

/*Analysis for subfamily Acc636-1*/
ods graphics on;
title 'Circ_15 ANOVA m4 for subfamily Acc636-1 with ID as random effect';
PROC glimmix data = TCC004_2 plots = residualpanel;
class m4 ID;
model CIRC_15 = m4;
random ID;
lsmeans m4/ alpha = .05 e cl tdiff pdiff plot;
RUN;
ods graphics off;

title "Descriptive Statistics m4 for subfamily Acc636-1";
proc means data = TCC004_2 N MEAN STD;
class m4;
var CIRC_15;
run;

/*Transformed data analysis */
ods graphics on;
title 'Log transformed Circ_15 ANOVA m3 with ID nested in subfamily';
PROC glimmix data = Tdata plots = residualpanel;
class m4 ID subfamily;
model lny = m4;
random subfamily ID(subfamily);
lsmeans m4/ alpha = .05 e cl tdiff pdiff plot;
RUN;
ods graphics off;

/* Clear current windows and start a new (clean) Results Viewer window (may need to run this twice) */
ods html close;
dm 'log; clear; output; clear' continue;
dm 'log; next results; clear; cancel' whostedit continue;
ods html newfile=none;
run;

/*Power plot code*/
proc power ;
 onewayanova
  groupmeans = 15 | 30
  stddev = 15 20 25
  alpha = 0.05
  npergroup = 2 to 10 by 1 12 to 20 by 2 25 to 55 by 5
  power = .;
  plot x=n min=2 max=50;
run;

/*Multiple Testing Calculations*/
title"Adjusted ANOVA p-values";
data a;
input Test$ Raw_P;
datalines;
test1 0.000358031
  test2 0.000000000013001
test4  0.455643833  
  test5  0.550530207  
  test6  0.58482888  
  test7  0.00000513348  
  test8  0.00000347092  
  test9  0.00000165521  
  test10  0.0000190332  
  test11  0.0000627371  
  test12  0.000139541  
  test13  0.004475474  
  test14  0.003289072  
  test15  0.057613804  
  test16  0.00000237733  
  test17  0.00000237733  
  test18  0.00000237733  
  test19  0.000332067  
  test20  0.000332067  
  test21  0.002553345;  

  title "Adjusted ANOVA p-values";  
  proc multtest pd=data=a holm dependentfdr out=new;  
  run;  
  proc print data=new;  
  run;  

  title "Adjusted Chi-square test for segregation p-values";  
  data b;  
  input Test$ Raw_P;  
  datalines;  
  test1  0.140772773  
  test2  0.84484622  
  test4  0.587936746  
  test5  0.0501995  
  test6  0.459336357  
  test7  0.13621687  
  test8  0.786457035  
  test9  0.308551395  
  test10  0.308551395  
  test11  0.359169208  
  test12  0.317310508  
  test13  0.699722857  
  test14  0.565298798  
  test15  0.025347319
test16 0.219491558
test17 0.219491558
test18 0.219491558
test19 0.008040685
test20 0.008040685
test21 0.009301633
;
proc multtest pdata=b holm dependentfdr out=new;
run;
proc print data=new;
run;

R code for colormap of field circumference
library(data.table)
library(xlsx)
library(fields)
#Read in file of map by plot#
FieldMap = read.table("FieldMapR.txt")
class(FieldMap)

#read in Richmond field data file from working directory
TCC = read.table('Richmond_Fieldbook.txt', header = T, nrows = 1765, blank.lines.skip = F, na.strings = ".")
TCC$Plot = as.integer(TCC$Plot)
summary(TCC$CIRC_15_cm)

#matrix of FieldMap
FM = as.matrix(FieldMap)
nr = nrow(FM)
nc = ncol(FM)

#replaces plot order with value of circumference
FM_Cir = matrix(ncol = nc, nrow = nr)
colnames(FM_Cir) = 10 + c(nc:1)
rownames(FM_Cir) = c(nr:1)
for (i in 1:nr) {
  for (j in 1:nc) {
    FM_Cir[i,j] = TCC$CIRC_15_cm[i]  when
  }
}

for (i in 1:nr) {
  for (j in 1:nc) {
    FM_Cir[i,j] = TCC$CIRC_15_cm[i]  when
  }
}
```r
FM_Cir[i,j] = TCC$CIRC_15_cm[TCC$Plot == FM[i,j]]
}
}
#creates randomized vector of circumference values
set.seed(4670)
N = length(TCC$CIRC_15_cm)
TCC$RandomCir = sample(TCC$CIRC_15_cm, size = N)
head(TCC$RandomCir)
summary(TCC$RandomCir)

#replaces plot order with value of circumference
Random_Cir = matrix(ncol = nc, nrow = nr)
colnames(FM_Cir) = 10 + c(nc:1)ownames(FM_Cir) = c(nr:1)

for (i in 1:nr) {
  for (j in 1:nc) {
    Random_Cir[i,j] = TCC$RandomCir[TCC$Plot == FM[i,j]]
  }
}

#Combined colormaps
windowsFonts(A = windowsFont("Times New Roman"))
set.panel()
par(oma=c(4, 0, 0, 0)) # margin of 4 spaces at the bottom
par(mfrow=c(2,1))
#set.panel(2, 1)
#blue/yellow colormap of circumference values
image(FM_Cir, col = topo.colors(119), breaks = c(seq(1,240, by = 2)), axes = FALSE,
     main = "Field Colormap of Rhizome Perimeters 2015", font.main = 2)

#randomized color map
image.plot(Random_Cir, col = topo.colors(119), breaks = c(seq(1,240, by = 2)), axes=FALSE,
           add = FALSE, horizontal = TRUE, legend.shrink = .5, legend.lab = "Centimeters",
           main = "Randomized Rhizome Perimeters", font.main = 2)
set.panel() # reset plotting device
```