ABSTRACT

QUANTITATIVE TRAIT LOCI AND GLOBAL GENE EXPRESSION PATTERNS OFFER INSIGHT INTO PHENOTYPIC PLASTICITY

By Matthew J. Rubin

Phenotypic plasticity, i.e. changes in phenotype with environment, may allow organisms to produce optimal phenotypes in all environments (adaptive plasticity). There is considerable discussion about the genetic mechanisms for phenotypic plasticity. However, there is some agreement that changes in gene expression must be involved. In order to examine how natural selection has acted on phenotypic plasticity on the trait bolting time trait (transition to reproduction) and gene expression patterns in past populations, a set of Recombinant Inbred Lines (RILs) of *Arabidopsis thaliana* were used. We tested the effect of cold-treating seeds (stratification) on bolting time in *Arabidopsis thaliana*. Cold stratification of seeds may be one environmental factor contributing to variation in spring versus fall germination in *A. thaliana* populations. Variation in both the direction and degree of plasticity was observed in a set of 120 RILs screened; genotypic selection analysis showed that past selection had favored bolting earlier in both environments. Three RILs that displayed extreme plasticities for bolting time in opposing directions across cold treatments were identified and global gene expression patterns were measured in a microarray experiment. A total of 294 genes were identified as being differentially expressed across cold treatments for the three extreme RILs (Fold change of >2; p-value<0.05). In addition, Quantitative Trait Loci (QTL) for bolting time were mapped in the complete set of RILs. Five QTL were mapped in the cold environment and three QTL were mapped in the no-cold environment explaining 40% and 30% of the observed phenotypic variation, respectively. QTL underlying variation in bolting time were screened for differentially expressed genes from the microarray study to identify candidate genes. Ninety-three of the identified genes co-localized to bolting time QTL, with kinesis and transferases being overrepresented in the expressed QTLs. Few studies have combined microarray and QTL data, and this study will offer insight into the genetic mechanism of phenotypic plasticity.

QUANTITATIVE TRAIT LOCI AND GLOBAL GENE EXPRESSION PATTERNS OFFER INSIGHT INTO PHENOTYPIC PLASTICITY

by

Matthew J. Rubin

A Thesis Submitted
In Partial Fulfillment of the Requirements
For the Degree of

Master of Science - Biology

at

The University of Wisconsin Oshkosh
Oshkosh WI 54901-8621

December 2011

COMMITTEE APPROVAL

_____ Advisor

14 Dec 2011 Date Approved

_____ Member

Dec. 14, 2011 Date Approved

_____ Member

12/14/2011 Date Approved

PROVOST
AND VICE CHANCELLOR

_____

12/14/2011
Date Approved

FORMAT APPROVAL

_____

12/7/11
Date Approved

ACKNOWLEDGEMENTS

There have been so many people who have not only helped me over my 6 years at the University of Wisconsin Oshkosh, but have also shaped me in to the person and researcher that I am today. First, I would like to thank my advisor Dr. Lisa Dorn who was willing to take on an undergraduate student with no experience in Genetics literally days after she moved to Oshkosh and my committee members: Dr. Toivo Kallas, and Dr. Todd Kostman for their support and guidance throughout my time at the UWO. Many of the projects I completed at UWO were at a very large scale and I could not have completed them on my own, therefore I would like to extend a huge thank you to everyone one who helped collect data on my thousands of plants, specifically all the members of the Gene Team, including Mohamed Yakub, Megan Nelson, Vai Lor, Jackie Mischka, George Weir IV, many REU students and undergraduate research technicians.

I also wish to thank the faculty, staff and administrative assistants of the Biology and Microbiology Department, it is because of each and every one of you that our department is so strong. I also extend my sincere gratitude to my home-away-from-home the Office of Science Outreach and all of the wonderful staff for developing my love of teaching; I want to extend a special thank you to Anne Murphy! Thank you to all my friends, both old and new, that made my time at UWO so memorable.

Lastly, but certainly not least, I wish to thank my family for supporting me throughout this time. Without their continuing positive encouragement none of this would have been possible.

TABLE OF CONTENTS

Page

LIST OF TABLES

Following Page 35

LIST OF FIGURES

Page

Following Page 35

CHAPTER I

INTRODUCTION

1.0 Phenotypic Plasticity

In heterogeneous environments the ability for a particular genotype to achieve the optimal phenotype in response to changing environmental cues is advantageous over genotypes with fixed phenotypes; the ability for a particular genotype to produce more than one phenotype is called phenotypic plasticity. If plasticity allows the organism to match its' phenotype to the optimum for each relevant environment, it may result in high fitness across the ecological conditions an organism may experience. Plasticity is therefore often assumed to be an adaptive evolutionary mechanism. In order for plasticity to be adaptive plasticity selection must differ across environments and plastic genotypes must alter their phenotype in the direction of the optimum for each environment. If plasticity is adaptive, we expect plastic genotypes to have increased fitness in two or more environment relative to non-plastic or homeostatic genotypes (Dudley and Schmitt 1995 & 1996). Present day populations are the product of past natural selection events, therefore selection may have culled out non-adaptive plastic genotypes due to reduced fitness. Therefore, assuming that there is genetic variation for selection to act on and if selection has culled out non-adaptive forms of plasticity (non-adaptive and maladaptive) it may be difficult to identify and study plasticity in natural populations. In addition, if no constraints (limits) or costs are associated with being plastic then individuals would

evolve perfect plasticity in which they have the trait mean that confers the highest fitness in all environments; however non-adaptive and mal-adaptive plasticities persist in natural populations providing evidence that there may be constraints on the evolution of adaptive plasticity (Dewitt et al. 1998; Dorn et al. 2000; Relyea 2002; Weinig et al. 2006)

1.1 Genetic mechanisms of plasticity

The genetic mechanism of phenotypic plasticity is not completely understood; however there are several proposed mechanisms. The first two mechanisms involve the expression of the same set of genes in all environments. In the first case the expression level of a suite of structural genes is constant, but the functionality of the gene product or protein is decreased in some environments; some argue that in this first case plasticity could arise from failure to initiate the appropriate response or functional gene products, i.e. this could result in the same phenotype regardless of environment (homeostasis) or could produce different phenotypes due to a strong G x E interaction. The second case revolves around gene regulation, in which the same genes are expressed at different levels, i.e. an organism has sensory mechanisms that feed into regulatory genes that control expression level of structural genes based on environment (Via et al. 1995). The third proposed mechanism underlying plasticity is the expression of different genes in different environments, implying the conservation of alternative phenotypic pathways which may or may not converge at a structural gene within the gene network responsible for the expression on phenotype. A forth hypothesized mechanism is that plastic organisms possess functional alleles at plasticity genes that interact with structural genes

to mediate expression of phenotype. This mechanism implies that plasticity is a trait on which selection can act, as opposed to plasticity being inherent in a trait (Via 1993a). This mechanism predicts that transcription factors, not structural genes, are the likely candidates for plasticity genes. Lastly, plasticity may result from a combination of two or more of the above described mechanisms and the mechanism may ultimately be genotype specific.

1.2 Costs of plasticity

Costs can be either local (costs associated with a single micro-environment) or global (costs associated with multiple micro-environments). Most studies attempt to identify local costs because they are looking at plasticity in genotypic trait means in two environments (i.e. snail shell morphology in presence vs. absence of predators or in presence of predator A vs. predator B). Even if selection has culled out only the extreme non-adaptive forms of plasticity it may be difficult to identify and study plasticity in natural populations. The genetic structure of Recombinant Inbred Lines of (RILs) *Arabidopsis* is likely to emulate that of natural populations, in which rare out-crossing events occur between colonists followed by subsequent inbreeding of individuals through selfing. Therefore, through the use of sets of RILs it is possible to examine plasticity and potentially detect costs of plasticity (reduced fitness relative to homeostatic or non-plastic genotypes) that are rarely detected in natural populations (Dorn et al. 2000; Van Kleunen & Fischer 2007). The increase in power to detect costs may be to due to the fact that extreme non-adaptive and maladaptive plastic genotypes may persist in recently created

RILs, whereas selection may have culled the extreme plastic genotypes from natural population. One of the most common costs associated with being plastic is the energetic cost of maintaining mechanisms to sense the environment and downstream pathways required to alter phenotype in response to some environmental cue; maintaining alternative pathways requires resources that could be allocated to reproduction, in turn increasing fitness. Encompassed in the first cost is the idea of production costs associated with producing a particular phenotype; however, some argue that even non-plastic genotypes experience production costs. Therefore, only when the production cost associated with plastic genotypes exceeds those of non-plastic genotypes is there truly a production cost (Dewitt et al. 1998; Relyea 2002). The actual process of acquiring information from the environment required to produce the optimal phenotype may also come with a cost. There may be an overall cost associated with being plastic, if the organism does not experience more than one environment. Lastly are the genetic costs, such as linkage, epistasis, and pleiotropy. Regulatory genes that induce a plastic response by regulating structural phenotypic genes may also modify the expression of other genes resulting in lower fitness. Plasticity genes may have negative pleiotropic effects on other traits or may be in linkage disequilibrium with genes conferring low fitness (Dewitt et al. 1998). Weinig et al. 2003 provide one example of a genetic trade-off (evolutionary constraint) where a specific allele increases fecundity of an annual plant in one environment, but decreases overwinter survivorship in another.

1.3 Constraints (limits) on the evolution of adaptive plasticity

If the phenotypic mean for a trait moves from an adaptive peak, then populations should experience directional selection to drive the population back to the adaptive peak. However, if being plastic is associated with costs that reduce fitness, only homeostatic individuals will have the capacity to reach the adaptive peak for any one environment. However, if the environment is highly heterogeneous, then plastic individuals may have the highest fitness across environments (i.e. reach the summit of the adaptive peak if the adaptive landscape incorporates environmental stability). Not reaching the adaptive peak as a result of producing the wrong phenotype due to inaccurate perception of environment resulting in non-adaptive or maladaptive plasticity is not a cost of plasticity, but rather a constraint on the evolution of adaptive plasticity, because homeostatic genotypes can also produce non-optimal phenotypes in any given environment. The other limits to plasticity are associated with development; plastic genotypes often lag developmentally behind non-plastic genotypes because they must first sense the environment, then initiate development. Fixed development in non-plastic genotypes may be more capable of producing adaptive, extreme phenotypes whereas plastic genotypes may not be able to produce the extreme phenotypes (Dewitt et al. 1998). Both the developmental lag time and inability to produce extreme phenotypes can result in reduced fitness, which may constrain the evolution of adaptive plasticity (Dewitt et al. 1998; Relyea 2002). In addition, just as trait correlations can confound the evolution of a

particular phenotypic trait; phenotypic plasticity may also be constrained by correlations among traits.

1.4 Evolution of plasticity

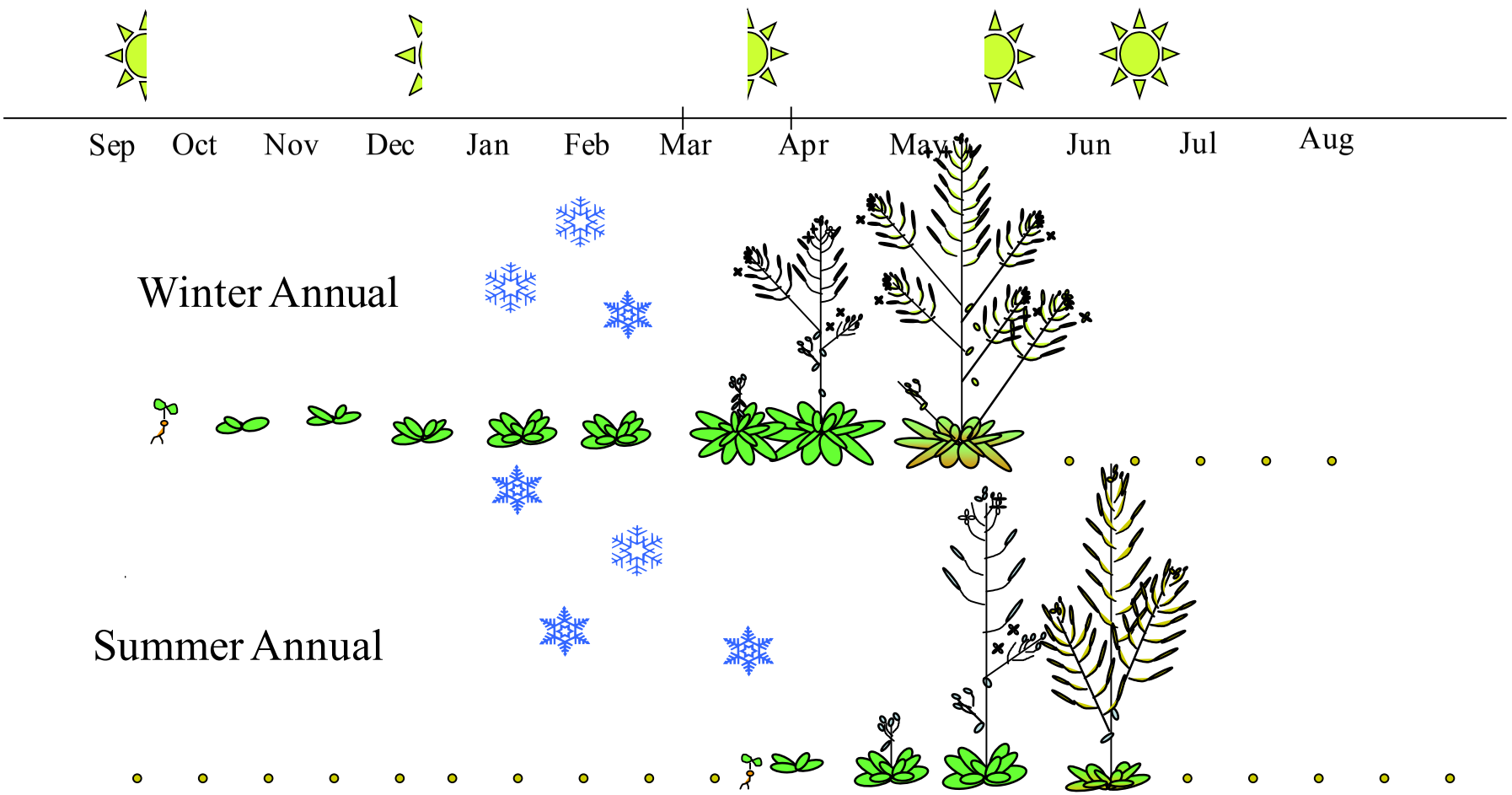
In the overall scheme of phenotype evolution a particular genotype can persist as plastic (varying degrees of plasticity present in populations) or homeostatic. Homeostatic genotypes can evolve as specialists (locally adapted) or generalists (Van Tienderen 1991; Van Tienderen 1997). Natural selection should select against canalization in heterogeneous environments due to large fitness costs; however both generalists and plastic environments should persist in heterogeneous environments. Generalists are often referred to as "a jack-of-all-trades is a master of none" meaning that many often have moderate relative fitness in several environments, but may not ever reach the fitness maximum (Van Tienderen 1997). On the contrary, adaptively plastic genotypes may reach or come closer to fitness maximums in several environments, increasing their fitness relative to generalists (Sultan & Spencer 2002). Since most organisms (or their offspring) will experience variable environments during their lifetime, adaptive plasticity should persist in spite of the costs; although the evolution of perfect plasticity may never be observed due to evolutionary constraints. Long term selection experiments such as the one proposed by Weinig et al. 2003, will examine the persistence of phenotypic plasticity in a set of *Arabidopsis thaliana* Recombinant Inbred Lines (RILs) that have evolved over 5 years under natural settings to ask how selection has acted on plasticity in past populations.

