Trait Variation and QTL Mapping in Early-Season Maize Populations

by
Raja Ram Khanal

A Thesis
Presented to
The University of Guelph

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in
Plant Agriculture

Guelph, Ontario, Canada

© Raja Khanal, October, 2011
ABSTRACT

Trait Variation and QTL Mapping in Early-Season Maize Populations

Raja Khanal
University of Guelph, 2011

Advisor: Dr. Lewis Lukens

Maize (Zea mays L.) inbred lines for hybrid breeding are usually developed within distinct heterotic groups. Breeders impose strong selection and maintain relatively small population sizes that are adapted to local environments, where the aim is to identify the desired recombinant types in the progeny. However, linkages between loci that control a trait may not permit breeders to obtain the desirable genetic recombination in these populations. It is hypothesized that different favorable and unfavorable alleles accumulate within the lines from different heterotic groups. In addition, within each inbred line, favourable alleles are linked with unfavourable alleles. Two early-season maize inbreds, CG60 (Iodent) and CG102 (Stiff Stalk), were used to develop a selfed recombinant inbred line (SRIL) and an intermated recombinant inbred line (IRIL) populations. Furthermore, individuals from within these populations were testcrossed with an inbred tester from the Lancaster Sure Crop heterotic group, to give rise to selfed SRIL testcross (SRIL-TC) and IRIL testcross (IRIL-TC) populations. The inbred and inbred-testcross populations were evaluated for trait variation and QTL mapping. The genetic variance was high in inbred populations (SRIL and IRIL) with transgressive segregation for
flowering time and agronomic traits. However, genetic variances and correlation coefficients did not significantly differ between the inbred populations. Results suggested that pleiotropic genes were prevalent for these traits. In addition, linkages between the loci that control these traits were not common within parental genomes. Genetic linkage maps developed from the IRIL population were larger than those of the SRIL population. In the inbred-testcross populations (SRIL-TC and IRIL-TC) high means and high levels of trait variation were observed for all traits. The genetic variances and correlation coefficients of hybrid traits did not significantly differ between the SRIL-TC and IRIL-TC populations. Twenty five significant small to moderate QTL were detected, but only one, for grain moisture, was shared between inbred-testcross populations. Overall, the two inbred parents from different heterotic groups have many distinct alleles that contribute to traits. The recombinant inbred line populations had high means and variances for grain yield and related traits, which opens the possibility of utilizing these lines for hybrid breeding.
Acknowledgements

I would like to express my appreciation to Dr. Lewis Lukens, thesis supervisor, for his excellent guidance during this research study. I appreciate his input and advice during thesis writing. I am really indebted to his scientific as well as social and personal support on difficult situations during thesis writing.

My special thanks are also extended to Drs. Elizabeth Lee, Dave Wolyn and Gustavo Garcia, members of thesis advisory committee, for their criticisms and suggestions. Appreciation also extended to Dr. Alireza Navabi for his scientific contribution during my thesis writing. I also enjoyed the academic discussions related to my thesis with Drs. Duane Falk, Hugh Earl and Istvan Rajcan.

I would also like to acknowledge and express my gratitude for financial support provided by Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) and the Natural Science and Engineering Research Council (NSERC), Canada. I also extend my thanks for the Mrs. Fred Ball scholarship and the Pride Seed Scholarship.

Thanks are due also to the members of Lukens Lab, I enjoyed working with you. Byron Good, I highly appreciate your assistance in field experiments. Thanks are also extended to Heather Peterson for her assistance in field and lab experiments. I am fortunate to be a part of plant agriculture family and the support I received during my critical time.

Finally I wish to thank my wife, Sarita, and daughter, Kinjal, for their love, moral support and understanding which proved invaluable in completing the present study.
# Table of Contents

Acknowledgements ........................................................................................................ iv

Table of Contents ........................................................................................................... v

List of Figures ................................................................................................................ vii
List of Tables .................................................................................................................. ix

Chapter 1 Introduction .................................................................................................... 1

1.1 Maize as a hybrid crop  .......................................................................................... 1
1.2 Trait variation in intermated crops ......................................................................... 3
1.3 Intermiting and linkage mapping .......................................................................... 4
1.4 Intermiting and QTL mapping ............................................................................... 6
Objectives of the study................................................................................................ 7
References ................................................................................................................10

Chapter 2 Genetic Basis of Flowering Time and Related Traits in Early-Season Maize
Lines ..........................................................................................................................13

Abstract .....................................................................................................................13
2.1 Introduction ..........................................................................................................14
2.2 Materials and methods .........................................................................................18
   2.2.1 Genetic materials ...........................................................................................18
   2.2.2 Experimental design ......................................................................................19
   2.2.3 Data collection ...............................................................................................20
   2.2.4 Statistical analyses ........................................................................................ 21
2.3 Results .................................................................................................................25
   2.3.1 Trait distributions in RIL populations and parental lines ................................. 25
   2.3.2 Genetic variance estimates in the IRIL and SRIL populations ........................26
   2.3.3 Phenotypic and genetic correlations among RILs ..........................................27
2.4 Discussion ...........................................................................................................28
   2.4.1 Genetic variation for flowering time and related traits.....................................28
   2.4.2 Effects of intermiting on means and variances..............................................30
   2.4.3 Linkage and pleiotropy ...................................................................................33
Acknowledgements ....................................................................................................36
References ................................................................................................................37

Chapter 3 Linkage Map and QTL Analysis Using Intermated vs. Selfed Recombinant
Inbred Line Populations .............................................................................................49

Abstract .....................................................................................................................49
3.1 Introduction ..........................................................................................................50
3.2 Materials and methods .........................................................................................53
   3.2.1 Empirical study ..............................................................................................53
   3.2.2 Computer simulation study ............................................................................58
3.3 Results .................................................................................................................62
   3.3.1 Empirical data analysis ..................................................................................62
   3.3.2 Simulation study ............................................................................................64
Chapter 4 Early Adapted Maize Genotypes Harbor Great Genetic Diversity with Little Evidence of Repulsion Phase Linkages ...........................................88

Abstract ..................................................................................................................88
4.1 Introduction .......................................................................................................89
4.2 Materials and methods ...................................................................................92
   4.2.1 Population development ...........................................................................92
   4.2.2 Experimental design and field evaluations ..............................................93
   4.2.3 Phenotypic data analysis .......................................................................94
   4.2.4 Quantitative trait locus mapping ............................................................96
4.3 Results ............................................................................................................99
   4.3.1 Hybrids of RILs generated from two early season genotypes from two heterotic groups vary greatly .................................................................99
   4.3.2 IRIL and SRIL hybrid trait variances are similar .....................................101
   4.3.3 QTL analyses revealed small effect QTL that were largely unique to one population .........................................................................................102
4.4 Discussion .....................................................................................................105
   4.4.1 RIL populations exhibited great genetic variation for all traits ..........105
   4.4.2 QTL within the RIL testcross populations ............................................108
Acknowledgements .................................................................................................111
References .............................................................................................................112

Chapter 5 Conclusions and Future Directions .........................................................138

   5.1.1 Summary and main conclusions ...............................................................138
   5.1.2 Limitations of the study ........................................................................140
   5.1.3 Future research and research contributions .........................................141
References .............................................................................................................143
Appendices ............................................................................................................144
List of Figures

Figure 2.1 Schematic of the IRIL and SRIL population development from two maize inbred lines. Details of population development are given in the text.

Supplemental Figure 2.1 Frequency distributions of days to anthesis, days to silking, leaf number, plant height, stay green, normalized difference red edge (NDRE) index, normalized difference vegetative index (NDVI) and NDRE/NDVI traits of intermated recombinant inbred line (IRIL) and selfed recombinant inbred line (SRIL) populations of maize grown in two environments in Ontario, Canada during 2007. The mean of parents (CG60 and CG102) are shown by arrows.

Supplemental Figure 2.2 Biplot of Principal Components Analysis (PCA) axis 2 versus axis 1 for eight traits of intermated recombinant inbred line (IRIL) and selfed recombinant inbred line (SRIL) populations of maize grown in two environments in Ontario, Canada during 2007.

Figure 3.1 Comparative linkage maps and QTL mapping of selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations. The QTL for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN, number), stay green (SG, scores), plant height (PH, cm), normalized difference in red edge (NDRE), normalized difference vegetative index (NDVI) and NDRE/NDVI (ratio) are given on the right side of each linkage group with a chart. The chart shows chromosome length (cM) in x-axis and LOD values y-axis. The dash-dot vertical line along the graph represents the threshold value for each population.

Figure 3.2 Contours of LOD scores, averaged over 1000 simulation runs for genetic models with a single QTL accounting for 5 to 30% of the phenotypic variance, placed at the centre of a 150 cM linkage group in selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations (n=140). The horizontal dash–dot line represents LOD 3 as an arbitrary threshold.

Figure 3.3 Contours of LOD score, averaged over 1000 simulation runs for a 150 cM linkage group with two QTL, with equal effect, linked 10, 20, or 30 cM apart in coupling (A and B) and in repulsion (C and D), on a 150 cM linkage group in selfed recombinant inbred line ((SRIL) and intermated recombinant inbred line (IRIL) populations (n= 140). The dash–dot horizontal line represents the arbitrary threshold value of LOD 3.

Figure 3.4 Contours of LOD score, averaged over 1000 simulation runs for a 150 cM linkage group with two QTL, with equal effect, linked 10, 20, or 30 cM apart in coupling (A and B) and in repulsion (C and D), on a 150 cM linkage group in selfed recombinant inbred line ((SRIL) and intermated recombinant inbred line (IRIL) populations (n= 420). The dash–dot horizontal line represents the arbitrary threshold value of LOD 3.

Figure 4.1 Frequency distributions of days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel
row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha\(^{-1}\)) traits of intermated recombinant inbred line (IRIL) and Selfed recombinant inbred line (SRIL) testcross populations of maize grown in six environments in Ontario, Canada during 2007 and 2008. The parental hybrids (CG60 and CG102) are shown by black arrows and testcross population means are shown by green arrows.

Figure 4.2 Maize genetic linkage maps of selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) with QTL for different traits. QTL for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha\(^{-1}\)) shown by a bar next to the chromosome. Graph shows the chromosome length (cM) on the vertical axis and LOD value on the horizontal axis. The dash–dot vertical line along the graph represents the threshold value for each population.

Supplemental Figure 4.1 Maize genetic linkage maps constructed in this study. Distances between markers are shown on the left of each linkage group. Putative QTL for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha\(^{-1}\)) shown by a bar next to the marker. Graph shows the chromosome length (cM) on the vertical axis and LOD value on the horizontal axis. The dash–dot vertical line along the graph represents the threshold value for each population.
List of Tables

Table 2.1 Mean squares from the analyses of variances for days to anthesis (DA), Days to silking (DS), leaf number (LN), plant height (PH), stay green (SG), normalized difference vegetative index (NDVI), normalized difference red edge (NDRE) index and NDRE/NDVI traits across two environments in Ontario, Canada during 2007.

Table 2.2 Mean trait value for two parents (CG60 and CG102) and means and ranges of the intermated recombinant inbred line (IRIL) and selfed recombinant inbred line (SRIL) populations for days to anthesis (DA), Days to silking (DS), leaf number (LN), plant height (PH), stay green (SG), normalized difference vegetative index (NDVI), normalized difference red edge (NDRE) index and NDRE/NDVI traits across two environments in Ontario, Canada during 2007.

Table 2.3 Variance components and broad-sense heritability (H²) estimates for days to anthesis (DA), days to silking (DS), leaf number (LN), plant height (PH), stay green (SG), normalized difference vegetative index (NDVI), normalized difference red edge (NDRE) index and NDRE/NDVI traits measured on intermated recombinant inbred line (IRIL) and selfed recombinant inbred line (SRIL) populations in two environments in Ontario, Canada during 2007.

Table 2.4 Genetic (above diagonal) and phenotypic correlations (below diagonal) for days to anthesis, dDays to silking, leaf number, plant height, stay green, normalized difference vegetative index (NDVI), normalized difference red edge (NDRE) index and NDRE/NDVI traits measured on intermated recombinant inbred line (IRIL) and selfed recombinant inbred line (SRIL) populations of maize lines grown in two environments in Ontario, Canada during 2007.

Table 3.1 Components of variance and QTL effect sizes for genetic models used in computer simulation of a quantitative trait locus (QTL) in selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations.

Table 3.2 Components of variance for genetic models used in computer simulation of a quantitative trait controlled by two linked QTL at 10, 20, and 30 cM in coupling and repulsion phase on a 150 cM linkage group in the selfed recombinant inbred line (SRIL) and the intermated recombinant inbred line (IRIL) populations.

Table 3.3 Number of marker loci and linkage map distance (cM) in selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations of corn, the ratio of map distance estimates in SRIL and IRIL populations.

Table 3.4 QTL detected on different linkage groups (LG) with the multiple QTL mapping approach for days to anthesis (DA), Days to silking (DS), leaf number (LN), plant height (PH), stay green (SG), normalized difference vegetative index (NDVI), normalized difference red edge (NDRE) and NDRE/NDVI (Ratio) in selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations.
Table 3.5 QTL support intervals, averaged over 1000 simulation runs (mean) and respective lower and upper confidence limits, for a 150 cM linkage group with one and two QTL for six genetic models in selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations (n= 140). The two QTL were linked in 10, 20, or 30 cM apart in coupling and in repulsion (italic) on a 150 cM linkage group.

Table 4.1 Mean trait values for two parental inbred lines (CG60 and CG102) and means, ranges and heritabilities (H2) across six environments for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, % ), and grain yield (GY, Mg ha-1) of the testcrosses involving the selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations. Standard errors (s.e.) of the means are in parentheses.

Table 4.2 Variance component estimates for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, units), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha-1) for selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL)testcross populations.

Table 4.3 Genetic correlation for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha-1) traits measured on selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) testcross populations of maize lines grown in six environments in Ontario, Canada. Values in italics bold are significant.$P=0.05$.

Table 4.4 QTL detected for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha-1) in selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) testcross populations combined across six environments.

Supplemental Table 4.1 Epistatic QTL for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha-1) for selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) testcross populations.
List of Abbreviations

ANOVA, analysis of variance;
BLUP, best linear unbiased predictor;
BSSS, stiff stalk synthetic;
CBD, Corn Belt Dent;
CIMMYT, international maize and wheat improvement center;
DA, days to anthesis;
DS, days to silking;
EL, ear length;
F, inbreeding coefficient;
GM, grain moisture;
GY, grain yield;
IBM, intermated B73 × Mo17;
IRIL, intermated recombinant inbred line;
KRN, kernel row number;
KW, 1000 kernel weight;
LN, leaf number;
LOD, logarithms of odds
LRT, likelihood ratio test;
LSC, Lancaster Sure Crop;
MQM, multiple QTL model;
NDRE, normalized difference red edge index;
NDVI, normalized difference vegetation index;
OPV, open pollinated varieties;
PCR, polymerase chain reaction;
PH, plant height;
QTL, quantitative trait locus;
RCBD; randomized complete block design;
REML, multivariate restricted maximum likelihood;
RIL, recombinant inbred lines;
SG, stay green;
SNP, single nucleotide polymorphism;
SRIL, selfed recombinant inbred lines;
SSR, simple sequence repeat;
Maize (Zea mays L.) is one of the most important food, feed and industrial crops. It originated from teosinte in the Western Hemisphere around 9000 years ago, and was domesticated 700 to 1000 years ago in south-central or south-western Mexico (Goodman and Brown, 1988). Its genetic diversity allows it to be cultivated from tropic to temperate regions of the world. The cultivation of maize in different environments has been possible due to adaptability for flowering time, the complex trait which to a large degree determines the reproductive success of crop plants. Maize inbred lines are developed for specific environments, and the lines adapted to Ontario, Canada are among the early flowering lines (Labate et al., 2003).

1.1 Maize as a hybrid crop

After domestication, maize diverged into over 250 landraces adapted to different environments stretching from Chile in the south to Canada in the north (Camus-Kulandaivelu et al., 2006). Among these landraces, the predominant source of commercial germplasm is Corn Belt Dent (CBD) (Goodman, 1985). The CBD landrace arose from a cross between the subtropical Southern Dent landrace and a Northern Flint landrace about 200 years ago (Doebley et al., 1988). This commercial germplasm pool is further divided into different heterotic groups (Lee and Tollenaar, 2007) and maize breeders have developed inbred lines suitable for hybrid breeding from these heterotic groups.
Hybrid breeding requires inbred lines developed from heterotic groups based on both the inbred line’s performance and its hybrid performance. A heterotic group comprises a set of genotypes that perform well when crossed with genotypes from an opposite heterotic group (Carena and Hallauer, 2009). Inbred lines for hybrid maize breeding are usually developed by crossing inbred lines from the same heterotic group. This approach has been successful; however, this breeding method has raised concerns regarding the narrowing of maize genetic diversity within heterotic groups (Troyer, 1999).

The main factors for increasing grain yield in maize are genetic improvement and enhanced agronomic practices (Duvick, 2005). Direct selection of the best individuals can result in the accumulation of repulsion phase linkages, in which linked favorable and unfavorable alleles are fixed within a population (Hill and Robertson, 1966). The trait may be altered by intermating between individuals, thereby breaking up the linked genes influencing it.

Intermating of individuals in a population prior to selection in a breeding program may have two possible outcomes. One view suggests that intermating among individuals in a population will break up the favorable coupling phase linkages. Another suggests that intermating increases the relative frequency of desirable recombinants if the linkages between the genes are predominantly in the repulsion phase in the population.
1.2 Trait variation in intermated crops

A key approach to understanding trait variation in an intermated maize population is to compare populations having identical genotypic frequencies at each individual locus but differing with respect to the joint frequencies of two and more loci (Melchinger et al., 2003). Covarrubias-Prieto et al. (1989) found small changes in the means for grain yield and grain moisture before and after six generations of intermating of maize F2 populations. The differences were attributed to epistatic effects or selection during the six generations of intermating. A significant reduction in maize grain yield in an intermated population was found when comparing an F2 population and eight generations of an intermated F2 population (Lamkey et al., 1995), and these differences were attributed to recombination and break-up of favourable allele combinations. Similar changes in population means of intermated as compared to non-intermated populations were attributed by others to epistatic effects (Altman and Busch, 1984; Humphrey et al., 1969) or to selection (Meredith and Bridge, 1971).

Intermating is expected to reduce genetic variance if the alleles are in coupling phase linkages in the F1 (Comstock and Robinson, 1948). For example, significant reduction in genetic variance for lint weight in cotton after intermating over that of a non-intermated population was reported (Miller and Rawlings, 1967). The effect of intermating in the (B73×B84) F2 population of maize were also evaluated by Han and Hallauer (1989) who found minor change in the genetic variance of the population after intermating. Lamkey et al. (1995) found
no change in genetic variance after intermating F$_2$ populations for eight generations. Conversely, it has been reported that genetic variance of grain yield reduced significantly in an intermated population as compared to a F$_2$ derived population of maize (Huang et al., 2010). These studies show overall that the types of results are varied and inconclusive based on these examples.

Correlation between traits has been the tool most frequently used to identify the changes in relationships among quantitative traits. Intermating results in breakup of linked alleles (Hanson, 1959). When Miller and Rowlings (1967) compared six generations of intermated F$_2$ population and non-intermated F$_2$ population in cotton developed from two contrasting inbred lines, they found that initial linkage blocks were disrupted with intermating. A negative genetic correlation between lint yield and fiber strength was lower in the intermated population when compared to the non-intermated population, suggesting that intermating in cotton has been effective to disrupt the linkage between the traits (Meredith and Bridge, 1971). An intermated maize population was developed by Dudley et al., (2004) and compared with an F$_2$ population for starch, protein, and oil traits. They reported that genetic and phenotypic correlations in the intermated population were lower than in the F$_2$ population.

1.3 Intermating and linkage mapping

Genetic linkage analysis provides maps that form important tools for applied plant breeding programs. A linkage map is based on the frequencies of recombination between markers during crossover of homologous chromosomes.
The greater the frequency of recombination between two markers, the farther apart the markers are in the chromosome, and the lower the frequency of recombination between the markers the smaller the distance between them on the chromosome.

Population development is the first step in a linkage mapping experiment. The choice of mapping population is an important contributor to success of any given QTL mapping experiment (Darvasi and Soller, 1995; Lee and Donald, 1995). Traditionally, segregating populations such as: F₂, doubled haploid, recombinant inbred lines (RILs) have been used for linkage mapping. The problem with these populations is that they have gone through only a single generation of meiosis and have limited recombination. If the genes are linked and close to each other in a chromosome, it is hard to separate them. Linkage is broken down more effectively by intermating an F₂ population for few generations, and more precise map resolution can be obtained in such populations (Darvasi and Soller, 1995).

Intermating to construct recombinant inbred line population, called IRIL population (Winkler et al., 2003) has been used extensively in recent studies. IRIL populations are randomly intermated for a number of generations after the F₂, followed by cycles of self fertilization to produce a population of homozygous lines. Using this method produces a permanent population of fixed lines similar to a mapping population developed by selfing from F₂, but the mapping resolution is higher in IRIL population compared to F₂ population (Balasubramanian et al., 2009). Lee et al., (2002) reported four-fold increased recombination frequencies by using intermated B73 × Mo17 (IBM) recombinant inbred population as
compared to an $F_2$ population in maize. The IBM population was later used for precise QTL mapping (Balint-Kurti et al., 2008; Rodriguez et al., 2008).

1.4 Intermating and QTL mapping

In QTL mapping, a segregating population derived from a bi-parental cross is common practice. The basic steps in QTL mapping are: 1) mapping population development, 2) genotyping of the individual population with molecular markers, 3) phenotyping the segregating population for the trait of interest, and 4) making statistical association between molecular and phenotypic data. Because the precision of the QTL mapping will depend on type of mapping population used, an intermated recombinant inbred line population is beneficial for increasing the precision of QTL mapping (Balasubramanian et al., 2009; Balint-Kurti et al., 2010; Darvasi and Soller, 1995). These mapping populations have been used successfully in plants to localize QTL to relatively narrow chromosomal segments (Balint-Kurti et al., 2007; Hazen et al., 2003). In maize, the IBM population has been used to map the QTL for disease resistance and flowering time traits (Balint-Kurti et al., 2008; Balint-Kurti et al., 2007). These authors have reported that more precise QTL mapping for disease resistance can be achieved with the intermated population than with the RIL population. They further reported that the use of the IRIL population provided between 5 and, in one case, 50 times greater mapping resolution of QTL than the RIL population.

