MECHANISMS OF ADAPTATION IN *ORYZA* AND *ARABIDOPSIS*

By

Michael A. Grillo

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Plant Biology – Doctor of Philosophy
Ecology, Evolutionary Biology & Behavior – Dual Major

2013

ABSTRACT

MECHANISMS OF ADAPTATION IN *ORYZA* AND *ARABIDOPSIS*

By

Michael A. Grillo

Adaptation by natural selection is a driving force in evolution. Understanding the ecological and genetic basis of adaptation to contrasting environments has been a major focus of biology since the origins of evolutionary thought. While significant gains have been made in our understanding of this process, the genetic basis of adaptation in particular remains poorly understood. To uncover the mechanisms of adaptation we ultimately need to identify the genes and mutations underlying adaptive traits and experimentally demonstrate their effect on fitness.

During the past century there have been noteworthy theoretical advances in understanding the genetic basis of evolution. The earliest theory dates back to Fisher's geometric model which demonstrates that infinitesimally small mutations have a 50% chance of being beneficial, whereas this probability quickly diminishes with mutations of increasing effect size. Thus, Fisher concluded that mutations of small effect are the raw material for adaptation; a viewpoint that aligned well with the long-standing 'micromutational' view of adaptation. Several decades later Kimura reexamined this work and pointed out that while small mutations are more likely to be beneficial, they are also highly prone to stochaistic loss by drift. Taking both factors into account, Kimura suggested that intermediate size mutations are most likely to contribute to adaptation. More recently, Orr revisited this work and identified that while Kimura's prediction is true for a particular step in adaptation it does not explain the distribution of mutations throughout an adaptive walk. Using analytical theory and computer simulations,

Orr has demonstrated that the effect sizes of beneficial mutations should follow an exponential distribution, with fewer mutations of large effect fixed early on and mutations of decreasing sizes fixed subsequently as populations approach the fitness optimum.

Although understanding the genetic basis of adaptation has been a long standing goal of evolutionary biology, data on the subject has for the most part been limiting. The advent of Quantitative Trait Loci (QTL) mapping, in particular, has provided a means for obtaining rigorous such data. QTL mapping has been primarily useful for characterizing the genetic architecture of phenotypic variation; that is estimating the number, effect size, and location of loci controlling trait variation. The most noteworthy outcome of this line of research has been the regular identification of large effect QTL. It should be noted that much of this research has been conducted in model systems or domesticated plants and animals, although studies in natural systems are beginning to accumulate.

Here I present a dissertation aimed at understanding the mechanisms of adaptation in two wild rice species and locally adapted populations of *Arabidopsis thaliana*. First, I assess the genetic architecture of adaptation in the wild progenitors of cultivated rice, by identifying QTL for a number of putative adaptive traits. Through this work flowering time was revealed as a key adaptation for habitat preference between these species. In the next chapter I attempt to elucidate the genetic basis of a major flowering time QTL through fine mapping. I continue my examination of flowering time genetics by examining the genetic basis of flowering time differentiation between locally adapted populations of *Arabidopsis thaliana*. Finally, I conduct a thorough study of comparative floral biology to identify key traits that control mating system divergence between the wild rice relatives. This work sets the stage for future efforts to understand the genetic basis of mating system evolution.

This dissertation is dedicated to my Grandfather, Arthur Fredrick Visnaw.

He passed away in 2008 during the early stages of my graduate school career.

I attribute my passion for biology to him.

ACKNOWLEDGMENTS

I am grateful to my graduate committee, Dr. Douglas Schemske, Dr. Jeff Conner, Dr. Shinhan Shui, and Dr. Steve van Nocker, as well as Dr. Tao Sang. In particular I am appreciative of Dr. Doug Schemske for his guidance and support. I would like to thank several undergraduate students for their support, Claire Lorts, Martha McCoy, and Emily Kramer. I also need to thank my family and friends for support.

TABLE OF CONTENTS

LIST OF TABLES……...vii

LIST OF FIGURES……viii

CHAPTER 1: Genetic architecture for the adaptive origin of annual wild rice, *Oryza nivara*.....1

Abstract……...……2
Introduction……..……..2
Materials and Methods…….……..…5
Results……....8
Discussion……..…16
Acknowledgements……..28
LITERATURE CITED……...29

CHAPTER 2: Fine mapping of a major flowering time QTL in *Oryza*…….....37

Introduction……..…38
Materials and Methods…….……...43
Results……...…46
Discussion……...…..49
LITERATURE CITED ……...……...58

CHAPTER 3: Genetic architecture of flowering time differentiation between locally adapted populations of *Arabidopsis thaliana*……...……..65

Abstract……...…..66
Introduction……..…66
Materials and Methods…….……..…71
Results……..…....76
Discussion……..…….....79
Acknowledgements……..90
LITERATURE CITED ……....…....91

CHAPTER 4: Functional traits underlying mating system differentiation between the wild relatives of rice……...…100

Introduction……....…101
Materials and Methods…….……..…107
Results…….……....111
Discussion……...…116
LITERATURE CITED ……130

LIST OF TABLES

Table 1.1. Effect of QTL alleles of *O. nivara* found in the combined analysis……...…...…10

Table 1.2. Epistatic interactions between QTL…….…...15

Table 2.1. Flowering time of Near-isogenic lines (NILs)…….……47

Table 2.2. Genes in the fine mapped region……..…48

Table 2.3. Polymorphisms in the *FTL1* coding region……...…52

Table 3.1. Results of quantitative trait locus (QTL) analysis of flowering time…….…..75

Table 3.2. Polymorphisms of candidate gene coding regions between the Italy and Sweden *Arabidopsis thaliana* mapping parents…….……...…..79

Table 4.1. Potential mating system traits and their putative function…….……111

Table 4.2. Trait means between *O. rufipogon* and *O. nivara*…….117

Table 4.3. Trait means for additional populations of each species……...……122

LIST OF FIGURES

Figure 1.1. Phenotypic divergence between *O. nivara* and *O. rufipogon*……..…….....13

Figure 1.2. Distribution of QTL Effect Size…….…26

Figure 2.1. Lifecycle of *Oryza* species…….....….50

Figure 2.2. Fine Mapping Results…….....52

Figure 2.3. Days to flowering in the recombinant mapping population……56

Figure 3.1. Proportion of plant flowering……..…….…77

Figure 3.2. Days to flowering in the F2 plants……...…...…82

Figure 3.3. Days to flowering in the Vernalization treatment……..……..…87

Figure 4.1. Functional traits of *Oryza* species…….....114

Figure 4.2. Pollen viability time course…….…...…118

Figure 4.3. Stigma receptivity time course……..…119

Figure 4.4. Autogamy and geitonogamy……..…125

Figure 4.5. Flower caricatures of *O. nivara* and *O. rufipogon*……..127

CHAPTER 1

Genetic architecture for the adaptive origin of annual wild rice, *Oryza nivara*

The following chapter was published in *Evolution.*

Grillo, M.A, C. Li, A.M. Fowlkes, T.M. Briggeman, A. Zhou, D.W. Schemske, and T. Sang. 2009. Genetic architecture for the adaptive origin of annual wild rice, *Oryza nivara*. *Evolution* 63: 870-883.

Abstract

The wild progenitors of cultivated rice, *Oryza* nivara and *Oryza* rufipogon, provide an experimental system for characterizing the genetic basis of adaptation. The evolution of annual *O. nivara* from a perennial ancestor resembling its sister species, *O. rufipogon*, was associated with an ecological shift from persistently wet to seasonally dry habitats. Here we report a quantitative trait locus (QTL) analysis of phenotypic differentiation in life history, mating system, and flowering time between *O. nivara* and *O. rufipogon*. The exponential distribution of effect sizes of QTL fits the prediction of a recently proposed population genetic model of adaptation. More than 80% of QTL alleles of *O. nivara* acted in the same direction of phenotypic evolution, suggesting that they were fixed under directional selection. The loss of photoperiod sensitivity, which might be essential to the survival of the ancestral populations of *O. nivara* in the new environment, was controlled by QTL of relatively large effect. Mating system evolution from cross- to self-fertilization through the modification of panicle and floral morphology was controlled by QTL of small-to-moderate effect. The lack of segregation of the recessive annual habit in the F2 mapping populations suggested that the evolution of annual from perennial life form had a complex genetic basis. The study captured the genetic architecture for the adaptive origin of *O. nivara* and provides a foundation for rigorous experimental tests of population genetic theories of adaptation.

Introduction

Darwin considered adaptation by natural selection a primary driving force for the origin of new forms and species. The following century witnessed noteworthy theoretical advances in elucidating the genetic basis of adaptation. The earliest theory traces back to Fisher's (1930)

geometric model predicting that favorable mutations fixed during adaptation should have small phenotypic effect, because mutations of large effect were more likely to generate negative, pleiotropic effects. Kimura later suggested that mutations of small effect could suffer from stochastic loss and concluded that mutations of intermediate sizes should be most frequently fixed. More recently, Orr (1998) proposed that the effect sizes of beneficial mutations should follow an exponential distribution, with fewer mutations of large effect fixed early on and mutations of decreasing sizes fixed subsequently as populations approach the fitness optimum. As molecular genetic markers and techniques became increasingly accessible to evolutionary biologists, empirical evidence for the genetic basis of adaptation, revealed largely by quantitative trait locus (QTL) mapping, has accumulated over the past two decades (Mackay, 2001a; Mackay, 2001b; Mauricio, 2001; Barton & Keightley, 2002; Rieseberg *et al.*, 2002; Rieseberg *et al.*, 2003; Orr, 2005; Slate, 2005; Mitchell-Olds & Schmitt, 2006). Perhaps the most remarkable progress has been made in elucidating the genetic basis of crop domestication under artificial selection (Doebley *et al.*, 2006). Genes and in some cases mutations responsible for major domestication transitions have been identified through molecular cloning of QTL in a number of important crops, such as maize, rice, wheat, barley, and tomato (Doebley *et al.*, 1997; Frary *et al.*, 2000; Wang *et al.*, 2005; Konishi *et al.*, 2006; Li, CB *et al.*, 2006a; Simons *et al.*, 2006; Komatsuda *et al.*, 2007). Similar insights were recently gained in natural systems such as sticklebacks and beach mice, where genes or mutations underlying adaptive traits were identified (Colosimo *et al.*, 2005; Hoekstra *et al.*, 2006). These successful evolutionary genetic analyses have benefited from a study system that consisted of an ancestor and its recent derivative with phenotypic traits of clear adaptive importance and evolutionary direction. With these attributes, the wild progenitors of rice, *Oryza* nivara and *O. rufipogon*, offer an excellent natural system for studying

the mechanisms of adaptive evolution. They are recently diverged sister species that have undergone marked adaptive differentiation to distinct habitats in tropical and subtropical Asia (Vaughan, 1994; Ge *et al.*, 1999; Vaughan *et al.*, 2008). Phylogenetic analysis has indicated that the annual species *O. nivara* was derived from a perennial ancestor resembling *O. rufipogon* (Zhu & Ge, 2005). A QTL analysis of phenotypic divergence between the two species thus should estimate the overall genetic architecture for the evolution of *O. nivara* from the *O. rufipogon*-like ancestor. *Oryza* nivara has differentiated from *O. rufipogon* in a number of traits that are considered of general importance for plant adaptation, including life history, mating system, and flowering time. Inhabiting deep-water swamps, *O. rufipogon* is perennial, outcrossing, and photoperiod sensitive. Adapted to the seasonally dry habitat, *O. nivara* became annual, predominately self-fertilized, and photoperiod insensitive. The derivation of an annual from a perennial life history and selfing from outcrossing mating system in *O. nivara* exemplifies not only the widely documented trends of life history and mating system evolution, but also the tendency of evolutionary association between annual habit and self-fertilization in flowering plants (Stebbins, 1974; Barrett *et al.*, 1996). The correlation between life history and habitat stability in these *Oryza* species, that is annual in seasonal environment and perennial in aseasonal environment, also represents a frequently observed phenomenon in plant adaptation (Stebbins, 1974). The ability to control flowering time by responding to photoperiod is another important adaptive mechanism in plants. It allows for the optimization of timing of pollination and seed dispersal and consequently maximizes reproductive success. Moreover, divergent flowering time can also promote assortative mating and act as a prezygotic barrier to gene flow (Kirkpatrick, 2000; Weis & Kossler, 2004), which may have served as a primary mechanism of reproductive isolation between the two *Oryza* species (Kuroda *et al.*, 2005). The adaptive

importance of these phenotypic differentiations between *O. nivara* and *O. rufipogon* was supported by a series of common garden experiments (Morishima *et al.*, 1961; Oka & Morishim.H, 1967; Oka, 1976), in which the two species were taxonomically treated as the annual and perennial strains of *O. perennis*. Here we report a QTL analysis of important adaptive traits involving life history, mating system, and photoperiod sensitivity of the two wild *Oryza* species. The availability of rice genome sequences has allowed the development of a high-quality linkage map with microsatellite markers evenly covering the 12 chromosomes. The results illustrate the genetic footprints of the adaptive origin of the annual species, *O. nivara*. The estimate of the number and magnitude of effect of QTL underlying the adaptive evolution provides empirical data for evaluating the predictions of the population genetic models of adaptation. The estimate of the dominant effect and direction of effect of QTL offers opportunities to address questions concerning the genetic consequences of the interplay between mating systems, standing genetic variation, and natural selection during adaptive evolution.

Materials and Methods

Oryza nivara and *O. rufipogon* are distributed from southeastern Asia to India. Based on our greenhouse observation of both species, two individuals representative of typical *O. rufipogon* (IRGC 80506 from India) and *O. nivara* (IRGC 80470 from India) were chosen as mapping parents (Fig. 1.1). Because the outcrossing parent *O. rufipogon* is highly heterozygous, we generated two F2 mapping populations (A, B) by self-fertilizing two F1 plants that were divergent in *O. rufipogon* alleles. Each mapping population consisted of 192 individuals. Both populations were grown in the same size pots under uniform environmental conditions and natural day length in a greenhouse at Michigan State University. The F2 seeds were germinated

in April 2004, and the mapping populations were grown for 2 years. Self-fertilized seeds of each parent were germinated at the same time, and grown under the same conditions as the F2 populations (N = 15, *O. nivara*; N = 30, *O. rufipogon*).

Simple sequence repeats (SSRs), or microsatellites, were designed based on previous publications and rice genome sequences (Temnykh *et al.*, 2001; McCouch *et al.*, 2002; Li, CB *et al.*, 2006a). SSRs were amplified by polymerase chain reaction (PCR) from DNA isolated using FastDNA kits (Q-BIOgene, Carlsbad, CA), and recorded on agarose gels stained with ethidium bromide. A total of 116 polymorphic markers with confirmed linkage using JoinMap 3.0 (Van Ooijen & Voorips, 2001) were chosen to construct the genetic map. The markers evenly covered the 12 chromosomes, with an average density of approximately 13 cM. Segregation distortion was tested using JoinMap 3.0.

Flowering time was determined as days from germination to flowering. The seeds of the F2 mapping populations and the selfed progenies of the parents germinated between April 30 and May 6, 2004.

Traits associated with mating system include panicle shape, panicle exsertion, anther length, and the timing of anther dehiscence (Fig. 1.1). Panicle shape was measured by the angles between primary branches and rachis. Panicle exsertion was measured by the distance from the flag leaf node to the first primary branch node. Each trait was measured twice when approximately 25% and 75% of flowers had opened on a panicle. The average of these two values was used to estimate panicle shape and exsertion around the peak flowering time. Three panicles were measured for a plant. Anther length was measured from three flowers that were about to open. In *O. nivara*, anthers become dehiscent immediately after the flower opens, whereas anther dehiscence is delayed from several minutes to more than half an hour in *O.*

rufipogon. This trait, however, was difficult to evaluate in the F2 populations because the timing of anther dehiscence was influenced by environmental factors, such as light, temperature, and humidity. As a result, we did not include it in the QTL study.

To determine whether F2 plants are annual like *O. nivara*, the longevity of the F2s were evaluated while the mapping populations were kept for 2 years. Other life-history traits involving resource allocation are tiller length, number of flowers (spikelets) per panicle, grain weight, and flowering duration. Tiller length was measured as the mean of the length of the three longest tillers. The number of spikelets per panicle was calculated as the mean of the five largest panicles. Grain weight was calculated as the mean of at least 30 grains. Flowering duration was measured as the proportion of flowering tillers in February and May, 2005 and scored as: 0% (0), 1–24% (1), 25–49% (2), 50–74% (3), 75–100% (4).

QTL were identified using composite interval mapping (CIM) implemented in the software package QTL Cartographer 2.5. A forward and backward stepwise regression procedure was used to identify cofactor for CIM.A10-cM scan window was employed, and the likelihood-ratio statistic was computed every 2 cM. Experiment-wide significance ($P < 0.05$) thresholds for QTL detection were determined with 1000 permutations. LOD values and R2 were determined based on likelihood-ratio tests under a hypothesis allowing both additive and dominance effects (H3:H0). Multiple interval mapping (MIM) was also conducted (Kao *et al.*, 1999). A Bayesian information criterion (BIC) model selection with the penalty function $c(n) = \ln(n)$ was used to perform MIM analysis at 1 cM steps. CIM and MIM analyses yielded very similar results with essentially the same number, chromosomal locations, and effect size of QTL identified. The main difference was that the LOD values of MIM were lower. Because the permutation test is still under development for MIM (Z. B. Zeng, per. comm.), we could not

compare the significance levels associated with the LOD values resulting from the two methods. Thus the results of CIM analysis were reported. Epistasis between QTL was tested using two-way analysis of variance (ANOVA) (Westerbergh & Doebley, 2004). The markers closest to the LOD peaks of QTL were tested for interaction. To test QTL clustering, we compared the distribution of the nearest neighbor QTL distances to a random distribution according to Protas *et al.* (2008).

Based on the genetic mapping and rice genome sequences, we identified a candidate gene, *Hd1*, for one of the flowering time QTL, DF6. The entire coding region of *Hd1* was PCR amplified from *O. rufipogon* using primers, AFH 5' CACGAAGCTTATGAATTATAATTTTGGTGGC and ARM 5' CGTGACGCGTTCAGAACCATGGAACAGTAC. The region was cloned into TA plasmids and fully sequenced. The verified sequence was then moved into a GA plasmid and driven by a 35S promoter. The plasmid was introduced into *Agrobacterium tumedaciens* strain LBA4404, which was then used to infect calli induced from germinating seeds of the *O. nivara* mapping parent. RNA was isolated from leaves of transgenic plants using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA). The expression of the introduced *O. rufipogon Hd1* coding region was confirmed by RT-PCR and sequencing. Transgenic plants that expressed the *O. rufipogon Hd1* were grown under long-day conditions (>14-h day time) for phenotypic evaluation.

Results

In this study, we evaluated morphological and physiological traits potentially important for the adaptation of *O. nivara* to a new habitat (Fig. 1.1). These include flowering time, continuation of flowering, tiller length, the number of flowers (spikelets) per panicle, grain

weight, panicle and floral morphology related to mating system differentiation, and longevity (see Discussion for details). For these phenotypic traits, the variation in the mapping populations as well as the mean and standard deviation in the parents is illustrated in Figure 1.1. Two traits associated with reproduction allocation, number of spikelets per panicle, and grain weight, show clearly transgressive segregation. As a short-day plant, the mapping parent of *O. rufipogon* did not flower when it was kept under a long-day condition (> 14-h day time). Under the natural day-length condition in the greenhouse, *O. rufipogon* began to flower in mid-October (~11-h day time), which corresponds to the day length experienced during early November in its native distributional range in India.

The *O. nivara* mapping parent was photoperiod insensitive and flowered after three to four months of vegetative growth under any day-length conditions. When grown together with the mapping populations in May of 2004, the self-fertilized progeny of the *O. nivara* parent flowered between the 9th and 16th of August, and the progeny of the *O. rufipogon* parent flowered between October 6th and November 7th. *Oryza* nivara flowered continuously throughout the growing season. By March 2005, all tillers of *O. nivara* had flowered and died, behaving as a true annual plant even with the unlimited supply of water in the greenhouse (Fig. 1.1). All of the mature tillers produced by *O. rufipogon* progeny completed flowering by February 2005. The remaining younger tillers did not flower until October 2005, the next short-day cycle. The maintenance of vegetative tillers during the current growing season allowed continuous growth into the next year, contributing to the perenniality of *O. rufipogon*. All *O. rufipogon* and F1 plants have survived for multiple years, indicating that the annual life history is recessive.

The F2 populations began to flower on July 31st, and the last F2 plant flowered on 15 November 2004. All but three F2 plants survived for 2 years in the greenhouse. The three individuals died during the first year probably as a result of hybrid weakness, showing short tillers and abnormal flowers with thin and aborted anthers. Thus, of the total 384 F2 individuals, there was no typical annual plant.

The *O. rufipogon* mapping parent was heterozygous (showing length polymorphism) for 66 of 116 or 59% SSR markers. The observation of such a high level of heterozygosity in the *O. rufipogon* genome indicated that the mapping parent was from a highly outcrossing population.

In contrast, no SSR polymorphism was found in the *O. nivara* parent. The two F1 plants used for developing the F2 mapping populations, A and B, had different *O. rufipogon* alleles at 38 of the 66 heterozygous SSR loci. Test for segregation distortion showed that significant deviation from the expected segregation ratio of 1:2:1 ($P < 0.01$) occurred at 15 SSR markers in mapping population A and 14 SSR markers in population B. Of these markers, eight showed segregation distortion in both populations, which covered two blocks on chromosomes 2 and 3.

Table 1.1. Effect of QTL alleles of *O. nivara* found in the combined analysis.

Effect of QTL alleles of *O. nivara* found in the combined analysis. R2, percent phenotypic variance explained by QTL; a, additive effect; d, dominance effect. Modes of gene action (GA): $d/a < -0.75$, recessive (R); $-0.75 \leq d/a \leq -0.25$, partially recessive (r); $-0.25 < d/a < 0.25$, additive (A); $0.25 \leq d/a \leq 0.75$, partially dominant (d); and $d/a > 0.75$, dominant (D). If the additive effect of an allele acted in the opposite direction of phenotypic variation, a "–" was added to the abbreviation for gene action (e.g., –r for TL7.2). LOD threshold determined by permutation for each trait in each is given in the parentheses following the trait name.

Table 1.1 (cont'd)

trait	QTL	population A					population B				
		LOD	R2	*a*	*d/a*	GA	LOD	R2	*a*	*d/a*	GA
Days to flowering A: 3.5, B:3.5	DF1	28.6	44.0	-21.3	0.23	A	16.7	24.8	-12.6	0.64	d
	DF6	9.0	9.3	-15.3	-0.64	r	11.0	13.3	-14.7	-0.71	r
	DF9	11.4	14.1	-17.6	-0.27	r	14.6	21.1	-15.6	0.13	A
Continued flowering A:3.5, B: 3.6	CF1	19.4	36.2	1.2	-0.15	A	16.8	34.0	0.63	0.70	d
	CF3	5.6	8.7	0.25	1.6	D	3.9	8.4	0.41	0.97	D
	CF10	4.9	9.3	0.36	0.89	D	8.5	16.5	0.55	0.84	D
Tiller length A: 3.5, B:3.5	TL1.1	17.1	27.3	-12.1	-0.06	A	11.1	20.6	-8.7	0.33	d
	TL1.2						3.9	6.9	-6.3	-0.76	R
	TL3						3.7	5.8	-6.0	-0.53	r
	TL4	3.8	7.7	-7.5	-0.40	r					
	TL6	4.8	9.3	-7.8	-0.46	r	7.8	11.1	-4.3	0.74	d
	TL7.1	11.9	25.1	-5.8	1.0	D	10.1	17.3	-4.5	1.3	D
	TL7.2	5.7	10.4	-9.0	-0.07	A					
Panicle exsertion A: 3.5, B: 3.5	PE1.1	4.1	9.8	0.9	0.58	-d					
	PE1.2						4.4	9.6	-0.84	0.06	A
	PE6	3.9	7.9	-1.3	-0.65	r					
	PE7						4.5	9.4	1.1	-0.06	-A
Panicle shape A: 3.6, B: 3.4	PS3						5.0	11.5	-3.6	0.31	d
	PS6.1						4.1	9.6	-4.4	-0.45	r
	PS6.2						3.9	10.1	-4.1	0.11	A
Anther length A:3.6, B: 3.6	AL2.1	6.0	13.8	-0.27	-0.48	r	5.7	12.0	-0.20	0.02	A
	AL2.2						3.9	7.9	-0.18	-0.44	r
	AL5	4.3	9.0	-0.22	-0.36	r	4.1	7.1	-0.11	1.00	D
	AL6	4.5	9.2	-0.16	-0.37	r	4.1	8.5	-0.18	-0.56	r
Spikelet number A: 3.3, B:3.5	SN1	7.1	13.8	-10.7	0.01	-A	15.0	23.3	-14.1	0.15	-A
	SN4						4.2	6.0	-8.2	-0.56	-r

A total of 22 and 28 QTL were identified from populations A and B, respectively. The QTL detection was based on LOD thresholds determined by permutation for each trait, ranging from 3.3 to 3.6.

Fifteen QTL had overlapping 1-LOD support intervals between the mapping populations, suggesting that they may be the same QTL detected independently from the two mapping populations. These included all of the QTL with effect sizes larger than 15% in at least one of the populations. QTL detected in only one population invariably had phenotypic effects lower than 15% in both mapping populations.

Of the 15 shared QTL, nine are found nearest to SSR markers with the same *O. rufipogon* allele in both populations and six are nearest to SSR markers with different *O. rufipogon* alleles between populations A and B. Of the seven QTL detected only in population A, four are found nearest to SSR markers with the same *O. rufipogon* allele in both populations and three are nearest to SSR markers with different *O. rufipogon* alleles between populations. The chi-square test indicates that there is no significant association between the detection of QTL unique to population A and the presence of the nearest *O. rufipogon* alleles unique to the population ($\chi 2 = 0.016$, df = 1, P = 0.90). No significant association is found for population B either ($\chi 2 = 0.080$, df = 1, P = 0.78), where five of 15 shared QTL and five of 13 unique QTL are found nearest to SSR markers with different *O. rufipogon* alleles between populations.

The results indicate that the allelic variation of the outcrossing parent does not seem to have a significant effect on QTL detection suggesting that *O. rufipogon* alleles showing SSR length polymorphism are functionally similar as far as interspecific phenotypic variation is concerned. Thus we combined the data of the two populations by treating two *O. rufipogon* alleles with length polymorphism at a locus as the same allele in the combined analysis.

The combined analysis detected a total of 30 QTL with effect sizes ranging from 2.9% to 36.5% (Table 1.1). The distribution of effect sizes (R2) of these QTL is illustrated in Figure 1.2. Of these QTL, 25 or 83.3% have additive effects in the same direction as the phenotypic evolution from *O. rufipogon* to *O. nivara*. The modes of action of the *O. nivara* alleles are broken down as: nine additive, 11 recessive or partially recessive, and 10 dominant or partially dominant.

Significant epistasis was detected between six pairs of QTL affecting four traits (Table 1.2). To test QTL clustering, 30 points were placed randomly on the linkage map. An average of three of 19 nearest neighbor distances were within 5 cM, compared to seven of 20 distances for the QTL dataset (Fisher's exact test, P = 0.96). This suggests that there is no significant clustering of QTL.

Figure 1.1. Phenotypic divergence between *O. nivara* and *O. rufipogon*.

For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation. Phenotypic divergence between *O. nivara* and *O. rufipogon*. (A) *Oryza rufipogon* mapping parent. (B) *Oryza nivara* mapping parent. (C) Open and exserted panicle of *O. rufipogon*. (D) Compact and partially inserted panicle of *O. nivara*. (E) An open flower of *O. rufipogon*, with large, indehiscent, pendant anthers. (F) An open flower of *O. nivara*, with smaller, immediately dehiscent, upright anthers. (G) *Oryza rufipogon* at the end of the first growing season, with dried tillers that had previously flowered and green tillers that would flower in the next season. (H) *Oryza nivara* at the end of the growing season, when all tillers had flowered and died.

Figure 1.1 (cont'd).

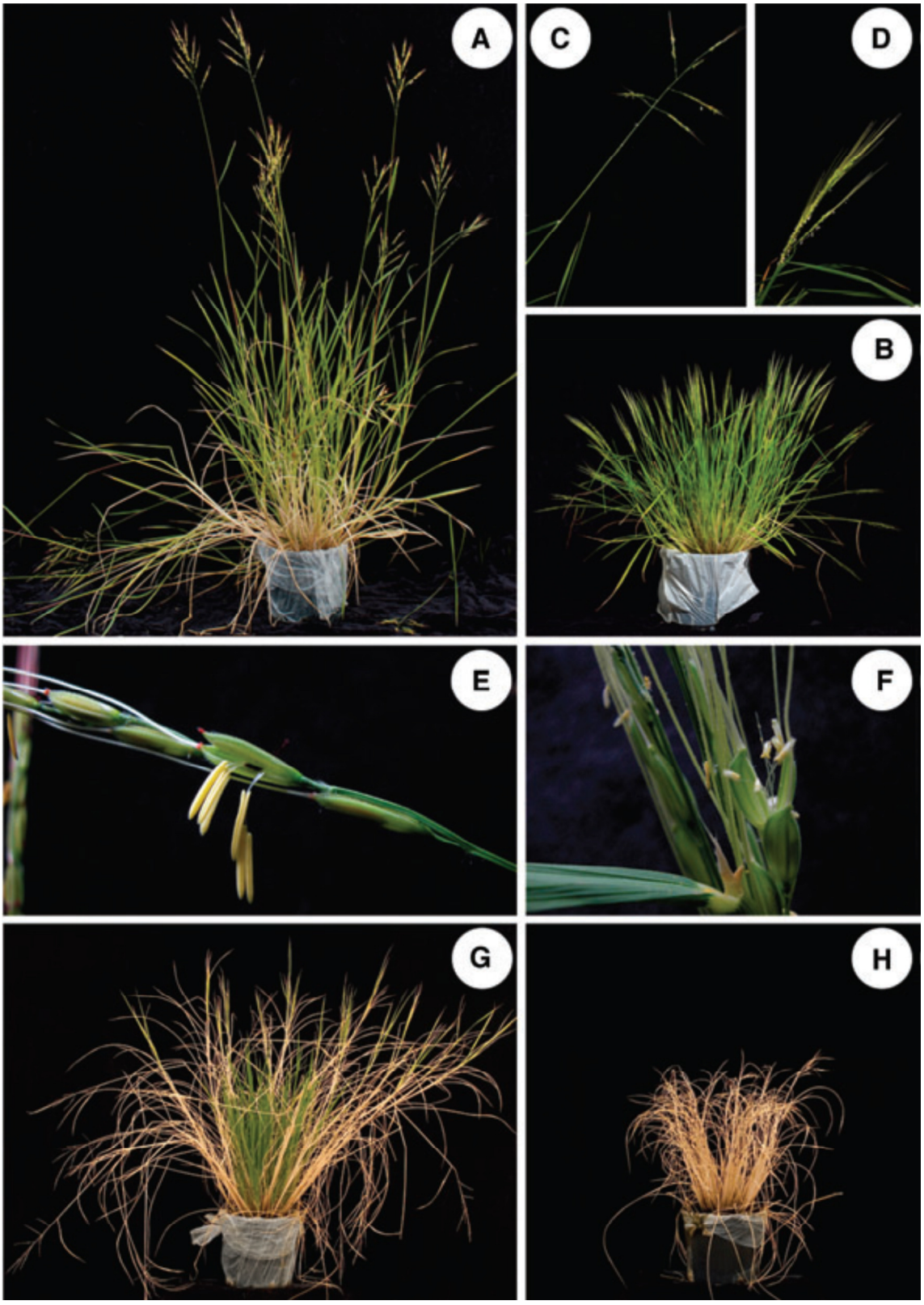


Table 1.2. Epistatic interactions between QTL.

