The Paradox of Seed Adaptation: Theoretical, Physiological and Genetic Studies of Long-Term Germination Delay in Arabidopsis thaliana seeds

Anne Harlan Prather

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

Doctor of Philosophy

University of Washington

2009

Program Authorized to Offer Degree: Department of Biology

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Anne Harlan Prather

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of the Supervisory Committee:

Rose Ann Cattolico

Reading Committee:

Rose Ann Cattolico

Elizabeth Van Volkenburgh

Elizabeth A. Thompson

Date: July 14, 2009

In presenting this dissertation in partial fulfillment of the requirements for the doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to Proquest Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, 1-800-521-0600, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Anne Harlan Peaths

Date 7-14-09

University of Washington

Abstract

The Paradox of Seed Adaptation: Theoretical, Physiological and Genetic Studies of Long-Term Germination Delay in Arabidopsis thaliana seeds
Anne Harlan Prather

Chairperson of the Supervisory Committee:
Professor Rose Ann Cattolico
Department of Biology

Examination of the relationship between germination efficiency and reproductive success using a simple model of seed population dynamics reveals that for short-lived seeds, opportunistic germination is required in order to facilitate population growth. For long-lived seeds, some inhibition of germination is required in order to achieve population growth. The simulation predicts that for plants growing in restrictive environments, germination frequencies should be low. For seeds which exhibit maternally-derived seasonal germination effects, physiological characteristics that facilitate growth under "off-season" conditions should be evident in the laboratory.

The germination response of *Arabidopsis thaliana* accessions is consistent with the predictions made by the simulation analysis. Cvi, which was collected from a restrictive climate, showed extreme germination delay. Col-O, L*er* and No-O all showed the high germination efficiency predicted for their more favorable habitat. No-O and Cvi both exhibited tolerance to high-temperature, high-humidity storage. In No-O, a positive correlation between season of seed production and germination efficiency was dependent on the maternal parent in F_1 hybrids in a Cvi X No-O reciprocal cross.

QTL analysis in 160 L*er*/Cvi recombinant inbred lines reveals that Cvi's delay phenotype is conferred in 12-week old seeds by a single locus on chromosome

5, called Germination Delay 1 (*GD*1). In 60-week seeds, an additional novel locus on Chromosome 3 (*GD*2) is required for the most extreme delay phenotype. The overall lower mean 48-hour germination frequency of the 60=week old lines is the result of a single L*er*-derived locus (*GD*3) that lies 3 cM to the left of the peak of the *GD*1 locus. Annotation of *GD*1 reveals that this locus is potentially distinct from the after-ripening gene *DOG*1.

The model of seed population dynamics presented here made two predictions regarding seed adaptation under limiting and intermittently catastrophic environmental conditions: a) low germination efficiency under restrictive conditions and b) the necessity for consistent off-season seed production for short-lived seeds. Each of these predictions is evident physiologically and genetically in a panel of four *Arabidopsis thaliana* accessions.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank the following people who supported this work with help in the laboratory: Jesara Schroeder, Jean Velupillai, Eileen Savoy, Elaina Foley, and Sarah Nielson. For editing help: Leslie Jordan and Callie Hills. For external review and technical support: Leonard Lipovich, Sarah Nielson, Arindam Roy Choudhury, and Martin Rosvall..

This investigation was supported in part by Public Health Service National Research Service Award, 2T32 GM007270 from the National Institute of General Medical Sciences:

CHAPTER 1

Introduction: Unlocking the Dormancy/Germination Balance

The *raison d'etre* for seeds, i.e., the dilemma seeds face, can be summed up by a parody of Hamlet's famous soliloquy: To germinate or not to germinate, that is the question—Whether it is better to germinate and risk death from adverse environmental conditions or remain dormant and risk death from failure to reproduce. The farmer understands the germination/dormancy balance from the perspective of the plants that he grows and the weeds he must remove from the soil (Dyer 1995). The ecologist views the regulation of germination versus dormancy as a consequence of environmental cueing (Allen and Meyer 1998). The physiologist and the molecular biologist understand the movement between dormancy and germination as a network of metabolic events driven by a complex choreographed array of physiological and molecular events (To et al. 2006). Whatever the perspective, the fact is that seeds have evolved mechanisms that support survival either by allowing germination and transition to vegetative growth or by insuring long term dormancy (Fenner 1985).

The impulse to understand and control the germination of seeds is partially a consequence of agricultural development (Koornneef and Karssen 1994). Numerous examples highlight the emphasis on germination synchronization for improvement of crop quality and production efficiency (Fenner 1985; Bradbeer 1988; Koornneef and Karssen 1994; Bewley 1997). The study of seed population dynamics, on the other hand (Ferrandis et al. 1996; Halpern et al. 1999; Morris et al. 2000; Alexander and Schrag 2003; Capon and Brock 2006) highlights the importance of a diverse array of germination delay mechanisms. These studies

lead to the idea of a balance between the two processes, in which some degree of germination delay is desirable and in which synchronous germination may be detrimental (Claus and Venable 2000).

The diversity of germination delay mechanisms that exist among different seed types is suggestive of the seed's critical role in plant survival (Flores 2002). Such mechanisms have been associated with each of the major tissue types present in seeds (Bewley 1997; Flores 2002; Garvin and Meyer 2003), as well as with distinct seed shapes within individual plants in some species (Garvin and Meyer 2003). These varied mechanisms act as a choreographed control network to block germination and preserve seeds, providing a hedge against germination in unfavorable years (Fenner 1985; Bradbeer 1988). How did this network of controls arise? What is the precise nature of the connection between environmental cues, molecular regulating mechanisms, and germination?

In the present work, I approach the issue of the dormancy/germination balance through an interdisciplinary suite of studies that encompasses mathematical modeling and simulation, physiological assays, and QTL analysis. To consider the unique nature of the dormancy-germination balance and its implications for experimental design, implementation, and data interpretation, I will begin this study with a brief discussion of seed development, physiology, and ecology.

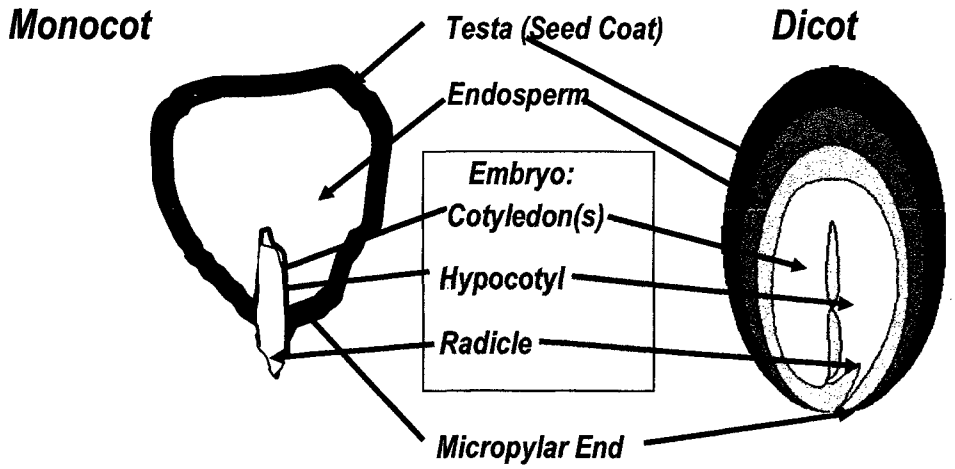
Seed Biology

Survival. Seeds are characterized by several components that facilitate their survival. Each seed has one or more protective seed coat layers, an embryo capable of developing into a fully-formed plant, and storage reserves which provide nutrients and an energy source to the developing embryo. The

arrangement of these components varies with the seed's taxonomic classification (Figure 1-1A) (Flores 2002). The outermost layer in angiosperm seeds, the testa, arises from the maternal tissue of the ovule. In large seeded plants, the testa is mechanically protective, while in smaller-seeded plants, it serves, in conjunction with the embryo, to control germination. Beneath the testa is the endosperm, which in monocot seeds contains the major nutrient source. In dicot seeds, the mass of the endosperm is greatly reduced; its main function appears to be to prevent the radicle from protruding before appropriate germination is programmed to begin. At the core of both monocot and dicot seeds is the embryo, which consists of the cotyledon(s), the shoot axis, and the root axis. In some dicot plants, such as *Arabidopsis*, the cotyledon becomes photosynthetic as the seed germinates, while in others such as legumes, the cotyledons remain below the ground during germination.

Figure 1-1: Seed Anatomy and Embryonic Development

A ***Seed Tissue Layers***



B ***Stages of Plant Embryo Patterning***

Globular Stage ***Heart Stage*** ***Torpedo Stage***

A: Diagram of monocot and dicot seeds showing the major tissue types and structures. **B:** Astra-blue stained media-longitudinal sections of critical embryonic stages in *Arabidopsis* seeds showing globular, mid-heart and mid-torpedo stages. Reprinted from Schere et al. 1994.

Genetics: The three tissue types present in the seeds, seed coat, endosperm and embryo (Fig. 1-1A), each have distinct genetic identities. The *testa*, or seed coat, is formed from the ovule within which fertilization occurs, and is thus maternal tissue. The endosperm is formed when a haploid nucleus in the pollen grain of the male gametophyte fuses with the two polar nuclei of the female gametophyte. The endosperm is therefore triploid, consisting of one paternal and

two maternal genomes. Finally, the embryo arises from the zygote which is the result of the union between a sperm and egg; the embryo is diploid, with one maternal and one paternal genome. The genetic identity of these tissues raises the possibility of localizing the action of germination-control mechanisms from genetic data alone (Papi et al. 2002).

Development. At the end of fertilization, cell division and patterning into both root and shoot axes mark the onset of embryogenesis (Figure 1-1B). The patterning process is mediated by the plant hormone auxin in a cell-autonomous manner (Chen et al. 2001). Patterning determines the number of cotyledons the seed will have. In dicot seeds, three proteins, LEC1, LEC2, and FUS3 have all been shown to be essential for establishing cotyledon identity during late embryogenesis (West et al. 1994; Curaba et al. 2004).

At the completion of embryo development, seeds follow one of two life history phases. Some seeds, particularly those of tropical plants, are ready for germination upon the completion of embryogenesis (Flores 2002). The primary purpose of such seeds appears to be immediate plant replenishment; they lack the protections that would be required to facilitate their long-term survival. Other seeds impose dormancy on the embryo, rendering them temporarily incapable of germination (Smith et al. 2002). Seeds are commonly called *recalcitrant* if their moisture content is equal to that of the surrounding environment and *intermediate* if, in addition to relatively high moisture contents, they develop some drought tolerance. *Orthodox* seeds undergo a maturation process at the end of embryogenesis whose major hallmark is active desiccation.

For both recalcitrant and intermediate seeds, internal water potential passively tracks that of the soil and is not actively influenced by internal developmental events. By contrast, the water potential of orthodox seeds is

determined by a cascade of events during the process of seed maturation, in which the synthesis of amphipathic protein and lipid storage reserves, phase shifts within internal cellular structures and sugar-mediated stabilization of cell membranes combine to impose extreme desiccation on the embryo and surrounding structures. The resulting extremely low internal water potential of orthodox seeds implies that their germination behavior can be predicted as a function of external water potential and time (Bradford 1997).

Seed maturation is initiated by the plant hormone abscisic acid (Parcy et al. 1994; Phillips et al. 1997) and with sugar signaling (specifically altering the ratio of glucose to hexose) (Gutierrez et al. 2007). In legume seeds, the process of maturation begins with the synthesis of storage proteins and lipids, starting with the embryo axes and proceeding outward (Figure 1-1B). As maturation proceeds, an increased level of dormancy is imposed, presumably by both physical and biochemical changes that are a consequence of increasing levels of storage reserves. Throughout this process, a substantial abscisic acid level is maintained, assuring continued inhibition of germination. Interestingly, some morphologically defective mutants of *Arabidopsis* will fail in germination if not allowed to proceed through maturation, but will germinate if maturation is allowed to proceed to completion (Hewitt 2007). However evidence for the idea that maturation and desiccation are essential to the germinability of orthodox seeds is not incontrovertible. The *abi*3 mutant of *Arabidopsis thaliana* is desiccation-sensitive and appears to behave as a non-dormant seed, exhibiting high germination frequencies upon harvest (Koornneef et al. 1989).