In a commercial hybrid breeding program, inbred lines are usually developed within a distinct heterotic group. Maize breeders impose strong selection and
maintain relatively small population sizes that are adapted to particular environments. The aim for the selection is to identify the best recombinant types in the resulting progeny. A number of fundamental questions arise, however, including whether or not inbred lines from different heterotic groups have fixed different sets of genes/alleles that control the same trait; if selection in a small population results in linkage between favourable and unfavourable genes, as predicted in Hill and Robertson (1966) theory; if intermating breaks the linkage between favourable and unfavourable genes; and finally, if it is possible to use maize inter-heterotic groups for inbred line development. Answering these quantitative genetics questions would increase our knowledge of inbred line development for hybrid breeding in maize.

Although studies on effect of intermating on trait variation in maize have been done previously, most of the studies focused on late season germplasm. Previous studies have focused on F₂ generation or backcross generation for this purpose. With the advancement of molecular markers and statistical methods of analysing molecular marker and phenotypic data during the last decade, a better understanding of the genetic basis of trait variation in a population seems possible. As Hill and Robertson’s theory (1966) predicts, selection may result in linkage between favourable and unfavourable alleles.

**Objectives of the study**

The main objectives of this study were to detect genetic differences between two early-season inbreds and to determine the effect of intermating on trait variation and QTL analysis in maize populations derived from these inbreds from
different heterotic groups. Previous studies have shown varied and conflicting results and I intend to generate conclusive results for effect of intermating on trait variation using early season adapted maize inbred lines.

**Hypothesis I**

Inbred lines for hybrid breeding are usually developed in an intensive selection process, which may result in accumulation of repulsion phase linkages that affect the trait. It is hypothesized that the inbred lines selected for an early-season environment have developed repulsion phase linkages, and intermating will break up the linked alleles and increase the trait variances and correlations.

The objectives were (i) to evaluate the genetic basis of flowering time and related traits in two early-season parental inbreds, (ii) to evaluate genotypic and phenotypic correlations between flowering time and related traits, and (iii) to evaluate whether or not disruption of putative coupling and repulsion phase linkage blocks causes changes in population genetic variances and genetic correlations.

**Hypothesis II**

Genetic resolution of linkage maps can be improved significantly by providing additional opportunities for recombination in a mapping population. It is hypothesized that three generations of intermating before selfing will increase recombination expand linkage maps, and thus provide more precise QTL estimation in the intermated recombinant inbred line (IRIL) population as compared to the selfed recombinant inbred line (SRIL) population.
The objectives of this study were (i) to study the changes in frequency of recombination between IRIL and SRIL populations, (ii) to compare QTL detection within the IRIL and the SRIL populations, and (iii) to discuss the empirical results relative to the expectations from simulations.

Hypothesis III

The success of breeding programs in maize depends mainly on the type of germplasm used for inbred line development. Inbred lines for hybrid breeding are usually developed from within a heterotic group, but breeders sometimes use inter-heterotic commercial hybrids as a source of germplasm for hybrid breeding. It is hypothesized that distinct favorable and unfavorable alleles accumulate within individuals from different heterotic groups. Within these individuals favorable alleles are linked with unfavorable alleles, a phenomenon termed the Hill-Robertson effect.

The objectives were (i) to determine the degree to which hybrid yield and flowering traits vary amongst hybrids between LH295 and RILs, (ii) to identify the loci polymorphic between CG60 and CG102 that contribute to this trait variation, and (iii) to determine if strong selection contributed to the development of genotypes with favorable alleles linked to unfavorable alleles, as predicted.
References


Hanson, W.D. 1959. The breakup of initial linkage blocks under selected mating systems. Genetics 44:857-868.


Chapter 2 Genetic Basis of Flowering Time and Related Traits in Early-Season Maize Lines

Abstract

Flowering time is the major factor in determining maize (Zea mays L.) maturities. Genetic bases of flowering time and other agronomically important traits were examined in a set of inter-heterotic pattern recombinant inbred lines (RILs). The RILs were developed from crossing the short season iodent inbred line CG60 with the short season Stiff Stalk inbred line CG102. RILs were derived through single-seed descent (SRILs) or intermated for three generations prior to inbreeding (IRILs), thereby increasing recombination. Genetic variation was high for all traits in both populations, and a large number of RILs had days to silking (DS), days to anthesis (DA), plant height (PH), visual stay green (SG), canopy reflectance, and leaf number (LN) trait values beyond the parental line ranges. Both the IRIL and SRIL populations had similar genetic variances and similar genetic and phenotypic correlations between traits. We conclude that although Northern Corn Belt Dents have been subjected to strong selection and drift and have diverged greatly from their source population, there is substantial cryptic variation for many traits. Our results also suggest that coupling and repulsion phase linkage blocks are not prevalent within parental genomes and that genetic correlations are caused by pleiotropic genes. Intermating may have little value in recovering extreme phenotypes for flowering time and other agronomically important traits.
2.1 Introduction

Flowering time and related traits are important agronomic traits in maize (Zea mays L.). Thus, flowering date differences are quite substantial in maize (e.g., Camus-Kulandaivelu et al., 2006), reflecting selection for genotypes adapted to different latitudes. Genotypes adapted to short season environments flower early because late season cold temperatures and frost damage negatively affect crop yield. Early flowering genotypes have shorter statures and fewer leaves than later flowering genotypes. These traits result in a low leaf area index, less dry matter accumulation and few kernels (c.f. Lee and Tollenaar, 2007). Maturity differences are also associated with visual stay-green (SG), the maintenance of green leaf area during the grain-filling period. The reflectance spectra from the crop provide a quantitative measure of SG. In combination reflectance spectra can be used to calculate the Canopy Chlorophyll Concentration Index (CCCi) which is strongly sensitive to leaf chlorophyll concentration, while being quite insensitive to differences in leaf area index (Barnes et al., 2000; Rodriguez et al., 2006). Stay green can increase the net carbon output of a leaf (Thomas and Howarth, 2000), indicate source-sink balance (Lee and Tollenaar, 2007; Rajcan and Tollenaar, 1999), and has been positively associated with historical improvements of maize hybrids (Duvick and Cassman, 1999).

Maize breeders typically self-pollinate F₂ plants or generate doubled-haploids to produce heterogenous populations of homozygous families (c.f. Lee and Tracy, 2009). Inbred line development via self-pollination or doubled haploid production quickly leads to homozygosity, leading to large, parental linkage
blocks in the inbred lines. Intermating methods such as full-sib mating maintain heterozygosity for long periods (Hallauer and Miranda, 1981). Slowing the rate of inbreeding provides an opportunity to form novel non-parental linkages. For example, in the intermated B73 x Mo17 (IBM) population, five generations of random-mating prior to inbreeding resulted in a nearly 4-fold increase in genetic map distance relative to the F2 population (Lee et al., 2002).

The phenotypic effect of increasing recombination during inbred development depends on the effects of linked genes within the parental lines. Recombination between linked, unfavourably epistatic genes and recombination between repulsion phase linkages increases phenotypic variation within a population of inbreds. Recombination between linked, favorably epistatic genes and recombination between coupling phase linkages decreases phenotypic variation. Lamkey et al. (1995) described a large but non-significant increase in genetic variance for grain yield, days to anthesis (DA), and days to silking (DS) in a random-mated, testcrossed population relative to the F2. In contrast, a different population generated by four generations of random-mating the F2 of an Illinois high protein × Illinois low protein maize cross had less variation for protein content than the F2. This significant decrease in genetic variance was ascribed to the break-up of coupling linkages (Dudley, 1994). Coupling linkages that affected a number of traits were also detected in a comparison of F2:3 populations and F6:7 populations from a cross between inbred lines Mo17 and H99 (Austin, 1995; Austin and Lee, 1996).
Despite the diversity of maize (e.g., Buckler et al., 2001; Gore et al., 2009), cultivated North American lines, including lines adapted to Ontario, Canada, are the products of only a small portion of the entire gene pool. First, of the 250 to 300 races of maize (Brown and Goodman, 1977), only one race, the Corn Belt Dent (CBD) is the predominant source of commercial germplasm (Goodman, 1985). CBD was generated by crosses between the subtropical Southern Dent landrace with a Northern Flint landrace about 200 yr ago (Camus-Kulandaivelu et al., 2006; Doebley et al., 1988). Second, the overwhelming majority of CBD inbred lines trace their pedigree back to only two open pollinated varieties, Reid Yellow Dent, and to a lesser extent, Lancaster Sure Crop (cf. Lee and Tracy, 2009). The proportion of Reid Yellow Dent derived germplasm in commercial North American germplasm is estimated to be 50% (Troyer, 1999). Third, the commercial germplasm pool is further stratified into heterotic groups (cf. Lee and Tollenaar, 2007). Finally, inbred lines have been developed for target environments, and CBD strains adapted to Ontario, Canada are among the earliest flowering cultivars (Labate et al., 2003).

In this study we examined the genetic means, co-variances, and variances of flowering time (DA and DS) and a number of flowering time associated traits, leaf number (LN), plant height (PH), SG, and canopy reflectance traits in a set of short-season, inter-heterotic pattern RILs representing two levels of intermating. Specifically our objectives were: (1) to evaluate the genetic basis of DA, DS, LN, PH, SG, and canopy reflectance traits within the short-season parental lines, (2) to evaluate genetic and phenotypic correlations between DA, DS, LN, PH, SG,
and canopy reflectance traits, and (3) to evaluate whether disruption of putative coupling and repulsion phase linkage blocks causes changes in population means, genetic variances, and genetic correlations. Maintaining or enhancing PH, LN, and SG while maintaining flowering time may enhance productivity in short season maize.
2.2 Materials and methods

2.2.1 Genetic materials

Two sets of inter-heterotic pattern recombinant inbred lines (RILs) were used in this study. The RILs were developed by crossing the Iodent inbred line CG60 (Lee et al., 2001a) with the Stiff Stalk inbred line CG102 (Lee et al., 2001b). One RIL subset was derived through single-seed descent (SRIL), while the other RIL subset was intermated for three generations prior to inbreeding (IRIL) (Figure 2.1). In total, 139 SRILs and 135 IRILs were developed and used for phenotypic evaluations.

The CG60 x CG102 F1 was self-pollinated to generate F2 seeds. Mating schemes to develop SRILs and IRILs each started with 140 F2 individuals. For the SRILs an ear-to-row method of single-seed descent was used. Briefly, each F2 individual was self-pollinated to generate F3 progeny; F3 seed from each ear was planted in individual rows and a single F3 individual was self-pollinated to generate F4 progeny. This was followed until the F6 generation was achieved. For the IRILs an ear-to-row version of chain sibbing was practiced. Briefly, each F2 individual was pollinated by an F2 individual, generally the adjacent plant. That same pollinated F2 individual was used as the pollen parent in a cross with a different F2 individual, generally the adjacent plant on the opposite side. Syn1 seed from each ear was planted in individual rows, and a plant from each row was used as the pollen and seed parent in crosses with the adjacent rows to form syn2 seed. Note that within a row one plant may have been used as the
seed parent and another may have been used as the pollen parent. This procedure was performed again to form syn3 seed. Following the third cycle of intermating, the syn3 seed was planted as ear-to-row, and inbreeding was practiced to achieve the $F_5$ generation using the ear-to-row method of single-seed descent described above (Figure 2.1). No intentional selection was imposed in developing either set of RILs. To avoid unintentional selection, particularly for flowering date, we performed pollinations when more than 50% of the plants in a row were silking and had pollen shedding. We estimated inbreeding coefficients ($F$) by determining the coancestry of plants within the pedigree. The coancestry of any two individuals is equal to their progeny’s inbreeding coefficient (Falconer and Mackay, 1996). The inbreeding coefficient ($F$) of individuals was estimated to be 0.95 and 0.97 for the IRIL and SRIL populations, respectively (Falconer and Mackay, 1996). The parental lines CG60 and CG102 were assumed to share no coancestry, thus inbreeding coefficients are conservative estimates.

**2.2.2 Experimental design**

The parental inbred lines, 139 SRILs and 135 IRILs were evaluated at two Ontario, Canada locations: Elora and Guelph. The field experiments were planted on May 9th 2007 at Elora and on May 15th 2007 at Guelph using a randomized complete block design (RCBD) with two replications. Because the blocks were large, the RCBD involved the randomization option for nearest neighbour analysis to control for within replication heterogeneity (Mulitze, 1990). Single-row plots with 20 plants per row with 4.0 m length and 76 cm between
rows were used. Weeds were controlled using pre-plant and post-emergence herbicides.

2.2.3 Data collection

Six traits were measured: DA, DS, LN, PH, SG, and canopy spectral reflectance. These traits were selected for phenotypic evaluation because they are associated with flowering time and hybrid yields. DA, DS, SG, and canopy reflectance were determined on a plot-basis. DA and DS were recorded when either 50% of the plants in the plot were shedding pollen or had visible silks. SG was recorded on a single date in each location after all genotypes had reached physiological maturity (i.e., black layer formation). SG was scored using a visual rating of 1-9 for an individual row, where 1 is complete senescence of plants in the row and 9 is a row of completely green plants. We also measured canopy spectral reflectance on all plots at physiological maturity using a dual channel reflectance spectrometer (Unispec DC, PP Systems, Amesbury MA). This instrument recorded the ratio of canopy reflectance to incoming solar radiation simultaneously at each 3-nm interval between 310 and 1100 nm. The downward-looking fibre optic was positioned approximately 70 cm above the tops of the plants, aimed downward at a 45-degree angle and parallel to the row direction. The downward- and upward-looking channels were cross calibrated just prior to field measurements by measuring reflectance of a barium sulphate standard using the same measurement geometry and orientation as used for the field measurements. Two reflectance scans were made per plot, and the mean was determined. From the reflectance data we calculated the Normalized
Difference Vegetation Index (NDVI), a widely-used indicator of green canopy cover, as

\[
\text{NDVI} = \frac{R_{790} - R_{670}}{R_{790} + R_{670}},
\]

where \( R_{790} \) and \( R_{670} \) are fractional canopy reflectance at 790 nm (near infrared) and 670 nm (red), respectively. We also calculated the Normalized Difference Red Edge Index (NDRE) (Barnes et al., 2000):

\[
\text{NDRE} = \frac{R_{790} - R_{720}}{R_{790} + R_{720}}.
\]

While both NDVI and NDRE are mostly indicators of canopy cover (Eitel et al., 2009), the simple ratio of \( \text{NDRE} / \text{NDVI} \) has a linear relationship with CCCi and has been used as a measure of leaf canopy greenness independent of leaf area (Eitel et al., 2009). We hypothesized that since the \( \text{NDRE} / \text{NDVI} \) ratio is reported to relate to greenness per se and not to plant size, it could serve as an effective alternative to visual SG assessments. LN was measured on three competitive plants (i.e., bordered by adjacent plants) from each plot. Six weeks after planting the seventh leaf was painted with enamel. In maize, the initial leaves often drop to the ground prior to reproductive stage. After pollination the total number of leaves was recorded. PH was measured on five competitive plants per plot as the distance from ground level to the collar of the flag leaf. Plot means for LN and PH were used for data analysis.

2.2.4 Statistical analyses

Analysis of variance for each trait was performed using PROC GLM in SAS (SAS Institute, 2008). Normality of residuals for each trait was tested with the
Shapiro-Wilk W statistic using PROC UNIVARIATE of SAS v9.1 (SAS Institute, 2008). Least-square mean values of genotypes over the two locations were computed using the LSMEANS statement, and PROC MEANS was used to calculate the range and standard error for each trait in both populations. The following linear model was used to perform the analyses:

\[ Y_{rge} = \mu + \alpha_g + \beta_e + \rho_r(\beta_e) + \alpha_g \beta_e + \epsilon_{rge}, \]

where \( Y_{rge} \) is the measured trait of genotype \( g \) in replication \( r \) at location \( e \); \( \mu \) is the grand mean; \( \alpha_g \) and \( \beta_e \) are the genotype and location main effects; \( \rho_r(\beta_e) \) is the replication effect nested within location; \( \alpha_g \beta_e \) is the interaction between genotype and location; and \( \epsilon_{rge} \) is the residual. Genotype sums of squares were partitioned into IRIL, SRIL, and the contrast IRIL vs. SRIL (i.e., population); the genotype × location sum of squares was partitioned accordingly. Genotype and location were considered random effects, and the contrast IRIL vs. SRIL was considered as a fixed effect. The genotype × location interaction mean square was used as the error term in the \( F \)-test for the genotype effects, whereas for the contrast IRIL vs. SRIL the overall error mean square was used. The analyses of variance were used to estimate genetic variances using the method of moments (Searle et al., 1992), i.e., the mean squares were equated to their respective expectations and the estimates of variance for each population were computed as follows: \( \sigma^2_G = \frac{(MS_G - MS_{GL})}{RL} \) as variance among genotypes within a population; \( \sigma^2_{GL} = \frac{(MS_{GL} - MS_e)}{R} \) as the variance of the interaction genotype ×
locations; \( \sigma^2_{\text{Phe}} = \frac{MS_G}{RL} \) as the phenotypic variance. In these expressions \( MS_G, MS_{GL} \) and \( MS_E \) refer to the mean squares of genotype, genotype by location interaction, and error; \( R \) and \( L \) refer to the number of replications and locations, respectively. Approximate 90% confidence intervals for \( \sigma^2_G \) were calculated as described by Knapp et al. (1987). Estimates of broad-sense heritability on a genotype-mean basis were computed as \( H^2 = \frac{\sigma^2_G}{\sigma^2_{\text{Phe}}} \). Ninety-percent confidence intervals were calculated for heritability using the method outlined by Knapp et al. (1985). Differences between populations for estimates of genetic variance components and heritability were declared significant when the confidence intervals for the two populations did not overlap. Least-square mean values of parents across environments were calculated to compare the means and standard errors of the two parental lines. The terms in the model were the same as above except that when the model was used to analyze parental lines genotype and genotype × location were considered fixed effects.

The phenotypic and genetic correlations among pairs of traits and their corresponding standard errors were computed across locations using multivariate restricted maximum likelihood (REML) estimation in PROC MIXED 9.1 (Holland, 2006). The estimated genetic \( \hat{r}_{g(xy)} \) and phenotypic \( \hat{r}_{p(xy)} \) correlations between traits x and y are given by:

\[
\hat{r}_{g(xy)} = \frac{\hat{\sigma}_{G(xy)}}{\sqrt{\hat{\sigma}^2_{G(x)} \cdot \hat{\sigma}^2_{G(y)}}}
\]
and

\[ \hat{r}_{p(xy)} = \frac{\hat{\sigma}_{p(xy)}}{\sqrt{\hat{\sigma}_{p(x)}^2 \cdot \hat{\sigma}_{p(y)}^2}} = \frac{\hat{\sigma}_{G(xy)} + \hat{\sigma}_{GE(xy)} + \hat{\sigma}_{e(xy)}}{\sqrt{\hat{\sigma}_{G(x)}^2 + \hat{\sigma}_{GE(x)}^2 + \hat{\sigma}_{e(x)}^2 \cdot \sqrt{\hat{\sigma}_{G(y)}^2 + \hat{\sigma}_{GE(y)}^2 + \hat{\sigma}_{e(y)}^2}}}, \]

where \( \hat{\sigma}_{G(xy)} \) and \( \hat{\sigma}_{p(xy)} \) are the estimated genetic and phenotypic co-variances between traits x and y; \( \hat{\sigma}_{G}^2 \) and \( \hat{\sigma}_{p}^2 \) are the estimated genetic and phenotypic variances; \( \hat{\sigma}_{GE(xy)} \) and \( \hat{\sigma}_{e(xy)} \) are the estimated genotype × environment and experimental error co-variances between traits x and y, while \( \hat{\sigma}_{GE}^2 \) and \( \hat{\sigma}_{e}^2 \) are the estimated genotype × environment and experimental error variances, respectively. Confidence intervals for genetic correlation coefficients were constructed as \( r \pm z_{(0.05)} \sigma_e \), where \( z_{(0.05)} \) is the value from the standardized normal distribution table at \( P = 0.05 \) and \( \sigma_e \) is the standard error of the correlation coefficient (Iqbal et al., 2007). Estimated correlation coefficients were regarded as significantly different from zero if their 95% confidence intervals did not include zero (Holland et al., 2003).
2.3 Results

2.3.1 Trait distributions in RIL populations and parental lines

The parental genotypes CG60 and CG102, both short-season adapted inbred lines, differed significantly ($P < 0.01$) for DA, DS, SG and PH but did not significantly differ for LN, NDRE or NDVI. However, the differences in parental trait values for all traits were very small compared to RIL trait value differences. Inbreds in both the IRIL and SRIL populations had highly variable and transgressive trait values. Genotype was a significant source of variation for all traits in both RIL populations ($P < 0.001$, Table 2.1). Location was a significant source of variation for all traits except PH and NDVI ($P < 0.05$), and location x genotype was significant for DA, DS, and LN ($P < 0.05$). CG60 flowers significantly earlier than CG102. The mean DA and DS for CG60 are respectively 2.25 and 2.5 days earlier than for CG102 (Table 2.2). Yet, within the IRIL and SRIL populations, DA differed up to 14 days (range 65 to 79 days), and DS differed up to 17 days (66 to 83 days). LN differences were as large as 6 leaves (12 to 18) within the RIL populations; PH ranged from 0.76 m to 1.94 m; SG differed by as much as seven visual units (1 to 8); NDVI ranged from 0.46 to 0.89; NDRE ranged from 0.23 to 0.51; and NDRE/NDVI ranged from 0.40 to 0.58 (Table 2.2, Supplemental Figure 2.1).

Linked loci may influence traits and be broken up by intermating. Mean trait values between the IRIL and SRIL populations significantly differed, but the differences were low in magnitude. Differences in DA, DS, LN, NDRE/NDVI, and
PH means between the two populations were highly significant ($P<0.001$). The NDRE difference was marginally significant ($P<0.05$) (Table 2.1). Mean DA, DS, LN, and PH, were all slightly lower within the IRIL versus the SRIL (Table 2.2). For example, DS was 71.89 days within the IRIL and 73.35 days within the SRIL. The mean DA differed by 1.36 days. In contrast, NDRE and NDRE/NDVI were slightly higher (0.40 vs. 0.39 and 0.50 vs. 0.49 respectively) in the IRIL population than the SRIL population. SG and NDVI did not significantly differ (Table 2.1).