Epistatic interactions between QTL revealed through two-way ANOVA with markers closest to the LOD peaks of QTL. ∗ P<0.05, ∗∗ P<0.01, ∗∗∗ P<0.001.

Trait	QTL	Markers	F-value	Significance
Days to flowering	DF1 x DF6	RC1–33×MRG5836	4.016	***
	DF1 x DF9	RC1–33×RC9–78	4.422	***
	DF6 x DF9	MRG5836×RC9–78	2.185	*
Tiller length	TL1.1 x TL7.2	RC1–42×RM248	2.812	*
Spikelet number	SN1 x SN9	RC1–33×MRG4840	3.985	**
Grain weight	GW2 x GW6	MRG2439×RM528	3.733	**

Of the three flowering time QTL identified in this study, the one with the smallest effect, DF6, mapped near a candidate gene, *Hd1*. *Hd1* is the rice ortholog of *Arabidopsis CONSTANS*, one of the central regulators of photoperiod sensitivity in both *Arabidopsis* and rice (Cremer & Coupland, 2003; Hayama *et al.*, 2003). We sequenced the coding region of the gene of the mapping parents and found that the *O. nivara* parent had a 55-bp insertion in the first exon of the gene. This was most likely a loss-of-function mutation. Previous studies showed that rice cultivars carrying a 2-bp insertion at the beginning of the second exon of *Hd1*, a loss-of-function mutation, had weaker photoperiod response than cultivars with the normal gene (Yano *et al.*, 2000).

To investigate the functional consequence of the loss-of-function mutation of *Hd1* in *O. nivara*, we introduced the *Hd1* coding region of *O. rufipogon* driven by the 35S promoter into

the *O. nivara* parent. A total of 12 T0 transgenic plants that expressed the *O. rufipogon* gene were obtained. All of the transgenic plants flowered when grown under a long-day condition (day length > 14 h). Seeds of five randomly chosen T0 lines together with those of the *O. nivara* parent were germinated and 10 T1 plants of each line were grown to maturity. All of the T1 plants flowered under the long-day condition, and their flowering time overlapped with that of *O. nivara*. These results showed that 35S-promoter driving expression of *O. rufipogon Hd1* coding region did not increase the sensitivity of photoperiod response of *O. nivara*.

Discussion

Because the genome of the outcrossing parent, *O. rufipogon*, is highly heterozygous, two F2 mapping populations were used to account for potential allelic effect on QTL detection. Taking advantage of the self-compatibility of the outcrossing parent and the F1s, we developed the F2 mapping populations through self-fertilization of separate F1 plants. In either mapping population we found<13% of SSR markers had experienced segregation distortion. The predominant Mendelian segregation in each mapping population allowed for straight forward QTL analyses.

With the two mapping populations, we could evaluate the extent to which different *O. rufipogon* alleles might affect the results of the QTL analysis. It is noteworthy that QTL with effect larger than 15% in at least one population are always detected in both populations in the separate analyses. The effect sizes of these shared QTL are significantly correlated between mapping populations (Spearman's Rho, $r = 0.83$, $P = 0.0002$). Furthermore, QTL that were identified in both A and B populations in the separate analyses were also found in the combined

analysis. The use of two mapping populations thus reinforces that these QTL are reliably identified.

In the separate analyses, the fact that only QTL of small effect lack the counterparts in the other population seems to suggest that this may be due primarily to limited statistical power for detecting small QTL in one of the mapping populations each with less than 200 F2 individuals (Beavis 1998). In the combined analysis, we were able to detect QTL with effect size as small as 2.9%, lower than 3.8% of the smallest effect of QTL found in the separate analyses. The combined analysis identified four new QTL, TL8, SN2, SN12, and GW10, with effect ranging 4.0–6.4%. However, nine QTL detected in either A or B population in the separate analyses were not found in the combined analysis, including two A population QTL with effect sizes of 7.7% and 9.8%, and seven B-population QTL of 5.8–10.1% effect (Table 1.1).

Apparently, the limited detection power of relatively small populations may explain the absence of small-effect QTL in one of the populations, but cannot explain the presence of QTL unique to one of the populations but absence in the combined analysis. One straightforward explanation for this is that the analysis of small populations could have created false QTL. However, we cannot rule out the possibility that the ability to detect small effect QTL in the combined analysis were undermined by subtle functional divergence between different *O. rufipogon* alleles. In fact, of nine QTL detected in one of the populations but not in the combined analysis, four are associated with the homozygous loci of *O. rufipogon* and five are associated the heterozygous loci. In comparison, of 10 QTL detected in one of the populations as well as in the combined analysis, seven and three are associated with homozygous and heterozygous loci of *O. rufipogon*, respectively. Thus, this raises the question of whether alleles

with minor functional differentiation might be easier to detect in separate analyses than in the combined analysis in which they are treated as the same functional allele.

Taken together, our analyses show that the detection of QTL of relative large effect is insensitive to the allelic variation, implying that QTL of large effect identified from two sampled individuals as the crossing parents may harbor genetic mutations important for the differentiation between the populations and species that they represent. Analyses of small populations could either omit QTL of small effect or create false QTL. However, the question of whether QTL present in one of the populations but not in the combined analysis are all or partially false cannot be fully addressed with the available data. In any event, QTL identified in both separate and combined analyses are considered to be reliable, and the results of the combined analysis will be the focus of further discussion.

Like the majority of *Oryza* species, *O. rufipogon* is a short-day plant that flowers near the end of the monsoon season. For perennials that live in relatively aseasonal habitats and rely primarily on vegetative reproduction, photoperiod sensitivity allows for the synchronization of flowering time and perhaps is advantageous also in preventing a prolonged flowering period and overproduction of seeds. Occurring in seasonal habitats, *O. nivara* is photoperiod insensitive and begins to flower three to four months after seed germination. This coincides with the middle of the monsoon season, allowing *O. nivara* to complete seed production before the onset of the dry season. Late flowering, however, is probably deleterious in the seasonally dry habitat because it does not allow enough time for seed production before desiccation. This study thus provides genetic evidence for the potential importance of the loss of photoperiod sensitivity in the adaptive evolution of a plant species.

Three QTL, DF1, DF6, and DF9, identified for flowering time have moderate-to-large effect (Table 1.1). Particularly, DF1 is the largest (R2 = 36.5%) of all QTL found in the combined analysis. Fixation of the allele of such a large effect could have substantially weakened photoperiod sensitivity of the *O. rufipogon*-like ancestor and consequently improved its fitness by permitting seed production in the seasonally dry habitat.

The complementation test for the candidate gene of DF6, Hd1, has the following implications for the evolution of photoperiod insensitivity of *O. nivara*. The failure to rescue photoperiod sensitivity of *O. nivara* by overexpressing the *O. rufipogon* gene, which is at least partially dominant over the *O. nivara* allele, is consistent with the finding that there are two other QTL of larger effect on photoperiod sensitivity. But we cannot rule out the possibility that the lack of phenotypic effect was due to the 35S promoter driving overexpression, because it was shown that in certain rice cultivars *Hd1* complementation with its endogenous promoter delayed flowering under long-day conditions (Yano *et al.*, 2000). Nevertheless, it would not be surprising if *Hd1* together with the *O. rufipogon* alleles at DF1 and DF9 are required to make *O. rufipogon* an obligated short-day plant. Interestingly, epistatic interactions have been detected between any two of the three flowering time QTL (Table 1.2), suggesting that they may function in the same or related pathways. The number and type of mutations selected at these three loci that led to the evolution of a day neutral plant from an obligated short-day plant is of great interest and can be investigated through molecular cloning of DF1, DF6, and DF9 followed by functional and ecological analyses of their influence on flowering time.

Previous studies of insect-pollinated plants have elucidated the genetic basis for the adaptive evolution of mating systems (Fenster & Ritland, 1994; Goodwillie, 1999). QTL of small to moderate effect were identified for floral traits associated with the differentiation of

selfing and outcrossing mating systems between *Mimulus* species (Lin & Ritland, 1997; Fishman *et al.*, 2002). Similarly, QTL of only moderate effect were found for the evolution of floral morphology associated with self-fertilization in *Leptosiphon* (Goodwillie *et al.*, 2006). However, QTL of relative large effect were found underlying floral morphology leading to higher outcrossing rates in wild tomatoes (Georgiady *et al.*, 2002).

For wind pollinated, self-compatible *Oryza* species, outcrossing rates were shown to be closely correlated with floral morphology (Oka & Morishim.H, 1967). *Oryza rufipogon* has larger anthers containing more pollen grains than *O. nivara*, consistent with the scenario that higher outcrossing rates are associated with higher male allocation of reproductive resources (Charlesworth & Charlesworth, 1981). The panicles of *O. rufipogon* are exserted from the flag leaves and open in shape, promoting pollen dispersal by wind (Fig. 1.1). In contrast, *O. nivara* has compact panicles that are partially enveloped by the flag leaves during pollination, reducing the chance of cross-pollination.

QTL identified for these panicle and floral traits had small to moderate effects that explained 4.9–10.5% of the phenotypic variance for two populations combined (Table 1.1). This adds another example, especially for wind-pollinated plants, to the growing body of empirical data suggesting that QTL of small to-moderate effect are often responsible for the evolution of self-fertilization through modification of floral morphology.

The differentiation between annual and perennial life histories is common in flowering plants and is often associated with traits influencing vegetative propagation and reproductive allocation. *Oryza rufipogon* inhabits deep-water swamps that support a rather closed cover of vegetation all year round. Because of relatively high seedling mortality in this type of habitat, perennials capable of vegetatively propagating are favored. In seasonally dry habitats in which

mature plants cannot survive through the dry season, annuals with high reproductive allocation are selectively advantageous. Field investigations confirmed that the majority of young tillers of *O. rufipogon* grew from underground stems produced in the previous growing season, and all individuals of *O. nivara* arose from seeds (Sano *et al.*, 1980).

Previous genetic studies of annual and perennial differentiation analyzed either multiple traits related to life history or a key trait, such as rhizome formation, responsible primarily for vegetative propagation or overwinter survival. They found that multiple QTL were involved in both cases (Hu *et al.*, 2003; Westerbergh & Doebley, 2004; Hall *et al.*, 2006). In this study, we grew the mapping populations for 2 years to allow the direct evaluation of annual versus perennial life form. All F2 plants survived during the 2-year period of experiment. There was no typical annual plant as the *O. nivara* parent found in either population. Given that the mapping parents showed the annual and perennial differentiation under the same growing conditions and the perennial habit is dominant in F1, the lack of segregation of the recessive annual habit suggests that the differentiation of annual and perennial life history, as evaluated by longevity, has complex genetic control. It may be determined by multiple traits associated with resource acquisition and allocation, such as tiller length, spikelet number, grain weight, and flowering duration. A total of 20 QTL were identified for these traits in the combined analysis (Table 1.1).

Oryza rufipogon reproduces largely through vegetative propagation by producing horizontal tillers up to 5-m long, which form new stems and adventitious roots at the nodes (Oka & Morishima, 1967; Xie *et al.*, 2001). In the greenhouse, *O. rufipogon* grew longer tillers with one to two more leaves than *O. nivara* before they flowered. *Oryza* nivara produces a larger number of seeds that weigh more than those of *O. rufipogon*. Thus, each tiller of *O. rufipogon*

can potentially yield more photosynthetic product but allocate a less amount of it to seed production. As a consequence, the perennial species stores much of the resource to support continuous vegetative propagation.

Furthermore, *O. rufipogon* was able to re-establish a vegetative growth phase after flowering in each growing season. In contrast, *O. nivara* was unable to halt flowering once initiated. Continuous flowering in *O. nivara* maximizes resource allocation to seed production until the plant is exhausted and dies. Although the developmental mechanisms underlying this trait are unknown, the loss of photoperiod sensitivity might have contributed to continuous flowering in *O. nivara*. Being photoperiod sensitive, the younger tillers of *O. rufipogon* were unable to flower from February to September while the day length was increasing. Taken together, the annual habit of *O. nivara* may have evolved as a result of a series of phenotypic modifications that shifted resource allocation from sustaining continuous vegetative growth to maximizing seed production.

It is noteworthy that 83.3% of the QTL alleles of *O. nivara* act in the direction of the phenotypic evolution from *O. rufipogon* to *O. nivara* in the combined analyses. For flowering time and duration, all six QTL act in the same direction to reduce photoperiod sensitivity and promote continuous flowering. Similarly, all seven QTL underlying mating system traits, panicle exsertion, panicle shape, and anther length, act in the same direction to potentially improve selfing rates. For the remaining three traits, tiller length, spikelet number per panicle, and grain weight, which are related to resource acquisition and allocation, a total of 17 QTL are identified and 12 act in the direction of phenotypic evolution of *O. nivara*.

The result is comparable to that of a QTL analysis of domestication transitions from *O. nivara* to an indica rice cultivar, in which the same *O. nivara* accession was used as the mapping

parent (Li, CB *et al.*, 2006a). It was found that 81% of QTL alleles of the indica cultivar acted to improve the domestication traits, consistent with the notion that the origin of cultivated rice was driven by strong directional artificial selection that should result in the fixation of favorable mutations largely in the same direction. Although the number of QTL detected in this study is too small for most of traits to provide enough power for a sign test, the finding that the proportion of total QTL acting in the evolutionary direction of *O. nivara* is even higher than that found in the case of rice domestication supports the hypothesis that the origin of *O. nivara* was driven by directional natural selection moving the ancestral populations toward a new fitness optimum (Orr, 1998; Rieseberg *et al.*, 2002; Carbone *et al.*, 2005).

Colocalization of QTL for different traits may be an indication of pleiotropic mutations or mutations of linked genes that form an adaptive gene complex. Of 30 QTL identified in the combined analysis, 15 or 50% have their 1-LOD supporting intervals overlapped with at least that of another QTL. But we failed to detect any significant clustering of the QTL. This degree of colocalization is not nearly as striking as was found for rice domestication where 70% of QTL for all morphological traits related to grain yield were colocalized and two large QTL clusters were clearly present on chromosomes 4 and 7 (Li, CB *et al.*, 2006a). During rice domestication, fixation of pleiotropic mutations or tightly linked mutations under strong artificial selection could have warranted rapid modifications in plant morphology for higher grain yield. We did not find similar evidence suggesting that such cohorts of traits were under natural selection during the adaptation of *O. nivara*.

Another potentially meaningful comparison between the origin *O. nivara* and domestication of rice concerns the dominant effect of QTL alleles. The previous analysis showed that the indica cultivar fixed 16 recessive and partially recessive alleles and only four

partially dominant alleles when domesticated from *O. nivara* (Li, CB *et al.*, 2006a). In this study, the combined analysis showed that there was nearly the equal number of dominant and recessive alleles fixed in *O. nivara*. The difference may be attributed to the different mating systems of the ancestral species.

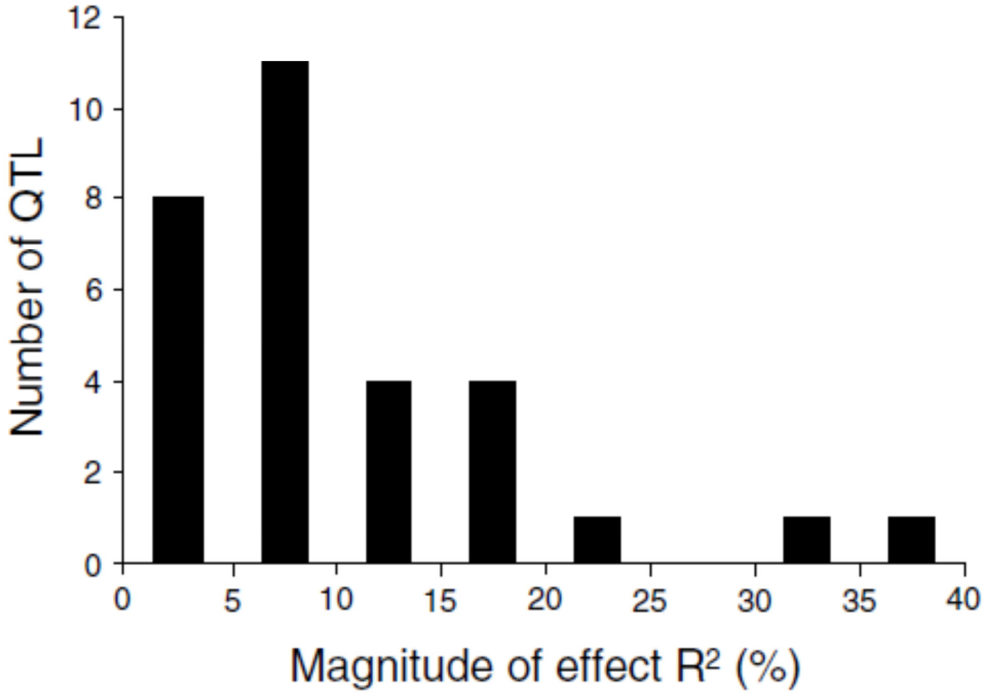
Although understanding the history of rice domestication has been a challenging evolutionary problem by itself, certain consensus began to emerge from recent molecular genetic and phylogenetic studies. That is, two major types of rice cultivars, namely indica and japonica, were domesticated from wild populations with divergent genomic backgrounds, and both *O. nivara* and *O. rufipogon* (sometimes treated as two ecotypes of *O. rufipogon* in domestication literature) contributed, probably to different extent, to the domesticated gene pool (Londo *et al.*, 2006; Sang & Ge, 2007a; Sang & Ge, 2007b; Sweeney & McCouch, 2007). *Oryza* nivara may have served as the primary genome donor for the indica cultivar (Li, CB *et al.*, 2006b). Thus, the progenitor of indica rice could be self-fertilized, whereas *O. nivara* had an outcrossing *O. rufipogon*-like ancestor. In this regard, the more abundant dominant mutations found in *O. nivara* than in the indica cultivar could be due to the fact that dominant mutations were more likely to be fixed in the outcrossing ancestral populations of *O. nivara*. Nevertheless, a total of 11 recessive alleles found in the combined analysis of *O. nivara* and *O. rufipogon* divergence are still larger in number than 10 dominant alleles. The fixation of relatively abundant recessive mutations in *O. nivara* can be explained by several reasons. First, beneficial mutations might have preexisted in the ancestral populations, in which recessiveness would allow these previously deleterious alleles to persist at the mutation selection balance in the outcrossing ancestor. The preexisting recessive mutations could be quickly fixed in response to a sudden environmental change (Orr & Betancourt, 2001). Second, the increase in selfing rates during the

evolution of *O. nivara* might have facilitated the fixation of at least partially recessive mutations as they occurred along the adaptive walk (Charlesworth, 1992). Finally, recessive mutations might have occurred more frequently than dominant ones in the *Oryza* species as suggested by the study of rice domestication (Li, CB *et al.*, 2006b).

The effect sizes of the QTL from the combined analysis follow roughly an exponential distribution (Fig. 1.2). That there are fewer QTL with effect sizes 0–5% than those with 5–10% effect is most likely because many QTL with small effect of 0–5% remain undetected given the mapping population size of less than 400 individuals (Beavis, 1998). This result thus fits the prediction of Orr's (1998a) population genetic model of adaptation. The similar distribution has been observed in other cases of plant adaptation under natural or artificial selection (Tanksley, 1993; Bradshaw *et al.*, 1998; Doganlar *et al.*, 2002; Doebley, 2004; Lauter *et al.*, 2004). There were also examples in which QTL effect did not follow an exponential distribution due to the lack of large effect QTL (Burke *et al.*, 2002; Fishman *et al.*, 2002; Lexer *et al.*, 2005). This might be explained by several possible reasons including the fixation of standing mutations of relative small effect, adaptation following a moving optimum such as a cline of intermediate habitats, and weak selection (Orr, 2005). The generalization of empirical support for the theoretical models of adaptation still awaits the accumulation of a larger body of genetic data.

Figure 1.2. Distribution of QTL Effect Size.

Distribution of magnitude of effect of QTL identified in the combined analysis of F2 mapping populations.



The consistency of our results with Orr's model of adaptation is not surprising because the adaptive evolution of *O. nivara* does seem to satisfy the basic assumptions of the model. That is, when a population is distantly off its adaptive optimum due to drastic environmental changes, an adaptive walk begins and moves the population progressively toward a new optimum. This is likely the case for the wild rice species. The *O. rufipogon*-like ancestor of *O. nivara* already adapted to the deep-water habitat could have been dispersed to the seasonally dry

habitat, or experienced severe drought as the original habitat dried up. Given the abundance of seasonal habitats in regions with a typical monsoon season, such sharp environmental changes would have repeatedly occurred and consequently imposed strong directional selection for the origin of *O. nivara*.

In this regard, Orr's model also predicts that initial mutations fixed during adaptation tend to have large phenotypic effect because they are less likely to overshoot the optimum than those fixed later when the population gets closer to the optimum. The QTL of the largest effect detected in this study is DF1, which explains 36.5% of the phenotypic variance for flowering time. This implies that the loss of photoperiod sensitivity could be an early step for the adaptive evolution of *O. nivara*. Early flowering in the seasonally dry habitat is no longer deleterious, as it would have been for the perennial ancestor living in the deep water habitat. Conversely, the loss of photoperiod sensitivity could have substantially improved seed production and consequently the fitness of ancestral populations of *O. nivara* in the seasonal habitat.

Testing this hypothesis requires molecular cloning of DF1 to determine whether the QTL indeed harbors a mutation of large effect (Phillips, 2005; Mitchell-Olds *et al.*, 2007). Further, the fitness effect of the QTL needs to be confirmed in the natural habitats using a near-isogenic line (Bradshaw & Schemske, 2003) that contains the *O. nivara* allele of DF1 in the genomic background of *O. rufipogon*. Both experiments are feasible especially with the availability of rice genome sequences. The genomic era of biology should soon witness rigorous experimental tests of the origin of species under natural selection.

Acknowledgments

We are grateful to the anonymous reviewers and T. Juenger for their valuable comments and suggestions that helped improve the previous versions of the manuscript. The research was supported by the National Science Foundation of USA.

LITERATURE CITED

LITERATURE CITED

Barrett SCH, Harder LD, Worley AC. 1996. The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **351**(1345): 1271-1280.

Barton NH, Keightley PD. 2002. Understanding quantitative genetic variation. *Nature Reviews Genetics* **3**(1): 11-21.

Beavis WD 1998. QTL analyses: power, precision, and accuracy. In: Paterson AH ed. *Molecular Dissection of Complex Traits*. Boca Raton, FL: CRC Press, 145-162.

Bradshaw HD, Otto KG, Frewen BE, McKay JK, Schemske DW. 1998. Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (Mimulus). *Genetics* **149**(1): 367-382.

Bradshaw HD, Schemske DW. 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* **426**(6963): 176-178.

Burke JM, Tang S, Knapp SJ, Rieseberg LH. 2002. Genetic analysis of sunflower domestication. *Genetics* **161**(3): 1257-1267.

Carbone MA, Llopart A, DeAngelis M, Coyne JA, Mackay TFC. 2005. Quantitative trait loci affecting the difference in pigmentation between Drosophila yakuba and D. santomea. *Genetics* **171**(1): 211-225.

Charlesworth B. 1992. EVOLUTIONARY RATES IN PARTIALLY SELF-FERTILIZING SPECIES. *American Naturalist* **140**(1): 126-148.

Charlesworth D, Charlesworth B. 1981. ALLOCATION OF RESOURCES TO MALE AND FEMALE FUNCTIONS IN HERMAPHRODITES. *Biological Journal of the Linnean Society* **15**(1): 57-74.

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM. 2005. Widespread parallel evolution

in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**(5717): 1928-1933.

Cremer F, Coupland G. 2003. Distinct photoperiodic responses are conferred by the same genetic pathway in Arabidopsis and in rice. *Trends in Plant Science* **8**(9): 405-407.

Doebley J. 2004. The genetics of maize evolution. *Annual Review of Genetics* **38**: 37-59.

Doebley J, Stec A, Hubbard L. 1997. The evolution of apical dominance in maize. *Nature* **386**(6624): 485-488.

Doebley JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication. *Cell* **127**(7): 1309-1321.

Doganlar S, Frary A, Daunay MC, Lester RN, Tanksley SD. 2002. Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. *Genetics* **161**(4): 1713-1726.

Fenster CB, Ritland K. 1994. QUANTITATIVE GENETICS OF MATING SYSTEM DIVERGENCE IN THE YELLOW MONKEYFLOWER SPECIES COMPLEX. *Heredity* **73**: 422-435.

Fisher RA. 1930. *The genetical theory of natural selection*. Chicago, IL: University of Chicago Press.

Fishman L, Kelly AJ, Willis JH. 2002. Minor quantitative trait loci underlie floral traits associated with mating system divergence in Mimulus. *Evolution* **56**(11): 2138-2155.

Frary A, Nesbitt TC, Grandillo S, van der Knaap E, Cong B, Liu JP, Meller J, Elber R, Alpert KB, Tanksley SD. 2000. fw2.2: A quantitative trait locus key to the evolution of tomato fruit size. *Science* **289**(5476): 85-88.

Ge S, Sang T, Lu BR, Hong DY. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proceedings of the National Academy of Sciences of the United States of America* **96**(25): 14400-14405.

Georgiady MS, Whitkus RW, Lord EM. 2002. Genetic analysis of traits distinguishing outcrossing and self-pollinating forms of currant tomato, Lycopersicon pimpinellifolium (Jusl.) Mill. *Genetics* **161**(1): 333-344.

Goodwillie C. 1999. Multiple origins of self-compatibility in Linanthus section Leptosiphon (Polemoniaceae): Phylogenetic evidence from internal-transcribed-spacer sequence data. *Evolution* **53**(5): 1387-1395.

Goodwillie C, Ritland C, Ritland K. 2006. The genetic basis of floral traits associated with mating system evolution in Leptosiphon (Polemoniaceae): An analysis of quantitative trait loci. *Evolution* **60**(3): 491-504.

Hall MC, Basten CJ, Willis JH. 2006. Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in Mimulus guttatus. *Genetics* **172**(3): 1829-1844.

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K. 2003. Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature* **422**(6933): 719-722.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**(5783): 101-104.

Hu FY, Tao DY, Sacks E, Fu BY, Xu P, Li J, Yang Y, McNally K, Khush GS, Paterson AH, Li ZK. 2003. Convergent evolution of perenniality in rice and sorghum. *Proceedings of the National Academy of Sciences of the United States of America* **100**(7): 4050-4054.

Kao CH, Zeng ZB, Teasdale RD. 1999. Multiple interval mapping for quantitative trait loci. *Genetics* **152**(3): 1203-1216.

Kirkpatrick M. 2000. Reinforcement and divergence under assortative mating. *Proceedings of the Royal Society B-Biological Sciences* **267**(1453): 1649-1655.

Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, Stein N, Graner A, Wicker T, Tagiri A, Lundqvist U, Fujimura T, Matsuoka M, Matsumoto T, Yano M. 2007. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proceedings of the National Academy of Sciences of the United States of America* **104**(4): 1424-1429.

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M. 2006. An SNP caused loss of seed shattering during rice domestication. *Science* **312**(5778): 1392-1396.

Kuroda Y, Sato YI, Bounphanousay C, Kono Y, Tanaka K. 2005. Gene flow from cultivated rice (Oryza sativa L.) to wild Oryza species (O-rufipogon Griff. and O-nivara Sharma and Shastry) on the Vientiane plain of Laos. *Euphytica* **142**(1-2): 75-83.

Lauter N, Gustus C, Westerbergh A, Doebley J. 2004. The inheritance and evolution of leaf pigmentation and pubescence in teosinte. *Genetics* **167**(4): 1949-1959.

Lexer C, Rosenthal DM, Raymond O, Donovan LA, Rieseberg LH. 2005. Genetics of species differences in the wild annual sunflowers, Helianthus annuus and H-petiolaris. *Genetics* **169**(4): 2225-2239.

Li CB, Zhou AL, Sang T. 2006a. Genetic analysis of rice domestication syndrome with the wild annual species, Oryza nivara. *New Phytologist* **170**(1): 185-193.

Li CB, Zhou AL, Sang T. 2006b. Rice domestication by reducing shattering. *Science* **311**(5769): 1936-1939.

Lin JZ, Ritland K. 1997. Quantitative trait loci differentiating the outbreeding Mimulus guttatus from the inbreeding M-platycalyx. *Genetics* **146**(3): 1115-1121.

Londo JP, Chiang Y-C, Hung K-H, Chiang T-Y, Schaal BA. 2006. Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, Oryza sativa. *Proceedings of the National Academy of Sciences of the United States of America* **103**(25): 9578-9583.

Mackay TFC. 2001a. The genetic architecture of quantitative traits. *Annual Review of Genetics* **35**: 303-339.

Mackay TFC. 2001b. Quantitative trait loci in Drosophila. *Nature Reviews Genetics* **2**(1): 11-20.

Mauricio R. 2001. Mapping quantitative trait loci in plants: Uses and caveats for evolutionary biology. *Nature Reviews Genetics* **2**(5): 370-381.

McCouch SR, Teytelman L, Xu YB, Lobos KB, Clare K, Walton M, Fu BY, Maghirang R, Li ZK, Xing YZ, Zhang QF, Kono I, Yano M, Fjellstrom R, DeClerck G, Schneider D, Cartinhour S, Ware D, Stein L. 2002. Development and mapping of 2240 new SSR markers for rice (Oryza sativa L.). *DNA Research* **9**(6): 199-207.

Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature* **441**(7096): 947-952.

Mitchell-Olds T, Willis JH, Goldstein DB. 2007. Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics* **8**(11): 845-856.

Morishima H, Chang WT, Oka HI. 1961. DIRECTIONS OF DIFFERENTIATION IN POPULATIONS OF WILD RICE, ORYZA PERENNIS AND O SATIVA F SPONTANEA. *Evolution* **15**(3): 326-&.

Oka HI. 1976. MORTALITY AND ADAPTIVE MECHANISMS OF ORYZA PERENNIS STRAINS. *Evolution* **30**(2): 380-392.

Oka HI, Morishim.H. 1967. VARIATIONS IN BREEDING SYSTEMS OF A WILD RICE ORYZA PERENNIS. *Evolution* **21**(2): 249-&.

Oka HI, Morishima H. 1967. VARIATIONS IN BREEDING SYSTEMS OF A WILD RICE ORYZA PERENNIS. *Evolution* **21**(2): 249-&.

Orr HA. 1998. The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution* **52**(4): 935-949.