2.0 *Arabidopsis thaliana*: A Model Organism

Arabidopsis thaliana is a small, flowering annual weed in the mustard family (Brassicaceae) found in discrete and genetically differentiated populations in Europe, Asia and now some places in North America. The *Arabidopsis* genus is in the same family as several agriculturally relevant crops such as cabbage, cauliflower, kohlrabi, broccoli, and kale (Meyerowitz 1987; Meyerowitz and Pruitt 1985). *Arabidopsis* is ideal for environmental studies because it is self-fertilizing which allows for the creation and maintenance of genetic lines (both natural ecotypes and mapping populations) and therefore replicates of each genotype can be grown in several environments to ask what effect environment has on phenotype and/or gene regulation. In addition, *Arabidopsis* is a rapid growing annual and fitness can be estimated from fruit number. The 125 mega-base genome of *Arabidopsis thaliana* was sequenced in 2000 and as a result many molecular tools are available.

Environmental heterogeneity that an organism encounters can exist at multiple spatial and temporal scales. For example populations *A. thaliana* encompass a large geographic range and as a result can experience large differences in duration of photoperiod and temperature regimes. Also, individuals within a single population can experience significant variation in abiotic factors as a result of dimorphic germination patterns. These dimorphic germination patterns result in two distinct life histories referred

to as winter annual and spring annuals. (Figure 1-1). In the fall annual life cycle, seeds are dispersed off the maternal plant in late summer, germinate in the fall and overwinter as a rosette of leaves. In the spring when conditions are favorable, the plants transition to reproduction, set seed, and senesce in mid-summer. In spring annuals seeds are dispersed off the maternal plant in late summer and overwinter a seed. In the spring the plant completes its' entire life cycle, from germination to producing seed, prior to mid-summer. Both of these life-history strategies ensure that reproduction has occurred prior to the summer drought (Donohue 2002). Natural populations *A. thaliana* can express either one or both of these life cycle (Figure 1-1).



re 1-1: *Arabidopsis thaliana* life histories as a result of dimorphic germination patterns. Top: *A. thaliana* as a winter annual: s are dispersed off the maternal plant in early to mid-summer and overwinter as a vegetative rosette and flowering early in the g with a large rosette and set seed; Below: *A. thaliana* as a summer annual: Seeds are dispersed off the maternal plant in early d-summer and overwinter as a dormant seed and complete entire life cycle in spring, also called a rapid cycling life cycle.

2.1 Recombinant Inbred Lines

Due to the ability of *Arabidopsis thaliana* to self-fertilize, it is possible to create and maintain genetic lines, such as Recombinant Inbred Lines (RILs). RILs are generated by crossing (i.e. collecting pollen from the paternal plant and using it to fertilize the ovule of the maternal plant) a single inbred individual from a natural population to an individual from a second population. Following the initial cross, the resulting offspring are subsequently inbred for 8 generations. After 8 generations of inbreeding the lines are homozygous at all genes, stabilizing the genome so that all seeds within a single line are identical to each other (Figure 1-2) Therefore, replicate seeds can be grown in different environments to determine if and how the environment influences any phenotype of interest. The structure of RILs is likely to emulate that of natural populations of *A. thaliana*, in which populations are initiated by rare out-crossing events followed by high levels of inbreeding (Jones 1971). Generation of RIL sets allows researchers to recreate allele combinations that have been culled from natural populations by natural selection.

2.2 Cal x Tac Recombinant Inbred Lines

The set of recombinant inbred lines used for this experiment were generated from a cross between two natural ecotypes of *Arabidopsis thaliana*: Cal (Calver, England; CS1062) and Tac-0 (Tacoma, Washington, USA; CS28754) (Huang 2010), where Tac was the maternal plant and Cal was the paternal pollen donor (Figure 1-2). These two

populations have experienced different selection regimes likely due to the differing abiotic and biotic conditions in their geographic locations of origin. Most pronounced is a requirement of a cold-treatment of the seeds (cold-stratification) by the Tac population to germinate that is not a requirement for seeds from the Cal population.

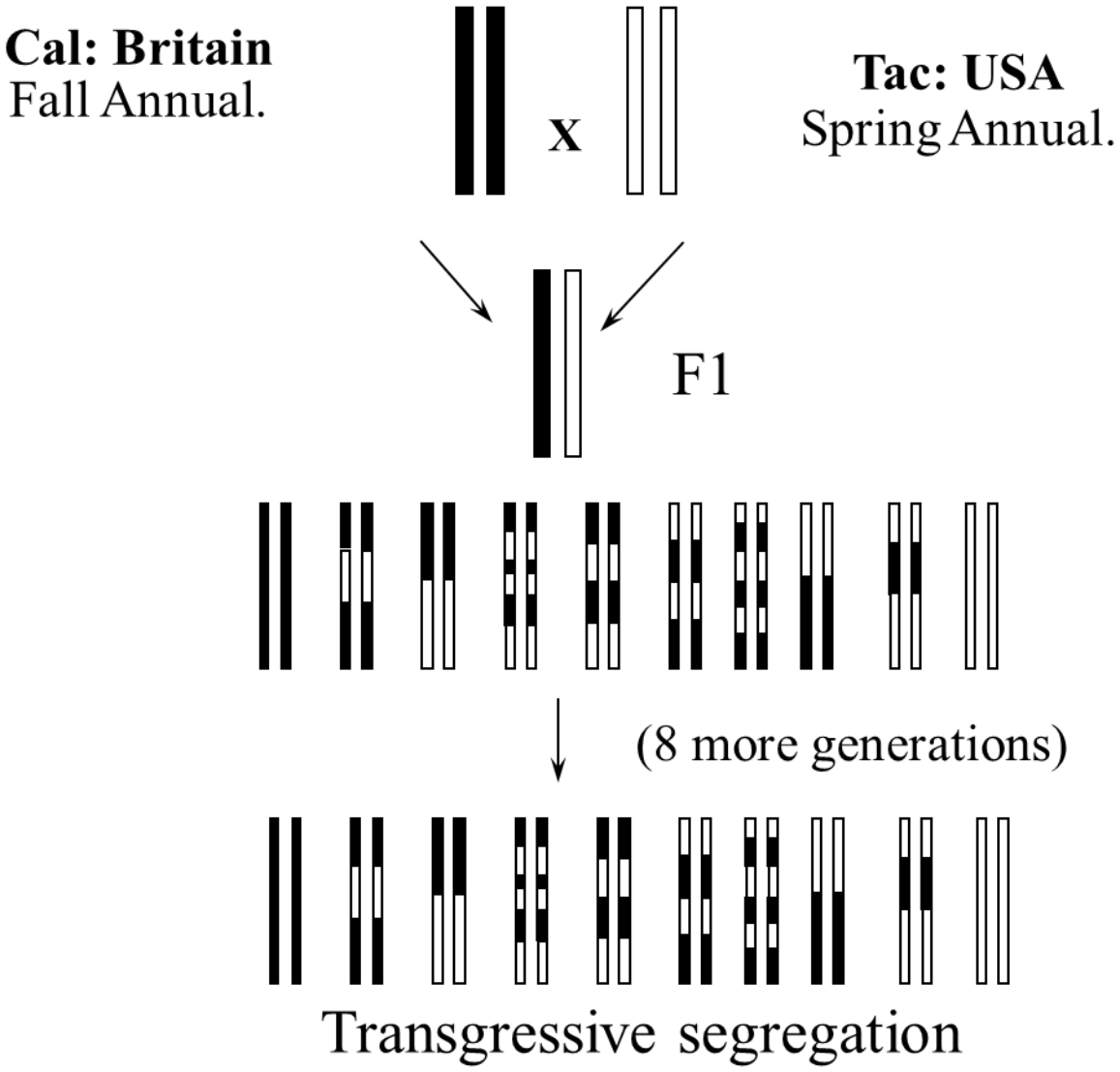


Figure 1-2: Generation of Cal x Tac Recombinant Inbred Lines. Black bars represent alleles derived from Cal parent and white bars represent alleles derived from the Tac parent. Homozygous individuals from two divergent natural populations are crossed to produce a heterozygous F1. The F1 progeny are then selfed to homozygosity over eight generations of single seed decent resulting in a panel of RILs that are novel combinations of the two parental alleles.

3.0 Microarrays and Quantitative Trait Loci Mapping

The central dogma of biology states that DNA is transcribed in RNA which is in-turn translated in proteins and ultimately any given phenotype is determined by protein present within an organism. Phenotypic variation can therefore be attributed to variation in the DNA sequence that results in different variants of the same gene product or protein, the concentration of proteins (as a result of transcriptional or translational regulation) or the interaction between the two above factors and the environment. Microarrays can be used to measure global gene expression patterns at a single time point. A microarray is a microscope slide that contains many single-stranded oligonucleotide probes each representing a single gene of a particular genome. RNA samples are collected, converted to complementary DNA (cDNA) through reverse-transcription polymerase chain reaction and the enzyme reverse transcriptase, and hybridized on the microarray where single-stranded cDNA will bind to the complementary probe and the expression levels for each gene can be calculated. More recently, next generation sequencing technology has been implemented to estimate gene expression levels, where again RNA samples can be collected, stabilized as cDNA, and labeled and the samples can be sequenced. The number of reads per transcript can be used to estimate the expression levels for each gene and compared to other genes or samples. To increase the throughput of microarrays, it is also possible to first label the cDNA with a fluorescent dye allowing hybridization of 2 labeled pools of cDNA from two distinct samples on a single microarray (Figure 1-3).

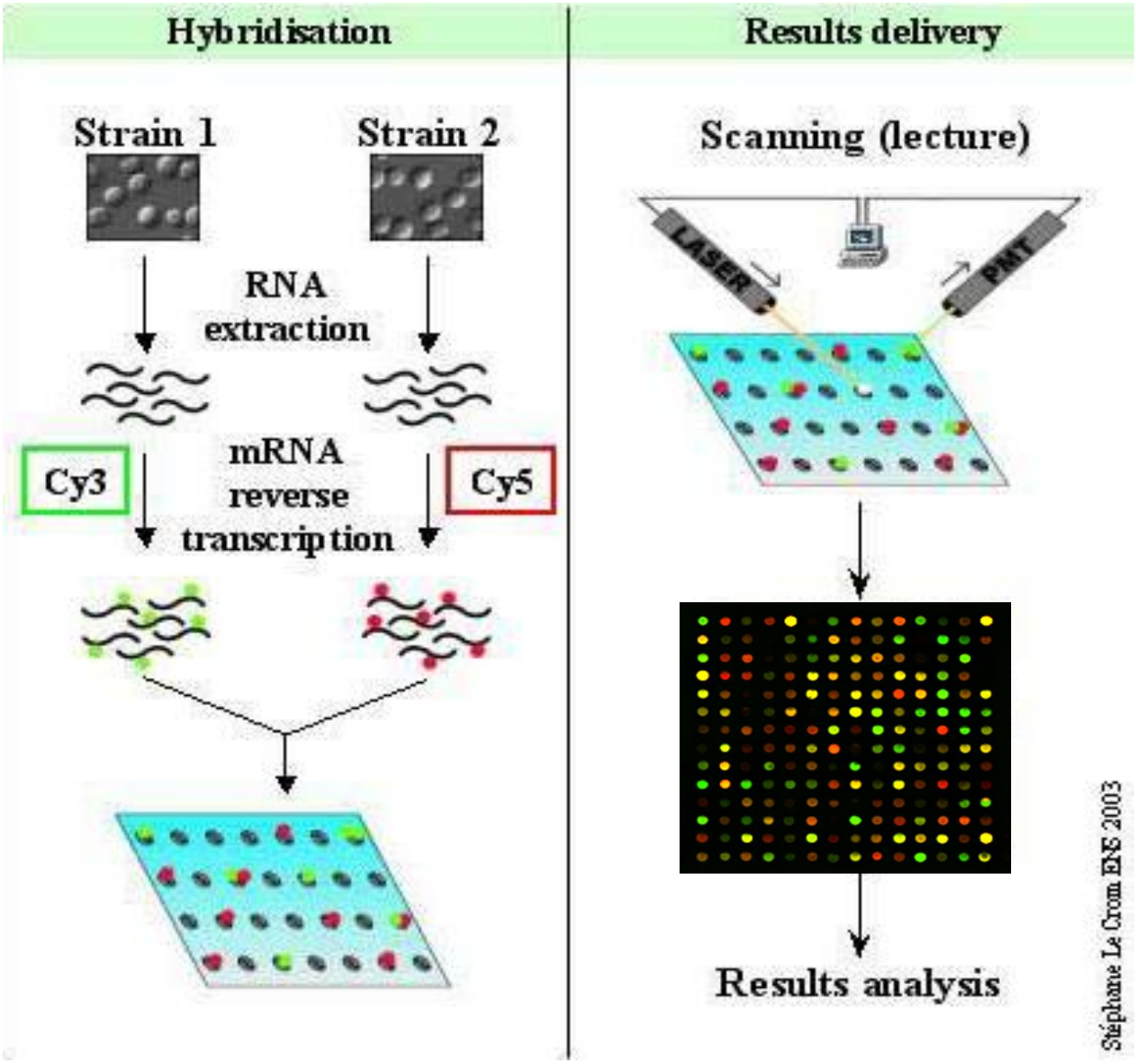


Figure 1-3: Summary of microarray procedure. RNA is extracted from each strain and mRNA was reverse transcribed to cDNA, which was then labeled with a fluorescent dye (either Cy3 or Cy5). Two labeled cDNA samples were pooled for each microarray and the sample was then hybridized onto the microarray. The microarrays were washed to remove unbound sample, scanned and the data was then analyzed.

