ABSTRACT

XIONG, JIN. Genetic Analysis of Forking Defects in Loblolly Pine. (Under the direction of Dr. Steven E. McKeand and Dr. Ross W. Whetten.)

Forking defects, including stem forking and ramicorn branching, are serious stem-quality problems in loblolly pine (*Pinus taeda* L.). The presence of forked stems and ramicorn branches greatly reduces wood quality, quantity, and the economic value of the wood. Assessing forking traits may enable breeders to successfully breed and deploy non-forked phenotypes in breeding populations. The objectives of this study were to assess the genetic basis of forking in loblolly pine by assessing trees in a large number of full-sib seedling tests throughout the Southeast and in a clonal test in South Carolina.

In the full-sib tests, the percentage of forking averaged 18%, and ranged from 4% to 80% across different test series. The individual-tree heritability was low (0.06), but family-mean heritabilities were moderately high (half-sib=0.76, narrow-sense full-sib=0.59, broad-sense full-sib =0.71), indicating that forking is partially under genetic control at the family level. By using half-sib family selection with a selection differential of 20%, genetic gain could be achieved with 12% to 23% reduction of forking across different regions. A weak unfavorable genetic correlation (0.18) was found between forking and height, suggesting that selection for growth alone will negatively impact forking in loblolly pine. A favorable genetic correlation (0.33) was found between forking and straightness.

In the clonal test, stem forking and ramicorn branching were serious problems in some clones. Forking averaged 17% with clone means ranging from 0% to 73%, and ramicorn branching averaged 24%, with clone means ranging from 3% to 50%. The estimated clone-mean repeatabilities were 0.86 and 0.67 for forking and ramicorn branching, respectively. Unfavorable genetic correlations between growth traits and forking defects were also found in this clonal test, suggesting that selection for either trait alone will negatively affect the genetic response for the other. A moderate positive genetic correlation between stem forking and ramicorn branching indicates that both traits may be partially controlled by the same genes and could be improved simultaneously.

Co-segregation analysis of markers and QTL for forking was conducted using the clonally replicated progeny of the outbred full-sib family from the clonal test. The linkage map was constructed with three types of markers containing various segregation patterns (F2, BC1 and BC2). A linkage map was achieved with18 linkage groups defined by 409 SNP markers. A single-marker analysis of 1257 loci identified 11 and 9 markers that were significantly associated with stem forking and ramicorn branching, respectively. Those SNP markers explained 7% to 10% of the total phenotypic variation for forking defects. By using the interval mapping method, two QTL were identified for forking, and three QTL were identified for ramicorn branching. Two QTL were detected for both traits using multiple traits analysis, suggesting there are pleiotropic effects on both traits. The

result from this study further confirmed that there is genetic control for forking defects, and genetic gain can be achieved through family and clonal selection. With the identified QTL influencing forking defects, marker-assisted selection can be considered in decreasing forking defect in the breeding population of loblolly pine.

Genetic Analysis of Forking Defects in Loblolly Pine

by
Jin Xiong

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Forestry

Raleigh, North Carolina

2010

APPROVED BY:

Dr. Steven E. McKeand Committee Co-Chair	Dr. Ross W. Whetten Committee Co-Chair
Dr. Fikret Isik Committee Member	Dr. Zhao-Bang Zeng Committee Member

BIOGRAPHY

Jin Xiong grew up in Guilin, one of the most famous tourist cities in China with the beautiful scenery of Li River and Karst hills. She spent three years in Guilin High School and graduated June of 1998. After that, she attended Guangxi University for undergraduate study. In the college of forestry, she have enjoyed her studies in the major of forestry and gained the fundamental skills and knowledge as a forester. In June of 2002, she graduated with a Bachelor degree in Forestry from Guangxi University.

With the interests in forestry and genetics, Xiong decided to pursue a master degree at Beijing Forestry University in September 2002, majoring forest genetic and tree breeding. Under the supervisor of Dr. Chen Xiaoyang, she engaged in a research project 'transformation of *Populus tomentosa* with *rolB* genes and the influence on rooting and growth ', sponsored by the National Science Foundation of China.

After achieving a master degree, Xiong went to North Carolina State University in 2006 to pursue her doctorate. Supervised by Dr. Steve McKeand and Dr. Ross W. Whetten of Cooperative Tree Improvement Program, she has been working on a project entitled 'Identification of Quantitative Trait Loci (QTL) influencing forking defect in loblolly pine by using high-throughput SNP markers'.

ACKNOWLEDGMENTS

First, I would like to acknowledge my co-advisers Dr. Steve McKeand and Dr. Ross Whetten for their invaluable guidance, constant inspiration and great support. With profound knowledge and serious attitude, they give me a model for the kind of the academic I hope to be. I am always grateful to my committees Dr. Fikret Isik and Dr. Zhao-Bang Zeng for their guidance in statistical analysis. Without the help of all my committees, I could not complete my dissertation successfully.

I want to take this opportunity to express my appreciation to Dr. Bailian Li and Dr. Barry Goldfarb for their generous help and advice during my PhD study. Special acknowledgments are to our research assistant Saul Garcia for his support and help in lab techniques. I also would like to thank all of the people in our program: Dr. James Grissom, Patrick Cumbie, Tori Batista-Brooks, and Josh Steiger for their assistance during my research.

I am indebted to those who helped me for my field work, especially Ben Smith, Lizzie Wood, Jesus Espinoza, George Khan, and Nhora Isaza. I deeply appreciate Mr. Dave Gerwig for his general assistance for assessing the clonal trials in MeadWestvaco (now ArborGen). I also want to thank Xiaohua Gong, Jianyun Zhu, and Xin Wang for helping me with my data analysis.

My deepest gratitude is to my parents, Shiguang Xiong and Bingmei Tang for their boundless love and constant encouragement. Last, I would like to thank my husband Chang Zuo for his love and strong support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Chapter 1. Genetic Variation of Stem Forking in Loblolly Pine

Jin S. Xiong, Fikret Isik, Steven E. McKeand, and Ross W. Whetten

(This chapter has been accepted by *Forest Science*)

Abstract

Forking defects were studied in loblolly pine (*Pinus taeda* L.) in 6-year-old genetic tests from the second cycle of breeding in the North Carolina State University Cooperative Tree Improvement Program. Data were available from 268 test series (each with two 6-parent half-diallel mating designs, with 12 parents, 30 crosses, and 144 progeny per cross) over 6 geographic regions. A subset of 123 series with average forking between 20% and 80% was used for genetic analysis. Forking differed significantly among regions, with the highest in the Upper Gulf Coastal Plain (24%) and the lowest in the Georgia-Florida Coastal Plain (12%). The individual-tree heritabilities were low (0.06), but the half-sib (0.76), narrow-sense full-sib (0.59), and broad-sense full-sib family-mean heritabilities (0.71) were moderately high. This suggests that forking is partially under genetic control at the family level, but at the individual tree level, it is mostly determined by the environment. Genetic gain was estimated for half-sib family selection with selection differential of 20%. Gain ranged from 12% to 23% reduction in forking across different regions. A weak unfavorable genetic correlation (0.18) was found between forking and height, suggesting that selection for growth alone will negatively impact forking in loblolly pine. A favorable genetic correlation (0.33) was found between forking and straightness.

Key words: Diallel mating design, heritability, genetic response, *Pinus taeda*

1.1. Introduction

Loblolly pine (*P. taeda* L.) is the most important commercial tree species in the southeastern United States, and is widely planted by landowners and forestry companies in the South (Schultz 1999). The NC State University Cooperative Tree Improvement Program has been breeding loblolly pine for more than 50 years with the objective of selecting and breeding loblolly pine trees with faster growth, disease resistance, and better stem straightness, but little attention has been paid to forking defects (McKeand et al. 2003, Li et al. 2000, McKeand and Bridgwater 1998).

Forking can be a serious stem-quality defect and has been found in many conifer species when grown in plantations. In seven-year-old open-pollinated patula pine (*P. patula*) progeny, bole forking was as high as 49.5% in Colombia (Ladrach and Lambeth 1991); Herrera's pine (*P. herrerae*), a potential plantation species for the subtropics, showed an average forking rate of 57% in Brazil (Dvorak et al. 2007); and 45% of coastal Douglas-fir (*Pseudotsuga menziesii* var. *menziesii* (Mirb.) Franco) individuals had one or more forking defects at age 10 (Adams and Bastien 1994). Forking defects include forked stems and major ramicorn or steep-angled branches (Schermann et al. 1997). Forked stems create large knots and irregular grain in lumber and also reduce the stem strength and timber uniformity. Ramicorn branches also produce large knots thus decreasing self-pruning and healing of branch scars (Barber 1964). Forking defects greatly decrease the wood quality, timber harvest and pulp yield, correspondingly reducing the economic value of the wood.

Stem defects have long been recognized by tree breeders. There is a general belief that environmental sources of variation are the major component of forking (Savill et al. 1999). Some foresters have assumed that forking is mainly caused by late-season frosts, because the frosts tend to damage the newly emerging buds and young leaves (Kerr 1995). Frost damage could lead to forking when late season flushing exposes the tender tissues of the main leader to early freezes. The death of the main leader could cause two or more buds in the topmost whorl to flush co-dominantly, thus creating forked stems (Franklin and Callaham 1970).

Similarly, leader breakage or dieback can cause forking. The origins of top breakage could be any accidental damage, such as birds perching, animal attack, wind damage, or pathogenic agents (Kerr and Boswell 2001). Additionally, physiological control of second flushing and apical dominance in terminal shoots were potentially involved in the causes of forking. Adams and Bastien (1994) found that the frequency of second flush of terminal buds had a strong positive genetic correlation with forking. From the study of lodgepole pine (*P. contorta* var. *murrayana* Engelm.), in the absence of mechanical damage, forks seemed to result from the loss of apical dominance under physiological control. Some silvicultural practices, such as wide spacing, heavy thinning, or soil nutrient status (in particular, a lack of boron) can also lead to stem form defects (Barber 1964, Zobel and Kellison 1978). A study of nutrition effects in loblolly pine showed that fertilized trees had a greater amount of forking than non-fertilized trees (Espinoza 2009).

Some investigators postulated that genetic control of forking may exist among trees. Assessments conducted in common ash (*Fraxinus excelsior*, L.) showed significant differences in forking among families (Mwase et al. 2008). Franklin and Callaham (1970) reported that family differences in the number of forks per tree and the number of forks per whorl were significant in lodgepole pine. Provenance differences and family variation in forking were reported in loblolly pine (Espinoza 2009). Overall there have been few studies of the genetic control of forking in loblolly pine.

Forking is a binary trait with dichotomous phenotypic characters (forked or not forked). For genetic analysis of such traits, it is common to assume an underlying continuously-distributed liability, with a threshold that imposes a discontinuity in the phenotypic expression (Lynch and Walsh 1998). An individual tree is assumed to have one phenotype (e.g. forked) when the underlying variable is above the threshold level; and the alternative phenotype (e.g. non-forked), when the underlying variable is below the threshold (Falconer and Mackay 1996). The value of each underlying liability is presumed to be normally distributed, with both genetic and environmental components (Dempster and Lerner 1950). The underlying variables for forking are not clear. They could include the levels of growth regulators controlling apical dominance and the contents of boron or other nutrients that could cause dieback. Genetic analysis of threshold traits can be realized by transforming the incident percentage to the mean liabilities, even without knowing what the underlying liabilities really are (Van Vleck 1972).

The techniques for estimating heritability in threshold traits are much more complex than for continuously variable traits. Several approaches have been developed to obtain the genetic parameters of threshold traits (Dempster and Lerner 1950, Falconer 1965, Van Vleck 1972, Gilmour et al. 2006, Littell et al. 1996, Krebs and Loeschcke 1997, Lourens et al. 1999, Gianola and Foulley 1983, Casellas and Piedrafita 2006). Observed-scale analysis deals with the raw data for p scale analysis (Robertson and Lerner 1949). The heritability is estimated by given genetic gains and the selection differential. Dempster and Lerner (1950) suggested converting the observed scale to an underlying continuous scale to estimate the heritability of the threshold trait. This method has been used in many genetic and breeding studies. Other approaches of data transformation have been developed by researchers, such as Bayesian methods (Damgaard and Korsgaard 2006), assuming random effects are distributed as beta-binomial, and a linear mixed model proposed by Gianola to evaluate the genetic merit for a binary trait in sires. In this model, the logits of plot probabilities are assumed to follow a linear mixed model (Gianola and Foulley 1983)

In this study, we examined the genetic and environmental components of stem forking in loblolly pine. The large dataset and wide range of test sites enabled us to address the following questions: 1) Does the geographic region have significant effects on forking? 2) How much genetic control exists for forking defects? 3) What are the genetic correlations between forking and other traits? 4) How much gain could be made by selecting against forking, and what breeding strategy should we use to gain the maximum genetic value?

1.2. Materials and methods

1.2.1 Materials and data

This forking study used data from the progeny tests of the second-cycle breeding program of the North Caroline State University Cooperative Tree Improvement Program (Xiang et al. 2003a). The mating design was a disconnected-half-diallel of 6 parents to produce 15 full-sib crosses in each half-diallel, without self and reciprocal crosses. Two disconnected half-diallels comprise a test series, consisting of 12 parents (30 full-sib crosses). Each test series was planted as a randomized complete block design on 4 test sites with 6 blocks per site and 6-tree row plots for each cross in each block.

At age six years, height, straightness and forking were assessed for each tree. Straightness was assessed using a score from 1 to 6, with 1 being the straightest trees and 6 the most crooked trees. Forking was scored as a binary trait, with a zero for trees without a forked stem, and a one for trees with a forked stem. Both ramicorn branches and stem forks were counted as forks. Ramicorn branches were defined as those with branch angle less than 30° from the main stem with a diameter distinctly less than the main stem but at least twice as big as the branches around it. Fork was defined as a separation of the trunk into two or more stems with similar sizes.

Overall, data from 268 test series were used to investigate the variation of forking over geographic regions. Those regions were grouped as 1) Virginia and Northern North Carolina Coastal Plain and Piedmont, 2 & 3) South Carolina and southern North Carolina

Coastal Plain, 4) Georgia-Florida Coastal Plain, 5) Lower Gulf Coastal Plain of Alabama and Mississippi, 6) Upper Gulf Coastal Plain of Alabama and Mississippi, and 7 & 8) Georgia, South Carolina, and southern North Carolina Piedmont (Figure 1.1). Of the 268 test series, 123 had forking rates between 20% and 80%. Rates of forking less than 20% or greater than 80% would contribute excessive environmental variance to the analysis and were excluded from the mixed model analysis (Lopes et al. 2000).

1.2.2 Statistical analysis

Testing for significant differences in the incidence of forking among regions was conducted using the ANOVA procedure of SAS (SAS Institute Inc. 1996). Product-moment correlations between site mean height and forking percentages were estimated by using the CORR procedure in SAS (SAS Institute Inc. 1996).

Height and straightness were analyzed as continuously distributed variable traits. Forking is a threshold trait with only two categories (forked or not forked), therefore the probability of forking was modeled with a generalized linear mixed model using a logit link function. For each test series, the following model was used:

$$\eta_{ijklm} = \log [P/(1-P)] = \mu + T_i + B(T)_{j(i)} + G_k + G_l + S_{kl} + TG_{ik} + TG_{il} + TS_{ikl} + P_{ijkl} + E_{ijklm}$$

where ,

η_{ijklm} is the *m*th observation for the *kl*th cross in *j*th block within *i*th test;

P is the proportion of forked trees;

μ is the conditional mean;

T_i is the *i*th fixed test environment (location) effect, *i*=1 to *t*;

$B(T)_{j(i)}$ is the fixed effect of the *j*th block within *i*th test, *j*=1 to *b*;

G_k, G_l are the random general combining ability (GCA) effect of the *k*th female or the *l*th male parent, *k,l*=1 to *p*, *k*<*l* ~ NID (0, σ^2_g);

S_{kl} is the random specific combining ability (SCA) effect of the *k*th and *l*th parents ~NID (0, σ^2_s);

TG_{ik}, TG_{il} are the random interaction effect of the *i*th test by the *k*th female or the *l*th male ~NID (0, σ^2_{gt});

TS_{ikl} is the random interaction effect of the *i*th test by the *kl*th cross ~NID (0, σ^2_{st});

P_{ijkl} is the random plot effect for the *kl*th cross within the *i*th block ~NID (0, σ^2_p);

E_{ijklm} is the random within plot error term ~NID (0, σ^2_e);

It is convenient to write the generalized linear mixed models in matrix notation as $\mathbf{y} = \boldsymbol{\eta} + \mathbf{e} = \mathbf{X\beta}+\mathbf{Zu}+\mathbf{e}$ to define the expectations of variances and covariances, where $\mathbf{y}$ is the vector of observations and $\boldsymbol{\eta}$ is a linear predictor ($\boldsymbol{\eta}=\mathbf{X\beta}+\mathbf{Zu}$, expressed through an inverse link function). The link function for the model is $\boldsymbol{\eta} = \mu/(1-\mu)$; which relates the

linear predictor ($\boldsymbol{\eta}$) scale to the observation ($\boldsymbol{\mu}$=E[y]) scale (Gilmour et al. 2006). **X** is the design matrix for the fixed effects; $\boldsymbol{\beta}$ is the vector of fixed effect parameters; **Z** is the design matrix for the random effects; **u** is the vector of random effects distributed MVN (0, **G**); **e** is vector of error distributed MVN (0, **R**) (Littell et al. 1996). Variance components and their functions for forking were obtained by using ASReml software (Gilmour et al. 2006).

1.2.3 Genetic parameters

Using the variance components, different types of heritabilities were calculated according to the following formulas (Xiang et al. 2003a, Jansson and Li 2004).

Narrow-sense individual-tree heritability:

$$h_i^2 = \frac{4\sigma^2_g}{2\sigma^2_g + \sigma^2_s + 2\sigma^2_{gt} + \sigma^2_{st} + \sigma^2_p + \sigma^2_e}$$

Broad-sense individual-tree heritability:

$$H_i^2 = \frac{4\sigma^2_g + \sigma^2_s}{2\sigma^2_g + \sigma^2_s + 2\sigma^2_{gt} + \sigma^2_{st} + \sigma^2_p + \sigma^2_e}$$

Half-sib family-mean heritability:

$$h^2_{hs} = \frac{\sigma^2_g}{\frac{p\sigma^2_g}{p-1} + \frac{\sigma^2_s}{p-1} + \frac{p\sigma^2_{gt}}{t(p-1)} + \frac{\sigma^2_{st}}{t(p-1)} + \frac{\sigma^2_p}{bt(p-1)} + \frac{\sigma^2_e}{bnt(p-1)}}$$

Narrow-sense full-sib family-mean heritability:

$$h^2_{fs} = \frac{2\sigma^2_g}{2\sigma^2_g + \sigma^2_s + \frac{2\sigma^2_{gt}}{t} + \frac{\sigma^2_{st}}{t} + \frac{\sigma^2_p}{bt} + \frac{\sigma^2_e}{bnt}}$$

Broad-sense full-sib family-mean heritability:

$$H^2_{fs} = \frac{2\sigma^2_g + \sigma^2_s}{2\sigma^2_g + \sigma^2_s + \frac{2\sigma^2_{gt}}{t} + \frac{\sigma^2_{st}}{t} + \frac{\sigma^2_p}{bt} + \frac{\sigma^2_e}{bnt}}$$

Where σ^2_g is genetic variance of GCA, σ^2_s is genetic variance of SCA, σ^2_{gt} is test by GCA interaction variance, σ^2_{st} is test by SCA interaction variance, σ^2_p is plot variance, σ^2_e is within plot error variance, *p* is the number of parents in the diallel set, *t* is the number of test sites within each series, *b* is the number of blocks at each test site and *n* is the number of trees within each plot for each cross. The logit transformation maps the probability to an underlying scale using the logistic distribution with the variance of 3.29. This is the variance of logit distribution and used for all the heritability estimates as the error variance (Lynch and Walsh 1998).

Genetic correlations were estimated between height and forking and between forking and stem straightness by fitting a bivariate model and using the following formula:

$$r_{gxy} = \frac{\sigma_{gxy}}{\sqrt{\sigma^2_{gx}\sigma^2_{gy}}}$$

Where σ_{gxy} is the genetic covariance between trait x and trait y; σ^2_{gx} and σ^2_{gy} are the genetic variances of trait x and trait y, respectively.

Variance components estimated for small numbers of parents might be biased and heritability estimates would be biased. Heritabilities and genetic correlations were calculated by using the means of the estimated variance components in the specific region to minimize the biases of estimation (Isik et al. 2005). The standard error of heritabilities and correlations were estimates using the Delta Method and using a SAS code (Isik et al. 2008a).

The predicted probabilities ($\hat{P}$) of forking were calculated by applying the inverse link function of the best linear unbiased prediction (BLUP) of random solution vector (μ). Genetic gains based on breeding values of half-sib families were estimated. A selection differential of 20% within region was used for all seven test regions.

1.3. Results and Discussion

1.3.1 Variation among regions

The forking percentage across all 268 test series averaged 18%, and it ranged from 4% to 80% across different test series. At specific test sites, the forking percentage varied more widely, with a maximum of 92% and a minimum of 0%. Forking differed significantly across regions (F-test $Pr \leq 0.001$). The highest forking was in region 6-Upper Gulf (24%), and the lowest was in Region 4-Coastal Georgia & Florida (12%) (Figure 1.2). Most of

the regions significantly differed with each other in forking; only regions 1 and 7 did not differ.

There is an apparent increase in forking from the coast to inland regions, and from south to the north. Possible causes could be soil fertility, weather effects such as total precipitation, late and early frosts during the growing seasons, or ice and storm damage, and genetic differences among seed sources. In order to see if the site fertility or other site productivity factors significantly influenced the forking rate, the average height for each site was used as an estimator of site quality to study the relationship between site quality and forking rate. Across all 268 series, the average height showed no relationship with forking rate (Pearson correlation $r=0.009$ $Pr>0.88$). We conclude that our data do not support the hypothesis that site quality plays a role in determining forking rate.

High forking in regions 6-Upper Gulf and 7-Piedmont could be explained by more cold damage or exposure to ice and snow, although we had no separate assessment of cold or ice and snow damage. These two regions are in the Piedmont and inland areas where snow and ice damage are more frequent and severe compared to more coastal populations. A similar speculation can also be valid for the region 1-Virginia and Northern North Carolina. This region is not ‘inland’ but represents the most northern distribution of loblolly pine populations with more frequent frosts and ice storms compared to the southern regions.

1.3.2 Heritability

Based on variance components across the 123 test series with forking rates between 20% and 80%, the narrow-sense and broad-sense individual-tree heritabilities for forking averaged 0.05 and 0.06, respectively (Table 1.1). Estimated average half-sib family-mean heritability for forking was 0.76 and ranged from 0.73 to 0.81 across the regions (Table1.1), while the average narrow-sense full-sib family-mean heritability was 0.59, ranging from 0.55 to 0.75. The average broad-sense full-sib family mean heritability was 0.71 ranging from 0.66 to 0.75. Here, additive effects (GCA) were much larger than dominance or SCA effects, with the ratio of 2.4 for GCA/SCA over all regions.

Regional individual-tree heritabilities were generally comparable to each other (Table 1.1). We observed slightly higher half-sib family mean heritability values for Region 1 and 6. The trends between regions for narrow-sense full-sib family-mean heritability were more pronounced. Regions 1, 3 and 6 had higher heritabilities than others. Higher family-mean heritabilities in regions 1 and 6 could be explained by higher forking incidence means in these regions. A mean incidence of 0.5 allows greater precision of expression of genetic and environmental factors for a binary trait in a population (Lynch and Walsh 1998). Variance components used to calculate heritabilities in this study were obtained from a generalized linear mixed model. In such models, the chosen canonical link function transforms that binary data to underlying continuous distribution (Gilmour et al. 1985). As the mean incidence approaches 0.5 in the population, the relationship between heritability on the observed and transformed scale increases (Isik et al. 2008b).

Many authors (e.g., Gilmour et al. 1985) do not recommend genetic analysis of threshold traits outside the range of mean incidence from 0.30 to 0.70, because environmental noise increases considerably beyond the suggested boundaries. In other words, because the variance of threshold trait is mean dependent, the relatively low forking incidences in some regions in this study might have caused lower estimates of the heritability in those regions. In this study, we included test series in analyses with minimum of 20% and a maximum of 80% forking, to have a large sample of sites across the regions.

In the literature, broad-sense and narrow-sense individual heritabilities reported for other conifer species were generally low, as we found in this study. In radiata pine (*P. radiata*), the narrow-sense heritability of malformation (forking) was found as 0.03 in polycross progenies and 0.04 in open-pollinated progenies (Jayawickrama and Low 1999). A similar heritability was found in Jayawickrama's later study in radiata pine as 0.1 within site and 0.03 across sites (Jayawickrama 2001). In a recent report for radiata pine, the individual heritability of forking was 0.04 in Spain (Codesido and Fernández-López 2008). In Herrera's pine, single-site heritability values were reported in South Africa (0.10) and Brazil (0.06) (Dvorak et al. 2007). In another study, the individual-tree heritability for forking was relatively high (0.32) for Chiapas pine (*P. chiapensis*) in Columbia (Dvorak et al. 1996).

The individual's underlying 'liability' to forking expresses the combination of genetic tendency to develop a fork and the external environment that makes it likely to develop the fork. Here, the low individual-tree heritability of forking suggested the random

environmental effects could be the major underlying cause for observing the trait on the normal scale. However, for threshold traits, an individual's phenotype on the underlying scale (liability) is estimated very imprecisely (Falconer and Mackay 1996). In contrast to individual tree phenotypic values, the mean of the family is much more precisely known from the proportion of the forking incidence. Family-mean heritabilities should therefore be more reliable than individual-tree heritabilities. The values of family mean heritabilities in this study are greater than the results reported in previous studies of forking. Family mean heritability for forking was 0.41 in radiata pine (Codesido and Fernández-López 2008) and was 0.38 in *Pinus oocarpa* (Moura et al. 1998). Ladrach and Lambeth (1991) found that the family-mean heritability for bole forking was 0.46 in patula pine progeny tests, and forking tendency on boron-deficient soils varied by genotypes. Family differences in forking were significant, and the genetic control of forking was assumed to be a primary causal factor for forking in lodgepole pine. (Franklin and Callaham 1970).

The broad-sense heritabilities (family-mean or individual) were higher than the corresponding narrow-sense heritabilities, because the broad-sense heritability included the dominance (SCA) component as part of the genetic variance. The ratio of GCA/SCA variance (2.4) in this analysis is in agreement with earlier studies of these diallel tests for growth traits (Xiang et al. 2003a, Isik et al. 2003). This ratio indicates that the additive effects are the major source of genetic variance in forking, which should be exploited in the breeding programs to reduce forking. High family-mean heritabilities suggest that

selection among families to reduce the average rate of forking in loblolly pine would be an efficient strategy.

1.3.3 Genetic correlations

Genetic correlations between forking and height were unfavorable and ranged from 0.12 to 0.29 across regions with the overall mean of 0.18 (Table 1.2). Correlation in the South Carolina Coastal Plain region (region 3) was relatively high (r_G=0.29) compared to other regions. The genetic correlation of forking and straightness showed more variation among different regions, ranging from 0.14 to 0.46. Overall, the genetic correlation of forking and straightness is favorable (i.e. straight trees tend to have fewer forks) with the value of 0.33.

The unfavorable genetic correlation between forking and height suggested that selection for height would result in a slight increase in forking. In a few studies on forking, unfavorable genetic correlations were reported. The correlation of height and forking was reported as 0.53 in Douglas-fir (Adams and Bastien 1994) and 0.26 in common ash (Mwase et al. 2008). Forking was found to have moderately unfavorable correlations with volume and height in loblolly pine (Cumbie et al. 2010). The correlation between height and forking suggested that selection based on height alone would indirectly increase forking defects in the next generation (Vargas-Hernandez et al. 2003). Franklin and Callaham (1970) reported no relationship between seedling height and forks in lodgepole

pine, but they did find highly significant differences among years for number of forks per major whorl, so that older trees had more forks .

The regional differences in the genetic correlation between forking and straightness might be caused by inconsistency of the measurement of forking and straightness over regions. Both of these traits depend on the judgment of the measurement crews, so differences between crews might have contributed to the error variance. In addition, sampling error in choice of the parents may contribute to the great variation of genetic variance and covariance. The favorable correlation suggested that straight trees have a tendency not to fork, but we suspect that measurement crews tended to call most forked trees as below average for straightness, which may have caused positive correlation between forking and straightness.