Orr HA. 2005. The genetic theory of adaptation: A brief history. *Nature Reviews Genetics* **6**(2): 119-127.

Orr HA, Betancourt AJ. 2001. Haldane's sieve and adaptation from the standing genetic variation. *Genetics* **157**(2): 875-884.

Phillips PC. 2005. Testing hypotheses regarding the genetics of adaptation. *Genetica* **123**(1-2): 15-24.

Protas M, Tabansky I, Conrad M, Gross JB, Vidal O, Tabin CJ, Borowsky R. 2008. Multi-trait evolution in a cave fish, Astyanax mexicanus. *Evolution & Development* **10**(2): 196-209.

Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan LA, Lexer C. 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**(5637): 1211-1216.

Rieseberg LH, Widmer A, Arntz AM, Burke JM. 2002. Directional selection is the primary cause of phenotypic diversification. *Proceedings of the National Academy of Sciences of the United States of America* **99**(19): 12242-12245.

Sang T, Ge S. 2007a. Genetics and phylogenetics of rice domestication. *Current Opinion in Genetics & Development* **17**(6): 533-538.

Sang T, Ge S. 2007b. The puzzle of rice domestication. *Journal of Integrative Plant Biology* **49**(6): 760-768.

Sano Y, Morishima H, Oka HI. 1980. INTERMEDIATE PERENNIAL-ANNUAL POPULATIONS OF ORYZA-PERENNIS FOUND IN THAILAND AND THEIR EVOLUTIONARY SIGNIFICANCE. *Botanical Magazine-Tokyo* **93**(1032): 291-305.

Simons KJ, Fellers JP, Trick HN, Zhang ZC, Tai YS, Gill BS, Faris JD. 2006. Molecular characterization of the major wheat domestication gene Q. *Genetics* **172**(1): 547-555.

Slate J. 2005. Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Molecular Ecology* **14**(2): 363-379.

Stebbins GL. 1974. *Flowering plants: evolution above the species level.* Cambridge, MA: The Belknap Press of Harvard University Press.

Sweeney M, McCouch S. 2007. The complex history of the domestication of rice. *Annals of Botany* **100**(5): 951-957.

Tanksley SD. 1993. Mapping polygenes. *Annual Review of Genetics* **27**: 205-233.

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S. 2001. Computational and experimental analysis of microsatellites in rice (Oryza sativa L.):

Frequency, length variation, transposon associations, and genetic marker potential. *Genome Research* **11**(8): 1441-1452.

Van Ooijen JW, Voorips RE. 2001. *Joinmap 3.0: software for the calculation of genetic linkage maps*. Wageningen, the Netherlands: Plant Research International.

Vaughan DA. 1994. *The wild relatives of rice: a genetic resources bankbook*. Philippines: International Rice Research Institute.

Vaughan DA, Lu B-R, Tomooka N. 2008. The evolving story of rice evolution. *Plant Science* **174**(4): 394-408.

Wang H, Nussbaum-Wagler T, Li BL, Zhao Q, Vigouroux Y, Faller M, Bomblies K, Lukens L, Doebley JF. 2005. The origin of the naked grains of maize. *Nature* **436**(7051): 714-719.

Weis AE, Kossler TM. 2004. Genetic variation in flowering time induces phenological assortative mating: Quantitative genetic methods applied to Brassica rapa. *American Journal of Botany* **91**(6): 825-836.

Westerbergh A, Doebley J. 2004. Quantitative trait loci controlling phenotypes related to the perennial versus annual habit in wild relatives of maize. *Theoretical and Applied Genetics* **109**(7): 1544-1553.

Xie ZW, Lu YQ, Ge S, Hong DY, Li FZ. 2001. Clonality in wild rice (Oryza rufipogon, Poaceae) and its implications for conservation management. *American Journal of Botany* **88**(6): 1058-1064.

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T. 2000. Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the arabidopsis flowering time gene CONSTANS. *Plant Cell* **12**(12): 2473-2483.

Zhu QH, Ge S. 2005. Phylogenetic relationships among A-genome species of the genus Oryza revealed by intron sequences of four nuclear genes. *New Phytologist* **167**(1): 249-265.

CHAPTER 2

Fine mapping of a major flowering time QTL in *Oryza*

Introduction

Adaptation is a driving force for the evolution of new forms and species. While significant gains have been made in our understanding of this process, the genetic basis of adaptation in particular remains poorly understood (Orr, 2005). To uncover the mechanisms of adaptation we ultimately need to identify the genes underlying adaptive phenotypes and experimentally demonstrate the effect of these genes on fitness (Barton & Keightley, 2002; Slate, 2005). Recent advances in genomic technologies now allow for the rigorous examination of the genetic mechanisms of adaptation in a broad range of organisms.

Throughout the last twenty years advancements in Quantitative Trait Loci (QTL) mapping have provided a reliable means for identifying regions of the genome associated with trait variation (Tanksley, 1993; Orr, 2001; Barton & Keightley, 2002; Orr, 2005; Mitchell-Olds & Schmitt, 2006; Mackay *et al.*, 2009). During this period a plethora of QTL studies have been carried out, with the majority of this work being conducted in genetic model systems or domesticated plants and animals; although examples from natural systems exist and continue to accumulate (Bradshaw *et al.*, 1998; Colosimo *et al.*, 2004; Goodwillie *et al.*, 2006; Hall & Willis, 2006; Bradshaw *et al.*, 2012; Leinonen *et al.*, 2013). QTL mapping has been primarily useful for characterizing the genetic architecture of phenotypic variation; that is estimating the number, effect size, and location of loci controlling trait variation. Perhaps the most noteworthy outcome of this line of research has been the regular identification of large effect QTL (Orr, 2005). This finding is surprising given that it is counter to the long standing dogma that mutations of small effect control adaptation (Fisher, 1930). In response, new theory has been developed in order to understand the role of large effect mutations (Orr, 1998; Orr, 2005).

Through these efforts it is now well established that loci of large effect can indeed influence trait variation (Barton & Keightley, 2002; Orr, 2005; Phillips, 2005; Mitchell-Olds *et al.*, 2007).

QTL mapping has been pivotal in our knowledge of the genetic architecture of adaptation, by identifying genomic regions associated with trait variation. These regions, however, can be quite large and may encompass hundreds of genes. Furthermore, it has been shown that separate mutations within a single gene can have independent effects on phenotypic differentiation (Rockman, 2012; Linnen *et al.*, 2013). Thus, a comprehensive understanding of the genetic basis of adaptation requires elucidating the causal genes and mutations that underlie QTL (Barton & Keightley, 2002; Orr, 2005; Phillips, 2005; Hoekstra & Coyne, 2007; Mitchell-Olds *et al.*, 2007; Rockman, 2012). Examination of such causal genetic elements can provide insight into several facets concerning the genetic basis of adaptation, including i.) the types of genes controlling adaptive variation (e.g. protein coding vs. regulatory genes), ii.) the role of *cis*-regulatory mutations, and iii.) the relative role of new mutations vs. standing genetic variation (Phillips, 2005; Mackay *et al.*, 2009).

The wild progenitors of cultivated rice, *Oryza nivara* and *Oryza rufipogon*, present an ideal system for studying the mechanisms of adaptation. *Oryza nivara* and *O. rufipogon* are recently diverged sister species that have undergone marked adaptive differentiation to distinct wetland habitats in the tropics and sub-tropics of Asia (Oka & Chang, 1960; Oka & Chang, 1961; Vaughan, 1994; Zhou *et al.*, 2008; Zheng & Ge, 2010). *Oryza rufipogon* thrives in lowland marshes and wetlands with continuous water availability throughout the year; whereas, *O. nivara* occurs in ephemeral pools that have standing water throughout the monsoon season and then desiccate upon the arrival of the dry season (Vaughan, 1994; Kariali *et al.*, 2008; Zhou *et al.*, 2008; Zheng & Ge, 2010). In addition to their distinct habitat preferences, these sister

species differ in their flowering time, mating system and life history. *Oryza rufipogon*, is perennial, out-crossing, and flowers late in the year under a short day length. In contrast, *O. nivara* is annual, predominately self-fertilized, and flowers earlier under a long day length (Oka & Chang, 1960; Vaughan, 1994). Phylogenetic analysis has indicated that *O. nivara* has recently diverged from a common ancestor resembling *O. rufipogon* (Barbier *et al.*, 1991; Zhu & Ge, 2005). The direction of evolution in this system is consistent with broader patterns in plant evolution with annual selfing species being derived from perennial outcrossers (Stebbins, 1974; Barrett *et al.*, 1996). The life cycle of each species is depicted in Figure 2.1. Germination occurs at the onset of the rainy season for both species. *O. rufipogon* is an obligate short day plant flowering in November and December. After flowering, *O. rufipogon* persists and survives throughout the dry season as it inhabits stable wetlands. Conversely, *O. nivara* flowers after a period of sufficient vegetative growth under any day length. *O. nivara* flowers in August and September, prior to the onset of the dry season. As an annual, *O. nivara* dies after seed maturation and seeds remain dormant during the dry season (Fig. 2.1).

Of the traits that differ between these species, flowering time stands out a crucial adaptation for survival in an *O. nivara* habitat. That is, *O. nivara* has adapted to survive in seasonal habitats by being able to flower and produce seeds prior to desiccation. Kariali *et al.* (2008) have monitored populations of these species throughout the growing season in India and documented that *O. nivara* habitats dessicate by early October, prior to the *O. rufipogon* flowering period. As a strict short day plant, *O. rufipogon*, would not be able to reproduce prior to the onset of the dry season in seasonal habitats.

The timing of reproduction is an important stage in the life history of any organism. Variation in the flowering time of plants is commonly observed in nature, and in many cases it

has been demonstrated that this variation has a genetic basis (McIntyre & Best, 1978; Marquis, 1988; Van Dijk & Hautekeete, 2007; Wilczek *et al.*, 2009). Additionally, there is substantial evidence that flowering time is adaptive, as has been revealed through reciprocal transplant experiments (Hall & Willis, 2006), the observation of latitudinal and altitudinal clines in flowering time (Lacey, 1988; Kalisz & Wardle, 1994; Olsson & Ågren, 2002), and contemporary studies of selection (Munguia-Rosas *et al.*, 2011). Plants that do not match their flowering time to the biotic and abiotic environment can incur severe fitness consequences. For these reasons, coupled with the relative ease of measuring, flowering time has likely received more attention than any other adaptive trait.

Furthermore, the molecular genetic pathways underlying the transition from a vegetative state to flowering have been well characterized (Simpson & Dean, 2002; Amasino, 2010). Many plants initiate flowering in response to environmental cues. For example, the photoperiod and vernalization (prolonged exposure to cold temperatures) pathways integrate their respective environmental signals to promote flowering. Moreover, plants can flower in response to endogenous signals including the gibberellin and autonomous pathways (Simpson & Dean, 2002). The molecular pathways controlling flowering have predominantly been characterized in the model plant, *Arabidopsis thaliana*. In *Arabidopsis* over 100 genes have been delineated as part of the flowering time network and many of the core genes are conserved across plant species (Kardailsky *et al.*, 1999; Kobayashi *et al.*, 1999; Böhlenius *et al.*, 2006; Tamaki *et al.*, 2007; Xia *et al.*, 2012). While the molecular underpinnings have been well described, only a subset of these genes have been shown to be associated with natural variation for flowering time (Stinchcombe *et al.*, 2004; Strange *et al.*, 2011). Although *Arabidopsis* is considered a genetic model system, its biology is surely not representative of all angiosperms. *Arabidopsis* is a long

day plant with flowering being promoted by increasing day length as well as vernalization (simpson and dean). In contrast, rice is a facultative short day plant and does not require vernalization. In general, rice and *Arabidopsis* share the same basic genetic machinery for flowering (Yano *et al.*, 2000; Hayama *et al.*, 2002; Kojima *et al.*, 2002; Hayama *et al.*, 2003) however, there are key distinctions that provide flowering under short day lengths (Hayama *et al.*, 2003; Komiya *et al.*, 2009; Takahashi *et al.*, 2009). For example in both rice and *Arabidopsis* the major floral regulator *CONSTANS* (*CO*), or *HEADING DATE1* (*Hd1)* in rice, is controlled by the circadian clock with increased expression under long days (Yano *et al.*, 2000; Izawa *et al.*, 2002). In *Arabidopsis* the increased expression of *CO* during long day lengths promotes *FLOWERING LOCUS T (FT)* expression, which allows for flowering (Kardailsky *et al.*, 1999). Conversely, in rice increased expression of *Hd1* under long days inhibits expression of the *FT* orthologue (*Hd3a*), and thereby flowering (Hayama *et al.*, 2003). While many flowering time genes are conserved between rice and *Arabidopsis*, there are flowering time genes that they do not share (Izawa, 2007). For example, the MADS-box gene *EARLY HEADING DATE 1* (*Ehd1*) activates *Hd3a* independently from *Hd1* (Takahashi *et al.*, 2009). Additionally, this gene has no clear orthologue in *Arabidopsis*. With an available genome sequence and abundant molecular genetic tools, rice has emerged as a model for understanding the basic framework for flowering in short day plants; however, the genetic basis of flowering time variation has not been as thoroughly explored as in *Arabidopsis*.

Previously, I have conducted QTL mapping for a suite of traits that differ between *O. rufipogon* and *O. nivara* (Grillo *et al.*, 2009). Three QTL were identified that control flowering time differentiation in this system. Here I build off this work and focus on the large effect flowering time QTL on chromosome 1. This QTL explained 36.5% of the variation in flowering

time between the focal species and had the largest effect size of any QTL in this mapping study. The accurate estimation of QTL effect sizes is quite challenging and can be influenced by a variety of factors, including the size of the mapping population, the amount of recombination in a given genomic region, and linkage disequilibrium between mutations within a QTL region (Mackay *et al.*, 2009; Rockman, 2012). The ability to accurately quantify QTL effect sizes is critical for our understanding of genetic models of adaptation. Near-isogenic lines (NILs) provide an empirical means to demonstrate the phenotypic effect of an individual QTL by introgressing it into a background genome (Bradshaw & Schemske, 2003; McKay *et al.*, 2003). The goals of the present work are to 1) create a near isogenic line to determine the effect on flowering time of *O. nivara* alleles at this QTL, 2) fine map this QTL region to identify the candidate causal gene, and 3) to determine if a candidate gene differs in coding region sequence.

Materials and Methods

In order to assess the phenotypic effect of individual *O. nivara* alleles, a near-isogenic line (NIL) was created through consecutive rounds of backcrossing and marker assisted selection. To begin this process, an F1 individual was backcrossed to *O. rufipogon*. Information on these accessions is provided in Grillo *et al.* (2009). *O. rufipogon* was selected as the F1 maternal parent as this mimics the direction of evolution, by which *O. nivara* alleles arose in an *O. rufipogon* like nuclear and cytoplasmic genome. To conduct the cross, *O. rufipogon* flowers were manually opened prior to anthesis and were emasculated to prevent selfing. Donor pollen from *O. nivara* was collected upon anthesis by shaking anthers over a glass microscope slide or my fingernail. The slide or fingernail was then brushed against an *O. rufipogon* stigma. After a 3 month period of dormancy the resulting BC1 seeds were germinated and DNA was extracted

from the seedlings with FastDNA kits. These seedlings were scored with a microsatellite marker (Ch1cM33) which was closely linked to the major flowering time QTL peak (Grillo *et al.*, 2009). Seedlings were selected that were heterozygous at this locus and were grown to maturity in the MSU greenhouse. A heterozygous genotype at this locus ensures that the cross was successful (i.e. not the product of geitonogamous selfing in the *O. rufipogon* maternal parent). Once the BC1 plants reached maturity, flowering was induced in growth chambers for another round of backcrossing to *O. rufipogon*. This process was continued for 5 total rounds of backcrossing at which point the background genome was expected be approximately 98.43% homozygous for *O. rufipogon* alleles, but be heterozygous at the QTL of interest. At this point a BC5 individual was self-fertilized to generate NILs that were homozygous for *O. rufipogon* and *O. nivara* alleles at the locus of interest. The BC5 individual was genotyped at 36 microsatellite markers distributed across the genome to ensure they were homozygous for *O. rufipogon* alleles. Three markers were selected on each of the 12 rice chromosomes, including markers that were closely linked to the QTL peak for the 2 other known flowering time QTL (Grillo *et al.*, 2009). In total it took approximately 5 years to generate the NILs. The resulting NILs, as well as *O. nivara* and *O. rufipogon* plants, were grown in the greenhouse with supplemental lights to sustain a long day condition of 16 hours day length. Plants were checked regularly and phenotyped for days to flower.

After 3 rounds of consecutive backcrossing, a BC2 individual was selected to fine map the QTL region. In the BC2 generation approximately 87.5% of the background genome is homozygous for *O. rufipogon* alleles. This BC2 individual was genotyped with microsatellite markers at the other 2 QTL on chromosome 6 and 9 to ensure that it was homozygous *O. rufipogon* at these loci. The BC2 plant was then grown in the greenhouse and vegetativly cloned

extensively, to produce a large number of seeds. The BC2 plants were isolated in a greenhouse room to prevent outcrossing, and were allowed to flower. Several thousand selfed progeny were generated through this process that served as a recombinant mapping population.

Microsattelite markers were then designed to flank the QTL region. These markers were selected by searching the *O. sativa* reference genome (RGAP7) for simple sequence repeats (SSRs). Approximately 30 microsatellite markers were generated and markers flanking the QTL were selected based on their position and ease of deciphering polymorphisms on a 3% agarose gel. The markers selected encompassed a 9 cM region and were located at approximately 27cM and 36cM, with the QTL peak being at approximately 31cM.

After a period of dormancy, the selfed BC2 seeds were germinated and screened at the flanking microsatellites (27 and 36cM). A high throughput DNA extraction protocol developed by Li *et al.* (2006) was used to quickly isolate DNA from over 6,000 seedlings. Seedlings were selected if genetic crossover occurred between the flanking markers. That is, the markers were heterozygous at one marker and homozygous at the other (i.e. RN-NN). DNA was re-extracted from recombinant plants using the FastDNA kits, and they were re-genotyped to confirm that they were true recombinants. Because of the dominant effect of *O. rufipogon* alleles at this locus, only *O. nivara* recombinants (RN-NN) were selected. In *O. rufipogon* recombinants it would be difficult to distinguish between the heterozygous and homozygous phenotype. The recombinant population, along with *O. rufipogon*, *O. nivara*, and F1 control plants, were grown in the greenhouse under a long day condition and phenotyped as described for the NILs.

Single nucleotide polymorphism (SNP) markers were designed to fine map the region. To identify SNPs, approximately 500bp regions of gene introns were sequenced in the region of interest. Sequencing primers were designed to amplify the target sequence based on the O.

sativa reference genome (RGAP7). Gene regions were PCR amplified using a high-fidelity polymerase (PfuUltra II Fusion HS DNA Polymerase, Stratagene; La Jolla, California, USA). Successful amplification was determined on 1.5% agarose gels, and DNA bands were then purified (Wizard SVGel and PCR Clean-Up, Promega). Sanger sequencing was conducted at the Michigan State University Genomics Facility. The position of the amplified fragment was confirmed by a BLAST search of the sequence to the *O. sativa* genome. The location of the SNP markers is provided in Figure 2.2.

The coding region of *FT-LIKE 1* (*FTL1*) was sequenced from both *O. rufipogon* and *O. nivara*. Sequencing primers were designed from the *FTL1* sequence of *O. sativa*. Sanger sequencing was conducted as described previously. The resulting Sequence data was analyzed with the Staden package Pre-gap4 and Gap4 (http://staden.sourceforge.net/). Polymorphisms were compared between the focal species in BioEdit (http://www.mbio.ncsu.edu/bioedit/page2.html). All polymorphisms were confirmed by re-amplification and sequencing of the specific region.

Results

Near-isogenic lines (NILs) are useful in demonstrating the phenotypic effect of a single locus (Bradshaw & Schemske, 2003; McKay *et al.*, 2003). In the present study the NILs that were homozygous for *O. nivara* alleles (i.e. NIL-NN), all flowered (n=21); whereas the NILs with *O. rufipogon* alleles (NIL-RR) did not flower (n=15; Table 2.1). Similarly the *O. nivara* control plants flowered (n=5), whereas *O. rufipogon* (n=5) did not. While both *O. nivara* and the NIL-NN plants flowered, they did so at distinct times. The mean NIL-NN flowering date was

38.3 days after the *O. nivara* mean (Table 2.1). The NIL results provide evidence that this single locus is sufficient to trigger flowering under a long day length.

Table 2.1. Flowering time of Near-isogenic lines (NILs).

O. nivara and *O. rufipogon* are parental plants that serve as controls. NIL-NN refers to a NIL with *O. nivara* alleles at the flowering time QTL. NIL-RR refers to a NIL with *O. rufipogon* alleles at the flowering time QTL.

Plant	Days to flower
O. nivara	104.0 (2.45)
O. rufipogon	did not flower
NIL-NN	142.3 (6.51)
NIL-RR	did not flower

High resolution fine mapping is a classic means for identifying the causal genetic elements influencing trait variation (Salvi *et al.*, 2002; Li *et al.*, 2006; Mackay *et al.*, 2009). The flowering time of the recombinant population (n=56), as well as the parental species and F1s is provided in Figure 2.3. Days to flowering was recorded as the days past germination. The *O. nivara* parents flowered between 98-109 days, the F1s between 147-182 days, and the *O. rufipogon* parents did not flower under these long day conditions. The recombinant plants displayed two distinct groups for flowering time dates. An early group flowered between 112 and 126 days and a later group from 138-155days. The early and late groups were scored as having the *O. nivara* or F1 phenotype, respectively. Integrating the phenotype and genotype of the recombinants allowed for fine mapping of the genetic region (Fig. 2.2). Based on this analysis, the causal gene is located between the SNP markers at 31.0cM and 31.58 cM (physical positions 6370921 and 6512441, respectively).

Within the 141kb sub-region there are 16 genes in *O. sativa* (Table 2.2). Only one of these genes stood out as a clear flowering time candidate based on the gene annotation. This gene, *FT-LIKE 1* (*FTL1*), is a close homolog of the *Arabidopsis* major flowering regulator, *FT*. Analysis of the coding region sequence revealed 5 polymorphisms between *O. rufipogon* and *O. nivara* (Table 2.3). Of these polymorphisms, 3 are nonsynonymous and are candidates for influencing gene function.

Table 2.2. Genes in the fine mapped region.

Genes in the fine mapped region. Gene ID refers to the accession number available from the MSU Rice Genome Annotation Project. Gene annotation provides information on the putative function of the gene.

Gene ID	Gene annotation
LOC_Os01g11790	GDSL-like lipase/acylhydrolase, putative, expressed
LOC_Os01g11800	retrotransposon protein, putative, unclassified, expressed
LOC_Os01g11810	powdery mildew resistant protein 5, putative, expressed
LOC_Os01g11820	expressed protein
LOC_Os01g11830	oryzain alpha chain precursor, putative, expressed
LOC_Os01g11840	xylem cysteine proteinase 2 precursor, putative, expressed
LOC_Os01g11850	expressed protein
LOC_Os01g11860	DJ-1 family protein, putative, expressed
LOC_Os01g11870	expressed protein
LOC_Os01g11880	DJ-1 family protein, putative, expressed
LOC_Os01g11900	hhH-GPD superfamily base excision DNA repair protein, expressed
LOC_Os01g11910	basic helix-loop-helix, putative, expressed
LOC_Os01g11920	retrotransposon protein, putative, unclassified, expressed
LOC_Os01g11940	osFTL1 FT-Like1 homologous to Flowering Locus T gene; contains Pfam profile PF01161: Phosphatidylethanolamine-binding protein, expressed
LOC_Os01g11946	ABC transporter, ATP-binding protein, putative, expressed
LOC_Os01g11952	histone-lysine N-methyltransferase ATX5, putative, expressed

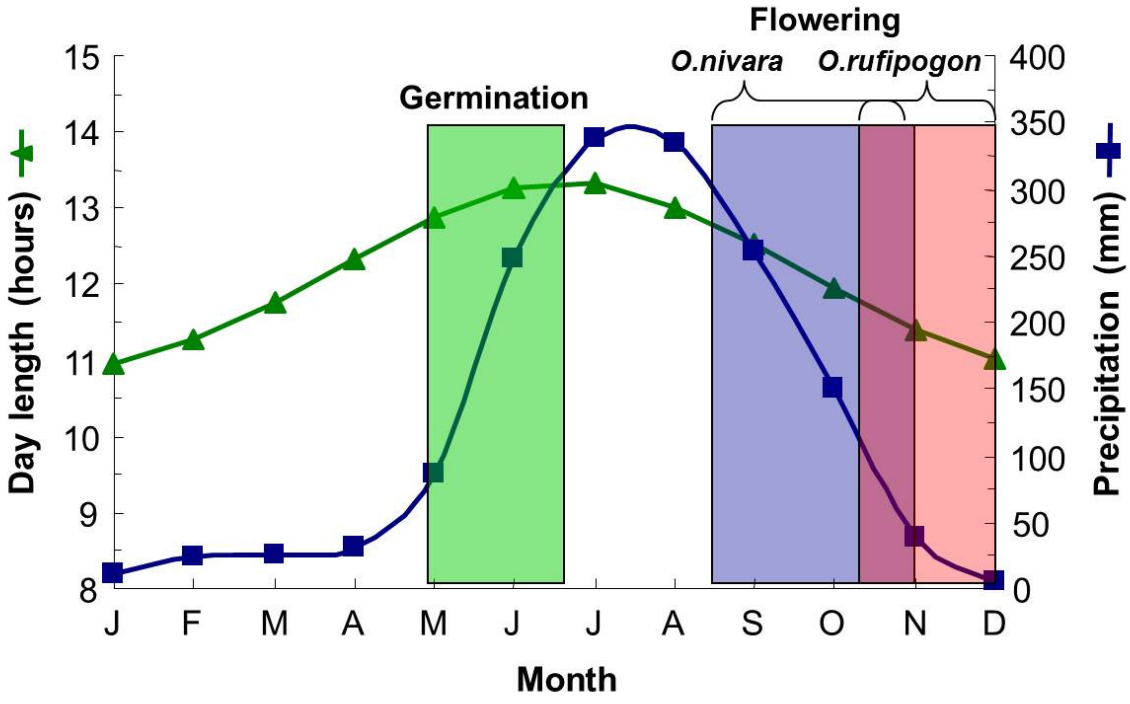
Discussion

Previously, I have identified three QTL that control flowering time differentiation between *O. rufipogon* and *O. nivara*, located on chromosomes 1, 6 and 9 (Grillo *et al.*, 2009). Through consecutive rounds of backcrossing I have introgressed *O. nivara* alleles for the major flowering time QTL into an *O. rufipogon* background, resulting in a near-isogenic line. Near-isogenic lines (NILs) provide a novel approach for determining the phenotypic effect of an individual allele. This procedure is especially relevant given the phylogenetic direction of evolution. Phylogenetic inference suggests that *O. nivara* has recently diverged from an *O. rufipogon*-like ancestor and that early flowering is a derived trait (Barbier *et al.*, 1991; Zhu & Ge, 2005). Therefore, the generation of a NIL closely mimics the process of *O. nivara* alleles arising in the background of the common ancestor resembling *O. rufipogon*. By growing these NILs under a constant long day condition, I have demonstrated that this allele is sufficient at providing photoperiod insensitive flowering. That is, NILs with *O. rufipogon* alleles at this locus do not flower under the long day length, whereas *O. nivara* alleles do flower. It should be noted that while the NILs with *O. nivara* alleles were able to flower under the long day length, they still flowered considerably later than the *O. nivara* parents (Table 2.1). This may be explained by the other QTL on chromosome 6 and 9 that control days to flowering, both of smaller effect than the QTL under investigation in the present study. *O. nivara* alleles at these loci likely promote earlier flowering, after flowering has been achieved by being photoperiod insensitive. One of these QTL harbors a candidate gene that lends support for this hypothesis. The QTL on chromosome 6 encompasses *Hd1* as a candidate, with *O. nivara* alleles possessing a loss of function mutation in this gene (Grillo *et al.*, 2009). Takahashi *et al.* (2009) found Hd1 to be one of the few flowering time genes associated with flowering time variation in cultivated rice, with

loss of function alleles providing early flowering. The NILs presented here contain *O. rufipogon* alleles at the two other QTL which act to delay flowering.

Figure 2.1. Lifecycle of *Oryza* species.

Yearly precipitation and day-length corresponding with the estimated flowering phenology. Data from Cuttack, India where CRRI is located (20ºN, 85ºE). Day-length is for the 1st day of the month (http://www.qpais.co.uk/modb-iec/dayleng.htm). Precipitation is the monthly average from 1867-1973 (www.worldclimate.com). Plant characteristics estimated from the literature, discussions with experts, and per observation (Oka & Chang 1960; Vaughn 1994)



The results of the NIL experiment suggest that flowering time in this system can be portioned into two distinct components: 1) the ability to flower given a photoperiod sensitivity requirement and 2) once flowering is achieved, variation in the days to flower.

A comprehensive understanding of the genetic basis of adaptation requires moving beyond QTL to identify the causal genes and mutations controlling trait variation (Barton & Keightley, 2002; Mackay *et al.*, 2009; Rockman, 2012). In addition to revealing the phenotypic effects of individual loci, NILs are useful starting material for fine mapping QTL regions (Salvi *et al.*, 2002; Werner *et al.*, 2005; Xia *et al.*, 2012). QTL regions can be quite large and encompass hundreds of genes. The QTL region under investigation in the present study was 9cM long (2.2Mbp) and includes over 300 genes. By screening a large recombinant fine mapping population I have substantially narrowed down the QTL region to an approximately 141kbp region that encompasses 16 genes (Table 2.2). The molecular function of these 16 genes has not been rigorously assessed, although they have been annotated with putative functions based on their sequence identity and expression patterns (RGAP7). Of these genes there is one strong candidate, *FT-like1* (*FTL1*) which is homologous to the *Arabidopsis* gene *FT* (Doi *et al.*, 2004; Chardon & Damerval, 2005). In *O. sativa* this gene is highly expressed in the young inflorescence and remains expressed in all floral tissues (www.plexdb.org). Furthermore, this gene is located at approximately 31.5cM which is closely aligned to the position of the QTL peak LOD, 31cM.

Table 2.3. Polymorphisms in the *FTL1* coding region.
Position, position of the polymorphism in *FTL1* based on the *O. sativa* sequence. AA, amino acid number of the polymorphism. *O. sativa* AA, *O. rufipgon* AA, *O. nivara* AA, amino acid abbreviations of these respective accessions.

Position in gene	Exon	AA #	R	N	*S*	*O. rufipogon* AA	*O. nivara* AA	*O. sativa* AA
72	1	24	A	C	C	T	T	T
121	1	41	G	A	G	D	N	D
385	1	129	T	C	C	L	L	L
449	1	150	A	C	C	D	A	A
696	4	232	G	C	G	Q	H	Q

The coding region of *FTL1* was sequenced from *O. rufipogon* and *O. nivara*, revealing 5 SNPs that varied between them. Of these polymorphisms, three are nonsynonmous resulting in a shift in amino acid (Table 2.3). All three of these polymorphisms could impact protein function as the substitutions changed the amino acid polarity and/or charge. Characterizing the role of protein coding and *cis*-regulatory variation is of significant interest in studying the genetic basis of adaptation.