Microarray have a variety of applications including comparing gene expression patterns from different tissues (Girke et al. 2000), comparing gene expression patterns of different genetic strains (Gilad and Borevitz 2006), quantifying gene expression profiles over a time-course (Girke et al. 2000), or accessing expression patterns across environments (Seki et al. 2001 and Seki et al. 2002).

One major problem with gene expression profiling is that mRNA transcript levels are used as a proxy for protein concentration, when in fact transcript levels could be high but ultimately not undergoing translation. This means though mRNA may be present at high concentrations for a gene, no proteins are being synthesize and are therefore that gene is not contributing to the expression of the measured phenotype of interest. Also, gene expression data are generally a point-estimate of gene expression (i.e. one time point for a certain tissue type) which can also be problematic because at least in *A. thaliana* it has been recently shown that 90 percent of all genes change expression levels on a diurnal cycle (Michael et al. 2008).

There are many important considerations when designing a microarray experiment including the number of biological and/or technical and dye bias (Churchill 2002; Yang and Speed 2002). Due to the fact the hybridization of the cDNA to the microarray slide is a stochastic process, it is imperative to have biological replicates to ensure an accurate estimate of the gene expression patterns for each sample. In addition, previous studies have detected differences both the binding efficiencies and intensities of

the two most commonly used dyes (Cy3 and Cy5). Therefore it is important to control for labeling dye bias through dye-swaps, in which all biological samples are labeled with both experimental dyes and hybridized in a factorial loop design. (Rosenzweig et al.2004; Yang and Speed 2002).

Quantitative Trait Loci (QTL) mapping tests for a significant association between a polymorphic marker within the genome and a phenotype of interest to identify chromosomal genes that underlie variation within the trait and is usually done in a segregating progeny (such as Recombinant Inbred Lines in *A. thaliana*). In addition to identifying the chromosomal regions that explain variation in the trait, the effect size of the QTL can also be determined. QTL mapping studies often do not have the power required to detect QTL of small effect size and it is only possible to detect QTL that contribute to a trait if the genetic marker linked to the QTL is polymorphic.

Microarray gene expression data can be combined with standard phenotypic QTL mapping results to identify candidate genes within a phenotypic QTL. Depending on the marker density of a given genetic map in *Arabidopsis* a single QTL can contain hundreds of genes. Identifying differentially expressed genes that colocalize with a QTL offer a list of candidate genes to be further investigated.

4.0 Present study

In the present study, I first screen a set of 120 Recombinant Inbred Lines of Arabidopsis thaliana in two environments (cold stratification of seeds and no-cold stratification of seeds) for Days to Bolting to better understand if the Cal x Tac RILs differ in both the degree and direction of plasticity in Days to Bolting in response to cold-stratification of seeds. Due to the differences in cold stratification in the parents of the RILs, I predict that there will be significant genetic variation for this trait in the segregating RIL progeny. To better understand the genetic mechanisms that drive the observed plasticity, I plan to select 3 extreme plasticity RILs for gene expression analysis and use the entire set of RILs to identify chromosomal regions that contribute to the expression of variation in Days to Bolting.

4.1 Present study objectives

The objectives of this study were to:

1. Determine if there is genetic variation present in the Cal x Tac RILs for days to bolting
2. Determine if there is variation in the direction and degree of phenotypic plasticity for days to bolting in response to seed stratification.
    a. Screen a set of 120 RILs.
    b. Identify a subset of genotypes that have extreme plasticities and a subset of genotypes that are non-plastic.
3. Ask how selection has acted on important life-history traits

   a. Estimate strength and direction of natural selection using phenotypic selection analysis.

4. Determine if there are costs associated with being phenotypically plastic.
5. Determine the genetic mechanism of phenotypic plasticity based on global gene expression patterns and QTL mapping results.

CHAPTER II – QUANTITATIVE TRAIT LOCI AND GLOBAL GENE EXPRESSION PATTERNS OFFER INSIGHT INTO THE GENETIC BASIS OF PHENOTYPIC PLASTICITY

1.0 Introduction

In heterogeneous environments the ability for a particular genotype to achieve the optimal phenotype in response to changing environmental cues is advantageous over genotypes with fixed phenotypes; the ability for a particular genotype to produce more than one phenotype is called phenotypic plasticity. If plasticity allows the organism to match its' phenotype to the optimum for each relevant environment, it may result in high fitness across the ecological conditions an organism may experience. Plasticity is therefore often assumed to be and has been demonstrated to be an adaptive evolutionary mechanism (Dorn, Hammond-Pyle, and Schmitt 2000; Dudley and Schmitt 1995 and 1996; Schmitt and Wulff 1993; Smith 1982). In order for plasticity to be adaptive, plasticity selection must differ across environments and plastic genotypes must alter their phenotype in the direction of the optimum for each environment. If plasticity is adaptive, we expect plastic genotypes to have increased fitness in two or more environment relative to non-plastic or homeostatic genotypes (Dudley and Schmitt, 1996). However, not all experimentally observed plasticity is adaptive (Dorn, Hammond-Pyle, and Schmitt 2000; Poulton and Winn 2002; Weinig 2000). Present day populations are the product of past natural selection events, therefore selection may have culled out non-adaptive plastic genotypes due to reduced fitness. Therefore, assuming that there is genetic variation for selection to act on and if selection has culled out non-adaptive forms of plasticity (non-adaptive and maladaptive) it may be difficult to identify and study plasticity in natural

populations. In addition, if no constraints (limits) or costs are associated with being plastic then individuals should evolve perfect plasticity in which they have the trait mean that confers the highest fitness in all environments; however non-adaptive and mal-adaptive plasticities persist in natural populations providing evidence that there may be constraints on the evolution of adaptive plasticity (Dechaine et al. 2007; Dewitt et al. 1998; Dorn, Hammond-Pyle, and Schmitt 2000; Relyea 2002; Stinchcombe, Dorn, and Schmitt 2004; van Kleunen, Fischer, and Schmid 2000; Weijschede et al. 2006; Weinig et al. 2004; Weinig et al. 2006).

There are three models relating to the genetic basis of phenotypic plasticity. First, the regulatory gene model predicts that environmental plasticity is caused by groups of genes, called structural genes that are expressed only in a specific environment while other groups of genes are only expressed in other environments (Scheiner and Lyman 1989 and 1991; Schlicting 1986; Schlichting & Pigliucci 1993). An alternative model argues that the same group of genes acts in all environments it is simply the amount of product each gene makes that changes with the environment or the functionality of the gene product may be environment specific. (Via 1993a and 1993b). Lastly, the overdominance model suggests that plasticity is in function of homozygosity, where individuals with high levels of homozygosisty will be highly plastic and will lose the ability to maintain a constant phenotype across differing environments. (Gillepsie & Turelli 1989).

Many studies have attempted to identify genes that contribute to phenotypic plasticity through the use of quantitative trait loci mapping with experimental mapping populations (Kliebenstein et al. 2002; Ungerer et al. 2003; Weinig et al. 2003, 2004 & 2006). Typically, identifying genes that influence quantitative traits or quantitative trait loci (QTL) is accomplished through the use of a genetic map that is generated by crossing two inbred individuals from divergent populations that differ in a phenotype of interest, that are polymorphic over regular intervals on their chromosomes. These polymorphisms act as markers that can be positioned on the chromosomes based on recombination rates among adjacent markers. Chromosomal regions with significant influence on the phenotype can them be identified by testing for associations between marker state and phenotype. However, contributing QTL to a phenotype cannot be identified using this approach if the parents of the segregating lack a polymorphism that co-segregates with the contributing gene. Furthermore, identified QTL can contain hundreds of genes and identifying candidate genes can be problematic.

Environmental conditions experienced early in the life cycle can contribute to the expression of later life-history traits. Seed stratification an important environmental cue for *Arabidopsis thaliana* for breaking seed dormancy (Baskin and Baskin 1983; Munir *et al.*2001) and can influence later phenotypes, such as flowering time (Nordborg and Bergelson 1999). Therefore, seed stratification or lack of this cue may play influence expression of traits in a single environment or plasticity in phenotype across environments.

The questions addressed in the current study are: 1) Is there genetic variation for phenotypic plasticity, 2) How has selection life-history traits, 3) Are there detectable costs associated with being phenotypic plastic, 4) What is the genetic basis of phenotypic plasticity based on QTL mapping and global gene expression patterns?

2.0 Materials and Methods

2.1 Genetic strains

Arabidopsis thaliana is a small, flowering annual weed in the mustard family (Brassicaceae) found in discrete and genetically differentiated populations in Europe, Asia and now some places in North America. We screened 120 recombinant inbred lines (RIL) of Arabidopsis thaliana developed from a cross between the natural ecotypes Cal (Calver, England; CS1062) and Tac-0 (Tacoma, Washington, USA; CS28754) in July 2005 for plasticity in days to bolting in response to cold stratification of seeds (Huang et al . 2010). The genetic structure of Recombinant Inbred Lines of (RILs) Arabidopsis thaliana is likely to emulate that of natural populations, in which rare out-crossing events occur between colonists followed by subsequent inbreeding of individuals through selfing (Jones 1971).

2.2 Pilot Study

In July 2005, 12 replicate seeds of a subset of the CalxTac RILs (106 RILs) were placed on top of soilless mix (Sunshine SB300 Universal Soilless Mix; Mfg. SunGro) in 13 98 cell flats in random arrays. The seeds were stratified by placing the flats in the

dark in a 4°C cooler for 11 days. To give the non-stratified seeds time to hydrate we planted the stratified seeds ten days after the beginning of the stratification treatment. In the non-stratified treatment 12 replicate seeds of the 120 RILs were again planted in random arrays in an additional 13 98 cell flats and immediately placed in the greenhouse at University of Wisconsin Oshkosh. The next day the stratified seeds were removed from the cold and also placed in the greenhouse. Approximately 5 days after germination, plants were thinned to one plant per cell. The average day length and temperature through the duration of the experiment were 14 hours and 24°C respectively. Plants were watered 2 times daily. Days to bolting was scored on all plants. Days to bolting is defined as the number of days from germination to differentiation of the reproductive apical meristem from the vegetative rosette. Figure 2-1A illustrates the plasticity in days to bolting for all 105 lines and Figure 2-2A illustrates the plasticities of the 3 RILs chosen for the microarray study based on their extreme phenotypic plasticity. The sensitive positive line flowered later when its seeds were cold treated (RIL30), the sensitive negative line flowered earlier when seeds were cold stratified (RIL 81) and the insensitive line flowered at the same time in both environments (RIL 43).

2.3 Environmental treatments

In July 2007, 30 replicate seeds of all 106 RILs used in preliminary study and an additional 14 RILs (120 RILs in total) were placed on top of soilless mix (Sunshine SB300 Universal Soilless Mix; Mfg. SunGro) in 37 98 cell flats in random arrays. The seeds were stratified by placing the flats in the dark in a 4°C cooler for 11 days. To give

the non-stratified seeds time to hydrate we planted the stratified seeds ten days after the beginning of the stratification treatment. In the non-stratified treatment 30 replicate seeds of the 120 RILs were again planted in random arrays in an additional 37 98 cell flats and immediately placed in the greenhouse at University of Wisconsin Oshkosh. The next day the stratified seeds were removed from the cold and also placed in the greenhouse. Following the same protocol as above, an additional 90 replicates per treatment were planted of the 3 RILs selected for the microarray study. Approximately 5 days after germination, plants were thinned to one plant per cell. The average day length and temperature through the duration of the experiment were 14 hours and 24°C. Plants were watered 2 times daily.

2.4 Phenotypic Traits

Plants were surveyed daily for the following traits: Germination Date, Bolting Date, and Flowering Date. From these dates, we calculated the following reproductive/developmental timing traits: number of Days to Bolting (Bolting Date- Germination Date), the number of Days to Flowering (Flowering Date- Germination Date) and Flowering Interval (Flowering Date- Bolting Date). As an estimate of fitness we counted the number of side and inflorescence branches and the number of siliques that fully developed on each plant after senescence.

2.5 Quantitative Genetics

Variance components within each environment were estimated using restricted maximum likelihood methods (PROC MIXED; (SAS 1999) from the following mixed model: trait= V_{Line} + $V_{Line*GH}$ + V_{error} and from this model Basic Linear Unbiased Predictors or BLUPs were calculated for each genotype within each environment. Environmental treatment significance was calculated using the following model: trait= V_{Line} + $V_{Line*TRT}$ + V_{TRT} +$V_{error.}$. These RIL BLUPs were subsequently used to estimate bivariate genotypic correlations within (r_G) and across (r_{GE}) environments (SAS PROC CORR, SAS ver.9.2).

2.6 Estimates of Selection

The number of siliques that fully developed on the side and inflorescence branches were counted as an estimate of fitness. Relative fitness was calculated by dividing the mean fitness for each genotype within each environment by the overall mean in that environment. A standard regression approach was used to estimate the strength and direction of linear and quadratic selection (Lande and Arnold 1983).