1.3.4 Genetic response

The highest predicted genetic gain was in region 6 with 23% expected reduction in forking (Figure 1.3). The lowest gain was 12% expected reduction of forking in region 3 (Figure 1.3). Selection at the family level against forking should be more effective in breeding programs, as for threshold traits (Falconer and Mackay 1996). Over all regions, half-sib family selection could reduce forking by 21%. The genetic gains derived from the BLUP analysis are not like the predicted genetic gain based on the genetic parameters (Xiang et al. 2003b), but similar to realized gains (Jansson and Li 2004). The result showed that it should be possible to decrease forking defects in loblolly pine through

family selection. A study in Douglas-fir (Vargas-Hernandez et al. 2003) also showed that selection could lower the forking percentage. Through selection at age 12, the number of whorls with steep-angle branches in Douglas-fir was reduced 24% at age 24.

1.4. Breeding implication

These results suggest that forking defects are under genetic control at the family level, and that forking should be included in selection indexes for loblolly pine improvement programs. Unfortunately, this has not generally done, because of the belief that the stem form defects are mostly under environmental control, and the inconsistency of the phenotypic measurement on these characters (Zobel and Jett 1995).

Drenou (2000) classified forks into four major categories: temporary forks, recurrent forks, main forks and accidental forks. For loblolly pine, it seems more reasonable to divide the forking into two categories: accidental forks, caused by unpredictable factors such as wind damage, animal attack or other physical injuries; and persistent forks, more probably under genetic control with multiple underlying causes. The underlying causes could be physiological control of apical dominance, insect or disease resistance, frost hardiness or stem size and strength, etc.

It is probably more difficult to decrease the incidence of accidental forks through breeding since this type of forking is assumed to be controlled by the environmental factors. However, it should still be possible to select families and progenies that do not respond negatively to these environmental factors. Conversely, the families that recover

more quickly or completely from the environmental influences can also be considered in the selection strategy. For persistent forking, it should be easier to select families and progenies with a lower genetic tendency to form these forks. Based on high family-mean heritability values, it seems likely that family-based selection against forking can provide good genetic gain. The positive correlation between forking and height means that selection against forking alone would slightly decrease the height, and selection for faster growth alone would increase the frequency of forking (Adams and Bastien 1994, Schermann et al. 1997). Fortunately, the genetic correlation between height growth and forking is weak enough that both traits can be improved simultaneously. With an index selection criterion considering both traits, we can maximize the height while reducing forking, or at least holding the forking to no change (Adams and Bastien 1994, Cumbie et al. 2010).

It is likely that traits such as frost hardiness, resistance to pathogens or insects, and physiological control of apical dominance genetically influence the occurrence of forking defects. Hence, improving resistance to diseases and insects such as pitch canker (*Fusarium circinatum*) and tip moth (*Rhyacionia frustranum*) may indirectly decrease the incidence of forking.

1.5. Conclusion

This analysis suggests that forking is under genetic control at the family level in loblolly pine but not at the individual tree level. The correlations between forking and height, and

forking and straightness, are relatively low. Half-sib family selection against forking would be an efficient strategy to reduce incidence of stem forking, which would be significant in improving stem quality in loblolly pine breeding programs.

Acknowledgements

This work was supported by the North Carolina State University Cooperative Tree Improvement Program of the Department of Forestry and Environmental Resources, NCSU. We appreciate the previous Cooperative staff members who contributed in test design and data collection.

1.6. References

Adams, W.T. and J.C. Bastien. 1994. Genetics of second flushing in a French plantation of coastal Douglas-fir. *Silvae Genet.* **43**(5):345-352.

Barber, J.C. 1964. *Inherent variation among slash pine progenies at the Ida Cason Callaway foundation.* U.S. For. Serv. Res. Pap. Stheast .For. Exp. Sta. **10**. 90 p.

Casellas, J. and J. Piedrafita. 2006. Bayes factor for testing the genetic background of quantitative threshold traits. *J. Anim. Breed. Genet.* **123**(5):301-306.

Codesido, V. and J. Fernández-López. 2008. Juvenile genetic parameter estimates for vigour, stem form, branching habit and survival in three radiata pine (*Pinus radiata* D.Don) progeny tests in Galicia, NW Spain. *Eur. J. For. Res.* **127**(4):315-325.

Cumbie, W P, F. Isik, and S E McKeand. 2010. Genetic variation in sawtimber potential in loblolly pine. *For. Sci.* In review.

Damgaard, L.H. and I.R. Korsgaard. 2006. A bivariate quantitative genetic model for a threshold trait and a survival trait. *Genet. Sel. Evol.* **38**(6):565-581.

Dempster, E.R. and M. Lerner. 1950. Heritability of threshold characters. *Genetics.* **35**:212-236.

Drenou, C. 2000. Pruning Trees: the Problem of Forks. *J. Arboric.* **26**(5):264-269.

Dvorak, W.S., J.K. Donahue and J.A. Vasquez. 1996. Provenance and progeny results for the tropical white pine, *Pinus chiapensis*, at five and eight years of age. *New Forest.* **12**(2):125-140.

Dvorak, W.S., E. Kietzka, G.R. Hodge, A. Nel, G.A.d. Santos and C. Gantz. 2007. Assessing the potential of *Pinus herrerae* as a plantation species for the subtropics. *For. Ecol. Manage.* **242**(2):598.

Espinoza,J.A.2009. *Genetic and nutritional effects on stem sinuosity in loblolly pine.* Ph.D. dissertation, North Carolina State University, Releigh, N.C., USA.102p.

Falconer, D.S. 1965. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann. Hum. Genet.* **29**:51-76.

Falconer, D.S. and T.F.C. Mackay, 1996. *Introduction to quantitative genetics.* 4th ed. Longman Group Ltd, Essex, UK. 464p.

Franklin, E.C. and R.Z. Callaham. 1970. Multinodality, branching, and forking in Lodgepole Pine (Pinus contorta var. murrayana Engelm.). *Silvae Genet.* **19**(5/6):180-184.

Gianola, D. and J.L. Foulley. 1983. Sire evaluation for ordered categorical data with a threshold model. *Genet. Sel. Evol.* **15**(2):201-224.

Gilmour, A.R., R.D. Anderson. and A.L. Rae, 1985. The analysis of binomial data by a generalized linear mixed model. *Biometrika* 72:593-599

Gilmour, A.R., B.J. Gogel, B.R. Cullis, S.J. Welham, and R. Thomson, 2006. *ASReml User Guide Release 2.0*. VSN International Ltd, Hemel Hempstead, HP1, 1ES, UK. 320p.

Isik F., H.V. Amerson, R.W. Whetten, S.A. Garcia, B. Li, and S. E. McKeand. 2008b. Resistance assessments of elite loblolly pine families to fusiform rust inocula in greenhouse testing. *Can. J. For. Res.* 38:2687-2696.

Isik, F., B.L. Li and J. Frampton. 2003. Estimates of additive, dominance and epistatic genetic variances from a clonally replicated test of loblolly pine. *For. Sci.* **49**(1):77-78.

Isik, F., D. Boos and B. Li. 2005. The distribution of genetic parameter estimates and confidence intervals from small disconnected diallels. *Theor. Appl. Genet.* **110**(7):1236-1243.

Isik, F., M. Gumpertz, B. Li, B. Goldfarb, and X. Sun. 2008a. Analysis of cellulose microfibril angle using a linear mixed model in *Pinus taeda* clones. *Can. J. For. Res.*38:1676-1689.

Jansson, G. and B. Li. 2004. Genetic gains of full-sib families from disconnected diallels in loblolly pine. *Silvae Genet.* **53**(2):60-64.

Jayawickrama, K.J.S. 2001. Genetic parameter estimates for radiata pine in New Zealand and New South Wales: a synthesis of results. *Silvae Genet.* **50**(2):45-53.

Jayawickrama, K.J.S. and C.B. Low. 1999. *Pinus Radiata* selections from different regions of New Zealand differ in branch habit, form, and growth rate. *N. Z. J. For. Sci.* **29**(1):3-24.

Kerr, G. and C. Boswell. 2001. The influence of spring frosts, ash bud moth (*Prays fraxinella*) and site factors on forking of young ash (*Fraxinus excelsior*) in southern Britain. *Forestry.* **74**(1):29-40.

Kerr, G. 1995. Silviculture of ash in southern England. *Forestry.* **68**(1):63-70.

Krebs, R.A. and V. Loeschcke. 1997. Estimating heritability in a threshold trait: heat-shock tolerance in Drosophila buzzatii. *Heredity.* **79**:252-259.

Ladrach, W.E. and C. Lambeth. 1991. Growth and heritability estimates for a seven-year-old open-pollinated *Pinus patula* progeny test in Colombia. *Silvae Genet.* **40**(5/6):169-173.

Li, B.L., S. McKeand and R. Weir. 2000. Impact of forest genetics on sustainable forestry -- results from two cycles of loblolly pine breeding in the U.S. *J. Sustain. Forest.* **10**(1):79-85.

Littell, R.C., G.A. Milliken, W.W. Stroup, and R.Wolfinger. 1996. *SAS System for Mixed Models.* SAS Institute Inc, Cary, USA.656p.

Lopes, U.V., D.A. Huber, T.L. White, E.R. Dempster, I.M. Lerner, R. Schall, R. Wolfinger and M. O'Connell. 2000. Comparison of methods for prediction of genetic gain from mass selection on binary threshold traits. *Silvae Genet.* **49**(1):50-56.

Lourens, A., G.J. Erasmus, J.B. van Wyk, S.J. Schoeman, F.W.C. Neser and M.G. Steyn. 1999. Evaluation of pelt traits in Karakul sheep applying linear and threshold models. *S. Afr. J. Anim. Sci.* **29**(1):11-14.

Lynch, M. and B. Walsh, 1998. *Genetics and analysis of quantitative traits.* Sinauer Associates Incorporated, Sunderland, MA, USA. 980p.

McKeand, S., T. Mullin, T. Byram and T. White. 2003. Deployment of genetically improved loblolly and slash pines in the South. *J. For.* **101**(3):32-37.

McKeand, S.E. and F.E. Bridgwater. 1998. A strategy for the third breeding cycle of loblolly pine in the Southeastern U.S. *Silvae Genet.* **47**(4):223-234.

Moura, V.P.G., W.S. Dvorak and G.R. Hodge. 1998. Provenance and family variation of *Pinus oocarpa* grown in the Brazilian cerrado. *For. Ecol. Manage.* **109**(1-3):315-322.

Mwase, W.F., P.S. Savill and G. Hemery. 2008. Genetic parameter estimates for growth and form traits in common ash (*Fraxinus excelsior*, L.) in a breeding seedling orchard at Little Wittenham in England. *New Forests.* **36**(3):225-238.

Robertson, A. and I.M. Lerner. 1949. The heritability of all-or-none traits: viability of poultry. *Genetics.* **34**:395-411.

SAS Institute Inc. 1996. SAS/STAT: *Changes and enhancement (through release 6.11)*. SAS Institute Inc., Cary, N.C.

Savill, P.S., R. Spencer, J.E. Roberts and J.D. Hubert. 1999. Sixth year results from four ash (*Fraxinus excelsior*) breeding seedling orchards. *Silvae Genet.* **48**(2):92-100.

Schermann, N., W.T. Adams, S.N. Aitken and J.C. Bastien. 1997. Genetic parameters of stem form traits in a 9-year-old coastal Douglas-fir progeny test in Washington. *Silvae Genet.* **46**(2/3):166-170.

Schultz, R.P. 1999. Loblolly--the pine for the twenty-first century. *New Forests.* **17/18**(1):71-88.

Van Vleck, L.D. 1972. Estimation of heritability of threshold characters. *J. Dairy Sci.* **55**:218-225.

Vargas-Hernandez, J.J., W.T. Adams and D.G. Joyce. 2003. Quantitative genetic structure of stem form and branching traits in Douglas-fir seedlings and implications for early selection. *Silvae Genet.* **52**(1):36-44.

Xiang, B., B. Li and F. Isik. 2003a. Time trend of genetic parameters in growth traits of *Pinus taeda* L. *Silvae Genet.* **52**(3/4):114-121.

Xiang, B., B. Li and S. McKeand. 2003b. Genetic gain and selection efficiency of loblolly pine in three geographic regions. *For. Sci.* **49**(2):196-208.

Zobel, B.J. and Jett, J.B. 1995. Genetics of wood production. Springer-Verlag. Berlin: 337p.

Zobel, B.J. and R.C. Kellison. 1978. The rate of growth syndrome. *Silvae Genet.* **27**(3):123-124.

Table 1.1 Individual-tree and family-mean heritabilities ± standard errors for forking in different regions.

Region	# of series	h^2_i	H^2_i	h^2_{fs}	H^2_{fs}	h^2_{hs}
1-VA & North NC	21	0.06±0.07	0.06±0.03	0.65±0.21	0.74±0.19	0.81±0.16
3-Coastal NC&SC	10	0.05±0.04	0.05±0.02	0.75±0.12	0.66±0.09	0.73±0.13
4-Coastal GA&FL	5	0.05±0.03	0.05±0.02	0.57±0.14	0.72±0.13	0.74±0.16
5-Lower Gulf	12	0.04±0.03	0.04±0.04	0.55±0.13	0.67±0.14	0.74±0.12
6-Upper Gulf	28	0.06±0.04	0.07±0.03	0.65±0.14	0.75±0.11	0.79±0.09
7-Piedmont	47	0.05±0.03	0.05±0.04	0.55±0.13	0.69±0.15	0.74±0.15
Overall	123	0.05±0.04	0.06±0.03	0.59±0.15	0.71±0.13	0.76±0.13

h^2_i = Narrow-sense individual-tree heritability

H^2_i = Broad-sense individual-tree heritability

h^2_{fs} = Narrow-sense full-sib family-mean heritability

H^2_{fs} = Broad-sense full-sib family-mean heritability

h^2_{hs} = Half-sib family-mean heritability

Table 1.2 Genetic correlations between forking & height, and forking & straightness (standard errors are in parentheses) in different regions.

Region	# of series	r_g_fork-height	r_g_fork-strt
1-VA & North NC	21	0.18 (0.12)	0.14 (0.12)
3-Coastal NC&SC	10	0.29 (0.15)	0.33 (0.11)
4-Coastal GA&FL	5	0.19 (0.13)	0.46 (0.10)
5-Lower Gulf	12	0.12 (0.10)	0.35 (0.11)
6-Upper Gulf	28	0.15 (0.12)	0.39 (0.10)
7-Piedmont	47	0.14 (0.13)	0.31 (0.11)
Overall	123	0.18 (0.16)	0.33 (0.11)

r_g_fork-height = genetic correlation between forking and height

r_g_fork-strt=genetic correlation between forking and straightness

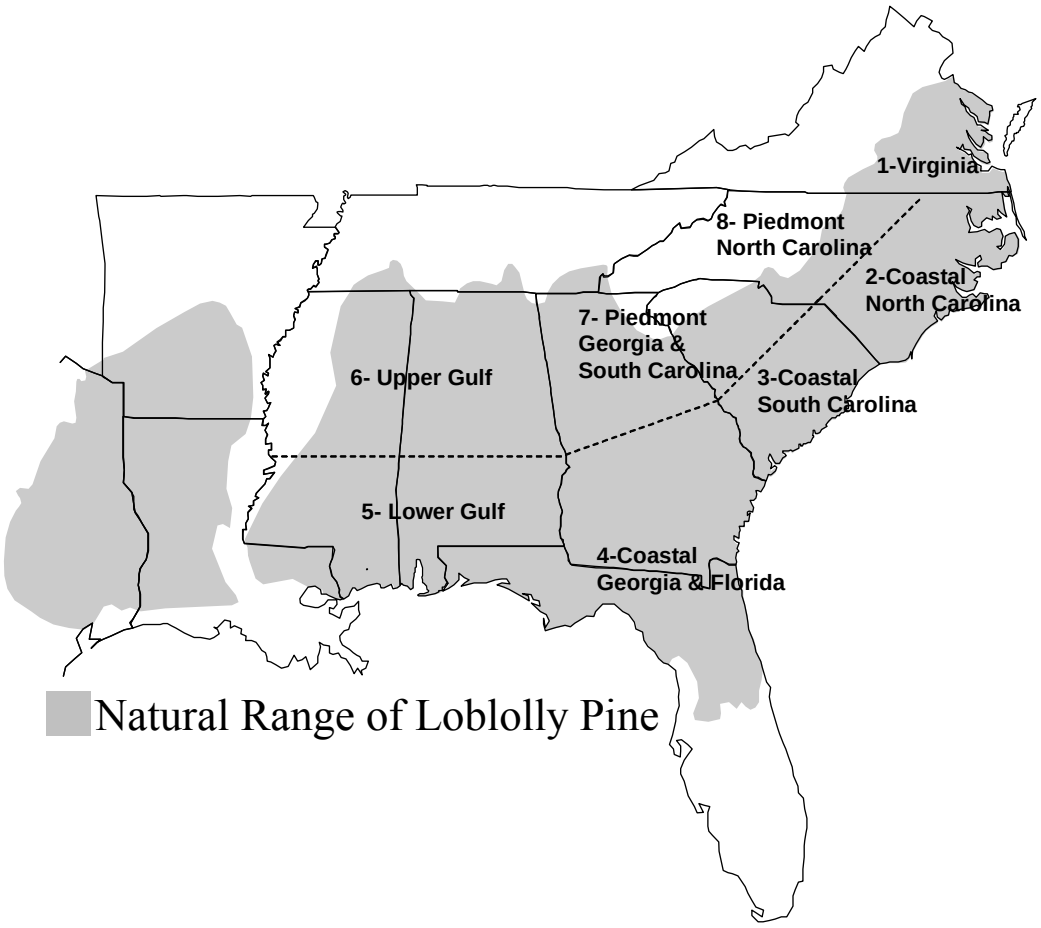


Figure 1.1 Natural range of loblolly pine showing the general locations of the test areas used in the second cycle of testing in the NC State Tree Improvement Cooperative. For these analyses, tests in areas 2 and 8 were combined with area 3 and 7, respectively.

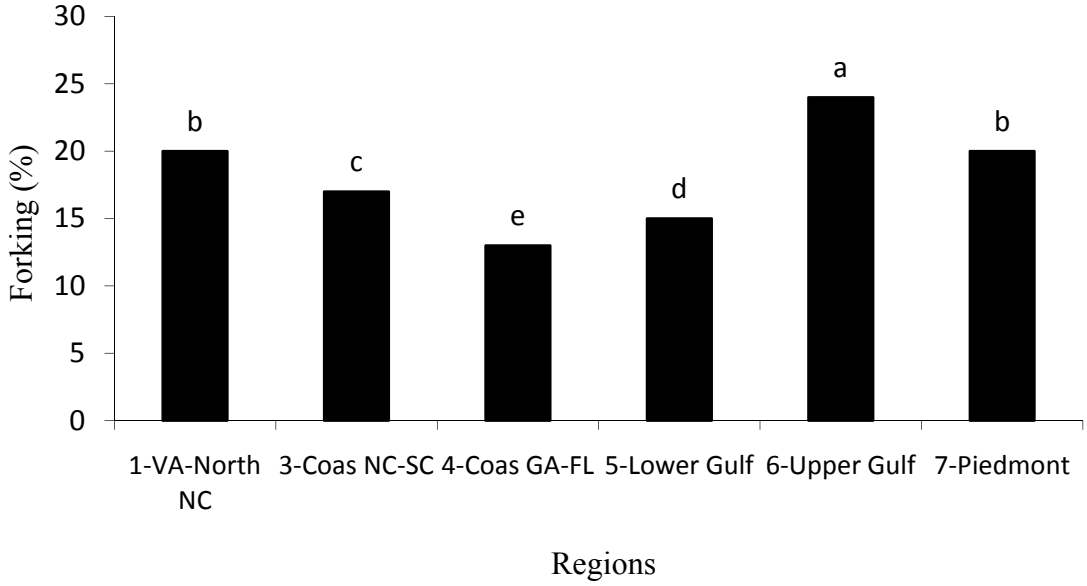


Figure 1.2 Forking (%) in different regions in the southeastern US. Values followed by different letters differ significantly at P=0.001.

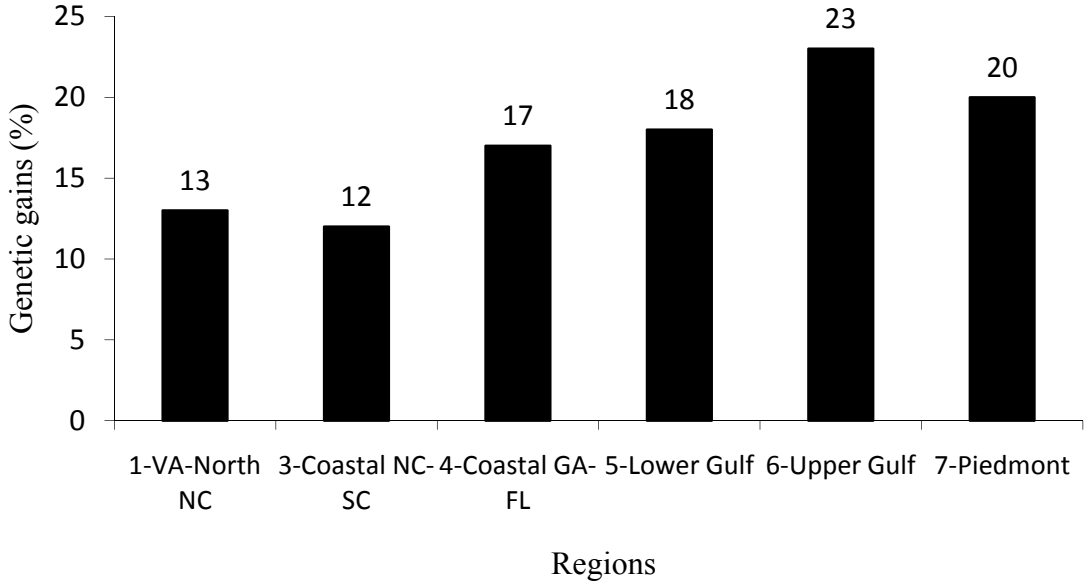


Figure 1.3 Genetic gain in forking from half-sib family selection with selection differential of 20% within regions.

2. Chapter 2. The Genetics of Growth and Forking from a Loblolly Pine Clonal Test

Jin S. Xiong, Steven E. McKeand, Ross W. Whetten, and Fikret Isik

Abstract

Forking defects and ramicorn branching were assessed in a cloned full-sib family of loblolly pine (*Pinus taeda* L.). The proportion of forking ranged from 13% to 21% across test sites with a mean of 17%. Ramicorn branching incidence averaged 24% and varied from 18% to 31% over all sites. There were significant differences among clones within this family for all traits studied. The estimated clone-mean repeatabilities were moderately high for forking defects (0.67 to 0.86). Through single trait selection, the proportion of trees with forks and ramicorn branches could be reduced 10% and 13%, respectively, with the selection differential of 4%. Unfavorable genetic correlations were found between growth traits and forking defects (r_g=0.12 to 0.45), suggesting that selection for either trait alone will negatively affect the genetic response for the other. No environmental correlations were found between forking and growth. A moderate positive genetic correlation between stem forking and ramicorn branching indicates that both traits may be partially controlled by the same genes and could be improved simultaneously. Strategies are discussed for selection to capture more gain for increased growth and reduced stem forking for the breeding of loblolly pine.

Key word: repeatability, clonal selection, genetic gain, *Pinus taeda*

2.1.Introduction

Forking defects, including forked stems and ramicorn branches, are a serious stem-quality problem in many conifer species (Ladrach and Lambeth 1991, Dvorak et al. 2007, Jayawickrama 2001). A stem forking is a separation of the trunk into two or more stems of similar sizes. Forked stems produce big knots and irregular grain thus increasing the difficulty of pruning and logging (Doede and Adams 1998). A ramicorn branch is a branch with twice the diameter and a substantially steeper angle than the other branches in the same whorl. Ramicorn branches also create large knots and increase the amount of compression wood in the stems (Schermann et al. 1997). Forking defects decrease stem strength and timber uniformity. Hence, the presence of forked stems and ramicorn branches greatly reduces wood quality and quantity, accordingly decreasing the economic value of the wood (Vargas-Hernandez et al. 2003).

Assessing forking in breeding populations may enable breeders to successfully breed and deploy non-forked phenotypes. Some earlier tree breeders postulated that genetic variation in forking may exist among trees (Franklin and Callaham 1970, Zobel and Jett 1995, Downs 1949), and later studies in some tree species have shown this to be true. Forking was significantly different among families in a *Pinus patula* progeny test in Colombia (Ladrach and Lambeth 1991), and provenance effects were also significant for the proportion of forked trees (Cornelius and Watt 2003). The individual-tree heritability of forking for tropical white pine (*Pinus chiapensis*) was 0.32 in Columbia (Dvorak et al. 1996); the within-family mean heritability for *Pinus oocarpa* was 0.39 (Moura et al.

1998); and the inheritance of forking for *Gmelina arborea* has been found to be under low to moderate control (Osorio 2004). In a recent study in loblolly pine, there was a moderate family-mean heritability for forking, and genetic gain could be obtained through family selection (Xiong et al. 2010a). This study was based on a large dataset, with 268 half-diallel test series over six geographic regions of the southeastern United States (total 3216 parents, 8040 crosses).

Most previous studies on forking defect used seedling trials with a single measurement per genotype, and resulted in low individual-tree heritabilities. The technology for vegetative propagation of forest trees provides the potential of using clonally replicated progeny to assess forking problems. When individual trees are replicated clonally, each individual contribution can be tested in different environments (Isik et al. 2005). The repeated observations provide more precise estimates of genetic parameters by reducing the environmental noise (Amerson et al. 2008, Baltunis et al. 2009). Additional precision in testing individuals in clonal trials allows better utilization of within family variation (both additive and non-additive) to capture more genetic gains (Stelzer and Goldfarb 1997, Shaw and Hood 1985). For threshold traits such as forking, clonal trials with many ramets per clone are particularly valuable for increasing the precision and reliability of the observations for individuals.

The current study is unique and uses a loblolly pine clonal trial with one family but a large number of clones and ramets per clone to assess forking traits in a cross. The objectives of the study were to examine the stem forking and ramicorn branches defects

in this cross, detect the genetic variation of forking traits within family, and develop selection strategies to breed trees with low frequencies of forks and ramicorn branches.

2.2. Material

2.2.1 Genetic materials and experimental design

Full-sib seedlings of family AF were hedged (cloned) to produce planting stock. Cuttings from each hedge were collected in the spring of 1999 and stuck in RLC7 Stubby Cells^{TM} (115cc) containing a peat-vermiculite-perlite medium. Rooted cuttings of 223 clones were transplanted at the Colleton County site in the fall of 1999. Cuttings from subsequent harvests and seedlings were used to establish the other three test sites in Berkeley, Allendale, and Charleston counties at South Carolina, U.S.A in the fall of 2000. The experimental design was a randomized complete block. Every test site had 10 blocks with a maximum of 232 clones per block using single-tree plot configuration.

2.2.2 Data Collection

All traits were measured in 2007, at age seven for the Colleton County site and at age six for the other three sites. Height was measured to the dominant leader, and DBH was measured at breast height (1.37 meters). Straightness was recorded on a 1 to 6 relative scale where 1 is straightest and 6 is most crooked. Branch angle was assessed in 15˚ increments where 1 = 15˚ (flat), 2 = 30˚, 3 = 45˚, 4 = 60˚, 5 = 75˚ (steep). If the tree was forked, it was scored as 1, otherwise as 0. Trees with ramicorn branches were scored as 1,

trees lacking ramicorn branches were scored 0. The numbers of stem forks and ramicorn branches per tree were also counted, and the height of forks was recorded for the lowest two forks. The first fork was the lowest fork from the bottom of the tree, and the second fork was the next higher fork than the first one. The volume of individual trees was calculated according to the equation developed by Goebel & Warner (Goebel and Warner 1966).

2.3.Statistical analysis

2.3.1 Growth, straightness and branch angle

Height, DBH, volume, straightness and branch angle were analyzed as quantitative traits. Variance components were estimated fitting a general linear mixed model in matrix form:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

Where **y** is the vector of individual phenotypic values; **X** is the design matrix for the fixed effects; β is the vector of fixed effect parameters; **Z** is the design matrix for the random effects; **u** is the vector of random effects distributed MVN (0, **G**); **e** is vector of error distributed MVN (0, **R**). Variance components and genetic parameters of these traits were estimated by using ASReml software (Gilmour et al. 2002) and the Mixed procedure of SAS (Littell et al. 1996). The genetic gains for continuously-distributed traits were predicted based on best linear unbiased prediction (BLUP). The significant differences of

growth traits among sites and clones were tested by F-test (alpha = 0.01) with the GLM procedure of SAS (SAS Institute Inc. 1996).