Figure 2.2. Fine Mapping Results.
Fine mapping results. Each line represents an individual recombinant. The numbers at the top indicate the location in cM of SNP and microsatellite markers used for fine mapping. The color indicates the genotype of the recombinant at that marker. The blue color represents a heterozygote and yellow represents an *O. nivara* homozygote. Green indicates the interval that recombination occurred in. The phenotype is given as being either that of the *O. nivara* homozygote (NN) or the heterozygote (RN).

Figure 2.2 (cont'd)

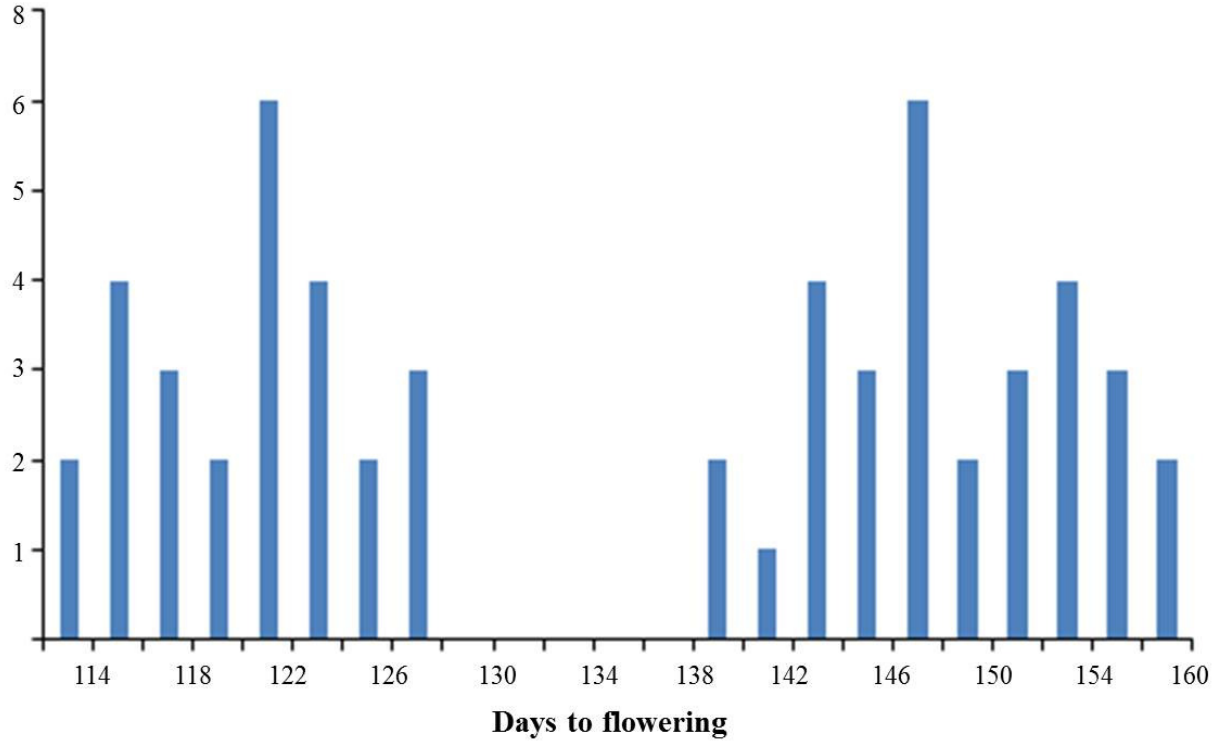
27	**29**	30.65	31	31.25	31.58	32.9	36	Phenotype
N	**N**	N	N	N	N	N	H	**NN**
N	**N**	N	N	N	N	N	H	**NN**
N	**N**	N	N	N	N	N	H	**NN**
N	**N**	N	N	N	N	N	H	**NN**
N	**N**	N	N	N	N	N	H	**NN**
N	**N**	N	N	N	N	N	H	**NN**
N	**N**	N	N	N	N	N	H	**NN**
N	**N**	N	N	N	N	N	H	**NN**
N	**N**	N	N	N	N	N	H	**NN**
N	**N**	N	N	N	N	N	H	**NN**
N	**N**	N	N	N	N	N	H	**NN**
N	**N**	N	N	N	N	H	H	**NN**
N	**N**	N	N	N	N	H	H	**NN**
N	**N**	N	N	N	N	H	H	**NN**
N	**N**	N	N	N	N	H	H	**NN**
N	**N**	N	N	N	H	H	H	**NN**
H	**H**	H	H	N	N	N	N	**NN**
H	**H**	H	H	N	N	N	N	**NN**
H	**H**	H	N	N	N	N	N	**NN**
H	**H**	H	N	N	N	N	N	**NN**
H	**N**	N	N	N	N	N	N	**NN**
H	**N**	N	N	N	N	N	N	**NN**
H	**N**	N	N	N	N	N	N	**NN**
H	**N**	N	N	N	N	N	N	**NN**
H	**N**	N	N	N	N	N	N	**NN**
N	**N**	N	N	H	H	H	H	**RN**
N	**N**	N	H	H	H	H	H	**RN**
N	**N**	N	H	H	H	H	H	**RN**
N	**N**	H	H	H	H	H	H	**RN**
N	**H**	H	H	H	H	H	H	**RN**
N	**H**	H	H	H	H	H	H	**RN**
N	**H**	H	H	H	H	H	H	**RN**
N	**H**	H	H	H	H	H	H	**RN**
N	**H**	H	H	H	H	H	H	**RN**
N	**H**	H	H	H	H	H	H	**RN**
H	**H**	H	H	H	N	N	N	**RN**

Hoekstra and Coyne (2007) highlighted that considerable emphasis has been placed on *cis*-regulatory evolution; however, sufficient data are lacking to make generalizations. In regards to in cultivated rice, both types of mutations have been documented in impacting flowering time variation. For example loss of function mutations in the *Hd1*coding region disrupt protein function, whereas promoter region variation impacts expression of *Hd3a* and thereby flowering time (Takahashi *et al.*, 2009). Functional studies are required to confirm that *FTL1* is the causal gene underlying the QTL, as well as to identify the specific mutation responsible.

FLOWERING LOCUS T, and its homolog *Hd3a*, have proven to be critical regulators of flowering time in both rice and *Arabidopsis* (Kardailsky *et al.*, 1999; Kojima *et al.*, 2002; Tamaki *et al.*, 2007). This gene is conserved across diverse angiosperms and has been documented as the mobile florigen signal (Izawa, 2007; Amasino, 2010). Both *FT* and *Hd3a* encode for a very similar phosphatidylethanolamine-binding protein (PEBP) (Chardon & Damerval, 2005). These proteins are found in both plants and animals, and while their molecular function is not entirely understood they have been shown to be highly involved with flowering in angiosperms. The PEBP family in *Arabidopsis* is quite small and consists of 5 members, all involved in flowering (Chardon & Damerval, 2005). Conversely, this family is much more complex in rice and other grass genomes (Doi *et al.*, 2004). The *FT-LIKE* genes are the largest PEBP subfamily in rice, consisting of 13 members (*FTL1-13*) including *Hd3a* and RICE *FLOWERING LOCUS T1* (*RFLT1/FTL3*) (Chardon & Damerval, 2005). Despite sharing a highly conserved protein, PEBP genes can have quite distinct effects on flowering. For example, in *Arabidopsis FT* promotes flowering in long day lengths, where its close homolog (TERMINAL FLOWER1) delays flowering under these conditions (Kobayashi 99).

Here, I have provided evidence for *FTL1* as a candidate gene within the fine mapped region. A strong argument can be made for the role of *FTL1* in flowering based on its close sequence similarity to *FT/Hd3a* including the PEBP domain, and its documented expression during the onset of flowering in cultivated rice. Moreover, Doi *et al.* (2004) have observed that *EARLY HEADING DATE 1*, a rice gene involved in short-day promotion of flowering, can induce *FTL1* expression in an *Hd1*-deficient background. In the present work, another *FT LIKE* gene, *FTL8*, was also a potential candidate within the large flowering time QTL on chromosome 1. The recombinant mapping results excluded *FTL8* as a causal gene as it does not reside within the fine mapped region. Interestingly, the remaining flowering time QTL on chromosome 9 also encompasses an *FT LIKE* gene (*FTL4*) (Chardon & Damerval, 2005). Functional experiments and high resolution mapping are needed to confirm the role of this gene in underlying flowering time variation at this QTL. Although *FT* and other members of PEBP family are crucial for flowering in *Arabidopsis*, they are not typically associated with variation for flowering time among natural accessions (Strange *et al.*, 2011). Compared to *Arabidopsis*, the *FT LIKE* genes are more diverse in rice and I have provided evidence for their potential role in governing flowering time variation in the wild rice species.

Figure 2.3. Days to flowering in the recombinant mapping population.



Flowering time is a key trait for survival in the seasonally dry habitats that *O. nivara* inhabits. Phylogenetic analysis suggests that the common ancestor of the focal species closely resembled *O. rufipogon* by being perennial and photoperiod sensitive (short day) (Barbier *et al.*, 1991; Zhu & Ge, 2005). It is likely that propagules of the common ancestor frequently dispersed into seasonally dry habitats along the periphery of wetlands. The dry season begins 2-3 months prior to the flowering period of *O. rufipogon* (Kariali *et al.*, 2008). Because of this, propagules of the common ancestor would be unable to persist in these habitats without a major shift in flowering time, allowing them to set seed prior to the onset of the dry season. A shift in

flowering time would thereby be a necessary initial step in order to adapt to this new habitat. Under this scenario minor adjustments in flowering time would not increase survival. *O. rufipogon* is an obligate short day plant, initiating flowering after an approximately 11 hour day length. Conversely, *O. nivara* can flower under any day length provided a sufficient period of vegetative growth (Oka & Chang, 1960). Thus flowering time divergence between these species is based on photoperiod sensitivity. Previously I have identified a large effect QTL that controls days to flowering in this system. The analysis of NILs has revealed that this locus is sufficient at reducing photoperiod sensitivity to allow flowering under a long day length. Therefore, it is highly likely that a mutation at this locus was an initial first step in adapting to seasonally dry habitats. This notion is in agreement with Orr's model of adaptation that predicts that large effect mutations occur early on during an adaptive walk(Orr, 1998).

Understanding the genetic basis of adaptation requires identifying the causal genes and mutations controlling adaptive traits (Mackay *et al.*, 2009; Rockman, 2012). Here, I have provided strong evidence for *FTL1* as a causal gene underlying a major QTL for flowering time, a key trait in this system. Additionally I have identified nonsynonymous polymorphisms in the coding region of this gene that could be the causal mutations. Further experiments are required to validate the role of *FTL1* and specific mutations that control phenotypic differentiation between the focal species. Such experiments could include conducting transgenic complementation, gene knockdown with VIGS or RNAi, assessing gene expression levels throughout floral initiation, and thoroughly examining the promoter region for cis-regulatory variation. Beyond this, variation in this gene could be assessed in numerous populations of each species. Lastly field experiments are required to confirm the role of flowering time and specific flowering time loci in influencing fitness.

LITERATURE CITED

LITERATURE CITED

Amasino R. 2010. Seasonal and developmental timing of flowering. *Plant Journal* **61**(6): 1001-1013.

Barbier P, Morishima H, Ishihama A. 1991. PHYLOGENETIC-RELATIONSHIPS OF ANNUAL AND PERENNIAL WILD-RICE - PROBING BY DIRECT DNA SEQUENCING. *Theoretical and Applied Genetics* **81**(5): 693-702.

Barrett SCH, Harder LD, Worley AC. 1996. The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **351**(1345): 1271-1280.

Barton NH, Keightley PD. 2002. Understanding quantitative genetic variation. *Nature Reviews Genetics* **3**(1): 11-21.

Böhlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O. 2006. CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* **312**(5776): 1040-1043.

Bradshaw HD, Otto KG, Frewen BE, McKay JK, Schemske DW. 1998. Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (Mimulus). *Genetics* **149**(1): 367-382.

Bradshaw HD, Schemske DW. 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* **426**(6963): 176-178.

Bradshaw WE, Emerson KJ, Catchen JM, Cresko WA, Holzapfel CM. 2012. Footprints in time: comparative quantitative trait loci mapping of the pitcher-plant mosquito, Wyeomyia smithii. *Proceedings of the Royal Society B-Biological Sciences* **279**(1747): 4551-4558.

Chardon F, Damerval C. 2005. Phylogenomic analysis of the PEBP gene family in cereals. *Journal of Molecular Evolution* **61**(5): 579-590.

Colosimo PF, Peichel CL, Nereng K, Blackman BK, Shapiro MD, Schluter D, Kingsley DM. 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *Plos Biology* **2**(5): 635-641.

Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M, Yoshimura A. 2004. Ehd1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-Iike gene expression independently of Hd1l. *Genes & Development* **18**(8): 926-936.

Fisher RA. 1930. *The genetical theory of natural selection*. Chicago, IL: University of Chicago Press.

Goodwillie C, Ritland C, Ritland K. 2006. The genetic basis of floral traits associated with mating system evolution in Leptosiphon (Polemoniaceae): An analysis of quantitative trait loci. *Evolution* **60**(3): 491-504.

Grillo MA, Li CB, Fowlkes AM, Briggeman TM, Zhou AL, Schemske DW, Sang T. 2009. Genetic architecture for the adaptive origin of annual wild rice, *Oryza nivara*. *Evolution* **63**(4): 870-883.

Hall MC, Willis JH. 2006. Divergent selection on flowering time contributes to local adaptation in Mimulus guttatus populations. *Evolution* **60**(12): 2466-2477.

Hayama R, Izawa T, Shimamoto K. 2002. Isolation of rice genes possibly involved in the photoperiodic control of flowering by a fluorescent differential display method. *Plant and Cell Physiology* **43**(5): 494-504.

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K. 2003. Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature* **422**(6933): 719-722.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: Evo devo and the genetics of adaptation. *Evolution* **61**(5): 995-1016.

Izawa T. 2007. Adaptation of flowering-time by natural and artificial selection in *Arabidopsis* and rice. *Journal of Experimental Botany* **58**(12): 3091-3097.

Izawa T, Oikawa T, Sugiyama N, Tanisaka T, Yano M, Shimamoto K. 2002. Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering of rice. *Genes & Development* **16**(15): 2006-2020.

Kalisz S, Wardle GM. 1994. Life-history variation in *Campanula americana* (Campanulaceae) - population differentiation. *American Journal of Botany* **81**(5): 521-527.

Kardailsky I, Shukla VK, Ahn JH, Dagenais N, Christensen SK, Nguyen JT, Chory J, Harrison MJ, Weigel D. 1999. Activation tagging of the floral inducer *FT*. *Science* **286**(5446): 1962-1965.

Kariali E, Kuanar SR, Mohapatra PK. 2008. Individual tiller dynamics of two wild Oryza species in contrasting habitats. *Plant Production Science* **11**(3): 355-360.

Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T. 1999. A pair of related genes with antagonistic roles in mediating flowering signals. *Science* **286**(5446): 1960-1962.

Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M. 2002. Hd3a, a rice ortholog of the Arabidopsis FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions. *Plant and Cell Physiology* **43**(10): 1096-1105.

Komiya R, Yokoi S, Shimamoto K. 2009. A gene network for long-day flowering activates RFT1 encoding a mobile flowering signal in rice. *Development* **136**(20): 3443-3450.

Lacey EP. 1988. Latitudinal variation in reproductive timing of a short-lived monocarp, *Daucus carota* (Apiaceae). *Ecology* **69**(1): 220-232.

Leinonen PH, Remington DL, Leppala J, Savolainen O. 2013. Genetic basis of local adaptation and flowering time variation in Arabidopsis lyrata. *Molecular Ecology* **22**(3): 709-723.

Li CB, Zhou AL, Sang T. 2006. Rice domestication by reducing shattering. *Science* **311**(5769): 1936-1939.

Linnen CR, Poh YP, Peterson BK, Barrett RDH, Larson JG, Jensen JD, Hoekstra HE. 2013. Adaptive Evolution of Multiple Traits Through Multiple Mutations at a Single Gene. *Science* **339**(6125): 1312-1316.

Mackay TFC, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics* **10**(8): 565-577.

Marquis RJ. 1988. PHENOLOGICAL VARIATION IN THE NEOTROPICAL UNDERSTORY SHRUB PIPER-ARIEIANUM - CAUSES AND CONSEQUENCES. *Ecology* **69**(5): 1552-1565.

McIntyre GI, Best KF. 1978. STUDIES ON FLOWERING OF THLASPI-ARVENSE L .4. GENETIC AND ECOLOGICAL DIFFERENCES BETWEEN EARLY-FLOWERING AND LATE-FLOWERING STRAINS. *Botanical Gazette* **139**(2): 190-195.

McKay JK, Richards JH, Mitchell-Olds T. 2003. Genetics of drought adaptation in Arabidopsis thaliana: I. Pleiotropy contributes to genetic correlations among ecological traits. *Molecular Ecology* **12**(5): 1137-1151.

Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature* **441**(7096): 947-952.

Mitchell-Olds T, Willis JH, Goldstein DB. 2007. Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics* **8**(11): 845-856.

Munguia-Rosas MA, Ollerton J, Parra-Tabla V, De-Nova JA. 2011. Meta-analysis of phenotypic selection on flowering phenology suggests that early flowering plants are favoured. *Ecology Letters* **14**(5): 511-521.

Oka HI, Chang WT. 1960. Survey of variations in photoperiodic response in wild Oryza species. *Bot. Bull. Acad. Sinica* **1**: 1-14.

Oka HI, Chang WT. 1961. HYBRID SWARMS BETWEEN WILD AND CULTIVATED RICE SPECIES, ORYZA PERENNIS AND O SATIVA. *Evolution* **15**(4): 418-&.

Olsson K, Ågren J. 2002. Latitudinal population differentiation in phenology, life history and flower morphology in the perennial herb *Lythrum salicaria*. *Journal of Evolutionary Biology* **15**(6): 983-996.

Orr HA. 1998. The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution* **52**(4): 935-949.

Orr HA. 2001. The genetics of species differences. *Trends in Ecology & Evolution* **16**(7): 343-350.

Orr HA. 2005. The genetic theory of adaptation: A brief history. *Nature Reviews Genetics* **6**(2): 119-127.

Phillips PC. 2005. Testing hypotheses regarding the genetics of adaptation. *Genetica* **123**(1-2): 15-24.

Rockman MV. 2012. THE QTN PROGRAM AND THE ALLELES THAT MATTER FOR EVOLUTION: ALL THAT'S GOLD DOES NOT GLITTER. *Evolution* **66**(1): 1-17.

Salvi S, Tuberosa R, Chiapparino E, Maccaferri M, Veillet S, van Beuningen L, Isaac P, Edwards K, Phillips RL. 2002. Toward positional cloning of Vgt1, a QTL controlling the transition from the vegetative to the reproductive phase in maize. *Plant Molecular Biology* **48**(5): 601-613.

Simpson GG, Dean C. 2002. Flowering - *Arabidopsis*, the rosetta stone of flowering time? *Science* **296**(5566): 285-289.

Slate J. 2005. Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Molecular Ecology* **14**(2): 363-379.

Stebbins GL. 1974. *Flowering plants: evolution above the species level.* Cambridge, MA: The Belknap Press of Harvard University Press.

Stinchcombe JR, Weinig C, Ungerer M, Olsen KM, Mays C, Halldorsdottir SS, Purugganan MD, Schmitt J. 2004. A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. *Proceedings of the National Academy of Sciences of the United States of America* **101**(13): 4712-4717.

Strange A, Li P, Lister C, Anderson J, Warthmann N, Shindo C, Irwin J, Nordborg M, Dean C. 2011. Major-effect alleles at relatively few loci underlie distinct vernalization and flowering variation in *Arabidopsis* accessions. *Plos One* **6**(5).

Takahashi Y, Teshima KM, Yokoi S, Innan H, Shimamoto K. 2009. Variations in Hd1 proteins, Hd3a promoters, and Ehd1 expression levels contribute to diversity of flowering time in cultivated rice. *Proceedings of the National Academy of Sciences of the United States of America* **106**(11): 4555-4560.

Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K. 2007. Hd3a protein is a mobile flowering signal in rice. *Science* **316**(5827): 1033-1036.

Tanksley SD. 1993. Mapping polygenes. *Annual Review of Genetics* **27**: 205-233.

Van Dijk H, Hautekeete N. 2007. Long day plants and the response to global warming: rapid evolutionary change in day length sensitivity is possible in wild beet. *Journal of Evolutionary Biology* **20**(1): 349-357.

Vaughan DA. 1994. *The wild relatives of rice: a genetic resources bankbook*. Philippines: International Rice Research Institute.

Werner JD, Borevitz JO, Warthmann N, Trainer GT, Ecker JR, Chory J, Weigel D. 2005. Quantitative trait locus mapping and DNA array hybridization identify an *FLM* deletion as a cause for natural flowering-time variation. *Proceedings of the National Academy of Sciences of the United States of America* **102**(7): 2460-2465.

Wilczek AM, Roe JL, Knapp MC, Cooper MD, Lopez-Gallego C, Martin LJ, Muir CD, Sim S, Walker A, Anderson J, Egan JF, Moyers BT, Petipas R, Giakountis A, Charbit E, Coupland G, Welch SM, Schmitt J. 2009. Effects of genetic perturbation on seasonal life history plasticity. *Science* **323**(5916): 930-934.

Xia ZJ, Watanabe S, Yamada T, Tsubokura Y, Nakashima H, Zhai H, Anai T, Sato S, Yamazaki T, Lu SX, Wu HY, Tabata S, Harada K. 2012. Positional cloning and characterization reveal the molecular basis for soybean maturity locus E1 that regulates photoperiodic flowering. *Proceedings of the National Academy of Sciences of the United States of America* **109**(32): E2155-E2164.

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T. 2000. Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the arabidopsis flowering time gene CONSTANS. *Plant Cell* **12**(12): 2473-2483.

Zheng XM, Ge S. 2010. Ecological divergence in the presence of gene flow in two closely related Oryza species (Oryza rufipogon and O. nivara). *Molecular Ecology* **19**(12): 2439-2454.

Zhou HF, Zheng XM, Wei RX, Second G, Vaughan DA, Ge S. 2008. Contrasting population genetic structure and gene flow between Oryza rufipogon and Oryza nivara. *Theoretical and Applied Genetics* **117**(7): 1181-1189.

Zhu QH, Ge S. 2005. Phylogenetic relationships among A-genome species of the genus Oryza revealed by intron sequences of four nuclear genes. *New Phytologist* **167**(1): 249-265.

CHAPTER 3

Genetic architecture of flowering time differentiation between locally adapted populations of *Arabidopsis thaliana*

The following chapter was published in *New Phytologist*

Grillo MA, Li CB, Hammond M, Wang LJ, Schemske DW. 2013. Genetic architecture of flowering time differentiation between locally adapted populations of *Arabidopsis thaliana*. New Phytologist 197(4): 1321-1331.

Abstract

To gain an understanding of the genetic basis of adaptation, we conducted quantitative trait locus (QTL) mapping for flowering time variation between two winter annual populations of *Arabidopsis thaliana* that are locally adapted and display distinct flowering times. QTL mapping was performed with large (n=384) F2 populations with and without vernalization, in order to reveal both the genetic basis of a vernalization requirement and that of variation in flowering time given vernalization. In the Non-Vernalization treatment, none of the Sweden parents flowered, while all of the Italy parents and 42% of the F2s flowered. We identified three QTL for flowering without vernalization, with much of the variation being attributed to a QTL co-localizing with *FLOWERING LOCUS C (FLC)*. In the Vernalization treatment all parents and F2s flowered, and six QTL of small to moderate effect were revealed, with underlying candidate genes that are members of the vernalization pathway. We found no evidence for a role of *FRIGIDA* in regulating flowering times. These results contribute to a growing body of evidence aimed at identifying ecologically relevant genetic variation for flowering time in *Arabidopsis,* and set the stage for functional studies to determine the link between flowering time loci and fitness.

Introduction

Throughout a species' range, populations adapt to local conditions. Understanding the genetic basis of such local adaptation has been a major goal in evolutionary biology for over a century. The last decade has witnessed profound advances in molecular genetic approaches for finding genomic regions associated with adaptive traits, allowing for the rigorous assessment of this long-standing question (Barton & Keightley, 2002).

Plants have developed sophisticated physiological, developmental and genetic mechanisms to optimize the time of flowering (Bastow & Dean, 2003). Because the transition from vegetative growth to flowering is crucial for plant reproductive success, plants integrate different environmental cues to achieve a flowering response that is adapted to local conditions. The observation that the flowering time of many plant populations varies with latitude or altitude suggests that this trait contributes to geographic adaptation (Lacey, 1988; Kalisz & Wardle, 1994; Olsson & Ågren, 2002). Hence, understanding the pathways and genetic mechanisms that contribute to natural variation in flowering time is a central goal of many studies in plant evolutionary genomics.

The model plant *Arabidopsis thaliana* (hereafter referred to as *Arabidopsis*), serves as a workhorse for the study of plant molecular biology. This species is distributed widely throughout northern latitudes and exhibits substantial natural variation in flowering time. Two different approaches have been used to characterize the life history and flowering behavior of *Arabidopsis* (Nordborg & Bergelson, 1999). The *ecological* criterion is based on the timing of seed germination and flowering; winter annuals germinate in fall, overwinter as rosettes, and flower in late spring, while summer annuals germinate in spring and flower in mid to late summer. Based on this criterion, *Arabidopsis* has a winter annual life history throughout most of its range, although summer annuals have also been described (Napp-Zinn, 1985; Pigliucci, 2003; Koornneef *et al.*, 2004; Shindo *et al.*, 2007). Climatic data from the southern edge of the native range, indicates that while soil temperatures very rarely fall below freezing they do fall below 4°C (Ågren & Schemske, 2012), which is a sufficient temperature for a vernalization response (Nordborg & Bergelson, 1999; Shindo *et al.*, 2005).

In contrast, the *physiological* criterion classifies plants as winter annuals if vernalization, i.e. a period of cold temperature at the seed or rosette stage, is required for flowering, and as summer annuals if plants can flower without vernalization. These categories are also referred to, respectively, as "early" or "late" flowering (Clarke *et al.*, 1995; Gazzani *et al.*, 2003) because the vegetative phase is shorter in plants that do not require vernalization. Some populations flower without vernalization in the lab, and would therefore be classified as summer annuals by the physiological criterion, but are winter annuals under the ecological criterion (Schemske and Ågren, pers. comm.). We suggest that the ecological criterion for classification of life history be adopted universally, because it best reflects the actual timing of growth and flowering in the field.

The genetic basis of flowering time in *Arabidopsis* has received considerable attention (Simpson & Dean, 2002; Sung & Amasino, 2004; Amasino, 2010). Analysis of *Arabidopsis* mutants has characterized multiple genetic pathways that regulate the transition from vegetative to reproductive growth. The photoperiod and vernalization pathway regulate the response to environmental signals, while the autonomous and gibberellin pathways respond to endogenous signals and are functionally independent from environmental cues (Simpson & Dean, 2002). These pathways form a complex network of more than 100 genes that regulate flowering. However, only a few genes identified from mutant screens have also been associated with natural variation in flowering time. Of these, *FRIGIDA (FRI)* and *FLOWERING LOCUS C (FLC)* are thought to be major regulators of flowering in natural populations (Burn *et al.*, 1993; Lee *et al.*, 1993; Clarke & Dean, 1994; Lempe *et al.*, 2005). *FLC*, a transcription factor encoding a MADs-box protein, represses the expression of other transcription factors promoting flowering (Michaels & Amasino, 1999). Functional alleles at *FRI* delay flowering by activating the strong

expression of *FLC* (Johanson *et al.*, 2000). Vernalization represses *FLC* expression by epigenetic modification, reducing its sensitivity to *FRI* and thus promotes flowering (Amasino, 2004).

Molecular analysis of natural accessions has revealed considerable allelic variation for *FRI* and *FLC*, and many mutations are known that cause inactivation of either *FRI* or *FLC*. Such mutants are often significantly associated with early flowering (Johanson *et al.*, 2000; Le Corre *et al.*, 2002; Gazzani *et al.*, 2003; Lempe *et al.*, 2005). Over 70% of early flowering accessions contain loss-of-function alleles at the *FRI* locus (Shindo *et al.*, 2005). In a geographic survey of the relationship between flowering time and *FRI*, active *FRI* alleles were associated with a latitudinal cline in flowering, but no such cline was detected in accessions with nonfunctional *FRI* (Stinchcombe *et al.*, 2004). In the presence of functional *FRI* alleles, *FLC* allelic variation contributes to a latitudinal cline in flowering, and *FLC* expression is most highly correlated with flowering time variation (Caicedo *et al.*, 2004; Lempe *et al.*, 2005; Shindo *et al.*, 2005). These results suggest that *FRI* and *FLC* are targets of natural selection (Stinchcombe *et al.*, 2004; Izawa, 2007). Although *FRI* and *FLC* are important determinants of natural variation in flowering time for some *Arabidopsis* populations, extensive molecular and genetic analysis of *Arabidopsis* accessions has indicated that other loci also contribute to flowering time (Lempe *et al.*, 2005; Werner *et al.*, 2005a; Werner *et al.*, 2005b; Li *et al.*, 2006)

Quantitative trait locus (QTL) mapping has proven to be a powerful method for identifying the genetic basis of quantitative trait variation (Tanksley, 1993; Barton & Keightley, 2002; Mitchell-Olds & Schmitt, 2006), and numerous QTL studies have been conducted examining flowering time variation in *Arabidopsis*. The majority of these studies include mapping populations derived from one or both of the laboratory strains Columbia (Col) and

Landsberg erecta (Ler), both of which could be classified as summer annuals based on a physiological criterion alone (Lister & Dean, 1993). Knowing where in the life cycle variation in flowering time is manifest is a critical step in assessing the mechanisms responsible for population differences in flowering time. For example, in plants requiring vernalization, the developmental and physiological pathways that contribute to variation in flowering time among populations may differ from those of populations that do not require vernalization. It is thus important to understand the life histories of populations grown under conditions of their native environments.

Recently, Ågren & Schemske (2012) have presented the results of a multi-year reciprocal transplant study between *Arabidopsis* populations originating from Sweden and Italy – the geographic limits of the native distribution. This experiment is the first of its kind involving *Arabidopsis* (Lowry, 2012), and has demonstrated strong adaptive differentiation of these populations to their source environment. In this study, freezing tolerance and flowering time were identified as putative adaptive traits conferring local adaptation. At the Italian field site, the Italy population flowered 33 and 50 days prior to Sweden, in the two years of study; at the Swedish field site the Italy population flowered 3 days prior to Sweden in both years (Ågren & Schemske, 2012). Here we report the results of extensive genetic mapping experiments designed to examine the genetic basis of variation in flowering time differences between these populations. Given the winter annual life history of the study populations, we employed a two tiered approach to dissect the genetic basis of flowering time. We first searched for QTL required for plants to flower in the absence of vernalization, and then identified QTL that contribute to variation in flowering following vernalization. In this way, we are able to decouple flowering time variation from a vernalization requirement.