Water Relations of Orthodox Seeds

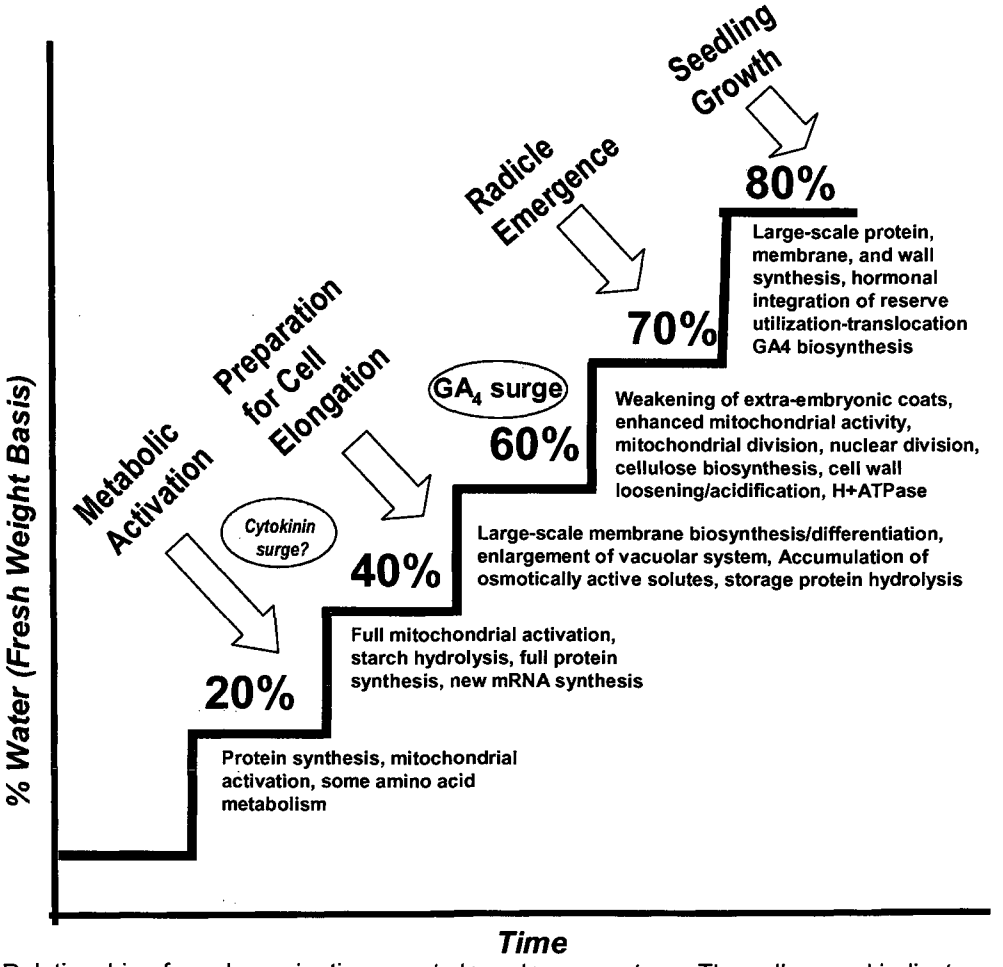
At the end of maturation, the moisture content of orthodox seeds is extremely low, ranging from 3 to 5% (Obroucheva 1999). For each species of orthodox seeds, it is possible to reach a species-specific and constant moisture equilibrium that is independent of the environment. The behavior of these seeds in germination studies therefore depends primarily upon the water availability of the environment (Claus and Venable 2000). This consistency with respect to seed water potential results in extremely predictable germination time course curves whose characteristics can be extrapolated from tests carried out in media of different water potentials (Bradford 1997).

The first steps in the germination process are governed by water uptake, or imbibition (Figure 1-2). Imbibition is typically triphasic, consisting of an initial log phase, followed by a stationary phase, then a final surge of water acquisition just prior to radicle emergence (Obroucheva 1999). During the initial phases of imbibition, active metabolism is initiated. Expression of germination specific genes is up-regulated (Comai and Harada 1990), protein synthesis begins (Schneider and Gifford 1994), and storage reserves are mobilized. During the stationary phase of imbibition, cells at the radicle end of the embryo undergo biochemical changes that cause cell walls to loosen, thus increasing their extensibility. The final phase of imbibition begins then, resulting in an increase in turgor, followed by radicle emergence.

One commercially useful feature of orthodox seeds is that at modest moisture percentages, no damage will occur if seeds are re-dried. Seeds hydrated to approximately 30% moisture content, re-dried and then rehydrated and allowed to germinate, will germinate more synchronously than control seeds that are imbibed for the first time. This process, called *priming*, is used

commercially to synchronize the germination of both agricultural and horticultural seeds (Akalehiywot and D 1977; Corbineau et al. 2000; Warren and Bennett 2000).

Figure 1-2: Water Relations of Seed Germination Stages



Relationship of seed germination events to water percentage. The yellow oval indicates a hormonal event that has been associated with other events at that level. Blue oval indicates an event which has been reported but not verified. (After Obrouscheva 1999 and Smith, 2002)

The Role of Hormones in the Dormancy/Germination Balance

Broadly speaking, the hormones abscisic acid (ABA) and gibberellin (GA) act antagonistically in seed development and germination (Nicolas et al. 1996).

Abscisic acid promotes the onset and maintenance of dormancy (Ni and Bradford 1993; Koornneef et al. 1998; Benech-Arnold et al. 2000; Bentsink 2002; Kucera et al. 2005; Toorop et al. 2005), while gibberellin is required for germination (Ni and Bradford 1993; Derkx et al. 1994; Debeaujon and Koornneef 2000). Most recent evidence suggests that it is the ratio between the two hormones which determines a seed's germination capability (Table 1-1). Light stimulates gibberellin biosynthesis; while changes in temperature increase gibberellin sensitivity as determined by dosage experiments conducted using exogenous gibberellins in the presence of gibberellin biosynthesis inhibitors. The *SPATULA* gene appears to be responsible for coordinating this temperature and light regulation of germination in Arabidopsis (Penfield et al. 2005).

Table 1-1: Genes Associated with Germination and Seed Development

Each gene is delineated with respect to its effect on dormancy, germination, ABA and GA. An L after the direction indicator indicates an increase in the level of the hormone, a P indicates direct up-regulation of genes in the synthesis pathway, and an R indicates an increase in sensitivity. Only when there is an observation regarding the direction of an effect in the specified reference is the direction indicated. ABA: abscisic acid; GA: gibberellin.

Gene Name	Dormancy	Germination	ABA	GA	Ref.
GPA1		Up		Up (R)	(Chen et al. 2004)
BR		Up	?	?	(Chen et al. 2004)
GCR1		Up		Up (R)	(Chen et al. 2004)
AtGA2ox6		Down	dep.	down	(Wang et al. 2004)
AtGA3ox2		Up		Up (P)	(Curaba et al. 2004)
BPBF		Up			(Mena et al. 2002)
HvNCED2	Up	Down	up		(Chono et

Gene Name	Dormancy	Germination	ABA	GA	Ref.
			(P)		al. 2006)
HvCYP707A1	down	Up	down (L)		(Chono et al. 2006)
Phytochrome		Up	down	Up (P)	(Chono et al. 2006)
DAG1		Down			(Papi et al. 2000)
DAG2		Up		Up (R)	(Gualberti et al. 2002)
NCED6			up (P)		(Lefebvre et al. 2006)
NCED9	up (req. NCED6)		up (P)		(Lefebvre et al. 2006)
CYP707A			down (L)		(Saito et al. 2004; Okamoto et al. 2006)
PA 4'-reductase			down (I)		(Saito et al. 2004)
LEC1			up (R)		(Parcy et al. 1997)
ABI1 PP2C	Down	Up	down (R)		(Gosti et al. 1999)
HvSPY			up (I)	down (I)	(Robertson et al. 1998)
ABI4	Up	Down	up (R)		(Finkelstein et al. 1998)
G-rich/LEA-like	Up				(Nicolas et al. 1997)
Ethylene	no effect	up (GA independent?)	down	Up	(Kucera et al. 2005)
BR			down	Up	(Kucera et al. 2005)
MARD 1	Up		up (R)		(Kucera et al. 2005)
Auxin	no effect	no effect			(Kucera et al. 2005)
Cytokinin		increase, ethylene req.			(Kucera et al. 2005)
SPATULA	Up			down (P)	(Penfield et al. 2005)

As the above paragraphs illustrate, there are a variety of germination-control mechanisms which contribute to germination behavior in *Arabidopsis thaliana* seeds. It will be helpful to define these mechanisms in order to clarify the distinct aspects of germination behavior studied in this work. Table 1-2 presents a brief definition of the relevant terms.

Table 1-2: Definition of Germination-Inhibition Mechanisms

MECHANISM	DEFINITION
Dormancy	Lack of germination in an otherwise viable seed.
After-ripening requirement	The length of time it takes post-harvest for a seed lot to achieve maximum germination efficiency. Measured in the laboratory by subjecting dry-stored seeds from a single planting to regular germination tests until maximum germination is achieved.
Imbibition	The taking up of water by dessicated seeds.
Germination	The point in seed imbibition when the radicle just protrudes from the seed coat.
Germination delay	Used in the literature to refer to all forms of germination inhibition. In this work, refers to reduced germination. efficiency in seed lots that do (eventually) reach completion.
Long-term germination delay	Inhibition of germination competence that persists beyond the after-ripening period.
Stratification	A period of moist chilling required for maximum germination efficiency. For the purposes of this work, the treatment conducted at 4° C in darkness.

The Evolutionary Ecology of Seed Germination Control

The evolution of seed dormancy is thought to date back to the Paleozoic period for gymnosperms (Mapes et al. 1989). These authors speculate that the evolution of dormancy was the result of climatic change that led to dryer conditions. Indeed, the development of desiccation tolerance would tend to favor survival under such conditions, and there are many regulatory pathways, genes and gene products common to seed dormancy, the development of desiccation

tolerance, and tolerance to osmotic stress (Gilmour et al. 1992; Campalans et al. 1999). For example, an increase in the amount of solutes within seed tissues, combined with larger amounts of oils and other hydrophobic compounds would elevate stored energy levels at the expense of water availability within the seed. The transition from an intermediate seed phenotype to an orthodox one would result in extremely low internal moisture content, but orthodox seeds gain the advantage of being able to precisely time germination based on external water availability and moderate temperature (Bauer et al. 1998; Claus and Venable 2000).

In annual plants, the evolution of germination strategy is dependent on whether the primary risk to seedlings is drought or frost (Fenner 1985; Allen and Meyer 1998; Loprz et al. 1999; Claus and Venable 2000). If frost is the primary risk factor, seeds will develop a "summer annual" growth habit, germinating in the spring and setting seed in the fall. Summer annual plants over-winter as seeds which are highly resistant to frost damage. By contrast, if drought is the primary risk to young seedlings, a "winter annual" life habit will result, with seeds germinating in the fall or winter. Plants will set seed in the spring and will survive the dry summer months as seeds which are highly drought-resistant.

The native life-habit of annual plants can be inferred by the length of their moist-chilling requirement. Winter annuals typically germinate very quickly in response to a moist chilling (stratification) treatment, while summer annuals exhibit a much greater stratification requirement. Thus, the difference in germination behavior between these two types of annual plants is reflected by the length of the stratification requirement.

The Genetics of Arabidopsis Adaptation

The use of quantitative trial locus (QTL) analysis to elucidate evolutionary processes is well-established (Westerman and Lawrence 1970). However, because each favored phenotype often arises at the expense of other (potentially favorable) phenotypes, the appropriate relationship between a given quantitative trait and its evolutionary antecedents is not always evident. This ambiguity particularly applies to those phenotypes which have arisen multiple times. The situation of convergent evolution is particularly relevant in inbred, self-fertilizing populations in which genetically distinct lineages are undergoing natural selection in locations very near one another and (presumably) experiencing similar selective pressures. Examples of just this situation are frequent in *Arabidopsis thaliana* (Weining et al. 2002).

The recent divergence of *Arabidopsis* across its broad range is suggestive of rapid adaptive change which, in turn, is suggestive of evolution by means of alterations in relatively few genes. Indeed, this hypothesis has been supported in studies of flowering time (Alonso-Blanco et al. 1998) and nitrate storage (Harada et al. 2004). Because *Arabidopsis* is primarily a self-fertilizing organism, it cannot "hide" harmful recessives in heterozygotes. Therefore, it is likely that a sizable number of traits evolve by balancing selection, in which both positive and negative effects combine to produce the favored intermediate phenotype (Malmberg et al. 2005). A most elegant example of this phenomenon is the study of leaf circadian movement in *Arabidopsis*, which demonstrates a classic model of intermediate phenotype in the parents and segregation to the extremes in the recombinant inbred lines (RILs) (Swarup et al. 1999).

Six reports of natural allelic variation affecting germination have been completed to date, culminating in the cloning of the first after-ripening-related QTL to be identified by natural allelic variation. Two of these studies were

conducted in Landsberg *erecta* (L*er*) and Columbia (Col) (van der Schaar et al. 1997; Donohue 2002; Donohue et al. 2005; Donohue et al. 2005; Malmberg et al. 2005), one in Cvi (Cape Verde Isles) (Alonso-Blanco et al. 2003) one in a set of accessions from Tacoma and Calgary (Donohue 2002; Weining et al. 2003; Donohue et al. 2005) and another involving L*er* as one parent and Shakdara (Sha) (Clerkx et al. 2004). Recently, the gene within the major Delay Of Germination associated QTL in Cvi, *DOG*1, has been cloned (Bentsink et al. 2006). Together these studies provide a powerful context in which to undertake the analysis of the critical adaptive changes in germination control that are predicted to result from rapid environmental change.

In the following chapters, data are presented which shed light on the nature of adaptive mechanisms underlying the regulation of long-term germination delay (Table 1-2). This study was undertaken in *Arabidopsis*, which presents several advantages as a study system: 1) the germination behavior of orthodox seeds is well understood, 2) the performance of multiple germination experiments in the laboratory is extremely straightforward and feasible, and 3) the predictive nature of germination with respect to other plant life history traits has been extensively studied (Hampton and TeKrony 1995).

In the present work, I examine seed germination control from three perspectives.

Chapter 2: First, I elaborate on currently-available models of soil seed bank dynamics with the goal of predicting the nature of adaptive phenotypes in seed germination control. The balance between seed number (i.e., per-plant seed yield) and seed size is a well-established observation in seed ecology (Leishman et al. 2000), but the use of equilibrium yields for the study of population dynamics in persistent soil seed banks (PSBs) has not been employed to date. In this

chapter, I define a quantity called "equilibrium yield" (y_{eq}) as the number of seeds per plant required to maintain a seed population at equilibrium, where the number of seeds in year N is equal to the number of seeds at year 0. I first describe how this quantity is calculated for both transient seed banks (TSBs), in which all seeds either germinate or die in the following year, and persistent seed banks in which a portion of seeds remain viable for more than one year. I then employ equilibrium yields for TSBs in the comparison of germination and establishment strategies under both constant year-to-year constraint and intermittently catastrophic years in which no plants survive to reproduce.

Chapter 3: I examine four *Arabidopsis* accessions to determine whether any of the predicted adaptive phenotypes are actually present and detectable in a system amenable to high throughput QTL analysis. The work on this research chapter builds on the work in Chapter 2 by providing physiological evidence of the phenotypes and responses predicted by the model: germination delay, timed germination and opportunistic germination. For each accession, I include a longitudinal analysis of multiple germination tests spaced at random intervals during a period of as long as 80 weeks post-harvest. For each accession, distinct components of the phenotypes detected by the simulation analysis in Chapter 2 are revealed. The heritability of two of these components is established. In addition, I validate the use of rapid-germination assay for the quick and definitive detection of QTLs associated with the different aspects of germination control. The results of this experimental chapter allow me to determine: (1) the best time at which to assay germination frequency for mapping purposes, (2) the mechanism by which germination-delay-related alleles are likely to transgress, and (3) the best strategy for detecting all possible germination-associated alleles

without confounding the detection process by the involvement of multiple pathways.