2.3.2 Stem forking and ramicorn branching

Stem forking and ramicorn branching are threshold characters with only two phenotypic classes, and are assumed to have an underlying continuous variation of liability (Falconer and Mackay 1996). The generalized linear mixed model with a canonical link function was used to analyze the forking traits in this study. Through this function, the forking incidence was transformed to a continuous distribution:

$$\eta_{ijkl} = \log [p/(1-p)] = \mu + T_i + B(T)_{j(i)} + C_k + TC_{ik} + E_{ijkl}$$

Where,

η_{ijkl} is the *l*th observation for the *k*th clone in *j*th block within *i*th test;

p is the proportion of trees showing at least one fork or one ramicorn branch;

μ is the conditional mean;

T_i is the *i*th fixed test environment (location) effect, i=1 to t;

$B(T)_{j(i)}$ is the fixed effect of the *j*th block within *i*th test, j=1 to b;

C_K is the random clone effect of the *k*th clone ~NID (0, σ^2_c);

CT_{ik} is the random interaction effect of the *i*th test by the *k*th clone

~NID (0, σ^2_{ct});

E_{ijkl} is the random error term ~NID (0, σ^2_e).

2.3.3 Genetic parameters

The repeatability of clone means ($H_{\bar{C}}^2$) was estimated by using the following formula:

$$H_{\bar{C}}^2 = \frac{\sigma^2_c}{\sigma_c^2 + \frac{\sigma_{ct}^2}{t} + \frac{\sigma_e^2}{tb}}$$

Where σ^2_c is the genetic variance among clones, σ_{ct}^2 is the interaction of test by clone, σ_e^2 is the error variance (which is assumed to be $\pi^2/3$ or 3.29) (Gilmour et al. 2002, Southey 2003), t is the number of tests and b is the number of blocks. The variance components and the standard error of clone-mean repeatability were estimated by using ASReml (Gilmour et al. 2002).

The Genetic correlations between trait x and y were computed as

$$r_{gxy} = \frac{\sigma_{gxy}}{\sqrt{\sigma^2_{gx}\sigma^2_{gy}}} ,$$

and the corresponding environmental correlation as

$$r_{exy} = \frac{\sigma_{exy}}{\sqrt{\sigma_{ex}^2 \sigma_{ey}^2}},$$

where σ_{gxy} is the clone mean genetic covariance between trait x and trait y, σ_{gx}^2 and σ_{gy}^2 are clone mean genetic variances for trait x and trait y, respectively; σ_{exy} is the environmental covariance between trait x and trait y, σ_{ex}^2 and σ_{ey}^2 are environmental variances for trait x and trait y, respectively. All correlations and approximated standard errors were estimated by ASReml software (Gilmour et al. 2002).

The predicted probabilities ($\hat{P}$) of forking and ramicorn branching were calculated by applying the inverse link function of the best linear unbiased prediction (BLUP) of random solution vector ($\hat{\boldsymbol{u}}$).

$$\hat{p} = \left[\exp(\mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\mathbf{u}})\right] / \left[1 + \exp(\mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\mathbf{u}})\right]$$

Where $\mathbf{X}$ is the design matrix for fixed effects, $\hat{\boldsymbol{\beta}}$ is the solution for fixed effects and $\hat{\mathbf{u}}$ is the solution for random effects.

Gains from multiple traits selection were predicted through index selection and independent culling selection (Van Vleck 1970, Lynch and Walsh 1997) by using the IML procedure of SAS (SAS Institute Inc. 1996, Xu and Muir 1991). The variance and covariance matrix from different traits were predicted by ASReml software (Gilmour et al. 2002). The selection differential was 3% for both selection criteria, and the economic

weight for the volume was given three times more than forking traits in index selection. In independent culling selection, the selected truncation for volume, forking and ramicorn branching were 106, 0.1 and 0.42, respectively.

2.4. Results

2.4.1 Tree growth, stem forking, and ramicorn branching

All growth traits showed significant differences across test sites (F-test Pr≤0.001). Height averaged 10.9 m and ranged from 9.4 m (test 3- Allendale County) to 12.4 m (test 1- Colleton County). The mean DBH was 14.9 cm, with the highest value at test 1 (16.6 cm). The averaged highest volume was also in test 1 (100.4 dm^3 /tree) and the lowest was in test 3 (56.2 dm^3 /tree).

Overall, 16 % of the trees had one fork and only 1 % of the trees had two forks on the stems, but ramicorn branching was more frequent than forking. About 18% and 5% of the trees had one and two ramicorn branches, respectively (Figure 2.1). Very few trees showed three or more forks (0.12%) or ramicorn branches (0.86%). The average height for the first-fork (close to bottom) was 5.2 m with the standard deviation of 2.4. The height of the second-fork (higher than the first-fork) averaged 6.4 m with the standard deviation of 1.83.

Test sites differed significantly in the frequency of forking and ramicorn branches (F-test Pr≤0.001). The site of Allendale had the highest proportion of forks (21%), and the

Berkeley County test had the lowest proportion (13%) (Figure 2.2). Similarly to forking, ramicorn branching was highest in Allendale (31%) and lowest in Berkeley (18%) (Figure 2.2).

2.4.2 Genetic variation and repeatability of clone means

Clones differed significantly for forking and ramicorn branching. The forking proportion among clones ranged from 0% to 73%, and ramicorn percentages varied from 3% to 50%. Frequency distributions of forking and ramicorn branching for clones are presented in Figure 2.3. Within the family, the phenotypic variance was mainly explained by clonal differences, and the interaction of clone by test was small (Table 2.1). The clone-mean repeatability for forking was 0.86, considerably higher than that for ramicorn branching (0.67). Estimated repeatabilities of clone means were reasonably high for all growth traits (HT and VOL) and other stem form traits (straightness and branch angles), and varied from 0.74 to 0.84(Table 2.1).

2.4.3 Trait-trait correlations

The pairs of growth traits (HT, DBH & VOL) show strong positive genetic correlations, ranging from 0.67 to 0.81 (Table 2.2). Forking was positively but only moderately correlated with HT and VOL and the correlations were unfavorable ($0.12 \leq r_g \leq 0.31$). Ramicorn branching showed similar associations with the growth traits, with moderately positive values. Straightness had favorable correlations with forking and ramicorns. Straightness appeared to be positively related with all other traits ($0.11 \leq r_g \leq 0.46$), while

branch angle seemed to be independent of the other traits. The two threshold traits, forking and ramicorn branching, had a moderately strong correlation of 0.68. All genetic correlations were associated with small standard errors.

The corresponding environmental correlations were very weak (close to 0) between forking defects and growth traits (Table 2.2). Strong to moderate environmental correlation was found between HT & VOL (0.73) and HT & DBH (0.57), while relatively low environmental association was detected between DBH and VOL (0.15).

2.4.4 Predicted genetic gain from clonal selection

Based on BLUP breeding values, the genetic gains from clonal selection against forking and ramicorn branching were estimated (Figure 2.4). With the high repeatability of clone means, selection resulted in a substantial response. The number of clones being selected greatly influenced the predicted gain in ramicorn branching, but was less so with forking. The highest selection intensity of 0.4% (1 out of 232 clones) could obtain the highest gain in reducing the incidence of ramicorns (17%), and the selection intensity of 2% resulted in the 11 % reduction of forks. When the selection intensity was decreased, the corresponding genetic gains for both traits were reduced. When the selection intensity was 4%, the predicted gains were 10% and 13% for forking and ramicorn branching, respectively.

The clonal selection also yielded considerable genetic gain for volume over the mean of all the clones (Figure 2.4). The response in genetic gain for volume from direct selection

could be maximized as 22% when the best one clone was selected. The predicted genetic gains for volume were more sensitive to the selection intensity than forking traits. When the selected clone number increased to 10 and 30, the genetic response for volume would be reduced to 17% and 12%, respectively.

Through index selection on volume and forking, both traits could be improved. Volume could be increased 15.5% and fork proportion could be decreased 6.3% simultaneously. When volume and ramicorn branching were put into the selection index, 15.5% gain could be made for volume and 3% decrease for ramicorn. In independent culling selection, forking and ramicorn frequency could be reduced 7% and 6%, respectively; volume could be increased 11%.

2.5. Discussion

2.5.1 Forking defects in clonal trials

This study showed that forked stems and major ramicorn branches can become serious stem problems in the clonal plantings. The high proportion of forks and ramicorns seen in this trial would greatly decrease the uniformity and the quality of an operational clonal forest. Forking differences among sites were more likely caused by environmental effects, such as site fertility, disease infection, or insect attacks, although we didn't have assessments for these factors. Higher growth in test 1- Colleton County is probably due to the older age of the trees compared to the other sites. Among the trees with forking defects, one fork per tree or one ramicorn per tree were the common cases. This was

similar to the previous study in ash trees, where 40% of the trees had one fork, and 29% had more than one fork (Kerr and Boswell 2001). The first fork was at an average height of 5.2 m, approximately half the average tree height (mean HT=10.9 m). The relatively low position for the occurrence of a fork would cause great economic losses, because the forks in the bottom of the stem have greater impacts on timber yield as compared to forks in the upper stem (Mwase et al. 2008).

2.5.2 Genetic variation of clones

There were significant differences for forking among clones within this full-sib family of loblolly pine. This was a single family study, so the results may not necessarily be representative of all families. However, our previous study based on thousands of families in diallel tests of loblolly pine have shown that forking is partially under genetic control at the family level (h^2_{fs}=0.59) (Xiong et al. 2010a). A weak individual-tree heritability (0.06) was found in the diallel test, probably because the seedling tests lacked the power to detect the genetic differences among individuals based on a single measurement per genotype. In this clonal test, a large number of ramets from each of the genotypes were assessed in different environments, which increases the precision of repeatability estimates. The finding of high clone-mean repeatabilities of forking and ramicorn branching in this study further supports the hypothesis that genetic control for forking is strong in clones.

Forks and ramicorn branches are both threshold (binary) traits, for which there is a common assumption that the character has an underlying continuity with a threshold, although their phenotypic values are discontinuous. The underlying variables, or liabilities, are influenced by both genetic and environmental effects (Falconer and Mackay 1996). The generalized linear mixed model (GLMM) with logit function was used to analyze the binomial data, which applied generalized linear model theory directly on the underlying scale (Gilmour et al. 1985). In this model, the assumptions of homogeneity of variances and normality for residual variance don't have critical impacts on the analysis, because the linearity is a property of the fitted values given by the model instead of data. Hence, GLMM provides relatively accurate predictions of variances and repeatabilities for threshold traits. However, the variances of threshold traits are mean dependent. When the mean incidence reaches 0.5 in the population, the error for the prediction of heritability is smallest (Lynch and Walsh 1997). In this clonal study, low proportions of forks in some test sites might bias the estimation of the clonal-mean repeatabilities, while large number of ramets per clones (n=40) helped to increase the precision of the estimations.

The distribution of frequencies of forking and ramicorn branching in clones was approximately normal. It is assumed that there is a continuous-distributed variable for forking. However, it is unclear whether the genetic control of forking and ramicorn branching in clonal loblolly pine will follow a typical quantitative model, or whether the variation will be accounted for by few genes of large magnitude. Our next study focusing

on the genetic mapping of QTL for forking defects will help to answer the questions about the genetic basis of forking.

Relatively high values of heritability were found in growth traits and other stem form traits (straightness and branch angles), suggesting that considerable genetic gains could be yielded through efficient selection for superior clones within family. The previous study (Isik et al. 2005) in two loblolly pine clonal trails also found moderately high within-family repeatability of clone means for growth traits (0.5 to 0.75).

2.5.3 Genetic gain

Fewer forked stems and ramicorn branches could be achieved through clonal selection, mainly due to the high clone-mean repeatabilities for the traits. These clonal tests demonstrate the advantages of repeated observations for the same genotype and the resulting high within-family heritabilities for binary traits. With the clonal replications and the clonal within-family selection being based on the same genotypes planted in different environments, much more gain can result (Shaw and Hood 1985).

Our findings with the growth traits agree with previous studies. The growth traits were found to be significantly different among clones within family in loblolly pine and slash pine clonal tests (Emhart et al. 2007). It has been reported that the clonal selection could achieve 30% of gain for volume over means of all the clones at two sites (Isik et al. 2005) and clone-based testing could greatly increase the gain for volume from within-family selection comparing to seedling test (Isik et al. 2005, Isik et al. 2003). The gains from

within-family selection of clonal tests were six times higher than the within-family selection of polymix and control-pollinated seedling testing (Isik et al. 2004).

2.5.4 Trait-trait correlations

Unfavorable genetic correlations were found between growth traits and forking defects, indicating that selecting higher growth trees will result in more forks and ramicorn branches. This result agreed with the previous study in Douglas-fir (*Pseudotsuga menziesii* var. *menziesii* (Mirb.) Franco) (Schermann et al. 1997, Adams and Bastien 1994) and loblolly pine (Cumbie et al. 2010). The genetic correlation between forking and straightness was desirable, which means straighter trees had less forking defects. Selection on straighter trees would correspondingly achieve the trees with low forks and ramicorns. Forks and major ramicorn branches were positively correlated (r_g=0.68), indicating that both forking defects may be partially controlled by the same genes. In our previous study (Xiong et al. 2010a), ramicorn branching and forking could not be distinguished and were considered as a single trait, which could bias the results (Jayawickrama 2001, Andersson et al. 2007). The positive correlation between forking and ramicorn branching makes it possible to reduce both forking defects simultaneously. The genetic correlations found in this clonal test were comparable with the earlier investigation in diallel tests of loblolly pine, where we found a moderately unfavorable correlation between height and forking and a favorable correlation between forking and straightness (Xiong et al. 2010a).

The high environmental correlation of HT & VOL and HT & DBH can be explained by the microenvironmental influences on the individual trees for growth traits. This indicates that the substantial environmental influences on HT could affect DBH and VOL as well. No environmental correlations were found between growth and forking, which suggested that the higher growth tree in a particular site won't tend to have more forks. It also suggested that stem form traits did not have as high sensitivity to microenvironmental conditions as do growth traits, and there is higher non-genetic plasticity to adjust to environmental conditions for growth than forking defects.

2.6. Deployment Implications

Our results suggest that significant gains can be achieved by the deployment of clones for loblolly pine, given the substantial genetic variation and high repeatabilities that existed among clones within family for forking defects. The unfavorable correlations between growth traits and forking defects imply that selection for either trait would have negative influences on the gain of the other. For forking defects, there is high genetic variation among clones and moderate genetic control at the family level (Xiong et al. 2010a). Hence, we can capture the family variation for growth traits and clonal differences for forking to obtain the maximum gains. Based on this point, a two-stage selection scenario is suggested for loblolly pine breeding. At first stage, by using polymix or control-pollinated testing (Isik et al. 2004), the best families are selected to achieve the high gain for volume. In the second stage, clones are generated from the full-sib families, which are selected from the first stage and all clones are tested for the forking traits. This method

maximizes the genetic gain for target traits and brings little negative impacts on the other economic traits at each step.

Because the weak to moderate associations between growth traits and forking, the high intensity selection by the scenario of index or independent culling levels could improve growth rate while limiting the frequency of forking defects. Index selection could make greater gain on volume while keeping forking defects relatively low. Compared to the index selection, volume gain was relatively low in the independent culling selection. The index selection is more accurate and powerful than independent culling selection by using the records of repeatability, variances and economic values of all traits (Cumbie et al. 2010). Index selection has been used in the breeding of Douglas-fir for the improvement of stem growth and forking defects (Schermann et al. 1997), and this type of selection enabled the breeder to obtain expected genetic gains for multiple traits.

This study showed the potential of using clones to achieve better stem forms for loblolly pine, but more issues should be considered when using clones in operational plantations. One major concern about deploying clones in plantations is the question of "gain versus risk" (Zobel and Talbert 1984). A narrow genetic base of clonal forest is more vulnerable to diseases and insects, and thus could increase the risk of losses in the forest. It is generally assumed that the risk will be lower if more genotypes are deployed in a clonal plantation (Roberds and Bishir 1997). However, when more clones are used in the plantations, lower gain will result. Normally, 7 to 30 clones are recommended as a safe and reasonable number for clonal forestry (Libby 1982). Fortunately, the predicted gain

for forking defect from clonal selection was not highly sensitive to the number of clones being selected, thus relatively more clones applied in the plantation will still achieve good gains in decreasing forking defects.

2.7. Conclusion

This study confirms that forking defects are under relatively strong genetic control at the clonal level. Clonal selection is expected to reduce the probability of forking and ramicorn branching over the population. Because the unfavorable genetic correlations between growth and forking, selection for stem growth alone would negatively impact wood stem quality. Growth potential in loblolly pine cannot be maximized without influencing stem form. Clones with better straightness tend to have less forks and ramicorn branches, although the correlation between them is relatively weak. Forking and ramicorn branching are highly interrelated, thus it is possible to develop both stem-quality traits simultaneously. Two-stage selection, index selection and independent culling selection could be the optional scenarios to achieve greater gains from growth traits and forking defects.

Acknowledgements

This work is supported by the Cooperative Tree Improvement Program in the Department of Forestry and Environmental Resources, NCSU. We are very thankful to MeadWestvaco (now ArborGen) for providing the clonal trials used in these analyses.

2.8. References

Adams, W.T. and J.C. Bastien. 1994. Genetics of second flushing in a French plantation of coastal Douglas-fir. *Silvae Genet.* **43**(5):345-352.

Amerson, H.V., R.W. Whetten, S.A. Garcia, B. Li, S.E. McKeand and F. Isik. 2008. Resistance of Pinus taeda families under artificial inoculations with diverse fusiform rust pathogen populations and comparison with field trials. *Can. J. Forest Res.* **38**(10):2687-2696.

Andersson, B., B. Elfving, T. Persson, T. Ericsson and J. Kroon. 2007. Characteristics and development of improved Pinus sylvestris in northern Sweden. *Can. J. Forest Res.* **37**(1):84-92.

Baltunis, B.S., T.J. Mullin, J.T. Brawner, H.X. Wu and H.S. Dungey. 2009. Comparisons of genetic parameters and clonal value predictions from clonal trials and seedling base population trials of radiata pine. *Tree Genet. Genomes.* **5**(1):269-278.

Cornelius, J.P. and A.D. Watt. 2003. Genetic variation in a Hypsipyla-attacked clonal trial of Cedrela odorata under two pruning regimes. *For.Ecol.Manage.* **183**(1):341-349.

Cumbie, W P, F. Isik, and S E McKeand. 2010. Genetic variation in sawtimber potential in loblolly pine. *For. Sci.* In review.

Doede, D.L. and W.T. Adams. 1998. The genetics of stem volume, stem form, and branch characteristics in sapling noble fir. *Silvae Genet.* **47**(4):177-183.

Downs, A.A. 1949. Low forking in white oak sprouts may be hereditary. *Jour. Forestry.* **47**:736.

Dvorak, W.S., J.K. Donahue and J.A. Vasquez. 1996. Provenance and progeny results for the tropical white pine, Pinus chiapensis, at five and eight years of age. *New Forest.* **12**(2):125-140.

Dvorak, W.S., E. Kietzka, G.R. Hodge, A. Nel, G.A.d. Santos and C. Gantz. 2007. Assessing the potential of Pinus herrerae as a plantation species for the subtropics. *For. Ecol. Manage.* **242**(2):598-605.

Emhart, V.I., T.A. Martin, T.L. White and D.A. and Huber. 2007. Clonal variation in crown structure, absorbed photosynthetically active radiation and growth of loblolly pine and slash pine. *Tree phys.* **27**(3):421-430.

Falconer, D.S. and Mackay, T.F.C. 1996. Introduction to quantitative genetics. 4 edn. Longman Group Limited. Harlow: 464p.

Franklin, E.C. and R.Z. Callaham. 1970. Multinodality, branching, and forking in Lodgepole Pine (Pinus contorta var. murrayana Engelm.). *Silvae Genet.* **19**(5/6):180-184.

Gilmour, A.R., R.D. Anderson and A.L. Rae. 1985. The analysis of binomial data by a generalized linear mixed model. *Biometrika.* **72**(3):593-599.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. and Thomson, R. 2002. ASReml User Guide, release 1.0. VSN International Ltd. Hemel Hempstead,UK.: .

Goebel, N.B. and J.R. Warner. 1966. Volume tables for small diameter loblolly, short leaf and Virginia pine in the upper South Carolina Piedmont. *For. Res. Serv.* **No. 7**.

Isik, F., B. Goldfarb, A. LeBude, B.L. Li and S. McKeand. 2005. Predicted genetic gains and testing efficiency from two loblolly pine clonal trials. *Can. J. Forest Res.* **35**(7):1754-1766.

Isik, F., B.L. Li and J. Frampton. 2003. Estimates of additive, dominance and epistatic genetic variances from a clonally replicated test of loblolly pine. *For.Sci.* **49**(1):77-78.

Isik, F., B.L. Li, J. Frampton and B. Goldfarb. 2004. Efficiency of seedlings and rooted cuttings for testing and selection in Pinus taeda. *For.Sci.* **50**(1):44-53.

Jayawickrama, K.J.S. 2001. Genetic parameter estimates for radiata pine in New Zealand and New South Wales: a synthesis of results. *Silvae Genet.* **50**(2):45-53.

Kerr, G. and C. Boswell. 2001. The influence of spring frosts, ash bud moth (Prays fraxinella) and site factors on forking of young ash (Fraxinus excelsior)in southern Britain. *Forestry.* **74**(1):29-40.

Ladrach, W.E. and C. Lambeth. 1991. Growth and heritability estimates for a seven-year-old open-pollinated Pinus patula progeny test in Colombia. *Silvae Genet.* **40**(5/6):169-173.

Libby, W.J.1982. What is a safe number of clones per plantation? P.342-360 *in the 3. International Workshop on the Genetics of Host-Parasite Interactions in Forestry*Anonymous PudocWageningen (Netherlands).

Littell, R.C., Milliken, G.A., Stroup, W.W. and Wolfinger, R. 1996. SAS System for Mixed Models. SAS Institute Inc. Cary ,N.C.: 656p.

Lynch, M. and Walsh, B. 1997. Genetics and analysis of quantitative traits. Sinauer Associates Incorporated. Sunderland, MA, U.S.A: 980p.

Moura, V.P.G., W.S. Dvorak and G.R. Hodge. 1998. Provenance and family variation of Pinus oocarpa grown in the Brazilian cerrado. *For.Ecol.Manage.* **109**(1-3):315-322.

Mwase, W.F., P.S. Savill and G. Hemery. 2008. Genetic parameter estimates for growth and form traits in common ash (Fraxinus excelsior, L.) in a breeding seedling orchard at Little Wittenham in England. *New Forests.* **36**(3):225-238.

Osorio, L.F. 2004. Provenance results of Gmelina arborea in southwest Colombia at three years of age. *New Forests.* **28**(2/3):179-185.

Roberds, J.H. and J.W. Bishir. 1997. Risk analyses in clonal forestry. **27**(3):425-432.

SAS INSTITUTE INC. 1996. SAS/STAT: Changes and enhancement (through release 6.11). Cary ,N.C..

Schermann, N., W.T. Adams, S.N. Aitken and J.C. Bastien. 1997. Genetic parameters of stem form traits in a 9-year-old coastal Douglas-fir progeny test in Washington. *Silvae Genet.* **46**(2/3):166-170.

Shaw, D.V. and J.V. Hood. 1985. Maximizing gain per effort by using clonal replicates in genetic tests. *Theor.Appl.Genet.* **71**(3):392-399.

Southey, B.R. 2003. Discrete time survival analysis of lamb mortality in a terminal sire composite population. *J. Anim. Sci.* **81**(6):1399-1405.

Stelzer, H.E. and B. Goldfarb. 1997. Implementing clonal forestry in the southeastern United States: SRIEG satellite workshop summary remarks. *Can. J. Forest Res.* **27**(3):442-446.

Van Vleck, L.D. 1970. Index selection for direct and maternal genetic components of economic traits. *Biometrics.* **26**(3):477-483.

Vargas-Hernandez, J.J., W.T. Adams and D.G. Joyce. 2003. Quantitative genetic structure of stem form and branching traits in Douglas-fir seedlings and implications for early selection. *Silvae Genet.* **52**(1):36-44.

Xiong, S.J., F. Isik, S.E. McKeand and R.W. Whetten. 2010. Genetic variation of stem forking in loblolly pine. *For.Sci.* (In Press).

Xu, S. and W.M. Muir. 1991. Multistage selection for genetic gain by orthogonal transformation. *Genetics.* **129**(3):963-974.

Zobel, B.J. and Jett, J.B. 1995. Genetics of wood production. Springer-Verlag. Berlin: 337p.

Zobel, B. and Talbert, J. 1984. Applied forest tree improvement. Wiley, John & Sons. New York: 505p.

Table 2.1 Estimated variance components and within family repeatabilities of clone means and their standard errors (SE). (Error variance was 1 for the threshold traits in this output, but 3.29 was used as error variance for the threshold traits in the genetic analysis.)

	Br. Angle		Ramicorn		Fork	
Source	estimate	(SE)	estimate	(SE)	estimate	(SE)
Clone	0.02	(0.003)	0.18	(0.035)	0.64	(0.090)
Clone × test	0.01	(0.002)	0.03	(0.036)	0.08	(0.044)
Error	0.23	(0.004)	1.00	(0.000)	1.00	(0.000)
$H_{\bar{C}}^2$	0.74	(0.030)	0.67	(0.053)	0.86	(0.022)

Table 2.1 Continued

	Height		Volume		Straightness	
Source	estimate	(SE)	estimate	(SE)	estimate	(SE)
Clone	0.13	(0.016)	72.00	(9.290)	0.15	(0.016)
Clone × test	0.07	(0.008)	61.47	(5.513)	0.03	(0.007)
Error	0.75	(0.012)	338.24	(5.591)	0.79	(0.013)
$H_{\bar{C}}^2$	0.79	(0.024)	0.75	(0.028)	0.84	(0.018)

Table 2.2 Estimates of genetic correlations (above the diagonal) and environmental correlations (below the diagonal) between traits and their standard errors in parentheses.

	Fork	HT	VOL	Strt	BrAngle	Ramicorn
Fork	-	0.12 (0.09)	0.26 (0.08)	0.23 (0.08)	-0.07 (0.09)	0.68 (0.09)
HT	-0.01 (0.01)	-	0.81 (0.03)	0.46 (0.09)	-0.24 (0.08)	0.30 (0.10)
VOL	0.05 (0.01)	0.73 (0.01)	-	0.11 (0.08)	-0.02 (0.09)	0.37 (0.09)
Strt	0.03 (0.01)	0.00 (0.01)	0.00 (0.01)	-	0.15 (0.07)	0.49 (0.08)
BrAngle	-0.06 (0.01)	-0.01 (0.01)	0.00 (0.01)	0.02 (0.01)	-	0.31 (0.10)
Ramicorn	-0.01 (0.01)	-0.05 (0.01)	0.01 (0.01)	0.08 (0.01)	0.02 (0.01)	-

Note: HT, height; DBH, Diameter at breast height; BrAngle: Branch angle; Strt, straightness; Ramicorn, ramicorn branches; Fork, forked stems.

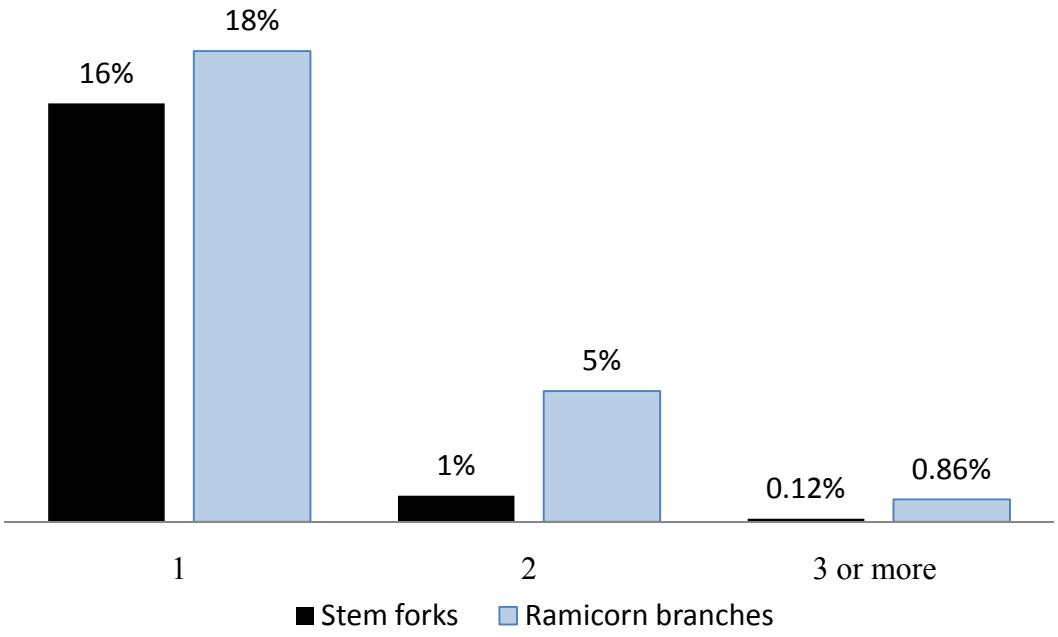


Figure 2.1 Occurrence (percent) of one or multiple stem forking and ramicorn branches per stem in loblolly pine. On average 16% of trees develop forked stems (one fork). Percent of trees developing two forks is about 1%. It seems that ramicorn branching is slightly more frequent in the species compared to forking.