Our study addressed the following questions: 1) what is the number, location, and magnitude of effect for QTL contributing to flowering time? 2) What are the genetic regions responsible for differences in vernalization response and do these contribute to flowering time variation after vernalization? 3) Do QTL co-localize with *FRI* and *FLC* as is seen in other studies? 4) Do candidate genes underlying *QTL* differ between parental populations in coding region sequence?

Materials and Methods

This experiment is part of an ongoing effort to understand the mechanisms of adaptation in natural populations of *Arabidopsis*. In this study, we used one population collected from central Italy, Castelnuevo di Porto (N 42°07' E 12°29'), and one from north-central Sweden, Rödåsen (N 62°48' E 18°12'), as mapping parents. These individuals are derived from the same maternal lines utilized by Ågren & Schemske (2012) -- additional information on these populations is presented there. Both the Italy and Sweden populations are winter annuals; plants germinate from late August to early September in Sweden and from late October to mid-November in Italy, and in both regions overwinter as rosettes. The flowering phenology differs between the two areas, with the Italy population flowering and fruiting much earlier (March-April) than the Sweden populations (May-June) (Ågren & Schemske, 2012). A plant from Italy was utilized as the maternal parent and was crossed to Sweden to produce an F1 generation. A large F2 population was then generated by autonomous self-pollination of an F1.

Arabidopsis seeds were cold stratified for 4 days at 6°C and then were sown on nutrient agar and then allowed to germinate in a growth chamber under long-day conditions (16h). Twenty-day old seedlings were transplanted into 2 inch pots and were allowed to grow for an

additional 4 weeks (total age ~ 7 week), at which time they were randomly assigned to Non-vernalization or Vernalization or treatments. The Non-vernalization treatment proceeded at 22°C under a long-day condition with 16 hours of light (total PAR= ~ 125 µmol m^{-2} s^{-1}). Plants in the Vernalization treatment experienced 5 weeks of vernalization at 6°C with a short day condition of 10 hours of light (Total PAR= ~ 50 µmol m^{-2} s^{-1}). After this 5 week period, the conditions were switched back to long days as described above in the Non-vernalization treatment. For each treatment there were 50 plants from each parental population and 384 F2 hybrids

The dates of bolting and of first flowering were recorded for each plant. The date of bolting was defined as the date the inflorescence extended beyond the rosette. Date of first flowering was the date on which the first flower opened, defined as the stage at which the petals reflexed to expose the stigma. In the Non-vernalization treatment, plants were coded as 1 if they flowered and 0 if they did not flower. For the Vernalization treatment, plants were coded as above, and the date that plants flowered was recorded. For data analysis, plants were scored as the number of days relative to the first plant that bolted or flowered.

Simple sequence repeats (SSR) were utilized for genotyping the mapping population. 99 SSR primer pairs were obtained from the TAIR website and another 202 SSR primer pairs were designed after searching for microsatellites in the *Arabidopsis* genome of the Columbia accession. The total, 301 SSR loci, were screened for polymorphisms between the Italy and Sweden mapping parents. 103 polymorphic markers were identified; of these 71 were used for segregation analysis in 96 F2 plants. Segregation distortion was examined by chi-square tests ($P < 0.05$), to identify markers that deviated from an expected 1:2:1 ratio (Lu *et al.*, 2002). In the end, 62 non-distorted markers were selected to construct the linkage map, providing even coverage of the *Arabidopsis* genome. DNA was isolated using FastDNA kits (Q-BIOgene,

Carlsbad, CA). SSRs were amplified by polymerase chain reaction (PCR) and were visualized on 3 % agarose gels stained with ethidium bromide.

The genetic map was constructed based on 62 SSR markers and 768 F2 plants from the combined Vernalization and Non-vernalization treatment. The linkage analysis was performed based on 706 F2 individuals using JoinMap 3.0 (Van Ooijen & Voorips, 2001). The linkage groups were separated using the highest logarithm of odds (LOD), 10.0. Recombination values were converted to genetic distances using the Kosambi mapping function. Five linkage groups were obtained corresponding to the five *Arabidopsis* chromosomes. For QTL analysis the mapping populations included 384 F2 plants in both the Vernalization and Non-vernalization treatments. In the Vernalization treatment QTL analysis was conducted with the raw flowering time data and quantile normalized flowering times. Both the raw and transformed data yielded essentially the same results, with the same number of QTL being identified and the effect sizes and peak LOD positions changing only slightly. Here we present QTL results with the raw flowering time data. QTL analysis was performed with the R/qtl package (Broman *et al.*, 2003). We used a Haley-Knott regression (Haley & Knott, 1992) in the "stepwiseqtl" function to detect QTL, with the normal and binary model being utilized in the Vernalization and Non-vernalization treatments, respectively. Genome-wide significance (P =0.01) thresholds for QTL detection were determined with 10,000 permutations. For all QTL, we determined 2-LOD confidence intervals, as well as estimates of additive and dominance allelic effects. Epistasis between QTL was tested using two-way analysis of variance (ANOVA) (Westerbergh & Doebley, 2002; Morgan & Mackay, 2006). The markers closest to the LOD peaks of QTL were tested for interaction.

A comprehensive list of potential candidate genes was compiled by searching TAIR 10 (Brachi *et al.*, 2010). Candidates were selected that contain "flowering" or "vernalization" in their name and/or description. This list includes a total of 165 unique genes. To reveal candidate genes underlying QTL, the sequence of SSR markers flanking QTL were located within the Columbia genome. The locations of the flowering/vernalization genes were located in the Columbia genome using the Chromosome Map Tool, available on TAIR. Genes that lie within markers flanking QTL were considered potential candidates. Candidate genes largely fell into three categories: 1) Genes that are involved in basic biological processes with mutants/transgenics showing a suite of abnormal phenotypes, including flowering; additionally published work is lacking or no direct evidence of a role in regulating flowering time or development has been documented. 2) Genes that are involved in development during the transition from vegetative growth to flowering; typically these genes have been well characterized, yet have no known role in controlling the initiation of flowering. 3) Genes that are members of the photoperiod or vernalization pathway in controlling flowering time, most of which have been well studied and identified as candidates in published QTL studies, or associated with natural variation in flowering time. These categories were classified as weak, medium or strong candidates respectively.

The coding regions of several strong candidate genes underlying QTL were sequenced in an effort to identify putative casual polymorphisms (Salome *et al.*, 2011). Sequencing primers were designed to amplify the gene regions based on the Columbia genome. Gene regions were PCR amplified using a high fidelity polymerase (*PfuUltra* II Fusion HS DNA Polymerase, Stratagene). Successful amplification was determined by running PCR products on a 1.5% agarose gel. DNA bands were then cut from the gel and purified (Wizard SV Gel and PCR

Clean-Up, Promega). PCR products were sequenced at the MSU Genomics Facility. To ensure the correct gene region was amplified, the sequence was compared to the Columbia genome via a BLAST search. Sequence data was analyzed with the Staden package Pregap4 and Gap4. Polymorphisms were compared between the mapping parents in BioEdit.

Table 3.1. Results of quantitative trait locus (QTL) analysis of flowering time.

Ch, chromosome on which the QTL is located. Position, peak logarithm of odds (LOD) position on the chromosome in cM. Support interval, 2 LOD. PVE, percentage phenotypic variance explained by the QTL. Nearest marker, simple sequence repeat (SSR) marker closest to the QTL

Treatment	QTL	Ch	Position	Support interval	LOD score	PVE	Additive effect (se)	Dominance effect	Nearest marker
NonVern	NVF1	1	70	60.00-82.00	9	5.9	-1.50 (0.25)	-0.16 (0.34)	F5114
	NVF2	5	8	5.96-10.00	41.1	32.8	-3.80 (0.49)	-0.42 (0.50)	NGA158
	NVF3	5	60	48.01-68.0	9.1	6	-1.30 (0.24)	0.78 (0.32)	C1W9
Vern	VF1	1	80.04	69.71-83.59	4.8	3.1	-1.50 (0.33)	0.36 (0.51)	MSAT1.5
	VF2	3	18.8	11.6-31.44	6.7	4.4	1.62 (0.36)	-1.84 (0.50)	MSAT3.2
	VF3	3	63.19	56.74-63.19	9.2	6.1	-2.0 (0.33)	-1.08 (0.47)	NGA6
	VF4	5	7.9	1.99-12.05	15.3	10.5	-3.42 (0.40)	0.27 (0.50)	NGA158
	VF5	5	37.88	26.09-44.49	4.5	2.9	-1.88 (0.45)	-0.95 (0.52)	T5E15
	VF6	5	68.54	62.71-74.0	4.5	2.9	-1.78 (0.40)	-0.39 (0.48)	MCK7-1

Results

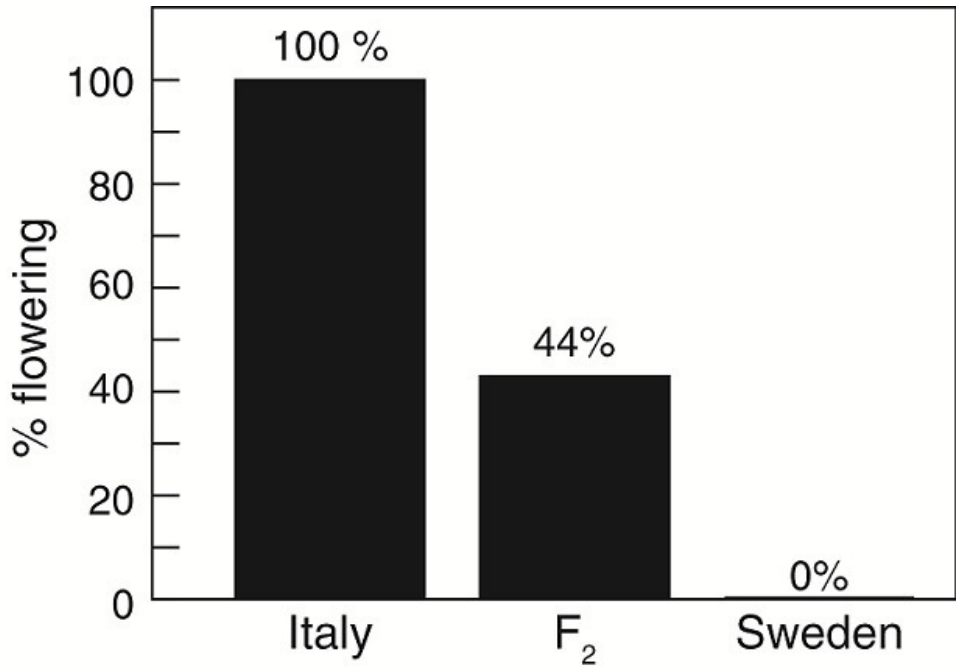
The days to bolting and flowering were shown to be highly correlated (R^2=0.94), with plants flowering an average of 6.2 days after bolting. Because of this, only days to flowering was used for analysis**.** In the Non-vernalization treatment, all (42/42) of the parental plants from Italy (Fig. 3.1) and 43% (165/384) of the F2 plants flowered (Fig. 3.2), but none of the Sweden parents flowered (0/42). Thus, the Italy population is able to flower without vernalization. For the Vernalization treatment, all of the parents from Italy (44/44) and Sweden (40/40), as well as all of the F2s, flowered. Since all plants flowered in this treatment, the five week cold treatment was sufficient for saturating vernalization. While all plants flower in this treatment, there is considerable flowering time differentiation between populations, with the mean flowering date for Italy being 28 days earlier than Sweden (Fig. 3.3). Transgressive segregation was not detected in this mapping population, which is in contrast to other studies (El-Lithy *et al.*, 2006; O'Neill *et al.*, 2008; Brachi *et al.*, 2010; Salome *et al.*, 2011; Seymour *et al.*, 2012).

A total of 9 QTL were identified in this study with 3, and 6, QTL being identified in the Non-vernalization and Vernalization treatment, respectively (Table 3.1). Of these, 8 were unique, as one of the QTL in the Vernalization (VF4) and Non-vernalization (NVF2) treatments mapped to the same chromosomal location, with overlapping LOD support intervals and identical QTL peaks. While this shared QTL co-localizes between treatments, the effect size differs, with a PVE of 32.8 and 10.5, in the Non-vernalization and Vernalization treatments, respectively. Two other pairs of QTL between treatments (NFF1/VF1 and NVF3/VF6) have slightly overlapping LOD support intervals, but distinct LOD peaks and so are considered separate QTL.

QTL effect sizes ranged from 2.9 to 32.8 (PVE) with 7 of the 9 QTL having effect sizes under 10 PVE. No QTL were detected on chromosome 2 or 4. With the exception of one QTL (VF2), the Italy alleles result in early flowering and the Sweden alleles in later flowering. Significant epistasis between QTL was identified in both treatments. In the Non-vernalization treatment, all three of the QTL were involved in an epistatic interaction. In this treatment the large effect QTL (NVF2) which co-localizes with FLC, interacts with both of the other QTL. In the Vernalization treatment there is a marginally significant interaction of the QTL that co-localizes with FLC (VF4) and VF1, as well as one other interacting pair of QTL.

Figure 3.1. Proportion of plant flowering.

Proportion of *Arabidopsis thaliana* populations that flowered in the Non- vernalization treatment. The number above each bar indicate the percentage of each group that flowered.



In this study, 7 of the 9 QTL intervals harbor at least one strong candidate gene (Table 3.1). Of these, all but two, *DAY NEUTRAL FLOWERING (DNF)* and *LATE FLOWERING (LATE),* have been discussed as candidates in previous QTL studies. The QTL VF2 encompasses *DNF*, as well as two other common candidates, *Vernalization Insensitive 3-like 1* (*VIL1*) and *REDUCED VERNALIZATION RESPONSE 1* (*VRN1*). *LATE* is the only strong candidate underlying NVF3, which has recently been characterized as controlling expression of certain flowering genes in leaf vasculature and floral meristem identity genes at the shoot apex (Weingartner *et al.*, 2011). Two QTL in the Vernalization treatment (VF3 and VF5) do not possess a strong candidate gene. A possible, yet unlikely candidate for VF5 is *SERRATED LEAVES AND EARLY FLOWERING (SEF); SEF* mutants show a suite of abnormal flowering phenotypes (March-Diaz *et al.*, 2007). Two developmental genes underlie VF3, *LATE MERISTEM IDENTITY2 (LMI1)* and *GLABROUS INFLORESCENCE STEMS (GIS),* but they have not been studied in regards to controlling flowering time.

Recently, Salome *et al.* (2011) identified flowering time QTL found across a range of *Arabidopsis* accessions and found that these genomic regions include the candidate genes *FRI*, *FLC*, *MAF1*, and *MAF2-5*. They sequenced and compared the coding regions of these candidates across accessions (Salome *et al.*, 2011). For comparison to their results, we sequenced the coding regions of the same genes (*FRI, FLC, MAF1*, and *MAF2*), and three other strong candidates underlying QTL in our mapping population (*VIN3, VIL1*, and *FT*). Polymorphisms between the mapping parents are presented in Table 3.2. All of the genes except for *FLC* have at least one non-synonymous substitution between the Italy and Sweden mapping parents. Many of the polymorphisms found between the mapping parents were also identified by Salome *et al.,* (2011). Only one indel was identified; a 3bp indel in exon 5 of *VIN3* (Table 3.2).

Discussion

Here we report a QTL mapping study of flowering time differentiation utilizing a large mapping population generated from two natural accessions of winter annuals. In the present study, QTL mapping was conducted with and without a vernalization treatment, allowing us to decouple a vernalization requirement from flowering time variation per se. In the Non-vernalization treatment all of the parental plants from Italy and none from Sweden flowered. Plants from Italy are therefore able to flower in the absence of vernalization, yet in nature they behave as true winter annuals-over wintering as rosettes and flowering in the spring (Ågren & Schemske, 2012). Shindo *et al*., (2005) examined vernalization response in accessions collected over a wide latitudinal range. When plants did not receive vernalization, all of the accessions collected above 62° N did not flower, all of the accessions collected below 45° N did flower, and intermediate latitudes displayed a range of responses. The Italy and Sweden populations in this study were collected at 42°07'N and 62°48'N, respectively.

Table 3.2 Polymorphisms of candidate gene coding regions between the Italy and Sweden *Arabidopsis thaliana* mapping parents.

Position, position of the polymorphism in the Col genome. AA, amino acid number of the polymorphism. Italy, Sweden and Col, sequence of these respective accessions. Italy, Sweden and Col AA, amino acid abbreviations of these respective accessions.

Table 3.2 (cont'd)

Gene	Position	Exon	AA	It	Sw	Col	Italy AA	Sw AA	Col AA
FLC	3173805	8	156	G	A	G	V	V	V
FRI	269462	1	146	C	G	A	A	G	E
	269469	1	148	G	G	A	M	M	I
	269962-269963	1	312-			16bp del			
FT	24331561	1	18	G	A	G	V	I	V
	24333548	4	129	G	G	C	G	G	G
MAF1	28958968	6	141	A	G	G	K	E	E
	28959821	9	189	A	G	G	P	P	P
MAF2	25982475	1	21	G	T	T	V	F	F
VIL1	8877617	2	159	T	G	G	V	G	G
	8876928	3	363	T	G	G	I	R	R
	8876600	3	472	G	A	A	R	R	R
	8876598	3	473	A	G	G	Q	R	R
	8876464	3	518	C	C	T	P	P	S
VIN3	23248316	3	79	C	A	A	L	M	M
	23248305	3	82	G	G	A	E	E	E
	23247644	4	274	A	C	A	M	L	M
	23247572	4	298	A	A	G	T	T	A
	23246987-23246893	5	502-503			3bp del	QG	R	QG

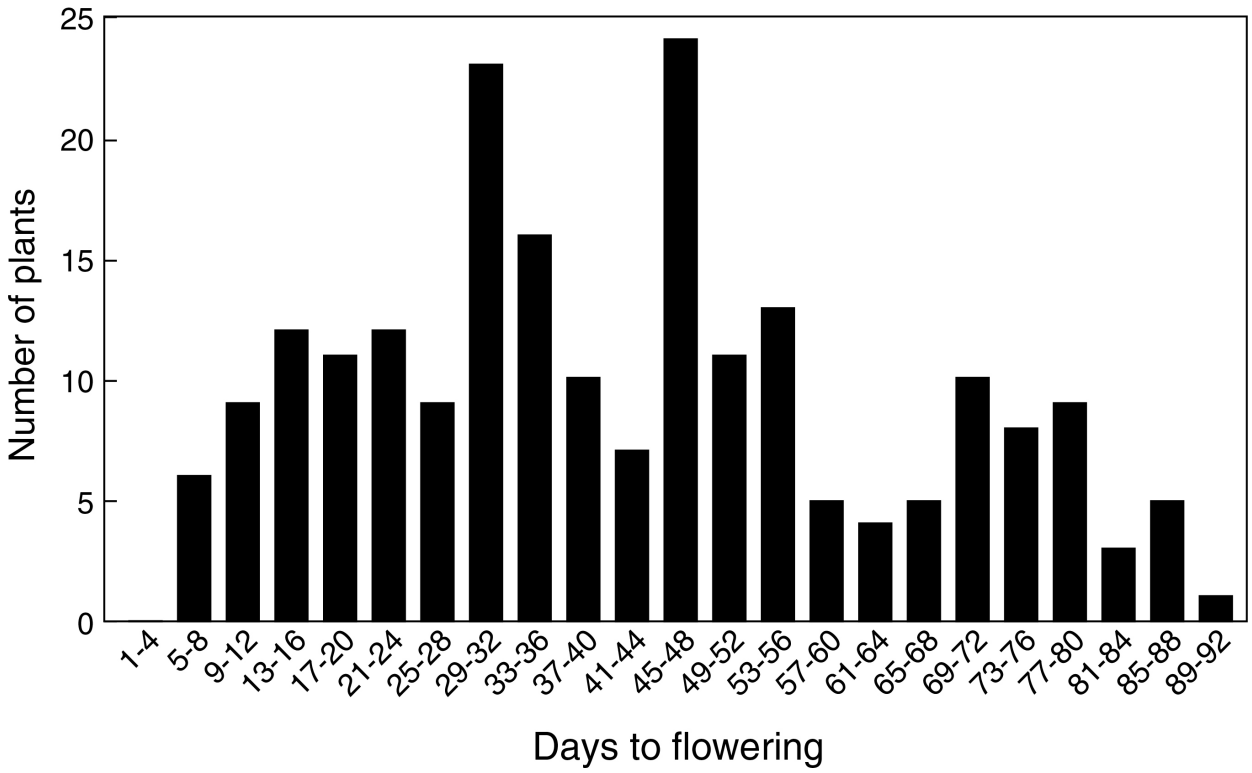
While all of the parents from Italy flowered, it should be noted that there was variation in the time required to initiate flowering, with plants flowering over a period of 36 days (Suppl. Fig. 3.1). Under the constant temperature and day length conditions of the Non-vernalization treatment, plants initiate flowering through an autonomous mechanism which likely does not

provide as consistent of a flowering response as environmental cues such as photoperiod and temperature. In the Vernalization treatment all of the Italy and Sweden plants flower, as do the F2s. Following vernalization, the parental plants from Sweden flower on average 28 days later than those from Italy (Figure 3.3). The earlier flowering of Italy after a vernalization treatment suggests the Italy plants are able to initiate the transition from vegetative to flowering stages after a much shorter period of cold exposure than Sweden. This is consistent with data from the Italian field site, where the Italy populations flower substantially earlier than Sweden (Ågren & Schemske, 2012).

A major goal of evolutionary biology has been to characterize the number and effect size of genes responsible for adaptive variation. Numerous QTL studies of adaptive traits have provided support for Orr's (1998) model; predicting an exponential distribution of effect sizes with few of large effect (Barton & Keightley, 2002). For nearly 20 years, Arabidopsis researchers have been dissecting the genetic basis of flowering time. We outline the traditional linkage (QTL) mapping publications; this includes 26 publications, with 98 different QTL experiments. These studies differ in the mapping parents utilized, size and type of mapping populations, vernalization treatment, photoperiod, and growing conditions. The average number of QTL identified is 4.03; we have detected a typical number of QTL, with studies ranging between one (O'Neill *et al.*, 2008; Balasubramanian *et al.*, 2009) and ten QTL (Weinig *et al.*, 2002). Our results are in agreement to most studies where at least one large effect QTL is identified, but in contrast to studies that detect a higher occurrence of large effect QTL (Loudet *et al.*, 2002; Salome *et al.*, 2011; Strange *et al.*, 2011). The ability to detect QTL of small effect is limited by the size of the mapping population (Beavis, 1998). Many previous QTL studies for flowering time involved mapping populations with less than 150 individuals, and therefore

limited ability to detect small effect QTL. Here we are able to identify QTL with effect sizes as low as 2.9% PVE. The distribution of QTL effect sizes in the present study provides additional support for Orr's (1998) model. Additionally, we have detected a moderate amount of epistasis between QTL. The QTL that co-localize with *FLC* (NVF2 and VF4) are involved in the majority of the significant QTL x QTL interactions. This may be expected due the central role of *FLC* in genetic pathways of floral initiation (Simpson & Dean, 2002).

Figure 3.2 Days to flowering in the F2 plants.



Understanding the role of genotype-environment interaction has been a long-standing goal in evolutionary biology as this can influence local adaptation (Kawecki & Ebert, 2004). Several studies in *Arabidopsis* have characterized significant QTL x environment interactions (Ungerer *et al.*, 2002; Weinig *et al.*, 2002; Li *et al.*, 2006). The ability to detect QTL co-localizing with *FRI* can depend on the environmental conditions. For example, in the extensively utilized mapping population, Col x Ler, detection of *FRI* varies with vernalization and photoperiod (Stratton, 1998; Weinig *et al.*, 2002; Weinig *et al.*, 2003; Juenger *et al.*, 2005). Here we report distinct QTL profiles between the Vernalization and Non-vernalization treatments, with only one QTL shared between them. The effect size of this shared QTL differs substantially between the treatments (Table 3.1). In the Non-vernalization treatment the QTL is of large effect, with a PVE of 32.8%, as opposed to 10.5% in the Vernalization treatment. *FLC* stands out as a candidate gene for the shared QTL (NVF2 and VF4). *FLC* is a strong inhibitor of flowering and is routinely identified in studies examining flowering time variation. Another candidate in the Non-vernalization treatment is *FT*; this QTL was not detected after vernalization.

In an analogous experiment, Strange *et al.,* (2011) examined the effect of different vernalization treatments on mapping populations of four accessions from Sweden crossed with Col. Overall, they also found that the effect size of QTL decreased after vernalization. More specifically, the effect size of QTL co-localizing with *FLC* and *FT* went from having a strong effect with no vernalization, to being not significant after 14 weeks of vernalization (Strange *et al.*, 2011). Our results suggest that differences between Italy and Sweden at the *FLC* locus allows for populations from Italy to flower in the absence of vernalization. Vernalization acts to repress *FLC* by epigenetic modification, allowing for the transition from vegetative to flowering

states (Amasino, 2004). *FLC* allelic variation has been shown to be associated with a latitudinal cline in flowering time variation. In southern populations alleles that weaken *FLC* allow for flowering without vernalization. It is to be expected that *FLC* would have a smaller effect size in the Vernalization treatment because vernalization would provide the strong repression of *FLC* in both the parental populations. The Non-vernalization treatment yields QTL with candidates involved in the photoperiod regulation of flowering (*FLC* & *FT*) where the Vernalization treatment involves candidates of the vernalization pathway (*VRN1, VIL1, VIN3*). In the present work only vernalization was varied between treatments; future studies should include photoperiods similar to those in the habitats of both populations.

Many of the QTL identified here have major candidate genes that map to the same chromosomal location (Table 3.1). It is possible that these candidate genes are the causal genetic elements underlying QTL, however, functional experiments are required to confirm their role in flowering time variation. As indicated above, *FLC* is a prime candidate for the shared QTL in both the Vernalization and Non-vernalization treatment. We did not identify any non-synonymous polymorphisms in the *FLC* coding region which is in agreement with other studies examining *FLC* variation (Gazzani *et al.*, 2003; Caicedo *et al.*, 2004; Strange *et al.*, 2011).

In addition to *FLC*, the Non-vernalization treatment yields *FT* as a possible candidate gene. *FT* has been identified as a key component of the photoperiod pathway of floral induction across plant species (Kardailsky *et al.*, 1999; Kobayashi *et al.*, 1999; Böhlenius *et al.*, 2006; Tamaki *et al.*, 2007) and is highly conserved (Kobayashi *et al.*, 1999). We have identified one non-synonmous polymorphism in the *FT* coding region; however, it is unlikely to have an effect as both amino acids are neutral and non-polar (Table 3.2). Recent studies have failed to detect non-synonomous polymorphisms in the coding region of *FT*, but rather find considerable cis-

regulatory variation (Schwartz *et al.*, 2009; Adrian *et al.*, 2010; Strange *et al.*, 2011). In our study, the QTL (NVF1) that encompasses *FT* is of relatively small effect size and only was identified in the Non-vernalization treatment. *FLC* acts to repress *FT* expression until it is inhibited by vernalization or an autonomous mechanism. It is logical that these candidates were detected in the Non-vernalization treatment as allelic variation between the mapping parents would provide differential sensitivity of *FT* to *FLC* in the absence of vernalization. With vernalization the effect of *FT* allelic variation could be negligible once *FLC* is sufficiently inhibited.

In the Vernalization treatment there are several candidate genes identified, including all members of the *FLC/MAF* clade, which Salome *et al.,* (2011) have highlighted as being common across QTL mapping studies. *MAF1*, also known as *FLM*, is a candidate underlying VF1. *MAF1* is a homologue of *FLC* and also acts to inhibit flowering in a similar fashion (Scortecci *et al.*, 2001). We have identified two synonymous substitutions and one non-synonymous substitution in the coding region (Table 3.2). Given that there is functional redundancy between *FLC* and *MAF1*, it would be expected that weakened alleles at both genes would be necessary to allow for earlier flowering.

The QTL, on the end of chromosome V, (VF6), encompasses seven potential candidate genes. *MAF2-4* act to repress flowering after short cold spells (Ratcliffe *et al.*, 2001), and polymorphisms at these loci have been shown to be associated with diverse flowering times in *Arabidopsis* accessions (Caicedo *et al.*, 2009). Other strong candidates at the locus include *VIN3, VIP4*, and *ELF5*. We have sequenced the coding regions of *MAF2* and *VIN3*. In *MAF2* there is one non-synonymous substitution between Italy in Sweden, however, there is no change

in polarity or charge. On the other hand, in the *VIN3* coding region there are 2 non-synonymous substitutions, as well as a 3bp deletion (Table 3.2).

Three genes, *VRN1, VIL1,* and *DNF* underlie the LOD support interval of VF2. *VRN2* and *VIL1* act to promote flowering and are upstream members of the vernalization pathway (Gendall *et al.*, 2001). *DNF* acts in the repression of *CO* during the early part of the day (Morris *et al.*, 2010). For successful vernalization it is important to note that two distinct processes must occur: 1) the initial repression of *FLC* and 2) the maintenance of *FLC* inhibition following vernalization. *VIL1* (along with *VIN3*) is necessary for the chromatin modification of *FLC/MAF1* repression.

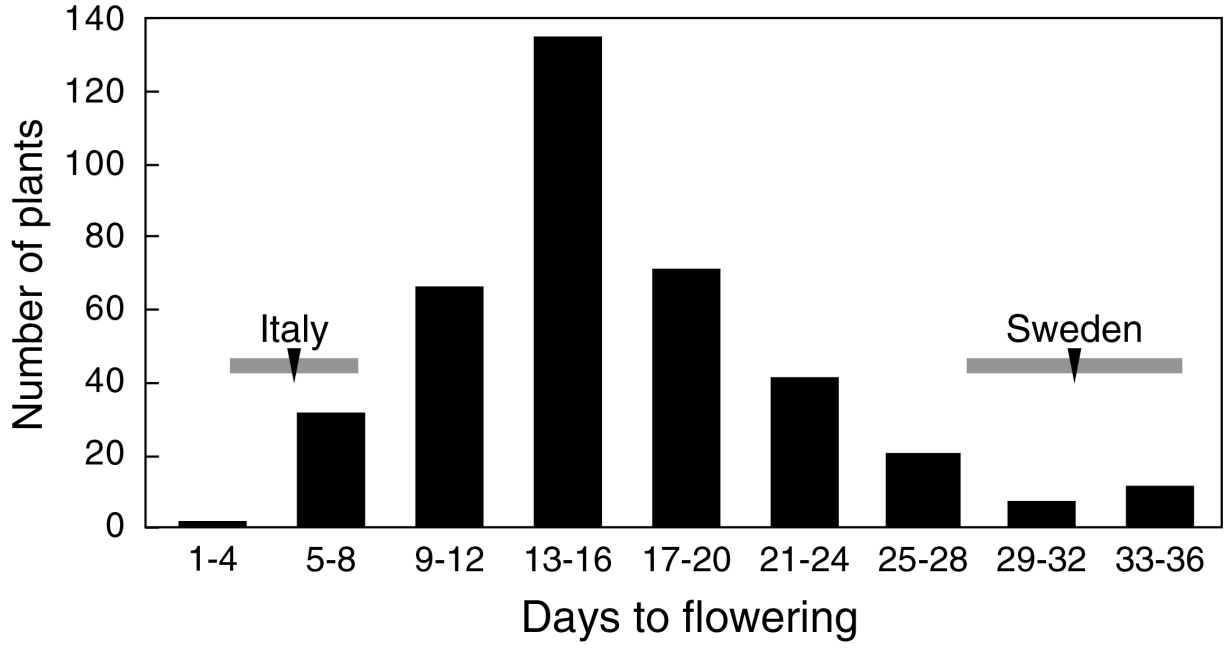
The genetics of flowering time in *Arabidopsis* has been examined extensively, and the molecular genetic pathways have been well characterized by a variety of means, including QTL mapping. However, almost all of these QTL studies have involved the lab strains Col, Ler or both as mapping parents and been performed under controlled laboratory conditions. In the present study we identify two QTL that do not have a clear candidate gene (VF3 and VF5). Two floral developmental genes underlie this QTL, but neither would be considered a strong candidate, or have been described in previous QTL studies. A possible candidate for VF5 is *SEF*, which has not been thoroughly studied to date. While it is rare to reveal novel loci involved in flowering time, it is possible; as is demonstrated in the present work through mapping natural populations, and a recent study that investigated natural accessions crossed with Col in the field (Brachi *et al.,* 2010).