2.7 Cost of Plasticity Analysis

Cost of Plasticity analysis asks if the ability to respond to environment (degree of plasticity) affects fitness within one environment; a negative plasticity coefficients implies a cost. Using the genotypic BLUPs, a multiple regression with the following model was used to test for costs associated with plasticity: WG = XG + PlXG, where WG

is the relative fitness, XG is the genotypic mean in each environment and PlXG is the absolute value of the difference in XG across environments.

2.8 Sampling for RNA Extractions

Three to 5 days after seedling emergence, seedlings of the 3 extreme RILS were harvested while still in an early cotyledon stage (i.e. lacking true leaves). The three selected RILs were harvested in pools of 15 replicate seedlings and immediately flash frozen in liquid nitrogen (4 replicate pools per line per treatment). RNA was extracted from the cotyledon samples using Qiagen© RNeasy® Plant Mini Kit (Catalog number 74904; www1.qiagen.com) as described in the RNeasy® Mini Handbook (Fourth Edition; April 2006) and stored at -80°C.

2.9 Microarray Design and Hybridization

Arabidopsis thaliana oligonucleotide microarrays were obtained from the University of Arizona. We used a factorial loop design of 12 microarrays (Figure 2-3). There were 4 biological replicates (i.e. pools of 15 cotyledons per replicate) of each combination of RILs and 2 stratification treatments half labeled with Cy3 and half with Cy5 . Arabidopsis oligonucleotide microarrays were obtained from the University of Arizona. Hybridization was done following the protocol provided with the microarrays (University of Arizona; http://www.ag.arizona.edu/microarray) and scanning of microarrays was done on a BioRad VersaRaay scanner at the National Science

Foundation / Robert E. Moore Proteomics and Functional Genomics Core Facility at the University of Wisconsin Oshkosh.

2.10 Microarray Analysis

To determine which genes were significantly differentially expressed for each genotype across cold stratification treatments, a mixed model analysis of variance ((PROC MIXED; (SAS 1999)), where dye and array were random variables, background was a continuous variable, and genotype was a fixed variable. Means were calculated by genotype for each gene and any gene than was up- or down-regulated by 2-fold with a significance level of p-value of 0.05 or less were characterized as significantly differentially expressed.

2.11 Genetic Map and QTL Mapping

The genetic map for Cal x Tac RILs currently contains 104 genetic markers (Figure 2-4) . Details regarding genetic map creation and markers are described in Huang et al. 2010. Quantitative trait loci were mapped using the composite interval mapping (Zeng 1994) procedure of QTL Cartographer (Basten et al. 1994; Basten et al. 1999) using a walking speed of 1.0 cM. Using a forward and backward regression method (probability of 0.5) a maximum of 10 background markers was implemented under the standard model. Significant threshold for each trait was calculated by 1000 permutations.

2.12 Colocalization of QTL and Differentially Expressed Genes

2-LOD support limits for each of the 8 Days to Bolting QTL were screened for differentially expressed genes with a fold change greater than 2, identified in the microarray analysis. A differential expressed gene that fell within the 2-LOD supports for a given QTL was said to co-localize with the QTL.

3.0 Results

3.1 Quantitative Genetics

All measured traits showed significant genetic variation with the sole exception of fitness in the cold stratification treatment. Heritability for each trait ranged from ranged from 2 percent for days to germination in the cold stratification treatment to 94 percent for days to bolting also in the cold stratification treatment. All measured traits showed genotype by environment interactions (Table 2-1).

3.2 Genotypic Correlations

Days to germination was significantly positively correlated with days to flowering (0.28, 0.0016; Table 2-2). Days to bolting and flowering were negatively correlated with fitness as estimated by fruit set in both environments (Table 2-2). Days to bolting was strongly correlated with days to flowering in both environments. There were significant across environment correlations for days to bolting, days to flowering and fitness (Table 2-2).

3.3 Selection Analysis

Selection favored early germination timing in the no cold stratification treatment and short flowering intervals in the cold-stratification treatment (Table 2-3). Early bolting and flowering were favored in both the cold stratification and no cold stratification treatments (Table 2-3). Significant quadratic selection was detected for flowering interval in the cold stratification treatment and flowering time in the no cold stratification treatment (Table 2-3).

3.4 Cost of Plasticity Analysis

No significant costs of plasticity were detected for days to bolting (Figure 2-5). The regression lines had a negative trend but there was no significant correlation between plasticity and relative fitness.

3.5 QTL Mapping

QTL were identified for all traits measured except for days to germination in the cold stratification treatment and flowering interval for both treatments, explaining between 6 percent and 32 percent of the observed genetic variation (Table 2-4). Three bolting time QTL were identified in the no cold stratification treatment and two of the QTL were environment specific. Similarly, five bolting time QTL were mapped for the cold stratification treatment and 4 were specific to that environment. Four fitness QTL were identified in the no cold stratification treatment and three were environment

specific. Also, three QTL were mapped for the cold stratification treatment and 2 were specific to that environment. Four fitness QTL colocalized with QTL for bolting time on chromosomes 2, 4 and 5. All three bolting time QTL colocalized with three flowering time QTL in the no cold stratification environment and two bolting time QTL colocalized with flowering time QTL in the cold stratification treatment (Figure 2-6).

3.6 Identification of Differentially Expressed Genes

Genes were identified for all three RILs that represent different types of plasticities, significantly differentially expressed genes included genes that were up- or down- regulated 2-fold with a p-value less than 0.05. For RIL 30, which bolted early with a cold stratification treatment, 54 genes were significantly up-regulated and 21 genes were significantly down-regulated in the cold stratification treatment (Figure 2-7A). RIL 81, which bolted later with a col stratification treatment, significantly down-regulated 31 genes in the cold stratification treatment (Figure 2-7B). The homeostatic RIL 43, which maintained its' bolting time across stratification treatments, significantly up-regulated 186 genes and down-regulated 2 genes in the cold stratification treatment (Figure 2-7C). Of the differentially expressed genes, a subset of the genes were differentially expresses by more than 1 RIL (plasticity type; Figure 2-8). When comparing the proportion of genes in each molecular function class for the entire genome to those genes that were differentially expresses, there is an increase in genes whose gene products are involved with transferase, kinase, and transporter activity and nucleotide binding function (Table 2-5).

3.7 Co-localization of Mapped QTL and Differentially Expressed Genes

93 of the 294 differentially expressed genes identified in the microarray study colocalized with a bolting time QTL (Table 2-5). Not only does this provide good candidate genes for subsequent studies, but it also offers insight into the organization of the *Arabidopsis* genome and gene regulation patterns. For example, bolting time QTL1, 4, 14, 15, 16 are highly enriched in transferases and bolting time QTL (Table 2-5).

4.0 Discussion

The dimorphic germination patterns of *Arabidopsis thaliana* provides very different abiotic conditions that allow for divergent selective regimes within a single population. Typically, natural selection in fall annuals should favor later bolting time due to increased size and to ensure that they and do not initiate reproduction (bolting) prior to the favorable spring conditions. In contrast, due to the much shorter lifespan of the spring annual, selection should favor early reproduction to ensure production of seeds before the summer drought. Early bolting was favored in both environments of this experiment that mimic the spring and fall life cycle of *Arabidopsis thaliana* (Table 2-3). This is consistent with increase fitness as a mechanism to avoid the summer drought and ensure that plants produce viable seeds that will make it to the next generation, however early bolting under the fall life cycle comes with the risk of transitioning to reproduction during the middle of winter which could be potentially lethal. This departure from the predicted natural selection regime could be due to the constant photoperiod and temperature regimes imposed on the plants in the greenhouse and the lack of a vernalization treatment that

would normally be experienced by fall annuals. However, the observed selection for early bolting time under fall conditions can also be explained by a genotype level correlation between bolting and fitness (-0.3467, <0.0001) in the cold stratification treatment (simulated fall life-cycle), where the most fit bolting time cannot be reached due to the negative correlation with fitness. Furthermore, there are three genetic trade-offs that exist in the QTL architecture between timing of reproduction and fitness where allelic effects of the Cal allele decrease bolting time but increases fitness for three contributing QTL that affect both fitness and bolting time (Table 2-4). This could be the result of a pleiotropic locus that contributes to the expression of both phenotypes or it could be simple due to adjacent loci that are in strong linkage disequilibrium. Bolting and flowering time were also highly correlated in both environments (0.89, <0.0001) likely do pleiotropic effects of genes that control the timing of two or more major life history transitions or genes that are in strong linkage, again this is further supported by the fact that there are so many overlapping QTL for bolting and flowering time where the additive allelic effects are in the same direction (Table 2-4).

No costs of plasticity in bolting time were detected in either environment (Figure 2-5). One possible explanation is due to the fact that the plants were grown under greenhouse conditions and to truly detect costs you need to use segregating populations under natural settings with selective pressures. In turn, we may have observed different patterns of selection on reproductive timing and genotypes that were not phenotypically plastic or plastic in the wrong direction may have suffered from reduced fitness and the costs would have been more prevalent. This is supported by the observation that despite

the fact that early flowering was favored in both environments, plastic genotype RIL 81 was adaptive allowing maintenance of fitness across environments similar to what we observed for RIL 43 which demonstrated adaptive homeostasis was the prediction based on the identical patterns of selection on reproductive timing in both environments.

Phenotypic plasticity, i.e. changes in phenotype with environment, may allow organisms to produce optimal phenotypes in all environments (adaptive plasticity). There is considerable discussion about the genetic mechanisms for phenotypic plasticity. (Bradshaw 1965). The genetic basis of phenotypic plasticity has been a topic of discussion for the past three decades and there are three models have advanced to the forefront: the regulatory gene model predicts that environmental plasticity by the expression of environment specific (Scheiner and Lyman 1991; Schlicting 1986; Schlichting and Pigliucci 1993), the allelic sensitivity model where genes are expressed in many environments but their expression levels or functionality is environment specific (Via 1993a and 1993b), and lastly the overdominance model suggests that plasticity is in function of homozygosity (Gillepsie & Turelli 1989). In term of general gene expression patterns, the three RILs with differing plasticities all responded to the same environments in very different ways. Plastic RIL30 up-regulated a subset of genes, plastic RIL81 down regulated a subset of genes and homeostatic RIL43 up-regulated the largest subset of genes. This suggests that plasticity results from a combination of the regulatory gene and allelic substitution models. Due to the use of recombinant inbred lines in this study, which have been selfed to homozygosity over 8 generations, the overdominance models can be excluded because all RILs should have a uniformly low level of heterozygosity.

Plasticity genes are likely genes that encode sensory proteins that sense the environmental an organism is in and then coordinate the gene expression patterns of all downstream phenotypic pathways. When examining the proportion of genes in each molecular function class, there is an increase in genes whose gene products are involved with transferase, kinase, and transporter activity and nucleotide binding function in the subset of genes that were differentially expressed in response to cold stratification of seeds relative to the entire genome. A similar pattern is observed for the 93 differentially expressed genes identified in the microarray study colocalized with a bolting time QTL. Kinases and tranferases have been implemented as signaling molecules to perceive environmental changes in many previous studies (Cheong et al. 2002; Shinozaki and Yamaguchi-Shinozaki 1997).

5.0 Conclusion

The results from this study indicate that the two models for the genetic basis of phenotypic plastic may not be mutually exclusive. We identified genes that were specific to environment, specific to plasticity type and genes that were differentially expressed in more than one plasticity type. Moreover, it would appear that plasticity results from the failure to change gene expression patterns evidenced by the fact that the homeostatic genotype studied here changed expression of three times more genes that either of the phenotypically plastic genotypes. In addition, as shown here not all plasticity can be adaptive.

Ultimately this experiment will provide starting data for many other experiments examine the expression patterns of genotypes with less extreme patterns of plasticity. Using the combined QTL mapping and global gene expression patterns will identify candidate genes whose expression can be quantified in the entire set of genotypes to ask how selection has acted on the expression patterns of genes involved in phenotypic plasticity.

6.0 Acknowledgements

I would like to thank M. Yakub, V. Lor, J. Mischka, G. Weir for helping with planting and data collection. I would also like to thank the University of Arizona for providing the microarrays at a reduced cost.

This work was supported by National Science Foundation-Research Experience for Undergraduates grant to LAD, UWO small grant to LAD and UWO Faculty/ Undergraduate Grant to MJR and LAD.

Table 2-1: Estimate of variance components and heritability. z-values are presented for random effects and f-values for fixed effects. A is the within treatment estimates for the cold stratified plants, B is the within treatment estimates for the non-cold stratified plants and c is the full model to test for and effect of treatment and genotype by treatment interaction. DtoGerm: Days to Germination, DtoBolt: Days to Bolting, Flrint: Flowering Interval, DtoFlr: Days to Flowering, and Fitness as estimated by fruit number. Significance thresholds: * : <0.05, **: <0.01, ***: <0.001, and ****: <0.0001.