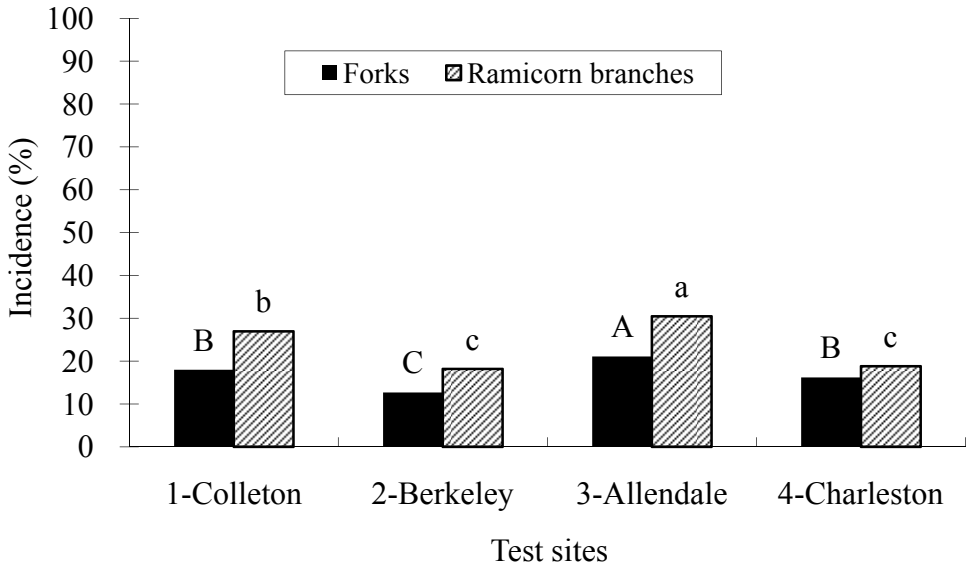


Figure 2.2 The proportion of forking and ramicorn branching at different test sites. The incidence for each trait with the same letter are not significantly different at α=0.05.

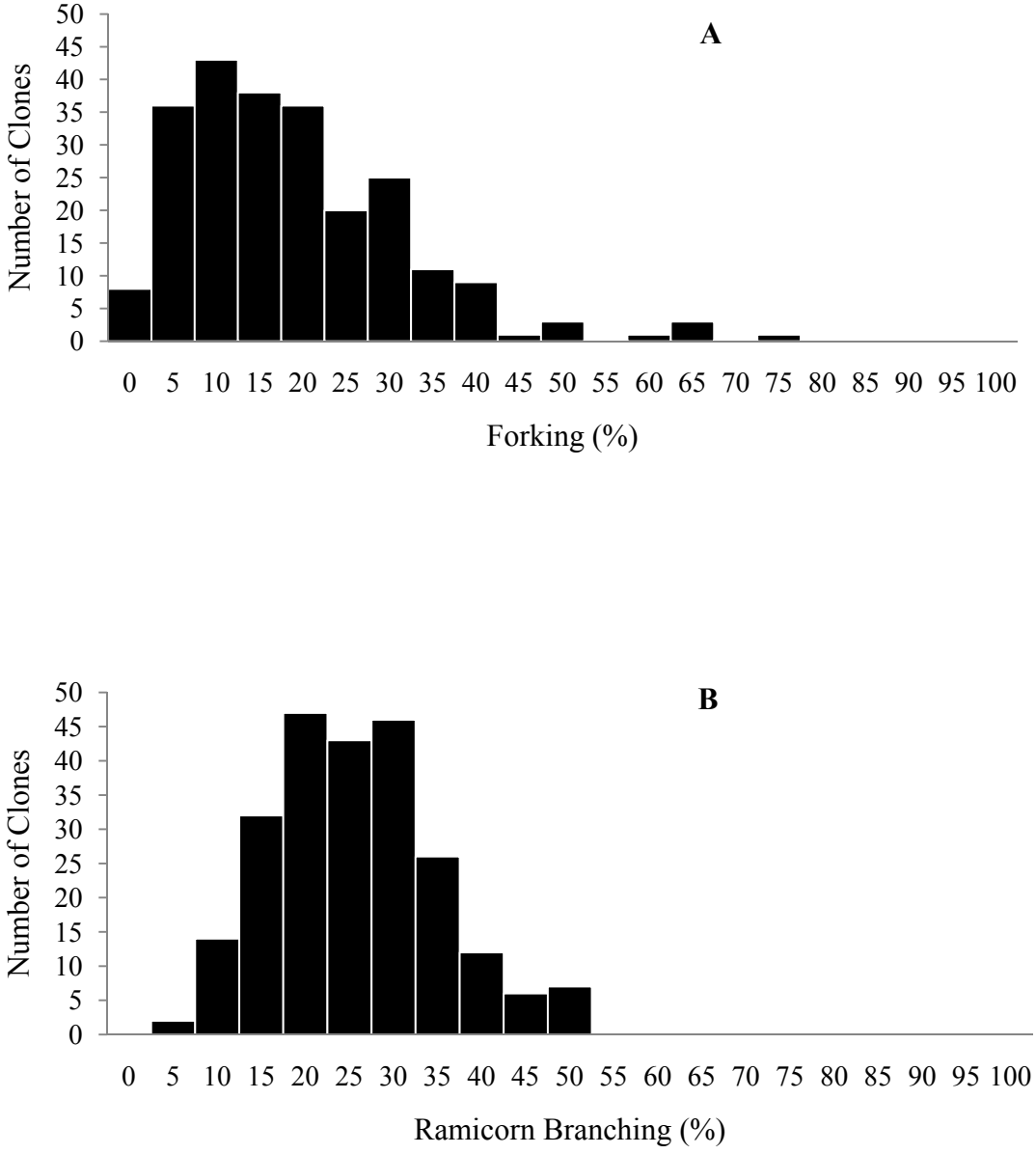
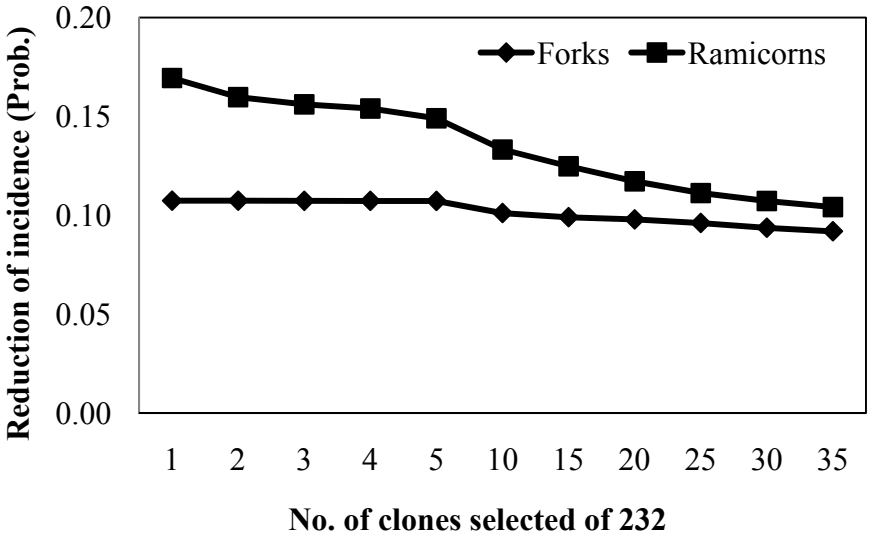


Figure 2.3 Frequency distributions for forking (A) and ramicorn branching (B) for clones.



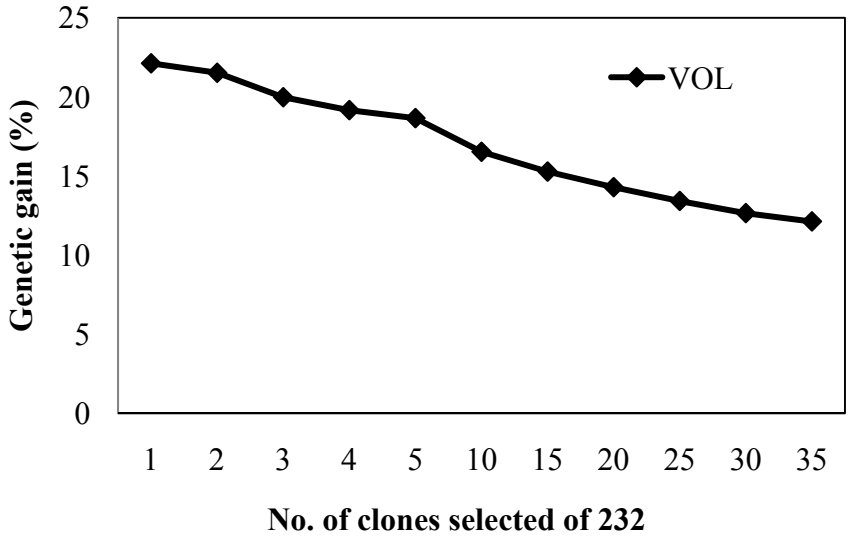


Figure 2.4 The genetic gain from clonal selection: reduction in probability of trees with forking and ramicorn branching versus the number of clone selected; genetic gain (%) for volume (VOL) from various clonal selections.

3. Chapter 3. Quantitative Trait Loci Influencing Forking Defects in Loblolly Pine

Jin S. Xiong, Ross W. Whetten, Steven E. McKeand, Zhao-Bang Zeng, and Fikret Isik

Abstract

Quantitative trait loci influencing forking defects were identified in an outbred pedigree of loblolly pine (*Pinus taeda* L.). Genetic markers segregated in the outbred full-sib family in F2 intercross and both backcross configurations. The linkage map was constructed with all types of markers. A total of 409 SNP markers distributed among 18 linkage groups covered 1703.7 cM. A single-marker analysis of 1257 loci identified 11 and 9 markers with significant associations with forking and ramicorn branch phenotypes, respectively, and clusters of markers were found associated with multiple traits ($q<0.05$). By using the interval mapping method, two and three QTL were identified for forking and ramicorn branch traits, respectively. Two QTL were detected for both traits using multiple traits analysis, suggesting there are pleiotropic effects on both traits. The result from this study is very encouraging for applying marker-assisted selection to improvement of forking defects in breeding populations of loblolly pine.

Key words: QTL, forking, outbred, *Pinus taeda,* MAS

3.1. Introduction

Tree improvement programs in loblolly pine aim to increase the wood quantity and quality for owners and managers of pine plantations. Conventional breeding programs have achieved great genetic gains in the productivity of loblolly pine (McKeand et al. 2003), but long breeding cycles limit the rate of genetic gain per year. Modern forest breeding, exploiting molecular markers and genomic maps to assess the genetic basis of phenotypic traits, may be able to achieve gains more efficiently. The application of marker-assisted selection (MAS) in breeding program could allow selection at early ages and correspondingly shorten breeding cycles (Ha et al. 2007). In addition, MAS using the genetic base of the traits instead of phenotypes makes the selection more precise and powerful.

3.1.1 Molecular markers

The successful application of MAS in tree breeding requires efficient and economical marker assays. Different kinds of molecular markers have been developed and applied to pine species. The first DNA-based molecular marker, restriction fragment length polymorphism (RFLP) (Southern 1975), are suited to identify different species and populations (Neale et al. 1994). The RFLP loci have been used to construct a genetic linkage map from a three-generation outbred pedigree loblolly pine (Devey et al. 1994, Devey et al. 1991, Sewell et al. 1999). A common set of RFLP markers were used in a single outbred loblolly pine pedigree to identify the QTL influencing wood specific

gravity (Groover et al., 1994, Knott et al. 1997). Later on, a PCR-based marker, randomly amplified polymorphic DNA (RAPD) (Williams et al. 1990) has been applied to genetic linkage and association study in pine species, such as longleaf pine (Nelson and Kubisiak 1994), white pine (*Pinus strobus* L.) (Echt and Nelson 1997), Japanese black pine (*Pinus thunbergii*) (Kondo et al. 2000) and loblolly pine (Wu et al. 1999). Amplified fragment length polymorphism (AFLP) is a more powerful marker with high reproducibility, and only needs a small number of oligonucleotide primers and small amounts of sample DNA (Vos et al. 1995). It has been used for genetic fingerprinting, genome mapping, and genetic variability studies. A complete genetic map of loblolly pine was constructed by using large number of AFLP makers (Remington et al. 1999). AFLP markers were also applied to rapidly generate a dense linkage map for Scots pine (*Pinus sylvestris*) (Lerceteau et al. 2000) and pinyon pine (*Pinus edulis*) (Travis et al. 1998). Another valuable marker, microsatellite or simple sequence repeat (SSR), has been widely used recently for the studies of paternity identification, estimation of outcrossing rates, monitoring of inbreeding, and genetic differentiation (Morgante and Olivieri 1993, Marques et al. 2002). Microsatellite markers were first used in forest trees in a study of *Pinus radiata* (Smith and Devey 1994), and were then applied for other tree species. Microsatellites have proved to be stable, abundant and versatile markers, particularly for population genetic analyses, but they also have potential drawbacks. One problem is the occurrence of null alleles that fail to amplify in PCR assays (Jarnea and Lagodab 1996, Dakin and Avise 2004).

Recent development in marker technologies has enabled low-cost and high-throughput markers to be used in tree species. Of all the molecular markers available today, single nucleotide polymorphism (SNP) markers have the best combination of characteristics such as polymorphism levels, codominance, and suitability for accurate, reproducible, and automated genotyping at high marker density (Gut 2001). The high density of SNPs are good for genome mapping, especially for the ultrahigh-density genetic maps, and the genetic maps help to identify the specific loci controlling the certain traits (Appleby et al. 2009). SNP markers have been applied to various studies, including genetic diversity analysis, cultivar identification, population evolution and genomic association study, because of the desirable characteristics (Rafalski 2002). Brown *et al* (2004) has applied SNP markers to study the nucleotide diversity and linkage disequilibrium in loblolly pine (Brown et al. 2004). Several candidate gene SNPs have been discovered that are associated with wood properties (González-Martínez et al. 2007, Neale et al. 2002, Brown et al. 2003) and drought-stress response (González-Martínez et al. 2008) in loblolly pine.

3.1.2 Genetic mapping

A layout of molecular markers by order of loci and the distance between them is called a genetic map or linkage map. The distance between two markers can be estimated as a recombination fraction through the method of maximum likelihood. The recombination fraction can be translated to map distance in cM by mapping function. The Haldane and Kosambi mapping functions are two most popular functions. Several software packages

are available for construction of genetic maps (Lander et al. 1987, Stam 1993, Iwata and Ninomiya 2006, Wu et al. 2008). Mapmaker has been applied to construct linkage maps in human and plants. It uses the Lander-Green algorithm to find locus order, and allows comparing different genetic orders (Lander et al. 1987). Iwata and Ninomiya (2006) provide software called AntMap that creates linkage maps based on the algorithm of Ant Colony Optimization. Presently, one of the most widely available genetic mapping software packages is JoinMap, which uses weighted least squares to estimate map distances and build up the map by numerical search for the best fitting order of markers (Stam 1993). JoinMap is useful for aligning the linkage groups from different experiments, as well as non-classic linkage experiments, such as outbred crosses. MSTMap is another recently-developed software tool that builds genetic linkage maps by first constructing a Minimum Spanning Tree (MST). It can handle the analysis of different experimental populations, such as BC1, DH, Hap and RIL. MSTMap has been reported to produce more consistent results than JoinMap, in terms of shorter running time and higher accuracy (Wu et al. 2008). Margarido et al (2007) developed a software package called OneMap to facilitate the linkage analysis in outcrossing species, by using Wu's methodology (Wu et al. 2002). OneMap uses a maximum-likelihood-based algorithm for simultaneously estimating linkage and linkage phases, and is capable of analyzing a mixed set of different marker types, including fully informative markers (segregating 1:1:1:1) and partially informative markers (or missing markers, segregating 1:2:1, 3:1, and 1:1). OneMap is extremely valuable for the genetic linkage study in

outbred plant and tree species, because it allows analysis of markers with different segregation patterns within the same dataset and the same analysis.

3.1.3 Quantitative trait loci and QTL mapping

Quantitative traits are the result of combined effects of genetics and environment (Falconer and Mackay 1996). They are usually considered to be controlled by many genes, each with a small effect on the trait. The genes or loci controlling the quantitative traits are called quantitative trait loci (QTL). QTL mapping is the relationship of phenotypes of traits to genotypes of QTL, including the location and number of the QTL, as well as the effects of the QTL and their interactions.

Different statistical methods and software for mapping QTL have been reported. Lander and Botstein (1989) proposed interval mapping (IM) in F2 or backcross populations from inbred line cross experiments by adapting the approach of LOD score analysis. In this approach, one marker interval is tested at a time to test for a putative QTL by likelihood ratio test (LRT) at each position in the interval, and the putative QTL position is identified as the position with the highest significant LRT. Based on the IM theory, the Mapmaker QTL program was designed to perform the LRT calculation to detect the candidate QTL. Compared to IM, the least squares (LS) method is relatively simpler and faster in computing, and it can extract the most information in multiple linked markers (Haley et al. 1994). LS is more powerful to detect the putative QTL by using all the markers in linkage group simultaneously. There are several software programs

implementing LS- simple interval mapping, such as MQTL (Tinker and Mather 1995), Qgene (Nelson 1997), and MCQTL (Jourjon et al. 2005).

Considering one QTL at a time can bias the estimation when multiple QTL are located in the same group. Composite interval mapping (CIM) was developed to deal with this problem by using the combination of IM with multiple regression analysis in mapping (Zeng 1994, Jansen 1993). Multiple interval mapping (MIM) is another mapping method developed by Kao and Zeng (1999), which is more powerful and precise in detecting QTL by using multiple marker intervals simultaneously to construct QTL mapping. Through MIM, epistatic QTL and individual genetic values can be estimated. Windows QTL Cartographer, created by Wang et al (2003), is a suite of program offering various tools for QTL mapping, including IM, CIM, and MIM.

Bayesian methods have been also proposed for the analysis of QTL (Satagopan et al. 1996, Fang et al. 2008, Hayashi and Awata 2005). Satagopan used Markov chain Monte Carlo (MCMC) techniques to assess multiple loci and their effects simultaneously, instead of fitting one locus at a time or using markers as cofactors to assess other putative loci (Satagopan et al. 1996). The improved Bayesian method has been applied in QTL analysis for outbred F2 families allowing inference about whether each of the grandparents is homozygous or heterozygous at QTL (Hayashi and Awata 2005).

Seaton et al (Seaton et al. 2002) developed QTL Express as a web-based software, the first application designed for QTL mapping in outbred populations. A robust two-step

procedure is used for QTL mapping, including determining the Identity-By- Descent (IBD) probabilities and fitting statistical models to the observations and IBD coefficients. QTL Express is suitable for half-sib outbred populations, and F2 populations from either inbred or outbred lines.

Recently, a SAS procedure, PROC QTL, was reported for QTL mapping (Hu and Xu 2009). It is able to perform QTL maps for continuous and discrete traits by various methods, including maximum likelihood (ML), least square (LS), iteratively reweighted least square (IRLS), Fisher scoring (FISHER), Bayesian (BAYES), and empirical Bayes (EBAYES). PROC QTL allows analysis for several different cross types, including F2, backcross, recombinant inbred lines (RIL), double haploid (DH) and four-way crosses (FW). The advantage of using PROC QTL is to use all existing features offered by the general SAS software.

3.1.4 Background of the study

Forking defects, including stem forking and ramicorn branching, greatly influence the wood quality and the commercial value of timber in loblolly pine plantations (Xiong et al. 2010a). In order to decrease forking defects in plantations, we want to study the genetic basis of forking traits and develop trees with reduced frequency of forking defects through a selective breeding program. Our previous studies have proven that forking defects are under moderate genetic control at the family level and under strong genetic control at the clone level (Xiong et al. 2010b). In this study, the genetic basis of forking

will be further dissected by using SNP markers to detect QTL controlling the traits. The objectives of this study include 1) construction of a genetic linkage map using SNP markers; 2) identification of QTL controlling forking defects in loblolly pine, and 3) detect significant SNPs associated with QTL controlling forking traits.

3.2. Materials and methods

3.2.1 Plant material and trait measurement

Clonally-replicated loblolly pine progeny from an outbred family AF provided by MeadWestvaco were used in this study. Phenotype traits of height (HT), diameter at breast height (DBH), ramicorn branching (rami), stem forking (fork), branch angle (BranAn), straightness (Strt) were measured at age 6 and age 7 (Xiong et al. 2010b). A stem fork is defined as a separation of the trunk into two or more stems of similar sizes, and a ramicorn branch is a branch with twice the diameter and a substantially steeper angle than the other branches in the same whorl. If the tree was forked, it was scored as 1, otherwise as 0. Trees with ramicorn branches were scored as 1 and trees lacking ramicorn branches were scored as 0. The volume of individual trees was calculated according to the equation developed by Goebel & Warner (Goebel and Warner 1966). A total of 218 full-sib progeny and two parents were genotyped, and 51 samples were genotyped twice to test the reproducibility of the genotyping.

3.2.2 Map construction

The segregation of the SNP loci were assessed for distortion from expected Mendelian ratios prior to constructing the linkage map. χ^2 tests were conducted by using SAS procedures (SAS Institute Inc. 2004), according to expected Mendelian ratios of 1:2:1(heterozygous in both parents) and 1:1(heterozygous in one parent) for SNPs. Loci showing significant segregation distortion were omitted from further analysis, by using a significance threshold of P=0.05.

The genotypic dataset with the three types of marker configurations (F2 intercross, Backcross from parent 1 and Backcross from parent 2) were used to construct the joint map by the software OneMap (Margarido et al. 2007). The data file for OneMap was coded to indicate the marker type, according to the segregation patterns described by Wu et al (2002). There were 3 types of marker segregation patterns in this analysis (Table 3.1), coded as B3.7, D1.10 and D2.15 in the data file of OneMap.

Two-point recombination fractions was estimated using the function "rf.2pts", with a minimum LOD-Score of 6 and maximum recombination fraction (MRF) of 0.45. The function "make.seq" was used to assign markers to linkage groups with the same LOD-Score and MRF thresholds. Marker orders within group were estimated by using the code of "rcd", using the Haldane map function. The "order.seq" function was used to construct the map through adding markers sequentially, where the five most informative markers were used in the compare step, the markers could be mapped into the sequence with

LOD-Score higher than 3. Final maps were checked using the "ripple" function. The initial map was checked for false linkages by examination of the recombination fraction matrix, and the map was reconstructed according to the diagnosis. SNPs that map to the same positions in initial maps were filtered prior to the construction of the final linkage maps to remove redundant information.

3.2.3 QTL analysis

QTL analysis was carried out utilizing PROC QTL in SAS (Hu and Xu 2009), by using a primary dataset and a map dataset. The primary dataset included phenotypic values, marker genotypes and pedigree information, while the map dataset included marker name, position and linkage group. Phenotypes were breeding values predicted by best linear unbiased prediction (BLUP) from ASReml software (Xiong et al. 2010b) and all traits were analyzed as continuously-distributed traits. The linkage map with 18 groups was applied to QTL mapping. The entire genome was scanned with 1cM increment using the Interval Mapping approach under single model QTL model (Han and Xu 2008), which is

$$\mathbf{y_j=X_j\beta+ e_j}$$

Where y_j is the phenotypic value of individual j (j=1 to n); $\mathbf{X_j}$ is the expectation of variables indicating the QTL genotype given marker information; $\boldsymbol{\beta}$ is a vector of QTL effects (including the population mean); and $\mathbf{e_j}$ is the residual effect.

The total phenotypic variance can be partitioned into four components,

$Var(y_j)=(\alpha^p)^2+(\alpha^m)^2+\delta^2+\sigma^2$

Where $(\alpha^p)^2$ is the variance component from parent one, $(\alpha^m)^2$ is the variance from parent two, δ^2 is the interaction from both parents, and σ^2 is the residual variance.

The methods of iteratively reweighted least square (IRLS) and maximum likelihood (ML) were used for interval mapping QTL (Lander and Botstein 1989, Xu 1998b, Xu 1998a), based on a four-way cross model of genetic segregation. The likelihood ratio test (LRT) statistic was converted to log of odds (LOD) score by the formula of

$LOD=\frac{1}{2}(\log_{10} e)\ LRT=0.217\ LRT$

Permutation tests were used to find the critical value for the LOD score (Churchill and Doerge 1994). The measurements of phenotype and genotype were shuffled among individuals to generate a permuted sample of the data (to simulate the null hypothesis of no association between genotype and phenotype). Each permuted dataset was used to perform interval mapping analysis, and the maximum test statistic over all analysis points for each of the shuffled analyses were recorded. These values were then ordered, and their 95% percentile is the estimated experimentwise critical value. When the test statistic in the original data in a genomic region is higher than this critical value, a QTL is declared. One thousand permutations were used in the estimation of the genome-wide threshold.

3.2.4 Genome-wide association analysis

The genome-wide association analysis with family data was implemented by an R package "GWAF" (Chen and Yang 2010). GWAF tested genetic association between genotype (SNPs) and traits by fitting a Linear Mixed Effects (LME) model accounting for within pedigree correlation. The function of "lme.batch" used Wald chi-square test to perform a global test of genotype effects. The SNP genotype was treated as a fixed effect, and a random effect correlated according to degree of relatedness within a family is also fitted. Three types of data file were used in the analysis: pedigree file, genotype file, and phenotype file. The p-values (α) from GWAF were adjusted by Bonferroni correction (β) and by estimation of the false discovery rate (Storey 2002). Here, β =α/n, where n is the number of the tests (here, the number of genetic markers) (Holm 1979). Q-value calculations and QQ plots were performed using the R package Q-value (Dabney and Storey 2009). Phenotypic data were permuted and put into 10 different random phenotype sets. With the permuted data, GWAF analysis was repeated to estimate the lambda (λ)

$$\lambda = \frac{\text{median}(\chi^2\text{stat})}{0.455}$$

Lambda measures the deviation of chi-square tests results from expected null distribution.

The proportion of phenotype variance was estimated by the formula of

$$h_q^2 = \max\left(0, \frac{\sigma^2_{G.null} + \sigma^2_{e.null} - \sigma^2_{G.full} - \sigma^2_{e.full}}{Var(y)}\right)$$

where $Var(y)$ is the total phenotypic variance, $\sigma^2_{G\ .null}$ and $\sigma^2_{e\ .null}$ are the polygenic variance and error variance when modeling without the tested SNP, $\sigma^2_{G\ .full}$ and $\sigma^2_{e\ .full}$ are the polygenic variance and error variance when modeling with the tested SNP.

In this analysis, 1257 SNP markers were used to test the association with the phenotypic traits. The identified significant markers associated with forking defects were tested for linkages by using OneMap.

3.3. Result

3.3.1 Data summary

The repeated sampling tests showed that the accuracy of sampling and genotyping is high. Genotyping results of 51 clones sampled twice showed that only one clone failed to show agreement between the two replicate samples. Of 51 repeats, 42 had 100% identity of the genotypes between the two samples (Figure 3.1). The failed samples were removed, and then a total of 217 full-sib individuals plus two parents were used in the further analysis.

Of the 4775 SNPs from genotyping, 1552 SNPs are informative, consisting of 552 backcross 1 configurations (BC1) from one parent, 696 backcross 2 (BC2) from the other parent, and 304 F2 intercross configurations (F2) (Table 3.1). There are 3223 SNP loci that are homozygous in both parents, so those markers are not informative for the

analysis. χ^2 tests for the segregation were performed according to the Mendelian ratios of 1:2:1 (F2) and 1:1 (BC). Markers of the former type were termed both informative, while those of the latter type were termed maternally informative or paternally informative depending on whether it was the mother or father tree that was heterozygous. Within the informative markers, 439 markers showed segregation distortion (P<0.05), including 67 F2 markers, 120 BC1 markers and 252 BC2 markers. Those markers were omitted from further analysis. A total of 1110 SNPs were used in linkage mapping analysis.