Elucidating the role of cis-regulatory and protein coding variation is a central goal in studying the genetic basis of adaptation. Hoekstra and Coyne (2007) highlight that considerable emphasis has been placed on cis-regulatory evolution, however, sufficient data is lacking to

make generalizations. Here, we examine only the protein coding regions of several known candidate genes underlying QTL. Rigorous functional experiments are required to identify the causal genes and mutations controlling flowering time adaptation. Nonetheless, in several cases a lack of protein coding variation suggests the role of cis-regulatory changes. In *FT* and *FLC* there are no structural polymorphisms between the mapping parents, which is in agreement with studies characterizing regulatory variation of these genes (Gazzani *et al.*, 2003; Michaels *et al.*, 2003; Schwartz *et al.*, 2009; Adrian *et al.*, 2010; Strange *et al.*, 2011). In contrast, *VIN3* could potentially represent protein coding changes responsible for flowering time variation, as has been well described in *FRI* (Johanson *et al.*, 2000; Gazzani *et al.*, 2003).

Figure 3.3. Days to flowering in the Vernalization treatment.

Days to flowering of the *Arabdiopsis thaliana* F2's in the Vernalization treatment. The gray bars represent the flowering duration of the respective parents, with the black triangle representing the mean flowering date.



While many genes control flowering time, few have been shown to be associated with natural flowering time variation. The regulatory gene *FRI* has received the most attention as a major determinant of flowering time in natural populations of *Arabidopsis* (Clarke & Dean, 1994; Johanson *et al.*, 2000; Le Corre *et al.*, 2002; Simpson & Dean, 2002). Functional *FRI* alleles result in the accumulation of *FLC* mRNA which inhibits flowering. Vernalization acts to reduce the sensitivity of *FLC* to *FRI* and thereby promotes flowering. Several lines of evidence support the role of *FRI* in controlling natural flowering time variation including 1) the identification of several naturally occurring, independently derived null *FRI* alleles (Le Corre *et al.*, 2002), 2) a large number of non-synonomous polymorphisms in the first exon of *FRI* (Le Corre *et al.*, 2002), 3) the identification of QTL co-localizing with *FRI* in several QTL studies with different parental populations varying in flowering time (Salome *et al.*, 2011), and 4) evidence for *FRI* alleles influencing fitness depending on the *FLC* allelic background and environmental conditions (Korves *et al.*, 2007). Despite this substantial body of research demonstrating the contribution of *FRI* in flowering time variation, we found no role of *FRI* allelic variation controlling flowering time differentiation in the Italy and Sweden populations.

It has been well established that coding region variation in exon 1 of *FRI* results in non-functional alleles and early flowering (Johanson *et al.*, 2000; Le Corre *et al.*, 2002; Gazzani *et al.*, 2003). We have identified one polymorphism in *FRI* between the Italy and Sweden mapping parents which results in a Glycine to Alanine substitution at amino acid 146 (Table 3.2). Both these amino acids are neutral and nonpolar, and are unlikely to affect protein function. Furthermore, no QTL were found to co-localize with *FRI*.

This may be surprising given that a latitudinal cline in flowering time has been predicted to be associated with *FRI* allele functionality (Karlsson *et al.*, 1993; Nordborg & Bergelson,

1999; Pigliucci & Marlow, 2001). That is, populations in southern latitudes are expected to harbor null *FRI* alleles providing flowering in the absence of vernalization, and northern populations will contain functional *FRI* alleles. However, several studies have been unable to detect such a cline in flowering time (Karlsson *et al.*, 1993; Nordborg & Bergelson, 1999; Pigliucci & Marlow, 2001; Stinchcombe *et al.*, 2004). In the most extensive study, Stinchcombe *et al.* (2004) identify functional *FRI* alleles to be associated with a cline in flowering, but no association was detected for non-functional *FRI* alleles, as is predicted. Furthermore, the populations from southern latitudes with functional *FRI* alleles actually flowered earlier than those from northern latitudes. The variation in flowering time is explained by an epistatic interaction between the functional *FRI* alleles and *FLC* (Caicedo *et al.*, 2004). Our results are in agreement with those of Stinchcombe *et al.* (2004) and Caicedo *et al.* (2004). In the present study, *FRI* is not identified as a candidate gene, but *FLC* is in both the Vernalization and Non-Vernalization treatment. This makes sense when taking life history into consideration. Given that both of the parental populations are winter annuals, *FRI* functional alleles would be favored to delay flowering until the appropriate time in the spring (Simpson & Dean, 2002). Weak alleles of *FLC* in the Italy mapping parent, reduce sensitivity to *FRI* and promote flowering after mild periods of cold compared to Sweden (Amasino, 2004).

In summary, we have presented a QTL mapping study of flowering time variation between two winter annual accessions of *Arabidopsis* collected from the northern and southern geographic limits of their native range. Our study benefits from a large mapping population size and is thus capable of accurately estimating effect sizes. Although we identified both large and small effect QTL, it is almost certain that we have missed many QTL of very small effect (≈1% PVE). We find the genetic architecture of adaptation to depend heavily on the environmental

conditions, namely the vernalization treatment. Previous QTL studies have emphasized the role of *FRI* and *FLC* in determining flowering time differentiation. While substantial variation is explained by a QTL co-localizing with *FLC*, we find no role of *FRI*. Furthermore, it should be noted that genetic basis of flowering time in this system is complex and includes several QTL with numerous other candidate genes. Future research should be conducted to thoroughly examine the causal genetic elements underlying QTL as well as field studies to determine the relationship of flowering time and fitness. An integration of these pursuits will provide a comprehensive understanding of the genetic basis of flowering time in *Arabidopsis*.

Acknowledgments

We would like to thank M. Cameron for assistance with figure development, K. Califf for assistance with sequence alignment, C. Oakley for help with R/qtl, and R. Amasino and D. Weigel for helpful comments on the manuscript as well as two anonymous reviewers.

LITERATURE CITED

LITERATURE CITED

Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, Turck F. 2010. cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of *FLOWERING LOCUS T* in *Arabidopsis*. *Plant Cell* **22**(5): 1425-1440.

Ågren J, Schemske DW. 2012. Reciprocal transplants demonstrate strong adaptive differentiation of the model organism *Arabidopsis thaliana* in its native range. *New Phytologist* **194**(4): 1112-1122.

Amasino R. 2004. Vernalization, competence, and the epigenetic memory of winter. *Plant Cell* **16**(10): 2553-2559.

Amasino R. 2010. Seasonal and developmental timing of flowering. *Plant Journal* **61**(6): 1001-1013.

Balasubramanian S, Schwartz C, Singh A, Warthmann N, Kim MC, Maloof JN, Loudet O, Trainer GT, Dabi T, Borevitz JO, Chory J, Weigel D. 2009. QTL mapping in new *Arabidopsis thaliana* advanced intercross-recombinant inbred lines. *Plos One* **4**(2).

Barton NH, Keightley PD. 2002. Understanding quantitative genetic variation. *Nature Reviews Genetics* **3**(1): 11-21.

Bastow R, Dean C. 2003. Plant sciences - Deciding when to flower. *Science* **302**(5651): 1695-1697.

Beavis WD 1998. QTL analyses: power, precision, and accuracy. In: Paterson AH ed. *Molecular Dissection of Complex Traits*. Boca Raton, FL: CRC Press, 145-162.

Böhlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O. 2006. CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* **312**(5776): 1040-1043.

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F. 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *Plos Genetics* **6**(5).

Broman KW, Wu H, Sen S, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**(7): 889-890.

Burn JE, Bagnall DJ, Metzger JD, Dennis ES, Peacock WJ. 1993. DNA methylation, vernalization, and the initiation of flowering. *Proceedings of the National Academy of Sciences of the United States of America* **90**(1): 287-291.

Caicedo AL, Richards C, Ehrenreich IM, Purugganan MD. 2009. Complex rearrangements lead to novel chimeric gene fusion polymorphisms at the *Arabidopsis thaliana MAF2-5* flowering time gene cluster. *Molecular Biology and Evolution* **26**(3): 699-711.

Caicedo AL, Stinchcombe JR, Olsen KM, Schmitt J, Purugganan MD. 2004. Epistatic interaction between *Arabidopsis FRI* and *FLC* flowering time genes generates a latitudinal cline in a life history trait. *Proceedings of the National Academy of Sciences of the United States of America* **101**(44): 15670-15675.

Clarke JH, Dean C. 1994. Mapping *FRI*, a locus controlling flowering time and vernalization response in *Arabidopsis thaliana*. *Molecular & General Genetics* **242**(1): 81-89.

Clarke JH, Mithen R, Brown JKM, Dean C. 1995. QTL analysis of flowering time in *Arabidopsis thaliana Molecular & General Genetics* **248**(3): 278-286.

El-Lithy ME, Bentsink L, Hanhart CJ, Ruys GJ, Rovito DI, Broekhof JLM, van der Poel HJA, van Eijk MJT, Vreugdenhil D, Koornneef M. 2006. New *Arabidopsis* recombinant inbred line populations genotyped using SNPWave and their use for mapping flowering-time quantitative trait loci. *Genetics* **172**(3): 1867-1876.

Gazzani S, Gendall AR, Lister C, Dean C. 2003. Analysis of the molecular basis of flowering time variation in *Arabidopsis* accessions. *Plant Physiology* **132**(2): 1107-1114.

Gendall AR, Levy YY, Wilson A, Dean C. 2001. The *VERNALIZATION 2* gene mediates the epigenetic regulation of vernalization in *Arabidopsis*. *Cell* **107**(4): 525-535.

Haley CS, Knott SA. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**: 315-324.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: Evo devo and the genetics of adaptation. *Evolution* **61**(5): 995-1016.

Izawa T. 2007. Adaptation of flowering-time by natural and artificial selection in *Arabidopsis* and rice. *Journal of Experimental Botany* **58**(12): 3091-3097.

Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C. 2000. Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. *Science* **290**(5490): 344-347.

Juenger TE, Sen S, Stowe KA, Simms EL. 2005. Epistasis and genotype-environment interaction for quantitative trait loci affecting flowering time in *Arabidopsis thaliana*. *Genetica* **123**(1-2): 87-105.

Kalisz S, Wardle GM. 1994. Life-history variation in *Campanula americana* (Campanulaceae) - population differentiation. *American Journal of Botany* **81**(5): 521-527.

Kardailsky I, Shukla VK, Ahn JH, Dagenais N, Christensen SK, Nguyen JT, Chory J, Harrison MJ, Weigel D. 1999. Activation tagging of the floral inducer *FT*. *Science* **286**(5446): 1962-1965.

Karlsson BH, Sills GR, Nienhuis J. 1993. Effects of photoperiod and vernalization on the number of leaves at flowering in 32 *Arabidopsis thaliana* (Brassicaceae) ecotypes. *American Journal of Botany* **80**(6): 646-648.

Kawecki TJ, Ebert D. 2004. Conceptual issues in local adaptation. *Ecology Letters* **7**(12): 1225-1241.

Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T. 1999. A pair of related genes with antagonistic roles in mediating flowering signals. *Science* **286**(5446): 1960-1962.

Koornneef M, Alonso-Blanco C, Vreugdenhil D. 2004. Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annual Review of Plant Biology* **55**: 141-172.

Korves TM, Schmid KJ, Caicedo AL, Mays C, Stinchcombe JR, Purugganan MD, Schmitt J. 2007. Fitness effects associated with the major flowering time gene *FRIGIDA* in *Arabidopsis thaliana* in the field. *American Naturalist* **169**(5): E141-E157.

Lacey EP. 1988. Latitudinal variation in reproductive timing of a short-lived monocarp, *Daucus carota* (Apiaceae). *Ecology* **69**(1): 220-232.

Le Corre V, Roux F, Reboud X. 2002. DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: Extensive nonsynonymous variation is consistent with local selection for flowering time. *Molecular Biology and Evolution* **19**(8): 1261-1271.

Lee I, Bleecker A, Amasino R. 1993. Analysis of naturally-occurring late flowering in *Arabidopsis thaliana*. *Molecular & General Genetics* **237**(1-2): 171-176.

Lempe J, Balasubramanian S, Sureshkumar S, Singh A, Schmid M, Weigel D. 2005. Diversity of flowering responses in wild *Arabidopsis thaliana* strains. *Plos Genetics* **1**(1): 109-118.

Li Y, Roycewicz P, Smith E, Borevitz JO. 2006. Genetics of local adaptation in the laboratory: flowering time Quantitative Trait Loci under geographic and seasonal conditions in *Arabidopsis*. *Plos One* **1**(2).

Lister C, Dean C. 1993. Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. *Plant Journal* **4**(4): 745-750.

Loudet O, Chaillou S, Camilleri C, Bouchez D, Daniel-Vedele F. 2002. Bay-0 x Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in *Arabidopsis*. *Theoretical and Applied Genetics* **104**(6-7): 1173-1184.

Lowry DB. 2012. Local adaptation in the model plant. *New Phytologist* **194**(4): 888-890.

Lu H, Romero-Severson J, Bernardo R. 2002. Chromosomal regions associated with segregation distortion in maize. *Theoretical and Applied Genetics* **105**(4): 622-628.

March-Diaz R, Garcia-Dominguez M, Florencio FJ, Reyes JC. 2007. SEF, a new protein required for flowering repression in *Arabidopsis*, interacts with *PIE1* and *ARP6*. *Plant Physiology* **143**(2): 893-901.

Michaels SD, Amasino RM. 1999. *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* **11**(5): 949-956.

Michaels SD, He YH, Scortecci KC, Amasino RM. 2003. Attenuation of *FLOWERING LOCUS C* activity as a mechanism for the evolution of summer-annual flowering behavior in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* **100**(17): 10102-10107.

Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature* **441**(7096): 947-952.

Morgan TJ, Mackay TFC. 2006. Quantitative trait loci for thermotolerance phenotypes in *Drosophila melanogaster*. *Heredity* **96**(3): 232-242.

Morris K, Thornber S, Codrai L, Richardson C, Craig A, Sadanandom A, Thomas B, Jackson S. 2010. *DAY NEUTRAL FLOWERING* represses *CONSTANS* to prevent *Arabidopsis* flowering early in short days. *Plant Cell* **22**(4): 1118-1128.

Napp-Zinn K. 1985. *Handbook of Flowering*. Boca Raton, FL: CRC.

Nordborg M, Bergelson J. 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**(4): 470-475.

O'Neill CM, Morgan C, Kirby J, Tschoep H, Deng PX, Brennan M, Rosas U, Fraser F, Hall C, Gill S, Bancroft I. 2008. Six new recombinant inbred populations for the study of quantitative traits in *Arabidopsis thaliana*. *Theoretical and Applied Genetics* **116**(5): 623-634.

Olsson K, Ågren J. 2002. Latitudinal population differentiation in phenology, life history and flower morphology in the perennial herb *Lythrum salicaria*. *Journal of Evolutionary Biology* **15**(6): 983-996.

Orr HA. 1998. The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution* **52**(4): 935-949.

Pigliucci M. 2003. Selection in a model system: Ecological genetics of flowering time in *Arabidopsis thaliana*. *Ecology* **84**(7): 1700-1712.

Pigliucci M, Marlow ET. 2001. Differentiation for flowering time and phenotypic integration in *Arabidopsis thaliana* in response to season length and vernalization. *Oecologia* **127**(4): 501-508.

Ratcliffe OJ, Nadzan GC, Reuber TL, Riechmann JL. 2001. Regulation of flowering in *Arabidopsis* by an *FLC* homologue. *Plant Physiology* **126**(1): 122-132.

Salome PA, Bomblies K, Laitinen RAE, Yant L, Mott R, Weigel D. 2011. Genetic architecture of flowering-time variation in *Arabidopsis thaliana*. *Genetics* **188**(2): 421-U313.

Schwartz C, Balasubramanian S, Warthmann N, Michael TP, Lempe J, Sureshkumar S, Kobayashi Y, Maloof JN, Borevitz JO, Chory J, Weigel D. 2009. Cis-regulatory changes at *FLOWERING LOCUS T* mediate natural variation in flowering responses of *Arabidopsis thaliana*. *Genetics* **183**(2): 723-732.

Scortecci KC, Michaels SD, Amasino RM. 2001. Identification of a MADS-box gene, *FLOWERING LOCUS M*, that represses flowering. *Plant Journal* **26**(2): 229-236.

Seymour DK, Filiault DL, Henry IM, Monson-Miller J, Ravi M, Pang AD, Comai L, Chan SWL, Maloof JN. 2012. Rapid creation of *Arabidopsis* doubled haploid lines for quantitative trait locus mapping. *Proceedings of the National Academy of Sciences of the United States of America* **109**(11): 4227-4232.

Shindo C, Aranzana MJ, Lister C, Baxter C, Nicholls C, Nordborg M, Dean C. 2005. Role of *FRIGIDA* and *FLOWERING LOCUS C* in determining variation in flowering time of *Arabidopsis*. *Plant Physiology* **138**(2): 1163-1173.

Shindo C, Bernasconi G, Hardtke CS. 2007. Natural genetic variation in arabidopsis: Tools, traits and prospects for evolutionary ecology. *Annals of Botany* **99**(6): 1043-1054.

Simpson GG, Dean C. 2002. Flowering - *Arabidopsis*, the rosetta stone of flowering time? *Science* **296**(5566): 285-289.

Stinchcombe JR, Weinig C, Ungerer M, Olsen KM, Mays C, Halldorsdottir SS, Purugganan MD, Schmitt J. 2004. A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. *Proceedings of the National Academy of Sciences of the United States of America* **101**(13): 4712-4717.

Strange A, Li P, Lister C, Anderson J, Warthmann N, Shindo C, Irwin J, Nordborg M, Dean C. 2011. Major-effect alleles at relatively few loci underlie distinct vernalization and flowering variation in *Arabidopsis* accessions. *Plos One* **6**(5).

Stratton DA. 1998. Reaction norm functions and QTL-environment interactions for flowering time in *Arabidopsis thaliana*. *Heredity* **81**: 144-155.

Sung SB, Amasino RM. 2004. Vernalization and epigenetics: how plants remember winter. *Current Opinion in Plant Biology* **7**(1): 4-10.

Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K. 2007. Hd3a protein is a mobile flowering signal in rice. *Science* **316**(5827): 1033-1036.

Tanksley SD. 1993. Mapping polygenes. *Annual Review of Genetics* **27**: 205-233.

Ungerer MC, Halldorsdottir SS, Modliszewski JL, Mackay TFC, Purugganan MD. 2002. Quantitative trait loci for inflorescence development in *Arabidopsis thaliana*. *Genetics* **160**(3): 1133-1151.

Van Ooijen JW, Voorips RE. 2001. *Joinmap 3.0: software for the calculation of genetic*

linkage maps. Wageningen, the Netherlands: Plant Research International.

Weingartner M, Subert C, Sauer N. 2011. *LATE*, a C2H2 zinc-finger protein that acts as floral repressor. *Plant Journal* **68**(4): 681-692.

Weinig C, Stinchcombe JR, Schmitt J. 2003. QTL architecture of resistance and tolerance traits in *Arabidopsis thaliana* in natural environments. *Molecular Ecology* **12**(5): 1153-1163.

Weinig C, Ungerer MC, Dorn LA, Kane NC, Toyonaga Y, Halldorsdottir SS, Mackay TFC, Purugganan MD, Schmitt J. 2002. Novel loci control variation in reproductive timing in *Arabidopsis thaliana* in natural environments. *Genetics* **162**(4): 1875-1884.

Werner JD, Borevitz JO, Uhlenhaut NH, Ecker JR, Chory J, Weigel D. 2005a. *FRIGIDA*-independent variation in flowering time of natural *Arabidopsis thaliana* accessions. *Genetics* **170**(3): 1197-1207.

Werner JD, Borevitz JO, Warthmann N, Trainer GT, Ecker JR, Chory J, Weigel D. 2005b. Quantitative trait locus mapping and DNA array hybridization identify an *FLM* deletion as a cause for natural flowering-time variation. *Proceedings of the National Academy of Sciences of the United States of America* **102**(7): 2460-2465.

Westerbergh A, Doebley J. 2002. Morphological traits defining species differences in wild relatives of maize are controlled by multiple quantitative trait loci. *Evolution* **56**(2): 273-283.

CHAPTER 4

Functional traits underlying mating system differentiation between the wild relatives of rice

Introduction

Angiosperms display a remarkable diversity of reproductive traits and strategies. Understanding the evolutionary mechanisms that have shaped this diversity has been a central goal in studies of plant evolution (Barrett, 2010a). Given the sessile nature of plants, reproduction is of particular significance, as it and subsequent seed dispersal are the only opportunity for alleles to move within and among populations, as well as the physical landscape (Levin & Kerster, 1974).

The mating system shift to predominantly selfing from outcrossing taxa is regarded as the most common transition in plant evolution (Stebbins, 1974). Outcrossing rates can vary considerably among closely related species and even populations, suggesting that it can respond quickly to natural selection (Jain, 1976; Goodwillie *et al.*, 2005). Additionally, the transition to selfing is associated with characteristic shifts in floral biology, life history, and ecology (Barrett *et al.*, 1996). There are two classic explanations for the advantage of self-fertilization: 1) selfing may be beneficial as it can provide a means for reproductive assurance when pollinators or mates are scarce; a hypothesis first proposed by Darwin (1876) and 2) an inherent genetic transmission advantage, demonstrated by Fisher (1941), as selfing individuals function as both the maternal and paternal parents. In contrast, the evolutionary consequences of self-fertilization can be profound, as selfing can drastically reduce the effective population size (Hamrick & Loveless, 1987; Schoen & Brown, 1991) and nucleotide diversity (Wright *et al.*, 2008), influence population genetic structure (Hamrick & Godt, 1996), and restrict broad scale rates of diversification (Charlesworth, 1992; Takebayashi & Morrell, 2001; Igic *et al.*, 2008). Most directly, selfed offspring can suffer from inbreeding depression, the primary genetic cost of self-fertilization (Lloyd, 1979; Charlesworth & Charlesworth, 1987; Husband & Schemske, 1996).

Lande and Schemske (1985) demonstrated the joint evolution of inbreeding depression and selfing, with reduced levels of inbreeding depression as selfing increases. This theoretical model predicts two stable endpoints to mating system evolution: predominant outcrossing and selfing. Subsequent analysis of empirical studies provided support for this bimodal distribution of outcrossing rates (Schemske & Lande, 1985). Taken together these findings prompted a flurry of theoretical and empirical investigation into the distribution of mating systems, with much emphasis on the evolutionary significance of mixed mating (Barrett, 2003). Recent surveys of outcrossing rates have demonstrated a higher occurrence of mixed mating and do not support a bimodal distribution, particularly for animal pollinated (entomophilous) taxa (Vogler & Kalisz, 2001; Goodwillie *et al.*, 2005). However, this finding should be interpreted with caution; for example there is a disproportionate lack of mating system estimates of obligate outcrossing species (self-incompatible or dioecious), as well as other potential sources of bias in available data (Igic & Kohn, 2006). Nonetheless, a bimodal distribution of outcrossing rates is well supported for wind-pollinated (anemophilous) plants (Aide, 1986; Vogler & Kalisz, 2001; Goodwillie *et al.*, 2005). This observation is attributed to the predictable nature of wind as a pollen vector, as opposed to animal vectors that can fluctuate over time (Aide, 1986; Friedman & Barrett, 2009b).

"Our knowledge of the reproductive ecology of most wind-pollinated plants is quite rudimentary, and the functional relationships between pollination and mating are poorly understood."-Friedman and Barrett 2009

For nearly 150 years, the vast body of literature devoted to plant mating systems and pollination biology has focused almost exclusively on animal-pollinated systems (Friedman & Barrett, 2009b; Barrett, 2010b). This is unfortunate given that approximately 10% of all angiosperms are wind-pollinated (Ackerman, 2000), including our most important crops (rice, wheat, maize, etc) and wind-pollination is a derived state, having evolved at least 65 times from animal-pollinated ancestors (Linder, 1998). There are also intrinsic differences in floral biology between wind- and animal-pollinated species that are particularly interesting to study within a mating system context, including: 1) the presence/absence of attractive floral structures that evolve with the outcrossing rate, 2) fundamental differences in floral anatomy that may determine the functional traits that control selfing, and 3) the role of geitonogamy (mating between flowers of the same plant) and mechanisms that provide this mode of selfing.

In animal-pollinated systems, comparative studies of floral biology between selfing and outcrossing groups have been conducted in a wide range of species (Darwin, 1876; Ritland & Ritland, 1989; Lyons & Antonovics, 1991; Goodwillie, 1999; Armbruster *et al.*, 2002; Vallejo-Marin & Barrett, 2009). Through these efforts it has been well established that selfing plants produce small flowers with highly reduced attractive floral traits (e.g. showy petals, nectar production, attractive scents, etc), when compared to outcrossers (Ornduff, 1969; Goodwillie *et al.*, 2010). To explain this syndrome, sex allocation theory has been developed; which predicts reduced allocation to both, male function and attractive structures, as selfing increases (Charlesworth & Charlesworth, 1987; Charlesworth & Morgan, 1991). Correlations between attractive floral characters and functional traits that provide selfing, are common and can influence, or even limit, an evolutionary response to selection (Conner, 2006). In wind-pollinated plants attractive floral traits are absent all together; therefore the transition to a selfing

syndrome may be relatively less complex than in animal-pollinated species. Comparative studies of floral biology between selfing and outcrossing groups are lacking in wind-pollinated plants.

In addition to comparative work, more mechanistic experiments have been aimed at elucidating the mechanisms of self-fertilization in animal-pollinated taxa. In hermaphroditic self-compatible plants, the most common means for achieving autogamous (fertilization within an individual flower) selfing, is through a reduction in herkogamy (the physical separation of stigma and anthers) (Ritland & Ritland, 1989; Dole, 1990; Holtsford & Ellstrand, 1992; Kohn & Barrett, 1994; Belaoussoff & Shore, 1995; Karron *et al.*, 1997; Chang & Rausher, 1999; Kalisz *et al.*, 1999; Barrett, 2003). Conversely, herkogamy in wind-pollinated groups has been unexplored, and its potential to contribute to selfing is unclear due to fundamental differences in floral biology; namely dry, unclumped pollen (Crane, 1986; Ackerman, 2000) and in the case of grasses, poricidal anthers that arrange pollen in a unique fashion and release it through terminal pores (Kirpes *et al.*, 1996). Therefore, physical contact between stigma and anthers may not provide a selfing mechanism as reliable as that seen in animal-pollinated plants. In wind-pollinated groups, inflorescence architecture is the functional trait that has received the most attention in regards to floral biology (Niklas, 1985; Niklas, 1987; Friedman & Harder, 2004; Friedman & Harder, 2005; Cresswell *et al.*, 2010). Grasses, in particular, display a diversity in 'openness' of inflorescences ranging from diffuse to compact (Doust & Kellogg, 2002). Research involving inflorescence architecture in grasses has focused solely on wind aerodynamics and strategies for pollen removal and receipt, and has not been investigated in terms of selfing rates. It is likely that close proximity of flowers in compact inflorescences could enhance the opportunity for geitonogamy.

Self-fertilization can arise from either geitonogamy or autogamy. Large floral displays function in attracting pollinators across long distances, but can result in geitonogamy when pollinators visit subsequent flowers on the same plant. Theoretical work has demonstrated that geitonogamy is unlikely to be favored by selection (Lloyd, 1992) and this mode of selfing can have profound costs through pollen and seed discounting (Lloyd & Schoen, 1992; Dejong *et al.*, 1993; Harder & Barrett, 1995; Barrett, 2003; Goodwillie *et al.*, 2005). Thus, geitonogamous selfing does not provide reproductive assurance in the same manner as autogamy; although there examples where this is the case, such as monecious *Begonia* species (Ågren & Schemske, 1993). Geitonogamy can be quite common in animal-pollinated systems (Harder & Barrett, 1995; Eckert, 2000; Dorken *et al.*, 2002; Karron *et al.*, 2004) and is suspected to promote the evolution of dichogomy or dioecy, particularly in large plants (Barrett, 2003). In a similar fashion, wind could provide geitonogamous selfing between flowers in close proximity, however, this has not been investigated in wind-pollinated taxa until recently. Friedman and Barrett (2009a) have identified substantial geitonogamous selfing in *Carex* spp., and propose that, unlike animal-pollinated species, geitonogamy in wind-pollinated systems may often provide reproductive assurance, particularly in plants with unisexual flowers. More experiments in wind-pollinated plants are needed that dissect the contributions of autogamy and geitonogamy, in order to understand their role in selfing.

Here, I present a comparative study of floral biology between two wind-pollinated sister species, *Oryza nivara* and *O. rufipogon*, that have recently diverged in their mating system (Morishima & Barbier, 1990; Vaughan, 1994). The goals of this study are to characterize differences in floral biology that may contribute to the mating system, and to partition the role of geitonogamy and autogamy in providing selfing. Additionally, I provide a mechanistic

framework for mating system divergence in these species. To my knowledge, this is the first study of comparative floral biology between a wind-pollinated selfing and outcrossing group. This line of research enhances my previous work with these species and provides a foundation for a comprehensive understanding of mating system evolution in this system.