2.3.2 Genetic variance estimates in the IRIL and SRIL populations

Increasing the recombination frequency of parental genes may cause a change in population genetic variances between the IRIL and SRILs. As expected, genetic variances of all traits were highly significant ($P<0.001$) for both IRIL and SRIL populations (Table 2.3). However, genetic variance estimates did not significantly differ ($P<0.05$) between populations (Table 2.3). Genetic variances were nominally high within the IRIL population with 1.04, 1.30, 1.07, 1.24, and 1.69 fold differences for DA, DS, LN, NDVI, and SG, respectively. Variances were nominally low for PH, NDRE and NDRE/NDVI with 0.73, 0.85 and 0.64 fold decreases, respectively (Table 2.3). Although five of eight traits, DA, DS, LN, SG, and NDVI, had higher variances in the IRIL, the traits are not independent. Broad-sense heritabilities of all traits, except SG in the SRIL population which was low (0.38) due to low genetic variance and high phenotypic variance, ranged from 0.70 to 0.90 (Table 2.3). There was no significant difference in heritabilities between populations (Table 2.3).
2.3.3 Phenotypic and genetic correlations among RILs

PH, stay green traits, and flowering time are all important for corn production, and we estimated the degree to which they are genetically and phenotypically correlated. Phenotypic correlations between two traits within a RIL population may occur because the traits are similarly influenced by changes in the environment and/or because the traits have a common genetic cause. In contrast, genetic correlations are derived from the genetic covariance and variance estimates (Holland, 2006). We defined significant positive correlations less than 0.3 as weak, significant positive correlations equal or greater than 0.3 but less than 0.6 as moderate, and significant positive correlations equal or greater than 0.6 as strong. The two flowering time traits, DA and DS, had strong phenotypic and genetic correlations with a mean phenotypic correlation of 0.69 and genetic correlation of 0.86 across both populations (Table 2.4). This result was expected because anthesis in maize is coordinated with the extrusion of silks on the developing ear. DA and DS also had moderate genetic correlations and moderate or weak phenotypic correlations with LN and PH (Table 4). LN and PH had moderate genetic correlations (mean 0.47) and weak phenotypic correlations (mean 0.28). The genetic correlations between DA and DS with SG in the SRIL population were strong and moderate, respectively, and both traits had weak phenotypic correlations (Table 2.4). SG measurements of the IRIL population had low genetic correlations with other traits perhaps because of high genetic variance. Both DA and DS were significantly negatively correlated with the NDRE/NDVI ratio (Table 2.4).
The canopy reflectance traits NDRE, NDVI, and NDRE/NDVI were positively correlated with visual stay green (SG). In the SRIL population, NDVI, NDRE, and NDVI/NDRE had genetic correlations with SG of 0.65, 0.71, and 0.57, respectively and phenotypic correlations of 0.30, 0.32, and 0.19, respectively. Both NDRE and NDRE/NDVI and NDVI and NDRE were strongly correlated (Table 2.4), as expected since NDVI and NDRE are both calculated using the 790nm reflectance. Interestingly, the phenotypic correlation of NDVI and NDRE exceeded the genetic correlation. The position of traits within a principle component analysis biplot reflected their correlation values (Supplemental Figure 2.2).

Genetic correlations between traits may be due to pleiotropic loci, linkage of loci with effects on the different traits, or a combination of both pleiotropy and linkage. No two traits had significantly different genetic or phenotypic correlations between the IRIL and SRIL populations (Table 2.4). In addition, if one removes SG correlations within the IRIL, there was no overall tendency for trait correlations to be higher within one population as compared to the other (Table 2.4).

2.4 Discussion

2.4.1 Genetic variation for flowering time and related traits

The RILs generated from a cross between CG60 and CG102, two short-season adapted CBD inbred lines, exhibited substantially greater phenotypic variation than would be expected given the parental values. For example,
although the parents’ DS differed by 3 days, the range for DS was 17 days within the RIL populations (Table 2.1). For comparison, among maize RILs generated by crossing B73, a CBD line adapted to Iowa, with 25 temperate and tropical lines, the greatest difference in number of days to silk within any one population was 17.5 days (B73 x Tzi8, a tropical line) (Buckler et al., 2009). The largest difference among any two RILs across the 25 populations was 28 days (Buckler et al., 2009). As flowering time is influenced by the environment, the short growing season (i.e., less total heat units) in Ontario, Canada likely revealed greater genetic variation in terms of earliness than would the sites used within the Buckler et al. (2009) analysis. Nonetheless the diversity of the CG60 x CG102 RILs for flowering is surprising.

The RILs were also highly variable for LN, PH, and the four stay green traits. The coefficients of variation for LN, PH and the four stay green traits among the CG60 x CG102 RILs were higher than the coefficients of variation for DA and DS (Table 2.2). A number of reports describe transgressive segregation for PH using diverse parental lines. In a cross between elite Chinese flint and dent inbreds that differed by 10 cm, Tang et al. (2007) found RIL heights had a range of 127 to 242 cm. Similar high levels of variation were observed by Wei et al. (2009) in a cross between two dent inbred lines with heights that differed by only 3 cm. For SG, Beavis et al. (1994) found that F_{2:4} lines arising from a cross between B73 and Mo17, two CBD lines with similar SG, varied for SG. In contrast, Zheng et al. (2009) evaluated a cross between a stay green Chinese inbred and Mo17, and transgressive segregation was infrequent. Because SG varies in the CG60 x
CG102 RIL populations and is only weakly correlated with DA and DS, further improvement of SG should be possible in short season germplasm.

Our study used two inbred lines that have extreme trait values relative to their source population, CBD from the mid-western United States. The high level of transgressive segregation among the RILs suggests that different sets of alleles became fixed in the Canadian lineages. Thus, the genetic architecture of flowering time among short season lines may be similar to that of lines adapted to Southern environments (e.g., Buckler et al., 2009). Alternatively, in the process of developing inbred lines from different heterotic groups for a short-season environment, non-optimal genes may have been fixed due to the small population sizes of the breeding lines, and selection favored the fixation of novel, compensatory mutations. Genetic analyses of DA, DS, LN, PH, and SG have found that loci that enhance and reduce the traits’ values often reside within a single parental genome. For example, Buckler et al.'s (2009) analysis of DA and DS found that every line contained alleles that both promoted and repressed flowering. Alleles contributing to SG were inherited from both B73 and Mo17 (Beavis et al., 1994). In contrast, SG positive alleles were inherited only from a short, stay green Chinese inbred in a cross with Mo17 (Zheng et al., 2009).

2.4.2 Effects of intermating on means and variances

If alleles have an additive effect on traits, then one would expect trait means within the IRIL to equal trait means within the SRIL. Here, DA, DS, LN, PH and stay green trait means significantly differed between the IRIL and SRIL.
populations (Table 2.1). Although significant, the differences between the populations were very small. For example DA and DS differed by one day. Inadvertent selection for short and early plants during intermating may have contributed to the small change in means. Alternatively, favourable linked epistatic genes that contribute to earliness may be present in the parents and were disrupted by intermating (Covarrubias-Prieto et al., 1989). Lamkey et al. (1995) reported no significant differences in either mean anthesis or silking dates between the non-intermated and intermated populations derived from B73 and B84 and testcrossing the two populations to Mo17. Buckler et al. (2009) also found little to no epistasis between alleles contributing to DA and DS.

Hanson and Hayman (1963) developed a model to associate genetic variance differences between IRIL and SRIL populations to the frequency of favourable alleles within one parent and to the number of genes segregating that affect a trait. We can use this model to predict gene numbers and the effects of linkage in this study. The model assumes that genes that affect a trait are uniformly distributed along a genetic map, all alleles have equal, additive effects, there is no additive × additive epistasis, and positive and negative alleles are not clustered within the genome. These assumptions are highly realistic for PH as well as DA and DS, assuming the genetic architecture of Northern adapted lines is similar to that of Southern adapted lines (Buckler et al., 2009; Wei et al., 2009). Under the model, if two parents with a number of different favorable alleles are crossed, a trait's genetic variance will increase slightly and gradually upon intermating because of the breakup of repulsion phase linkages. However, to
observe an increase in variance, a very large number of polymorphic loci must affect a trait. This number of polymorphic genes is unlikely between our parents. In contrast, a trait's genetic variance will decrease upon intermating under the model if one parent contains a greater number of favorable alleles than the other parent. Thus, in this work, the nominal decline in PH variance (Table 2.3) could be due to recombination among a moderate number of genes with coupling phase linkages. This scenario does not likely explain the nominal reduction of the canopy reflectance traits NDRE and NDVI/NDRE because parents have similar trait values.

Hill and Robertson (1966) predicted that favorable alleles have a higher chance of being in repulsion phase linkage disequilibrium if recombination has been limited, as was likely the case in the development of short-season lines. Recently, Flint and Mackay (2009) wrote that increasing mapping resolution typically shows closely linked loci that often have opposite effects. Interestingly, the most parsimonious explanation for our results is that trait variation among populations is due to genes that segregate independently. Increases or decreases in genetic variances in the intermated vs. non-intermated populations were not significant and very small relative to the differences between parental lines and the F2. An intermated population does greatly assist in the identification of genes allelic to QTL (Balasubramanian et al., 2009).
2.4.3 Linkage and pleiotropy

Variation in anthesis and silking dates is largely due to genetic differences in the timing of the apical meristem’s transition from the vegetative to the reproductive stage (Irish and Nelson, 1991). Thus, we and others have noted that lines that flower late tend to be relatively tall and have a large number of leaves, and lines that flower early tend to be relatively short and have a small number of leaves (Rood and Major, 1981). Within the CG60 x CG102 RILs, we observed that DA and DS were significantly correlated with LN and PH, with moderate genetic correlations and weak and moderate phenotypic correlations (Table 2.4). These correlation estimates are consistent with other studies (Tang et al., 2007), although Zheng et al. (2009) did not detect a significant correlation between leaf number and time to flower. We also expected a positive correlation of flowering time with the stay green traits because the late flowering lines would initiate the leaf senescence program later than the early flowering lines. In addition, across studies, flowering time is consistently associated with SG (e.g., Beavis et al., 1994; Zheng et al., 2009). In the CG60 x CG102 SRIL population, SG had moderate to strong genetic correlations with DA and DS, although SG phenotypic correlations with DA and DS were weak or non-significant (Table 2.4).

Moderate, significant genetic correlations between traits suggest the presence of loci that contribute pleiotropically to DA, DS, LN, PH and SG, and the presence of loci that influence one trait but not a second. Both pleiotropic loci and loci with specific effects have been characterized in maize, and both
classes of genes may be segregating in this cross. For example Vgt1, a major effect QTL, maps to a 2-kb non-coding region that lies 70kb upstream of the flowering time gene Zmrap2.7 on chromosome 8. This segment introgressed from Gaspe Flint into inbred N28 affects DA, LN, and PH, and Vgt1 alleles are associated with earliness (Ducrocq et al., 2008; Salvi et al., 2002; Salvi et al., 2007; Vladutu et al., 1999). A mutant likely allelic to vgt1, vfgt-f7p, decreases both time to flowering and the number of nodes (Chardon et al., 2005). Like vgt1, the maize mutants indeterminate1 and delayed flowering1 have pleiotropic effects. Both mutants cause increases in DA, DS, LN, and PH (Colasanti and Coneva, 2009; Muszynski et al., 2006). Other studies have shown DA, DS, LN, and PH to be genetically separable. Tang et al. (2007) mapped QTL for PH, LN and internode length in a population of 294 RILs in a Chinese dent x flint population. Four of the six QTL detected for internode length were in the same region as PH QTL suggesting that internode length and PH are controlled by the same genetic factors. However, separate QTL explained LN and PH variation (Tang et al., 2007). The early phase change (epc) mutant affects flowering time but not LN (Vega et al., 2002), and we have identified a mutation that affects PH and LN with little effect on flowering time (Avila and Lukens, unpublished data).

Previous work indicated that the NDRE/NDVI ratio was sensitive to leaf chlorophyll content but not to leaf area (Eitel et al., 2009). Thus, we hypothesized that this ratio would correlate with SG and serve as an effective substitute for visual SG ratings. In all cases, however, the NDRE/NDVI ratio was less strongly correlated with SG than were either of the two components of the
ratio (Table 2.4). One possible explanation for this result is that visual SG ratings were biased in favour of plots that were not only green, but that also had high leaf area; however, this idea seems inconsistent with the weak and insignificant genetic and phenotypic correlations between NDVI or NDRE and LN (Table 2.4). Nonetheless, visual SG is strongly, genetically correlated with the NDVI and NDRE reflectance indices (Table 2.4).

Pleiotropy, linkage, or both could explain trait correlations within the RIL populations. If linked loci positively contribute to two traits within one parental genome and negatively affect two traits in the second parental genome, then intermating decreases genetic correlations. Conversely, intermating increases genetic correlations if one locus positively affects one trait and a linked locus negatively affects a second. Genetic correlations between traits were highly similar in the IRIL and SRIL populations, and not one genetic correlation significantly differed between them (Table 2.4). Thus, genetic correlations are likely due to pleiotropy and not to linkage of loci with distinct effects.

Despite their similar trait values and adaptation to a relatively extreme environment, CG60 and CG102 harbor considerable genetic differences. Flowering time traits, stay green traits, LN, and PH are highly transgressive within the IRILs and SRILs. These results suggest that early-season lines can be selected for growth in still shorter seasons and that LN, PH, and stay green traits may be genetically altered while maintaining flowering time. Intermating these lines prior to inbreeding did not significantly affect genetic variation, phenotypic variation and correlations. Thus, although a number of studies suggest that
linked genes may contribute to genetic variation; our data seem best explained by independent segregation of unlinked genes with pleiotropic and additive effects. It will be interesting to determine the genetic basis for transgressive segregation and to identify the polymorphic loci within the CG60 x CG102 RILs.

Acknowledgements

This research was supported by a grant from the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) and the Natural Sciences and Engineering Research Council (NSERC) of Canada. We would like to thank Dr. Alireza Navabi for assistance in analyses, and Dr. Joseph Colasanti and Dr. Cortland Griswold for their comments on the manuscript. Byron Good and Heather Peterson provided for technical assistance to conduct field experiments.
References

Austin, D.F. 1995. Genetic analysis of quantitative trait loci with recombinant inbreds in maize. Iowa State University, Ames, IA.


Muszynski, M.G., T. Dam, B. Li, D.M. Shirbroun, Z. Hou, E. Bruggemann, R. Archibald, 
E.V. Ananiev, and O.N. Danilevskaya. 2006. delayed flowering1 encodes a basic 
leucine zipper protein that mediates floral inductive signals at the shoot apex in 

Rajcan, I., and M. Tollenaar. 1999. Source : sink ratio and leaf senescence in maize:: II. 

nitrogen deficiency in wheat from spectral reflectance indices and basic crop eco-

Rood, S.B., and D.J. Major. 1981. Diallel Analysis of leaf number, leaf development rate, 

Salas Fernandez, M.G., P.W. Becraft, Y. Yin, and T. Lübberstedt. 2009. From dwarves 
to giants? Plant height manipulation for biomass yield. Trends Plant Sci. 14:454-
461.

Salvi, S., R. Tuberosa, E. Chiapparino, M. Maccaferri, S. Veillet, L. van Beuningen, P. 
QTL controlling the transition from the vegetative to the reproductive phase in 

Salvi, S., G. Sponza, M. Morgante, D. Tomes, X. Niu, K.A. Fengler, R. Meeley, E.V. 
Ananiev, S. Svitashev, E. Bruggemann, B. Li, C.F. Hainey, S. Radovic, G. Zaina, 
Conserved noncoding genomic sequences associated with a flowering-time 


& Sons, New York.

Genetic dissection of plant height by molecular markers using a population of 
recombinant inbred lines in maize. Euphytica 155:117-124.


linked quantitative trait loci involved in the transition of the maize apical meristem 

Figure 2.1 Schematic of the IRIL and SRIL population development from two maize inbred lines. Details of population development are given in the text.
Supplemental Figure 2.1 Frequency distributions of days to anthesis, days to silking, leaf number, plant height, stay green, normalized difference red edge (NDRE) index, normalized difference vegetative index (NDVI) and NDRE/NDVI traits of intermated recombinant inbred line (IRIL) and selfed recombinant inbred line (SRIL) populations of maize grown in two environments in Ontario, Canada during 2007. The mean of parents (CG60 and CG102) are shown by arrows.
Supplemental Figure 2.2  Bioplot of Principal Components Analysis (PCA) axis 2 versus axis 1 for eight traits of intermated recombinant inbred line (IRIL) and selfed recombinant inbred line (SRIL) populations of maize grown in two environments in Ontario, Canada during 2007.
Table 2.1 Mean squares from the analyses of variances for days to anthesis (DA), Days to silking (DS), leaf number (LN), plant height (PH), stay green (SG), normalized difference vegetative index (NDVI), normalized difference red edge (NDRE) index and NDRE/NDVI traits across two environments in Ontario, Canada during 2007.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>DA</th>
<th>DS</th>
<th>LN§§</th>
<th>PH§</th>
<th>SG§</th>
<th>NDVI§§§</th>
<th>NDRE§§§</th>
<th>NDRE/NDVI§§§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loc</td>
<td>1</td>
<td>3399.94*</td>
<td>4635.37*</td>
<td>40.70</td>
<td>28.90*</td>
<td>120.44*</td>
<td>3996.47</td>
<td>8257.49*</td>
<td>5438.43**</td>
</tr>
<tr>
<td>Rep(Loc)</td>
<td>2</td>
<td>40.04***</td>
<td>49.17***</td>
<td>0.50***</td>
<td>794.63</td>
<td>4.43</td>
<td>505.72***</td>
<td>193.10***</td>
<td>35.54***</td>
</tr>
<tr>
<td>Genotypes</td>
<td>273</td>
<td>16.90***</td>
<td>18.16***</td>
<td>23.44***</td>
<td>68.49***</td>
<td>20.56***</td>
<td>31.03***</td>
<td>26.10***</td>
<td>20.23***</td>
</tr>
<tr>
<td>Pop (1)</td>
<td>501.94***</td>
<td>577.12***</td>
<td>61.65***</td>
<td>493.95***</td>
<td>10.16</td>
<td>14.64</td>
<td>18.53*</td>
<td>46.81***</td>
<td></td>
</tr>
<tr>
<td>IRIL (134)</td>
<td>15.62***</td>
<td>17.85***</td>
<td>23.84***</td>
<td>59.31***</td>
<td>19.24***</td>
<td>33.64***</td>
<td>23.92***</td>
<td>16.50***</td>
<td></td>
</tr>
<tr>
<td>SRIL (138)</td>
<td>14.60***</td>
<td>14.38***</td>
<td>22.76***</td>
<td>74.26***</td>
<td>21.91***</td>
<td>28.60***</td>
<td>28.27***</td>
<td>23.64***</td>
<td></td>
</tr>
<tr>
<td>Loc*Genotype</td>
<td>273</td>
<td>2.04**</td>
<td>2.48*</td>
<td>5.04</td>
<td>8.78*</td>
<td>9.54</td>
<td>8.80</td>
<td>3.87</td>
<td>2.69</td>
</tr>
<tr>
<td>Loc*Pop (1)</td>
<td>545</td>
<td>1.50</td>
<td>2.02</td>
<td>4.07</td>
<td>8.65</td>
<td>8.78</td>
<td>8.06</td>
<td>3.71</td>
<td>2.98</td>
</tr>
</tbody>
</table>

§, §§ and §§§ Variance estimates multiplied by 10, 10³ and 10⁴.

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.
Table 2.2 Mean trait value for two parents (CG60 and CG102) and means and ranges of the intermated recombinant inbred line (IRIL) and selfed recombinant inbred line (SRIL) populations for days to anthesis (DA), Days to silking (DS), leaf number (LN), plant height (PH), stay green (SG), normalized difference vegetative index (NDVI), normalized difference red edge (NDRE) index and NDRE/NDVI traits across two environments in Ontario, Canada during 2007.

<table>
<thead>
<tr>
<th>Traits</th>
<th>CG60</th>
<th>CG102</th>
<th>IRIL population</th>
<th>SRIL population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean (SE)</td>
<td>range</td>
<td>CV</td>
<td>mean (SE)</td>
</tr>
<tr>
<td>DA</td>
<td>71.75 ± 0.21†</td>
<td>74.00 ± 0.30</td>
<td>70.91 (0.12)</td>
<td>72.27 (0.12)</td>
</tr>
<tr>
<td>DS</td>
<td>71.00 ± 0.42</td>
<td>73.50 ± 0.59</td>
<td>71.89 (0.14)</td>
<td>73.35 (0.13)</td>
</tr>
<tr>
<td>LN</td>
<td>15.92 ± 0.25</td>
<td>15.42 ± 0.36</td>
<td>15.21 (0.04)</td>
<td>15.36 (0.04)</td>
</tr>
<tr>
<td>PH</td>
<td>1.31 ± 0.04</td>
<td>1.55 ± 0.06</td>
<td>1.39 (0.01)</td>
<td>1.44 (0.01)</td>
</tr>
<tr>
<td>SG</td>
<td>5.50 ± 0.35</td>
<td>7.00 ± 0.49</td>
<td>6.19 (0.04)</td>
<td>6.25 (0.05)</td>
</tr>
<tr>
<td>NDVI</td>
<td>0.77 ± 0.015</td>
<td>0.79 ± 0.021</td>
<td>0.80 (0.002)</td>
<td>0.80 (0.002)</td>
</tr>
<tr>
<td>NDRE</td>
<td>0.37 ± 0.009</td>
<td>0.38 ± 0.013</td>
<td>0.40 (0.001)</td>
<td>0.39 (0.001)</td>
</tr>
<tr>
<td>NDRE/NDVI</td>
<td>0.48 ± 0.008</td>
<td>0.49 ± 0.011</td>
<td>0.50 (0.001)</td>
<td>0.49 (0.001)</td>
</tr>
</tbody>
</table>

† Standard error of the parental mean
Table 2.3 Variance components and broad-sense heritability ($H^2$) estimates for days to anthesis (DA), days to silking (DS), leaf number (LN), plant height (PH), stay green (SG), normalized difference vegetative index (NDVI), normalized difference red edge (NDRE) index and NDRE/NDVI traits measured on intermated recombinant inbred line (IRIL) and selfed recombinant inbred line (SRIL) populations in two environments in Ontario, Canada during 2007.