3.3.2 Linkage map

When using a minimum LOD value of 6.0 and a maximum recombination fraction of 0.45 in grouping the markers, 12 linkage groups were obtained (Figure 3.2). The redundant SNPs in the same position were removed prior to the construction of the linkage map. Overall, 427 markers were defined at a linkage map with total length of 3771.6 cM in 12 groups. It seems reasonable to have 12 groups for the linkage maps, because pine has 12 pairs of chromosomes. However, this map has major flaws, including gaps larger than 100 cM within some linkage groups (Figure 3.2). In group 2, for example, a gap exists between marker M21055 and M21003 and another between marker M21094 and P512. The plots of the recombination fraction matrix generated from OneMap were used to make diagnostics about the map. The color scale varies from red (small distance) to dark blue (big distance), representing different recombination fraction. Based on the diagnostic plotting, subgroups of closely-linked markers can be detected within the larger linkage groups. From the matrix of group 2 (Figure 3.2 A), it is possible

to see a sub-division in three groups. Some of the other groups showed the obvious subgroups as well (Figure 3.3). The reason that sub-groups were placed together in a group might be a false linkage (false positive) detected between markers.

According to diagnostic plotting, the subgroups were divided by breaking those markers with distances more than 25 cM. Finally, 18 groups were formed and used for reconstructing the linkage map. The new linkage map had a total length of 1703.7 cM, defined by 409 markers (Figure 3.4). The largest group was group 14 with 44 markers, while the smallest group 17 only had 5 markers. The average distance between two markers was 4.2 cM over all groups. The markers in the new linkage map showed the same orders as those in the old map, although they are presented in different groups in some cases.

3.3.3 Association analysis

The QQ plots produced using the GWAF package in R from the original data set showed over dispersion, with high values of lambda. The estimated lambda is 2.40 and 2.43 for the stem forking and ramicorn branching traits, respectively. The values of lambda for other traits were also much higher than 1. When the data were permuted, the QQ plots did not show the dispersion. A QQ plot of permuted data is shown in Figure 3.5. The values of lambda from 10 permutation tests were close to 1 for all traits, suggesting the observed P-values do not strongly depart from the expected P-value distribution by chance. It also indicates the over-dispersion might come from true genetic effects, which happen to be

segregating within this full-sib family and are detected by linkages over much greater distances than would typically be the case in the type of association genetics study for which the GWAF package was designed

From the additive model, the p-value of the chi-square test for each marker was estimated. After Bonferroni correction of p-value ($\alpha<0.05$), 11 markers showed significant association with forking. Those SNPS each individually explain 7% to 10% of the total phenotypic variations (Table 3.2). Nine SNPs had association with ramicorn, and the proportion of phenotypic variation explained by each SNP was 7% ~ 9.5% (Table 3.3). A larger number of SNPs were found to be significantly associated with growth traits, including 24 and 16 SNPs for volume and height, respectively (data not shown). For both traits, the proportion of the phenotypic variation explained by individual SNPs ranged from 7% to 14%.

From the q-plots of the forking trait (Figure 3.6.A), π_0 showed a trend toward smaller values when the value of λ increased, except the end of the right tails. The estimated π_0 was 0.745, which means the proportion of true null hypotheses among all tests was 74.5%. When the q-value cut-off increased, the number of the significant test and the expected false positives increased (Figure 3.6.C&D). When q-value cut-off was chosen as 0.05, the number of significant tests was 59. Correspondingly, the p-value was about 0.003 and the number of the expected false positives was 3 at this point (Figure 3.6. B&D). For the trait ramicorn branching, the plots showed similar trends as for the forking

trait (Figure 3.7). When q-value cut-off was 0.05, the number of significant tests was 83 and the expected number of false positive results was 4.

According to q-plots, it seems reasonable to have q-value≤0.05 for achieving ideal number of significant markers. Under this cut-off, 59 and 83 SNPs were found significantly associated with the stem forking and ramicorn branching traits, respectively. Those 59 and 83 significant markers were used to construct linkage maps, respectively. The results showed that those markers identified as associated with stem forking fell into four linkage groups (Figure 3.8). The significant markers associated with ramicorn branching were clustered in 8 linkage groups (Figure 3.9).

For growth traits, 234 SNPs showed association with volume and 80 SNPs were associated with height (q-value≤0.05). There were some common SNPs shared by different traits (Table 3.4). Forking shared 19 associated SPNs with ramicorn branching and 13 with height. Ramicorn branching had 20 common associated SNPs with height and 45 with DBH. Height and DBH had 44 SNP associations in common.

3.3.4 Detection of QTL

In contrast with the SNP association analysis, linkage analysis made use of the linkage map for detecting QTL controlling the traits. By the method of Interval Mapping, the log of odds (LOD) score profiles for fork from two methods (ML and IRLS) were obtained and compared. The profiles from ML and IRLS overlap, giving almost identical results for QTL mapping (Figure 3.10). We therefore will only discuss the results from ML

analysis in the rest of this paper. The critical LOD score for fork was 3.43 from the permutation tests at a significant level of α=0.05, so 3.43 was used as the threshold to detect genome-wide QTL. Across the whole genome, there are two peaks over the threshold: one at the position of 1204.41 cM and the other at 1695.01 cM (Figure 3.10). The closest markers to these two QTL are P2096 and M21186, with LOD scores of 6.038 and 3.483, respectively (Table 3.5). The two QTL account for 84-86% of the variation in estimated clonal genetic values for forking, when the residual variance from the PROC QTL analysis is used in the estimate of total variation.

For ramicorn branching, three QTL exceed the critical LOD score (3.53) at positions of 772.76 cM, 852.99 cM and 872.38 cM, respectively (Figure 3.11). The SNP markers near to the QTL are M21144, P62 and M21191from group 9 and 10 (Table 3.6). The QTL explained 82-87% of the variation in clonal genetic values for ramicorn branching, based on the residual variance from the PROC QTL analysis.

QTL are also found for growth traits. Six QTL are identified for volume, from linkage group 2, 6, 8, 11 and 13 (Figure 3.12), using a LOD threshold of 3.61. The significant SNPs close to the QTL are M21033, P2012, P107, P419, M21266 and P2368. Of the total variation of volume, QTL effects explained 2.7-5.7%. In terms of height, three QTL are found, at positions 656.95 cM, 425.82 cM and 1506.52 cM (Figure 3.13). The linked SNP markers are P107, P2038 and P416. Additionally, one QTL is found for trait straightness, and the nearest marker is M21144 (Figure 3.14). No QTL was found for branch angle.

Through the multiple trait analysis, QTL controlling pairs of traits were identified. From the joint analysis on forking and ramicorn branching, two QTL are detected at the position of 772.76 cM and 1204.41 cM (Figure 3.15). Marker M21144 and P2096 are the closest SNPs, which can used to detect the QTL controlling both traits. Four QTL are identified for stem forking and height, which located at 656.95cM, 774.10cM, 1055.37cM and 1203.01 cM of the genome (Figure 3.16). Four QTL are found for stem forking and volume, as well. They are at positions 656.95 cM, 773.43 cM, 975.75 cM and 1205.11 cM, respectively (Figure 3.17). Three QTL (773.43 cM, 872.38 cM, 974.93 cM) are identified in the joint mapping for ramicorn branching and height (Figure 3.18). For ramicorn branching and volume, seven QTL are found for the two traits, and the QTL are in linkage groups 2, 6 8, 9, 10, 11, and 13 (Figure 3.19).

3.4. Discussion

3.4.1 Statistical techniques

It is more complicated to construct a linkage map in a full-sib family from a cross of outbred parents than in well designed pedigrees from two inbred lines. When performing a linkage analysis in an outcrossing population based on a mixed set of different marker types, some of the genetic assumptions made in traditional linkage analysis do not hold true. There are uncertainties, including a variable number of segregating alleles and differences in parental linkage phases between different markers. In this study, we successfully constructed a linkage map for a full-sib family derived from two outbred

parents by using a new genetic mapping software, OneMap. Based on Wu's methodology (Wu et al. 2002), OneMap performs simultaneous maximum likelihood estimation of linkage and linkage phases between markers (Margarido et al. 2007), while allowing the analysis of a mixed set of different marker types containing various segregation patterns. The algorithm developed by Wu is capable of characterizing the linkage phases based on the posterior probability distribution of the assignment of alternative alleles at given markers to two homologous chromosomes of each parent, conditional on the observed phenotypes of the markers (Wu et al. 2002). This algorithm uses two- and multi-point analyses to estimate the recombination fraction and determine the most likely linkage phase between all possible segregating markers at the same time, so it is very effective compared to the traditional methods (Ritter and Salamini 1996, Ridout et al. 1998). Garcia et al. (2006) reported Wu's approach allowed more accurate construction of co-segregation groups with higher number of markers, than could be achieved using JoinMap. Those authors reported that the simultaneous maximum likelihood estimation of linkage and linkage phases using OneMap was more efficient than JoinMap in building the genetic map of sugarcane. Our research reports the application of OneMap for genetic mapping of outcrossing tree species. OneMap is powerful and user-friendly software for constructing linkage maps for outbred families, but the major disadvantage of this software is the long computational time. It took more than 72 hours to obtain the order of 100 markers in a linkage group using the function "order.seq". In addition, the software doesn't provide the function to draw the map, and the quality of the recombination fraction diagram is low.

GWAF implements association tests between markers and traits by using the family data. By making use of the kinship coefficient matrix in the analysis, false positives due to unexpected familial correlation can be avoided. Association analysis based on family data allows more accurate modeling of the residual correlation within families in the test of genotype-phenotype association. However, there was only one family involved in this study, which violates an assumption on which the software analysis is based; that SNP markers are to be tested for association with traits in a population including multiple families.

Through the SAS procedure PROC QTL, interval mapping was used to detect the QTL .The method implemented in interval mapping was maximum likelihood (ML), in which information on the presence of a QTL was derived from both the mean difference between the flanking marker genotype classes and the distribution of the trait within each marker genotype class (Lander and Botstein 1989). The approach of interval mapping considers one QTL at a time using a single QTL model, so it will be biased if multiple QTL located in the same linkage group. However, the IM approach can still detect multiple QTL, simply by evaluating the number of peaks in the test statistic profile that pass the critical value of the test statistic. Multiple Interval Mapping (MIM) and Composite Interval Mapping (CIM) are probably more powerful to detect QTL when there are multiple QTL in the same linkage group, but CIM approach is not commonly used in outbred populations due to the variable information content along the genome and inconsistent linkage phase among families (Gao and Yang, 2006).

The SAS procedure PROC QTL is convenient to use in analyzing genetic data, because data input and manipulation are done in the environment of SAS. With the powerful functions of SAS for diverse studies, such as PROC SORT, PROC MACRO, and PROC MIXED, the QTL estimation can be integrated with other genetic analysis. Additionally, PROC QTL provides different methods to perform QTL mapping for all line crossing experiments. However, it is unable to calculate the threshold for QTL, which is a shortcoming of this procedure.

3.4.2 QTL for forking defects

Few studies have reported results of QTL mapping for stem form or branch architecture in forest trees. Shepherd et al. (2002) reported that no QTL were detected for ramicorns and double leaders (forking) in *Pinus elliottii var. elliottii*× *Pinus caribaea var. hondurensis* hybrids. In Shepherd's study, markers showing a segregation ratio of 3:1 were excluded from the analysis, in order to use the pseudotestcross strategy. As a result, the power to detect effects of QTL could be low, as a backcross design has only half the power of the F2 intercross design to detect small effect QTL (Soller et al. 1976). In our study, all markers from various cross types were used in constructing the genetic map and providing the most informative linkage map for QTL detection. To our knowledge, this is the first analysis of QTL mapping for forking defects by using different cross types of markers (F2 and backcross) in loblolly pine.

The significant SNPs found using single-marker analysis fall into linked clusters, pinpointing regions of the linkage map that are associated with the genetic control of these traits. The significant SNPs associated with stem forking are largely in linkage groups 11 and 13, and the SNPs associated with ramicorn branching cluster in linkage groups 3, 5 and 9. However, some significant SNPs identified in single-marker analysis are not in linkage groups of the linkage map. The advantage of single-marker analysis is that it can identify the association between QTL and unlinked markers as it does not need a genetic map in the analysis. Although the single marker analysis by GWAF can capture QTL when there is limited recombination or markers are sufficient at that map region, this approach has many limitations. It cannot predict the most probable location of the QTL, and the QTL effects are likely to be underestimated because of the confounding of effects with the recombination frequency (Lander and Botstein 1989). The single-marker analysis in this study was used for preliminary estimation of QTL, and identification of the unlinked SNPs associated with phenotypic variation.

The interval mapping method uses a linkage map to gain more power in detecting QTL (Lander and Botstein 1989). The QTL analysis through interval mapping has identified two and three QTL for forking and ramicorn branching, respectively, with high value of additive QTL effects. It indicates that the forking traits may be controlled by a few genes with large effects, although QTL mapping in a family of 200 individuals is likely to both underestimate the true number of QTL and overestimate the magnitude of the effects of detected QTL (Xu, 2003). Our previous studies have shown that forking is under genetic

control at the family level, with moderate family-mean heritability of (0.59-0.76) (Xiong et al. 2010a). Forking defects were also confirmed to be under strong genetic control at the clonal level in the clonal study, with moderate to high clone mean repeatability (0.67 to 0.86) (Xiong et al. 2010b). The QTL for forking explain 84-86% of the variance in the estimated clonal genetic values used as phenotypes for the QTL analysis. According to Beavis effect, the average estimates of phenotypic variances associated with corrected QTL were greatly overestimated if only 100 progenies were used in the analysis, slightly overestimated when 500 progenies were evaluated, and fairly close to the true magnitude if 1000 progenies were tested (Beavis, 1998). In this study, the detected QTL effects might be biased upward, because the number of progenies evaluated is not big enough.

Heritability plays an important role in determining the value of additive effects from the significant QTL (Grattapaglia et al. 1995). Using clonal replicates of the individuals from a full-sib family in our study should increase the power of QTL mapping, because the clonal replication essentially increases the heritability of the forking traits. It was reported that a sample size of 200 would be needed to detect half of the additive genetic variance at α=0.01, when the within-family heritability is 0.5 and five effective QTL control the trait (Strauss et al. 1992). With the sample size of 217 and clonal clone-mean repeatability of 0.67 to 0.86 for forking traits in this clonal test, the power of estimation of QTL location and effects should be relatively strong in our study.

3.4.3 Clustering of QTL affecting different traits

Two QTL were detected for fork and ramicorn in joint mapping. The co-segregation of QTL for these traits was not surprising because of the moderate genetic correlation between the two traits (r=0.68) (Xiong et al 2010b). It suggests that there are pleiotropic effects for these two forking traits, and will make it possible to decrease both forking defects simultaneously. Forking defects (forks and ramicorns) are clearly associated with height as well as volume, with common QTL identified in the joint mapping. The QTL for forking and ramicorn branching could be used for marker assisted selection for both traits. However, the unfavorable correlations between forking defects and growth will cause negative effects on one trait when selecting for the other.

3.4.4 Marker assisted selection

Forking traits are usually difficult to measure because it is hard to keep the consistency of the measurement by different crews. Additionally, the individual heritability of forking is very low, so conventional selection carried out based on observed phenotypes may not achieve desirable gains in decreasing forking defects. The detection of QTL and the associated molecular markers make it possible to apply marker-assisted selection in tree improvement programs to decrease the incidence of stem forks and ramicorn branches. If the detected QTL really do account for most of the genetic variation of forking traits, selection based on markers for these QTL will likely result in greater genetic gain than can be obtained through phenotypic selection. The two parents of this full-sib family have

been used in the NCSU Cooperative Tree Improvement loblolly pine breeding program, so the significant markers identified from this full-sib family will have future applications in the breeding population. In the third cycle of loblolly pine breeding, more focus has been put into within-family selection by using clonal tests (McKeand et al. 2003, McKeand and Bridgwater 1998). Large amount of linkage disequilibrium by hybridization and non-additive genetic variation can be captured through vegetative propagations, which provides favorable conditions for MAS in loblolly pine (Lande and Thompson 1990). Favorable loci segregating within families could efficiently tagged by markers and used for marker-assisted within-family selection for elite individuals. Clonal testing could be an efficient way to track the inheritance and segregation of important QTL on an individual basis. Using the trait-associated markers to pre-screen individual clones would allow early selection of good germplasm for field testing or operational plantations.

Marker-assisted selection could accelerate introgression of desired genes into breeding populations and shorten the breeding cycles, which would be most relevant for traits such as resistance to fusiform rust disease in loblolly pine. This trait has been shown to be controlled by individual genes with major effects on phenotype (Wilcox et al., 1996). Moreover, the development of new technologies for high throughout genotyping makes the cost of MAS much lower than before, allowing more cost-effective applications of marker-assisted selection in tree breeding program. However, limitations of marker-assisted selection are still exist and cannot be ignored.

The QTL mapping strategies for outbred populations described to date have used individual families within a single population to increase the degree of linkage disequilibrium, but single family analysis may limit the utility of marker-trait associations for MAS in other families. The recombination distance will change among different genetic background, which will affect the confidence of the markers used for QTL detection (Strauss et al. 1992). Linkage phase relationships between markers and QTL will differ among pedigrees, leading to questions about the generalization of QTL findings from experimental populations to breeding populations. In addition, loblolly pine is a long-lived organism, so the consistency of QTL may vary across ages. The modification in QTL expression over ages has already been reported in poplar for basal area (Bradshaw and Stettler 1995) and *Eucalyptus* for multiple traits (Verhaegen et al. 1997). Plomion et al. (1996) also found that different loci may be involved in the genetic control of height-growth at different ages in *Pinus pinaster*. Another problem for MAS is the interaction of QTL and environment. A QTL study in chemical contents found that there were significant differences among the populations from North Carolina and Oklahoma, and it was suggested that the QTL interacts with environmental location (Sewell et al. 2002).

We had assessed forking traits at four different sites in this study and did not find the interaction of site by clones. There was not strong evidence of G×E for forking, but if it is necessary, QTL may need to be verified in different environments to increase the reliability of MAS. The correlation of forking and ages wasn't assessed in this study, but

a study in ash (*Fraxinus excelsior*) found that the proportion of forked trees was much higher in later ages than in the preceding 4 years (Kerr and Boswell 2001). If there is significant QTL× age for forking traits, QTL may need to be identified at different ages.

In this study, the putative QTL controlling forking defects were identified from a single family, so the application of the QTL is limited in this family. In order to know the general utility of QTL to different populations, more families need to be involved in the further study. A good strategy is using mating designs to uncover the QTL, combining classic genetic analysis and QTL study by analyzing multiple mapping families (Verhoeven et al. 2006). Through a mating design with multiple families, this approach can explore the genetic architecture, as well as the underlying QTL variation in quantitative traits. Verhoeven (2006) made use of classic diallel analysis for QTL mapping and found that the QTL detection power was greatest for the mating design with the fewest but largest families. Hence, using multiple families with large number of individuals per family will raise the efficiency to detect the QTL and increase the general utility of the QTL from experimental populations to breeding populations.

As the cost of marker genotyping is going down, the number of families that can be used for QTL mapping is no longer the big constraint for MAS. The power of detecting the true association between marker and trait will be greatly increased as efficient high-throughput genotyping assays and powerful statistical analysis tools become accessible. As loblolly pine improvement programs enter into new breeding cycles and reach a

relatively high level of sophistication, MAS should be considered as one of the important strategies in achieving more genetic gains.

Acknowledgements

We thank all the members of Neale Lab at Department of Plant Sciences, University California at Davis for assistance with genotyping. We also acknowledge MeadWestvaco for providing the clonal trials used in these analyses.

3.5. References

Appleby, N., D. Edwards and J. Batley. 2009. New technologies for ultra-high throughput genotyping in plants. *Methods Mol.Biol.* **513**:19-39.

Beavis, W.D. QTL analysis: power, precision, and accuracy, in: Paterson A.H. (Ed.), Molecular Dissection of Complex Traits, CRC Press, Boca Raton, 1998, pp. 145–162.

Bo-Keun Ha, R.S. Hussey and H.R. Boerma. 2007. Development of SNP Assays for Marker-Assisted Selection of Two Southern Root-Knot Nematode Resistance QTL in Soybean. *Crop Sci.* **47**:S73-S82.

Bradshaw, H.D., J. and R.F. Stettler. 1995. Molecular genetics of growth and development in Populus. IV. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. *Genetics.* **139**(2):963-973.

Brown, G.R., D.L. Bassoni, G.P. Gill, J.R. Fontana, N.C. Wheeler, R.A. Megraw, M.F. Davis, M.M. Sewell, G.A. Tuskan and D.B. Neale. 2003. Identification of quantitative trait loci influencing wood property traits in loblolly pine (*Pinus taeda L.*). III. QTL verification and candidate gene mapping. *Genetics.* **164**(4):1537-1546.

Brown, G.R., G.P. Gill, R.J. Kuntz, C.H. Langley and D.B. Neale. 2004. Nucleotide diversity and linkage disequilibrium in loblolly pine. *Proc.Natl.Acad.Sci.U.S.A.* **101**(42):15255-15260.

Chen, M. and Q. Yang. 2010. GWAF: an R package for genome-wide association analyses with family data. *Bioinformatics.* **26**(4):580-581.

Churchill, G.A. and R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. *Genetics.* **138**(3):963.

Dabney, A. and J.D. Storey. 2009. Q-value estimation for false discovery rate control. R package version 1.20.0.

Dakin, E.E. and J.C. Avise. 2004. Microsatellite null alleles in parentage analysis. *Heredity.* **93**(5):504-509.

Devey, M.E., K.D. Jermstad, C.G. Tauer and D.B. Neale. 1991. Inheritance of RFLP loci in a loblolly pine three-generation pedigree. *Theor.Appl.Genet.* **83**(2):238-242.

Devey, M.E., T.A. Fiddler, B.H. Liu, S.J. Knapp and D.B. Neale. 1994. An RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. *Theor.Appl.Genet.* **88**(3):273-278.

Echt, C.S. and C.D. Nelson. 1997. Linkage mapping and genome length in eastern white pine (Pinus strobus L.). *Theor.Appl.Genet.* **94**(8):1031-1039.

Falconer, D.S. and Mackay, T.F.C. 1996. Introduction to quantitative genetics. 4 edn. Longman Group Limited. Harlow: 464p.

Fang, M., D. Jiang, J.P. Li, J.G. Hui, P. Ji, H.Y. Wang and Q.Y. Run. 2008. Multitrait analysis of quantitative trait loci using Bayesian composite space approach. **9**:1-11.

Gao, H., and R. Yang. 2006. Composite interval mapping of QTL for dynamic traits. **51**(15):1857-1862.

Garcia, A.A.F., E.A. Kido, A.N. Meza, H.M.B. Souza, L.R. Pinto, M.M. Pastina, C.S. Leite, J.A.G. Silva, E.C. Ulian, A. Figueira and A.P. Souza. 2006. Development of an integrated genetic map of a sugarcane (Saccharum spp.) commercial cross, based on a maximum-likelihood approach for estimation of linkage and linkage phases. *Theor. Appl. Genet.* 112(2):298-314.

Goebel, N.B. and J.R. Warner. 1966. Volume tables for small diameter loblolly, short leaf and Virginia pine in the upper South Carolina Piedmont. *For. Res. Serv.* **No. 7**.

González-Martínez, S.C., D. Huber, E. Ersoz, J.M. Davis and D.B. Neale. 2008. Association genetics in Pinus taeda L. II. Carbon isotope discrimination. *Heredity.* **101**(1):19-26.

González-Martínez, S.C., E. Ersoz, G.R. Brown, N.C. Wheeler and D.B. Neale. 2006. DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in Pinus taeda L. *Genetics.* **172**(3):1915-1926.

González-Martínez, S.C., N.C. Wheeler, E. Ersoz, C.D. Nelson and D.B. Neale. 2007. Association genetics in Pinus taeda L. I. Wood property traits. *Genetics.* **175**(1):399-409.

Grattapaglia, D., F.L. Bertolucci and R.R. Sederoff. 1995. Genetic mapping of QTLs controlling vegetative propagation in Eucalyptus grandis and E. urophylla using a pseudo-testcross strategy and RAPD markers. *Theor.Appl.Genet.* **90**(7-8):933.

Groover, A. M. Devey, T. Fiddler, J. Lee, R. Megraw, T. Mitchel-Olds, B. Sherman, S. Vujcic, C. Williams, and D. Neale. 1994. Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. *Genetics.* **138**(4):1293-1300.

Gut, I.G. 2001. Automation in genotyping of single nucleotide polymorphisms. *Hum.Mutat.* **17**(6):475-492.

Haley, C.S., S.A. Knott and J.M. Elsen. 1994. Mapping quantitative trait loci in crosses between outbred lines using lease squares. *Genetics.* **136**(3):1195-1207.

Han, L. and S. Xu. 2008. A Fisher scoring algorithm for the weighted regression method of QTL mapping. *Heredity.* **101**(5):453-464.

Hayashi, T. and T. Awata. 2005. Bayesian mapping of QTL in outbred F2 families allowing inference about whether F0 grandparents are homozygous or heterozygous at QTL. *Heredity.* **94**(3):326-337.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Statist*. 6(2):65-70.

Hu, Z. and S. Xu. 2009. PROC QTL-A SAS Procedure for Mapping Quantitative Trait Loci. *Int J Plant Genomics.* **2009**:141234-141234.

Iwata, H. and S. Ninomiya. 2006. AntMap: constructing genetic linkage maps using an ant colony optimization algorithm. *Breed. Sci.* **56**(4):371-377.

Iwata, H. and S. Ninomiya. 2006. AntMap: constructing genetic linkage maps using an ant colony optimization algorithm. *Breed. Sci.* **56**(4):371-377.

Jansen, R.C. 1993. Interval mapping of multiple quantitative trait loci. *Genetics.* **135**: 205-211

Jarnea, P. and P.J.L. Lagodab. 1996. Microsatellites, from molecules to populations and back. **11**(10):424-429.

Jourjon, M., S. Jasson, J. Marcel, B. Ngom and B. Mangin. 2005. MCQTL: multi-allelic QTL mapping in multi-cross design. *Bioinformatics.* **21**(1):128-130.

Kao, C. and Z. Zeng. 1999. Multiple interval mapping for quantitative trait loci. *Genetics.* **152**:1203-1216.

Kerr, G. and C. Boswell. 2001. The influence of spring frosts, ash bud moth (*Prays fraxinella*) and site factors on forking of young ash (*Fraxinus excelsior*) in southern Britain. *Forestry.* **74**(1):29-40.

Knott, S.A., D.B. Neale, M.M. Sewell and C.S. Haley. 1997. Multiple marker mapping of quantitative trait loci in an outbred pedigree of loblolly pine. *Theor.Appl.Genet.* **94**(6-7):810-820.

Kondo, T., K. Terada, E. Hayashi, N. Kuramoto, M. Okamura and H. Kawasaki. 2000. RAPD markers linked to a gene for resistance to pine needle gall midge in Japanese black pine (Pinus thunbergii). *Theor.Appl.Genet.* **100**(3-4):391-395.

Lander, E.S. and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics.* **121**(1):185-199.

Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daly, S.E. Lincoln and L. Newburg. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics.* **1**(2):174-181.

Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daly, S.E. Lincoln and L. Newburg. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics.* **1**(2):174-181.

Lerceteau, E., C. Plomion and B. Andersson. 2000. AFLP mapping and detection of quantitative trait loci (QTLs) for economically important traits in Pinus sylvestris: a preliminary study. *Mol.Breed.* **6**(5):451-458.

Maheswaran, M., P.K. Subudhi, S. Nandi, J.C. Xu, A. Parco, D.C. Yang and N. Huang. 1997. Polymorphism, distribution, and segregation of AFLP markers in a doubled haploid rice population. *Theor.Appl.Genet.* **94**(1):39-45.

Margarido, G.R.A., A.P. Souza and A.A.F. Garcia. 2007. OneMap: software for genetic mapping in outcrossing species. *Hereditas.* **144**(3):78-79.

Marques, C.M., R.P.V. Brondani, D. Grattapaglia and R. Sederoff. 2002. Conservation and synteny of SSR loci and QTLs for vegetative propagation in four Eucalyptus species. *Theor.Appl.Genet.* **105**(2):474-478.

McKeand, S., T. Mullin, T. Byram and T. White. 2003. Deployment of genetically improved loblolly and slash pines in the South. *J.For.* **101**(3):32-37.

McKeand, S.E. and F.E. Bridgwater. 1998. A strategy for the third breeding cycle of loblolly pine in the Southeastern U.S. *Silvae Genet.* **47**(4):223-234.

Morgante, M. and A.M. Olivieri. 1993. PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal*. **3**(1):175-187.

Neale, D.B., C.S. Kinlaw and M.M. Sewell. 1994. Genetic mapping and DNA sequencing of the loblolly pine genome. *For Genet.* **1**(4):197-206.