A

	Cold			
	V_{line}	$V_{LinexGH}$	V_{Error}	h^2
DtoGerm	2.79**	0.15	40.38****	0.02
DtoBolt	7.7****	0.54	40.00****	0.94
FlrInt	7.5****	0.56	37.91****	0.60
DtoFlr	7.68****	NE	38.68****	0.91
Fitness	0.62	5.51****	37.73****	0.01

B

	No Cold			
	V_{line}	$V_{LinexGH}$	V_{Error}	h^2
DtoGerm	6.05****	NE	39.58****	0.12
DtoBolt	7.65****	NE	38.98****	0.81
FlrInt	7.26****	NE	37.17****	0.42
DtoFlr	7.65****	NE	37.16****	0.91
Fitness	5.77****	2.93**	35.83****	0.05

C

Random Effects			Fixed Effect
V_{line}	$V_{LinexTRT}$	V_{Error}	V_{TRT}
NE	7.41****	57.01****	1.33
3.81****	7.66****	56.38****	3.54
NE	10.47****	53.65****	0.37
3.44***	7.63****	53.64****	5.91*
5.12****	4.16****	53.15****	17.01****

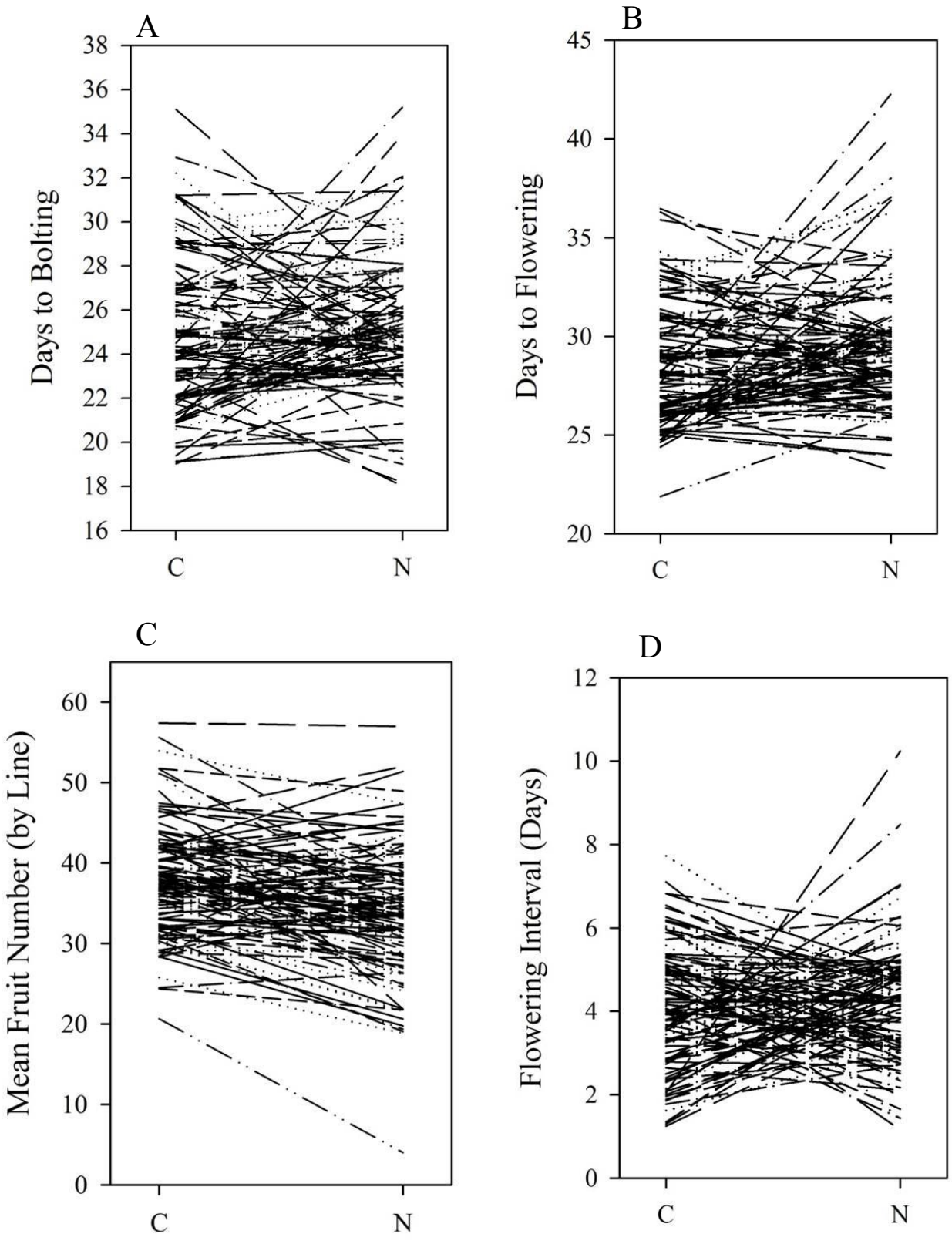


Figure 2-1: Reaction Norms illustrating the strength and direction of plasticity. C = stratification treatment, N= no cold stratification treatment.

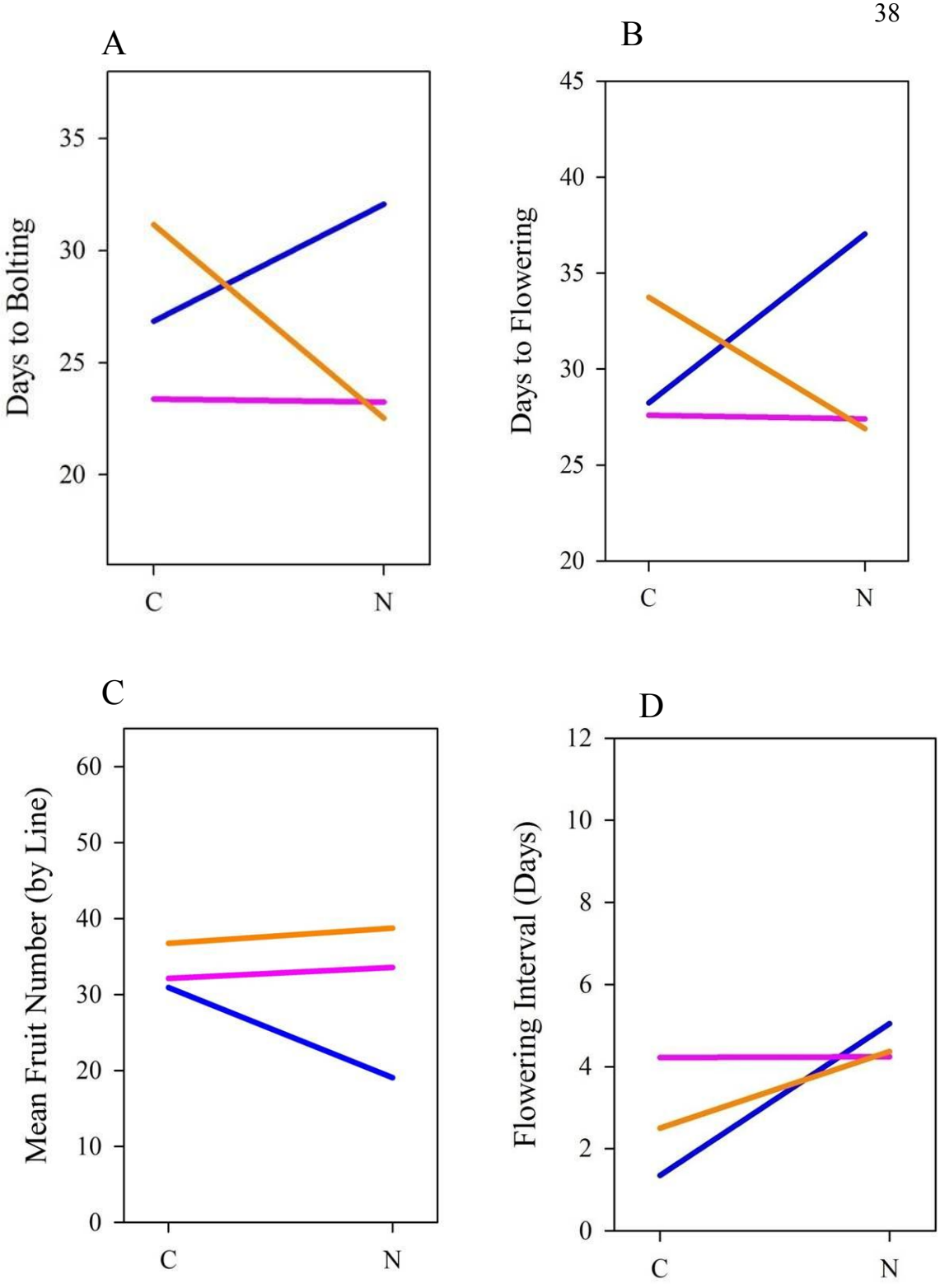


Figure 2-2: Reaction Norms illustrating the strength and direction of plasticity for the three selected RILs. C = stratification treatment, N= no cold stratification treatment. Environmental sensitivity is calculated by substracting the mean phenotype for a RIL in the N from the mean phenotype in the C. Orange is sensitive positive RIL 81, Blue is the sensitive negative RIL30, and Pink is the homeostatic RIL 43.

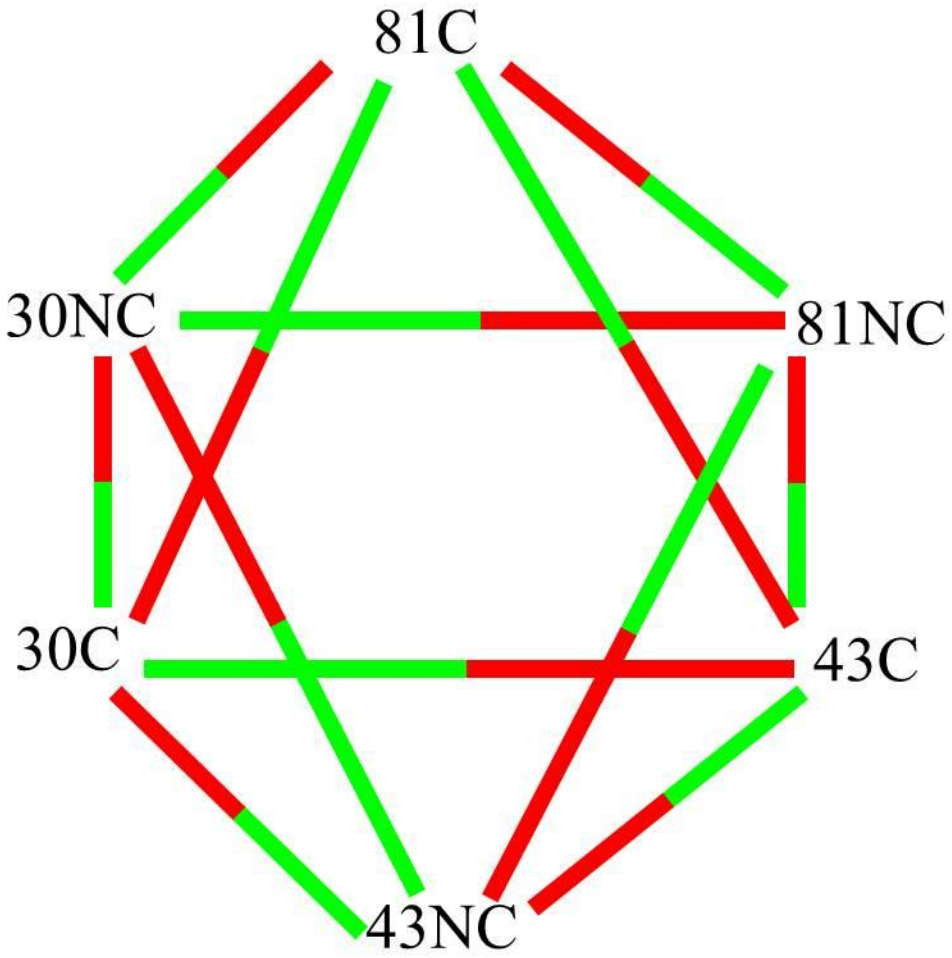


Figure 2-3: Factorial loop design for the microarray study where each genotype is represented four times per environment labeled twice with red dye and twice with green to achieve an orthogonal experimental design. Each Red/Green bar represents aone of the twelve microarrays. The colors are indicative of the dye used to label each sample per microarray.

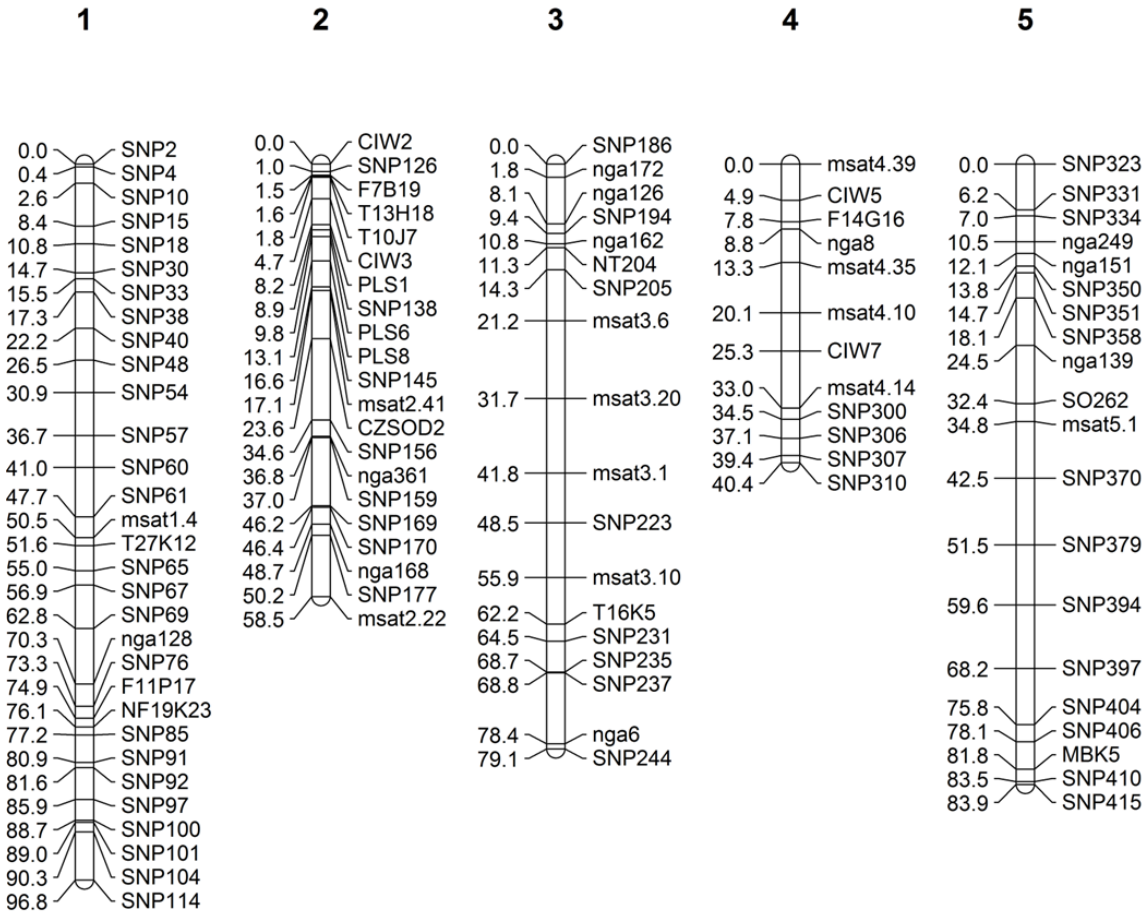


Figure 2-4: Cal x Tac RIL genetic map. Numbers 1-5 represent the 5 chromosomes of A. thaliana. Numbers to the left of each chromosome and cM distances based on recombination frequencies. Names to the right of each chromosome are the genetic markers.