<table>
<thead>
<tr>
<th>Traits</th>
<th>Population</th>
<th>$\sigma^2_{Ph}$</th>
<th>$\sigma^2_G$</th>
<th>$\sigma^2_{G+L}$</th>
<th>$\sigma^2_E$</th>
<th>$H^2$</th>
<th>$\dagger$ Ratio $\sigma^2_G$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DA</td>
<td>IRIL</td>
<td>3.91</td>
<td>3.33 (2.65; 4.26) $\dagger$</td>
<td>0.42</td>
<td>1.46</td>
<td>0.85 (0.82; 0.89) $\dagger$ $\ddagger$</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>3.65</td>
<td>3.21 (2.56; 4.07)</td>
<td>0.14</td>
<td>1.54</td>
<td>0.88 (0.85; 0.90) $\dagger$ $\ddagger$</td>
<td>1.30</td>
</tr>
<tr>
<td>DS</td>
<td>IRIL</td>
<td>4.46</td>
<td>3.86 (3.08; 4.92)</td>
<td>0.25</td>
<td>1.90</td>
<td>0.87 (0.84; 0.90)</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>3.59</td>
<td>2.97 (2.33; 3.82)</td>
<td>0.20</td>
<td>2.14</td>
<td>0.83 (0.79; 0.86)</td>
<td>1.07</td>
</tr>
<tr>
<td>LN</td>
<td>IRIL</td>
<td>5.96</td>
<td>4.75 (3.70; 6.17)</td>
<td>0.41</td>
<td>4.02</td>
<td>0.80 (0.75; 0.84)</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>5.69</td>
<td>4.44 (3.43; 5.80)</td>
<td>0.44</td>
<td>4.14</td>
<td>0.78 (0.73; 0.83)</td>
<td>1.07</td>
</tr>
<tr>
<td>PH</td>
<td>IRIL</td>
<td>1.48</td>
<td>1.22 (0.96; 1.58)</td>
<td>0.13</td>
<td>0.79</td>
<td>0.83 (0.79; 0.86)</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>1.86</td>
<td>1.68 (1.35; 2.12)</td>
<td>-0.11</td>
<td>0.94</td>
<td>0.90 (0.88; 0.92)</td>
<td>0.73</td>
</tr>
<tr>
<td>SG</td>
<td>IRIL</td>
<td>4.81</td>
<td>3.48 (2.62; 4.63)</td>
<td>-0.09</td>
<td>5.51</td>
<td>0.72 (0.67; 0.78)</td>
<td>1.69</td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>5.48</td>
<td>2.06 (0.97; 3.42)</td>
<td>0.83</td>
<td>12.02</td>
<td>0.38 (0.24; 0.51)</td>
<td>1.24</td>
</tr>
<tr>
<td>NDVI</td>
<td>IRIL</td>
<td>8.41</td>
<td>6.20 (4.71; 8.22)</td>
<td>-0.63</td>
<td>10.08</td>
<td>0.74 (0.68; 0.80)</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>7.15</td>
<td>4.99 (3.69; 6.69)</td>
<td>1.28</td>
<td>6.14</td>
<td>0.70 (0.63; 0.76)</td>
<td>0.85</td>
</tr>
<tr>
<td>NDRE</td>
<td>IRIL</td>
<td>5.98</td>
<td>5.11 (4.07; 6.54)</td>
<td>-0.28</td>
<td>4.03</td>
<td>0.85 (0.82; 0.89)</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>7.07</td>
<td>6.01 (4.76; 7.68)</td>
<td>0.44</td>
<td>3.41</td>
<td>0.85 (0.82; 0.88)</td>
<td>0.85</td>
</tr>
<tr>
<td>NDRE/NDVI</td>
<td>IRIL</td>
<td>4.13</td>
<td>3.41 (2.69; 4.39)</td>
<td>-0.22</td>
<td>3.30</td>
<td>0.83 (0.79; 0.87)</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>5.91</td>
<td>5.30 (4.26; 6.69)</td>
<td>-0.10</td>
<td>2.70</td>
<td>0.90 (0.87; 0.92)</td>
<td>0.64</td>
</tr>
</tbody>
</table>

$\dagger$ Ratio of IRIL and SRIL genetic variance
$\ddagger$ 90% confidence interval of genetic variance
$\ddagger\ddagger$ 90% confidence interval of heritability
$§$, $§§$ and $§§§$ Variance estimates multiplied by 10, 10² and 10⁴
Table 2.4 Genetic (above diagonal) and phenotypic correlations (below diagonal) for days to anthesis, dDays to silking, leaf number, plant height, stay green, normalized difference vegetative index (NDVI), normalized difference red edge (NDRE) index and NDRE/NDVI traits measured on intermated recombinant inbred line (IRIL) and selfed recombinant inbred line (SRIL) populations of maize lines grown in two environments in Ontario, Canada during 2007.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Population</th>
<th>Days to anthesis</th>
<th>Days to silking</th>
<th>Leaf number</th>
<th>Plant height</th>
<th>Stay green</th>
<th>NDVI</th>
<th>NDRE</th>
<th>NDRE/NDVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days to anthesis</td>
<td>IRIL</td>
<td>0.87</td>
<td>0.58</td>
<td>0.43</td>
<td>0.38</td>
<td>0.25</td>
<td>-0.12</td>
<td>-0.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>0.85</td>
<td>0.50</td>
<td>0.34</td>
<td>0.74</td>
<td>0.23</td>
<td>-0.13</td>
<td>-0.31</td>
<td></td>
</tr>
<tr>
<td>Days to silking</td>
<td>IRIL</td>
<td>0.67</td>
<td>0.41</td>
<td>0.42</td>
<td>0.12</td>
<td>0.12</td>
<td>-0.17</td>
<td>-0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>0.71</td>
<td>0.40</td>
<td>0.55</td>
<td>0.26</td>
<td>0.30</td>
<td>-0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leaf number</td>
<td>IRIL</td>
<td>0.35</td>
<td>0.22</td>
<td>0.47</td>
<td>0.20</td>
<td>0.23</td>
<td>-0.07</td>
<td>-0.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>0.32</td>
<td>0.24</td>
<td>0.47</td>
<td>0.16</td>
<td>0.17</td>
<td>-0.23</td>
<td>-0.41</td>
<td></td>
</tr>
<tr>
<td>Plant height</td>
<td>IRIL</td>
<td>0.23</td>
<td>0.18</td>
<td>0.27</td>
<td>0.02</td>
<td>0.06</td>
<td>-0.29</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>0.13</td>
<td>0.17</td>
<td>0.29</td>
<td>-0.08</td>
<td>-0.09</td>
<td>-0.30</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>Stay green</td>
<td>IRIL</td>
<td>0.20</td>
<td>0.11</td>
<td>0.05</td>
<td>0.01</td>
<td>0.38</td>
<td>0.33</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>0.22</td>
<td>0.12</td>
<td>0.04</td>
<td>0.02</td>
<td>0.65</td>
<td>0.71</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>NDVI</td>
<td>IRIL</td>
<td>0.10</td>
<td>-0.05</td>
<td>0.12</td>
<td>0.13</td>
<td>0.15</td>
<td>0.69</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>-0.01; 0.21</td>
<td>-0.06; 0.16</td>
<td>-0.01; 0.23</td>
<td>-0.17</td>
<td>0.12</td>
<td>0.10</td>
<td>0.14</td>
<td>0.23</td>
</tr>
<tr>
<td>NDRE</td>
<td>IRIL</td>
<td>-0.10</td>
<td>-0.12</td>
<td>-0.06</td>
<td>-0.07</td>
<td>0.19</td>
<td>0.74</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>-0.10</td>
<td>0.21</td>
<td>0.09</td>
<td>-0.08</td>
<td>0.32</td>
<td>0.70</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>NDRE/NDVI</td>
<td>IRIL</td>
<td>-0.22; 0.02</td>
<td>-0.24; 0.00</td>
<td>-0.18; 0.06</td>
<td>-0.19; 0.05</td>
<td>0.09; 0.29</td>
<td>0.69; 0.79</td>
<td>0.81; 0.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SRIL</td>
<td>-0.22; 0.02</td>
<td>-0.32; -0.10</td>
<td>-0.21; 0.03</td>
<td>-0.02; 0.04</td>
<td>0.24; 0.40</td>
<td>0.64; 0.76</td>
<td>0.85; 0.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.22; -0.11</td>
<td>-0.30; -0.08</td>
<td>-0.30; -0.08</td>
<td>-0.28; -0.06</td>
<td>0.02; 0.22</td>
<td>0.01; 0.20</td>
<td>0.16</td>
<td>0.66</td>
</tr>
</tbody>
</table>

† Confidence interval of the correlation coefficient.
Chapter 3 Linkage Map and QTL Analysis Using Intermated vs. Selfed Recombinant Inbred Line Populations

Abstract

Intermating of individuals in a population increases the possibility of genetic recombination which is useful for linkage mapping and quantitative trait loci (QTL) analysis. The objectives of this study were to compare the recombination frequencies and precision of QTL detection in an intermated recombinant inbred line (IRIL) population with those of a selfed recombinant inbred line (SRIL) population. Both IRIL and SRIL populations were developed from maize inbreds CG60 and CG102. The recombinant inbred lines were grown in two environments for phenotypic evaluations, and individual lines from both populations were genotyped with SSR and SNP markers for linkage map construction and QTL identification. In addition, we simulated a selfed recombinant inbred line population and an intermated recombinant inbred line population from two inbred parents. In both empirical and simulated studies, the recombination event was higher in the IRIL population compared to the SRIL population, and the QTL support interval was reduced in the IRIL population compared to the SRIL population. We mapped 21 quantitative trait loci (QTL) for all traits both SRIL and IRIL populations. All QTL had small to moderate effects and QTL detected for days to silking (DS) and leaf number (LN) were common between SRIL and IRIL populations. In IRIL population lower number of QTL were detected than SRIL population. This study supports the importance of intermated RIL populations for map expansion and precise QTL mapping.
3.1 Introduction

Genetic linkage maps are important tools for genetic research and plant breeding programs. A genetic linkage map is a linear order of gene and/or marker loci on a chromosome of a species that shows the position of known genes or genetic markers relative to each other in terms of recombination frequency (Terwilliger and Ott, 1994). The distances on a genetic map are based on the frequencies of recombination between markers due to crossover of homologous chromosomes, estimated as the frequency of recombination between pairs of loci. The increased frequency of recombination between two markers results in increased recombination between the markers on the chromosome.

Developing populations is the first step in linkage mapping and these mapping populations can be used for quantitative trait loci (QTL) analysis. Often, mapping populations are derived from bi-parental crosses of homozygous parents, followed by a few generations of selfing. These selfed populations are referred to as recombinant inbred line populations.

Recombinant inbred line (RIL) populations are useful in QTL mapping as they provide unique, immortalized combinations of parental alleles and can be used in different environments for phenotypic evaluation (Burr and Burr, 1991). In addition, during the process of selfing from the F2 generation, there are additional opportunities for recombination between markers in the RIL population compared to F2 populations. The main problem with RIL populations in mapping studies is
the limited number of effective recombination events that occur. Hence, it is often difficult to map tightly linked QTL to a precise interval on a chromosome.

Intermating leads to improved genetic maps, with more opportunities to order closely linked markers. In order to effectively map tightly linked markers a population of RILs derived from advanced intercrossing, referred to as an intermated recombinant inbred line (IRIL) population, has been recommended (Darvasi and Soller, 1995). In such populations, unlike the RIL populations beginning with the F2 generation, individuals are randomly intercrossed, thus increasing the frequency of effective recombination between linked markers before genotypes are made homozygous by selfing (Darvasi and Soller, 1995; Liu et al., 1996). The high frequency of recombination between tightly linked loci results in a small standard error (σr) of the estimated recombination frequency (r) per meiosis and, therefore, increases the chance of detecting the order of tightly linked loci.

The intermated B73 × Mo17 (IBM) population is an IRIL maize population developed by four generations of intermating following the formation of the F2 generation prior to the development of inbred lines (Lee et al., 2002). The increased recombination between linked loci has had the effect of expanding the genetic map approximately four-fold compared to the map generated using the F2 population (Lee et al., 2002). The IBM population has been used for fine mapping of traits such as disease resistance (Balint-Kurti et al., 2008). Balint-Kurti et al. (2008; 2010) compared the precision of QTL detection in RIL and IRIL populations, both derived from B73 × Mo17 parents, for disease resistance traits.
in maize. These authors reported reductions of QTL support intervals (the intervals defining the region within two log-of-odds levels of the peak value) in the IRIL population compared to the RIL population.

The IRIL population also allows the separation of closely linked QTL. Balint-Kurti et al. (2008) compared the IRIL population with the RIL population for resistance to Southern Leaf Blight (SLB) disease and found two distinct QTL linked in repulsion phase linkage in the IRIL population, whereas only one QTL was identified in the RIL population developed from B73 × Mo17 parents. Huang et al. (2010) compared an intermated population and a F₂ -derived population from F₂ and F252 parental lines of maize and found two distinct, linked QTL for grain yield in coupling phase linkage in the intermated population, whereas only one QTL was detected in the F₂ -derived population.

The objectives of this study were (i) to study the changes in the frequency of recombination between the IRIL and SRIL populations, (ii) to compare QTL detection within the IRIL and the SRIL populations, and (iii) to discuss empirical results relative to the expectations from simulations.
3.2 Materials and methods

3.2.1 Empirical study

3.2.1.1 Plant materials

Two recombinant inbred line populations, described in detail by Khanal et al. (2011), were developed from crosses between the Iodent inbred line CG60 (Lee et al., 2001a) and the Stiff Stalk inbred line CG102 (Lee et al., 2001b). One RIL population (n = 136) was derived through four generations of inbreeding, starting from the F₂, using a single-seed descent approach. This population is hereafter referred to as a selfed recombinant inbred line (SRIL) population. The other RIL population (n=135) was derived from three generations of intermating, starting from F₂, followed by three generations of selfing. This population is hereafter referred to as the intermated recombinant inbred line (IRIL) population. The details of population development are given in Chapter Two.

3.2.1.2 Field trials and experimental design

The parental inbreds and recombinant inbred lines from both populations were evaluated at two field locations in Elora and Guelph, Ontario, Canada. Trials were planted on May 9th 2007 at Elora and on May 15th 2007 at Guelph using a randomized complete block design (RCBD) with two replications. Each plot consisted of a single-row with 20 plants per row, 4.0 m long and 76 cm row spacing. Fertilizer was applied on the basis of soil tests performed prior to seeding at the rate of 124, 36, 18 kg ha⁻¹ of N, P₂O₅ and K₂O, respectively, which was supplemented with 15 t ha⁻¹ of solid cattle manure. Weeds were controlled
as necessary with conventional herbicide applications using recommended products.

3.2.1.3 Phenotypic data collection

Each line was evaluated for number of days to anthesis (DA), days to silking (DS), leaf number (LN), plant height (PH), visual stay green (SG), and canopy spectral reflectance. DA, DS, and SG were determined on a plot-basis. DA and DS were recorded when 50% of the plants in the plot were shedding pollen or had visible silks, respectively. SG was recorded at physiological maturity (i.e., black layer formation) at a fixed date using a visual rating of 1 to 9, where 1 is complete senescence of plants in the row and 9 is a row of completely green plants. Canopy spectral reflectance on all plots at physiological maturity was measured using a dual channel reflectance spectrometer (Unispec DC, PP Systems, Amesbury MA) as described in Khanal et al. (2011). Briefly, the ratio of canopy reflectance to incoming solar radiation was recorded simultaneously at 3-nm intervals between 310 and 1100 nm. The downward-looking fibre optic was positioned approximately 70 cm above the row, aimed downward at a 45-degree angle and parallel to the row direction of the plot. The downward- and upward-looking channels were cross calibrated just prior to field measurements by measuring reflectance of a barium sulphate standard using the same measurement geometry and orientation as used for the field measurements. Two reflectance scans were made per plot, and the mean was calculated. From the canopy spectral reflectance data, the Normalized Difference Vegetation Index (NDVI), Normalized Difference Red Edge Index (NDRE) and the ratio of NDRE /
NDVI (referred to as RATIO hereafter) were calculated. The LN was measured on three representative plants bordered by adjacent plants from each plot. The total number of leaves was counted two times. Six weeks after planting, the seventh leaf was painted with enamel paint. Then, after pollination, the total number of leaves was recorded starting from the seventh leaf. The PH was measured on five representative plants in each plot, as the distance from ground level to the collar of the flag leaf. Plot means of LN and PH were used for data analysis.

3.2.1.4 Genotyping

Leaf tissues were harvested from individual plants of parental lines and one plant from each recombinant inbred line within the SRIL and IRIL populations. Harvested leaves were freeze-dried and ground to powder using the FastPrep system (Qbiogene, Irvine, CA). Genomic DNA extraction was performed according to the CTAB method of Saghai-Maroof et al. (1984) with modifications following CIMMYT Applied Biotechnology Center’s Manual of Laboratory Protocols (http://www.cimmyt.org/ABC/Protocols/manualABC.htm; accessed 2006/01/18).

SSR genotyping. Parental lines, CG60 and CG102, were screened with 748 publicly available simple sequence repeat (SSR) markers from the MaizeGDB (http://www.maizedb.org/ssr.php; accessed 2006/01/18). Out of the 132 polymorphic SSR markers, 90 SSR markers, evenly distributed across the genome were selected to genotype both populations. The PCR amplification and size-fractionation of PCR products, using electrophoresis, were conducted as
described in MaizeGDB (http://www.maizemap.org/ssr_methods.htm; accessed 2006/01/18). The PCR reactions were conducted in a 15 µl volume containing 10 ng template genomic DNA, 0.15 mM of each dNTP, 2.5 mM MgCl₂, 0.25 µM of each primer, 1x Taq DNA polymerase buffer, and 0.5 U of Taq DNA polymerase (Invitrogen, GmbH, Karlsruhe, Germany). The PCR temperature cycles included: (1) activation of AmpliTaq for 10 min at 95°C; (2) 30 cycles of denaturation at 94°C for 45 s, annealing at 55°C for 45 s, and extension at 68°C for 45 s; and (3) a final extension at 72°C for 10 min. The PCR products were stored at 4°C. The PCR products were size-fractionated using electrophoresis performed on 3% (w/v) Metaphor agarose (Bio-Whitaker Molecular Applications, Rockland, ME) gels in 1x TBE buffer using Sunrise 96 gel electrophoresis units (Invitrogen) running at 150 V for 2 hours. PCR fragments were visualized using an imaging system (UVP, Upland, Calif.) following staining with ethidium bromide.

**SNP genotyping.** Marker genotyping was performed with an additional 51 polymorphic SNP markers, evenly distributed in the maize genome. Putative SNPs between CG60 and CG102 were identified from several public sources (Gore et al., 2009). To identify SNPs, custom perl scripts were used. First, BLASTN was used to associate a gene of interest with the predicted transcript from B73. Second, maize HapMap data was used to determine the position of predicted SNPs within the transcript (Gore et al., 2009). Third, the SNP and 100nt of sequence context on either side of the polymorphic site was used to BLAST against the pseudo-chromosomes to localize the SNP within the maize genome. Finally, a list of SNPs was further filtered for ones which the maize
HapMap project observed as polymorphic between B73-MO17 or B73-OH43. The selected SNPs were first screened to identify the SNP polymorphism between the parental genotypes, CG60 and CG102. Fifty-one SNPs were found to be polymorphic between the parents and selected for further genotyping of the RIL lines. Genotyping of the SNPs was performed using the iPLEX Gold SNP genotyping assay on the Sequenom MassARRAY Platform. A total of 49 SNP loci were successfully amplified and used in linkage mapping.

### 3.2.1.5 Linkage Mapping and QTL Analyses

Linkage maps comprising 71 SSR and 49 SNP loci were constructed using the software package Joinmap 4.0 (Van Ooijen, 2006) separately for the SRIL and IRIL populations. Each marker locus was tested for conformity to the expected allelic frequency of 0.5 in both populations by $\chi^2$ tests in Joinmap 4.0 (Van Ooijen, 2006). Markers were assigned to linkage groups using the logarithm of odds ratio, the ratio of the probability that two loci are linked with a given recombination value over a probability that the two loci are not linked. The markers were first ordered into linkage groups using the “group” command (parameter value LOD > 5). The remaining markers were added to the respective linkage groups, based on previous mapping information for these markers (Lawrence et al., 2004), using the “assign” command. A marker was discarded during the mapping stage if its presence caused inconsistencies in the map. The final maps included 120 markers in each population. The units of distance in the intermated recombinant inbred population are not, strictly speaking,
centimorgans (cM). “Intermated centimorgans” (IcM) are therefore used as a measure of genetic distance for this study.

QTL analyses were performed using MapQTL6 (Van Ooijen, 2004). Multiple QTL mapping (MQM) was carried out with cofactors initially identified by simple interval mapping (Lander and Botstein, 1989) and subsequently by an initial round of MQM. In each analysis, a LOD threshold for declaring a significant QTL for each trait was computed using 1000 permutations (Churchill and Doerge, 1994). The LOD threshold corresponded to a 10% genome-wide type I error. The QTL support interval was estimated from the LOD curves calculated by the interval mapping approach executed in MapQTL 6.0. The 2-LOD fall method was used to estimate the 95% confidence interval (Van Ooijen, 1992). The graphical representation of the linkage maps and QTL were prepared using MapChart (Voorrips, 2002)

3.2.2 Computer simulation study

3.2.2.1 Genetic models

Perl scripts (kindly provided by N. Tinker and H. Pham) were used to simulate genetic models. For simulations, we assumed a hypothetical linkage group with a length of 150 cM with a marker density of 5 cM, corresponding to 31 evenly distributed markers. Two genetic models were considered. In the first model, the linkage group had a single QTL with different effect sizes (Table 3.1). In the second model, the linkage group had two QTL linked with different effect sizes, linkage distances, and linkage phases (Table 3.2).
For each genetic model, two sets of recombinant inbred line populations (n=140) were simulated. In the first set we simulated an intermated recombinant inbred line population (IRIL) with three generations of intermating from the F2 generation followed by three generations of selfing. In the second set, we simulated a selfed recombinant inbred lines (SRIL) population with four generations of selfing from the F2 generation. In addition, for the linked QTL model, the population size was increased 3 fold (n= 420) to study the effect of population size on the likeliness of separating linked QTL. We simulated 1000 SRIL and 1000 IRIL populations. Each population was derived from the F1 of a cross between two parental inbreds, homozygous for alternative alleles at the QTL and at all marker loci. The model for the phenotypic variance ($\sigma_p^2$) of a trait in a population of recombinant inbred lines was:

$$\sigma_p^2 = \sigma_{QTL}^2 + \sigma_{BG}^2 + \sigma_E^2$$  \hspace{1cm} (1)$$

where, $\sigma_p^2$ is the phenotypic variance, $\sigma_{QTL}^2$ is the variance explained by the QTL, $\sigma_{BG}^2$ is the variance of the background genetic effect, $\sigma_E^2$ is the environmental variance.

**Single QTL models.** In the single QTL model, we tested six QTL with different levels of additive effects on a quantitative trait of interest. The phenotypic variation explained by the QTL ($R^2_{QTL}$) ranged from 0.05 to 0.30. With complete additive effects and with equal marker allele frequencies ($P = q = 0.5$), the additive effect of the QTL ($a$), is half the difference between alternative
homozygotes at the QTL (Mather and Jinks, 1971). In the F₅ population, the additive effect is:

\[ a = \sqrt{\frac{\sigma^2_{QM}}{15/16}} \]  

(2)

Using equations 1 and 2, values of \( a \) and the sum of \( \sigma^2_{BG} \) and \( \sigma^2_E \) were set to provide models, in which the simulated QTL was responsible for 5, 10, 15, 20, 25, and 30\% of \( \sigma^2_p \) (Table 3.1).