Chapter 4: Finally, I use composite-interval (CI) mapping, epistasis analysis and genome annotation to determine the location and nature of genomic regions associated with germination delay, the major adaptive phenotype predicted by the simulation analysis. The mapping studies in this chapter confirm that the primarily adaptive phenotype predicted by the model in Chapter 2 is in fact the phenotype exhibited by the Cvi accession which is resident in the allopatric environment of the Cape Verde Isles. The large amount of mapping information available for L*er* and Cvi accessions is utilized to conduct a QTL mapping study on germination delay. This study is distinct from all others published to date because: 1) the use of the radicle-extension criteria and the germination incubation curve for detection of delay alleles limits the number of physiological pathways potentially at play, 2) the extremely long storage periods between mapping studies allows detection of both after-ripening associated delay and long-term, constituitive delay, and 3) the use of a stratification-rescue treatment distinguishes between dormant and non-viable seeds. My observations concerning the major germination delay loci confirm those of Alonso-Blanco et al. and Clerkx et al. with respect to after-ripening (Alonso-Blanco et al. 2003; Clerkx et al. 2004). Moreover, I have identified two additional loci that appear to be implicated in long-term germination delay. The inheritance of a single L*er*-derived delay allele is predicted by the data in Chapter 2. The identification of three well-isolated QTLs allow the annotation of the three critical regions in the *Arabidopsis* genome and the identification of a small list of potential candidate genes for further study.

Chapter 5: This chapter presents conclusions and insight into the contribution made by this study. A critical problem in understanding the interaction between organisms and their environment is the difficulty of integrating information obtained at the molecular, cellular, organismal, and ecological levels. By employing an interdisciplinary suite of strategies, I have demonstrated how an appropriate choice of simulations, physiological work, and genetic studies can help tie these multiple levels together. This tie-in is achieved for the following reasons: First, the intense predictability of germination behavior in orthodox seeds provides a quantifiable link between the application of an environmental cue and a germination time course curve. Second, the production of recombinant inbred lines closely parallels the probable evolutionary scenario in Arabidopsis. Third, the recent divergence of Arabidopis across its broad habitat range reduces the amount of time within which adaptive mutations can occur. This reduction, in turn, implies that many adaptive phenotypes present within Arabidopsis inbred lines will eventually turn out to be the result of the interaction between a relatively small number of genes.

Ultimately, the ability to fully utilize the vast data base being accumulated on the genomic, molecular, and physiological bases of heritable traits will depend on my understanding of the processes and the circumstances that gave rise to them through the action of evolution. It is expected that this work will contribute to my understanding of seed germination by adding to my information concerning the loci affecting seed germination. Additionally, the particular experimental simulation approach employed in this work has the potential to identify other critical factors that may prove significant in other complex phenotypes .

In summary, the data provided in this dissertation give insight into the relationship between environmental variability, ecology, physiology and genetics

with respect to seed germination success. My research provides both a model for further experiments that unite these various levels and the context in which to place data that is rapidly accumulating in this field of study.

CHAPTER 2

Use of Equilibrium Yield Requirements for the Analysis of Adaptive Survival Strategies in the Maintenance of Persistent Soil Seed Banks

ABSTRACT

There are two essential strategies for ensuring the reproductive efficacy of seeds. One is to increase the number of plants that bear offspring (here referred to as "established" plants) and the other is to increase the length of time for which ungerminated seeds remain viable to form a *persistent seed bank* (PSB). Using a series of simple calculations for determining the per-plant seed yield required to maintain plant and seed populations at equilibrium, I demonstrate that increases in both seed longevity and in the number of seed-bearing plants lower equilibrium yield requirements substantially. Comparison of low-germination/high establishment and high-germination/low establishment scenarios with identical transient seed bank (TSB) equilibrium yields reveals that low-germination strategies result in a nearly 100-fold advantage in short-lived PSBs and a nearly 10,000-fold advantage in long-lived PSBs over high germination/low establishment strategies with identical TSB equilibrium yield requirements. The occurrence of off-season germination increases the population almost 100-fold for short-lived PSBs but has little effect on long-lived PSBs. Random, infrequent catastrophic conditions (1 year in fifteen with no reproduction) resulted in population decrease in the high-germination/low-establishment scenario in the 1-generation PSB. This effect was mitigated if some seeds could germinate and reproduce successfully in the "off-season" in one year out of four. When the

frequency of catastrophe was increased to one year in every five, a population crash resulted from both low and high germination strategies. For the low-germination strategy, off-season germination one year in four was sufficient to allow population increase. However, for the high-germination strategy, off-season germination was required in every year. For a 20-generation PSB, substantial population increase resulted from every strategy and was maintained even in the presence of frequent catastrophic conditions. However, the "off-season" scenario resulted in only a 10-fold increase in seed population growth when no catastrophe occurred. Results for scenarios involving catastrophic conditions were proportionally smaller, suggesting that this strategy is much more relevant in short-lived PSBs. These results suggest that maximizing the portion of seeds that establish even in the presence of low germination frequency is more effective for population maintenance for both short- and long-lived PSBs than is an increase in germination efficiency alone.

INTRODUCTION

The high per-plant seed yields observed for many plant species are a consequence of the failure of a majority of seeds to reproduce (Fenner 1985). Thus, reducing the frequency of reproductive failure in seeds is a large factor in lowering the per-plant yield requirement. Leaving aside the issue of dispersal (which results in loss to the system independent of life-history status) and immigration (which results in gain to the system independent of life-history status), the only way to improve the population's reproduction efficacy is to increase the portion of seeds that yield seed bearing plants. One way to accomplish this goal is to increase germination efficiency. Assuming that the portion of germinating seeds which reproduce remains constant, improving

germination efficiency will result in improved reproduction efficacy. A second strategy for improving reproduction efficacy is to increase the length of time for which seeds remain viable and capable of germination, thus increasing the chance that each individual seed will be able to reproduce successfully.

Annual plants, which undergo only one reproductive opportunity in their lifetimes, face a particular challenge regarding population maintenance. Once reproduction has occurred and the parent plant dies, population maintenance is completely dependent on the number of seeds which can germinate and reproduce successfully in the following year. If seeds have no longevity, and the season immediately following seed dispersal is not favorable, the population will quickly go extinct. Numerous ecological and physiological studies suggest that, in fact, annual plants have evolved an intricately regulated alternating pattern of life history phases that circumvent this challenge (Fenner 1985; Bradbeer 1988; Koornneef and Karssen 1994; Bewley 1997; Allen and Meyer 1998). Germination control is a central component of this regulated transition between phases.

Two kinds of germination response patterns have been documented for annual plants. One is timed germination, in which seeds stay dormant for a length of time equal to that of the "unfavorable" season (Bauer et al. 1998). A second is opportunistic germination, in which a few plants establish and reproduce during unfavorable times, yielding seeds that are less dormant and thus capable of rapid germination in the immediately following favorable season (Derkx and Karssen 1994; Munir et al. 2001).

Seeds from both annual and perennial species persist in the soil as "seed banks" (Dyer 1995). Evidence from a range of studies highlights the importance of seed banks in the maintenance of plant populations (McCue and Holtsford 1998; Morris et al. 2000). Seed banks allow germination of new plants to take

place, both under "normal" growth conditions and sometimes under a variety of unusual circumstances, such as natural disasters (Reyes and Casal 2001; Capon and Brock 2006), during periods of increased rainfall (Claus and Venable 2000), or after years in which no seeds are produced (Botto et al. 1998).

The structure of seed banks is measured by three main methods: (1) direct counting of large seeds after separation from the surrounding soil, (2) germination testing of soil samples from defined transects, and (3) DNA fingerprinting. For large-seeded plants such as sunflower (*Helianthus annuus*), the seeds can be separated from the soil, weighed and counted (Lorthe and Turkington 2002). By taking soil cores from defined transects within a particular location, the number and structure of the seed bank for a selected plant type can be assayed. The problem with this method is that it is almost impossible to separate tiny seeds from the surrounding soil in the sample. For this reason, seed bank diversity is most often determined by a soil germination test. In this test, soil samples are gathered and placed in a greenhouse or growth chamber under conditions designed to facilitate germination. After an appropriate time, the resulting seedlings are counted and identified with respect to species. The germination-test method will identify seeds of all sizes in the seed bank, but will miss any seeds not capable of germinating because of their dormancy status. This method also does not directly address the number of seeds; it simply asks what species are present. Finally, DNA fingerprinting techniques have recently been used to measure the structure and diversity of seed banks (Schmidt and Jensen 2000; Federici et al. 2001; Huh and Huh 2001; Bahulikar et al. 2004). This method will answer questions about the species present within a seed bank and (if quantitative techniques can be applied) may yield some insight about seed number. Multiple issues, e.g., variation in seed size, toughness, and shape,

involved in DNA retrieval from heterogeneous seed populations potentially affect the estimation of seed diversity and species-specific population size.

To assess the adaptive constraints on the genetics of seed germination regulation, it would be useful to understand the quantitative relationship between seed yield, establishment percentage (the number of germinants which go on to reproduce successfully), and germination percentage (the number of seeds which germinate, regardless of whether the resulting plants go on to reproduce successfully). By comparing outcomes from different combinations of these parameters, it would be possible to determine the optimum germination percentage at which, for example, a low-yield plant could survive. It would also be possible, given seed yields, to predict the nature of germination-control strategies. Although several field studies of germination have been conducted in a range of annual plants, and the presence of these plants' seeds in seed banks has been documented (Allen and Meyer 1998; Venable et al. 1998; Claus and Venable 2000), no study to date has attempted to establish a quantitative relationship among seed yield, establishment percentage, and germination percentage parameters.

In the present work, I first employ a simple set of calculations for determining the equilibrium yield (the number of seeds per plant required to maintain a seed population at equilibrium, where the number of seeds in year *N* is equal to the number of seeds at year 0 -- designated "y_{eq}") for *transient seed banks* (TSBs), in which all seeds either germinate or die in the next favorable season and *persistent seed banks* (PSBs), in which a portion of seeds neither germinate nor die but remain viable for germination in succeeding years. I demonstrate that increases in both germination and establishment frequency lower equilibrium yield requirements. I next employ scenarios whose TSB y_{eq}s are identical, but

whose germination and establishment frequencies differ, to determine the cost of high germination under environmental conditions which limit the number of reproducing plants. I compare the results from scenarios involving no off-season germination to those that involve off-season germination. Finally, I examine the effect of random, intermittent catastrophic years on the same combination of dormancy and off-season scenarios. The results of this work shed light on measureable traits which are essential for the maintenance of persistent soil seed banks.

RATIONALE

Consider a hypothetical annual plant, which reproduces once and then dies, storing its seeds in a soil seed bank. For the purposes of this work, I will consider two types of seed banks: TSBs and PSBs. In a TSB, all seeds present at the beginning of the year either yield reproducing plants or die without producing offspring. In a PSB, some portion of the seed population contributes to the number of germination-capable seeds present at the beginning of the following year. I wish to understand the relationship between the per-plant yield (a quantity which can be directly measured) and the characteristics of the seeds that survive through multiple years in PSBs. I begin my discussion with seed banks in which no net change in seed population occurs between succeeding years. A hypothetical seed bank exhibiting this behavior is said to be in equilibrium.

The initial number of seeds in the seed bank, D, is represented by D_0 while values of D for succeeding years are represented by D_N where N is the number of the year under consideration. For a seed bank at equilibrium

$$D_0 = D_1 = D_N$$

meaning that the number of seeds in the seed bank remains constant from year 0 to year N.

In a TSB, all the seeds present at the beginning of the year must be replaced in that year:

$$D_N = Y_N$$

The population yield Y is a function of the per-plant yield y and the number of plants that reproduce successfully. I can represent the portion of seeds that germinate and establish to form successfully reproducing plants as G_E. The population yield can be written as:

$$Y = yG_E$$

and because $D_N = Y_N$, the equilibrium yield for the TSB is equal to

$$\text{TSB } y_{eq} = D_N/G_{E(N)} \qquad \text{Eq. 1}$$

where TSB y_{eq} is the number of seeds per plant required to maintain the transient seed bank at equilibrium.

Notice that the equilibrium yield is a function of the number of reproducing plants and not a function of the number of seeds that germinate. In this simple case, all seeds are "lost" to the bank and must be replaced.

In a PSB, a portion of the seeds present at the beginning of year N neither germinate nor die, but remain viable for a period of time after seed set in year N. This quantity of remaining seeds can be represented by S_R and assuming that no seeds are lost to the bank for any reason other than germination, the portion of seeds from previous years remaining in the PSB at the end of year N can be written as:

$$S_{R(N)} = D_N - G_N$$

and because all the seeds lost to the bank are lost due to germination in the

simplest case, the number of seeds that must be replaced by the population yield is:

$$Y = D_N - S_{R(N)}$$

and the equilibrium yield for a PSB is therefore:

$$\text{PSB } y_{eq} = (D_N - S_{R(N)})/G_{E(N)}. \qquad \text{Eq. 2}$$

In other words, the number of seeds lost divided by the number of reproducing plants gives the per-plant equilibrium yield in the simplest case of a PSB.

The above equation assumes that all seeds lost from the seed bank are lost due to germination. In reality, a certain portion of seeds in the S_R term will also be lost. Adding a survival coefficient to the S_R term, I obtain:

$$\text{PSB } y_{eq} = (D_N - (H_{SR}S_{R(N)}))/G_{E(N)} \qquad \text{Eq. 3}$$

where H_{SR} is a survival hazard coefficient applied to the seeds in the seed bank which remain dormant. If H_{SR} is constant from year to year, the viability of seeds in the seed bank follows an exponential decay model, and their longevity can be predicted from Equation 3. Taking the natural log of the exponential expression for decay, and because the hazard coefficient refers to the portion of seeds that remain viable during the year, the half-life is:

$$S_{HL} = ln2/(1-H_{SR})$$

where S_{HL} is the seed half-life.