Neale, D.B., M.M. Sewell and G.R. Brown. 2002. Molecular dissection of the quantitative inheritance of wood property traits in loblolly pine. *Ann.For.Sci.* **59**(5):595-605.

Nelson, C.D. and T.L. Kubisiak. 1994. A genetic linkage map of longleaf pine (*Pinus palustris* Mill.) based on random amplified. *J.Hered.* **85**(6):433-439.

Nelson, J.C. 1997. QGENE: software for marker-based genomic analysis and breeding. *Mol. Breed.* 3(3):239-245.

Plomion, C., C.E. Durel and D.M. O'Malley. 1996. Genetic dissection of height in maritime pine seedlings raised under accelerated growth conditions. *Theor.Appl.Genet.* 93(5):849-858.

Rafalski, A. 2002. Applications of single nucleotide polymorphisms in crop genetics. *Curr.Opin.Plant Biol.* 5(2):94-100.

Remington, D.L., R.W. Whetten, B.-H. Liu and D.M. O'Malley. 1999. Construction of an AFLP genetic map with nearly complete genome coverage in Pinus taeda. *Theor.Appl.Genet.* **98**(8):1279-1292.

Ridout, M.S., S. Tong, C.J. Vowden, K.R. Tobutt and E. Ritter. 1998. Three-point linkage analysis in crosses of allogamous plant species. *Genet.Res.* **72**(2):111-121.

Ritter, E. and F. Salamini. 1996. The calculation of recombination frequencies in crosses of allogamous plant species with applications to linkage mapping. *Genet.Res.* **67**(1):55-65.

SAS INSTITUTE INC. 2002-2004. SAS 9.1.3 Help and Documentation. Cary, NC:SAS Institute Inc.

Satagopan, J.M., B.S. Yandell, M.A. Newton and T.C. Osborn. 1996. A Bayesian approach to detect quantitative trait loci using Markov chain Monte Carlo. *Genetics.* **144**(2):805-816.

Seaton, G., C.S. Haley, S.A. Knott, M. Kearsey and P.M. Visscher. 2002. QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics.* **18**(2):339-340.

Sewell, M.M., B.K. Sherman and D.B. Neale. 1999. A consensus map for loblolly pine (Pinus taeda L.). I. Construction and integration of individual linkage maps from two outbred three-generation pedigrees. *Genetics.* **151**(1):321-330.

Sewell, M.M., M.F. Davis, G.A. Tuskan, N.C. Wheeler, C.C. Elam, D.L. Bassoni and D.B. Neale. 2002. Identification of QTLs influencing wood property traits in loblolly pine (Pinus taeda L.). II. Chemical wood properties. *Theor.Appl.Genet.* **104**(2):214-222.

Shepherd, M. 2002. Branch architecture QTL for Pinus elliottii var. elliottii× Pinus caribaea var. hondurensis hybrids. *Ann.For.Sci.* **59**(5-6):617.

Smith, D.N. and M.E. Devey. 1994. Occurrence and inheritance of microsatellites in Pinus radiata. *Genome.* **37**(6):977-983.

Soller, M., T. Brody and A. Genizi. 1976. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theor.Appl.Genet.* **47**(1):35-39.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J.Mol.Biol.* **98**(3):503-517.

Stam, P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: JOINMAP. *Plant J.* **3**(5):739-744.

Storey, J.D. 2002. A direct approach to false discovery rates. *J. R. Statist. Soc. B* **64**(3):479-498.

Strauss, S.H., R. Lande and G. Namkoong. 1992. Limitations of molecular-marker-aided selection in forest tree breeding. *Can. J. For. Res*. **22**(7):1050-1061.

Tinker, N.A. and D.E. Mather. 1995. MQTL: software for simplified composite interval mapping of QTL in multiple environments. J. QTL. 1:1-5.

Travis, S.E., K. Ritland, T.G. Whitham and P. Keim. 1998. A genetic linkage map of Pinyon pine (Pinus edulis) based on amplified fragment length polymorphisms. *Theor.Appl.Genet.* **97**(5):871-880.

Verhaegen, D., C. Plomion, J.-. Gion, M. Poitel, P. Costa and A. Kremer. 1997. Quantitative trait dissection analysis in Eucalyptus using RAPD markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. *Theor.Appl.Genet.* **95**(4):597-608.

Verhoeven, K.J.F., J.-. Jannink and L.M. McIntyre. 2006. Using mating designs to uncover QTL and the genetic architecture of complex traits. *Heredity.* **96**(2):139-149.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T.v.d. Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman and M. Kuiper. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**(21):4407-4414.

Wang, S., C.J. Basten, P. Gaffney and Z.-. Zeng. 2003. Windows QTL Cartographer version 2.0. Statistical Genetics,North Carolina State University, Ralgeigh, NC.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski and S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* **18**(22):6531-6535.

Wu, R., C.X. Ma, I. Painter and Z.B. Zeng. 2002. Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. *Theor.Popul.Biol.* **61**(3):349-363.

Wu, R.L., D.L. Remington, J.J. MacKay, S.E. McKeand and D.M. O'Malley. 1999. Average effect of a mutation in lignin biosynthesis in loblolly pine. *Theor.Appl.Genet.* **99**(3-4):705-710.

Wu, Y., T.J. Close and S. Lonardi. 2008. On the accurate construction of consensus genetic maps. *Comput Syst Bioinformatics Conf.* **7**:285-296.

Xiong, S.J., F. Isik, S.E. McKeand and R.W. Whetten. 2010a. Genetic variation of stem forking in loblolly pine. *For.Sci.* (in press).

Xiong, S.J., S. E. McKeand, R. W. Whetten, F. Isik. 2010b. The genetics of growth and forking from a loblolly pine clonal test. Ph.D Dissertation, Chapter2.

Xu, S. 1998a. Further investigation on the regression method of mapping quantitative trait loci. *Heredity.* **80**(3):364-373.

Xu, S. 1998b. Iteratively reweighted least squares mapping of quantitative trait loci. *Behav.Genet.* **28**(5):341-355.

Zeng, Z. 1994. Precision mapping of quantitative trait loci. *Genetics.* **136**:1457-1468.

Table 3.1 Different cross types of markers, three types of markers B3.7, D1.10 and D2.15 were used for constructing linkage mapping, coded for OneMap.

Crosstype	Parent cross	Offspring observed band	Segregation	# of markers	Configurations
1	aa×aa	a		3075	non-informative
2	aa×bb	ab		148	non-informative
B3.7	ab×ab	a,2ab,b	1:2:1	304	F2 intercross
D1.10	aa×ab	a,ab	1:1	552	Backcross 1
D2.15	ab×aa	a,ab	1:1	696	Backcross 2

Table 3.2 The significant markers for forking using Bonferroni correction with α=0.05. (Phen: trait phenotype; h^2_q: the portion of phenotypic variation explained by the SNP; beta: the regression coefficient of SNP covariate; se: standard error of beta; Chi-square statistic for testing beta not equal to zero; p-val: p-value of the Chi-square statistic.)

Marker	Phen	h^2_q	beta	se	chisq	model	p-val
M21127	fork	0.105	-0.042	0.008	25.415	additive	4.62E-07
P233	fork	0.083	0.052	0.012	19.497	additive	1.01E-05
P370	fork	0.081	0.050	0.012	18.694	additive	1.53E-05
P502	fork	0.083	0.053	0.012	18.136	additive	2.06E-05
P513	fork	0.077	-0.050	0.012	18.092	additive	2.11E-05
P636	fork	0.077	-0.050	0.012	18.092	additive	2.11E-05
P101	fork	0.077	0.050	0.012	17.944	additive	2.28E-05
P470	fork	0.076	0.050	0.012	17.751	additive	2.52E-05
P645	fork	0.076	0.050	0.012	17.666	additive	2.63E-05
P509	fork	0.074	0.049	0.012	17.296	additive	3.20E-05
M21246	fork	0.072	-0.036	0.009	16.823	additive	4.10E-05

Table 3.3 The significant markers for ramicorn branching using Bonferroni correction with α=0.05. (Phen: trait phenotype; h^2_q: the portion of phenotypic variation explained by the SNP; beta: the regression coefficient of SNP covariate; se: standard error of beta; Chi-square statistic for testing beta not equal to zero; p-val: p-value of the Chi-square statistic.

Marker	Phen	h^2_q	beta	se	chisq	model	p-val
P591	rami	0.095	0.045	0.010	21.431	additive	3.67E-06
P244	rami	0.088	-0.044	0.010	20.870	additive	4.92E-06
P592	rami	0.084	0.043	0.010	19.749	additive	8.83E-06
P409	rami	0.078	0.041	0.010	18.436	additive	1.76E-05
P2093	rami	0.076	-0.040	0.010	17.810	additive	2.44E-05
P479	rami	0.076	-0.041	0.010	17.682	additive	2.61E-05
P2408	rami	0.074	0.040	0.010	17.400	additive	3.03E-05
P2257	rami	0.077	0.040	0.010	17.253	additive	3.27E-05
P258	rami	0.077	0.041	0.010	17.229	additive	3.31E-05

Table 3.4 The number of significant SNPs shared by different traits at false discovery rate of 5% ($q<0.05$).

	HT	**DBH**	**VOL**	**Strt**	**BrAngle**	**Ramicorn**
Fork	13	18	17	5	9	19
HT		44	49	24	0	20
DBH			188	22	15	45
VOL				31	15	47
Strt					11	18
BrAngle						13

Table 3.5 Estimated QTL parameters for forking from the method of Interval Mapping. (μ is the mean; a_1 is the additive effects from parent1; a_2 is the additive effects from prarent2; The standard errors of estimated additive effects are given in parentheses;σ^2_e is the residual variance.)

QTL	Nearest marker	Group	LOD	Position(cM)	μ	a_1	a_2	σ^2_e
FQ1	P2096	13	6.038	1205.11	0.154	0.026(0.006)	0.020(0.006)	0.007
FQ2	M21186	18	3.483	1695.01	0.156	-0.019(0.006)	-0.017(0.006)	0.007

Table 3.6. Estimated QTL parameters for ramicorn branching. (μ is the mean; $a_{\_1}$ is the additive effects from parent1; $a_{\_2}$ is the additive effects from prarent2; The standard errors of estimated additive effects are given in parentheses;σ^2_e is the residual variance.)

QTL	Nearest marker	Group	LOD	Position(cM)	μ	$a_{\_1}$	$a_{\_2}$	σ^2_e
RQ1	M21144	9	3.773	772.09	0.442	-0.004(0.005)	-0.020(0.005)	0.005
RQ2	P62	10	3.645	852.79	0.441	-0.015(0.005)	0.016(0.005)	0.005
RA3	M21191	10	4.848	873.33	0.443	-0.016(0.005)	0.017(0.005)	0.005

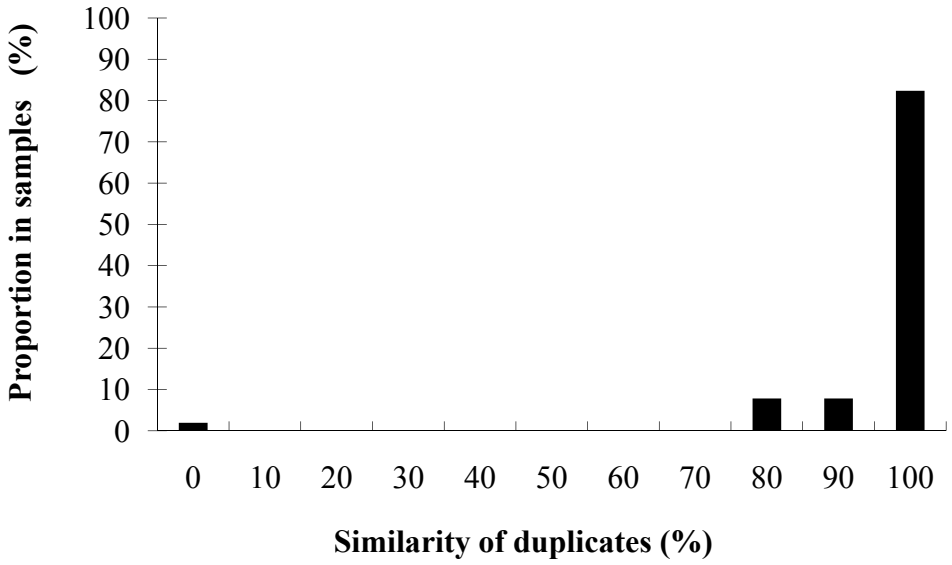
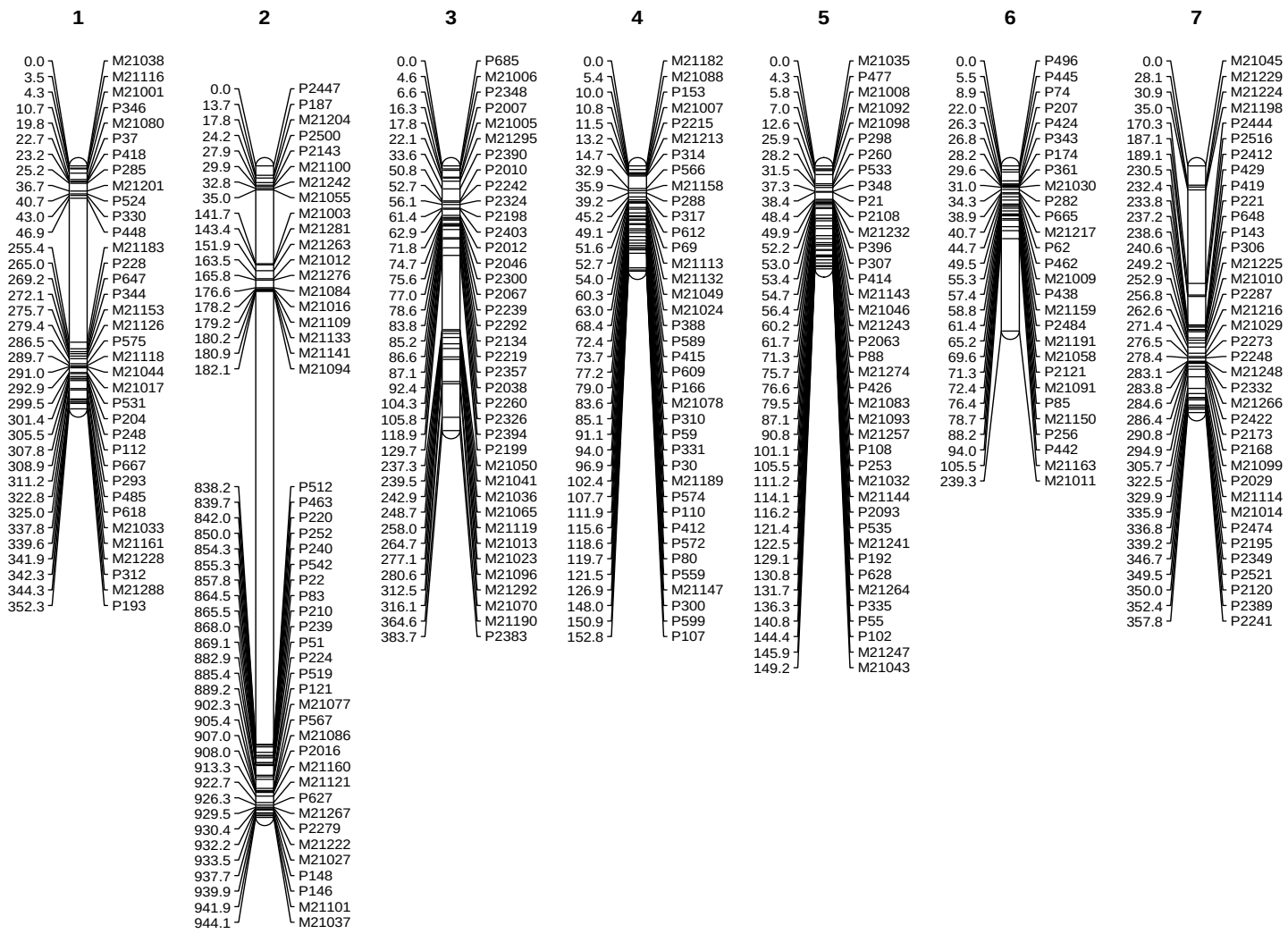
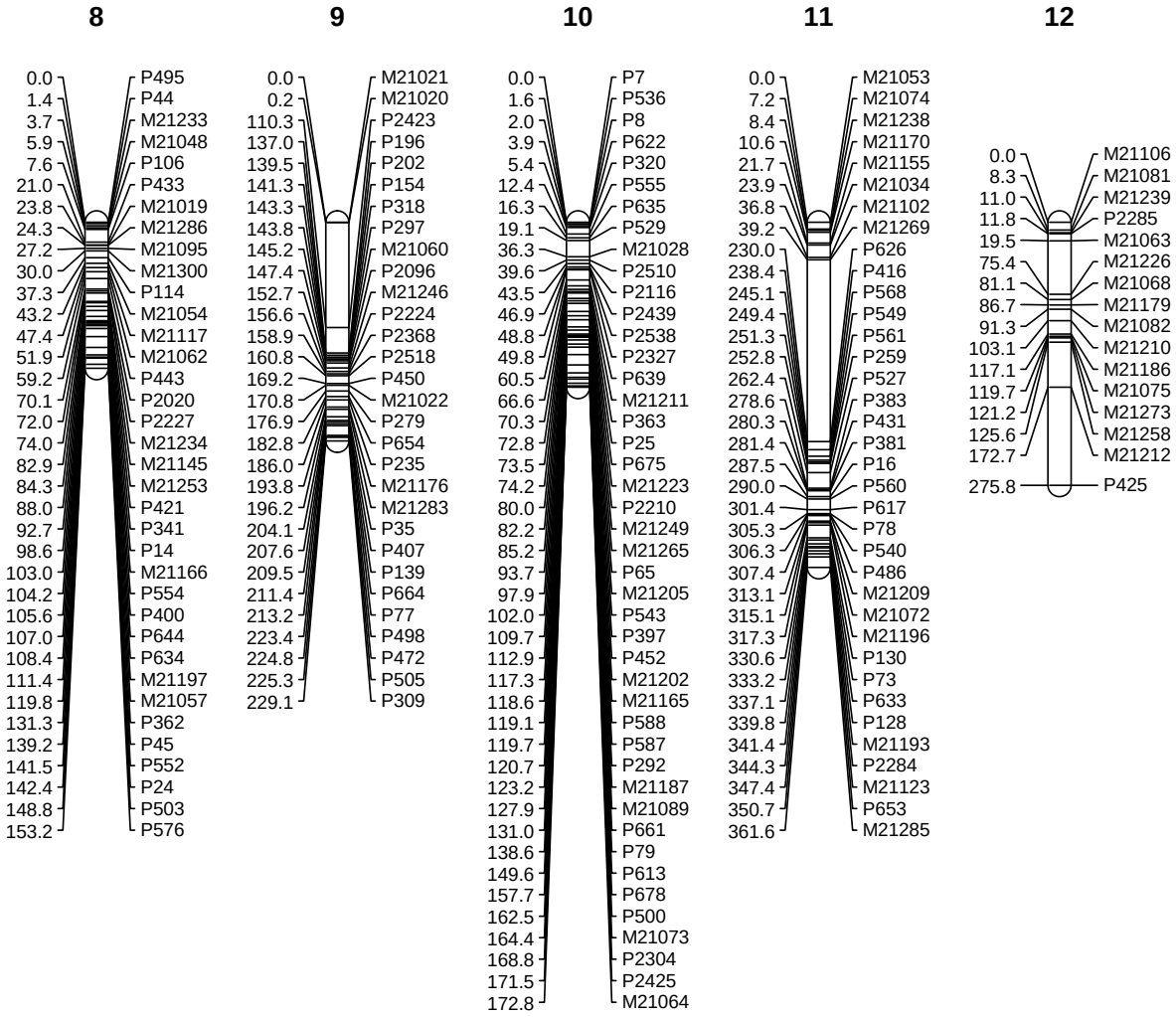


Figure 3.1 Similarity of repeated samples in genotyping.

Figure 3.2 First constructed genetic map of 12 groups.


1
0.0 M21038
3.5 M21116
4.3 M21001
10.7 P346
19.8 M21080
22.7 P37
23.2 P418
25.2 P285
36.7 M21201
40.7 P524
43.0 P330
46.9 P448
255.4 M21183
265.0 P228
269.2 P647
272.1 P344
275.7 M21153
279.4 M21126
286.5 P575
289.7 M21118
291.0 M21044
292.9 M21017
299.5 P531
301.4 P204
305.5 P248
307.8 P112
308.9 P667
311.2 P293
322.8 P485
325.0 P618
337.8 M21033
339.6 M21161
341.9 M21228
342.3 P312
344.3 M21288
352.3 P193
2
0.0 P2447
13.7 P187
17.8 M21204
24.2 P2500
27.9 P2143
29.9 M21100
32.8 M21242
35.0 M21055
141.7 M21003
143.4 M21281
151.9 M21263
163.5 M21012
165.8 M21276
176.6 M21084
178.2 M21016
179.2 M21109
180.2 M21133
180.9 M21141
182.1 M21094
838.2 P512
839.7 P463
842.0 P220
850.0 P252
854.3 P240
855.3 P542
857.8 P22
864.5 P83
865.5 P210
868.0 P239
869.1 P51
882.9 P224
885.4 P519
889.2 P121
902.3 M21077
905.4 P567
907.0 M21086
908.0 P2016
913.3 M21160
922.7 M21121
926.3 P627
929.5 M21267
930.4 P2279
932.2 M21222
933.5 M21027
937.7 P148
939.9 P146
941.9 M21101
944.1 M21037
3
0.0 P685
4.6 M21006
6.6 P2348
16.3 P2007
17.8 M21005
22.1 M21295
33.6 P2390
50.8 P2010
52.7 P2242
56.1 P2324
61.4 P2198
62.9 P2403
71.8 P2012
74.7 P2046
75.6 P2300
77.0 P2067
78.6 P2239
83.8 P2292
85.2 P2134
86.6 P2219
87.1 P2357
92.4 P2038
104.3 P2260
105.8 P2326
118.9 P2394
129.7 P2199
237.3 M21050
239.5 M21041
242.9 M21036
248.7 M21065
258.0 M21119
264.7 M21013
277.1 M21023
280.6 M21096
312.5 M21292
316.1 M21070
364.6 M21190
383.7 P2383
4
0.0 M21182
5.4 M21088
10.0 P153
10.8 M21007
11.5 P2215
13.2 M21213
14.7 P314
32.9 P566
35.9 M21158
39.2 P288
45.2 P317
49.1 P612
51.6 P69
52.7 M21113
54.0 M21132
60.3 M21049
63.0 M21024
68.4 P388
72.4 P589
73.7 P415
77.2 P609
79.0 P166
83.6 M21078
85.1 P310
91.1 P59
94.0 P331
96.9 P30
102.4 M21189
107.7 P574
111.9 P110
115.6 P412
118.6 P572
119.7 P80
121.5 P559
126.9 M21147
148.0 P300
150.9 P599
152.8 P107
5
0.0 M21035
4.3 P477
5.8 M21008
7.0 M21092
12.6 M21098
25.9 P298
28.2 P260
31.5 P533
37.3 P348
38.4 P21
48.4 P2108
49.9 M21232
52.2 P396
53.0 P307
53.4 P414
54.7 M21143
56.4 M21046
60.2 M21243
61.7 P2063
71.3 P88
75.7 M21274
76.6 P426
79.5 M21083
87.1 M21093
90.8 M21257
101.1 P108
105.5 P253
111.2 M21032
114.1 M21144
116.2 P2093
121.4 P535
122.5 M21241
129.1 P192
130.8 P628
131.7 M21264
136.3 P335
140.8 P55
144.4 P102
145.9 M21247
149.2 M21043
6
0.0 P496
5.5 P445
8.9 P74
22.0 P207
26.3 P424
26.8 P343
28.2 P174
29.6 P361
31.0 M21030
34.3 P282
38.9 P665
40.7 M21217
44.7 P62
49.5 P462
55.3 M21009
57.4 P438
58.8 M21159
61.4 P2484
65.2 M21191
69.6 M21058
71.3 P2121
72.4 M21091
76.4 P85
78.7 M21150
88.2 P256
94.0 P442
105.5 M21163
239.3 M21011
7
0.0 M21045
28.1 M21229
30.9 M21224
35.0 M21198
170.3 P2444
187.1 P2516
189.1 P2412
230.5 P429
232.4 P419
233.8 P221
237.2 P648
238.6 P143
240.6 P306
249.2 M21225
252.9 M21010
256.8 P2287
262.6 M21216
271.4 M21029
276.5 P2273
278.4 P2248
283.1 M21248
283.8 P2332
284.6 M21266
286.4 P2422
290.8 P2173
294.9 P2168
305.7 M21099
322.5 P2029
329.9 M21114
335.9 M21014
336.8 P2474
339.2 P2195
346.7 P2349
349.5 P2521
350.0 P2120
352.4 P2389
357.8 P2241


8
0.0 P495
1.4 P44
3.7 M21233
5.9 M21048
7.6 P106
21.0 P433
23.8 M21019
24.3 M21286
27.2 M21095
30.0 M21300
37.3 P114
43.2 M21054
47.4 M21117
51.9 M21062
59.2 P443
70.1 P2020
72.0 P2227
74.0 M21234
82.9 M21145
84.3 M21253
88.0 P421
92.7 P341
98.6 P14
103.0 M21166
104.2 P554
105.6 P400
107.0 P644
108.4 P634
111.4 M21197
119.8 M21057
131.3 P362
139.2 P45
141.5 P552
142.4 P24
148.8 P503
153.2 P576
9
0.0 M21021
0.2 M21020
110.3 P2423
137.0 P196
139.5 P202
141.3 P154
143.3 P318
143.8 P297
145.2 M21060
147.4 P2096
152.7 M21246
156.6 P2224
158.9 P2368
160.8 P2518
169.2 P450
170.8 M21022
176.9 P279
182.8 P654
186.0 P235
193.8 M21176
196.2 M21283
204.1 P35
207.6 P407
209.5 P139
211.4 P664
213.2 P77
223.4 P498
224.8 P472
225.3 P505
229.1 P309
10
0.0 P7
1.6 P536
2.0 P8
3.9 P622
5.4 P320
12.4 P555
16.3 P635
19.1 P529
36.3 M21028
39.6 P2510
43.5 P2116
46.9 P2439
48.8 P2538
49.8 P2327
60.5 P639
66.6 M21211
70.3 P363
72.8 P25
73.5 P675
74.2 M21223
80.0 P2210
82.2 M21249
85.2 M21265
93.7 P65
97.9 M21205
102.0 P543
109.7 P397
112.9 P452
117.3 M21202
118.6 M21165
119.1 P588
119.7 P587
120.7 P292
123.2 M21187
127.9 M21089
131.0 P661
138.6 P79
149.6 P613
157.7 P678
162.5 P500
164.4 M21073
168.8 P2304
171.5 P2425
172.8 M21064
11
0.0 M21053
7.2 M21074
8.4 M21238
10.6 M21170
21.7 M21155
23.9 M21034
36.8 M21102
39.2 M21269
230.0 P626
238.4 P416
245.1 P568
249.4 P549
251.3 P561
252.8 P259
262.4 P527
278.6 P383
280.3 P431
281.4 P381
287.5 P16
290.0 P560
301.4 P617
305.3 P78
306.3 P540
307.4 P486
313.1 M21209
315.1 M21072
317.3 M21196
330.6 P130
333.2 P73
337.1 P633
339.8 P128
341.4 M21193
344.3 P2284
347.4 M21123
350.7 P653
361.6 M21285
12
0.0 M21106
8.3 M21081
11.0 M21239
11.8 P2285
19.5 M21063
75.4 M21226
81.1 M21068
86.7 M21179
91.3 M21082
103.1 M21210
117.1 M21186
119.7 M21075
121.2 M21273
125.6 M21258
172.7 M21212
275.8 P425