**Linked QTL models.** In the model with two linked QTL, two linkage phases, coupling and repulsion, and three linkage distances, 10, 20, and 30 cM, were considered. The additive effects of QTL1 and QTL2 were the same (\( a=30 \)) but with different signs to simulate the two linkage phases (Table 3.2). The proportion explained by QTL (R²) equal to the genotypic variance because phenotypic variance is 1. The proportion of variance that was accounted for by each model (R²) was estimated as:

\[ V_G = a_1^2 + a_2^2 + 2(1-2r)a_1a_2 \]  

(3)

where \( a_1 \) and \( a_2 \) are the additive effects of QTL1 and QTL2, respectively, and \( r \) is the recombination frequency between QTL1 and QTL2 (Li et al., 2010). The Haldane mapping function was used to convert genetic distances to recombination frequencies. These resulted in phenotypic variances explained by QTL (\( R^2_p \)) for coupling linkages as 0.32, 0.29, and 0.26 and for repulsion linkages as 0.04, 0.07, and 0.10. Based on the estimated \( R^2_p \) value, the sum of \( \sigma^2_{BG} \) and \( \sigma^2_E \) was estimated for linkage distance 10, 20 and 30 cM as 0.68, 0.71 and 0.74,
respectively, for coupling linkages. Similarly the sum of $\sigma_{BG}^2$ and $\sigma_{E}^2$ was estimated for linkage distance 10, 20 and 30 cM as 0.96, 0.93 and 0.90, respectively, for repulsion linkages. These parameters were then used in the simulation.

3.2.2.2 QTL mapping

The entire genome was scanned using the Interval Mapping approach with PROC QTL in SAS (Hu and Xu, 2009) under the single QTL model (Han and Xu, 2008), which is

$$Y_j = X_j \beta + \varepsilon_j$$

(4)

where $Y_j$ is the phenotypic value of individual $j$ ($j = 1$ to $n$); $X_j$ is the expectation of variables indicating the QTL genotype; $\beta$ is a vector of QTL effects, and $\varepsilon_j$ is the residual effect. The maximum likelihood method was implemented with 1.0 cM steps in interval mapping. The output from PROC QTL was used to obtain the QTL position and to perform a likelihood ratio test (LRT) value of the QTL. The LRT statistics were divided by 4.61 to obtain their corresponding LOD scores. The PROC MEANS procedure in SAS v9.2 (SAS Institute, 2008) was used to estimate the mean, upper limit and lower limit LOD values for each marker locus with LOD scores higher than the LOD = 3 empirical threshold over 1000 simulations for each genetic model. The upper and lower limit values were used as confidence interval for respective QTL.
3.3 Results

3.3.1 Empirical data analysis

3.3.1.1 Comparison of linkage maps between SRIL and IRIL populations

A total of 120 molecular markers, including 71 SSR and 49 SNP loci, were mapped on ten linkage groups, totaling 1788.74 cM in the SRIL and 2516.79 IcM in the IRIL populations (Figure 3.1, Table 3.3). Thus, there was an overall map expansion of 1.41 fold in the intermated RIL population. The average linkage distance between adjacent loci was 16.26 cM in the SRIL population, significantly less than the average distance of 23.04 IcM in the IRIL population ($P <0.001$). The marker interval distances were 1.17 to 1.76 fold higher in the IRIL population relative to interval distances in the SRIL population. The sum of the estimated recombination events in the SRIL population was 493.50. The sum of the estimated recombination events in the IRIL population was 688.90.

The order of loci was 96% identical between the linkage maps for the two populations. Only five loci (4%) found on linkage groups 1, 2, 3 and 6 had different orders between the two maps. The loci, umc1076 and ug2, umc1065 and umc1028, ug11 and mmc0132, ug23 and umc1133, and ug25 and umc1887 were inverted between the two maps (Figure 3.1).

3.3.1.2 Marker genotypic frequencies differ between populations

A chi-square goodness-of-fit test was performed for each marker locus in each population to test the conformity of genotypic segregation ratios with the expected 0.5 frequency. Of the 120 loci, 24 loci in SRIL population had significant
(P < 0.01) segregation distortion. In the IRIL population, 32 loci out of 120 loci had significant (P < 0.01) segregation distortion. Among the distorted markers, 10 distorted marker loci were common between SRIL and IRIL population.

### 3.3.1.3 Comparison of QTL in SRIL and IRIL populations

We identified a total of 14 QTL in the SRIL population and seven QTL in the IRIL population (Table 3.4, Figure 3.1). Within the SRIL population, two QTL were detected on chromosomes 2 and 6 for DA, which explained 10 and 9.7% of the phenotypic variation, respectively. Three QTL were detected on chromosomes 1, 2, and 8 for DS, which accounted for 6.5, 6.6, and 9.3% of phenotypic variance, respectively. One QTL was detected for LN on chromosome 7 and accounted for 12.5% of phenotypic variance. One QTL was detected for PH on chromosome 8 and explained about 17.6% of the phenotypic variance. One QTL was identified for SG on chromosome 8 accounted for 10.6% of phenotypic variance. One QTL was identified for NDRE on chromosome 1 and accounted for 10.8% of variation. One QTL for NDVI detected on chromosome 2 explained 19.3% of the phenotypic variation. Four QTL were detected for RATIO on chromosomes 1, 1, 2 and 8 which explained 12.3, 6.4, 5.8, and 9.5% of the phenotypic variance, respectively. The QTL overlapped for NDRE and RATIO on chromosome 1, DA, DS and NDVI on chromosome 2 and for DS, PH, SG, and RATIO on chromosome 8.

Multiple QTL mapping detected a total of seven QTL for DA, DS and LN in the IRIL population (Table 3.4, Figure 3.1). Four QTL were detected for DA on chromosomes 1, 4, 5, and 10. These QTL explained 8.4, 9.1, 9.9, and 7.8% of
phenotypic variance, respectively. Two QTL were detected for DS on chromosomes 1 and 10 explained 8.5 and 8.5% of phenotypic variance, respectively. One QTL was detected for LN on chromosome 7 that explained 10.1% of phenotypic variance. QTL were not detected for PH, SG, NDRE, NDVI, and RATIO. The QTL overlapped for DA and DS on chromosome 1.

The average support interval of all detected QTL was 26 cM and 20 IcM for the SRIL and the IRIL populations, respectively, corresponding to a support interval reduction by a factor of 1.30 in the IRIL compared to the SRIL population. QTL for DS and LN detected on chromosomes 1 and 7, respectively, were identified in both two populations.

3.3.2 Simulation study

3.3.2.1 Comparison of recombination frequency between SRIL and IRIL populations

Frequencies of recombination are the events of physical exchanges of chromosomal portions from the different parents of a cross. In order to estimate the recombination events, we randomly selected five simulated populations from each of the SRIL and IRIL population types. The hypothetical population had a linkage group with 31 markers separated by 5 cM. The genotypic data were analyzed in Joinmap 4.0 to estimate the recombination events in SRIL and IRIL populations. The average recombination events between linked markers were 752 and 1221 in the SRIL and IRIL populations, respectively. The result showed that the recombination events increased by 1.62-fold when population development included three generations of intermating.
3.3.2.2 Effect of intermating on single QTL

In a single QTL model, QTL with high phenotypic variance (\( \sigma^2_{\text{QTL}} \)) had high peak LOD scores at the QTL position in both the simulated SRIL and IRIL populations. The LOD score was the highest at the QTL position, and there was no significant difference for the LOD scores at the peak between the simulated IRIL and SRIL populations (Figure 3.2 and Table 3.5). The size of the support interval was always lower for the IRIL population than the SRIL population (Table 3.5).

3.3.2.3 Effects of intermating on the detection of linked QTL

In two-QTL models with QTL placed 10, 20, and 30 cM apart, interval mapping in both simulated SRIL and IRIL populations identified single QTL (Figure 3.3, A and B) in coupling phase linkage. The LOD scores of QTL in the SRIL population were higher than the LOD scores in the IRIL population. The QTL support interval was reduced in the IRIL population as compared to the SRIL population. Whereas in repulsion phase linkage with two QTL placed in 10, 20, and 30 cM apart, interval mapping in both SRIL and IRIL populations identified two peaks. However, due to the fact that the two QTL in repulsion cancel each other’s effects, we did not detect a significant QTL (Figure 3.3, C and D).

We simulated SRIL and IRIL populations with two linked QTL with repulsion and coupling phase in a linkage group, but the population size was increased by three-fold (\( n = 420 \)). The results showed that large population size resulted in
separation of linked QTL in coupling phase linkage within simulated IRIL population (Figure 3.4). However, two significant QTL were detected when the linked QTL were separated by 30 cM distance. In repulsion phase linkage, large population size resulted in increased LOD scores in both populations but the IRIL population had higher LOD score values than the SRIL population (Figure 3.4). We detected two QTL in the SRIL population when the linked QTL were separated by 30 cM. In the IRIL population, two separate QTL were detected when the linked QTL were separated by 20 and 30 cM distances.
3.4 Discussion

We report here the effectiveness of an IRIL population for linkage map expansion and QTL mapping compared to a SRIL population. The simulation study provided the opportunity to examine the issues in replications of multiple populations. This study highlights a number of advantages of utilizing IRIL populations for QTL mapping.

The genetic resolution of linkage maps can be improved significantly by providing additional opportunities for recombination through intermating during the development of mapping populations. In the simulation study, the recombination frequencies increased by 1.66-fold in the IRIL population compared to the SRIL populations. In the empirical study, the recombination frequencies increased by 1.23-fold which represent the accumulated recombination frequencies in the IRIL population as compared to the SRIL population. This corresponds to a 1.41-fold increase in map length in the IRIL population as compared to the SRIL population. The increases in map length are due to the use of map function (Haldane’s mapping function) to convert the recombination frequencies to map distance. Map expansion in an IRIL population was also reported in previous studies (Lee et al., 2002). Lee et al. (2002) randomly intermated the cross of maize inbreds B73 and Mo17 for five generations after the F2, followed by five generations of self-fertilization to produce a mapping population and compared it with the F2 mapping population. These authors reported an almost four-fold increase in map distance in the IRIL population as compared to F2 population. In our study, we used three
generations of intermating before selfing in order to develop the IRIL population. The reason for the small expansion of the linkage maps in the IRIL population compared to the SRIL population may be due to the fewer generations of intermating after the F$_2$ to develop the intermated RIL population.

Disturbing influences may cause a deviation from the expected segregation ratio of parental alleles within the RIL populations. In this study, the two populations were developed from the same two parental inbreds and there was no intentional selection during the development of recombinant inbred lines. We observed segregation distortion for 20% and 27% of the loci in the SRIL and the IRIL populations, respectively. Distortion of marker frequency is often more frequent in IRIL populations than in SRIL populations. Sharapova et al. (2002) observed 43% significantly distorted loci among 983 loci mapped in an intermated RIL maize population. Lee at al. (2002) and Lu et al. (2002) also found a high frequency of segregation distortion in intermated maize populations. In one study, only 12% of the loci showed significant distortion in RIL populations (Causse et al., 1996). Lu et al. (2002) reported that segregation distortion accumulated with additional generations of meiosis.

Precision of QTL mapping can be measured by the reduction of the QTL support interval (Darvasi and Soller, 1995). We found a significant gain in the precision of QTL detected in the IRIL population compared to the SRIL population. The QTL support interval was reduced 1.66-fold in the simulated IRIL population as compared with the simulated SRIL population. Whereas in the empirical study, the average QTL support interval was reduced by 1.30-fold in
the IRIL population compared with the SRIL population. The lower support interval reduction in the empirical study may be due to the random nature of recombination during the development of simulation populations. Balint Kurti et al. (2008) compared the IBM population with the RIL population for disease resistance QTL and reported improved precision of QTL mapping in the IBM population with up to 50-fold smaller support intervals. The authors conducted the phenotypic evaluations in separate environments for each population and population sizes were variable (158 individuals for RILs and 258 for intermated RILs). In addition, large effect QTL were segregating in their mapping population.

Intermating helps to separate linked QTL. In the simulation study, with an arbitrary threshold LOD score of 3, two QTL linked in coupling phase linkages were detected as a single significant QTL in both the SRIL and IRIL populations. When the two QTL are linked in repulsion phase, no significant QTL were detected in the simulated SRIL and IRIL populations, even at the maximum linkage distance tested of 30 cM. Population size is crucial for QTL detection (Beavis, 1994; Li et al., 2010; Melchinger et al., 1998) and in our simulation study we used 140 lines in each population type. When, we increased the population size by three-fold (n=420) in both populations, two separate QTL were detected significantly. In coupling phase, linked QTL were significantly separated in the IRIL population, when the QTL were placed at 30 cM. In repulsion phase linkages, linked QTL were significantly separated in both SRIL and IRIL populations.
In empirical study, QTL for LN and DS were common between the SRIL and IRIL populations. We attribute this result to the relatively small population sizes (n=136 and 135) of the SRIL and IRIL populations. In previous simulation study with 100 individuals with 10 segregating, unlinked QTL with equal, additive effects that explained 63% of the variability only 12 percent of the simulated QTL were correctly identified (Beavis, 1994).

In summary, although developing IRIL populations may take significant time, this study emphasizes the advantages of intermated RIL populations to precisely map QTL. The empirical IRIL population has increased map length as compared to the SRIL population. Furthermore, our simulation study confirms that the mapping population size is crucial for detection of multiple QTL in linkage, especially in repulsion phase linkage.

Acknowledgements

This research was supported by a grant from the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) and the Natural Sciences and Engineering Research Council (NSERC) of Canada. We would like to thank Gregory Downs for identifying SNPs. We also would like to thank Byron Good and Heather Peterson for technical assistance in conducting field experiments.
References


Figure 3.1 Comparative linkage maps and QTL mapping of selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations. The QTL for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN, number), stay green (SG, scores), plant height (PH, cm), normalized difference in red edge (NDRE), normalized difference vegetative index (NDVI) and NDRE/NDVI (ratio) are given on the right side of each linkage group with a chart. The chart shows chromosome length (cM) in x-axis and LOD values y-axis. The dash-dot vertical line along the graph represents the threshold value for each population.
### Chromosome 9

<table>
<thead>
<tr>
<th>SRIL</th>
<th>IRIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>umc1697</td>
<td>bng12a</td>
</tr>
<tr>
<td>7.0</td>
<td>7.9</td>
</tr>
<tr>
<td>37.9</td>
<td>44.4</td>
</tr>
<tr>
<td>56.2</td>
<td>59.8</td>
</tr>
<tr>
<td>78.0</td>
<td>76.5</td>
</tr>
<tr>
<td>88.1</td>
<td>85.1</td>
</tr>
<tr>
<td>95.2</td>
<td>93.7</td>
</tr>
<tr>
<td>103.3</td>
<td>103.9</td>
</tr>
<tr>
<td>114.7</td>
<td>114.3</td>
</tr>
<tr>
<td>127.3</td>
<td>122.9</td>
</tr>
<tr>
<td>157.3</td>
<td>145.2</td>
</tr>
<tr>
<td>154.5</td>
<td>ugc43</td>
</tr>
<tr>
<td>177.3</td>
<td>ugc42</td>
</tr>
<tr>
<td>188.1</td>
<td>ugc41</td>
</tr>
<tr>
<td>200.5</td>
<td>ugc40</td>
</tr>
<tr>
<td>210.8</td>
<td>ugc39</td>
</tr>
</tbody>
</table>

### Chromosome 10

<table>
<thead>
<tr>
<th>SRIL</th>
<th>IRIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>umc578</td>
<td>ugc56</td>
</tr>
<tr>
<td>6.0</td>
<td>6.0</td>
</tr>
<tr>
<td>27.3</td>
<td>27.3</td>
</tr>
<tr>
<td>32.1</td>
<td>32.2</td>
</tr>
<tr>
<td>40.8</td>
<td>40.9</td>
</tr>
<tr>
<td>47.9</td>
<td>47.8</td>
</tr>
<tr>
<td>67.3</td>
<td>67.3</td>
</tr>
<tr>
<td>82.8</td>
<td>82.8</td>
</tr>
<tr>
<td>86.0</td>
<td>86.0</td>
</tr>
<tr>
<td>122.7</td>
<td>122.7</td>
</tr>
<tr>
<td>145.2</td>
<td>145.2</td>
</tr>
<tr>
<td>154.5</td>
<td>ugc43</td>
</tr>
<tr>
<td>177.3</td>
<td>ugc42</td>
</tr>
<tr>
<td>188.1</td>
<td>ugc41</td>
</tr>
<tr>
<td>200.5</td>
<td>ugc40</td>
</tr>
<tr>
<td>210.8</td>
<td>ugc39</td>
</tr>
</tbody>
</table>

---

This page contains diagrams and tables related to Chromosomes 9 and 10, likely illustrating genetic markers or positions. The content seems to be part of a scientific or research document, possibly related to genetics or genomics.
Figure 3.2 Contours of LOD scores, averaged over 1000 simulation runs for genetic models with a single QTL accounting for 5 to 30% of the phenotypic variance, placed at the centre of a 150 cM linkage group in selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations (n=140). The horizontal dash–dot line represents LOD 3 as an arbitrary threshold.
Figure 3.3 Contours of LOD score, averaged over 1000 simulation runs for a 150 cM linkage group with two QTL, with equal effect, linked 10, 20, or 30 cM apart in coupling (A and B) and in repulsion (C and D), on a 150 cM linkage group in selfed recombinant inbred line ((SRIL) and intermated recombinant inbred line (IRIL) populations (n= 140). The dash–dot horizontal line represents the arbitrary threshold value of LOD 3.
Figure 3.4 Contours of LOD score, averaged over 1000 simulation runs for a 150 cM linkage group with two QTL, with equal effect, linked 10, 20, or 30 cM apart in coupling (A and B) and in repulsion (C and D), on a 150 cM linkage group in selfed recombinant inbred line ((SRIL) and intermated recombinant inbred line (IRIL) populations (n= 420). The dash–dot horizontal line represents the arbitrary threshold value of LOD 3.
Table 3.1 Components of variance and QTL effect sizes for genetic models used in computer simulation of a quantitative trait locus (QTL) in selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations.

<table>
<thead>
<tr>
<th>$\sigma^2_{QTL}$</th>
<th>$a$</th>
<th>$\sigma^2_{BG} + \sigma^2_E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.22</td>
<td>0.95</td>
</tr>
<tr>
<td>0.10</td>
<td>0.31</td>
<td>0.90</td>
</tr>
<tr>
<td>0.15</td>
<td>0.38</td>
<td>0.85</td>
</tr>
<tr>
<td>0.20</td>
<td>0.43</td>
<td>0.80</td>
</tr>
<tr>
<td>0.25</td>
<td>0.48</td>
<td>0.75</td>
</tr>
<tr>
<td>0.30</td>
<td>0.53</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Abbreviations: $\sigma^2_{QTL}$: variance of the effect of a simulated QTL; $a$: additive effect of the simulated QTL; $\sigma^2_{BG}$: variance of the background genetic effect (i.e., genetic variance due to loci other than the simulated QTL); $\sigma^2_E$: environmental variance.
Table 3.2 Components of variance for genetic models used in computer simulation of a quantitative trait controlled by two linked QTL at 10, 20, and 30 cM in coupling and repulsion phase on a 150 cM linkage group in the selfed recombinant inbred line (SRIL) and the intermated recombinant inbred line (IRIL) populations.

<table>
<thead>
<tr>
<th>Linkage Phase</th>
<th>Linkage distance between QTL</th>
<th>$\sigma^2_G$</th>
<th>$\sigma^2_{BG} + \sigma^2_E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coupling</td>
<td>10</td>
<td>0.32</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.29</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.26</td>
<td>0.74</td>
</tr>
<tr>
<td>Repulsion</td>
<td>10</td>
<td>0.04</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.07</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.10</td>
<td>0.90</td>
</tr>
</tbody>
</table>

$\sigma^2_G$, Genetic variance of the model; $\sigma^2_{BG}$, variance of the background genetic effect (i.e., genetic variance due to loci other than the simulated QTL); $\sigma^2_E$, environmental variance.
Table 3.3 Number of marker loci and linkage map distance (cM) in selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations of corn, the ratio of map distance estimates in SRIL and IRIL populations.

<table>
<thead>
<tr>
<th>Linkage Group</th>
<th>No. of Loci</th>
<th>Linkage map distance (cM)</th>
<th>Ratio (IRIL/SRIL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SRIL</td>
<td>IRIL</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>326.8</td>
<td>408.9</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>119.2</td>
<td>177.2</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>202.7</td>
<td>245.9</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>137.5</td>
<td>232.8</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>192.1</td>
<td>224.5</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>177.6</td>
<td>353.6</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>147.5</td>
<td>215.4</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>129.9</td>
<td>156.5</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>210.5</td>
<td>246.3</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>145.2</td>
<td>255.9</td>
</tr>
<tr>
<td>Mean</td>
<td>12</td>
<td>178.9</td>
<td>251.7</td>
</tr>
<tr>
<td>Sum</td>
<td>120</td>
<td>1788.7</td>
<td>2516.8</td>
</tr>
</tbody>
</table>
Table 3.4 QTL detected on different linkage groups (LG) with the multiple QTL mapping approach for days to anthesis (DA), Days to silking (DS), leaf number (LN), plant height (PH), stay green (SG), normalized difference vegetative index (NDVI), normalized difference red edge (NDRE) and NDRE/NDVI (Ratio) in selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations.

<table>
<thead>
<tr>
<th>Trait</th>
<th>SRIL population</th>
<th>IRIL population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LG^a</td>
<td>Marker interval^b</td>
</tr>
<tr>
<td>DA</td>
<td>2</td>
<td>ug8 – umc1065</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>umc1023 – ug22</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>umc1524 – bngl118</td>
</tr>
<tr>
<td>DS</td>
<td>1</td>
<td>umc1917 – umc1076</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>ug7 – ug8</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>umc1807 – ug36</td>
</tr>
<tr>
<td>LN</td>
<td>7</td>
<td>ug33 – umc1036</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>umc1807 – ug36</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>umc1917 – umc1076</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>umc1076 - umc1028</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>ug2 – umc1035</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>mmc041 – ug3</td>
</tr>
<tr>
<td></td>
<td>Markers</td>
<td>LOD</td>
</tr>
<tr>
<td>---</td>
<td>---------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>2</td>
<td>mmc0271 – umc1042</td>
<td>94.37</td>
</tr>
<tr>
<td>8</td>
<td>umc1913 – umc1807</td>
<td>29.09</td>
</tr>
</tbody>
</table>

a linkage group in maize linkage map  
b marker interval of QTL peak LOD value  
c QTL location on the chromosome was estimated by MpaQTL  
d additive effect in reference to CG60 inbred
Table 3.5  QTL support intervals, averaged over 1000 simulation runs (mean) and respective lower and upper confidence limits, for a 150 cM linkage group with one and two QTL for six genetic models in selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) populations (n= 140). The two QTL were linked in 10, 20, or 30 cM apart in coupling and in repulsion (italic) on a 150 cM linkage group.