Environmental carrying capacity: Vegetative growth is more costly in terms of environmental resources than the maintenance of viable but dormant seeds. In the case of populations growing at or near environmental carrying capacity, the equilibrium yield becomes a function of the limiting value on establishment, and increases in proportion to the frequency of germination,

$$y_{eq} = (D - S_R)/e_L \qquad \text{Eq. 4}$$

where e_L is the maximum number of reproducing plants the environment can

support. Thus, the increase in y_{eq} results from the smaller number of seeds in S_R and the correspondingly larger number of germinated seeds that must be replaced. The extent to which the above equation applies to a population growing under limiting environmental conditions depends on what happens when the limiting value is reached. If growth is logistic, the above equation applies once the population growth converges on the limiting value. If, however, reaching the limiting value results in a "crash" in the year following limiting growth, the equilibrium yield will be year-specific and dependent on the number of reproducing plants during each year. In either case, the equilibrium yield for a PSB is dependent on both the number of successfully reproducing plants and the half-life of the seeds in S_R.

Germination-control scenarios: Besides scenarios involving increasing dormancy and seed longevity, two other types of scenarios are relevant for understanding adaptive traits for the maintenance of PSBs. In timed germination, germination capability is maximized at a season which either increases the percent of plants that can successfully reproduce or increases the environmental carrying capacity, or both. Calculation of the equilibrium yield for timed-germination scenarios is identical to that for dormancy scenarios except that they are calculated using the increased number of established plants resulting from increased establishment percentage and/or environmental carrying capacity.

For scenarios involving multiple germination opportunities within the year, the calculation of equilibrium yield becomes complex. The yield from the year immediately preceding the "off-season", as well as S_R from that year both contribute to off-season seed production. In some cases, the equilibrium yield for "off-season" scenarios falls below one, because reproduction in the off-season

would completely compensate for the lack of reproduction in the favorable season.

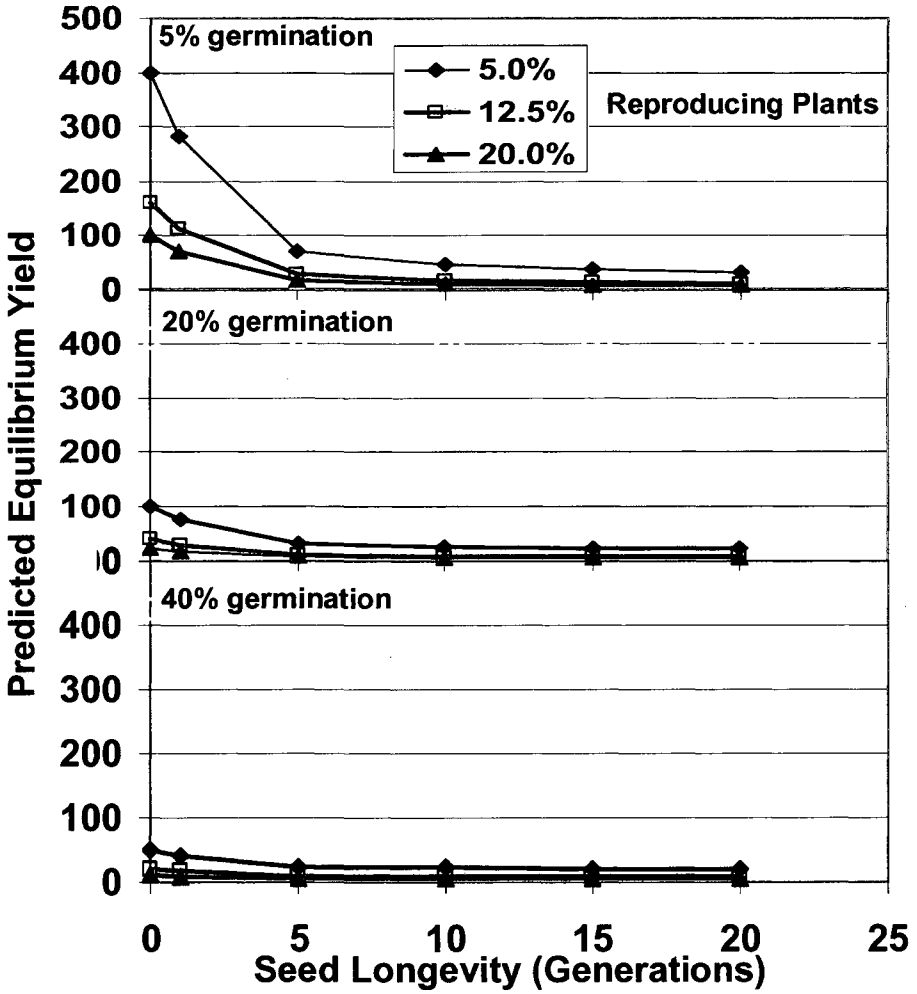
RESULTS

Increases in seed longevity, germination frequency and establishment frequency all lower equilibrium yield requirements. Lowering of per-plant seed yield requirements allows plant resources to be diverted to other life-history phases, such as vegetative growth (Leishman et al. 2000). To investigate the relationship between per-plant yield and population growth, I used Equations 1 and 3 to calculate precise equilibrium yields for both TSBs and PSBs. When germination and establishment frequency are both 5%, equilibrium yield requirements are strongly dependent on seed longevity, with the greatest difference occurring with S_{HL}s of 1-5 generations. At higher germination frequencies, equilibrium yields are much less dependent on seed longevity, even at low establishment frequencies. These results suggest that the greatest impacts on equilibrium yields at low germination frequencies result from a 5-generation seed half-life, while at higher germination frequencies, seed longevity has a proportionally smaller effect on equilibrium yields.

Low-germination/high-establishment strategies have the greatest positive effect on seed population growth. The maintenance of adult plants is much more likely to be a limiting factor in seed/plant population growth than is the maintenance of a PSB. Thus, limits on establishment exert a cost on the seed bank proportional to the number of seeds that germinate. If many seeds

germinate, but few seedlings live to bear fruit, more seeds are lost to the seed bank, and the equilibrium yield requirement increases. This feature of soil seed banks has the interesting property that several combinations of termination and establishment frequencies have identical TSB y_{eq}s, but their persistent equilibrium yields differ (Table 2-1).

Figure 2-1: Effect of Seed Longevity on Equilibrium Yield



Impact on equilibrium yield requirements of increased seed longevity, germination frequency and establishment frequency. Equilibrium yields assume that no seeds are lost to the seed bank except due to germination or decrease in seed viability.

Table 2-1: Relationship of Germination and Establishment Frequencies to Equilibrium Yield

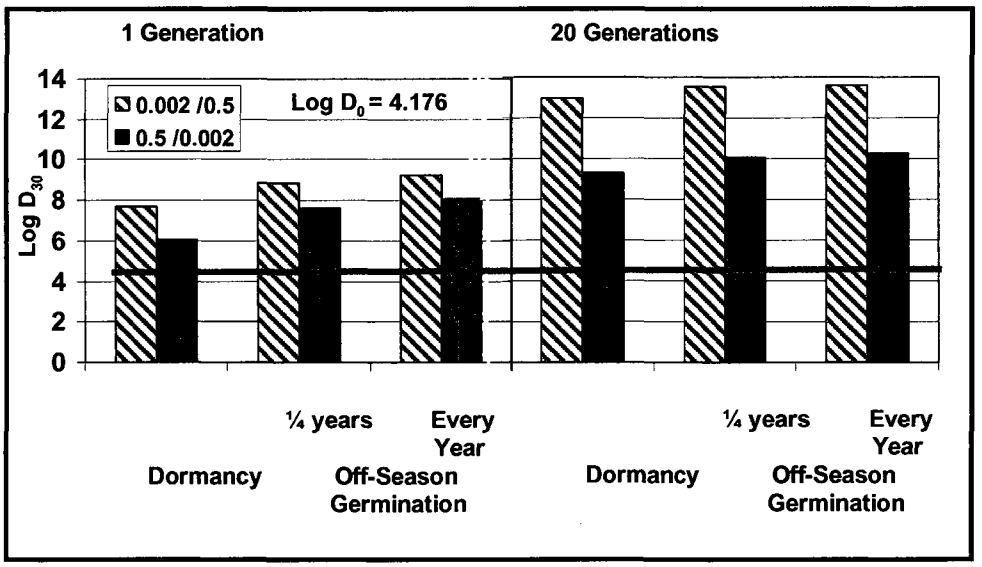
Germination and establishment frequency combinations with identical TSB y_{eq}s. Shaded cells represent scenarios compared in Figures 2-2 and 2-3. All yields assume no loss to the seed bank except through exponential decrease in seed viability and germination. Headers are: g = germination coefficient, g_e = establishment coefficient.

g (%)	g	g_e (%)	g_e	TSB y_{eq}	PSB 20 gen. y_{eq}	PSB 1 gen. y_{eq}
.2%	0.002	50%	0.5	1000	36.59	693.76
2%	0.02	5%	0.05	1000	53.96	699.28
20%	0.2	0.5%	0.005	1000	227.73	754.52
.5%	0.005	20%	0.2	1000	39.48	694.68
5%	0.05	2%	0.02	1000	82.92	708.49
50%	0.5	0.2%	0.002	1000	517.33	846.57

As Table 2-1 suggests, if maintenance of a particular equilibrium yield requirement were the sole requirement for adaptive success, many possible combinations of germination and establishment frequencies would suffice for a TSB. Furthermore, if low equilibrium yield were required, seed longevity confers a clear advantage for PSBs. However, I wished to understand the role of off-season germination in the maintenance of seed populations under low-germination/high-establishment and high-germination/low-establishment behavior. Accordingly, I generated 30-generation population (D_{30}) figures for the scenarios in Rows 1 and 6 of Table 1, which had the lowest and highest PSB equilibrium yield requirements, respectively. As suggested by the table, the 50% germination frequency/0.2% establishment frequency combination had the smallest population increase in both 1- and 20-generation PSBs (Figure 2-2). When germination in the off-season was added, the populations in the 1-generation seed bank experienced a 100-fold advantage in the high-germination/low establishment scenario. This effect was much less noticeable in the 20-generation seed bank. These findings suggest that under constant

environmental limitations, a short-lived PSB is sufficient to confer a 10,000-fold increase for the low-germination strategy but only a 100-fold increase for the high-germination strategy. Off-season germination is particularly important for population growth in 1-generation PSBs, and its effect is roughly equivalent to the effect of increased longevity in a high-germination/low establishment scenario. The high-germination/low-establishment scenario is a good representation of what happens as a high-germination population converges on the limiting value imposed by the environment. These data demonstrate that low germination frequency has a substantial advantage in resource-limited environments, even if the seeds are short-lived.

Figure 2-2: Effect of Environmental Carrying Capacity



Comparison of low germination/high establishment and high-germination/low establishment scenarios for 1-generation and 20-generation soil seed banks. All scenarios had identical equilibrium yields for a transient seed bank. Solid blue line represents the seeds in the bank at the beginning of year 0. Germination probability is given first in graph legend.

Both off-season germination strategies and increases in seed longevity compensate for seed loss during intermittent catastrophic years. Although seed

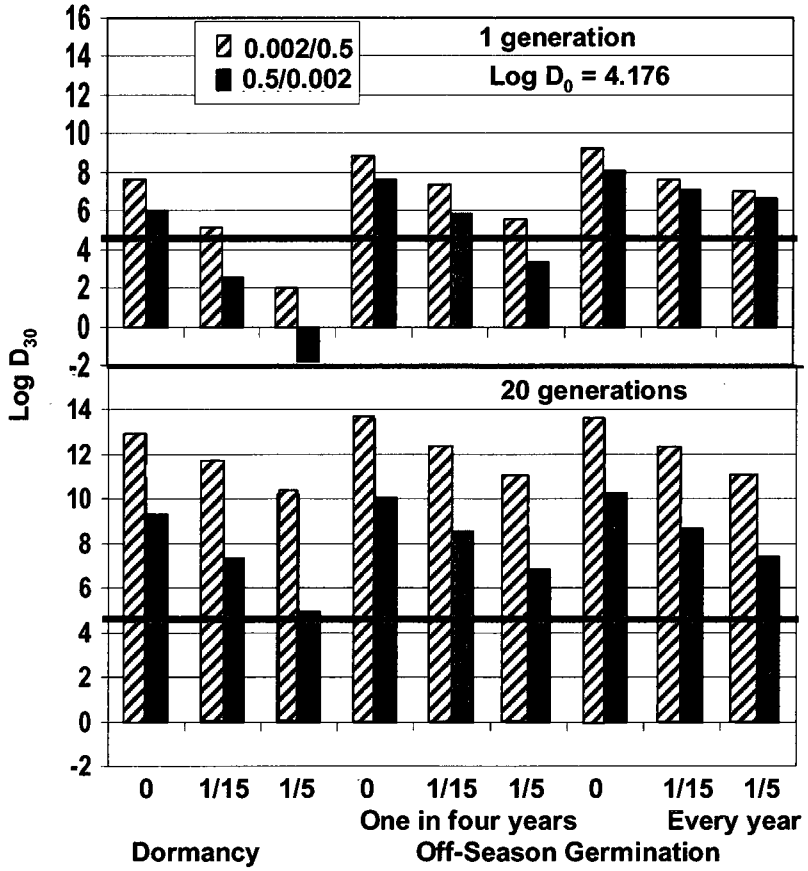
longevity and off-season germination both facilitate population growth even in the context of environmentally-limited reproduction, I wished to determine whether there was a difference in the efficacy of these strategies under variable environmental conditions that occasionally result in complete reproductive failure during the normal growth season. To examine the effect of such conditions on PSBs, I used the germination and establishment frequencies in Rows 1 and 6 of Table 2-1, together with the corresponding off-season germination scenarios to examine predicted population numbers at the end of 30 generations. Off-season in every year was required to compensate for reproductive failure in the high-germination scenario of the 1-generation PSB for reproduction failure that occurred approximately every 5 years. In this scenario, only the low-germination model predicts any increase above equilibrium, and this increase is slightly less than 10-fold (Figure 2-3). In a 20-generation PSB, the high-germination scenario is just barely above equilibrium, even though the per-plant yield of 1000 is nearly twice the yield equilibrium requirement for the non-catastrophic version of this scenario (Figure 2-3). These results suggest that it is the intermittent occurrence of complete reproductive failure that differentiates the effect on population growth of germination-control strategies, changes in seed longevity, and off-season strategies, rather than their effect on equilibrium yield requirements.