The sister species, *Oryza rufipogon* and *O. nivara* , are wetland grasses distributed throughout the tropics and subtropics of Asia. *Oryza nivara*, is an annual that inhabits ephemeral pools containing standing water throughout the monsoon season, but then desiccates during the dry season (Vaughan, 1994). Conversely, *O. rufipogon*, is a perennial that inhabits lowland marshes and wetlands that retain sufficient water levels for the plants to persist throughout the year. Both species are self-compatible and hermaphroditic, but differ in their mating system. Barbier (1989) measured outcrossing rates for one *O. nivara* and two *O. rufipogon* populations in Thailand. Multilocus outcrossing rates were estimated with 8 isozyme loci based on the mixed-mating model (Ritland & Jain, 1981). *O. rufipogon* had substantially higher outcrossing rates (t_m=0.50 and 0.56) than *O. nivara* (t_m=0.04), which is consistent with earlier estimates based on other approaches (Oka & Morishima, 1967). Phylogenetic analysis has indicated that *O. nivara* has recently diverged from a common ancestor resembling *O. rufipogon* (Barbier *et al.*, 1991; Zhu & Ge, 2005), which is consistent with broader observations of annual selfers evolving from perennial outcrossers (Barrett *et al.*, 1996). Previously, I have mapped the genetic basis of several adaptive traits that differ between these species, including three mating system traits: anther length, inflorescence architecture, and inflorescence exsertion (Grillo *et al.*, 2009). In the present study, I expand upon this work by examining a broad range of floral morphological and physiological traits that could potentially influence the mating system. A list of traits and their putative function are outlined in Table 1. These traits are compiled from preliminary

observations of floral biology in the focal species and a review of relevant literature on the subject.

Materials and Methods

An exhaustive survey of potential traits that may influence mating system divergence between *O. rufipogon* and *O. nivara* was conducted. Unless otherwise mentioned, these data were obtained from one accession of each species collected in India and described in Grillo et al (2009). Based on previous greenhouse observations of both species, these two individuals were selected as being representative of typical *O. rufipogon* and *O. nivara* populations. Statistical differences between trait means were determined through one-way ANOVA, implemented in R. Multivariate analysis of variance (MANOVA), implemented in JMP 10, was performed to identify overall differences in floral phenotypes. Plants were grown and studied over a three year period at Michigan State University in greenhouses and growth chambers. All the mating system traits require plants to be in flower, and this was achieved by exposing plants to short day length in the growth chamber.

Anther length, anther basal pore size, pollen size, stigma length, and stigma papillae number were measured from freshly collected samples with a compound light microscope. Length measurements were obtained with a lens ocular micrometer and were converted to the appropriate units (µm or mm). Stigma color was measured qualitatively as purple or white. Anther position was scored as '1' if anthers fell below the flower in a pendent-like position, or as '0' if they remained upright above the stigma. Inflorescence openness was measured by the angles between primary branches and rachis. Inflorescence exsertion was measured by the distance from the flag leaf node to the first primary branch node. For all traits the number of

measurements is given in Table 2. A subset of the traits (anther length, inflorescence exsertion, and inflorescence openness) was measured from multiple populations of each species, including five populations of *O. nivara* and two populations of *O. rufipogon* from Thailand, and one population of each species from India (Table 3).

The timing of anther dehiscence was examined before anthesis. Under the constant conditions of the growth chamber, both species began flowering 4 hours after first light. All flowers of an inflorescence typically complete flowering within a 4 day period, thus each day a subset of flowers open (Grillo per. obs.). On the final day of this period, approximately 10 flowers remain and there is a very high probability that they will open on that day. Thus, it is reasonably possible to predict when a flower will open on the last day of flowering. When this was the case I waited until the first flower opened for that day and then promptly opened the remaining flowers. Anthers and stigmas of these flowers were collected and examined under a microscope. Early dehiscence was scored as 0, 1, 2, 3,or 4 based on the relative number of pollen grains on the stigma prior to flower opening (0 = 0, 1 = <20, 2 = 20-40, 3 = 40-80, 4 = >80 pollen grains, respectively).

Pollen viability was measured by pollen staining with a fluorescent dye (Rodriguez-Riano & Dafni, 2000). Pollen was collected at the following time intervals (minutes following anthesis), 0, 10, 20, as well as prior to anthesis as described for 'early dehiscence'. To store pollen, anthers were harvested with forceps and then shaken over a microcentrifuge tube containing 1mL of 10% sucrose. Storage of pollen in sucrose solution prevents against further degradation and maintains pollen grains at the collection state (Rodriguez-Riano & Dafni, 2000). A Fluorescein Diacetate (FDA) solution was created with 60mg of FDA in 30mL of deionized water. Droplets of FDA solution were added to the pollen collection tubes (suspended in 1mL

10% sucrose) until the solution turned milky gray (Rodriguez-Riano & Dafni, 2000). The stained pollen were then allowed to incubate for 5 minutes, before being viewed on a fluorescent microscope (purple-indigo beam). Pollen grains were scored as being viable if they absorbed the stain. For each measurement approximately 100 pollen grains were scored. Measurements at each time point were repeated 5 times by collecting pollen from fresh flowers. Pollen staining was attempted with Potassium Iodide, as has been conducted in *O.* sativa (Song *et al.*, 2004); however, this effort was unsuccessful due to consistent over-staining. Additionally, pollen germination was attempted with a standard pollen germination media (Khatun & Flowers, 1995), yet this was unsuccessful as well. Pollen staining indicates the presence of a cytoplasm and is merely a proxy for pollen viability (Rodriguez-Riano & Dafni, 2000). To validate the pollen staining results a pollination study was conducted. For this, approximately 30 minute-old pollen was deposited on fresh stigmas (at anthesis) of *O. rufipogon*. The donor *O. rufipogon* flowers were emasculated at anthesis prior to anther dehiscence, to prevent self-pollination. Accomplishing this in *O. nivara* is highly unlikely and so this experiment was attempted only with *O. rufipogon* as the maternal plant. This experiment was conducted 10 times for each species.

Stigma receptivity was measured by the peroxide test (Kearns & Inouye, 1993). Stigmas were collected and placed in a droplet of 15% hydrogen peroxide on a microscope slide. The slides where then viewed under a dissecting microscope and the number of bubbles emitted over a 30 second period was recorded. Hydrogen peroxide reacts with enzymes on the surface the stigma, generating bubbles of oxygen gas. Stigma receptivity is indicated by the number bubbles emitted (i.e. high receptivity yields lots of bubbles). Stigmas were collected at the following time intervals (minutes following anthesis), 0, 15, 30, 45, 60, as well as prior to anthesis as

described above. To verify actual stigma receptivity, stigmas of *O. rufipogon* were hand pollinated with fresh pollen 45 minutes post anthesis. The flowers were emasculated prior to anther dehiscence and the experiment was carried out 10 times. *O. nivara* was not included in this experiment for reasons indicated above (see pollen viability). As a control fresh *O. rufipogon* stigmas were pollinated with fresh pollen at the time of anthesis.

In order to distinguish between geitonogamy and autogamy, manipulations of floral number were conducted (Schoen & Lloyd, 1992). To determine the rate of autogamy inflorescences were trimmed so that only one flower remained. These plants were then isolated in separate greenhouse rooms and allowed to flower. Plants were marked on the day of anthesis, and approximately one week later, flowers were collected to determine if there was successful self-fertilization. To determine the rate of geitonogamy, plants were cut back so that only a single inflorescence was allowed to flower, and again plants were isolated from each other in separate greenhouse rooms. After all flowers had opened, the inflorescence was bagged, and seed set was scored once they had matured. Under these conditions, there is the opportunity for both autogamy and geitonogamy. Additionally, plants with several inflorescences were allowed to flower in a greenhouse room with dozens of simultaneously flowering plants and abundant wind generated by fans.

Table 4.1. Potential mating system traits and their putative function.

These traits were compiled from examination of literature and detailed observations of floral biology in *O. rufipogon* and *O. nivara*.

Trait	Putative Function
Inflorescence openness	Open inflorescences disperse and capture pollen in the wind.
Inflorescence exsertion	Exserted inflorescences from the foliage promote pollen dispersal/capture
Anther length	Long anthers contain more pollen to promote outcrossing
Filament length	Long filaments may allow anthers to fall below the flower
Anther position	Position in relation to the stigma controls opportunity for selfing
Early dehiscence	Anthers release pollen prior to anthesis promoting autogamous selfing
Basal pore length	Pollen deposited out the basal pore fall onto the stigma causing selfing
Apical pore length	Pollen dispensed by apical pore may promote outcrossing
Pollen size	Can influence pollen viability and aerodynamic properties in the wind
Pollen viability	Prolonged pollen viability can promote outcrossing over the large distance
Stigma length	Large stigmas for capturing pollen in the wind
Stigma papillae density	Increases surface area for pollen capture to promote outcrossing
Stigma color	Anthocyanins may protect against desiccation and promote outcrossing
Stigma receptivity	Duration of receptivity restricts the ability for outcrossing
Dichogamy	Temporal separation of male and female functions

Results

Trait means are provided in Table 2 for all of the morphological and physiological traits under investigation. Additionally, Table 3 gives the values for anther length, inflorescence exsertion, and inflorescence openness for additional populations of each species. Overall, there was a significant difference between the multivariate floral phenotypes of *O. nivara* and *O. rufipogon* (MANOVA: $F_{8,45} = 7.5$; $P<0.0001$).

Unlike *O. rufipogon*, early dehiscence occurred prior to floret opening in *O. nivara,* resulting in pollen deposited on the stigma (Fig. 1). Early dehiscence was scored based on the relative number of pollen grains deposited on the stigma prior to anthesis. *O. nivara* had an early dehiscence value of 2.33, compared to 0.1 in *O. rufipogon*. While this difference is significant, it is likely to be an underestimate for *O. nivara.* A measurement of '0' could either indicate that the anthers did not dehisce and deposit pollen on the stigma, or that the particular flower was sampled too early to witness early dehiscence. Alternatively, the proportion of flowers with any pollen deposited on the stigma was examined. In this case, 55% of *O. nivara* flowers (n=20) exhibited early dehiscence, compared to 6.4% (n=31) in *O. rufipogon*. In *O. rufipogon,* early dehiscence was observed in only two flowers, and in both cases few pollen grains were deposited. These instances could have been caused by contamination from pollen in the air, or physical damage of the stigma during manual flower opening. Nevertheless, even if this is not the case, early dehiscence occurs rarely in *O. rufipogon*.

In a time course study, pollen viability did not differ between species until after 20 minutes, at which time *O. rufipogon* had an average of 62% pollen viability, compared to only 21% in *O. nivara* (Fig. 2). This observation is consistent with results presented by Song et al (2004), in which *O. rufipogon* was approximately 38% viable after 20 minutes, compared to 0% in *O.* sativa. The ability for pollen grains to absorb a stain is merely a proxy for the actual pollen viability, and is likely to be an overestimate (Song et al 2004). A functional experiment was conducted by placing 30-minute old pollen on fresh stigmas. In this experiment there was no significant difference between species (60% and 70% in *O. nivara* and *O. rufipogon*, respectively). The large number of pollen grains deposited on the stigma may have allowed for seed production, despite the presumably low pollen viability of *O. nivara* .

Oryza rufipogon and *O. nivara* had similar rates of decrease in stigma receptivity following anthesis (Fig. 3). As indicated by the assay, stigmas have low levels of receptivity after 45 minutes. To test the functional significance of the stigma receptivity values, a hand-pollination study was conducted with stigmas 45 minutes post anthesis. Only one *O. rufipogon* stigma that was hand pollinated resulted in the production of a seed (n=10), validating the stigma receptivity assay. In contrast, all of the controls resulted in seed production when stigmas were hand-pollinated at the time of anthesis (n=10). Similar experiments were not conducted in *O. nivara* due to the high occurrence of prior selfing (see discussion) and the fact that both species generate identical rates of decline in stigma receptivity. This rapid decrease in receptivity is not unexpected as flowers are open for just 1 hour. In contrast, stigma receptivity did differ between species prior to anthesis, with *O. nivara* being considerably more receptive (Fig. 3).

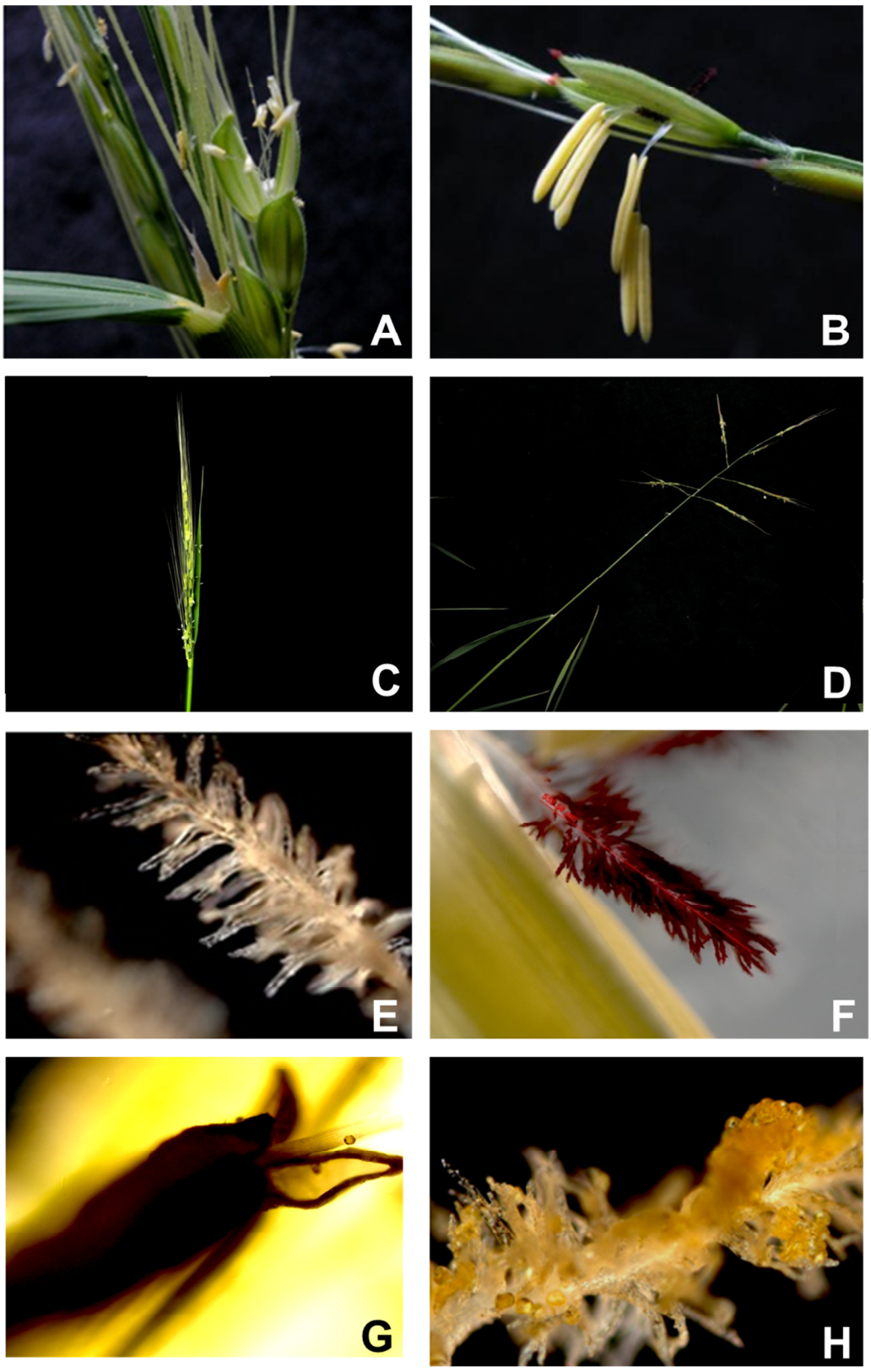
By manipulating flower number, I can partition the selfing rate into contributions from geitonogamy and autogamy (Fig. 4). If autogamous selfing is predominant, one would predict a consistent level of seed set regardless of flower number. Alternatively, if geitonogamy is predominant, seed set should increase with increases in flower number/pollen availability (Schoen & Lloyd, 1992). In the treatments with only a single flower, seed set averaged 72.5% and 10.3% in *O. nivara* and *O. rufipogon*, respectively. Since there is only one flower, all resulting seed set is generated through autonomous selfing. In the treatments with a single inflorescence, there is the opportunity for geitonogamy and autogamy. Under these conditions there is a significant increase in the average seed set for both species, however, seed set of *O. rufipogon* (37.7%) did not approach levels comparable to *O. nivara* (85.4%). These results clearly demonstrate that autogamy is the predominant form of selfing in *O. nivara.* While geitonogamy did provide additional seed set, this may be an artifact of damage induced by the

floral manipulations. In the single-flower experiments the inflorescence was cut back, which could have had detrimental effects on the remaining flower. Conversely, there were no such manipulations in the single-inflorescence experiment. This does not impact interpretation of the role for autogamy between species, as both experienced the same manipulations in the single-flower experiments, but yielded drastically different results. In the case of *O. rufipogon*, geitonogamy increased seed set in the single-inflorescence experiment, although the effects were limited - likely due to pollen limitation. This is demonstrated in the whole-plant experiments, where there was vastly more available pollen. It should be noted that there is opportunity for autogamy, geitonogamy, and outcrossing in this treatment, however; this is inconsequential as pollen limitation is the key factor being addressed. Under these conditions, seed set in *O. rufipogon* did not differ significantly from *O. nivara.*

Figure 4.1. Functional traits of *Oryza* species.

Functional traits controlling mating system differentiation in *O. nivara* and *O. rufipogon*. *O. nivara* traits are in the left column and *O. rufipogon* on the right. A) *O. nivara* floret demonstrating anther size. B) *O. rufipogon* floret demonstrating anther size. C) *O. nivara* panicle shape and exertion. D) *O. rufipogon* panicle shape and exertion. E) *O. nivara* stigma length, papillae density and stigma color. F) *O. rufipogon* stigma length, papillae density and stigma color. G. Basal pore of *O. nivara*. H. Early dehiscence of *O. nivara*, stigma coated in pollen prior to floret opening.

Figure 4.1 (cont'd).



Discussion

The development of polymorphic markers and statistical methodologies has allowed for the reliable estimation of outcrossing rates (t) in plants (Ritland & Jain, 1981; Ritland, 2002). It is widely accepted that outcrossing rates below 20% ($t \leq 0.2$) and above 80% ($t \geq 0.8$), represent predominantly selfing and outcrossing mating systems, respectively (Goodwillie *et al.*, 2005). The estimate of multilocus outcrossing rate in *O. nivara* (tm=0.04) confirms its status as being predominantly selfing. As an annual with small population sizes (Morishima & Barbier, 1990; Vaughan, 1994; Zhou *et al.*, 2008) selection for reproductive assurance likely drove the evolution of selfing in *Oryza nivara*. Despite being substantially more outcrossing than *O. nivara;* mating system estimates from *O. rufipogon* (tm=0.50 and 0.56) would be classified as mixed-mating (Barbier, 1989). It is possible that these estimates may be less than the actual outcrossing rate, due to the occurrence of geitonogamy and/or biparental inbreeding. Geitonogamy is thought to be responsible for a significant amount of the variation in outcrossing rates in plants (Barrett, 2003). This may be the case in *O. rufipogon*, as I have demonstrated that autogamy is rare in this species and large clonal ramets are common in nature (Xie *et al.*, 2001), providing ample opportunity for geitonogamy. Similarly, outcrossing events between close relatives (biparental inbreeding) can result in 'apparent selfing', as is observed in self-incompatible species with outcrossing estimates less than 1 (Ritland, 1984; Goodwillie *et al.*, 2005). Populations of *O. rufipogon* can be quite large (Oka & Morishima, 1967; Vaughan, 1994; Zheng & Ge, 2010), however increasingly their habitats are being fragmented and disturbed. For example, Akimoto *et al.* (1999) have monitored one such population in Thailand that became severely fragmented. During a ten year period the observed outcrossing rate decreased from 0.538 to 0.236, due in part to an increase in biparental inbreeding. Outcrossing rates should be

measured in more populations of each species to better estimate the mating system. Regardless, I am confident that *O. rufipogon* possess attributes consistent with an outcrossing strategy, despite the observed estimates that would be considered mixed-mating.

Table 4.2. Trait means between *O. rufipogon* and *O. nivara*.

Trait means between *O. nivara* and *O. rufipogon*, standard deviation given in parenthesis, n=the number of measurements. Pollen viability is measured at 20 minutes post anthesis. Stigma receptivity is at an 'early dehiscence' stage. Significance tested with one-way anova, ∗, P<0.05; ∗∗, P<0.01; ∗∗∗, P<0.001.

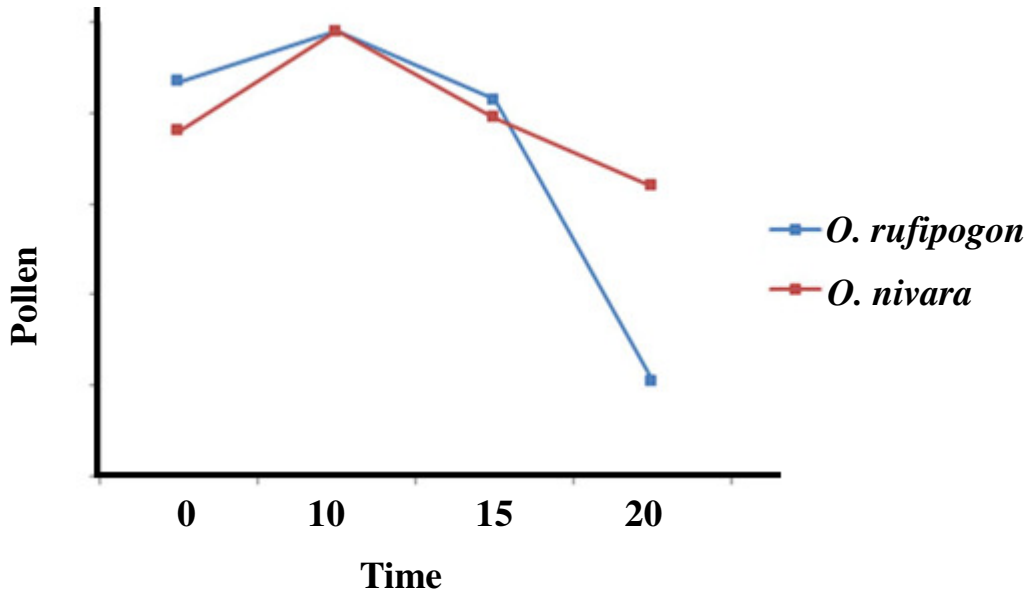
Trait	***O. nivara***	***O. rufipogon***	**Significance**
Anther length (mm)	2.19 (0.22) n=30	4.39 (0.46) n=30	***
Anther position (%)	40 (n=20)	71.4 (n=20)	**
Basal pore length (mm)	0.35 (0.11) n=32	0.14 (0.07) n=30	***
Early dehiscence	2.33 (2.0) n=20	0.1 (0.3) n=31	**
Pollen size (mm)	0.047 (0.002) n=54	0.031 (0.003) n=50	***
Pollen viability (%)	19.2 (2.5) n=5	42.2 (2.4) n=5	*
Stigma length (mm)	1.2 (0.19) n=30	3.02 (0.3) n=30	***
Stigma papillae density	29.8 (3.43) n=20	38.3 (3.15) n=20	**
Stigma pigmentation	white	purple	***
Stigma receptivity	3.25 (0.5) n=5	0.428 (0.46) n=5	***
Inflorescence openness (°)	5.1 (4.1) n=20	45.6 (10.2) n=20	***
Inflorescence exsertion (cm)	1.2 (0.8) n=20	13.9 (6.3) n=20	***

In wind-pollinated plants, male fitness is thought to be directly related to investment in pollen production (Charlesworth & Charlesworth, 1981), and because of this male-biased sex allocation is commonly observed (Burd & Allen, 1988). As an illustration, anthers of *O.*

rufipogon are approximately 60% of total flower length. Anthers of *O. nivara* are half the length of *O. rufipogon* (Table 2) and likely house half the amount of pollen, although there are slight differences in pollen size (see below). This decrease in anther length fits predictions of sex allocation theory, with decreased investment in male function as selfing increases (Charlesworth & Morgan, 1991).

Figure 4.2. Pollen viability time course.

Pollen viability time course for *O. nivara* and *O. rufipogon*. Percent pollen viability measured using FDA florescent staining. Pollen was collected at anther dehiscence and was fixed in 10% sucrose at 10 minute intervals. For the 'Pre-anthesis' timepoint, anthers from unopened florets were collected at peak flowering.

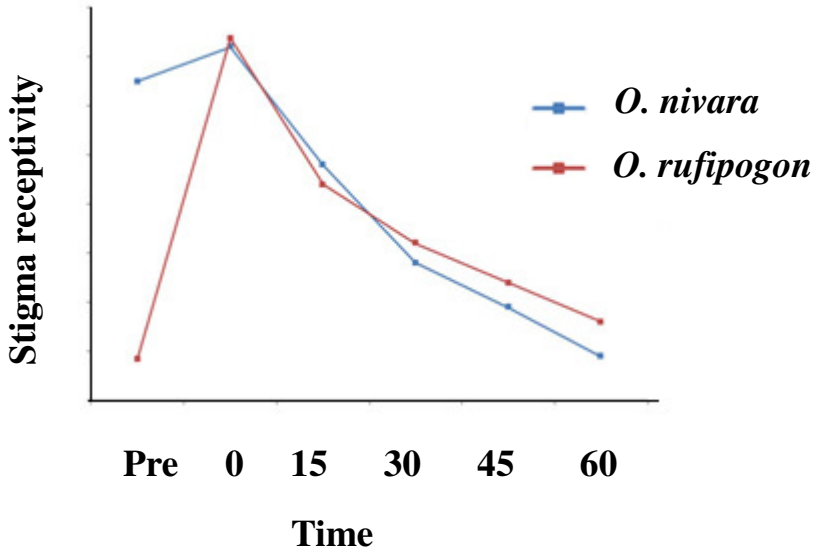


Oryza rufipogon and *O. nivara* possess poricidal anthers that open at pores on each end. There is no difference in apical pore size between these species (Table 2). However, there is a considerable difference in basal pore size, with *O. nivara* having a basal pore twice that of *O. rufipogon*. In a study examining different breeds of cultivated rice, large basal pore size resulted in increased self-pollen deposited on the stigma (Matsui & Kagata, 2003). In *O. rufipogon*, anthers hang in a pendent-like fashion, below the flower (Fig. 5 and Table 2). In this case, pollen is predominantly dispensed through the apical pore due to gravity. Alternatively, anthers of *O. nivara* remain upright and pollen falls through the basal pore. The basal pore of *O. rufipogon* is quite small and often does not open, as it likely serves no function. In these species, the stamen filaments are the same length and the position of anthers upon anthesis (pendent or upright) is a likely result of pollen mass.

Figure 4.3. Stigma receptivity time course.

Stigma viability time course for *O. nivara* and *O. rufipogon*. Stigmas were collected at 15 minute time intervals after floret opening. Stigma receptivity was measured through a peroxidase detection assay by counting the number of bubbles emitted from the stigma surface during a 30 second time interval. Stigma viability scored as: 0=0 1= <50, 2=50-100, 3=100-200, 4=>200 bubbles emitted.

Figure 4.3 (cont'd).



Furthermore, in *O. nivara,* the basal pore opens and pollen is released before the flower opens, contributing to prior selfing (Fig. 5). Prior selfing is autonomous selfing that occurs before the opportunity for the receipt of outcross pollen (Lloyd & Schoen, 1992). In animal-pollinated systems, this happens due to early spatial and developmental overlap between male and female functions and can occur prior to anthesis (Fishman & Wyatt, 1999; Fishman *et al.*, 2002). Matsui and Kagata (2003) observed early dehiscence in cultivated rice. Similarly, *O. nivara* flowers that are artificially opened reveal considerable pollen deposited on the stigma (Fig. 1). This phenomenon is not observed in *O. rufipogon*, and is likely a necessary step for autogamous selfing in *O. nivara.*

Unlike animal-pollinated plants, pollen in wind-pollinated taxa lack ornamentation and exhibit a narrower range of pollen sizes (Harder, 1998; Ackerman, 2000). Pollen size evolution is dictated by the aerodynamic properties of wind, in order to promote successful pollen dispersal and receipt (Friedman & Barrett, 2009b). For example small pollen grains have low settling velocity, allowing them to travel farther, and low inertia to facilitate removal from anthers

(Niklas, 1985). *Oryza rufipogon* has smaller pollen grains which may be beneficial for outcrossing over long distances. The larger pollen grains of *O. nivara* may be the result of relaxed selection on pollen size. Although, it is possible that larger pollen grains would settle out of the airstream more quickly; better ensuring deposition on the stigma within the same flower.

Typically, anthesis in wind pollinated taxa is highly synchronous and abbreviated when compared to animal pollinated systems (Dowding, 1987). Accordingly, pollen of wind-pollinated species is only viable for a short duration (Dafni & Firmage, 2000). Pollen viability in *O. nivara* decreased rapidly, and after 20 minutes (post anther dehiscence) was only 1/3 as viable as *O. rufipogon* (Fig. 2). The loss of viability in *O. nivara* closely resembled that observed in *O.* sativa by Song *et al.* (2004). Both, *O. nivara* and *O.* sativa are highly selfing, and are unlikely to experience selection for prolonged pollen viability. Conversely, *O. rufipogon* pollen retains high viability after 20 minutes, allowing for successful fertilization after pollen dispersal (Fig. 2). Song (2001) identified similar levels of pollen viability for *O. rufipogon*.

Grasses possess unique feathery stigmas that are exserted from the flower (Friedman & Harder, 2004). Wind tunnel experiments have demonstrated that feathery stigmas have higher pollen collection efficiencies than solid stigmas; as they maintain a high surface area, but produce a thinner boundary layer (Niklas, 1985). Stigmas in the focal species differ in their length, papillae number, pigmentation, and receptivity (Table 2). *Oryza rufipogon* has larger stigmas with a higher density of papillae, than *O. nivara,* which may aid in capturing outcross pollen. Additionally stigmas of *Oryza rufipogon* are dark purple in coloration, due the presence of anthocyanins (Fig. 1). Anthocyanins can function to protect plant tissue from environmental stress, including drought and oxidative damage from UV exposure (Stapleton & Walbot, 1994; Hughes *et al.*, 2010). In order to ensure successful pollen germination, stigmas must maintain a

specialized enzymatic surface. Wind-pollinated stigmas are likely prone to desiccation and UV radiation by being exserted outside the flower. Stigma pigmentation in *O. rufipogon* may be an adaptation to promote successful pollen germination and outcrossing. Stigmas of *O. nivara* lack pigmentation and are white. As a selfer, *O. nivara* would not experience selection to maintain pigmentation and allocating resources to anthocyanins could incur a cost. Variation in stigma color is quite common across grass species (Grillo per. obs.) and I hypothesize is associated with the mating system. Stigma receptivity also differed between the focal species, prior to anthesis. *O. nivara* has considerably higher stigma receptivity at this stage which is expected if fertilization occurs due to prior selfing. Overall, there is a diversity of stigma traits in grasses that have not been examined in regards to pollination and mating system (Friedman & Harder, 2004). *O. rufipogon* and *O. nivara* encompass much of this diversity and serve as a potential model for understanding the functional relevance of stigma characteristics.