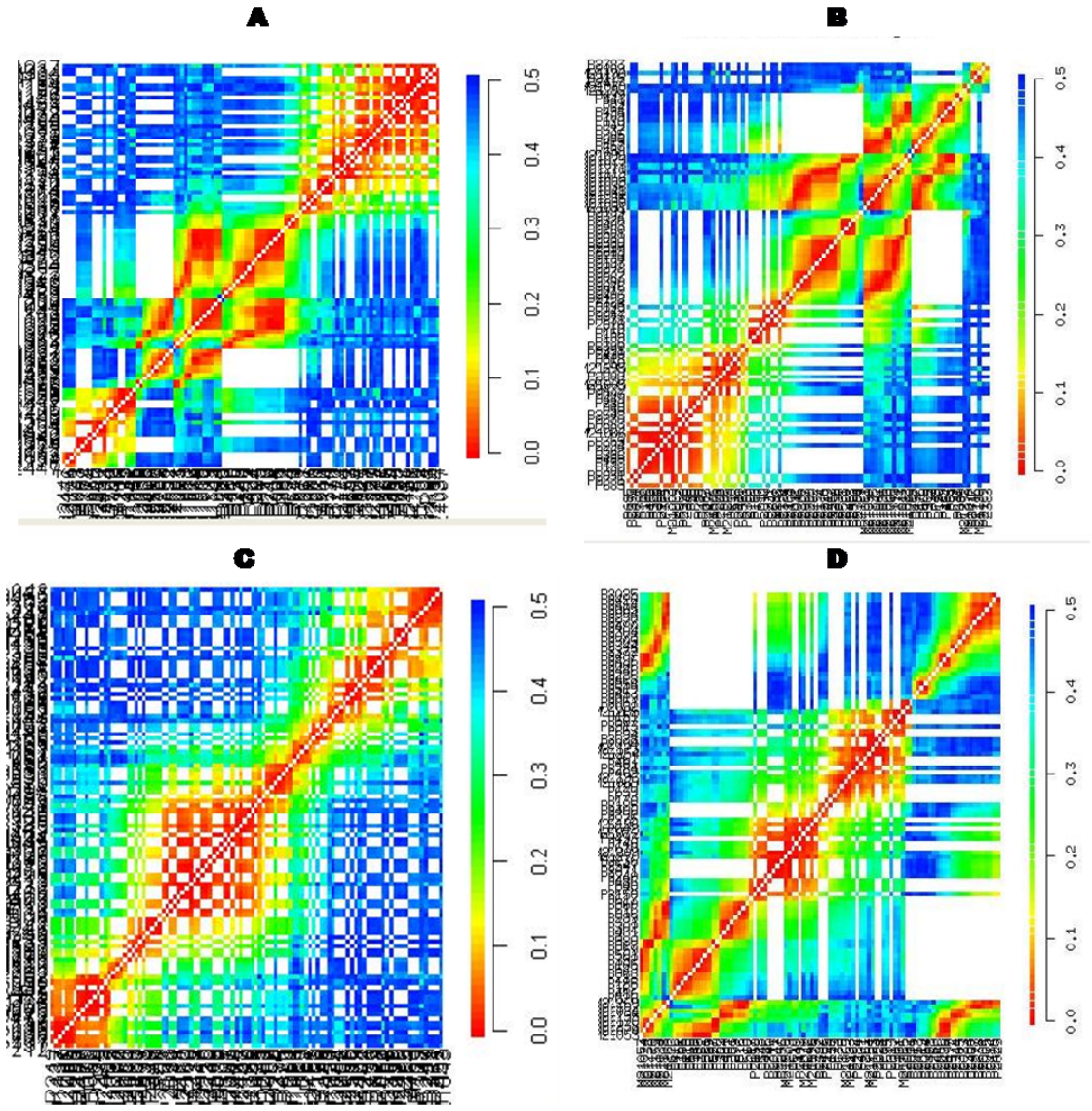
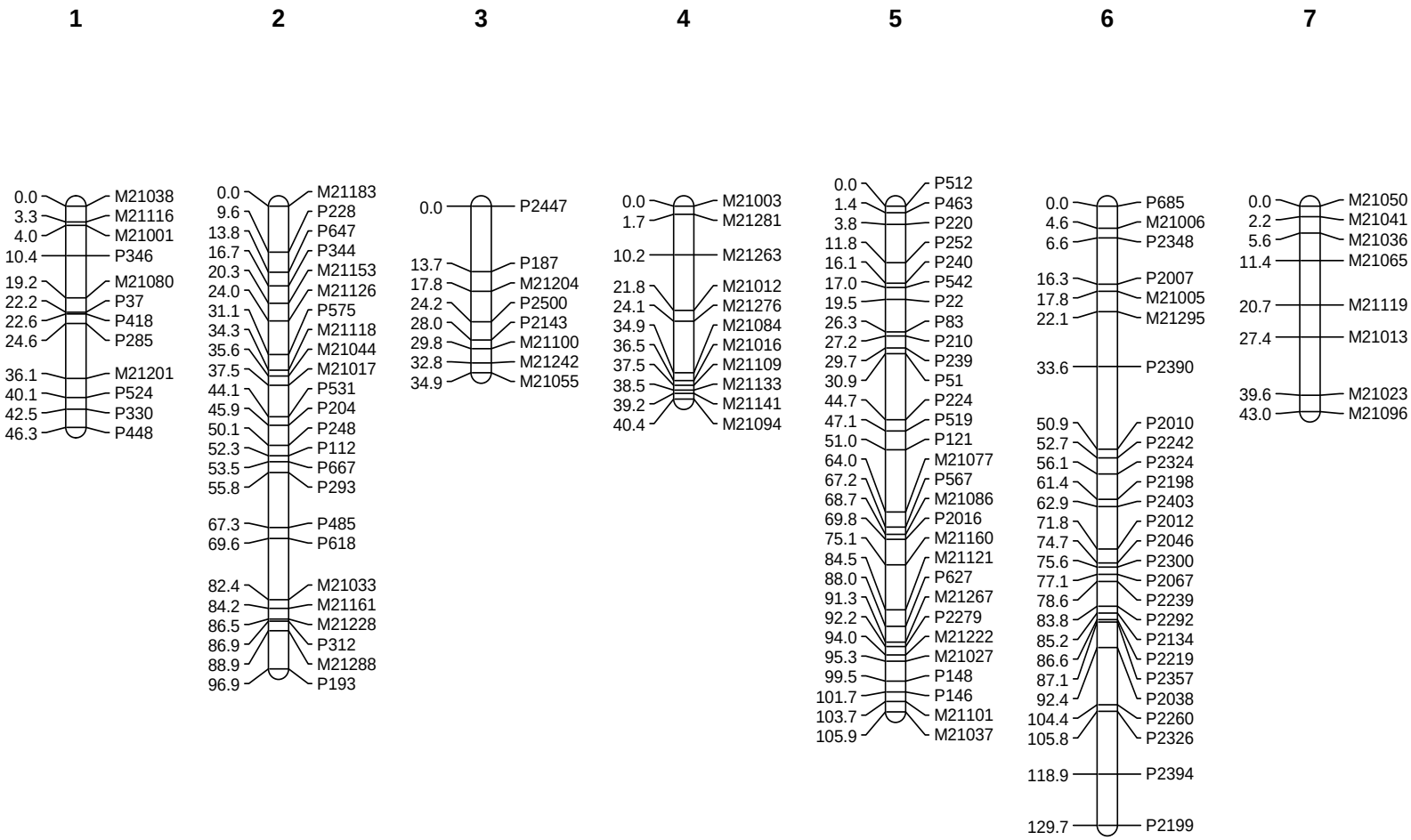
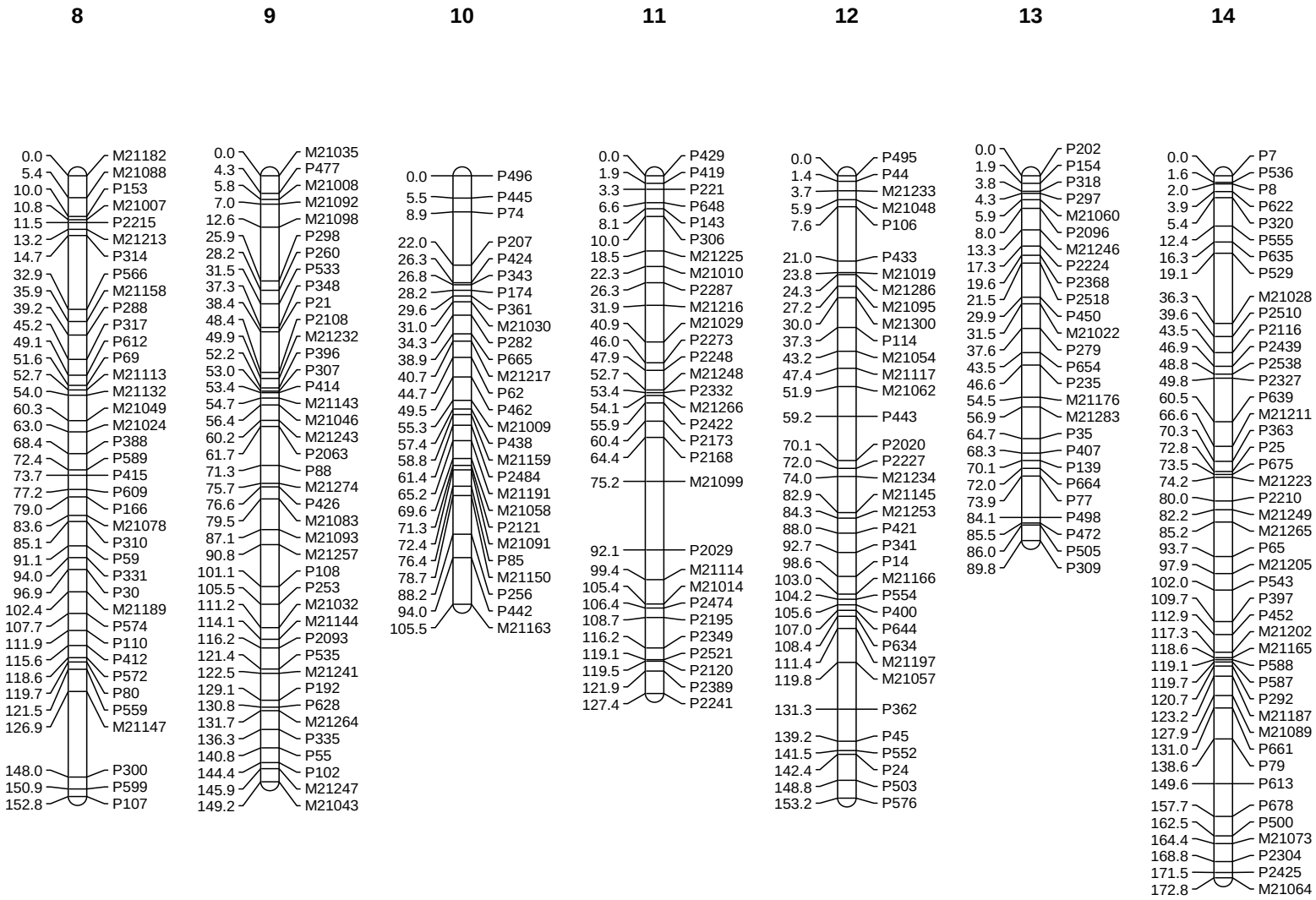
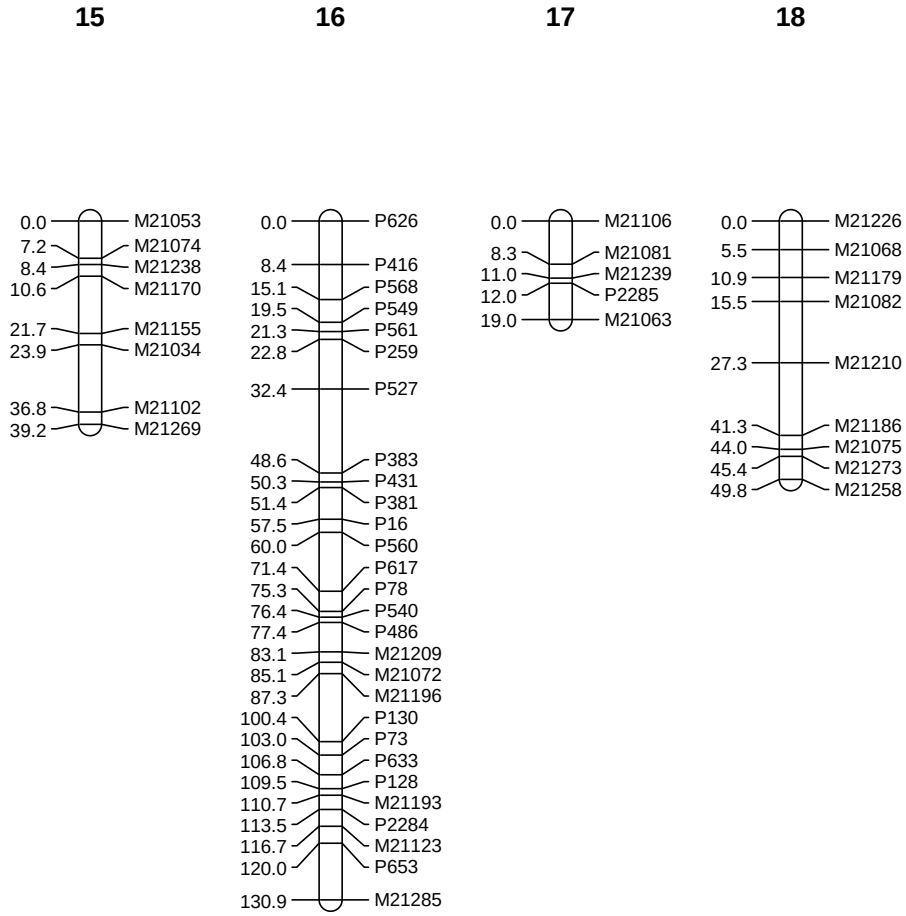


Figure 3.3 Recombination fraction diagram for some of the groups in first genetic mapping (A: Group 2; B: Group 3;C: Group 5;D Group 11).

Figure 3.4 Genetic maps of 18 groups.


1
0.0 M21038
3.3 M21116
4.0 M21001
10.4 P346
19.2 M21080
22.2 P37
22.6 P418
24.6 P285
36.1 M21201
40.1 P524
42.5 P330
46.3 P448
2
0.0 M21183
9.6 P228
13.8 P647
16.7 P344
20.3 M21153
24.0 M21126
31.1 P575
34.3 M21118
35.6 M21044
37.5 M21017
44.1 P531
45.9 P204
50.1 P248
52.3 P112
53.5 P667
55.8 P293
67.3 P485
69.6 P618
82.4 M21033
84.2 M21161
86.5 M21228
86.9 P312
88.9 M21288
96.9 P193
3
0.0 P2447
13.7 P187
17.8 M21204
24.2 P2500
28.0 P2143
29.8 M21100
32.8 M21242
34.9 M21055
4
0.0 M21003
1.7 M21281
10.2 M21263
21.8 M21012
24.1 M21276
34.9 M21084
36.5 M21016
37.5 M21109
38.5 M21133
39.2 M21141
40.4 M21094
5
0.0 P512
1.4 P463
3.8 P220
11.8 P252
16.1 P240
17.0 P542
19.5 P22
26.3 P83
27.2 P210
29.7 P239
30.9 P51
44.7 P224
47.1 P519
51.0 P121
64.0 M21077
67.2 P567
68.7 M21086
69.8 P2016
75.1 M21160
84.5 M21121
88.0 P627
91.3 M21267
92.2 P2279
94.0 M21222
95.3 M21027
99.5 P148
101.7 P146
103.7 M21101
105.9 M21037
6
0.0 P685
4.6 M21006
6.6 P2348
16.3 P2007
17.8 M21005
22.1 M21295
33.6 P2390
50.9 P2010
52.7 P2242
56.1 P2324
61.4 P2198
62.9 P2403
71.8 P2012
74.7 P2046
75.6 P2300
77.1 P2067
78.6 P2239
83.8 P2292
85.2 P2134
86.6 P2219
87.1 P2357
92.4 P2038
104.4 P2260
105.8 P2326
118.9 P2394
129.7 P2199
7
0.0 M21050
2.2 M21041
5.6 M21036
11.4 M21065
20.7 M21119
27.4 M21013
39.6 M21023
43.0 M21096


8
0.0 M21182
5.4 M21088
10.0 P153
10.8 M21007
11.5 P2215
13.2 M21213
14.7 P314
32.9 P566
35.9 M21158
39.2 P288
45.2 P317
49.1 P612
51.6 P69
52.7 M21113
54.0 M21132
60.3 M21049
63.0 M21024
68.4 P388
72.4 P589
73.7 P415
77.2 P609
79.0 P166
83.6 M21078
85.1 P310
91.1 P59
94.0 P331
96.9 P30
102.4 M21189
107.7 P574
111.9 P110
115.6 P412
118.6 P572
119.7 P80
121.5 P559
126.9 M21147
148.0 P300
150.9 P599
152.8 P107
9
0.0 M21035
4.3 P477
5.8 M21008
7.0 M21092
12.6 M21098
25.9 P298
28.2 P260
31.5 P533
37.3 P348
38.4 P21
48.4 P2108
49.9 M21232
52.2 P396
53.0 P307
53.4 P414
54.7 M21143
56.4 M21046
60.2 M21243
61.7 P2063
71.3 P88
75.7 M21274
76.6 P426
79.5 M21083
87.1 M21093
90.8 M21257
101.1 P108
105.5 P253
111.2 M21032
114.1 M21144
116.2 P2093
121.4 P535
122.5 M21241
129.1 P192
130.8 P628
131.7 M21264
136.3 P335
140.8 P55
144.4 P102
145.9 M21247
149.2 M21043
10
0.0 P496
5.5 P445
8.9 P74
22.0 P207
26.3 P424
26.8 P343
28.2 P174
29.6 P361
31.0 M21030
34.3 P282
38.9 P665
40.7 M21217
44.7 P62
49.5 P462
55.3 M21009
57.4 P438
58.8 M21159
61.4 P2484
65.2 M21191
69.6 M21058
71.3 P2121
72.4 M21091
76.4 P85
78.7 M21150
88.2 P256
94.0 P442
105.5 M21163
11
0.0 P429
1.9 P419
3.3 P221
6.6 P648
8.1 P143
10.0 P306
18.5 M21225
22.3 M21010
26.3 P2287
31.9 M21216
40.9 M21029
46.0 P2273
47.9 P2248
52.7 M21248
53.4 P2332
54.1 M21266
55.9 P2422
60.4 P2173
64.4 P2168
75.2 M21099
92.1 P2029
99.4 M21114
105.4 M21014
106.4 P2474
108.7 P2195
116.2 P2349
119.1 P2521
119.5 P2120
121.9 P2389
127.4 P2241
12
0.0 P495
1.4 P44
3.7 M21233
5.9 M21048
7.6 P106
21.0 P433
23.8 M21019
24.3 M21286
27.2 M21095
30.0 M21300
37.3 P114
43.2 M21054
47.4 M21117
51.9 M21062
59.2 P443
70.1 P2020
72.0 P2227
74.0 M21234
82.9 M21145
84.3 M21253
88.0 P421
92.7 P341
98.6 P14
103.0 M21166
104.2 P554
105.6 P400
107.0 P644
108.4 P634
111.4 M21197
119.8 M21057
131.3 P362
139.2 P45
141.5 P552
142.4 P24
148.8 P503
153.2 P576
13
0.0 P202
1.9 P154
3.8 P318
4.3 P297
5.9 M21060
8.0 P2096
13.3 M21246
17.3 P2224
19.6 P2368
21.5 P2518
29.9 P450
31.5 M21022
37.6 P279
43.5 P654
46.6 P235
54.5 M21176
56.9 M21283
64.7 P35
68.3 P407
70.1 P139
72.0 P664
73.9 P77
84.1 P498
85.5 P472
86.0 P505
89.8 P309
14
0.0 P7
1.6 P536
2.0 P8
3.9 P622
5.4 P320
12.4 P555
16.3 P635
19.1 P529
36.3 M21028
39.6 P2510
43.5 P2116
46.9 P2439
48.8 P2538
49.8 P2327
60.5 P639
66.6 M21211
70.3 P363
72.8 P25
73.5 P675
74.2 M21223
80.0 P2210
82.2 M21249
85.2 M21265
93.7 P65
97.9 M21205
102.0 P543
109.7 P397
112.9 P452
117.3 M21202
118.6 M21165
119.1 P588
119.7 P587
120.7 P292
123.2 M21187
127.9 M21089
131.0 P661
138.6 P79
149.6 P613
157.7 P678
162.5 P500
164.4 M21073
168.8 P2304
171.5 P2425
172.8 M21064


15
0.0 M21053
7.2 M21074
8.4 M21238
10.6 M21170
21.7 M21155
23.9 M21034
36.8 M21102
39.2 M21269
16
0.0 P626
8.4 P416
15.1 P568
19.5 P549
21.3 P561
22.8 P259
32.4 P527
48.6 P383
50.3 P431
51.4 P381
57.5 P16
60.0 P560
71.4 P617
75.3 P78
76.4 P540
77.4 P486
83.1 M21209
85.1 M21072
87.3 M21196
100.4 P130
103.0 P73
106.8 P633
109.5 P128
110.7 M21193
113.5 P2284
116.7 M21123
120.0 P653
130.9 M21285
17
0.0 M21106
8.3 M21081
11.0 M21239
12.0 P2285
19.0 M21063
18
0.0 M21226
5.5 M21068
10.9 M21179
15.5 M21082
27.3 M21210
41.3 M21186
44.0 M21075
45.4 M21273
49.8 M21258

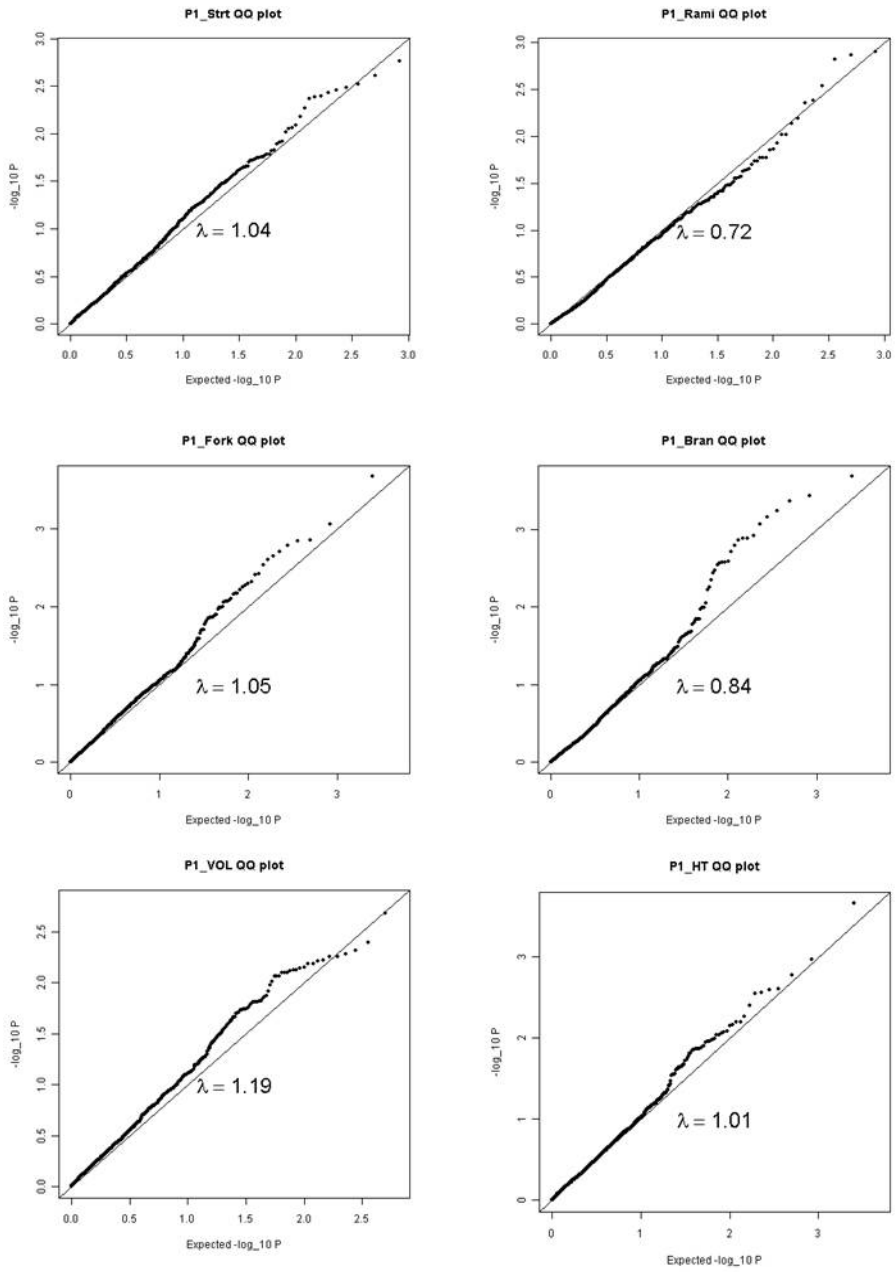


Figure 3.5 QQ Plot from permutation data sets. (λ measures deviation of chi-square test results from expected null distribution.)

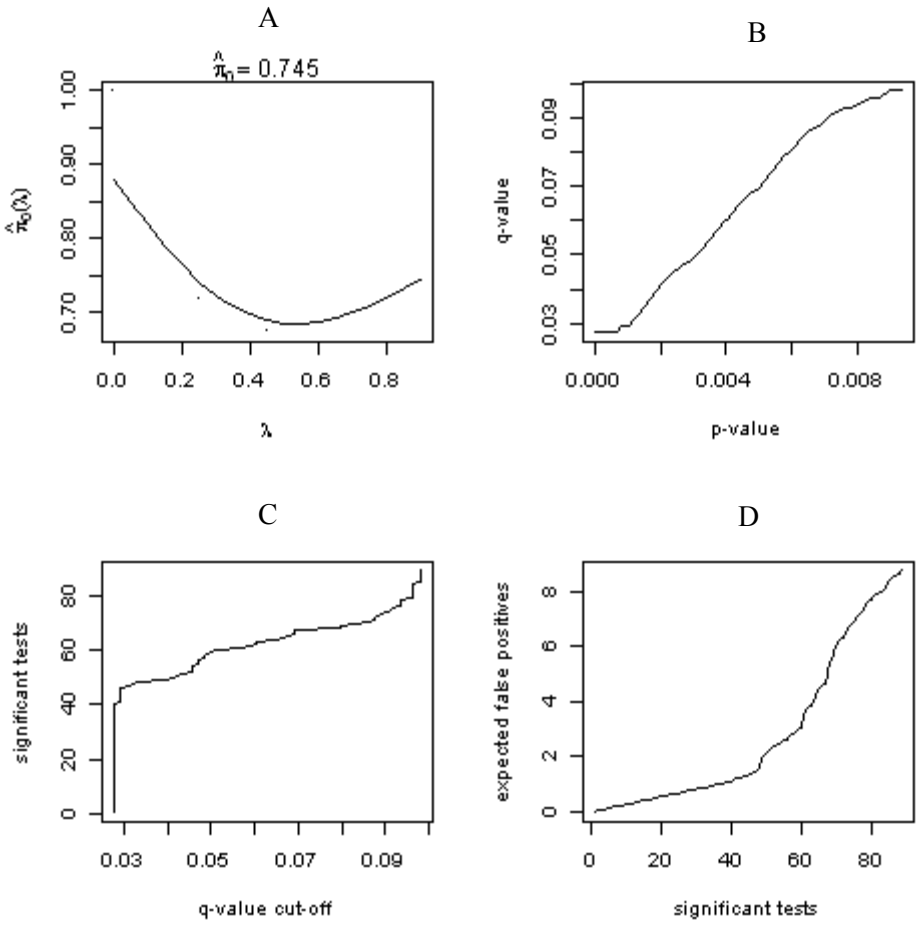


Figure 3.6 Q-plots for fork, including the estimated π versus the tuning parameter λ (A); the q-values versus the p-values (B); The number of significant tests versus each q-value cut-off (C); the number of expected false positives versus the number of significant tests (D). π_0 is the overall proportion of true null hypotheses among all tests. The p-value of a test measures the minimum false positive rate that is incurred when calling the test significant. The q-value of a test measures the minimum false discovery rate that is incurred when calling the test significant.