DISCUSSION

Although germination frequency is a good predictor of crop plant growth efficiency, its relationship to reproductive efficiency for wild plants is complicated by seasonal changes in seed dormancy, limitations on reproduction imposed by resource scarcity and variability in environmental conditions leading to intermittent reproductive failure. I addressed this complex relationship by

developing a simple model of an equilibrium seed bank, determining the per-plant yields required to maintain population equilibrium in the presence or absence of seed dormancy, and calculating predicted population levels for limiting and variable environmental conditions.

Figure 2-3: Effect of Intermittently Catastrophic Environmental Conditions



Comparison of high and low-germination scenarios under intermittently catastrophic environmental conditions. Transient seed bank equilibrium yields for all scenarios were identical. Heavy solid line represents the number of seeds in the seed bank at year 0. For off-season scenarios, the per-plant yield was 400 (below that for the on-season), germination was 5% and establishment was 5%, with a limiting value of 5000. Germination frequency coefficient is given first in graph legend. Fractions in top line of X-axis label indicate the frequency at which catastrophic reproduction failure occurs.

It is clear from this work that both seed germination delay and seed longevity are essential components of adaptation to non-permissive environments. Seed longevity allows the persistence of a majority of seeds produced in any one year for multiple generations, resulting in a build-up of reserves which compensates for intermittent reproductive failure. Germination delay facilitates survival in these environments by preventing excess vegetative growth that would result in fewer seeds in the seed bank. As suggested by Figure 2-3, both germination delay and seed longevity are required for robust survival in allopatric environments.

The role of off-season germination proved to be especially important in protecting short-lived seed populations from the costs imposed by intermittent in-season reproductive failure. Indeed, if a limited number of seeds can germinate and establish successfully in the off-season, the effect of a one-in-fifteen year reproductive failure rate can be completely mitigated, even in high-germination situations. However, off-season establishment and reproduction must occur *every* year to achieve this effect when the frequency of reproductive failure increases to 1 in five years, making it likely that sub-species that exhibit off-season germination will exhibit physiological differences consistent with growth in the unfavorable times of year.

The data suggest three hallmarks of adaptation of seeds that persist in soil seed banks. For seeds growing in restrictive environments, there should be evidence of both increased longevity and extremely delayed germination. For shorter-lived seeds, some evidence of capability for off-season growth should be detectable. For these seeds, the ability to germinate opportunistically becomes paramount because the seed bank by itself cannot maintain sufficient reserves to guard against environmental hazards except at very high yields. Thus, the effectiveness of the opportunistic germination scenario depends on the ability to

grow under a wide range of conditions, as well as the ability to germinate synchronously.

The program developed for this work has the potential to provide a flexible environment for the analysis of seed bank function in situations where actual assessment of the seed bank is impossible due to seed size. By relating seed bank strategy to fitness parameters and germination strategies that can be measured, the program will greatly facilitate the detection of genes involved at all stages of seed life history, particularly those whose phenotypes are subtle or difficult to detect experimentally.

With the exception of seed longevity, all of the traits suggested by these simulations can be tested in the laboratory. The detection of germination strategies can be accomplished simply by carrying out germination tests under controlled environmental conditions using the same panel of accessions in each experiment. Once the "norm of reaction" of the desired accessions has been determined for the environmental conditions to be used, a single time point can be measured; the germination frequency at this time point can then be considered to be reflective of overall germination efficiency. The result is an extremely simple, high-throughput assay that yields readily to the rigorous demands of quantitative trait locus (QTL) analysis while preserving a connection between environment and phenotype.

CHAPTER 3

Accession-Specific Effects in Seed Germination of Four *Arabidopsis* Accessions

ABSTRACT

Models of soil seed bank dynamics suggest that long-term germination delay is adaptive for both short- and long-lived seeds. Statistical analysis of data drawn from multiple independent experiments revealed that accession-specific germination delay was most evident at 48 hours of incubation, with an overall effect size of nearly 40%. The strongest positive correlation between 48-hour germination frequency and accession identity was for the European accession Col-O, while the strongest negative correlation was for Cvi. Col-O, L*er*, and No-O seeds reached germination frequencies of >80% by 31 weeks post-harvest (PH), while Cvi remained at <50% germination after 60 weeks. Cvi seeds responded to stratification (moist chilling at 4°C in darkness) in an extremely linear manner with an r^2 value greater than 0.85 in three independent experiments. When the stratification treatment was applied after 96 hours of incubation, stratified seeds reached completion, where non-stratified seeds reached only 25% germination. The two related European accessions (Col-O and L*er*) exhibited reduced germination after 8d. of treatment at 37°C, 99% relative humidity (RH), with values of 10% and 2% respectively. Both No-O (the genetically distinct European accession) and Cvi (from the Cape Verde Isles) reached >60% of control under this treatment. Logistic-regression analysis on the association between accession identity, germination frequency and season of harvest (as measured

by the number of days from harvest to the nearest winter solstice) reveal a sizable "Season" effect in No-O, with a positive correlation of 0.2 at 48 hours. This effect is reciprocal, with No-O X Cvi F_1 hybrids exhibiting positive correlation with summer harvest dates while their reciprocal hybrids exhibit the Cvi phenotype of positive correlation with seed age. These results suggest that genetic control of germination is modulated by seed age, seed storage conditions, and maternal growth environment.

INTRODUCTION

The maintenance of a soil "seed bank" consisting of quiescent, ungerminated seeds is an essential component of survival strategy in seed-bearing plants. Seed banks facilitate species survival by encouraging seedling establishment under optimal environmental conditions (Allen and Meyer 1998), by providing replacement of the parental population after disasters such as fire or drought (Venable et al. 1995; Allen and Meyer 1998; Venable et al. 1998) and by preserving genetic diversity (McCue and Holtsford 1998; Mahy et al. 1999). The ability to maintain a persistent seed bank is fundamentally dependent upon the relationship between germination efficiency and reproductive success.

Events throughout a seed's life history can alter its ability to respond to germination cues. In particular, seed age (Bauer et al. 1998), maternal growth environment (Derkx and Karssen 1993), and burial/storage conditions (Botto et al. 1998) have all been shown to play critical roles in determining germination efficiency. The variety of signals that modulate germination behavior in both winter and summer annuals implies a large degree of phenotypic plasticity in this response.

Like other annual plants, the model winter annual *Arabidopsis thaliana* exhibits both (a) accession-specific and (b) environmentally-driven variability in germination efficiency. Seeds of the accession Landsberg *erecta (Ler)* show differences in germination efficiency in response to seed age, season of harvest, and seed storage conditions (Derkx and Karssen 1993; Derkx and Karssen 1994). The temperature- and light-dependent components of variation in germination rate of this accession have been shown to be related to gibberellin sensitivity and production, respectively (Derkx and Karssen 1993).

The relationship between seed age and germination efficiency has also been studied extensively in *Arabidopsis.* Typically, the effect of seed age on germination rate is examined in terms of the "after-ripening" requirement—that is, the time it takes freshly-harvested seeds stored under warm, dry conditions (see Table 1-2) to reach maximum germination efficiency (Christensen et al. 1996; Bradford 1997; Cheng and Bradford 1999; Allen et al. 2000). Accession-specific variation in the after-ripening requirement has been well-documented (Koornneef and Karssen 1994) and has been shown to be due to a complex array of quantitative-trait loci in L*er*/Col (van der Schaar et al. 1997; Malmberg et al. 2005) and L*er*/Cvi recombinant inbred lines (Alonso-Blanco et al. 2003).

Although after-ripening studies have culminated in the positional cloning of a major locus associated with this phenomenon, identification of the genetic component of long-term germination delay response remains elusive (Leon-Kloosterziel et al. 1996; Parcy et al. 1997; Finkelstein et al. 1998; Debeaujon et al. 2000; Papi et al. 2000; Russell et al. 2000). The failure to detect any single locus associated with long-term germination delay is likely due to (a) the length of the experiments required to detect long-term delay phenotypes, (b) the difficulties

posed by extreme variability in these responses, (b) and the binary nature of the phenotypic data required.

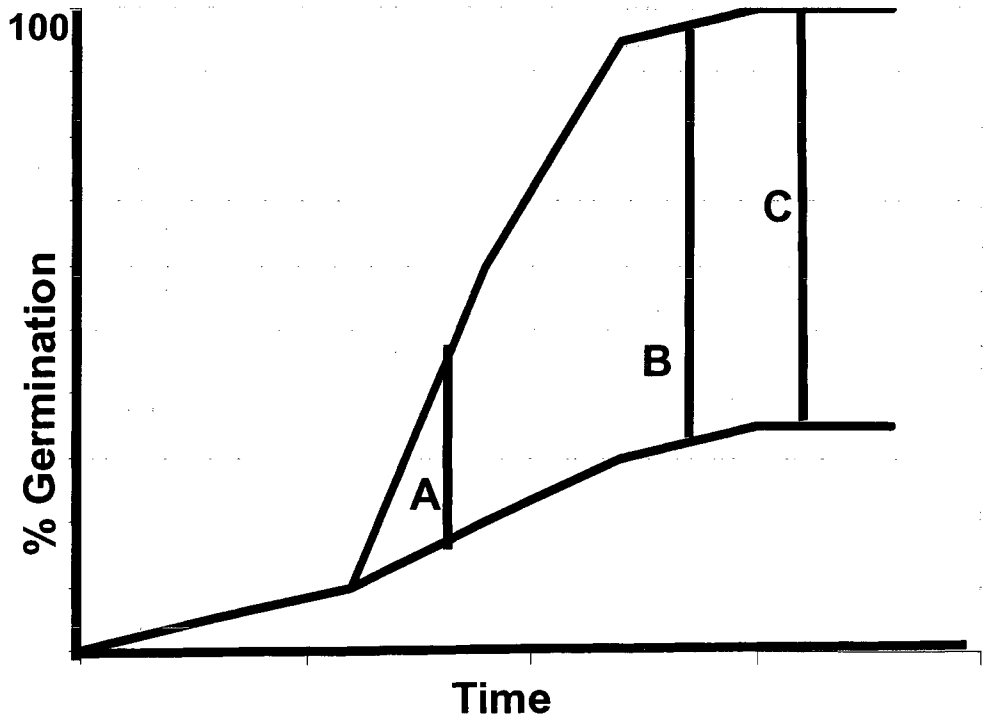
The three difficulties mentioned above--experiment efficiency, phenotypic variability, and binary data--can be addressed by a two-phase experimental approach. In the first phase, a maximally-informative time point for germination frequency is chosen on the basis of effect size and percent completion. In the second phase, a "norm of reaction" in the parental accessions can be used to determine the environments which yield the largest accession-specific variation (Lynch and Walsh 1998). The binary nature of germination data can be addressed by employing logistic-regression methods for estimation of variance.

The use of binary data drawn from a single time-point in a germination time course has several advantages. One can improve experimental efficiency by selecting a standard germination medium that optimizes germination efficiency while preserving variability in germination response. Germination delay can be distinguished from dormancy by mapping response on a maximally-informative time point in the germination time course and then allowing the seeds to proceed to completion. Population totals are then normalized to remove ungerminated seeds from consideration.

The choice of a mapping time point cannot be arbitrary. In cases where one lineage proceeds to completion too quickly to allow for estimation of the time to 50% completion, a two-pronged approach can be employed in choosing the most informative point. Examination of individual time-course curves reveals two potential time points where both lineages are informative (Figure 3-1). At Point A, neither lineage is complete, but the difference between them is small. At point B, the effect size is large, but one lineage is near completion. At Point C, the effect size is large, but one line is complete, rendering that line uninformative because

one can no longer tell whether there might be delay alleles present in the completed line. The goal is to identify the point where the effect size is largest, but neither line is complete. Thus, the largest effect size where neither line is complete would appear to be the best choice. The choice among these points can be resolved by subjecting time-point data from multiple independent experiments to logistic-regression analysis and maximum-likelihood estimation. The relevant environmental conditions can be treated in three ways: (1) they can be randomized, (2) they can be tracked, but not controlled, or (3) they can be controlled. The decision about which of these three strategies to pursue directly affects statistical design and QTL analysis

Figure 3-1: Effect Size at Single Time Points



Strategy for identifying appropriate time points for mapping of germination phenotypes. Solid lines represent effect sizes between two genetically distinct lineages.

In the present work, I examined the role of seed age, seed storage conditions, and maternal growth environment (as determined by the season of harvest) on seed germination in four commonly-used *Arabidopsis* accessions. Three of the four accessions (Col-O, L*er* and No-O) were collected in a temperate climate (Europe). Two of the three European accessions (Col-O and L*er*) were collected in a single sample at Landsberg, in southern Germany, while the third (No-O) was collected from Nossen, also in southern Germany. The fourth (Cvi) was collected from the dry tropical environment of the Cape Verde Isles. The results presented here highlight the similarities and the differences among the four accessions. I report that one accession (Col-O) exhibits strongly synchronous germination under most environmental conditions tested, while L*er* consistently lags behind Col-O by a minimum of 24 hours. Cvi is the most germination-delayed of all accessions. Interestingly, No-O is capable of both strongly synchronous germination (with a time course resembling that of Col-O) and delayed germination (with a value at 48 hours less than L*er* but greater than Cvi). The two closely-related European accessions (Col-O and L*er*) are both sensitive to elevated temperature/humidity pre-treatments, while the remaining two (No-O and Cvi) are relatively resistant. Together these results suggest that there are accession-specific aspects to all three of the variables which contribute to both short- and long-term germination regulation.