<table>
<thead>
<tr>
<th>Linkage phase position (cM)</th>
<th>One QTL</th>
<th>Two QTL</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R^2_{QTL} )</td>
<td>SRIL</td>
<td>IRIL</td>
</tr>
<tr>
<td>0.05</td>
<td>10.54</td>
<td>6.4</td>
</tr>
<tr>
<td></td>
<td>9.05</td>
<td>5.38</td>
</tr>
<tr>
<td></td>
<td>12.03</td>
<td>7.41</td>
</tr>
<tr>
<td>0.1</td>
<td>15.07</td>
<td>8.92</td>
</tr>
<tr>
<td></td>
<td>14.05</td>
<td>8.23</td>
</tr>
<tr>
<td></td>
<td>16.08</td>
<td>9.61</td>
</tr>
<tr>
<td>0.15</td>
<td>20.32</td>
<td>11.46</td>
</tr>
<tr>
<td></td>
<td>19.41</td>
<td>10.84</td>
</tr>
<tr>
<td></td>
<td>21.23</td>
<td>12.08</td>
</tr>
<tr>
<td>0.2</td>
<td>26.01</td>
<td>14.16</td>
</tr>
<tr>
<td></td>
<td>24.99</td>
<td>13.61</td>
</tr>
<tr>
<td></td>
<td>27.04</td>
<td>14.71</td>
</tr>
<tr>
<td>0.25</td>
<td>31.98</td>
<td>16.64</td>
</tr>
<tr>
<td></td>
<td>30.99</td>
<td>16.14</td>
</tr>
<tr>
<td></td>
<td>32.96</td>
<td>17.13</td>
</tr>
<tr>
<td>0.3</td>
<td>38.31</td>
<td>20.52</td>
</tr>
<tr>
<td></td>
<td>37.31</td>
<td>19.96</td>
</tr>
<tr>
<td></td>
<td>39.31</td>
<td>21.07</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Linkage phase position (cM)</th>
<th>Mean (cM)</th>
<th>Lower limit (cM)</th>
<th>Upper limit (cM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coupling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65 - 75</td>
<td>46.75</td>
<td>45.78</td>
<td>47.72</td>
</tr>
<tr>
<td>60 - 80</td>
<td>50.2</td>
<td>49.17</td>
<td>51.23</td>
</tr>
<tr>
<td>55 - 85</td>
<td>53.26</td>
<td>52.25</td>
<td>54.28</td>
</tr>
<tr>
<td>Repulsion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65 - 75</td>
<td>6.35</td>
<td>1.16</td>
<td>11.53</td>
</tr>
<tr>
<td>60 - 80</td>
<td>5.18</td>
<td>4.27</td>
<td>6.09</td>
</tr>
<tr>
<td>55 - 85</td>
<td>7.6</td>
<td>6.21</td>
<td>9</td>
</tr>
</tbody>
</table>

\( R^2_{QTL} \), Proportion of the phenotypic variance accounted for by the simulated QTL.

The two QTL models were simulated to place two QTL with same effect size 10, 20 and 30 cM apart.
Chapter 4 Early Adapted Maize Genotypes Harbor Great Genetic Diversity with Little Evidence of Repulsion Phase Linkages

Abstract

To produce adapted maize hybrids, inbred lines are often independently selected from distinct populations. We investigate the genetic basis of two elite maize inbreds’ adaptation to a short-season environment to determine if selection for inbreds in different populations leads to the development of inbred lines with large but hidden genetic differences. We developed selfed recombinant inbred lines (SRILs) and intermated recombinant inbred lines (IRILs) from two early-season adapted lines and examined trait variation within hybrids between these lines and a tester. Hybrids’ flowering time, leaf number, stay green, yield, and yield component traits were measured in multiple locations. Recombinant inbred line (RIL) hybrids were highly variable, and a large proportion of genotypes were superior to the high parent, indicating that the genetic basis for local adaptation differed between lines. We mapped 25 quantitative trait loci (QTL) for all traits in both SRIL and IRIL hybrid populations. All QTL had small to moderate effects, and only one QTL was shared between populations. Furthermore, we investigated if selection resulted in an excess of repulsion phase linkages among the adapted lines. Genetic variances and genetic correlations did not significantly differ between intermated and non-intermated populations suggesting that repulsion phase linkages play a small, if any, role in generating variation. Our results show that parallel selection regimes to adapt cultivars to a common environment can cause the fixation of genetically divergent genotypes with similar trait values. The capacity for further improvement is high and little influenced by linkage disequilibrium. Our results also suggest that a large proportion of genetic variation may be caused by higher order epistatic interactions or epigenetic phenomena.
4.1 Introduction

Major cereal crops such as maize, rice, and wheat cover a large area of the earth's arable land, with each crop grown on 159, 161, 225 million hectares (FAOSTAT, 2009), respectively. In the western hemisphere, after domestication, maize (Zea mays L.) diverged into over 250 landraces and has been grown as far south as Chile and as far north as Canada in a diverse range of elevations, and environments (Camus-Kulandaivelu et al., 2006). The Corn Belt Dent (CBD) maize landrace is the source of most of the current, commercial hybrid maize in North America, and this landrace arose from a cross between two of these landraces Southern Dent and Northern Flint (Doebley et al., 1988). Although originally found in the temperate midwestern United States, CBD genotypes have been adapted to a wide range of latitudes and are grown in diverse environments (Labate et al., 2003). Breeders have selected for high yielding CBD lines within these environments. For Northern latitudes, selection has been for early flowering because of the short growing season. Plants that initiate pollination early will have mature seed by the time of damaging cold and frost (Salvi et al., 2007). In Southern latitudes, plants that flower late yield more than plants that flower early.

In maize, new genotypes adapted for specific environments, such as early season environments, are often derived from distinct CBD germplasm groups, termed heterotic groups. Thus, a number of independently selected inbred lines may be developed that are adapted to a specific environment (Troyer, 1990). These heterotic groups, include Iowa Stiff Stalk Synthetic (BSSS), lodent, and
Lancaster Sure Crop (LSC) (Carena and Hallauer, 2009; Mikel and Dudley, 2006). Inbreds are developed from different heterotic groups because hybrid maize is generated by crossing inbred lines from different heterotic groups.

The elite inbreds selected for a specific environment are phenotypically similar but may harbor cryptic genetic variation. Indeed, previous studies of late flowering and relatively unimproved CBD lines found that these lines often had different genetic bases for traits. Within a set of 26 highly diverse genotypes, Buckler et al. (2009) showed that each genotype contained both positive and negative flowering time alleles. Khanal et al. (2011) also showed that many traits per se are highly variable. A number of flowering time and yield component QTL studies have also shown that members of the different heterotic groups, with similar traits, have different alleles.

Selection for earliness would be expected to cause the fixation of some suboptimal alleles within the resulting inbred lines. The strong artificial selection during breeding could cause hitchhiking of slightly deleterious, linked alleles (repulsion phase linkages), and the low-level of recombination within the populations may cause their fixation. In a groundbreaking study, Hill and Robertson (1966) used simulations to demonstrate that that selection at one favorable allele interferes with selection at a second, linked beneficial allele, reducing its probability of fixation. The Hill and Robertson effect has been confirmed in a number of instances, but a few studies have failed to find evidence of repulsion linkages.
Three elite and early-season genotypes from the Iodent, Stiff Stalk, and Lancaster heterotic groups are CG60, CG102, and LH295, respectively. Using recombinant inbred lines (RILs), derived from intermating from CG60 and CG102, we examined the effects of selection on genetic variation. The first objective was to determine the degree to which hybrid yield and flowering traits vary amongst hybrids between LH295 and RILs. We also sought to identify the loci polymorphic between CG60 and CG102 that contribute to this trait variation. Finally, we sought to determine if strong selection contributed to the development of genotypes with favorable alleles linked to unfavorable alleles, as predicted.
4.2 Materials and methods

4.2.1 Population development

Two recombinant inbred line (RIL) populations were developed from a pair of elite maize inbred lines, CG60 and CG102. The inbred line CG60 belongs to the Iodent heterotic group (Lee et al., 2001a) and the inbred line CG102 belongs to the Stiff Stalk heterotic group (Lee et al., 2001b). One RIL population was derived through repeated selfing of F2 individuals for four generations, and this population is called the selfed recombinant inbred line (SRIL) population. The other RIL population was intermated for three generations prior to selfing and is called the intermated recombinant inbred line (IRIL) population. Inbred RILs derived from CG60 and CG102 are highly variable for a number of traits including flowering time, stay green, and plant height (Khanal et al., 2011). Individuals from both populations (136 from SRIL and 135 from IRIL) were crossed with a tester, LH295, to develop selfed recombinant inbred line testcross (SRIL-TC) and intermated recombinant inbred line testcross (IRIL-TC). The inbred tester was also crossed with parental lines CG60 and CG102. LH295 is a member of the Lancaster Sure Crop (LSC) heterotic group, belonging to the LH82 family of lines. The testcross progenies were produced in isolation at Woodstock, Ontario, Canada in 2006.
4.2.2 Experimental design and field evaluations

Testcross individuals (136 SRIL-TC, 135 IRIL-TC, and 2 parental hybrids) and 10 check hybrid varieties were evaluated at three locations (Elora, Alma and Waterloo) in Ontario, Canada in 2007 and 2008. Location × year combinations were considered as environments. The entries were evaluated in the six environments using a randomized complete block design (RCBD) with two replications at each location. Experimental units were two-row plots with 5.79 m rows, 0.76 m between rows, and 0.91 m between ranges. Plots were planted early- to mid-May with an ALMACO cone-planter (Allan Machine Company, Nevada, IA). Each plot was overplanted and uniformly thinned by hand to 66 plants per plot around the sixth leaf-tip stage resulting in a 75,000 plants per hectare. Weeds were controlled as necessary with conventional herbicide applications. Plots were harvested in October or November with a New Holland split-plot combine (Allan Machine Company, Nevada, IA). Machine-harvestable grain yield (adjusted to 15% grain moisture) was measured at harvest using a HM2200 HarvestMaster high capacity dual GrainGage (Juniper Systems Inc., Logan, UT). Data were collected on a single plot basis for the following traits: (1) days to anthesis (DA, days); (2) days to silking (DS, days); (3) leaf number (LN); (4) stay green (SG, visual score); (5) ear length (EL, cm); (6) kernel row number (KRN); (7) 1000 kernel weight (KW, g); (8) grain moisture (GM, %) at harvest; (9) grain yield (GY, Mg ha⁻¹) adjusted to 15.5% moisture. DA and DS were recorded when 50% of the plants in the plot were shedding pollen or had visible silks, respectively. LN was measured on three competitive plants (i.e., plants bordered
by adjacent plants) from each plot. For this, six weeks after planting the leaves were counted, and the seventh leaf was painted with enamel paint. Five weeks after pollination the total number of leaves was recorded. SG was recorded on a single date in each location after all genotypes had reached physiological maturity (i.e., black layer formation). SG was scored using a visual rating of one to nine for an individual plot, where one corresponds to complete senescence of all plants in the plot, and nine corresponds to a plot of completely green plants. A sample of three ears was collected randomly from each plot for EL, KRN and KW. The ears were dried for 48 hours prior to measuring traits. For EL, the three ears were measured with a ruler and the plot mean EL was calculated. Similarly, kernel row number was counted from three ears, and the plot mean KRN was calculated. For KW, three ears were shelled together, and 1000 seeds were counted and weighed to measure the plot value.

4.2.3 Phenotypic data analysis

Data were analyzed using the PROC MIXED procedure in SAS v9.2 (SAS Institute, 2008) with environment, replication, genotype and the interaction of genotype with environment considered as random using the following model:

\[ Y_{ir} = \mu + \alpha_i + \beta_e + (\alpha\beta)_i + \gamma_{r(e)} + \varepsilon_{ir} \]

where \( \mu \) = overall mean; \( \alpha_i \) = effect of genotype \( i \); \( \beta_e \) = effect of environment \( e \); \( (\alpha\beta)_i \) = effect of an interaction between genotype \( i \) and environment \( e \); \( \gamma_{r(e)} \) = effect of replication \( r \) nested within environment \( e \); and \( \varepsilon_{ir} \) = residual error of plot containing genotypes \( i \) in replication \( r \) and environment \( e \). Variance
components and their 95% confidence intervals were estimated from PROC MIXED in SAS v9.2 (SAS Institute, 2008). Best linear unbiased predictions (BLUPs, Littell et al., 2006) of genotypes within the SRIL-TC and IRIL-TC populations were estimated for all traits across environments by PROC MIXED in SAS v9.2 (SAS Institute, 2008).

Least-square mean estimates of all traits of the testcross populations and parental hybrids across six environments were computed using the LSMEANS statement in SAS v9.2. PROC MEANS in SAS was used to calculate the ranges and standard errors for each trait in both testcross populations. The parental line effects and the testcross populations’ effects were considered as fixed.

Broad sense heritabilities ($H^2$) across environments for each trait and their corresponding standard errors were calculated on testcross mean basis in SAS v9.2 from the equation as follows (Holland et al., 2003):

$$H^2 = \frac{\sigma^2_g}{\sigma^2_g + \left(\frac{\sigma^2_{ge}}{e}\right) + \left(\frac{\sigma^2}{re}\right)}$$

where $H^2$ = broad-sense heritability; $\sigma^2_g$ = total genetic variance; $\sigma^2_{ge}$ = genotype by environment variance; $\sigma^2$ = error variance; $r$ = number of replications, and $e$ = number of environments.

The genotypic and phenotypic correlations among pairs of traits and their corresponding standard errors were computed across environments using multivariate restricted maximum likelihood (REML) estimation in PROC MIXED in SAS v9.2 with the ASYCOV option to calculate genetic variance and covariance between traits (Holland, 2006). The output was then used to calculate genetic
correlations and their standard errors using the procedure of Mode and Robinson (1959). Confidence intervals for correlation coefficients were constructed as $r \pm 1.96 \times \text{standard error (SE)}$, where $r$ is the correlation coefficient, 1.96 is the (alpha=0.05 two-tailed) critical value from the standard normal distribution and SE is the standard error of the correlation coefficient. Estimated correlation coefficients were regarded as significantly different from zero if their 95% confidence intervals did not include zero. If the confidence intervals for correlation coefficients between the two populations did not overlap, differences between correlation coefficients were declared significant.

4.2.4 Quantitative trait locus mapping

Linkage maps comprising 71 simple sequence repeats (SSR) and 49 SNP loci were constructed using the software package Joinmap 4.0 (Van Ooijen, 2006) using the SRIL and IRIL population genotypic data. The markers were divided into linkage groups using the “group” command (parameter value LOD > 5). The remaining markers were added to the respective linkage groups, based on previous mapping information for these markers (Lawrence et al., 2004), via the “assign” command (parameter value LOD > 5). We removed markers that caused large increases in genetic map distances. The final linkage maps had a total length of 1788.74 and 2526.79 centimorgans (cM) and an average interval length of 16.26 and 23.04 cM in the SRIL and IRIL populations, respectively.

For each trait in both testcross populations, a QTL analysis was performed with the MapQTL6 software package (Van Ooijen, 2004). Initially, the
conventional interval mapping method (Lander and Botstein, 1989) was performed to identify possible QTL. Second, the multiple QTL mapping (MQM) method (Jansen and Stam, 1994) was used. Markers were chosen using the command “automatic cofactor selection”. In this method, background markers were selected to take over the role of the putative QTL as cofactors to reduce the residual variance. With this MQM method, a one-dimensional search over the genome is performed by testing for segregating QTL as in interval mapping while simultaneously fitting the selected cofactors in the model (Jansen and Stam, 1994). The genome-wide significance logarithms of odds (LOD) thresholds for QTL were obtained by performing a permutation test in MapQTL6 with a set of 1,000 iterations. Permutation testing randomly assigns the observed trait values to the individuals in the mapping population.

The support interval of each QTL was estimated from the interval mapping LOD curves. The 2-LOD fall method was used to estimate the 95% support interval (Van Ooijen, 1992), which represents the map interval in which the LOD score is within one unit of its maximum. Graphs of the linkage maps and QTLs were prepared using MapChart (Voorrips, 2002).

Epistatic interactions and the variance explained by the interactions ($R^2$) were calculated by the EPISTACY 2.0 macro in SAS v9.2 (Holland, 1998). We used a critical $P$-value filter of 0.0001. This threshold was determined because 45 independent combinations exist among the ten linkage groups of maize. A comparison-wise error rate of $10^{-4}$ corresponds approximately to an experiment-wise error rate of 0.01. All interactions in the Epistacy output in which two
markers were within 50 cM of each other were removed to avoid linkage effects (Holland, 1998).
4.3 Results

4.3.1 Hybrids of RILs generated from two early season genotypes from two heterotic groups vary greatly

The inbred lines CG60 and CG102 are two early flowering Iodent and Stiff Stalk inbred lines that were developed for hybrid maize production in short season environments. Hybrids between the LSC tester inbred LH295 and CG60 and LH295 and CG102 had similar trait values (Table 4.1). The means of CG60 x LH295 and CG102 x LH295 hybrids did not significantly differ for DA, KRN, GM, GY, and LN. The hybrids differed significantly ($P < 0.05$) for DS, EL, SG, and KW (Table 4.1). However, the differences were small for all but SG in real terms. For example, the CG60 x LH295 hybrid extrudes silks (DS) 1.06 days earlier than the CG102 x LH295 hybrid. The CG60 x LH295 hybrid had 36.02 g higher KW than the CG102 x LH295 hybrid. EL was 1.19 cm greater within the CG60 x LH295 hybrid than the CG102 x LH295 hybrid. The SG score of 4.17 the CG102 x LH295 was substantially larger than the mean (2.67) of the CG60 x LH295 hybrid.

The SRIL-TC and IRIL-TC population means did not significantly differ for LN, EL, KW, KRN and GY (Table 4.1). The RIL x LH295 testcross means were also similar to parental means. GM ($P \leq 0.05$), DA, DS, and SG ($P \leq 0.01$) significantly differed between the two testcross populations, but the differences were small: 0.29 %, 0.33 days, 0.38 days, and 0.44 units, respectively (Table 4.1). The SRIL-TC population mean was higher for DA, DS, SG and GM.
Despite the similarity of parental hybrids, RIL testcross hybrids were highly variable. In both testcross populations, DA and DS ranged from 67 to 87 days (Table 4.1). For example, over 20% of the SRIL-TC population had days to anthesis and silking dates that deviated by more than 10 days from the population mean (Figure 4.1). GM ranged from 14 to 32 %; LN ranged from 12 to 18; and SG differed by as much as eight visual units (1 to 9). Yield and yield components also varied greatly. EL ranged from 10 to 20 cm; KRN ranged from 12 to 24 rows; KW ranged from 153 to 555 g; and GY ranged from 6 to 15 Mg ha\(^{-1}\) (Table 4.1). Ranges within each location were similar to ranges across all locations, and these differences were largely due to genetic effects. Genotypic variance components were highly significant (\(P \leq 0.001\)) for all traits within each testcross population (Table 4.2). The genotype × environment variance was significant for GY (\(P \leq 0.05\)), DS, SG and GM (\(P \leq 0.001\)) in the SRIL testcross population and was significant for GY (\(P \leq 0.01\)), SG, EL, GM (\(P \leq 0.001\)) in the IRIL testcross population (Table 4.2). The estimates of genetic variance were approximately two to eight times greater than the genotype × environment variances across all traits (Table 3.2). Broad-sense heritability (\(H^2\)) values for the traits ranged from 0.60 to 0.90. The heritability values were high (greater than 0.70) for DA, DS, SG, EL, KRN, and GM, whereas GY, KW and LN had moderately high heritability (between 0.60 and 0.70) in both testcross populations (Table 4.1).

Trials were run with commercial checks as controls, hybrids developed and released by commercial breeding programs. These hybrid checks had a mean
grain yield of 11.80 Mg ha\(^{-1}\), and 12\% of the RIL hybrids exceeded the hybrid checks' mean grain yields.

4.3.2 IRIL-TC and SRIL-TC trait variances are similar

As described above, the parental inbred lines are products of strong selection within small populations. Thus, we investigated if RILs derived from intermating were more variable than RILs derived from selfing the F\(_2\) because of greater recombination between linked alleles. Hybrid traits had extremely similar frequency distributions within the SRIL-TC and IRIL-TC populations (Figure 4.1), and the genetic variance for all traits did not significantly differ between two testcross populations.

We similarly wondered if selection for one beneficial trait, such as high yield, could have resulted in the fixation of a linked allele with a negative effect on another trait, such as a late flowering time. Within both populations, genetic correlations of most traits were as expected. Plants that flowered late tended to have a high LN (r = 0.52), SG (r = 0.44), and high GM (r = 0.62). Genetic correlations between LN and SG were (r = 0.36) and (r = 0.26), respectively. Grain yields may be high in late flowering lines in early-season environments during mild summers such as those in this study. Grain yield had positive genotypic correlations with DA (r = 0.34), DS (r = 0.27), EL (r = 0.24) and KW (r = 0.24). The negative genotypic correlations of KRN with KW (r = -0.66) and EL (r = -0.20) suggested a trade-off (Table 4.3). Nonetheless, not one correlation coefficient significantly differed between the IRIL-TC and SRIL-TC populations.
4.3.3 QTL analyses revealed small effect QTL that were largely unique to one population.

Because of the large number of testcross hybrids with trait values that differed from the population mean, we speculated that CG60 and CG102 differed for a small number of large effect loci that contributed to the hybrid traits. Such a trait distribution would be unlikely if only small effect, additive QTL were present. For example, if CG60 and CG102 had ten different, unlinked favorable alleles each with a small effect (e.g. 0.5 days, as suggested in Buckler et al., 2009) that contributed to early flowering, the probability of a single plant with all favorable alleles would be 0.5\(^{10}\).