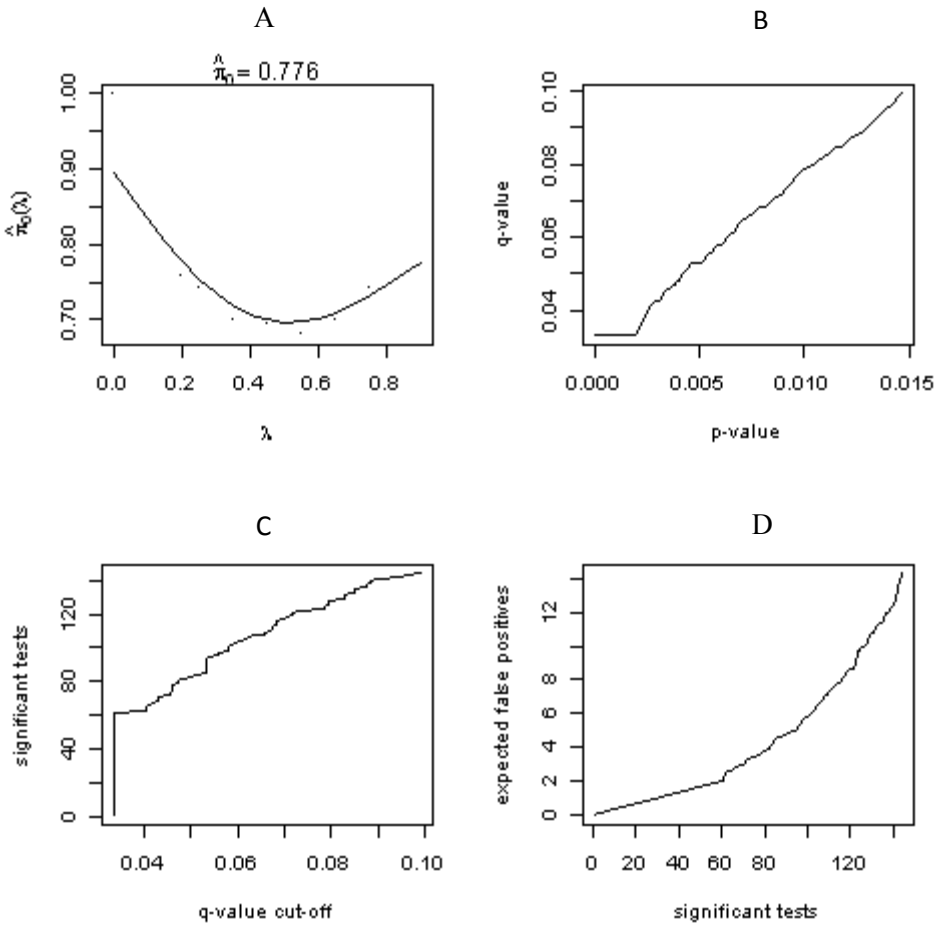


Figure 3.7 Q-plots for ramicorn branching, including the estimated π versus the tuning parameter λ (A); the q-values versus the p-values (B); The number of significant tests versus each q-value cut-off (C); the number of expected false positives versus the number of significant tests (D).

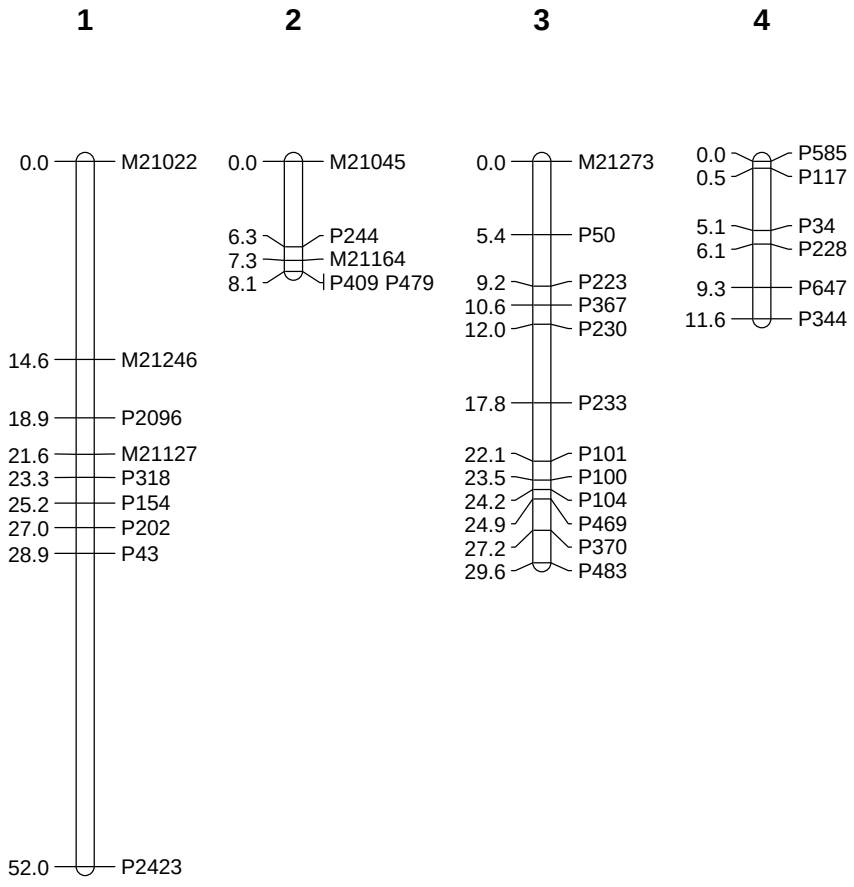
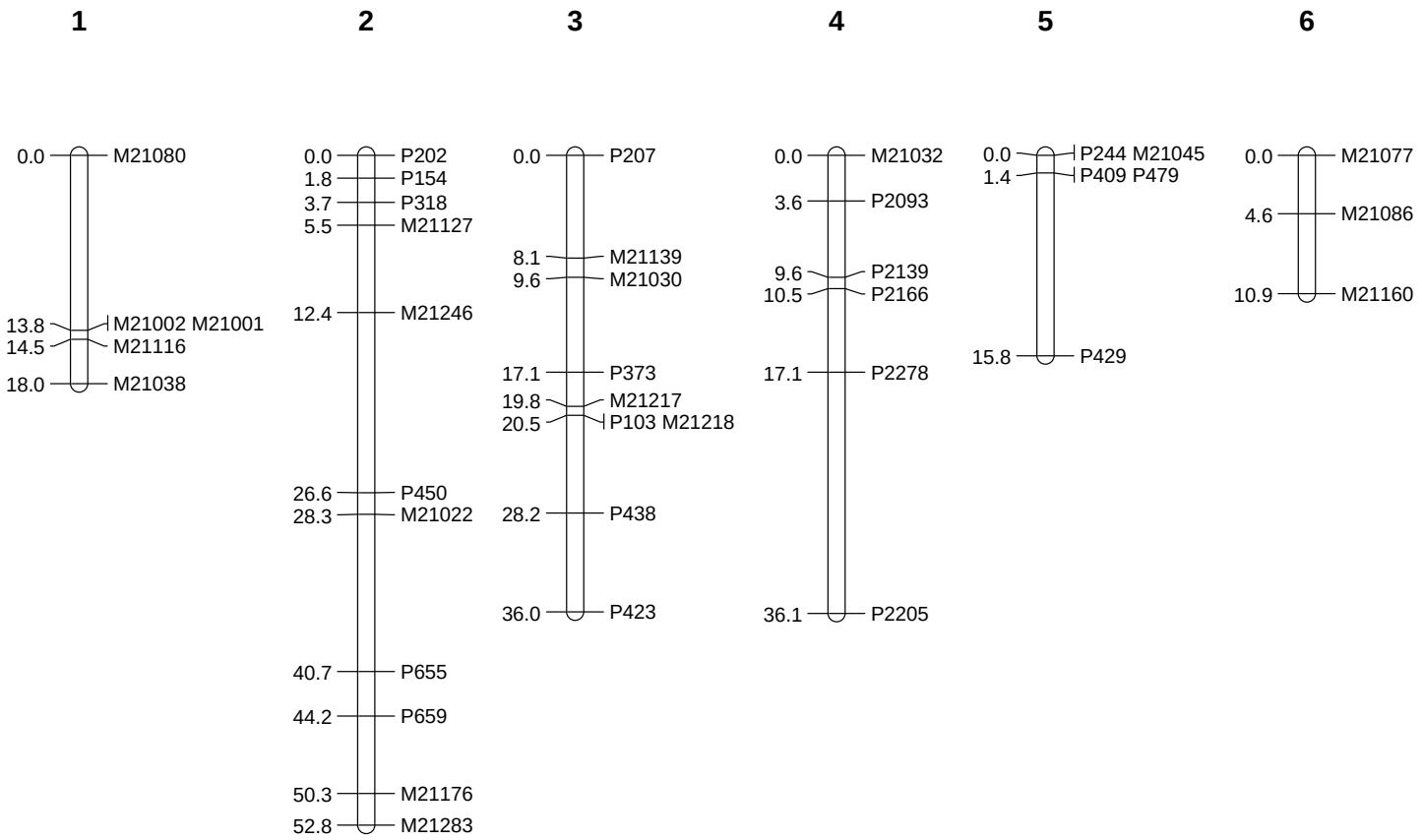
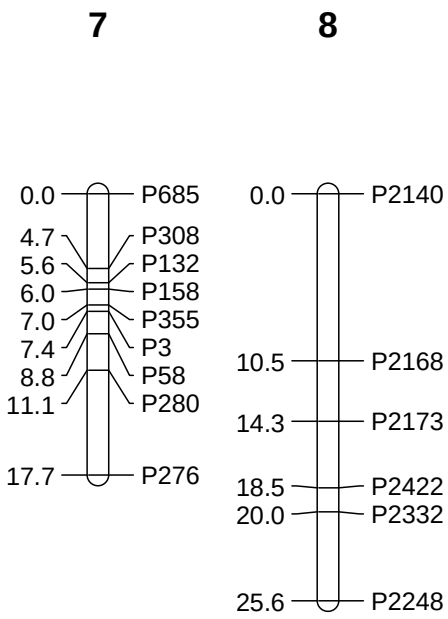


Figure 3.8 Linkage groups constructed by significant markers associated with forking in single marker analysis ($q<0.05$).

Figure 3.9 Linkage groups constructed by significant markers associated with ramicorn branching in single marker analysis ($q<0.05$).


1
0.0 M21080
13.8 M21002 M21001
14.5 M21116
18.0 M21038
2
0.0 P202
1.8 P154
3.7 P318
5.5 M21127
12.4 M21246
26.6 P450
28.3 M21022
40.7 P655
44.2 P659
50.3 M21176
52.8 M21283
3
0.0 P207
8.1 M21139
9.6 M21030
17.1 P373
19.8 M21217
20.5 P103 M21218
28.2 P438
36.0 P423
4
0.0 M21032
3.6 P2093
9.6 P2139
10.5 P2166
17.1 P2278
36.1 P2205
5
0.0 P244 M21045
1.4 P409 P479
15.8 P429
6
0.0 M21077
4.6 M21086
10.9 M21160


7
8
0.0
P685
4.7
P308
5.6
P132
6.0
P158
7.0
P355
7.4
P3
8.8
P58
11.1
P280
17.7
P276
0.0
P2140
10.5
P2168
14.3
P2173
18.5
P2422
20.0
P2332
25.6
P2248

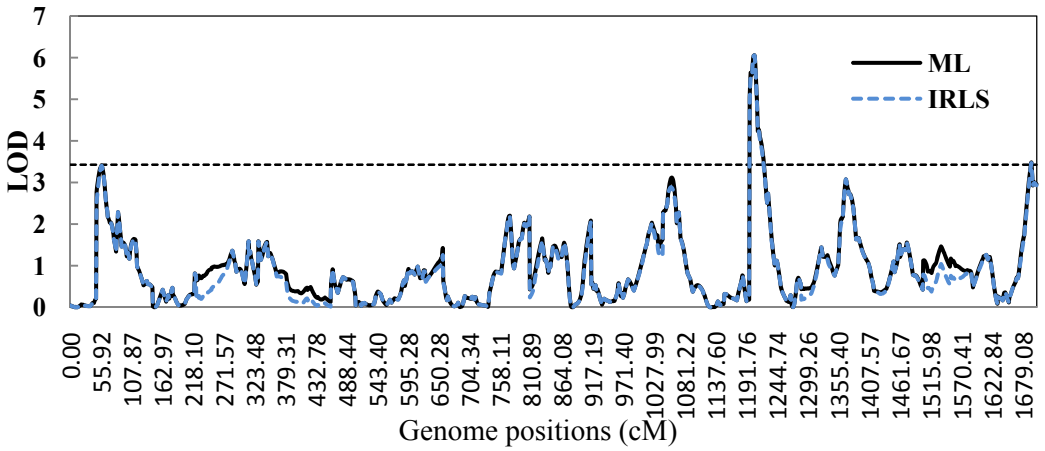


Figure 3.10 Log of Odds (LOD) test statistic profiles from two methods (ML and IRLS) for forking. The 18 groups are merged into a single genome. The dash line along the top of the graph presents the threshold value of LOD 3.43.

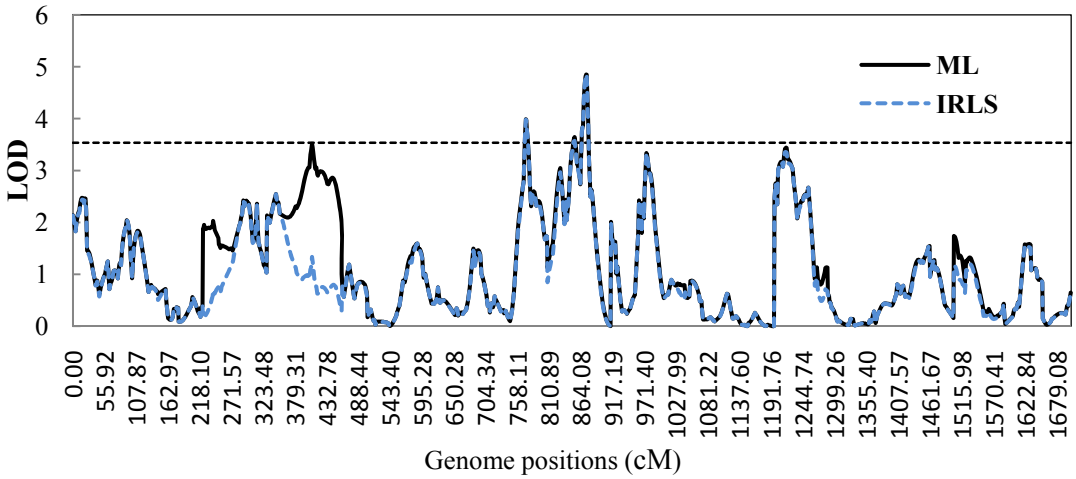


Figure 3.11 Log of Odds (LOD) test statistic profiles from two methods (ML and IRLS) for ramicorn branching. The 18 groups are merged into a single genome. The dash line along the top of the graph presents the threshold value of LOD 3.53.

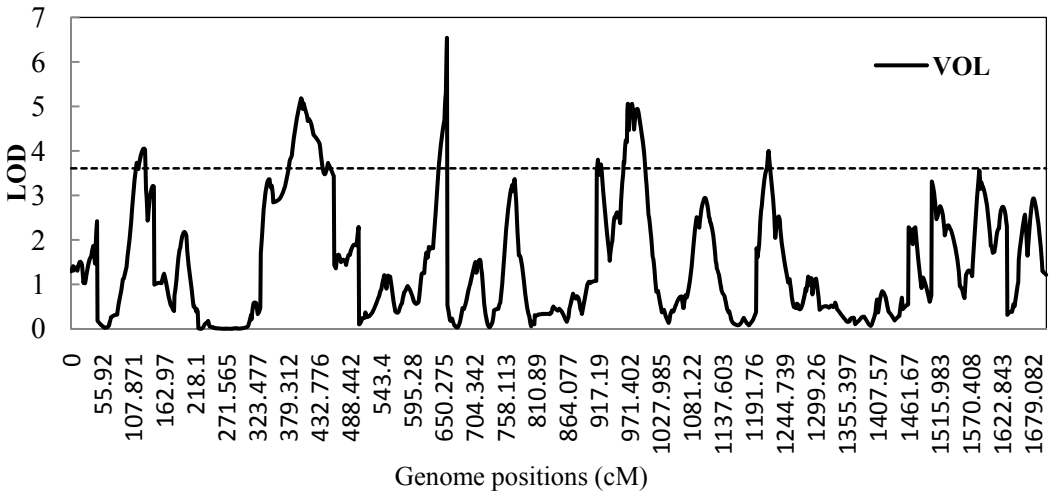


Figure 3.12 Log of Odds (LOD) test statistic profiles for volume. The 18 groups are merged into a single genome. The dash line along the top of the graph presents the threshold value of LOD 3.61.

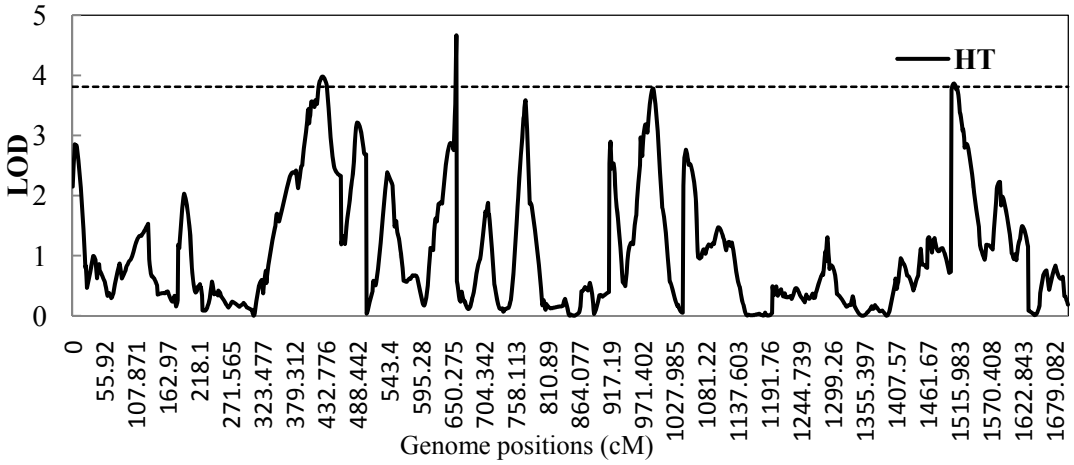


Figure 3.13 Log of Odds (LOD) test statistic profiles for height. The 18 groups are merged into a single genome. The dash line along the top of the graph presents the threshold value of LOD 3.82.

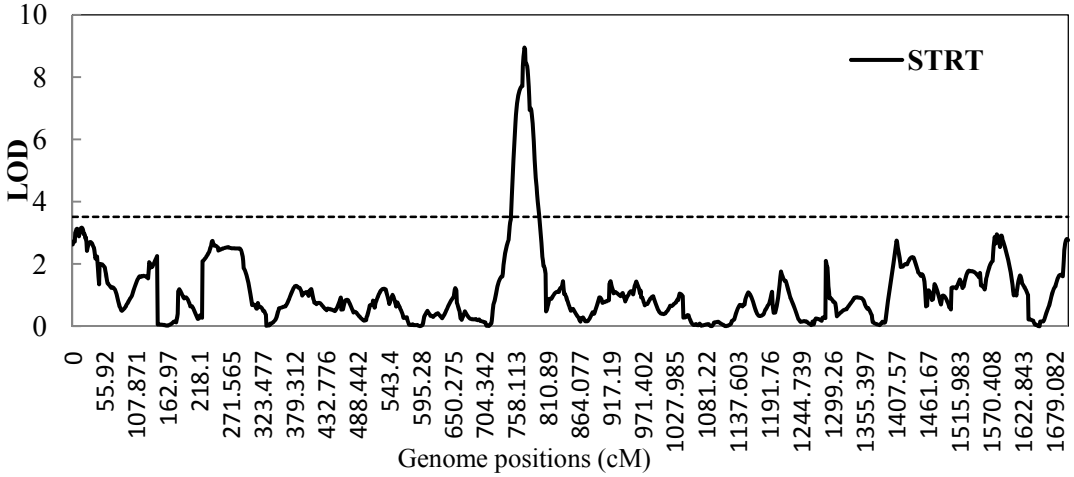


Figure 3.14 Log of Odds (LOD) test statistic profiles for straightness. The 18 groups are merged into a single genome. The dash line along the top of the graph presents the threshold value of LOD 3.52.

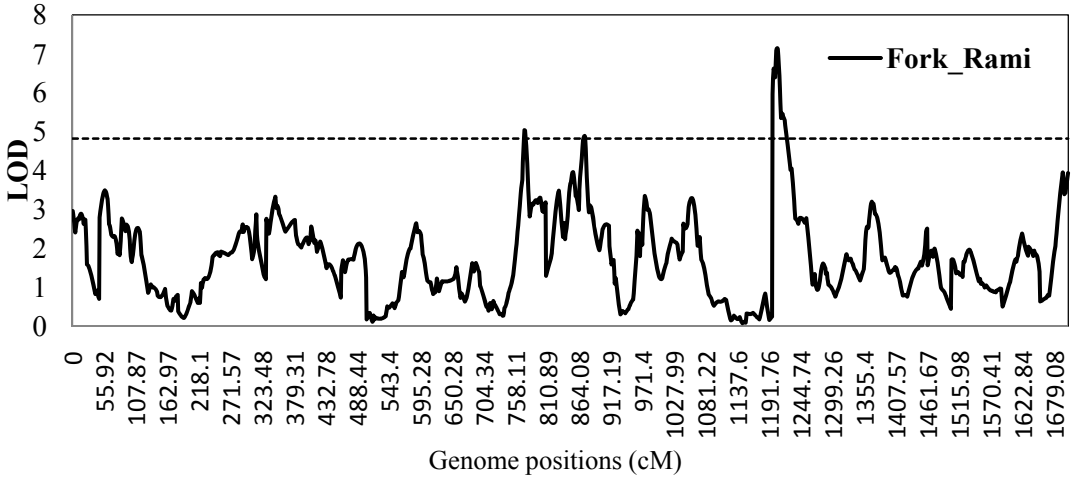


Figure 3.15 Log of Odds (LOD) test statistic profiles for multiple traits analysis (forking and ramicorn branching). The 18 groups are merged into a single genome. The dash line along the top of the graph presents the threshold value of LOD 4.82.

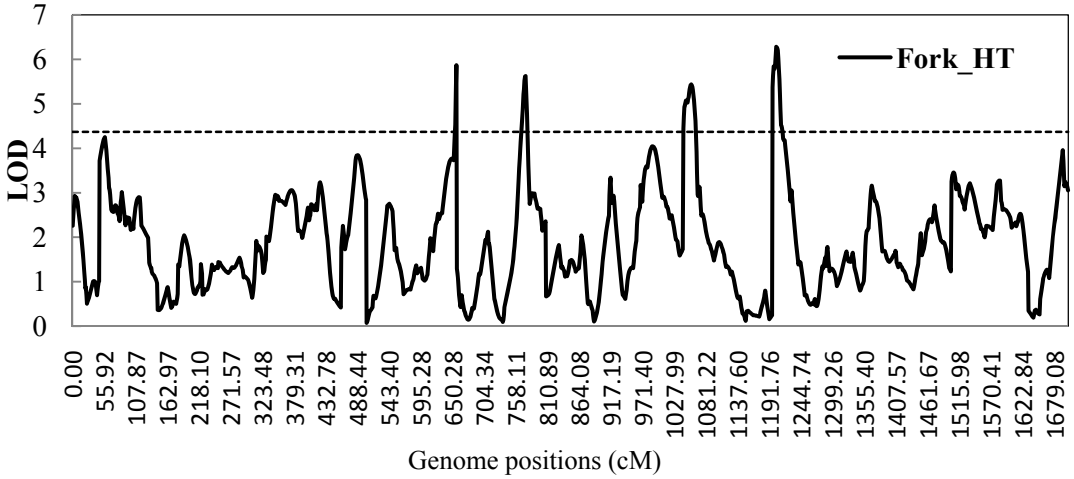


Figure 3.16 Log of Odds (LOD) test statistic profiles for two traits analysis (forking and HT). The 18 groups are merged into a single genome. The dash line along the top of the graph presents the threshold value of LOD 4.37.

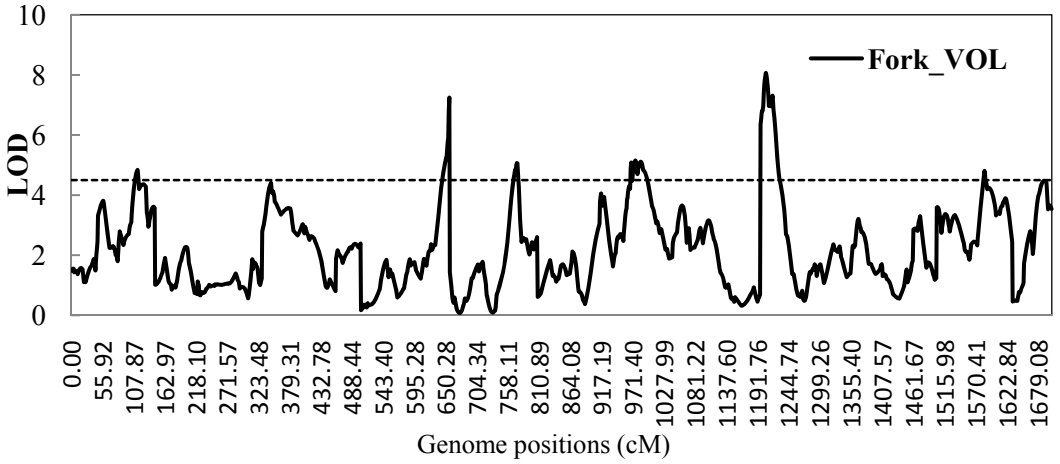


Figure 3.17 Log of Odds (LOD) test statistic profiles for two traits analysis (forking and volume). The 18 groups are merged into a single genome. The dash line along the top of the graph presents the threshold value of LOD 4.5.

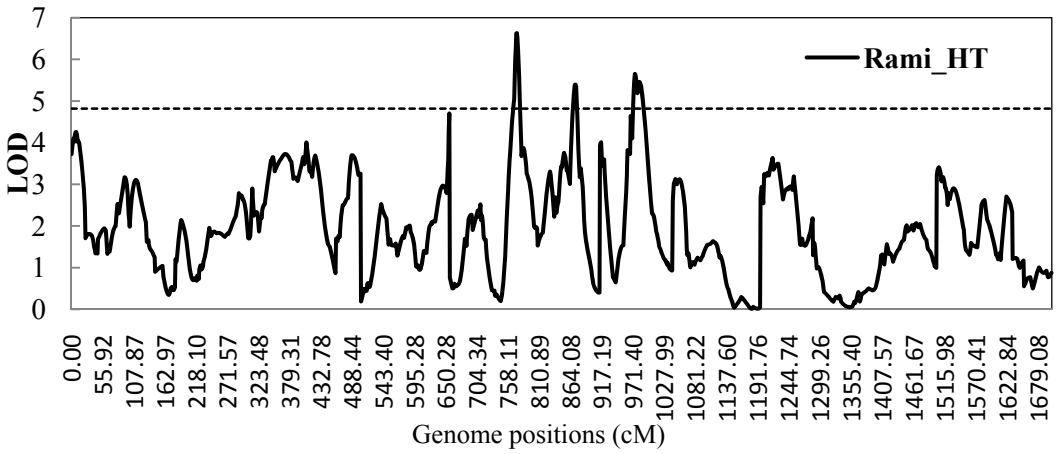


Figure 3.18 Log of Odds (LOD) test statistic profiles for two traits analysis (ramicorn branching and HT). The 18 groups are merged into a single genome. The dash line along the top of the graph presents the threshold value of LOD 4.82.

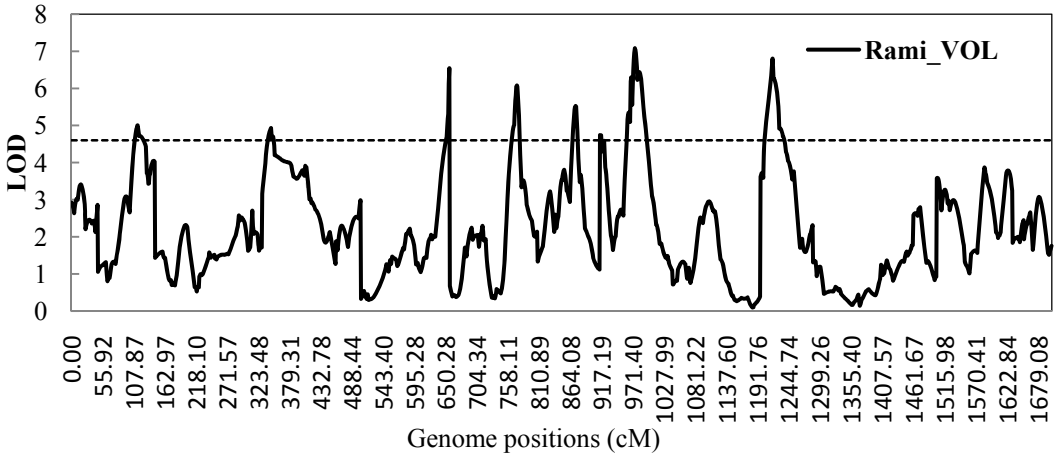


Figure 3.19 Log of Odds (LOD) test statistic profiles for two traits analysis (ramicorn branching and volume). The 18 groups are merged into a single genome. The dash line along the top of the graph presents the threshold value of LOD 4.6.