MATERIALS AND METHODS

Plant growth and handling

Seed Stock: Seeds from the four accessions Columbia (Col-O), Landsberg *erecta (* L*er),* Cape Verde Isles (Cvi), and Nossen (No-O), were initially obtained

either from the *Arabidopsis* Stock Center (Columbus, Ohio, USA) or from Lehle Seeds (Round Rock, Texas, USA). For Lehle seed stock used in germination tests, harvest dates and growth conditions were obtained from Lehle Seeds. As with my seeds, Lehle seed stocks were produced in a temperature-regulated greenhouse. Seeds used in all experiments of this study were produced in a greenhouse, in which ambient light was supplemented from October to May to simulate a 14 hour light, 10 hour darkness cycle. Plants were randomized across benches to compensate for slight variations in fluence rate. Seeds were sown on water-saturated soil (SunShine potting mixture) and thinned after 1.5-2 weeks to 1 - 3 plants per pot. To isolate plants and promote uniform yield, once inflorescence stalks were well-established, individual plants were either staked, placed in perforated plastic bags, or isolated using Arasystem™ cones and tubes (Arasystem, Inc., Belgium). Seeds were harvested when a majority of the siliques were golden-brown.

Surface sterilization and seed sowing: 50-200 seeds of each accession were placed in 1.5 ml plastic microcentrifuge tubes for surface sterilization. Between 750 and 900 μL of a 30% $HClO_4$ solution (made in water) were added to each tube. The tubes were shaken at 20°C for 15 minutes with reciprocal agitation. At the end of the agitation period, seeds were rinsed five times with 1000 μL of sterile double-distilled water. The tubes were filled to 1.5 mL with a solution of 0.1% Phyta-Agar (Sigma Chemical Co., Munich, Germany) made with double-distilled sterile water. The resultant suspension ensured even spacing of seeds when sown onto plates with a pipeter. Seeds were sown on ½ strength Murashige and Skoog basal salt mixture (Sigma Chemical Co., Munich, Germany) made with double-distilled water (Murashige and Skoog 1962) and 0.8% PhytaGel (Sigma Chemical Co., Munich, Germany). The 10-cm square, 6

by 6 gridded plastic plates (Phenix Research Products, Chandler, North Carolina) were sealed with gas-permeable paper surgical tape (3M Corporation, Maplewood, Minnesota) to prevent desiccation.

Germination incubation conditions: All germination tests were performed with seeds plated on a sterile medium, under continuous light. Plates were placed either in a constant-temperature chamber (Model CMP4030, Conviron, Winnipeg, Canada) maintained at 21°C or in temperature-controlled growth chambers with temperature ranges of 17-20°C or 21-24°C. All studies took place under light intensities of between 20 and 100 μE. At regular intervals during the germination incubation period, seeds were observed using a Leica MZ12 stereomicroscope equipped with a dark-field base at 40x magnification. At the beginning of each experiment, the number of seeds on each plate was counted to obtain a plate total. As the experiments progressed, seeds with emerged radicles were counted as germinated, and scored as a percentage of the original total, yielding a "germination frequency". Plates were scored every 12-24 hours until germination reached completion.

Experimental Design and Treatment Protocols

Stratification: Stratification treatment was carried out by placing seeds, sown on agar (as previously described), into the dark at 4° C. To generate a standard curve for Cvi stratification response, a total of 15 plates were sown concurrently and then treated for 0 to 96 hours at 4°C in darkness before being placed in incubation conditions. Each time treatment was carried out in triplicate. Plates were scored once every 24 hours upon being placed in incubation conditions. For the stratification rescue protocol, non-stratified Cvi seeds were sown as above. 96 hours into the incubation period, a subset of plates was incubated at 4°C in

darkness for 24 hours. At the end of this treatment, seeds were returned to germination incubation conditions, and counting continued for another 72 hours, at which point seeds showed no further increase in germination.

Seed storage tests: For seed storage tests, 300-400 dry seeds were placed in 1.5 ml microcentrifuge tubes. Tubes were then placed horizontally into covered Pyrex dishes with the tube caps open. In one of the Pyrex dishes, a small tray containing 100 mL of sterile double-distilled water was added to produce saturating humidity. In the "humidity" Pyrex dish, tubes were placed beside the tray containing double-distilled water. The dishes were placed side by side in a 37°C incubator where they remained for 8 days. Temperatures and relative humidity inside the Pyrex dishes were measured with an archival min/max thermometer/hydrometer (Light Impressions, Portland, OR).

Reciprocal crosses: Crosses were preferentially performed on newly bolted plants using buds 1-2 days from opening. First, any open flowers on the inflorescence were removed, and then the largest 3-6 buds were dissected to remove all sepals, petals, and stamens without allowing immature stamens to contact the stigma. Dehiscing stamens from the pollen parent were promptly dabbed onto the stigma of the dissected bud. Crossed plants were labeled and spaced well apart to prevent accidental pollen transfer to the naked pistil, though experience suggests that the stigma is no longer receptive after several hours. For each set of hybrids, a set of self-fertilized plants was grown in identical pots on the same greenhouse bench, but isolated in order to prevent contamination of crossed plants.

Statistical Analysis and Data Management

Seed inventory tracking and maintenance: Each individual seed lot was assigned a unique inventory number consisting of the year and day of harvest followed by a sequential number reflecting the order in which the seeds were harvested. Pooled seed lots were created either by growing more than one plant per pot, collecting seeds from multiple plants simultaneously, or by mixing seeds from multiple lots of the same planting into experimental tubes. After harvest, seeds were stored in envelopes for between 1 and 15 weeks at room temperature (between 18 and 24°C) before being transferred to microcentrifuge tubes for permanent storage in a 4°C refrigerator. Both date of harvest and date of transfer to refrigerator were noted.

Evaluation of seed handling and treatment variables: Values for the seed handling variables used in this analysis were determined as follows: (a) "Season": the difference between the nearest winter solstice and the harvest day (only values of <182 were used); (b) "Transfer": the difference between the harvest date and the date of transfer to refrigerator storage; (c) "Age": the difference between the experimental date and the harvest date; (d) "Temperature": average incubation temperature for each experiment (the temperature range for all experiments was between 17 to 25° C); (e) "Storage": the treatment length in days (for details, see "Seed Storage Tests" above); (f) "Lot": the inventory number for each individually harvested batch of seeds; and (g) "Accession": the name of the accession/ecotype as supplied by the *Arabidopsis* Stock Center (Columbus, Ohio, USA) or Lehle Seeds (Round Rock, Texas, USA).

Statistical Analysis: Logistic regression analysis was used to assess the relationship between germination frequency and the covariates season of harvest, transfer, age, induction temperature, storage, lot, and accession. For full

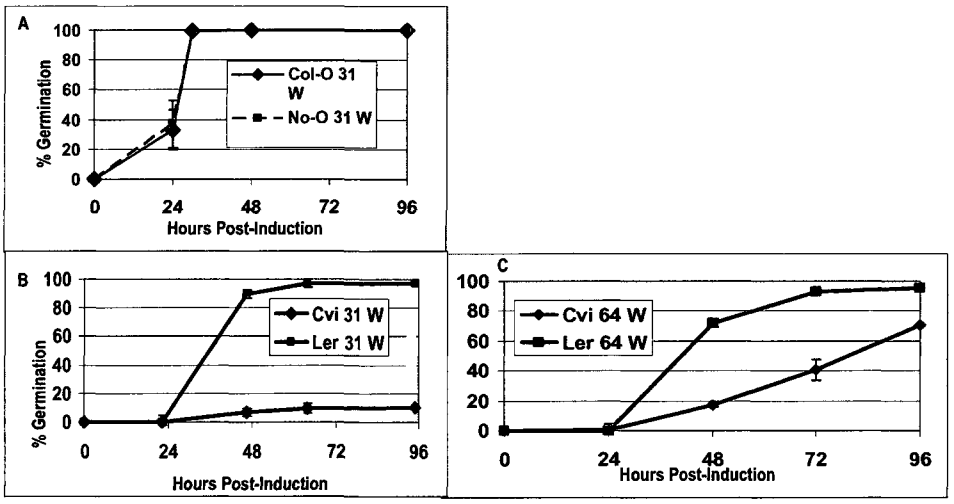
data, analysis both considering and ignoring the storage type and lot used was carried out. For both full data and accession datasets, the covariate "Lot," which refers to individually numbered seed lots (see Inventory section, above) was nested within the covariate "Accession". When considering the full data, "Accession" was used as one of the covariates. Computation was done using the R package ("glm" function with "binomial" family) (McCullaugh and Nelder 1989).

RESULTS

Variation in germination frequency within the germination time-course is accession-specific. The role of accession-specific variation in germination control is well-established (Bradbeer 1988; Bewley 1997). To determine whether the germination efficiency of *Arabidopsis* seeds is accession-specific, I examined the profiles of three European lineages (Col-O, L*er*, and No-O) along with the Cvi (Cape Verde Isles) lineage. The germination profiles of the three European accessions were very similar with respect to final percentage of germination (Figures 3-2A and B). However, at 24 and 48 hours, the three accessions exhibited obvious differences in germination frequency. L*er* consistently had the lower 24-hour germination percentage of the two related European accessions (L*er* and Col-O), while that of the genetically distinct accession, No-O, was variable (see Figure 3-4). Though Col-O exhibited modest germination at 24 hours, germination was consistently complete at 48 hours. L*er* germination values at 48 hours were between 80 and 95%. In contrast to the three European accessions, Cvi seeds routinely took more than 72 hours to reach completion (Figures 3-2B and C). To determine how germination of L*er* and Cvi seeds changed over time within a single rotation, I tested the same rotations used in

Figure 3-2B at 64 weeks post-harvest (Figure 3-2C). L*er* and Cvi seeds attain maximum germination of 100% and 70%, respectively, by 96 hrs. (Figure 3-2C).

Figure 3-2: Incubation Profiles of the Four Accessions



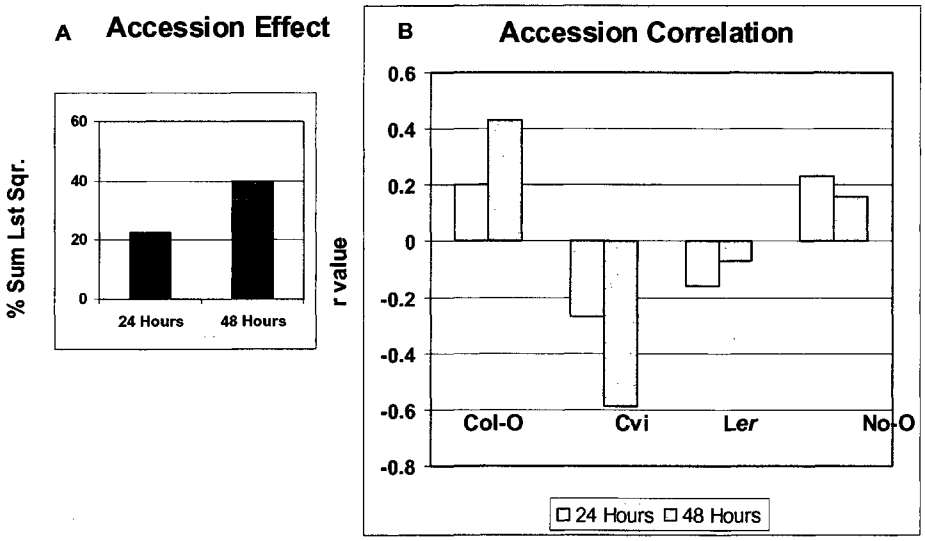
Germination incubation profile of Col-O and No-O at 31 weeks post harvest. Error bars represent standard deviation of three replicates of 100-150 seeds each. Values for standard deviation were from 0 to 13. **B** and **C:** Incubation profiles of L*er* and Cvi at 31 weeks **(B)** and 64 weeks **(C)** post-harvest. Error bars on **B** represent the standard deviation of six replicates. Error bars on **C** represent the standard deviation of three replicates N on all replicates is 100-120 seeds each. Standard deviations from 0.44 to 7.1.

Statistical analysis of data drawn from multiple independent experiments revealed that accession-specific germination delay was most evident at 48 hours of imbibition. I analyzed the data from 40 independent experiments using maximum-likelihood estimation on a logistic-regression analysis. At 48 hours of incubation, 35% of the variance is explained by accession alone (Figure 3-3A). I examined the correlation between germination frequency and each accession individually as a measure of the ability of a specific genotype to determine germination success. The strongest positive correlation is between the Col-O accession and 48-hour germination frequency, while Cvi at 48 hours had the

strongest negative correlation (Figure 3-3B). These data suggest that a major germination effect segregates in the Cvi population. Additionally, there is a slight but detectible delay effect associated with the L*er* accession. Like Col-O, No-O showed a positive correlation with germination percentage at 48 hours. The strong positive correlation between the Columbia phenotype and 48-hour germination frequency allows the succession to be used as the operational standard for the fastest possible germination time course.