Multiple QTL mapping within the SRIL population identified a total of 16 QTL for DA, DS, LN, SG, KRN, KW, GM, and GY (Figure 4.2, Table 4.4). No QTL were found for EL. All QTL explained a small proportion of phenotypic variance within the SRIL and had small additive effects. KRN QTL explained 7.9 to 9.2% of the phenotypic variance and KW QTL explained 9.2 to 10.5% of the phenotypic variance. Each flowering time QTL affected DA and DS by less than 0.5 day. The QTL overlapped for DA, DS and GY on chromosome 1, DS, SG, GM and KW on chromosome 2, KW and GY on chromosome 4, and DA, DS, and SG on chromosome 8, suggesting a common genetic basis for these traits. The largest effect QTL for GM, explaining 17.6% of the phenotypic variance, was found on chromosome 2. The CG60 allele caused a 0.5% reduction in GM. The QTL with the smallest R\(^2\) value was for GY on linkage group 4 (R\(^2\) = 5.6%). The CG60 allele had a positive, additive effect of 120 kg/ hectare. The average support
interval, of all detected QTL was 33 cM (data not shown) in the SRIL-TC population. LOD thresholds for genome-wide significance ranged from 2.5 to 2.6 ($\alpha = 0.10$).

In the IRIL-TC population, nine QTL were identified for EL, KRN, KW, and GM (Figure 4.2, Table 4.4). No QTL was identified for SG or flowering time traits, e.g. DA, DS, and LN. All IRIL-TC QTL also explained a small proportion of the phenotypic variation. QTL support intervals overlap for KRN and KW on chromosome 5, GM and KW on chromosome 7, and EL and GM on chromosome 9. The GM QTL explained the highest proportion of variance (Table 4.4). The QTL for GM on chromosome 9 explained the lowest proportion of variance. QTL support intervals averaged 20 cM (data not shown). The LOD thresholds for genome-wide significance of QTL at a type I error rate of $\alpha = 0.10$ ranged between 2.6 to 2.8.

QTL analyses indicated that multiple intervals of morphological QTL overlapped with yield QTL (Figure 4.2). For example, QTL for DA, DS, and GY were detected from same region on chromosome 1 in SRIL-TC population. In IRIL-TC population, QTL for KW and KRN on chromosome 5, QTL for GM and KW on chromosome 7 and QTL for EL and GM on chromosome 9 were detected from same region.

A single QTL for GM on chromosome 2 was shared between the SRIL-TC and IRIL-TC populations (Figure 4.2). There was little evidence that significant QTL in one population had a corresponding QTL in the second population that were close to the significance (Supplemental Figure 4.1). Within the IRIL-TC
population, QTL with LOD score values above one overlapped ten of the 15 unique, significant QTL in the SRIL-TC population. Similarly, three of the QTL within the SRIL-TC population corresponded to three of the nine IRIL-TC QTL.

Epistasis could explain the large number of genotypes with extreme trait values within the IRIL-TC and SRIL-TC populations. The analysis of epistasis conducted with EPISTACY (Holland, 1998) lead to detection of 16 and 18 pairs of epistatic QTL for all measured traits in the SRIL-TC and IRIL-TC populations, respectively (Supplemental Table 4.1). The epistatic interactions accounted for higher proportions of phenotypic variances than did individual loci, but the proportions of phenotypic variance were nonetheless still moderate. The proportion of phenotypic variation explained by the epistatic interaction (R^2) ranged between 11.30 to 22.05%, with the highest value corresponding to an interaction between two KRN QTL. One of the loci for GY (umc1076) was significant as a main effect.
4.4 Discussion

4.4.1 RIL populations exhibited great genetic variation for all traits

Hybrids involving the LSC inbred line LH295 and CG60 and CG102, the Iodent and Stiff Stalk, inbred lines, respectively, exhibited very similar trait values for DA, DS, EL, GM, GY, KRN, LN, SG and KW. Nonetheless, hybrids between the CG60 × CG102 derived RILs and LH295, exhibited substantial phenotypic variation for all traits, suggesting that the similar CG60 and CG102 phenotypes come from different sets of positively and negatively acting loci. The parental hybrids’ grain yield differed by 0.12 Mg ha\(^{-1}\), and yet the ranges for grain yield were 5.54 to 14.48 and 5.93 to 14.90 Mg ha\(^{-1}\) for the SRIL-TC and IRIL-TC populations, respectively. Transgressive segregation is common for both grain yield and flowering related traits, but this variability from two phenotypically similar parents is exceptional. For example, GY tends to be a highly transgressive trait. Austin et al. (2001) compared the hybrid progeny of F\(_{2:3}\) and F\(_{6:8}\) lines from a Mo17×H99 cross and reported that the mean parental difference for grain yield was 0.42 Mg ha\(^{-1}\), and the range of testcross yields was 5.64 – 8.58 Mg ha\(^{-1}\). Both Mo17 and H99 are classified as members of the LSC heterotic group (Melchinger et al., 1991). Similarly, Lamkey et al. (1995) compared maize testcross progeny in F\(_2\) population and eight generations of intermated F\(_2\) population developed from B73 × B84 crossed with Mo17. They reported that parental hybrid grain yield difference was 0.95 Mg ha\(^{-1}\). Approximately 95% of hybrids had values within a range of 1.2 Mg ha\(^{-1}\). Huang et
al. (2010) reported that the range of grain yield in testcross hybrids developed from European flint (F2) and US dent (F252), crossed to another dent (MBS847) was 2.84 Mg ha$^{-1}$ within an F$_3$ population and 2.22 Mg ha$^{-1}$ within an intermated population, whereas the parental difference was 0.70 Mg ha$^{-1}$. The high variation in grain yield observed in our study was not caused by poor environments, as the average grain yields of commercial check varieties were high.

The variability of flowering time traits was also exceptional. For example, the days to anthesis and silking, DA and DS, differed by one day between the parental hybrids, but the range for flowering time was 20 days within the testcross RIL populations (Table 4.1). Huang et al. (2010) reported that the range for days to silking in a population developed from crossing two inbred lines of maize was 8.42 days while parental difference for days to silking was 2.93 days. These results indicate a large number of hidden genetic differences between CG60 and CG102.

The RIL hybrids also had unusually high means relative to parental hybrids. In our study, the mean testcross population grain yield was 10.54 Mg ha$^{-1}$. The high parent grain yield was 10.70 Mg ha$^{-1}$, with 16% of the RILs exceeding the high parent trait value. It is rare to have a large number of RILs generated from elite × elite crosses that significantly outperform the best parent in testcrosses (Austin et al., 2001; Lamkey et al., 1995). Singh et al. (2011) used hybrids developed from RILs of CG60 and CG108 (both from the Iodent heterotic group) and crossed to an inbred CG102 from Stiff Stalk Sythentic heterotic group. All or the large majority of testcross individuals had lower GY than the high parent across environments.
Because CG60 and CG102 were developed separately from different heterotic groups, we conclude that the genotypes have different positive alleles that affect traits.

As described above, selection for positive alleles in small populations can lead to the fixation of linked deleterious alleles (Hill and Robertson, 1966). If this is the case in CG60 and CG102, then the genetic variance should increase and correlations among selected traits should increase with intermating. Not a single genetic variance significantly differed between the SRIL-TC and IRIL-TC testcross populations. Thus, we conclude that disruption of linkage blocks with positive and deleterious alleles, should they exist, contributes little to expanding functional genetic variation. Empirically, studies have often reported lower genetic variances or unchanged genetic variances within intermated populations compared to non-intermated populations for yield and flowering traits (Huang et al., 2010; Lamkey et al., 1995). For example, Lamkey et al. (1995) found a large but non-significant difference in GY between intermated populations and the F$_2$ populations. The Illinois high oil and Illinois low oil as well as the Illinois high protein and the Illinois low protein populations have undergone divergent selection over 70 generations since 1896, resulting in phenotypes that differ by four times higher oil and protein content than the starting population. Intermating significantly reduced genetic variance for oil content in the cross of Illinois High Oil × Illinois Low Oil (Moreno-Gonzalez et al., 1975), and for protein content in the cross of Illinois High Protein × Illinois Low Protein (Dudley et al., 2004). These results suggest that coupling phase linkages contribute to these traits. It is
difficult to reconcile these results with a number of studies of populations that have undergone both natural and artificial selection. These studies have shown that the efficacy of selection is limited by linkage (e.g. Betancourt and Presgraves, 2002).

These results suggest that repulsion phase linkages are not frequent within these lines. Hanson and Hayman (1963) performed simulations assuming that the location of a favorable or unfavorable allele at a position within a progenitor genome is random. For parents with equal frequencies of favorable and unfavorable alleles, variances remain almost constant with intermating. Alternatively, given that segregating loci have a small effect on traits, the effect of recombination of linked, repulsion phase loci would be masked by other segregating loci. In the present study, for eight of the nine traits (DA, DS, SG, EL, KRN, KW, GM, GM), the intermated population had nominally lower genotypic variances. Large population sizes may be necessary to detect these recombinants.

4.4.2 QTL within the RIL testcross populations

The percentages of phenotypic variance explained by the QTL were small for all traits in both testcross populations ranging from 5.6 to 17.6%. For all traits, the sum of all positive QTL effects was less than the observed values. Likewise the predicted trait values of a genotype with all negative QTL alleles were much higher than observed genotypes with low trait values. Digenic epistasis between pairs of unlinked markers explained a higher proportion of the variance, but the
combination of epistatic and additive effects also did not account for observed trait values. QTL that contribute to yield and flowering time traits in a number of maize populations typically have effects of a similar magnitude (Frascaroli et al., 2009; Frascaroli et al., 2007; Huang et al., 2010). Although one could ascribe the unexplained variance to small effect QTL that are below the level of detection, as described above, the probability of such small effect of the QTL causing the observed trait distributions is very low. One possibility is that higher order epistatic interactions generated extreme yield and flowering time traits.

A number of the detected QTL mapped to previously identified loci, suggesting that the same loci contribute to genetic differences between CG60 and CG102 as in other genotypes. The QTL identified on chromosome 8 between markers ug37 and umc1309 maps Vegetative to generative transition 1 (Vgt1), a major QTL involved in floral transition (Vladutu et al., 1999), suggesting that it is polymorphic between CG60 and CG102, despite its role in the adaptation of short season maize (Ducrocq et al., 2008). We identified a QTL for DS on chromosome 1S in the same region as a QTL identified for days to silking in a cross between the late flowering lines B73 and NC264 (Koester et al., 1993).

Only one QTL for GM was common between the SRIL-TC and IRIL-TC populations. We attribute this result to the relatively small population sizes (n=136 and 135) of the SRIL and IRIL populations. Beavis (1994) simulated populations of 100 individuals with 10 segregating, unlinked QTL with equal, additive effects that explained 63% of the variability. Twelve percent of the simulated QTL were correctly identified using interval mapping. Empirically,
Huang et al. (2010) used populations of 300 and 322 lines from intermated and non-intermated origin, respectively. Three QTL detected in the intermated population had overlapping confidence intervals with 9 GY QTL found within the non-intermated population. Two QTL detected in the intermated population had overlapping confidence intervals with 9 QTL for GM, and 3 of 12 DS QTL were found. Similarly, Austin et al. (2000) compared testcrosses between F2:3 and F6:8 populations from Mo17 x H99 with three testers and identified 5 shared QTL among 58 detected.

It is unlikely that QTL differed between the SRIL-TC and IRIL-TC populations because of the break-up of repulsion phase linkages, thereby generating new IRIL-TC QTL, or because of the break-up of coupling phase linkages, thereby losing the SRIL-TC QTL. In simulations, two linked loci in repulsion phase (30 cM between the QTL and population size of 100), each accounting for 10% of the phenotypic variation, are difficult to detect (Li et al., 2010). Other segregating loci would increase this difficulty (Huang et al., 2010). Similarly, coupling phase linkages are difficult to resolve with small to moderate population sizes.

This study shows that genotypes that have undergone strong selection for productive growth in environments far from their ancestral germ pool can harbor substantial cryptic genetic variation. Both genetic means and variances of the testcross progeny were high. The genetic basis for this variation remains an enigma. The magnitudes of additive and digenic QTL effects were small for all traits, and it is unlikely that small effect QTL could account for the distribution of traits within the RIL testcross populations. Further experiments with different
testers and different inbred lines will clarify the applicability of our findings across different testers. Our results suggest that elite maize has ample genetic diversity for adaptation far from its ancestral home.

Acknowledgements

This research was supported by a grant from the Ontario Research Fund and the Natural Sciences and Engineering Research Council (NSERC) of Canada. We would like to thank Byron Good and Heather Peterson for technical assistance to conduct field experiments.
References


Figure 4.1 Frequency distributions of days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha\(^{-1}\)) traits of Selfed recombinant inbred line testcross (SRIL-TC) and intermated recombinant inbred line testcross (IRIL-TC) populations of maize grown in six environments in Ontario, Canada during 2007 and 2008. The parental hybrids (CG60 and CG102) are shown by black arrows and testcross population means are shown by green arrows.

SRIL testcross population

IRIL testcross population
SRIL testcross population

IRIL testcross population

![Graphs showing frequency distribution of ear length and kernel row number in SRIL and IRIL testcross populations.](image)

Ear length (cm)

Kernel row number (rows)
Figure 4.2 Maize genetic linkage maps of selfed recombinant inbred line testcross (SRIL-TC) and intermated recombinant inbred line testcross (IRIL-TC) with QTL for different traits. QTL for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha⁻¹) shown by a bar next to the chromosome. Graph shows the chromosome length (cM) on the vertical axis and LOD value on the horizontal axis. The dash–dot vertical line along the graph represents the threshold value for each population.
Supplemental Figure 4.1  Maize genetic linkage maps constructed in this study. Distances between markers are shown on the left of each linkage group. Putative QTL for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha\(^{-1}\)) shown by a bar next to the marker. Graph shows the chromosome length (cM) on the vertical axis and LOD value on the horizontal axis. The dash–dot vertical line along the graph represents the threshold value for each population.
Table 4.1 Mean trait values for two parental hybrids and means, ranges and heritabilities (H2) across six environments for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, % ), and grain yield (GY, Mg ha⁻¹) of the testcrosses involving the selfed recombinant inbred line testcross (SRIL-TC) and intermated recombinant inbred line testcross (IRIL-TC) populations. Standard errors (s.e.) of the means are in parentheses.

<table>
<thead>
<tr>
<th>Trait</th>
<th>CG60 hybrid</th>
<th>CG102 hybrid</th>
<th>SRIL-TC</th>
<th>IRILTC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean (s.e.)</td>
<td>Range</td>
<td>H²</td>
<td>mean (s.e.)</td>
</tr>
<tr>
<td>DA</td>
<td>75.11 (0.37)</td>
<td>76.17 (0.37)</td>
<td>76.22 (0.121)a</td>
<td>67 - 87</td>
</tr>
<tr>
<td>DS</td>
<td>75.50 (0.24)a</td>
<td>76.50 (0.24)b</td>
<td>76.86 (0.116)a</td>
<td>67 - 87</td>
</tr>
<tr>
<td>LN</td>
<td>16.50 (0.33)</td>
<td>15.90 (0.33)</td>
<td>16.06 (0.019)</td>
<td>13.33 - 18.00</td>
</tr>
<tr>
<td>SG</td>
<td>2.67 (0.17)a</td>
<td>4.17 (0.17)b</td>
<td>3.58 (0.039)a</td>
<td>1 - 9</td>
</tr>
<tr>
<td>EL</td>
<td>15.76 (0.24)a</td>
<td>14.57 (0.24)b</td>
<td>15.27 (0.031)</td>
<td>10.43 - 19.50</td>
</tr>
<tr>
<td>KRN</td>
<td>16.64 (0.11)</td>
<td>16.83 (0.11)</td>
<td>17.39 (0.035)</td>
<td>13.33 - 22.00</td>
</tr>
<tr>
<td>KW</td>
<td>295.89 (8.60)a</td>
<td>259.87 (8.60)b</td>
<td>267.99 (1.01)</td>
<td>161.24 - 555.47</td>
</tr>
<tr>
<td>GM</td>
<td>20.53 (0.26)</td>
<td>21.24 (0.26)</td>
<td>21.55 (0.081)a</td>
<td>13.70 - 31.60</td>
</tr>
<tr>
<td>GY</td>
<td>10.58 (0.30)</td>
<td>10.70 (0.30)</td>
<td>10.54 (0.035)</td>
<td>5.54 - 14.48</td>
</tr>
</tbody>
</table>

a Mean values with a different letter are significantly different
Table 4.2 Variance component estimates for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, units), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha⁻¹) for selfed recombinant inbred line testcross (SRIL-TC) and intermated recombinant inbred line testcross (IRIL-TC) populations.

<table>
<thead>
<tr>
<th>Trait</th>
<th>population</th>
<th>$V^a_g$</th>
<th>confidence interval</th>
<th>$V^{b}_{ge}$</th>
<th>$V^{c}<em>{gSRIL}/V^{c}</em>{gIRIL}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DA</td>
<td>SRIL-TC</td>
<td>0.61***</td>
<td>(0.42; 0.81)</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.35***</td>
<td>(0.22; 0.48)</td>
<td>0.04</td>
<td>1.74</td>
</tr>
<tr>
<td>DS</td>
<td>SRIL-TC</td>
<td>0.95***</td>
<td>(0.67; 1.24)</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.54***</td>
<td>(0.36; 0.73)</td>
<td>0.22**</td>
<td>1.76</td>
</tr>
<tr>
<td>LN</td>
<td>SRIL-TC</td>
<td>0.04***</td>
<td>(0.02; 0.05)</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.05***</td>
<td>(0.03; 0.06)</td>
<td>0.02</td>
<td>0.80</td>
</tr>
<tr>
<td>SG</td>
<td>SRIL-TC</td>
<td>0.39***</td>
<td>(0.27; 0.50)</td>
<td>0.15***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.27***</td>
<td>(0.18; 0.35)</td>
<td>0.18***</td>
<td>1.44</td>
</tr>
<tr>
<td>EL</td>
<td>SRIL-TC</td>
<td>0.41***</td>
<td>(0.28; 0.53)</td>
<td>0.10**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.29***</td>
<td>(0.19; 0.38)</td>
<td>0.06</td>
<td>1.41</td>
</tr>
<tr>
<td>KRN</td>
<td>SRIL-TC</td>
<td>0.61***</td>
<td>(0.44; 0.78)</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.42***</td>
<td>(0.29; 0.55)</td>
<td>0.09</td>
<td>1.45</td>
</tr>
<tr>
<td>KW</td>
<td>SRIL-TC</td>
<td>168.94***</td>
<td>(107.72; 230.17)</td>
<td>60.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>157.13***</td>
<td>(96.13; 218.13)</td>
<td>44.45</td>
<td>1.08</td>
</tr>
<tr>
<td>GM</td>
<td>SRIL-TC</td>
<td>1.09***</td>
<td>(0.79; 1.38)</td>
<td>0.19***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.87***</td>
<td>(0.63; 1.11)</td>
<td>0.16***</td>
<td>1.25</td>
</tr>
<tr>
<td>GY</td>
<td>SRIL-TC</td>
<td>0.14***</td>
<td>(0.09; 0.20)</td>
<td>0.10**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.13***</td>
<td>(0.08; 0.18)</td>
<td>0.06*</td>
<td>1.08</td>
</tr>
</tbody>
</table>

*a* genetic variance  
*b* interaction variance due to genotype and environment  
*c* genetic variance due to selfed recombinant inbred line and genetic variance due to intermated recombinant inbred line

*P< 0.05 level; **P< 0.01 level; ***P< 0.001 level.
Table 4.3 Genetic correlation for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha\(^{-1}\)) traits measured on selfed recombinant inbred line testcross (SRIL-TC) and intermated recombinant inbred line testcross (IRIL-TC) populations of maize lines grown in six environments in Ontario, Canada. Values in italics bold are significant. \((P < 0.05)\)

<table>
<thead>
<tr>
<th>Trait</th>
<th>Pop</th>
<th>DS</th>
<th>LN</th>
<th>SG</th>
<th>EL</th>
<th>KRN</th>
<th>KW</th>
<th>GM</th>
<th>GY</th>
</tr>
</thead>
<tbody>
<tr>
<td>DA</td>
<td>SRIL-TC</td>
<td>0.96(0.013)</td>
<td>0.48(0.117)</td>
<td>0.51(0.090)</td>
<td>0.18(0.108)</td>
<td>0.10(0.106)</td>
<td>0.24(0.119)</td>
<td>0.68(0.065)</td>
<td>0.31(0.120)</td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.98(0.015)</td>
<td>0.55(0.115)</td>
<td>0.36(0.110)</td>
<td>0.11(0.124)</td>
<td>0.27(0.116)</td>
<td>0.17(0.134)</td>
<td>0.56(0.087)</td>
<td>0.37(0.130)</td>
</tr>
<tr>
<td>DS</td>
<td>SRIL-TC</td>
<td>0.35(0.122)</td>
<td>0.49(0.088)</td>
<td>0.12(0.107)</td>
<td>0.19(0.101)</td>
<td>0.12(0.010)</td>
<td>0.63(0.067)</td>
<td>0.19(0.119)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.50(0.116)</td>
<td>0.32(0.108)</td>
<td>0.13(0.120)</td>
<td>0.22(0.113)</td>
<td>0.19(0.016)</td>
<td>0.60(0.079)</td>
<td>0.35(0.126)</td>
<td></td>
</tr>
<tr>
<td>LN</td>
<td>SRIL-TC</td>
<td>-0.09(0.133)</td>
<td>0.01(0.133)</td>
<td>-0.15(0.127)</td>
<td>0.49(0.133)</td>
<td>0.36(0.111)</td>
<td>0.42(0.139)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.11(0.128)</td>
<td>-0.07(0.132)</td>
<td>0.11(0.128)</td>
<td>0.17(0.143)</td>
<td>0.36(0.116)</td>
<td>0.40(0.132)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SG</td>
<td>SRIL-TC</td>
<td>0.25(0.102)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.19(0.111)</td>
<td></td>
<td>-0.01(0.119)</td>
<td>0.18(0.121)</td>
<td>0.24(0.099)</td>
<td>0.36(0.116)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EL</td>
<td>SRIL-TC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>-0.22(0.099)</td>
<td></td>
<td>0.11(0.116)</td>
<td>0.02(0.101)</td>
<td>0.18(0.114)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KRN</td>
<td>SRIL-TC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>-0.17(0.111)</td>
<td></td>
<td>0.16(0.016)</td>
<td>-0.07(0.107)</td>
<td>0.30(0.115)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KW</td>
<td>SRIL-TC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>-0.65(0.033)</td>
<td></td>
<td></td>
<td>0.13(0.097)</td>
<td>-0.12(0.114)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM</td>
<td>SRIL-TC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>-0.68(0.039)</td>
<td></td>
<td></td>
<td></td>
<td>0.12(0.121)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GY</td>
<td>SRIL-TC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRIL-TC</td>
<td>0.23(0.107)</td>
<td></td>
<td></td>
<td>0.25(0.123)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.36(0.005)</td>
<td>0.23(0.130)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.19(0.111)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.42(0.090)</td>
<td></td>
</tr>
</tbody>
</table>

133
Table 4.4 QTL detected for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha\(^{-1}\)) in selfed recombinant inbred line testcross (SRIL-TC) and intermated recombinant inbred line testcross (IRIL-TC) populations combined across six environments.