Table 2-2: Genotypic correlations for life-history traits and fitness. Correlation coefficients are the top value and the p-value significance values are below them. The bolded diagonal values are the across stratification treatment correlations. Above the diagonal are the within treatment correlations for the cold stratification treatment. Below the diagonal are the within treatment correlations for the non-cold stratification treatment.

	DtoGerm	DtoBolt	FlrInt	DtoFlr	Fitness
DtoGerm	**-0.6560**	-0.0927	0.1220	-0.0461	0.1263
	0.4747	0.3117	0.1824	0.6156	0.1673
DtoBolt	0.2233	**0.3730**	-0.4942	0.8986	-0.3467
	0.0130	**<.0001**	<.0001	<.0001	<.0001
FlrInt	0.1362	-0.1763	**-0.1190**	-0.0630	0.1855
	0.1330	0.0512	**0.1935**	0.4923	0.0417
DtoFlr	0.2823	0.8992	0.2551	**0.3332**	-0.3055
	0.0016	<.0001	0.0044	**0.0002**	0.0007
Fitness	0.1366	-0.4154	-0.1115	-0.4635	**0.5445**
	0.1318	<.0001	0.2195	<.0001	**<.0001**

Table 2-3: Estimates of genotypic selection by stratification treatment. Cold= cold stratification treatment and No cold= no stratification treatment. β represents linear selection and ϒ represents quadratic selction terms. DtoGerm: Days to Germination, DtoBolt: Days to Bolting, Flrint: Flowering Interval, DtoFlr: Days to Flowering, and Fitness as estimated by fruit number. Significance thresholds: * : <0.05, **: <0.01, ***: <0.001, and ****: <0.0001.

Trt	Trait	β (SE)	γ (SE)	R^2
Cold	DtoGerm	0.013 (0.015)	0.002 (0.005)	0.02
	DtoBolt	-0.048 (0.013)***	0.004 (0.009)	0.12
	FlrInt	0.025 (0.011)*	-0.026 (0.010)**	0.09
	DtoFlr	-0.039 (0.012)**	-0.001 (0.009)	0.09
No Cold	DtoGerm	0.072 (0.033)*	0.006 (0.004)	0.04
	DtoBolt	-0.077 (0.016)****	0.001 (0.009)	0.17
	FlrInt	-0.012 (0.018)	-0.005 (0.008)	0.02
	DtoFlr	-0.107 (0.018)****	0.017 (0.008)*	0.24

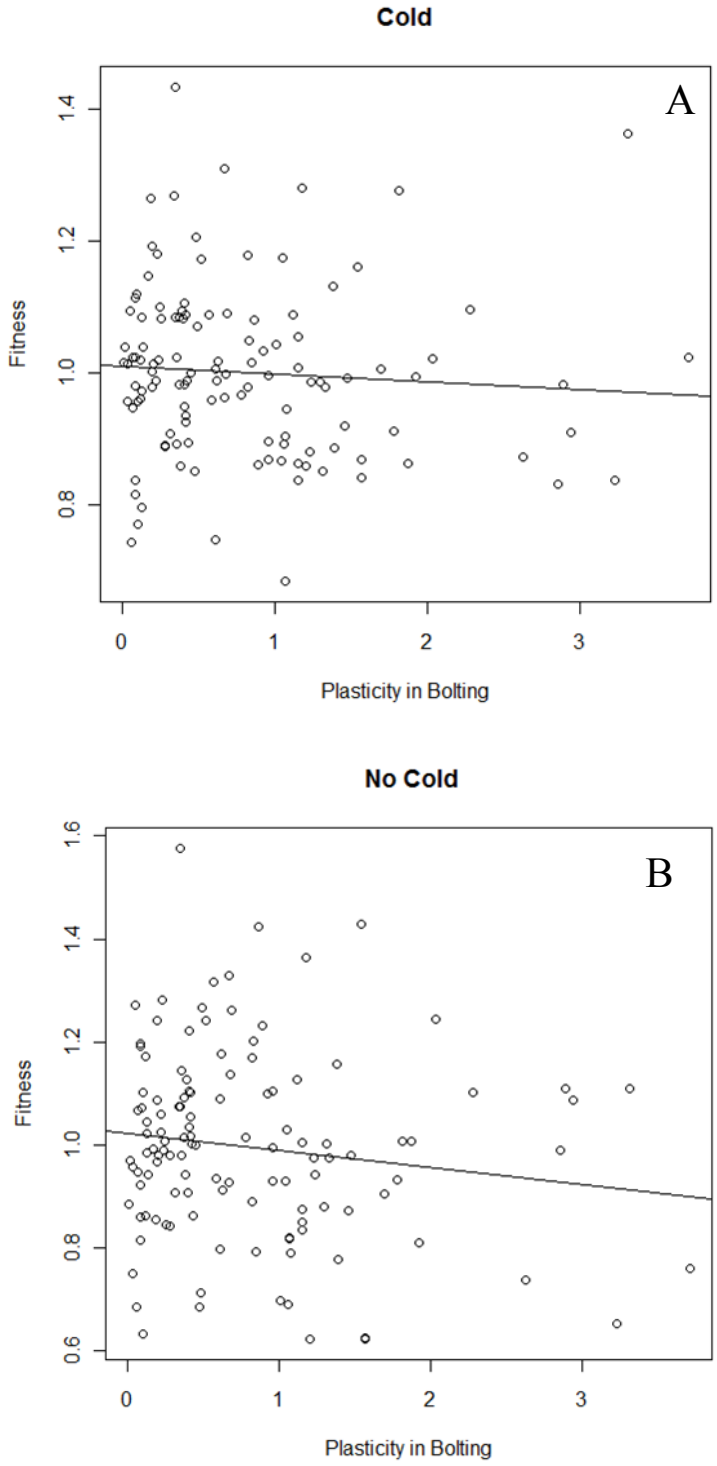


Figure 2-5: Cost of plasticity scatter plots. A is testing for an association between plasticity in bolting time and relative fitness in the col stratification treatment and B is for the no cold stratification treatment. Plasticity in days to bolting is on the x axis and is calculated by subtracting the mean phenotype for a RIL in the no cold from the mean phenotype in the cold.

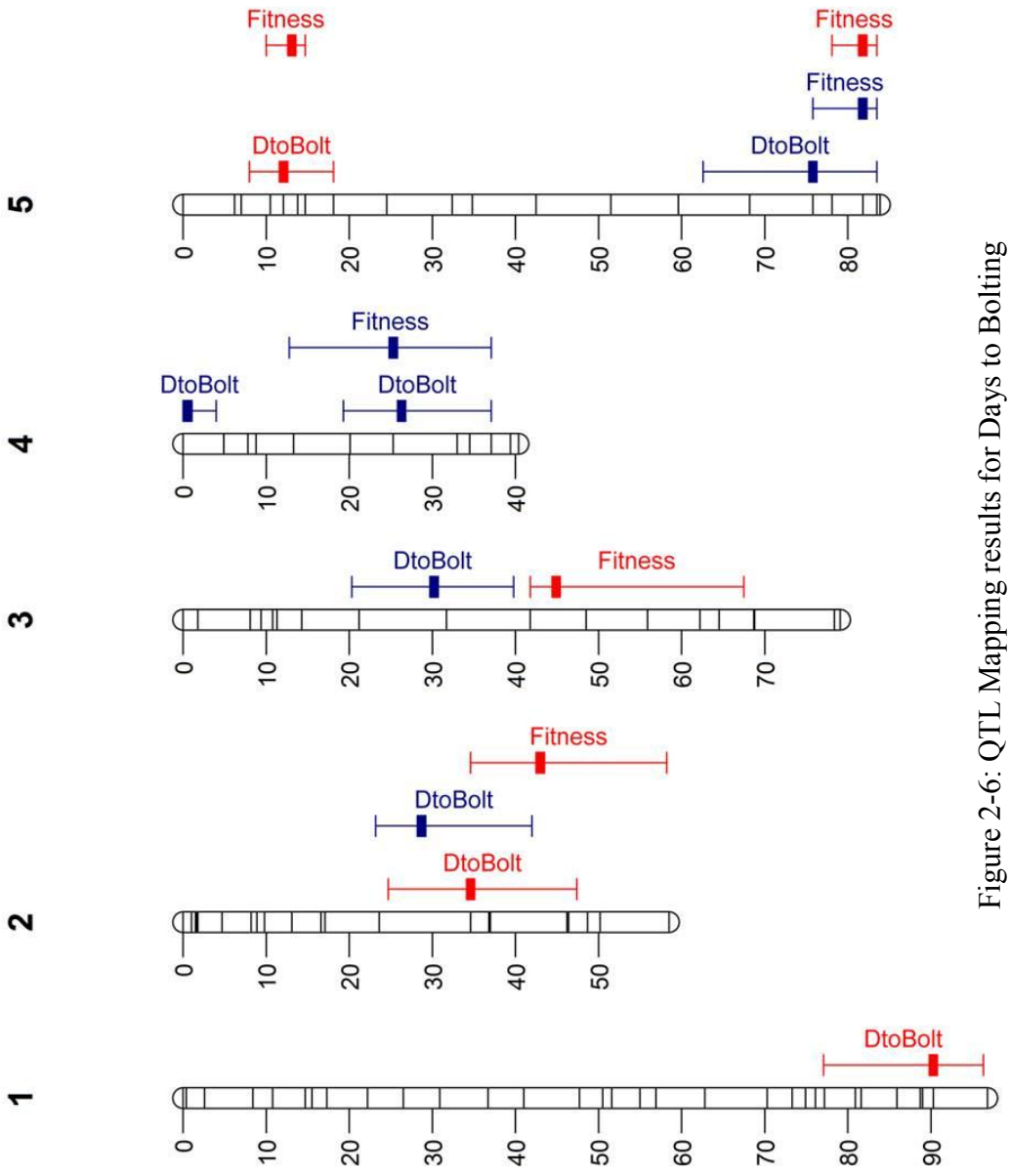


Figure 2-6: QTL Mapping results for Days to Bolting and Fitness. Numbers 1-5 represent chromosomes of A. thaliana. Numbers to the left of each chromosome and cM distances based on recombination frequencies. Blue QTL are from the cold stratification the 5 treatment and red are from the no cold stratification treatment. Bold bar within each QTL represent the QTL peak and the whiskers are the 2-LOD support limits.

Table 2-4: Mapped QTL for Life-history traits and Fitness. QTL # is an arbitrary number given to each QTL. Traits are abbreviated as: DtoGerm: Days to Germination, DtoBolt: Days to Bolting, Flrint: Flowering Interval, DtoFlr: Days to Flowering, and Fitness as estimated by fruit number. Chr: Chromosome the QTL is located on. Significance threshold was calculated for each trait by 1000 permutations. Additive effect indicated the effect of the Cal allele. Stnd. Add. Effect is the additive effect standardized by the trait variance. r2 is the percent variance explained by the QTL.

QTL #	Trt	Trait	Chr.	QTL map position (cM) (closest marker)	cM range of 2-LOD support limits (closest markers)	Sign. threshold	Likelihood ratio	Additive effect	Stnd. Add. effect	r2
1	C	DtoBolt	2	28.65 (CZSOD2)	23.15-42.02 (CZSOD2, SNP169)	11.39	11.24	-1.073	-0.099	0.07
2	C	DtoBolt	3	30.16 (msat3.20)	20.27-39.75 (msat3.6, msat3.1)	11.39	19.98	1.206	0.111	0.11
3	C	DtoBolt	4	0.01 (msat4.39)	0.01-4.01 (msat4.39, CIW5)	11.39	12.91	-1.069	-0.099	0.07
4	C	DtoBolt	4	25.31 (CIW7)	19.29-37.08 (msat4.10, SNP306)	11.39	13.63	-1.075	-0.099	0.07
5	C	DtoBolt	5	75.83 (SNP404)	62.57-83.53 (SNP394, SNP410)	11.39	18.55	-1.070	-0.099	0.09
6	C	DtoFlr	3	64.5 (SNP231)	47.85-76.79 (SNP223, nga6)	11.46	12.16	0.779	0.094	0.06
7	C	DtoFlr	4	23.09 (CIW7)	19.29-40.43 (msat4.10, SNP310)	11.46	11.34	-0.955	-0.116	0.07
8	C	DtoFlr	5	11.53 (nga151)	6.01-20.11 (nga151, SNP358)	11.46	14.32	1.046	0.127	0.08
9	C	DtoFlr	5	71.26 (SNP397)	54.46-81.77 (SNP379, MBK5)	11.46	16.51	-0.986	-0.119	0.10
10	C	Fitness	4	25.31 (CIW7)	12.79-37.08 (msat4.35, SNP306)	11.37	11.33	0.114	0.005	0.07
11	C	Fitness	4	34.5 (SNP300)	12.79-37.09 (msat4.35, SNP306)	11.37	10.74	0.100	0.004	0.07
12	C	Fitness	5	81.77 (MBK5)	75.83-83.53 (SNP404, SNP410)	11.37	24.99	0.157	0.007	0.17
13	NC	DtoGerm	4	21.09 (msat4.10)	13.29-33.02 (msat4.35, msat4.14)	11.59	11.76	-0.069	-0.529	0.10
14	NC	DtoBolt	1	90.26 (SNP104)	77.09-96.26 (SNP85, SNP114)	11.75	13.38	-0.867	-0.096	0.08
15	NC	DtoBolt	2	34.58 (SNP155)	24.65-47.37 (CZSOD2, SNP170)	11.75	13.75	-0.833	-0.092	0.08
16	NC	DtoBolt	5	12.08 (nga151)	8.03-18.11 (SNP334, SNP358)	11.75	27.30	1.483	0.164	0.18
17	NC	DtoFlr	1	90.26 (SNP104)	85.92-96.26 (SNP97, SNP114)	11.04	13.91	-0.773	-0.081	0.07
18	NC	DtoFlr	2	36.85 (nga361)	23.65-46.02 (CZSOD2, SNP169)	11.04	13.19	-0.725	-0.076	0.06
19	NC	DtoFlr	5	12.08 (nga151)	10.53-15.72 (nga249, SNP351)	11.04	53.39	1.959	0.204	0.32
20	NC	Fitness	2	43.02 (SNP169)	34.58-58.17 (SNP155, msat2.22)	11.4	16.33	0.149	0.004	0.08
21	NC	Fitness	3	44.85 (msat3.1)	41.75-67.5 (msat3.1, SNP235)	11.4	24.22	-0.180	-0.004	0.11
22	NC	Fitness	5	13.08 (SNP350)	10.03-14.72 (nga249, SNP351)	11.4	53.52	-0.370	-0.009	0.28
23	NC	Fitness	5	81.77 (MBK5)	78.11-83.53 (SNP406, SNP410)	11.4	25.35	0.178	0.004	0.10