Table 4.3. Trait means for additional populations of each species

Species	Location	Anther length (mm)	Inflorescence openness (°)	Inflorescence exsertion (cm)
O. nivara	Thailand (Changmai)	2.60 (0.20) n=8	21.3 (3.1) n=10	6.1 (1.3) n=10
	Thailand (Changmai)	2.31 (0.28) n=15	12.4 (3.8) n=10	5.8 (1.8) n=10
	Thailand (Kahnken)	2.12 (0.19) n=20	15.4 (6.1) n=9	7.4 (2.3) n=9
	Thailand (Rayong)	2.43 (0.32) n=6	27.2 (4.3) n=10	5.5 (1.1) n=10
	Thailand (Rayong)	2.52 (0.31) n=11	20.7 (6.8) n=11	6.3 (2.1) n=11
	India (Cuttack)	2.24 (0.23) n=6	8.2 (2.1) n=4	3.7 (0.9) n=4
O. rufipogon	Thailand (Changmai)	4.93 (0.51) n=10	71.4 (12.3) n=4	16.3 (1.1) n=4
	Thailand (Changmai)	4.51 (0.43) n=10	62.3 (10.7) n=4	17.8 (1.7) n=4
	India (Cuttack)	4.42 (0.44) n=6	50.6 (4.8) n=3	16.5 (1.3) n=3

Grasses display a range of inflorescence architectures, ranging from diffuse to compact, representing different means for pollen removal and receipt (Friedman & Harder, 2005; Friedman & Barrett, 2009b). Wind tunnel experiments have demonstrated that compact panicles act to obstruct airflow, thereby creating eddies on the leeward side of the panicle. This allows pollen to settle out of the airstream and become deposited on stigmas. Alternatively, diffuse panicles disturb airflow less, and pollen simply comes in direct contact by windward stigmas (Niklas, 1987). By disturbing airflow less, diffuse panicles may be for more effective at pollen dispersal (Friedman & Harder, 2005). Surveys of *O. rufipogon* populations revealed consistent openness of panicles that may promote pollen dispersal in such a way (Table 3). On the whole *O. nivara* panicles are quite compact, although there is variation for this trait (Table 3). Certainly, the close proximity of simultaneously open flowers within a compact panicle presents an opportunity for geitonogamy.

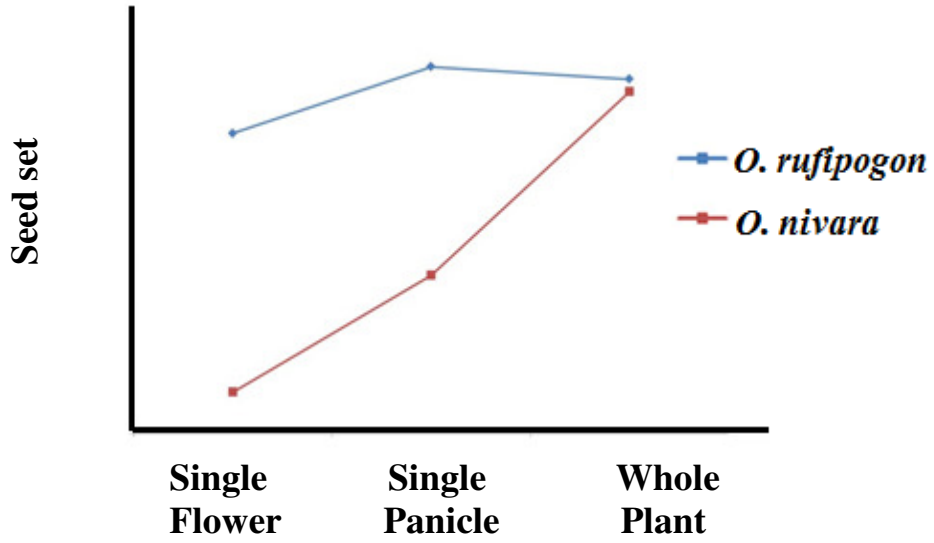
In addition to inflorescence architecture, plant height can influence pollination efficiency. In general, horizontal wind speed increases with height above ground (Burd & Allen, 1988). Therefore, taller plants experience higher wind speeds than shorter ones in the same habitat. For this reason, grasses generally protrude their inflorescence above the boundary layer of their foliage (i.e. panicle exsertion) to promote pollen dispersal (Friedman & Harder, 2004; Friedman & Harder, 2005). Tiller length is greater in *O. rufipogon* (Grillo *et al.*, 2009), although this may not affect plant height as *O. rufipogon* inhabits marshes and tillers essentially float. Panicle exsertion is likely more important in this regard for the focal species. Panicles of *O. nivara* have limited exsertion (Table 2, Fig. 1) and are essentially nested within foliage, which is unlikely to provide effective pollen dispersal. The reverse is true for *O. rufipogon*, consistent with an outcrossing mating system.

Self-compatible plants with many flowers in anthesis simultaneously present an opportunity for geitonogamous selfing (Dejong *et al.*, 1993). This is surely the case in grasses that display highly synchronous bouts of flowering and often have large plant sizes (Friedman & Barrett, 2009b). For this reason, it is suspected that avoidance of geitonogamy has driven the common occurrence of dioecy and self-incompatibility in grasses (Friedman & Harder, 2005; Friedman & Barrett, 2009b). Nonetheless, there are certainly numerous self-compatible hermaphroditic grasses that are prone to this mode of selfing. In order to distinguish between geitonogamy and autogamy, manipulative experiments of floral number are needed (Schoen and Lloyd 1992). Such experiments are rarely carried out, but the few examples reveal high levels of geitonogamous selfing (Harder & Barrett, 1995; Eckert, 2000; Dorken *et al.*, 2002; Karron *et al.*, 2004). I have demonstrated that in *O. nivara,* high levels of selfing are achieved through autogamy. In contrast self-fertilization in *O. rufipogon* is due to geitonogamy and increases with flower number (Fig. 4). As described previously, geitonogamy in *O. rufipogon* could influence the observed outcrossing rate.

The transition from outcrossing to selfing is regarded as the most common trend in plant evolution (Stebbins, 1974; Barrett, 2003). This is exemplified by the recent divergence of *O. nivara* (selfing) from an *O. rufipogon*-like (outcrossing) ancestor (Barbier *et al.*, 1991; Zhu & Ge, 2005). Therefore, it is important to identify the functional traits that underlie this shift, and to consider the means by which they have evolved. In *O. nivara* a suite of traits act in concert to provide autogamous self-fertilization: the timing of anther dehiscence, anther basal pore size, and stigma receptivity. Before the flower even opens the anthers of *O. nivara* become dehiscent with the opening of the basal pore (Fig. 5). Pollen is released through the basal pore, which is quite large in this species. In turn, the stigma is receptive at this time as well. Upon anthesis the

anthers in *O. nivara* remain upright and continue to deposit pollen onto the stigma of the same flower. Thus, autogamy is achieved in *O. nivara* through prior selfing.

Figure 4.4. Autogamy and geitonogamy.
Partionioning of selfing into autogamy and geitonogamy through floral number manipulations. In the 'single flower' treatment inflorescences were trimmed to a single flower, allowing only for autogamous selfing. In the 'single inflorescence' treatment there is opportunity for autogamy and geitonogamy. In the 'many plants' treatment numerous plants flowered simultaneously and so there was no pollen limitation and high seed sets were achieved.



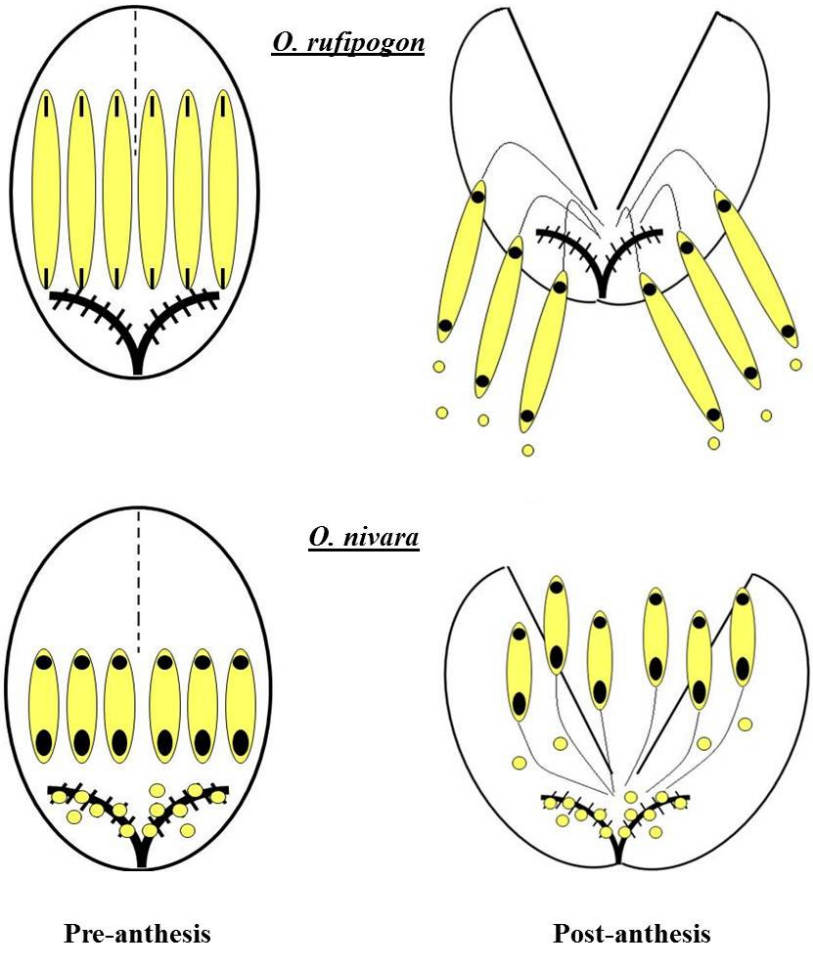
Based on the properties of the mating system traits, it is possible to predict the order in which they may have evolved. Here I describe two possible alternatives for the evolution of autogamy: 1) an indirect path where an increase in selection for geitonogamy provides a bridge to autogamy, or 2) the direct evolution of autogamy.

Selfing likely became beneficial in this system concurrently with a shift to seasonally dry habitats that do not allow plants to persist as perennials. Populations of *O. nivara* are often quite small and potential mates can be scarce today, as they likely were for the original founders of these habitats. Under such conditions there would be selection for reproductive assurance through selfing. A similar situation is observed in populations that exhibit increased selfing rates along range margins, where there are few mates available (Barrett, 2003). For the *O. rufipogon*-like ancestor, selfing would be achieved through geitonogamy in these circumstances. Changes in inflorescence architecture could influence the rates of geitonogamy and variation for this trait is common in natural populations (Table 3). Selection for compact panicles would thereby provide close proximity of flowers and aerodynamic properties favorable for high geitonogamous selfing. Once appreciable levels of self-fertilization were achieved through geitonogamy, there would be selection for reduced investment in allocation to male function. Large anthers are likely to be energetically costly and excess pollen production (for outcrossing) would be unnecessary under these conditions. As anthers began to shorten they would weigh less and maintain a more upright position. At this point gravity would no longer allow pollen to exit through the apical pore; rather the basal pore would be more important in this regard. Selection for large basal pores would allow pollen to fall directly onto the stigma of the same flower, thus providing autogamous selfing. After this point, alterations in the timing of anther

dehiscence and stigma receptivity would result in prior selfing. In this way geitonogamy would serve as a bridge to autogamous selfing.

Figure 4.5. Flower caricatures of *O. nivara* and *O. rufipogon*

Floral caricatures of *O. rufipogon* and *O. nivara*, pre- and post-anthesis. *O. nivara* has smaller anthers than *O. rufipogon* and the anther pores open, releasing pollen through the basal pore prior to anthesis. Upon anthesis the anthers remain upright and dispense pollen through the basal pore onto the stigma. *O. rufipogon* has much larger anthers that do not dehisce prior to anthesis. Upon anthesis the large basal pores hang pendent-like below the flower and pollen is dispensed through the apical pore.



Alternatively, autogamy may have evolved directly in an *O. rufipogon*-like ancestor. In this scenario selfing would have to occur prior to anthesis, as the large basal pores of *O. rufipogon* would fall below the flower upon opening. Therefore, anther dehiscence would be the first trait to evolve. Early anther dehiscence would be unlikely to provide reliable autogamous selfing alone and would have to be accompanied closely by an increase in basal pore size to deposit pollen on the stigma. Additionally, earlier stigma receptivity would enhance the rate of autogamy. Once autogamous selfing became predominant there would be selection for reduced anther size.

In either case traits that promote outcrossing are expected to evolve after autogamy has been established. These traits include pollen viability and stigma characteristics (length, papillae density, pigmentation).

In the work presented here, I have identified traits associated with a shift in mating system between *O. nivara* and *O. rufipogon*. Additionally, I have demonstrated that autogamy is the mode of selfing displayed by *O. nivara.* This research sets the stage for future experiments to gain a more comprehensive understanding of mating system evolution in this system. In the summer of 2013 I will conduct experiments aimed at 1) identifying the specific functional traits controlling the degree of autogamous selfing and 2) mapping the genetic basis of mating system trait variation including the autogamous selfing rate. To date I have generated a mapping population with 200 F2 lines and at least 4 clones of each line. This summer I will finish measuring the mating system traits in the F2 population. Additionally, I will determine the autogamous selfing rate by manipulating flower number as described here. I will also complete genotyping these F2s with microsatellite markers that I have previously developed (Grillo *et al.*, 2009). Through these pursuits I will be able to determine which traits are associated with the

autogamous selfing rate. I will also characterize the genetic architecture of mating system trait differentiation and identify which specific trait loci co-localize with mating system (autogamy) loci.

LITERATURE CITED

LITERATURE CITED

Ackerman JD. 2000. Abiotic pollen and pollination: ecological, functional, and evolutionary perspectives. Plant Systematics and Evolution **222**(1-4): 167-185.

Ågren J, Schemske DW. 1993. OUTCROSSING RATE AND INBREEDING DEPRESSION IN 2 ANNUAL MONOECIOUS HERBS, BEGONIA-HIRSUTA AND B-SEMIOVATA. Evolution **47**(1): 125-135.

Aide TM. 1986. THE INFLUENCE OF WIND AND ANIMAL POLLINATION ON VARIATION IN OUTCROSSING RATES. Evolution **40**(2): 434-435.

Akimoto M, Shimamoto Y, Morishima H. 1999. The extinction of genetic resources of Asian wild rice, Oryza rufipogon Griff.: A case study in Thailand. Genetic Resources and Crop Evolution **46**(4): 419-425.

Armbruster WS, Mulder CPH, Baldwin BG, Kalisz S, Wessa B, Nute H. 2002. Comparative analysis of late floral development and mating-system evolution in Tribe Collinsieae (Scrophulariaceae SL). American Journal of Botany **89**(1): 37-49.

Barbier P. 1989. GENETIC-VARIATION AND ECOTYPIC DIFFERENTIATION IN THE WILD RICE SPECIES ORYZA-RUFIPOGON .1. POPULATION DIFFERENTIATION IN LIFE-HISTORY TRAITS AND ISOZYMIC LOCI. Japanese Journal of Genetics **64**(4): 259-271.

Barbier P, Morishima H, Ishihama A. 1991. PHYLOGENETIC-RELATIONSHIPS OF ANNUAL AND PERENNIAL WILD-RICE - PROBING BY DIRECT DNA SEQUENCING. Theoretical and Applied Genetics **81**(5): 693-702.

Barrett SCH. 2003. Mating strategies in flowering plants: the outcrossing-selfing paradigm and beyond. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **358**(1434): 991-1004.

Barrett SCH. 2010a. Darwin's legacy: the forms, function and sexual diversity of flowers. Philosophical Transactions of the Royal Society B-Biological Sciences **365**(1539): 351-368.

Barrett SCH. 2010b. Understanding plant reproductive diversity. Philosophical Transactions of the Royal Society B-Biological Sciences **365**(1537): 99-109.

Barrett SCH, Harder LD, Worley AC. 1996. The comparative biology of pollination and mating in flowering plants. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **351**(1345): 1271-1280.

Belaoussoff S, Shore JS. 1995. FLORAL CORRELATES AND FITNESS CONSEQUENCES OF MATING-SYSTEM VARIATION IN TURNERA-ULMIFOLIA. Evolution **49**(3): 545-556.

Burd M, Allen TFH. 1988. SEXUAL ALLOCATION STRATEGY IN WIND-POLLINATED PLANTS. Evolution **42**(2): 403-407.

Chang SM, Rausher MD. 1999. The role of inbreeding depression in maintaining the mixed mating system of the common morning glory, Ipomoea purpurea. Evolution **53**(5): 1366-1376.

Charlesworth B. 1992. EVOLUTIONARY RATES IN PARTIALLY SELF-FERTILIZING SPECIES. American Naturalist **140**(1): 126-148.

Charlesworth D, Charlesworth B. 1981. ALLOCATION OF RESOURCES TO MALE AND FEMALE FUNCTIONS IN HERMAPHRODITES. Biological Journal of the Linnean Society **15**(1): 57-74.

Charlesworth D, Charlesworth B. 1987. INBREEDING DEPRESSION AND ITS EVOLUTIONARY CONSEQUENCES. Annual Review of Ecology and Systematics **18**: 237-268.

Charlesworth D, Morgan MT. 1991. ALLOCATION OF RESOURCES TO SEX FUNCTIONS IN FLOWERING PLANTS. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **332**(1262): 91-102.

Conner JK 2006. Ecological genetics of floral evolution. In: Harder LD, Barrett SCH eds. The Ecology and Evolution of Flowers: Oxford University Press.

Crane PR 1986. Form and function in wind dispersed pollen. In: Blackmore S, Ferguson IK eds. Pollen and Spores: Form and Function. London: Academic, 179-202.

Cresswell JE, Krick J, Patrick MA, Lahoubi M. 2010. The aerodynamics and efficiency of wind pollination in grasses. Functional Ecology **24**(4): 706-713.

Dafni A, Firmage D. 2000. Pollen viability and longevity: practical, ecological and evolutionary implications. Plant Systematics and Evolution **222**(1-4): 113-132.

Darwin C. 1876. The effects of cross and self fertilisation in the vegetable kingdom. London, UK: John Murray Publishers.

Dejong TJ, Waser NM, Klinkhamer PGL. 1993. GEITONOGAMY - THE NEGLECTED SIDE OF SELFING. Trends in Ecology & Evolution **8**(9): 321-325.

Dole JA. 1990. ROLE OF COROLLA ABSCISSION IN DELAYED SELF-POLLINATION OF MIMULUS-GUTTATUS (SCROPHULARIACEAE). American Journal of Botany **77**(11): 1505-1507.

Dorken ME, Friedman J, Barrett SCH. 2002. The evolution and maintenance of monoecy and dioecy in Sagittaria latifolia (Alismataceae). Evolution **56**(1): 31-41.

Doust AN, Kellogg EA. 2002. Inflorescence diversification in the panicoid "bristle grass" clade (Paniceae, Poaceae): Evidence from molecular phylogenies and developmental morphology. American Journal of Botany **89**(8): 1203-1222.

Dowding P. 1987. WIND POLLINATION MECHANISMS AND AEROBIOLOGY. International Review of Cytology-a Survey of Cell Biology **107**: 421-437.

Eckert CG. 2000. Contributions of autogamy and geitonogamy to self-fertilization in a mass-flowering, clonal plant. Ecology **81**(2): 532-542.

Fisher RA. 1941. Average excess and average effect of a gene substitution. Annals of Eugenics **11**: 53-63.

Fishman L, Kelly AJ, Willis JH. 2002. Minor quantitative trait loci underlie floral traits associated with mating system divergence in Mimulus. Evolution **56**(11): 2138-2155.

Fishman L, Wyatt R. 1999. Pollinator-mediated competition, reproductive character displacement, and the evolution of selfing in Arenaria uniflora (Caryophyllaceae). Evolution **53**(6): 1723-1733.

Friedman J, Barrett SCH. 2009a. The consequences of monoecy and protogyny for mating in wind-pollinated Carex. New Phytologist **181**(2): 489-497.

Friedman J, Barrett SCH. 2009b. Wind of change: new insights on the ecology and evolution of pollination and mating in wind-pollinated plants. Annals of Botany **103**(9): 1515-1527.

Friedman J, Harder LD. 2004. Inflorescence architecture and wind pollination in six grass species. Functional Ecology **18**(6): 851-860.

Friedman J, Harder LD. 2005. Functional associations of floret and inflorescence traits among grass species. American Journal of Botany **92**(11): 1862-1870.

Goodwillie C. 1999. Multiple origins of self-compatibility in Linanthus section Leptosiphon (Polemoniaceae): Phylogenetic evidence from internal-transcribed-spacer sequence data. Evolution **53**(5): 1387-1395.

Goodwillie C, Kalisz S, Eckert CG 2005. The evolutionary enigma of mixed mating systems in plants: Occurrence, theoretical explanations, and empirical evidence. Annual Review of Ecology Evolution and Systematics. Palo Alto: Annual Reviews, 47-79.

Goodwillie C, Sargent RD, Eckert CG, Elle E, Geber MA, Johnston MO, Kalisz S, Moeller DA, Ree RH, Vallejo-Marin M, Winn AA. 2010. Correlated evolution of mating system and floral display traits in flowering plants and its implications for the distribution of mating system variation. New Phytologist **185**(1): 311-321.

Grillo MA, Li CB, Fowlkes AM, Briggeman TM, Zhou AL, Schemske DW, Sang T. 2009. Genetic architecture for the adaptive origin of annual wild rice, Oryza nivara. Evolution **63**(4): 870-883.

Hamrick JL, Godt MJW. 1996. Effects of life history traits on genetic diversity in plant species. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **351**(1345): 1291-1298.

Hamrick JL, Loveless MD. 1987. ASSOCIATIONS BETWEEN THE BREEDING SYSTEM AND THE GENETIC-STRUCTURE OF TROPICAL TREE POPULATIONS. American Journal of Botany **74**(5): 642-642.

Harder LD. 1998. Pollen-size comparisons among animal-pollinated angiosperms with different pollination characteristics. Biological Journal of the Linnean Society **64**(4): 513-525.

Harder LD, Barrett SCH. 1995. MATING COST OF LARGE FLORAL DISPLAYS IN HERMAPHRODITE PLANTS. Nature **373**(6514): 512-515.

Holtsford TP, Ellstrand NC. 1992. GENETIC AND ENVIRONMENTAL VARIATION IN FLORAL TRAITS AFFECTING OUTCROSSING RATE IN CLARKIA-TEMBLORIENSIS (ONAGRACEAE). Evolution **46**(1): 216-225.

Hughes NM, Reinhardt K, Feild TS, Gerardi AR, Smith WK. 2010. Association between winter anthocyanin production and drought stress in angiosperm evergreen species. Journal of Experimental Botany **61**(6): 1699-1709.

Husband BC, Schemske DW. 1996. Evolution of the magnitude and timing of inbreeding depression in plants. Evolution **50**(1): 54-70.

Igic B, Kohn JR. 2006. The distribution of plant mating systems: Study bias against obligately outcrossing species. Evolution **60**(5): 1098-1103.

Igic B, Lande R, Kohn JR. 2008. Loss of self-incompatibility and its evolutionary consequences. International Journal of Plant Sciences **169**(1): 93-104.

Jain SK. 1976. EVOLUTION OF INBREEDING IN PLANTS. Annual Review of Ecology and Systematics **7**: 469-495.

Kalisz S, Vogler D, Fails B, Finer M, Shepard E, Herman T, Gonzales R. 1999. The mechanism of delayed selfing in Collinsia verna (Scrophulariaceae). American Journal of Botany **86**(9): 1239-1247.

Karron JD, Jackson RT, Thumser NN, Schlicht SL. 1997. Outcrossing rates of individual Mimulus ringens genets are correlated with anther-stigma separation. Heredity **79**: 365-370.

Karron JD, Mitchell RJ, Holmquist KG, Bell JM, Funk B. 2004. The influence of floral display size on selfing rates in Mimulus ringens. Heredity **92**(3): 242-248.

Kearns CA, Inouye DW. 1993. Techniques for pollination biologists. Niwot, CO: University Press of Colorado.

Khatun S, Flowers TJ. 1995. THE ESTIMATION OF POLLEN VIABILITY IN RICE. Journal of Experimental Botany **46**(282): 151-154.

Kirpes CC, Clark LG, Lersten NR. 1996. Systematic significance of pollen arrangement in microsporangia of Poaceae and Cyperaceae: Review and observations on representative taxa. American Journal of Botany **83**(12): 1609-1622.

Kohn JR, Barrett SCH. 1994. POLLEN DISCOUNTING AND THE SPREAD OF A SELFING VARIANT IN TRISTYLOUS EICHHORNIA-PANICULATA - EVIDENCE FROM EXPERIMENTAL POPULATIONS. Evolution **48**(5): 1576-1594.

Lande R, Schemske DW. 1985. THE EVOLUTION OF SELF-FERTILIZATION AND INBREEDING DEPRESSION IN PLANTS .1. GENETIC MODELS. Evolution **39**(1): 24-40.

Levin DA, Kerster HW. 1974. Gene flow in seed plants. Evol. Biol. **7**: 139-220.

Linder HP 1998. Morphology and the evolution of wind pollination. In: Owens SJ, Rudall PJ eds. Reproductive biology in systematics, conservation and economic botany: Kew: Royal Botanic Gardens, 123-135.

Lloyd DG. 1979. SOME REPRODUCTIVE FACTORS AFFECTING THE SELECTION OF SELF-FERTILIZATION IN PLANTS. American Naturalist **113**(1): 67-79.

Lloyd DG. 1992. SELF-FERTILIZATION AND CROSS-FERTILIZATION IN PLANTS .2. THE SELECTION OF SELF-FERTILIZATION. International Journal of Plant Sciences **153**(3): 370-380.

Lloyd DG, Schoen DJ. 1992. SELF-FERTILIZATION AND CROSS-FERTILIZATION IN PLANTS .1. FUNCTIONAL DIMENSIONS. International Journal of Plant Sciences **153**(3): 358-369.

Lyons EE, Antonovics J. 1991. BREEDING SYSTEM EVOLUTION IN LEAVENWORTHIA - BREEDING SYSTEM VARIATION AND REPRODUCTIVE SUCCESS IN NATURAL-POPULATIONS OF LEAVENWORTHIA-CRASSA (CRUCIFERAE). American Journal of Botany **78**(2): 270-287.

Matsui T, Kagata H. 2003. Characteristics of floral organs related to reliable self-pollination in rice (Oryza sativa L.). Annals of Botany **91**(4): 473-477.

Morishima H, Barbier P. 1990. MATING SYSTEM AND GENETIC STRUCTURE OF NATURAL POPULATIONS IN WILD RICE ORYZA-RUFIPOGON. Plant Species Biology **5**(1): 31-40.

Niklas KJ. 1985. THE AERODYNAMICS OF WIND POLLINATION. Botanical Review **51**(3): 328-386.

Niklas KJ. 1987. POLLEN CAPTURE AND WIND-INDUCED MOVEMENT OF COMPACT AND DIFFUSE GRASS PANICLES - IMPLICATIONS FOR POLLINATION EFFICIENCY. American Journal of Botany **74**(1): 74-89.

Oka HI, Morishima H. 1967. VARIATIONS IN BREEDING SYSTEMS OF A WILD RICE ORYZA PERENNIS. Evolution **21**(2): 249-&.

Ornduff R. 1969. Reproductive biology in relation to systematics. Taxon **18**: 121-133.

Ritland C, Ritland K. 1989. VARIATION OF SEX ALLOCATION AMONG 8 TAXA OF THE MIMULUS-GUTTATUS SPECIES COMPLEX (SCROPHULARIACEAE). American Journal of Botany **76**(12): 1731-1739.

Ritland K. 1984. THE EFFECTIVE PROPORTION OF SELF-FERTILIZATION WITH CONSANGUINEOUS MATINGS IN INBRED POPULATIONS. Genetics **106**(1): 139-152.

Ritland K. 2002. Extensions of models for the estimation of mating systems using n independent loci. Heredity **88**: 221-228.

Ritland K, Jain S. 1981. A MODEL FOR THE ESTIMATION OF OUTCROSSING RATE AND GENE-FREQUENCIES USING N INDEPENDENT LOCI. Heredity **47**(AUG): 35-52.

Rodriguez-Riano T, Dafni A. 2000. A new procedure to asses pollen viability. Sexual Plant Reproduction **12**(4): 241-244.

Schemske DW, Lande R. 1985. THE EVOLUTION OF SELF-FERTILIZATION AND INBREEDING DEPRESSION IN PLANTS .2. EMPIRICAL OBSERVATIONS. Evolution **39**(1): 41-52.

Schoen DJ, Brown AHD. 1991. WHOLE-FLOWER AND PART-FLOWER SELF-POLLINATION IN GLYCINE-CLANDESTINA AND G-ARGYREA AND THE EVOLUTION OF AUTOGAMY. Evolution **45**(7): 1651-1664.

Schoen DJ, Lloyd DG. 1992. SELF-FERTILIZATION AND CROSS-FERTILIZATION IN PLANTS .3. METHODS FOR STUDYING MODES AND FUNCTIONAL-ASPECTS OF SELF-FERTILIZATION. International Journal of Plant Sciences **153**(3): 381-393.

Song ZP, Lu BR, Wang B, Chen JK. 2004. Fitness estimation through performance comparison of F-1 hybrids with their parental species Oryza rufipogon and O-sativa. Annals of Botany **93**(3): 311-316.

Stapleton AE, Walbot V. 1994. FLAVONOIDS CAN PROTECT MAIZE DNA FROM THE INDUCTION OF ULTRAVIOLET-RADIATION DAMAGE. Plant Physiology **105**(3): 881-889.

Stebbins GL. 1974. Flowering plants: evolution above the species level. Cambridge, MA: The Belknap Press of Harvard University Press.

Takebayashi N, Morrell PL. 2001. Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. American Journal of Botany **88**(7): 1143-1150.

Vallejo-Marin M, Barrett SCH. 2009. Modification of flower architecture during early stages in the evolution of self-fertilization. Annals of Botany **103**(6): 951-962.

Vaughan DA. 1994. The wild relatives of rice: a genetic resources bankbook. Philippines: International Rice Research Institute.

Vogler DW, Kalisz S. 2001. Sex among the flowers: The distribution of plant mating systems. Evolution **55**(1): 202-204.

Wright SI, Ness RW, Foxe JP, Barrett SCH. 2008. Genomic consequences of outcrossing and selfing in plants. International Journal of Plant Sciences **169**(1): 105-118.

Xie ZW, Lu YQ, Ge S, Hong DY, Li FZ. 2001. Clonality in wild rice (Oryza rufipogon, Poaceae) and its implications for conservation management. American Journal of Botany **88**(6): 1058-1064.

Zheng XM, Ge S. 2010. Ecological divergence in the presence of gene flow in two closely related Oryza species (Oryza rufipogon and O. nivara). Molecular Ecology **19**(12): 2439-2454.

Zhou HF, Zheng XM, Wei RX, Second G, Vaughan DA, Ge S. 2008. Contrasting population genetic structure and gene flow between Oryza rufipogon and Oryza nivara. Theoretical and Applied Genetics **117**(7): 1181-1189.

Zhu QH, Ge S. 2005. Phylogenetic relationships among A-genome species of the genus Oryza revealed by intron sequences of four nuclear genes. New Phytologist **167**(1): 249-265.