There are accession-specific components to age-related germination behavior. To examine the relationship between seed age and germination frequency, multiple rotations of each accession were tested at 6 to 30 weeks post-harvest (Col-O and No-O) and 14 to 80 weeks post harvest (L*er* and Cvi) (Figure 3-4). At 6 weeks post-harvest, Col-O 48-hour germination values

Figure 3-3: Accession-Specific Effect and Correlation



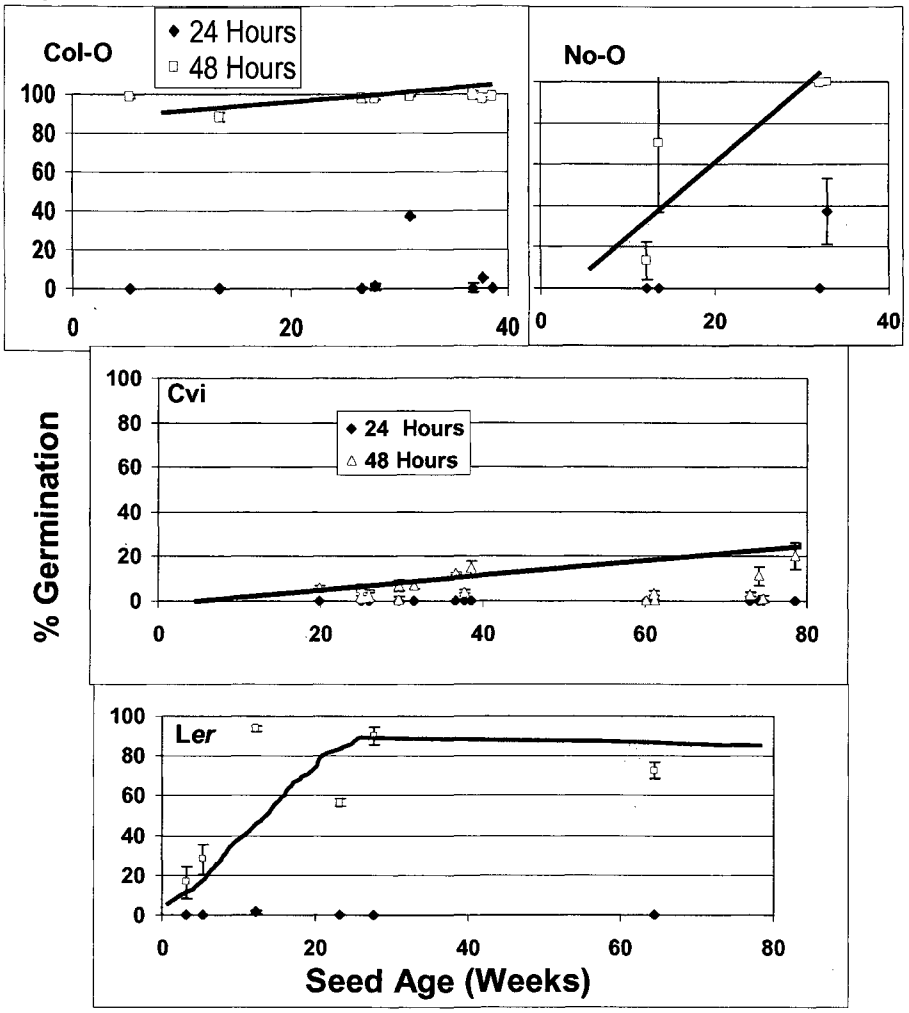
A: Effect of accession on germination frequency at 24 and 48 hours expressed as a percent of the sum of least squares on a logistic regression. **B**: Correlation between each accession and the germination frequency at 24 and 48 hours.

ranged from 95-100%. In contrast, Cvi germination measured at 48 hours continuously increased to 20% by 80 weeks, while L*er* germination increased until 31 weeks post-harvest and then leveled off (Figure 3-4). No-O at 12 weeks showed 15% germination for one rotation, and 70% for another of equal age, suggesting that additional cues contribute to the germination response of this accession (Figure 3-4). Because these data represent multiple experiments across varied incubation environments, the changes observed are likely due to a genetically-based difference in germination behavior over time. These data suggest: first, that there are age-dependent components to germination behavior; second, that two of the four accessions (Cvi and L*er*) are associated with germination delay (See Table 1-2); and third, that a major age-independent germination response is present in No-O. Of these three effects, the most striking is the high degree of germination delay present in the Cvi accession.

Cvi shows a positive response to moist chilling (stratification) both prior to incubation and at 96 hours of incubation. The high degree of germination delay present in the Cvi accession is suggestive of a strong dormancy potential and an increased requirement for dormancy release (Figure 3-5). A classic signal for dormancy release in winter annuals is a brief period of moist chilling (stratification) (Allen and Meyer 1998). In the laboratory, the stratification treatment is normally carried out at 4°C in darkness (Koornneef and Karssen 1994). Discovery of a Cvi response to stratification would allow for experimental synchronization of Cvi seeds to their much less dormant counterparts (L*er*, Col-O and No-O). Two distinct types of stratification treatments are potentially informative. The first is the establishment of a "standard curve" for stratification-induced dormancy release in which the treatment time required to restore full germination capability is determined. The second stratification treatment is

designed to distinguish seeds capable of germination from those rendered non-viable due to handling or other environmental factors. In this second

Figure 3-4: Effect of Seed Age on Germination



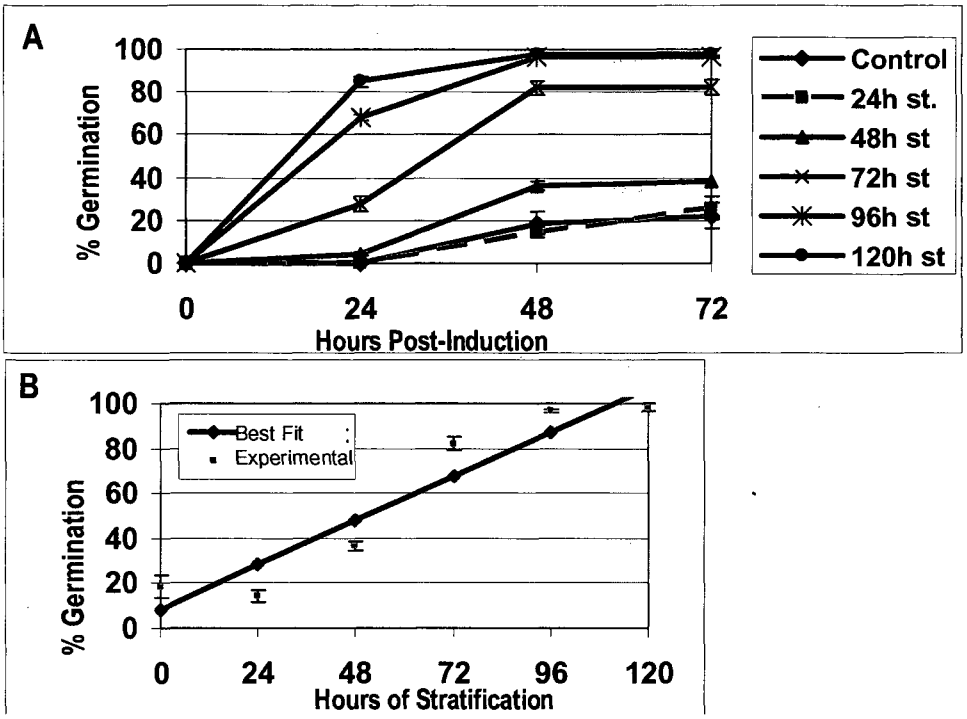
Age profiles of germination in four *Arabidopsis* accessions: Col-O, No-O, L*er,* and Cvi. Error bars represent standard deviation from the mean of between three and six replicates of 100-120 seeds each. Lines represent "best-fit" curves for the data sets. 24-hour and 48-hour time points are represented.

stratification rescue protocol, the cold treatment is applied to Cvi at a time when seeds of L*er*, Col-O and No-O would normally be expected to have completed germination. The portion of Cvi seeds which successfully germinate after the

application of a stratification pulse then becomes the viable population total. Experimental results can be normalized to this total, thus excluding any non-viable seeds that are present.

To determine whether Cvi seeds respond in a predictable way to stratification prior to incubation, I subjected them to a 4°C treatment for 0 to 96 hours. Increasing the duration of stratification treatment augmented germination. When no stratification treatment was administered, Cvi seeds reached 20% germination in 72 hours. In contrast, complete germination was achieved after only 48 hours of incubation when seeds received a 96-hour stratification treatment (Figure 3-5A). Response to pre-incubation chilling treatment is strikingly linear, with r^2 values above 0.85 in all rotations tested (Figure 3-5B). The solid line represents the best-fit curve when the data are fitted to a linear model. The resulting data provides a useful standard curve for the determination of appropriate stratification treatment length for seeds in experiments involving the Cvi accession.

Figure 3-5: Effect of Stratification Treatment

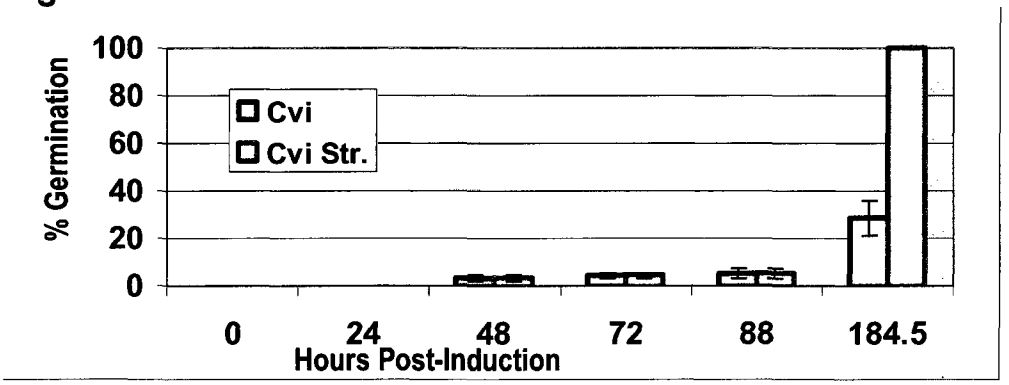


A: Family of germination time-course curves for Cvi resulting from increasing lengths of moist chilling (stratification). Each line represents three replicates of 80-120 seeds. Error bars represent standard deviation from the mean for three replicates. Seeds were 46 weeks old at time of experiment. **B:** Correlation between hours of stratification and germination frequency at 48 hours of incubation for curves in **A**. Solid line represents best-fit curve from linear regression ($r^2=0.85$). Error bars on "experimental" points represent standard deviation from the experimental mean.

To determine whether Cvi seeds would respond to a stratification rescue protocol, I subjected non-stratified Cvi seeds to a 24-hour 4° C stratification treatment after 88 hours of germination incubation. The 88-hour time point was chosen because in other experiments this time point coincided with a marked increase in germination efficiency. I wished to determine whether stratification was still effective in the context of this increased response. As shown in Figure 3-6, stratified seeds reached complete germination 72 hours after being returned to the light, while their non-stratified counterparts remained at 30% germination.

These data demonstrate that the use of a stratification rescue treatment is an effective experimental technique for distinguishing between viable and non-viable seeds.

Figure 3-6: Stratification Rescue in 41 Week-old Cvi Seeds



Initial points represent three replicates. One randomly-chosen replicate was placed in a 4° C refrigerator for 24 hours after 96 hours of germination. The other two seed samples were maintained in the growth chamber (20°C). Error bars on 0 to 88 hour points represent standard deviation from the mean of three replicates; error bars on the "Cvi" line represent the standard deviation from the mean of the two plates that remained in the growth chamber.

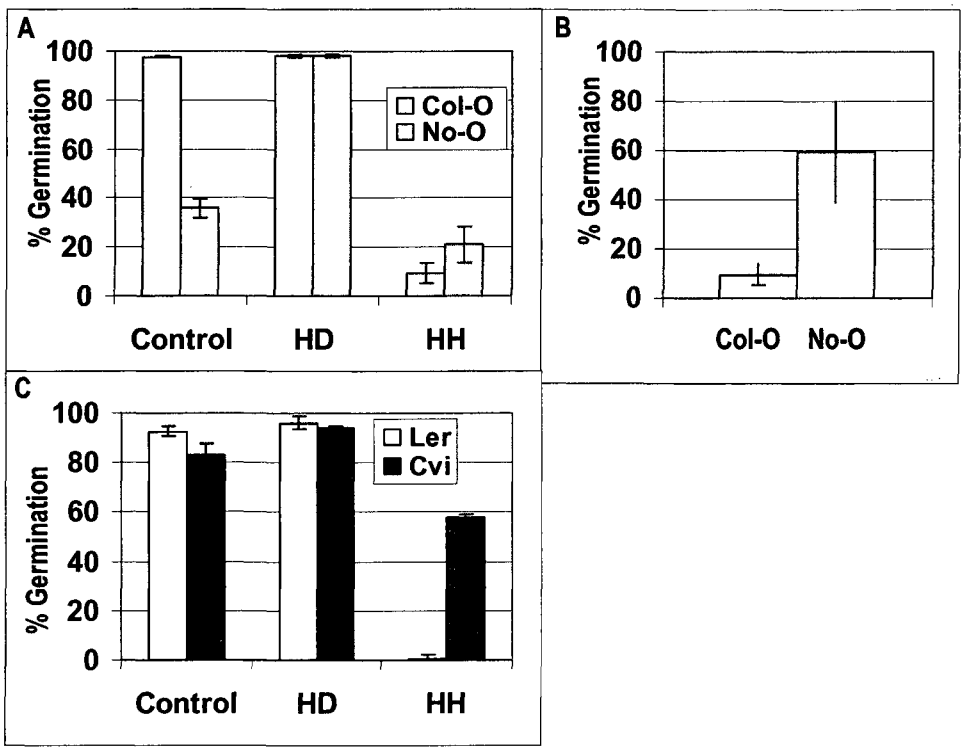
Two of the four accessions exhibit tolerance to brief periods of storage stress. Seeds in soil seed banks experience a wide range of temperature and moisture levels which impact their long term viability. The ability of seeds to withstand elevated temperatures may be adaptive, particularly in a tropical habitat. Furthermore, pre-treatments at elevated temperature and relative humidity (RH) are used as a standard test of seed vigor because of the sensitivity of seeds to these conditions (Hampton and TeKrony 1995). To test whether there was an accession-specific variation in response to storage stress pre-treatments (Figures 3-7A to C), I subjected seeds of all four accessions to 8-day pre-treatments at 37°C 20% RH (hot and dry, HD) or 37°C 99% RH (hot and humid, HH). For Col-O and No-O, a 24-hour stratification treatment was administered just prior to incubation to remove any residual dormancy (as described by Derkx and Karssen, 1993b). This stratification treatment was increased to 72 hours for

L*er* and Cvi, to compensate for the substantial dormancy present in both accessions at 4 weeks post-harvest. Seeds were then returned to normal incubation temperature of approximately 22°C, and their germination scored after 96 hours.

Results show (Figure 3-7A&B) that only the Col-O control reached completion at 96 hours of incubation in these experiments. No-O attained 35%, L*er* 92%, and Cvi 83% germination under control conditions. Because the Col-O control went to completion, it was not possible to determine whether the HD treatment had a positive effect on Col-O germination. The HD treatment improved the efficiency of L*er* germination by only 8%, while Cvi showed 13% improvement. By contrast, the HD treatment improved No-O germination by >60%.