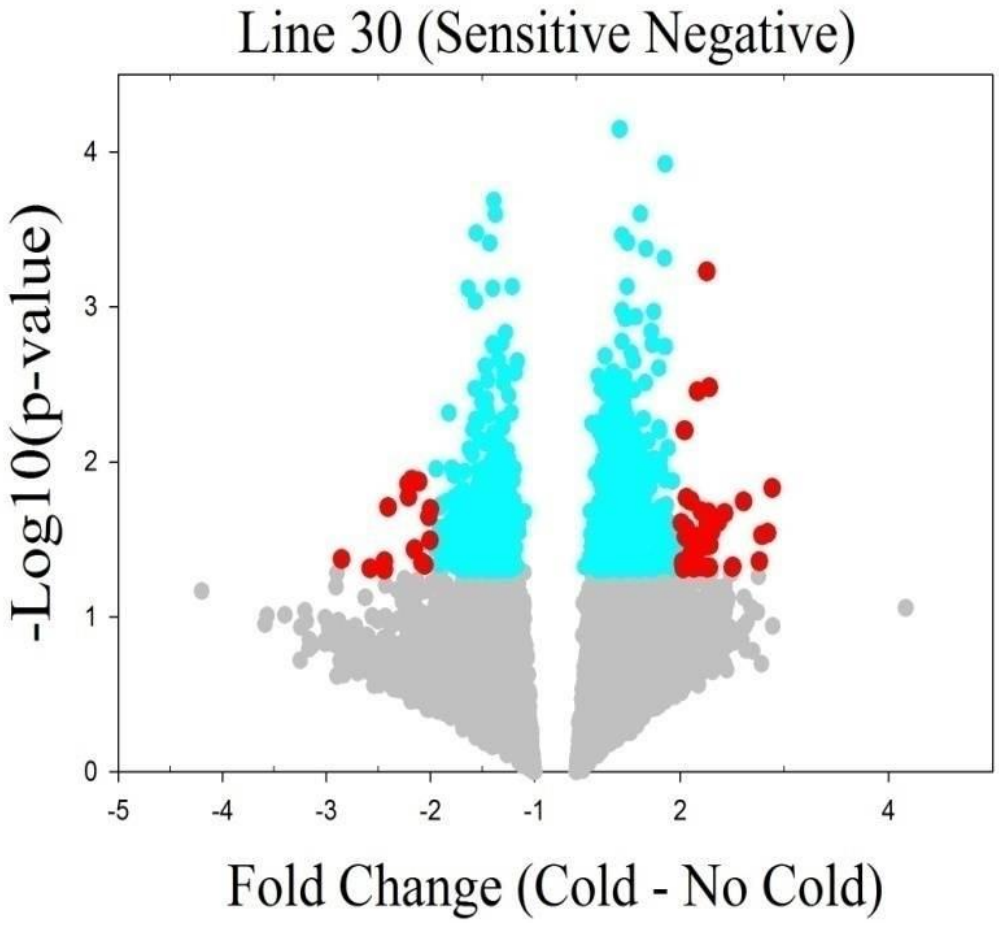


Figure 2-7A: Volcano plots illustrating overall patterns of differential gene expression in the RIL30 in response to cold stratification. The y-axis is negative log value of the p-value and the x-axis is the log2 value of the fold change (cold-no cold). Gray dots indicates a gene that is not significantly expressed, teal dots indicate genes that are significantly differentially expressed with $p<0.05$, and red dots indicate the significantly differentially expressed genes ($p<0.05$) that have a 2-fold change in either direction.

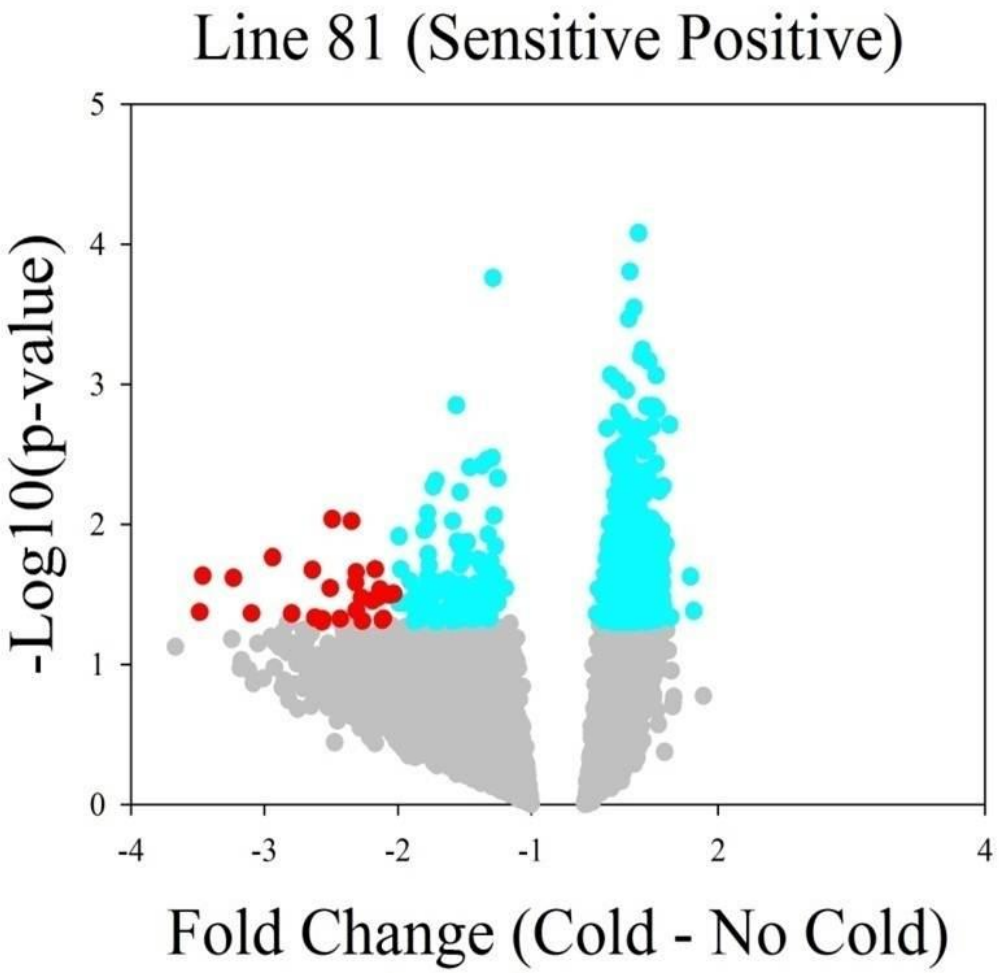


Figure 2-7B: Volcano plots illustrating overall patterns of differential gene expression in the RIL81 in response to cold stratification. The y-axis is negative log value of the p-value and the x-axis is the log2 value of the fold change (cold-no cold). Gray dots indicates a gene that is not significantly expressed, teal dots indicate genes that are significantly differentially expressed with $p<0.05$, and red dots indicate the significantly differentially expressed genes ($p<0.05$) that have a 2-fold change in either direction.

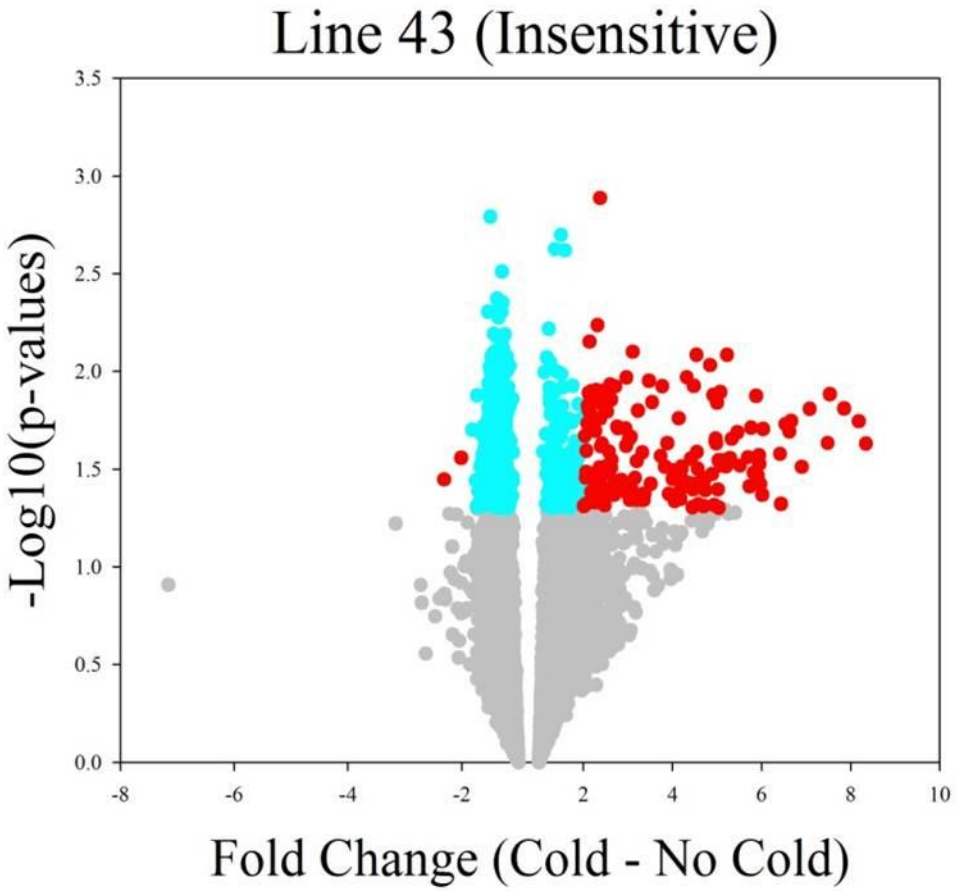
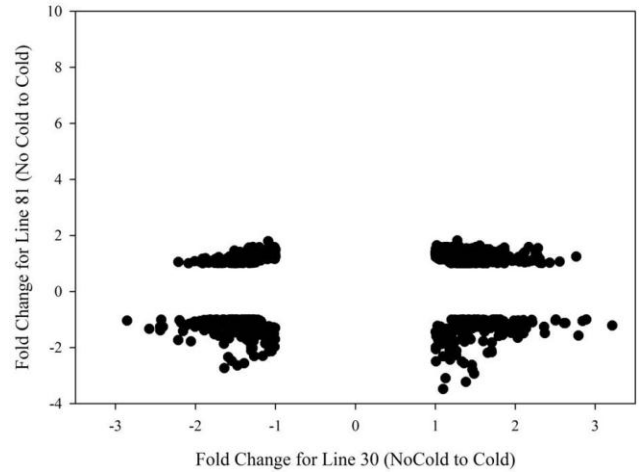
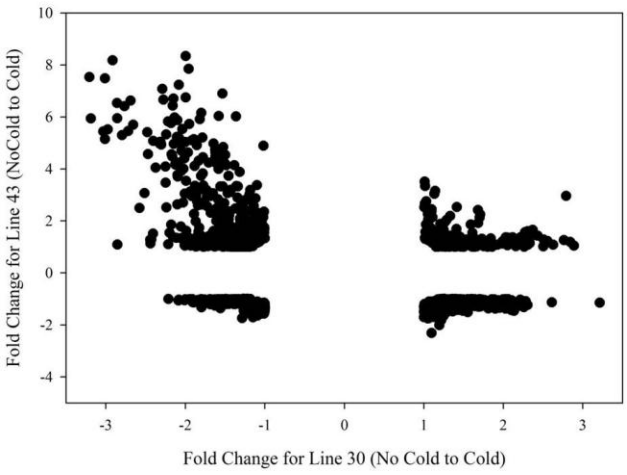


Figure 2-7C: Volcano plots illustrating overall patterns of differential gene expression in the RIL43 in response to cold stratification. The y-axis is negative log value of the p-value and the x-axis is the log2 value of the fold change (cold-no cold). Gray dots indicates a gene that is not significantly expressed, teal dots indicate genes that are significantly differentially expressed with $p<0.05$, and red dots indicate the significantly differentially expressed genes ($p<0.05$) that have a 2-fold change in either direction.