<table>
<thead>
<tr>
<th>Chromo</th>
<th>Marker interval</th>
<th>Pos (cM)</th>
<th>LOD</th>
<th>R²</th>
<th>Add</th>
<th>Trait</th>
<th>Chromo</th>
<th>Marker interval</th>
<th>Pos (cM)</th>
<th>LOD</th>
<th>R²</th>
<th>Add</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ug2 - umc1035</td>
<td>154.8</td>
<td>3.12</td>
<td>8.7</td>
<td>0.32</td>
<td>DA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>ug37 - umc1141</td>
<td>70.17</td>
<td>3.02</td>
<td>8.4</td>
<td>-0.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>ug2 - umc1035</td>
<td>153.8</td>
<td>3.82</td>
<td>7.8</td>
<td>0.4</td>
<td>DS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ug6 - ug7</td>
<td>28.13</td>
<td>3.82</td>
<td>7.8</td>
<td>-0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>ug37 - umc1141</td>
<td>72.17</td>
<td>3.81</td>
<td>7.8</td>
<td>-0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>umc1139 - umc1807</td>
<td>29.09</td>
<td>3.3</td>
<td>8.3</td>
<td>0.05</td>
<td>LN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ug8 - umc1065</td>
<td>37.09</td>
<td>4.36</td>
<td>11.6</td>
<td>-0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>ug37 - umc1309</td>
<td>92.6</td>
<td>4.31</td>
<td>11.5</td>
<td>-0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ug21 - umc1524</td>
<td>184.48</td>
<td>2.53</td>
<td>7.9</td>
<td>-0.21</td>
<td>KRN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>bnlg2235 - umc1139</td>
<td>7</td>
<td>2.98</td>
<td>9.2</td>
<td>-0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ug7 - ug8</td>
<td>33.08</td>
<td>3.36</td>
<td>9.2</td>
<td>-3.5</td>
<td>KW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>ug14 – nc005</td>
<td>30.06</td>
<td>3.79</td>
<td>10.5</td>
<td>-6.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>ug4 - umc1774</td>
<td>323.75</td>
<td>5.66</td>
<td>14.2</td>
<td>-0.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>ug38 - phi022</td>
<td>65.81</td>
<td>4.07</td>
<td>9.8</td>
<td>-0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>umc1031 – umc1155</td>
<td>143.98</td>
<td>3.79</td>
<td>11.4</td>
<td>-0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>umc1031 – umc1155</td>
<td>148.97</td>
<td>3.38</td>
<td>9.8</td>
<td>3.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>ug31 - ug32</td>
<td>0</td>
<td>2.86</td>
<td>8.3</td>
<td>3.28</td>
<td>GM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

134
<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Marker Interval</th>
<th>QTL Peak LOD Value</th>
<th>QTL Location on the Chromosome</th>
<th>Additive Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>ug7 - umc1065</td>
<td>36</td>
<td>6.02</td>
<td>-0.46</td>
</tr>
<tr>
<td>2</td>
<td>ug8 - umc1028</td>
<td>77.67</td>
<td>3.25</td>
<td>11.3</td>
</tr>
<tr>
<td>7</td>
<td>ug31 - ug32</td>
<td>0</td>
<td>5.25</td>
<td>16.4</td>
</tr>
<tr>
<td>8</td>
<td>ug37 - umc1141</td>
<td>99.47</td>
<td>2.9</td>
<td>6.1</td>
</tr>
<tr>
<td>9</td>
<td>bnlg1724 – umc1033</td>
<td>44.37</td>
<td>2.75</td>
<td>5.8</td>
</tr>
<tr>
<td>1</td>
<td>umc1917 - umc1076</td>
<td>118.49</td>
<td>4.9</td>
<td>11.2</td>
</tr>
<tr>
<td>3</td>
<td>mmc0132 - ug11</td>
<td>64.47</td>
<td>3.35</td>
<td>7.2</td>
</tr>
<tr>
<td>4</td>
<td>umc1117 – ug14</td>
<td>18.78</td>
<td>2.63</td>
<td>5.6</td>
</tr>
</tbody>
</table>

*a* chromosome number in Maize linkage map  
*b* marker interval of QTL peak LOD value  
*c* QTL location on the chromosome was estimated by MapQTL  
*d* additive effect in reference to CG60 inbred
Supplemental Table 4.1 Epistatic QTL for days to anthesis (DA, days), days to silking (DS, days), leaf number (LN), stay green (SG, score), ear length (EL, cm), kernel row number (KRN), 1000 kernel weight (KW, g), grain moisture (GM, %), and grain yield (GY, Mg ha⁻¹) for selfed recombinant inbred line (SRIL) and intermated recombinant inbred line (IRIL) testcross populations.

<table>
<thead>
<tr>
<th>Interaction group</th>
<th>marker</th>
<th>LG&lt;sup&gt;a&lt;/sup&gt;</th>
<th>marker</th>
<th>LG&lt;sup&gt;a&lt;/sup&gt;</th>
<th>P value</th>
<th>R&lt;sup&gt;2&lt;/sup&gt;&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Trait</th>
<th>Interaction group</th>
<th>marker</th>
<th>LG&lt;sup&gt;a&lt;/sup&gt;</th>
<th>marker</th>
<th>LG&lt;sup&gt;a&lt;/sup&gt;</th>
<th>P value</th>
<th>R&lt;sup&gt;2&lt;/sup&gt;&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>DA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>bnlg2235</td>
<td>8</td>
<td>umc2148</td>
<td>5</td>
<td>0.0002</td>
<td>12.45</td>
<td></td>
<td></td>
<td>bnlg2235</td>
<td>8</td>
<td>umc1518</td>
<td>2</td>
<td>0.0001</td>
<td>15.82</td>
</tr>
<tr>
<td>2</td>
<td>mmc0041</td>
<td>1</td>
<td>ug23</td>
<td>6</td>
<td>0.0002</td>
<td>13.6</td>
<td></td>
<td></td>
<td>mmc0241</td>
<td>6</td>
<td>ug17</td>
<td>4</td>
<td>0.0002</td>
<td>14.04</td>
</tr>
<tr>
<td></td>
<td>ug10</td>
<td>2</td>
<td>ug39</td>
<td>9</td>
<td>0.0001</td>
<td>12.93</td>
<td></td>
<td></td>
<td>ug35</td>
<td>7</td>
<td>ug7</td>
<td>2</td>
<td>0.0001</td>
<td>13.71</td>
</tr>
<tr>
<td>DS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>ug10</td>
<td>2</td>
<td>ug39</td>
<td>9</td>
<td>0.0001</td>
<td>12.42</td>
<td></td>
<td></td>
<td>dupssr10</td>
<td>5</td>
<td>umc1045</td>
<td>10</td>
<td>0.0001</td>
<td>13.76</td>
</tr>
<tr>
<td></td>
<td>ug35</td>
<td>7</td>
<td>ug7</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ug35</td>
<td>7</td>
<td>ug7</td>
<td>2</td>
<td>0.0001</td>
<td>13.56</td>
</tr>
<tr>
<td>SG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ug13</td>
<td>3</td>
<td>ug40</td>
<td>9</td>
<td>0.0002</td>
<td>12.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>ug1</td>
<td>1</td>
<td>ug47</td>
<td>10</td>
<td>0.0004</td>
<td>10.34</td>
<td></td>
<td></td>
<td>bnlg619</td>
<td>9</td>
<td>phi113</td>
<td>5</td>
<td>0.0008</td>
<td>10.47</td>
</tr>
<tr>
<td></td>
<td>ug33</td>
<td>7</td>
<td>umc1065</td>
<td>2</td>
<td>0.0009</td>
<td>9.43</td>
<td></td>
<td></td>
<td>mmc0271</td>
<td>2</td>
<td>ug38</td>
<td>9</td>
<td>0.0009</td>
<td>9.55</td>
</tr>
<tr>
<td></td>
<td>umc1031</td>
<td>5</td>
<td>umc2105</td>
<td>3</td>
<td>0.0009</td>
<td>10.01</td>
<td></td>
<td></td>
<td>ug31</td>
<td>7</td>
<td>ug37</td>
<td>9</td>
<td>0.0007</td>
<td>11.88</td>
</tr>
<tr>
<td></td>
<td>umc1269</td>
<td>1</td>
<td>umc1656</td>
<td>6</td>
<td>0.0005</td>
<td>20.47</td>
<td></td>
<td></td>
<td>ug49</td>
<td>10</td>
<td>umc1155</td>
<td>5</td>
<td>0.0001</td>
<td>9.86</td>
</tr>
<tr>
<td>EL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>bnlg589</td>
<td>4</td>
<td>bnlg619</td>
<td>9</td>
<td>0.0001</td>
<td>16.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>ug46</td>
<td>10</td>
<td>umc1141</td>
<td>8</td>
<td>0.0001</td>
<td>15.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KRN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>mmc0132</td>
<td>3</td>
<td>umc1366</td>
<td>9</td>
<td>0.0001</td>
<td>15.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

136
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>ug12</td>
<td>3</td>
<td>umc1028</td>
<td>2</td>
<td>0.0002</td>
</tr>
<tr>
<td>13</td>
<td>bnlg1724</td>
<td>9</td>
<td>ug3</td>
<td>1</td>
<td>0.0002</td>
</tr>
<tr>
<td>14</td>
<td>umc1518</td>
<td>2</td>
<td>umc2105</td>
<td>3</td>
<td>0.0002</td>
</tr>
<tr>
<td>14</td>
<td>phi113</td>
<td>5</td>
<td>umc1033</td>
<td>9</td>
<td>0.0002</td>
</tr>
<tr>
<td>14</td>
<td>ug13</td>
<td>3</td>
<td>umc1103</td>
<td>7</td>
<td>0.0001</td>
</tr>
<tr>
<td>15</td>
<td>umc1103</td>
<td>8</td>
<td>umc1148</td>
<td>3</td>
<td>0.0001</td>
</tr>
<tr>
<td>GM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>ug31</td>
<td>7</td>
<td>ug43</td>
<td>9</td>
<td>0.0001</td>
</tr>
<tr>
<td>GY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>umc1076</td>
<td>1</td>
<td>umc1844</td>
<td>3</td>
<td>0.0001</td>
</tr>
<tr>
<td>16</td>
<td>ug27</td>
<td>6</td>
<td>umc1336</td>
<td>10</td>
<td>0.0001</td>
</tr>
<tr>
<td>17</td>
<td>ug36</td>
<td>8</td>
<td>umc1269</td>
<td>1</td>
<td>0.0002</td>
</tr>
<tr>
<td>18</td>
<td>umc1077</td>
<td>10</td>
<td>umc1269</td>
<td>1</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

*a Linkage group
*b Percentage of phenotypic variance explained by epistatic effects of the mapped QTL
5.1.1 Summary and main conclusions

In this study, two early-season inbred lines from different Corn Belt Dent heterotic groups (CG60 - Iodent and CG102 - Stiff Stalk) were used to develop a selfed recombinant inbred line (SRIL) population and an intermated recombinant inbred line (IRIL) population. The lines within these populations were testcrossed with an inbred tester from a different heterotic group, Lancaster Sure Crop (LSC). The inbred populations (SRIL and IRIL) and SRIL testcross (SRIL-TC) and IRIL testcross (IRIL-TC) populations were used to estimate trait means, variances and correlations. In addition, these populations were used for QTL mapping for grain yield and agronomic traits.

The genetic difference between lines from different heterotic groups has not been previously reported in maize inbreds selected for early-season environment. The usefulness of recombinant inbred lines derived from inter-heterotic crosses for maize hybrid breeding will enhance our knowledge of breeding method for inbred line development. Moreover, the use of intermating, to break up the repulsion linkage as predicted by Hill and Robertson (1966), to develop the desired recombinant types in early season maize has not been reported. This study reported large amount of trait variation in recombinant inbred line populations derived from two phenotypically similar parents. The genetic variances and genotypic correlations did not change after intermating, suggesting that the genes responsible for these traits are segregating independently. The
inbred lines within the populations had an extensive amount of transgressive segregation relative to parental inbred lines suggesting that these inbred lines are an important resource for hybrid maize breeding programs.

This study focused also on the linkage mapping and QTL comparison between the SRIL population and the IRIL population. In addition, SRIL and IRIL populations were also simulated with different levels of QTL effects and compared for QTL detection. As expected, the intermating resulted in genetic map expansion. The usefulness of the IRIL population for precise QTL detection shown by reduced QTL support intervals for a variety of traits. The support intervals, in both the empirical study and the simulation study, were reduced in the IRIL population as compared to the SRIL population.

Hybrid trait variation and QTL mapping of yield and other traits were also studied using the SRIL-TC and IRIL-TC populations. The results showed that traits greatly varied in both testcross populations. Genetic variances and genetic correlations, however, did not significantly differ between SRIL-TC and IRIL-TC populations. In total, 25 QTL were identified for different traits in both testcross populations. Only one QTL, for grain moisture, was shared between populations. The results of this study suggest that two inbred parents have different positive and negative alleles that contribute to trait values.

In a comparison of QTL for flowering time, leaf number and stay green between the inbred populations and testcross populations, seven regions were identified that contained QTL in inbred SRIL population (Chapter Three). One of
these QTL (2L) for days to silking seems to be associated with SRIL testcross population QTL (Chapter four). Parental contributions were the same in both inbred and testcross progeny for this QTL. The QTL detected for DA, DS and LN in IRIL inbred population were not detected in IRIL-TC population. The QTL in the inbred population explaining the largest portion of variation for days to anthesis was in the SRIL population, which was located on 2L (between ug8 – umc1065) and was not detected in the testcross populations. In a comparison of F2:4 progeny per se and testcross progeny, Beavis et al (1994) identified four common QTL for grain yield and several morphological traits. In another study, Schön et al (1994) reported only one QTL for kernel weight and two QTL for protein content were common in inbred population and testcross population. On the basis of previous studies and the evidence reported in this study, QTL identified for hybrid traits would not sufficient for inbred population performance. Breeding programs would need to identify and incorporate QTL for both inbred population and testcross population.

5.1.2 Limitations of the study

Like most studies, this study is not without limitations. For example, QTL studies require very large sample sizes, and only those molecular differences can be mapped that are polymorphic between the initial parental lines. In this study, I used 135 recombinant inbred lines in each population for QTL mapping. The small sample size reflects the maximum number of progeny generally employed per cross in a breeding program (Beavis et al., 1994). Similarly, I used
120 polymorphic markers for genotyping the SRIL and IRIL populations. In the inbred populations, the phenotypic trait variations for all traits were exceptionally high. I was expecting major effect QTL were segregating for these traits and used 120 markers with an average marker density of 19.65 cM for QTL analysis.

5.1.3 Future research and research contributions

The results of this study suggested that the use of inter-heterotic group crossing for inbred development can help to increase the genetic diversity of germplasm used for inbred line development. However, as heterotic groups were originally created to make heterosis more predictable, some strategic crossing of heterotic groups is probably more feasible than complete loss of heterotic groups. The results of this study are based on single tester. Hence further experiments with different testers are required to clarify whether the findings of this study are applicable across different testers. In addition, the genetic material used in this study belongs to maize inbred lines adapted to early-season environments. It would be more informative to conduct similar experiments with other maize inbred lines and growing environments.

This research encourages the use of inter-heterotic group crosses for inbred line development in a breeding program. The success of the breeding program depends on the population mean and the genetic variation for a trait of interest (Bernardo, 2002). Because the mean and trait variation for grain yield are high in the RIL populations, the inbred lines can be valuable resource for hybrid maize breeding. It is important to know if the use intermating would enhance the
chances to select new lines resulting in improved hybrids. In the present study, intermating did not significantly increase the trait variances and trait correlations.
References


Appendices

Appendix 1 SAS code for the estimation of best linear unbiased predictors (BLUPs) for lines grown in replicated multi-environment trials. The code estimates BLUPs for grain yield (GY) in the intermated recombinant inbred line population.

```sas
proc import DATAFILE="C:\SAS\HYBRID\HYBRIDSAS.xls" out = a replace;
run;
data B; set a;
if pop='P1' then delete;
if pop='P2' then delete;
if pop='C' then delete;
if pop='NIM' then delete;
%macro all_BLUPs(trait);
proc mixed data=B;
class env rep LINE;
model &trait = /ddfm=satterth solution;
random env rep(env) LINE env*LINESolution;
ods listing exclude solutionF solutionR;
ods output solutionF = &trait.F solutionR = &trait.R;
run;
data m; set &trait.F; if Effect = "Intercept"; grd_mn = Estimate; dummy = 1;
proc print;
run;
data e; set &trait.R; if Effect = "LINE"; rnd_effect = Estimate; dummy = 1;
proc print;
run;
data &trait.blup;
merge e m; by dummy;
BLUP = grd_mn + rnd_effect;
&trait=BLUP;
keep LINE grd_mn md_effect &trait;
proc sort data= &trait.blup; by LINE;
proc print data= &trait.blup;
run;
proc export data=&trait.blup outfile='C:\SAS\hybrid\BLUPLNIM.xls'
DBMS=EXCEL REPLACE;
SHEET = &trait;
run;
%mend all_BLUPs;
%all_BLUPs (GY)
run;
```
Appendix 2 SAS code for the estimation of heritabilities on a plot and an entry means basis using variance components according to Holland et al. (2003). The code refers to the heritability of grain yield (GY) content in the IRIL testcross population.

/*SAS program to estimate heritability from random lines evaluated in a randomized complete block design with 2 replications at each of 6 environments, with some missing data*/
proc import DATAFILE="C:\SAS\hybrid\hybridsas.xls" out = a replace;
run;
data one; set a;
if pop='P1' then delete;
if pop='P2' then delete;
if pop='C' then delete;
if pop='NIM' then delete;
run;
/*eliminate checks and lines from different population from data set*/
ods listing;
data two; set one;
proc sort; by env;
proc mixed asycov; class env rep line; model GY= ;
random env rep(env) line line*env;
ods listing exclude AsyCov CovParm; ods output asycov = covmat covparms = estmat;
proc iml;
start seh(V, C, LG, LP, H, SE);
Vp = LP`*V;
Vg = LG`*V;
H = VG/Vp;
d = (1/Vp)*(LG - (LP*H));
VH = d`*C*d;
SE = sqrt(VH);
finish seh;
use estmat; read all into v; use covmat; read all into c;
/* Note that SAS introduces an extra first column into the matrix which must be removed*/
C = C(1:nrow(C), 2:ncol(C));
/*order of variance components in v and c matrices is V(E), V(R), V(G), V(GE), V(error)*/
LG = {0, 0, 1, 0, 0};
LP = {0, 0, 1, 1, 1};
call seh(V, C, LG, LP, H, SE);
print "Heritability on a Plot Basis", H, SE;
quit; run;
%mend heritability;
%
heritability
run;
Appendix 3 SAS code for the estimation of genotypic and phenotypic correlations and their standard errors according to Holland (2006). The code refers to the genotypic and phenotypic correlation of days to anthesis (DA) with days to silking (DS) in the IRIL testcross population.

```
proc import DATAFILE="c:\SAS\HYBRID\hybridSAS.xls" out = a replace;
run;
data B; set a;
   if pop='P1' then delete;
   if pop='P2' then delete;
   if pop='C' then delete;
   if pop='NIM' then delete;
run;
/*first, estimate variance components for each trait separately to compare to multivariate analysis*/
%macro varcomp(trait);
   proc mixed data = B;
      class env rep LINE;
      model &trait = ;
      random env rep(env) LINE LINE*env;
   /*also check effect of setting environments and reps fixed on other variance components*/
   proc mixed data = B;
      class env rep LINE;
      model &trait = env rep(env) ;
      random LINE LINE*env;
   run;
%mend;
%varcomp(DA);
%varcomp(DS);
/*restructure data set for multivariate reml analysis*/
data two; length trait $ 3; set B;
   trait = "DA"; y = DA; output;
   trait = "DS"; y = DS; output;
   drop DA DS;
/*analyze variables pair-wise*/
%macro corr(trait1, trait2);
   data traits; set two; if trait = "&trait1" or trait = "&trait2";
   proc mixed asycov data = traits;
      class trait env rep LINE;
      model y = env(trait) rep(env*trait);
      random trait/subject = LINE type = un;
      random trait/subject = LINE*env type = un;
      repeated trait/ sub = rep*LINE(env) type = un;
      ods output covparms = estmat; ods output asycov = covmat;
   run;
   proc iml;
      use estmat; read all into e;
      use covmat; read all into cov;
      /* Note that SAS introduces an extra first column into the covariance matrix which must be removed*/
      C = cov(1:nrow(cov), 2:ncol(cov));
   /*obtain genotypic and phenotypic covariance and variance components*/
      CovG = e(2,1);
      VG1 = e(1,1);
      VG2 = e(3,1);
      CovP = CovG + e((5,1)) + e((8,1));
      VP1 = VG1 + e((4,1)) + e((7,1));
      VP2 = VG2 + e((6,1)) + e((9,1));
   /*create a module called "correl" that will estimate genotypic and phenotypic correlations and their standard errors*/
   start correl(C, CovG, VG1, VG2, CovP, VP1, VP2, RG, RP, SERG, SERP);
```

RG = CovG/sqrt(VG1*VG2);
/*Make the derivative vector for rg, note that the order of the rows and columns of the variance
 covariance matrix is VG1, CovG, VG2, VGE1, CovGE, VGE2, VError1, CovError, VError2*/
dg = (-1/(2*VG1))/(1/CovG)/(-1/(2*VG2))/0/0/0/0/0/0;
varrg = (RG**2)*dg*C*dg; serg = sqrt(varrg);
RP = CovP/sqrt(VP1*VP2);
/*Make the derivative vector for rp*/
d1p = -1/(2*VP1);
d2p = 1/CovP;
d3p = -1/(2*VP2);
dp= d1p/d2p//d3p//d1p//d2p//d3p//d1p//d2p//d3p;
varrp = (RP**2)*dp*C*dp; serp = sqrt(varrp);
finish correl;
call correl(C, CovG, VG1, VG2, CovP, VP1, VP2, RG, RP, SERG, SERP);
print "Genotypic Correlation Between &trait1 and &trait2";
print RG serg;
print "Phenotypic Correlation Between &trait1 and &trait2";
print RP serp;
quit; run;
%mend;
%corr(DA,DS);
run;