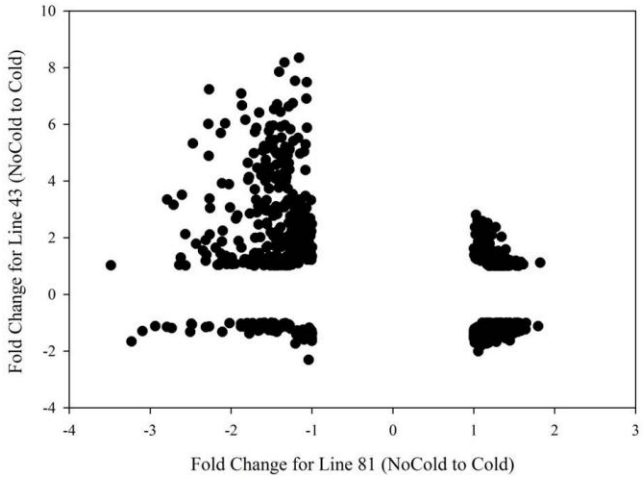


Figure 2-8: Pairwise scatter plots of differentially expressed genes by genotype. Expression levels for all significantly expressed genes were plotted by genotype pairs to illustrate how the subset of genes related by one RIL related to the other RILs.

Table 2-5: Molecular characterization of differentially expressed genes and colocalization of differentially expressed genes from the microarray study and mapped QTL for bolting time. All numbers represent percentages of the genes that fall into each molecular class.

Molecular Function	Genome	All Sig	SigFold2	All QTL	QTL1	QTL15	QTL3	QTL4	QTL5	QTL14	QTL2	QTL16
unknown molecular function	30.12	21.046	20.32	12.644	7.407	7.407	NA	15.385	20.833	7.143	16.667	17.143
Other Binding	10.589	11.126	8.904	8.046	0	0	NA	7.692	8.333	10.714	0	11.429
Other Enzyme Activity	8.657	9.875	8.904	6.322	3.704	3.704	NA	0	0	16.071	0	2.857
Transferase Activity	7.779	9.316	9.361	12.069	18.519	18.519	NA	19.231	4.167	10.714	0	11.429
Hydrolase Activity	6.372	7.685	7.763	10.345	3.704	3.704	NA	11.538	20.833	7.143	33.333	8.571
DNA or RNA Binding	5.666	5.228	5.708	4.598	3.704	3.704	NA	3.846	8.333	1.786	16.667	5.714
Protein Binding	5.61	6.501	6.621	8.046	3.704	3.704	NA	11.538	8.333	10.714	0	5.714
Kinase Activity	5.523	7.038	7.991	13.793	25.926	25.926	NA	15.385	8.333	8.929	0	17.143
Nucleotide Binding	4.796	6.077	7.991	9.77	11.111	11.111	NA	7.692	8.333	10.714	0	11.429
Transporter Activity	3.67	4.133	4.566	4.598	3.704	3.704	NA	3.846	0	5.357	33.333	2.857
Transcription Factor Activity	3.327	4.178	3.425	1.724	3.704	3.704	NA	0	4.167	1.786	0	0
Nucleic Acid Binding	3.199	2.234	3.196	2.874	7.407	7.407	NA	0	4.167	1.786	0	2.857
Other Molecular Function	3.151	3.887	3.653	2.874	3.704	3.704	NA	0	4.167	5.357	0	0
Structural Molecule Activity	1.066	1.139	0.913	1.149	3.704	3.704	NA	0	0	0	0	2.857
Receptor Binding or Activity	0.476	0.536	0.685	1.149	0	0	NA	3.846	0	1.786	0	0
Total Number of Genes	33,597	3293	294	93	11	12	0	17	14	24	5	21

REFERENCES

Baskin, J.M. and C.C. Baskin. 1983. Seasonal changes in the germination responses of buried seeds of *Arabidopsis thaliana* and ecological interpretation. *Botany Gazette* 144:9 540-543.

Basten, C. J., B. S. Weir and Z.-B. Zeng. 1994. Zmap-a QTL Cartographer, pp. 65-66 in *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production: Computing Strategies and Software*, edited by W. Fairfull. Organizing Committee, 5th World Congress on Applied Genetics Applied to Livestock Production, Guelph, Ontario.

Basten, C. J., B. S. Weir and Z.-B. Zeng. 1999. QTL Cartographer, Version 1.13.

Bradshaw, A. D. 1965. Evolutionary significance of phenotypic plasticity in plants. *Advances in Genetics* 13:115-155.

Cheong , Y.H. , H.S. Chang, R. Gupta, X. Wang, T. Zhu, and S. Luan. 2002. Transcriptional Profiling Reveals Novel Interactions between Wounding, Pathogen, Abiotic Stress, and Hormonal Responses in Arabidopsis. *Plant Physiol*. 129:661-677.

Churchill, G. A. 2002. Fundamentals of experimental design for cDNA microarrays. *Nature Genetics* 32:490-495.

Dechaine, J.M., J.A. Johnston, M.T. Brock, and C. Weinig. 2007. Constraints on the evolution of adaptive plasticity: costs of plasticity to density are expressed in segregating progenies. *New Phytologist* 176:874-882.

Dewitt, T.J., A. Sih , and D.S. Wilson. 1998. Costs and limits of phenotypic plasticity. *Trends in Ecology and Evolution* 13:77–81.

Donohue, K. 2002. Germination timing influences natural selection on life-history characters in *Arabidopsis thaliana*. *Ecology* 83(4): 1006-1016.

Dorn, L. A., E. Hammond-Pyle, and J. Schmitt. 2000. Plasticity to light cues and resources in *Arabidopsis thaliana*: testing for adaptive value and costs. *Evolution* 54:1982-1994.

Dudley, S. A., and J. Schmitt. 1995. Genetic differentiation in morphological responses to simulated foliage shade between populations of *Impatiens capensis* from open and woodland sites. *Functional Ecology* 9:655-666.

---. 1996. Testing the adaptive plasticity hypothesis: density-dependent selection on manipulated stem length in *Impatiens capensis*. *The American Naturalist* 147: 445-465.

Galbraith, D. 2010. University of Arizona. http://ag.arizona.edu/microarray/

Gilad, Y., and J. Borevitz. 2006. Using DNA microarrays to study natural variation. *Current Opinions in Genetics and Development* 16:553-558.

Gillespie, J. H. and M. Turelli. 1989. Genotype-environment interaction and the maintenance of polygenic variation. *Genetics* 121:129-138.

Girke, T., J. Todd, S. Ruuska, J. White, C. Benning, and J. Ohlrogge. 2000. Microarray analysis of developing *Arabidopsis* seeds. *Plant Physiology* 124:1570-1581.

Huang, D.Q., W.R. Wu, S.R. Abrams, and A.J. Cutler. 2008. The relationship of drought-related gene expression in *Arabidopsis thaliana* to hormonal and environmental factors. *Journal of Experimental Botany* 59:2991-3007.

Jones, M. E. 1971. The population genetics of *Arabidopsis thaliana*.II. Population structure. *Heredity* 27:51-58.

Kliebenstein, D., D. Pedersen, B. Barker, and T. Mitchell-Olds. 2002. Comparative Analysis of Quantitative Trait Loci Controlling Glucosinolates, Myrosinase and Insect Resistance in *Arabidopsis thaliana*. *Genetics* 161:325-332.

Lande, R. and S.J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37(6):1210-1226.

Meyerowitz, E. 1987. *Arabidopsis thaliana*. *Annual Reviews in Genetics* 21:93-111.

Michael, T.P., Mockler, T.C., Breton, G., McEntee, C., Byer, A., Trout, J.D., Hazen, S.P., Shen, R., Priest, H.D., Sullivan, C.M., Givan, S.A., Yanovsky, M., Hong, F., Kay, S.A., and J.Chory. 2008. Network Discovery Pipeline Elucidates Conserved Time-of-Day-Specific cis-Regultory Modules. *PLoS Genetics* 4(2):1-17.

Meyerowitz, E. M., and R. E. Pruitt. 1985. *Arabidopsis thaliana* and plant molecular genetics. *Science* 229:1214-1218.

Munir, J., L.A. Dorn, K. Donohue, and J. Schmitt. 2001. The effect of maternal photoperiod on seasonal dormancy in *Arabidopsis thaliana* (Brassicaceae). *American Journal of Botany* 88:1240-1249.

Nordborg, M., and J. Bergelson. 1999. The effect of seed and rosette cold treatment on germination and flowering time in some *Arabidopsis thaliana* (Brassicaceae) ecotypes. *American Journal of Botany* 86:470-475.

Poulton, J. and A. A. Winn. 2002. Costs of canalization and plasticity in response to neighbors in *Brassica rapa*. *Plant Species Biology* 17:109-118.

Relyea, Rick A. 2002. Costs of phenotypic plasticity. *American Naturalist* 159(3):272-282.

Rosenzweig, B. A., P. S. Pine, O. E. Domon, S. M. Morris, J. J. Chen, and F. D. Sistare. 2004. Dye-bias correction in dual-labeled cDNA microarray gene expression measurements. *Environmental Health Perspectives*. 112:480-487.

SAS. 1999. *SAS/STAT User's Guide*. Cary, NC: SAS Institute.

Scheiner, S.M. and R.F. Lyman. 1989.Genetics of phenotypic plasticity. I. Heritability. *Journal of Evolutionary Biology* 2:95-107.

---. 1991. The genetics of phenotypic plasticity. II. Response to selection. *Journal of Evolutionary Biology* 4:23-50.

Schlicting, C. D. 1986. The evolution of phenotypic plasticity in plants. *Annual Review of Ecology and Systematics* 17:667-693.

Schlicting, C. D., and M. Pigliucci. 1993. Control of phenotypic plasticity via regulatory genes. *American Naturalist* 142:366-370.

Schmitt, J., and R. D. Wulff. 1993. Light spectral quality, phytochrome and plant competition. *Trends in Ecology and Evolution* 8:47-51.

Seki, M., M. Narusaka, H. Abe, M. Kasuga, K. Yamaguchi-Shinozaki, P. Carninci, Y. Hayashizaki, and K. Shinozaki. 2001. Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. *Plant Cell* 13:61-72.

Seki, M., M. Narusaka, J. Ishids, T. Nanjo, M. Fujita, Y. Oono, A. Kamiya, M. Nakajima, A. Enju, T. Sakurai, M. Satou, K. Akiyama, T. Taji, K. Yamaguchi-Shinozaki, P. Carninci, J. Kawai, Y. Hayashizaki, and K. Shinozaki. 2002. Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold, and high-salinity stresses using a full-length cDNA microarray. *Plant Journal* 31:279-292.

Shinozaki, K. and K. Yamaguchi-Shinozaki. 1997. Gene Expression and Signal Transduction in Water-Stress Response. *Plant Physiol*. 115(2):327-334.

Smith, H. 1982. Light quality, photoreception, and plant strategy. *Annual Review of Plant Physiology* 33: 481-518.

Stinchcombe, J.R., L.A. Dorn, and J. Schmitt. 2004. Flowering time plasticity in *Arabidopsis thaliana:* a reanalysis of Westerman & Lawrence (1970). *Journal of Evolutionary Biology* 17:197-207.

Sultan, S.E. and H.G. Spencer. 2002. Metapopulation structure favors plasticity over local adaptation. *Interg. Comp. Biology* 43:459-469.

Ungerer, M. C., S. S. Halldorsdottir, M. D. Purugganan, and T. F. C. Mackay. 2003. Genotype-Environment Interactions at Quantitative Trait Loci Affecting Inflorescence Development in *Arabidopsis thaliana*. *Genetics* 165:353-365.

van Kleunen, M. and M. Fischer**.** 2007**.** Progress in the detection of costs of phenotypic plasticity in plants. *New Phytologist* 176:727-730.

Van Tienderen, P. H. 1991. Evolution of generalists and specialists in spatially heterogeneous environments. *Evolution* 45 (6):1317-1331.

Van Tienderen, P. H. 1997. Generalists, specialists, and the evolution of phenotypic plasticity in sympatric populations of distinct species. *Evolution* 51 (5):1372-1380.

Via, S. 1993a. Adaptive phenotypic plasticity: Target or by-product of selection in a variable environment? *American Naturalist* 142:352-365.

---. 1993b. Regulatory genes and reaction norms. *American Naturalist* 142:374-378.

Via, S., R. Gumulkiewicz, G. De Jong, S.M. Scheiner, C.D. Schlichting and P.H. Van Tienderen. 1995. Adaptive phenotypic plasticity: consensus and controversy. *TREE* 10(5):212-217.

Via, S. and R. Lande 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* 39:505-522.

---. 1987. Evolution of genetic variability in a spatially heterogeneous environment: effects of genotype-environment interaction. *Genetical Research* 49:147-156.

Weijschede, J., J. Martinkova, H. de Kroon, and H. Huber. 2006. Shade avoidance in *Trifolium repens:* costs and benefits of plasticity in petiole length and leaf size. *New Phytologist* 172:655-666.

Weinig, C. 2000. Plasticity versus canalization: population differences in the timing of shade-avoidance responses. *Evolution* 54:441-451.

Weinig, C., K.A. Graveur, N.C. Kane, and J. Schmitt. 2004. Testing adaptive plasticity to UV: Costs and benefits of stem elongation and light-induced phenolics. *Evolution* 58:2645-2656.

Weinig, C., L.A. Dorn, N.C. Kane, Z.M. German, S.S. Halldorsdottir, M.C. Ungerer, Y. Toyonaga, T.F.C. Mackay, M.D. Purugganan and J. Schmitt. 2003. Heterogeneous selection at specific loci in natural environments in *Arabidopsis thaliana*. *Genetics* 165:321-329.

Weinig, C., Z.M. German, L.M. Demink, and J.A. Johnston. 2006. Local and global costs of adaptive plasticity. *American Naturalist* 126:826-836

Yang, H., Y. Li, and J. Hua. 2006. The C2 domain protein BAP1 negatively regulates defense responses in *Arabidopsis*. *The Plant Journal* 48:238-248.

Yang, H., J. McLeese, M. Weisbart, J. L. Dionne, I. Lemaire, and R. A. Aubin. 1993. Simplified high throughput protocol for Northern hybridization. *Nucleic Acids Research* 21:3337-3338.

Yang, Y. H., and T. Speed. 2002. Design issues for cDNA microarray experiments. *Nature Review Genetics* 3:579-588.