With respect to the HH treatment, Col-O and No-O had distinctly different germination responses. Col-O achieved slightly more than 10% of the control (Figure 3-7B) and No-O was 60% of control. L*er* germination was essentially eliminated, and Cvi germination reached approximately 75% of control (Figs. 3-7A-7C). The ability of Cvi to withstand the HH treatment is not surprising, given the latitude and climate of its tropical habitat. However, the finding that No-O is somewhat resistant to HH was unexpected and suggests that the heat-tolerant phenotype may not be entirely climate specific.

Figure 3-7: Storage Stress Response



Effect of 8 days of pretreatment at 37°C 20% RH (HD) and 37°C, 99% RH (HH) on Col-O, No-O, L*er* and Cvi. Bars represent two replicates; error bars are standard deviation from the mean. These are 96-hour values. **A:** Col-O and No-O, incubated at 21-24°C. Seeds were 31 weeks old, and were stratified for 24 hours. **B:** Col-O and No-O HH treatment represented as % of control. **C:** L*er* and Cvi incubated at 21-24°C for 96 hours. Seeds were 4 weeks old and were stratified for 72 hours.

Arabidopsis seeds respond to age and maternal growth environment (harvest season) in an accession-specific manner. Both seed age and maturation season play important roles in plant population maintenance (Fenner 1985; Allen and Meyer 1998). However, these two variables are difficult to resolve experimentally because if "season" is matched in a single experiment, the seeds are by definition close to a year apart in age, e.g. seeds harvested in August 2003 and August 2004 would be matched in season and a year apart in age. While season alone can be matched in a single experiment, it is impossible to match age alone (e.g., for an experiment conducted in October, all 90-day old seeds would have had to have been harvested in June, thus matching "season" by definition). Accordingly, I wished to clarify the role of "Age" and "Season" by examining the results of multiple experiments in a two-step approach using maximum-likelihood estimation. I first wished to clarify the relative importance of two major control variables: "Temperature" and "Transfer". The Temperature variable represents the average temperature of the three different incubation regimes used. For Temperature, No-O at 48 hours had the largest effect with a sizable negative correlation between temperature and germination frequency (Table 3-1). The effect size for the three other accessions was less than 10%. With the exception of No-O, all of the correlations for temperature were less than 0.12.

The "Transfer" variable represents the number of days seeds were stored at room temperature before being transferred to the 4°C refrigerator for permanent storage. This variable is therefore a reasonable representation of an "after-ripening" treatment, though no attempt was made to determine the progression of germination through these periods. As expected, Cvi has the greatest positive correlation, but the effect size is small (Table 3-1)

Table 3-1: Values for Handling and Control Variables

Effect is the percent of deviance explained by the variable, correlation is the Spearman correlation (r) value. Time: Time of incubation. NS: Not significant.

VAR	Temperature				Transfer			
	Effect		Correlation		Effect		Correlation	
Time	24 h.	48 h.	24 h.	48 h.	24 h.	48 h.	24 h.	48 h.
All	0.21	0.34	0.19	0.25	0.82	4.37	-0.03	0.05
Col-O	NS	6.52	NS	-0.07	NS	3.45	0.22	-0.26
Cvi	4.16	2.11	-0.09	0.07	2.77	4.73	0.51	0.47
Ler	NS	2.11	0.09	0.11	5.26	1.68	-0.04	0.32
No-O	4.13	19.05	0.01	-0.27	4.7	1.15	0.37	-0.28

To examine the role of the seed age variable within a germination time-course, I assayed between 5 and 10 independent rotations of each accession under the conditions outlined in Materials and Methods (for details, see Appendix 1).

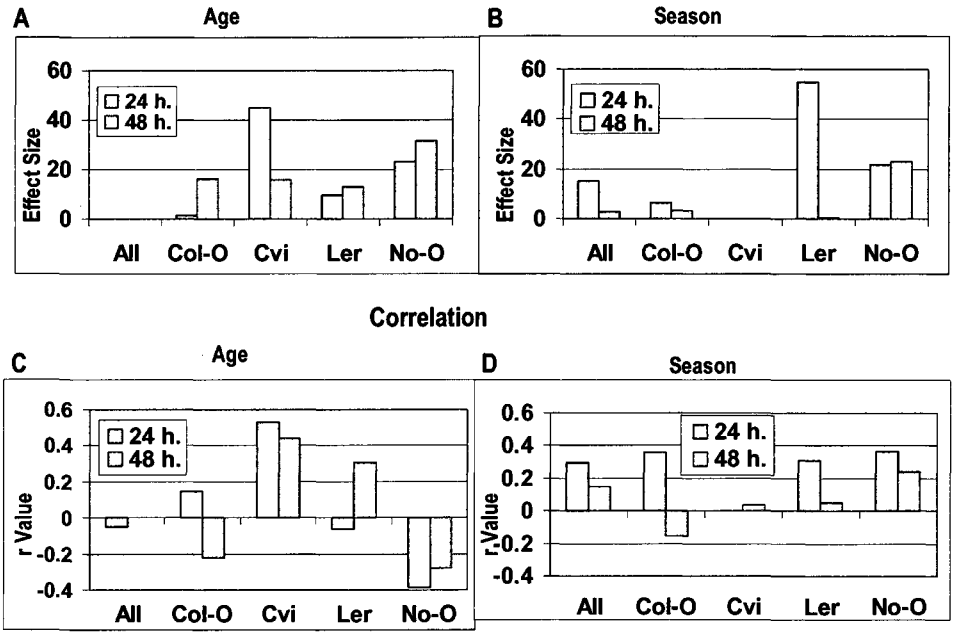
When all accessions are taken together, the age variable is not detectable (Figures 3-8A and C). The absence of a detectable global age effect might suggest that there is no age effect in these accessions. However, the presence of a substantial accession-specific age effect (Fig. 3-8A) and the bi-directionality of age correlation (Fig. 3-8C) among the four accessions reveal a set of effects which cancel each other out globally. The correlation between seed age and germination frequency is greatest in Cvi at 24 hours (Figure 3-8C). At 48 hours, there is still a substantial positive r-value for Cvi. The magnitude of the age variable in *Ler* at 24 and 48 hours is modest (15% and 18%, respectively) (Figure 3-8A). At 24 hours, the correlation between age and germination frequency for *Ler* is slightly negative; at 48 hours, the correlation is positive (Figure 3-8C). No-O had the strongest negative relationship between age and germination frequency, with a 24-hour effect size of 21% and a negative correlation of 0.4. At 48 hours, No-O again had a negative correlation between age and germination

frequency, but the degree of correlation was reduced (Figures 3-8A and C). Nevertheless, the age variable explained 35% of the total variance for No-O. The smallest 24-hour age effect was observed in Col-O, with a modest positive correlation. At 48 hours, the effect size increased to 18%, and the correlation was negative. These results suggest that there are a number of accession-specific aspects to the relationship between seed age and germination frequency, particularly at early points in the germination time-course.

I then analyzed the effect of harvest season on germination. When all accessions were considered, there was a modest positive correlation between season and 24-hour germination frequency, with an effect size of 15% (Figures 3-8B and D). For Col-O at 24 hours, the effect size was modest (<10%), but the correlation value for season was substantial (0.35). At 48 hours, the direction of the season correlation changes, but the effect size remains minimal at <5%. A harvest season effect was not detectable in Cvi at either 24 or 48 hours. For L*er* at 24 hours, harvest season alone accounted for 55% of the variance, with a respectable positive correlation of 0.31. At 48 hours, the harvest season variable for L*er* became marginally detectable. Interestingly, No-O showed the greatest 48-hour effect size for harvest season (21%) with a positive correlation of 0.24. Because of the method by which the "Harvest Season" variable was evaluated, a positive correlation is suggestive of more efficient germination in summer-harvested seeds. These data imply that harvest season, as measured by the number of days from the nearest winter solstice, plays a role in germination efficiency. The role appears to be accession-specific. Two of the four accessions (L*er* and No-O) show a positive correlation between the "Season" variable and germination frequency at 24 hours. At 48 hours, only No-O shows a significant correlation between season and germination frequency. Moreover, the No-O

observation encompasses seeds of ages up to 245 weeks (Appendix 1-1) obtained from two different sources.

Figure 3-8: Effect of Season and Age on Germination



Top: Effect size values for age **(A)** and season (**B**). Each bar represents the percent of the deviance. **Bottom:** Correlation coefficients (r-values) for age (**C**) and season (**D**). Each bar represents the r-value. Direction of the value indicates the direction of correlation. Season value was the number of days to the nearest winter solstice.

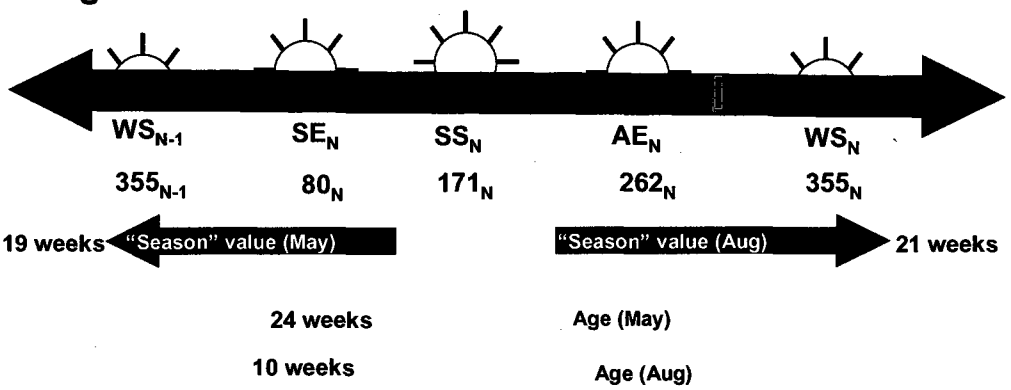
Multiple reciprocal effects are evident in Cvi X No-O F_1 reciprocal hybrids.

The Cvi and No-O accessions differed from each other with respect to germination speed and the effect of harvest "season" on germination efficiency (Figure 3-9). As a preliminary step towards developing a strategy for detecting the genetic components of the "Season" phenotype, I first needed to know whether this phenotype was detectable in age-matched seeds (as opposed to a global analysis). To determine whether "Harvest Season" affected No-O and Col-O germination in a manner independent of age, I compared the germination profile from one winter and one summer rotation for seeds that were aged 13

weeks. Although all four seed lots reached 100% germination by 72 hours, No-O-December lagged Col-O December, as well as No-O August and Col-O August (Figure 3-10A). The difference in the 24-hour germination frequency was significant ($p = 0.015$) in an unpaired t-test assuming unequal variance. By contrast, neither the difference between the August and December Col-O rotations, nor the difference between Col-O August and No-O August rotations was significant ($p>0.1$). These data support the determination made in the logistic regression analysis that No-O has a significant, although variable, "Season" effect.

In experiments involving germination phenotypes, the maternal effect on germination control must be examined due to the influence of the maternally-derived seed coat on germination response (Papi et al. 2000; Russell et al. 2000; Bentsink 2002). To determine whether germination frequency was correlated with the maternal parent, I crossed Cvi, the most delayed accession, reciprocally to No-O, which had relatively fast germination. The relationship between the Julian harvest dates, ages, and "Season" values (as measured in the global statistical analysis) are shown in Figure 3-9.

Figure 3-9: Schematic of Reciprocal Age/Season Experimental Design



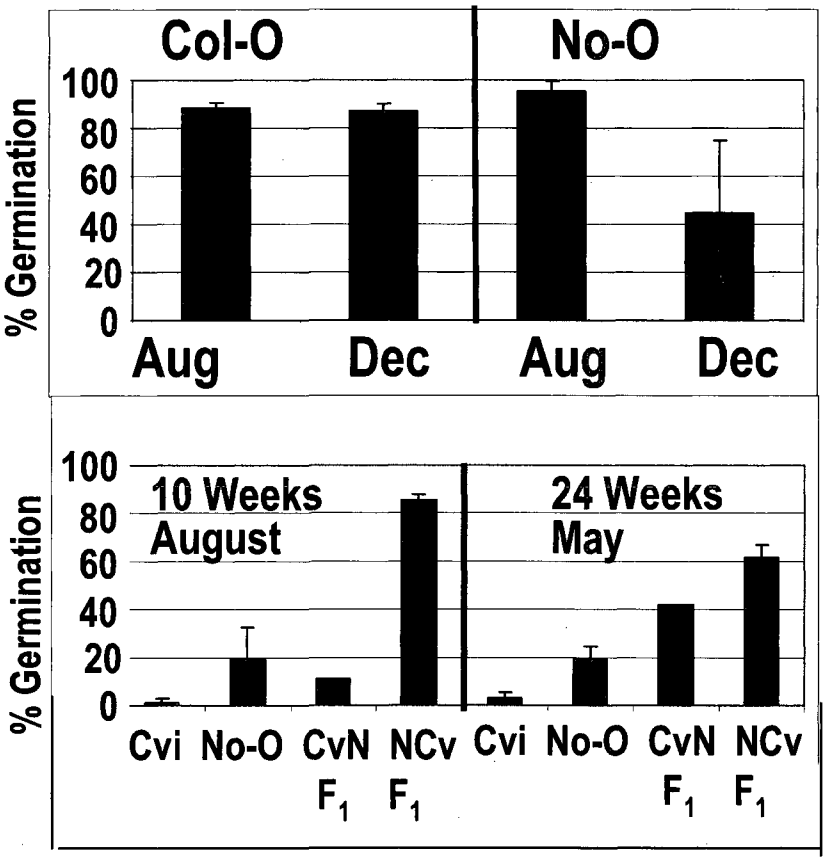
Schematic diagram of the relationship between the age of the seeds, their "season" values (as measured for the statistical analysis) and the harvest day code for the rotations used in the reciprocal-cross experiment. Numbers below year line represent the Julian date of the winter solstice, spring equinox, summer solstice, and autumn equinox. N indicates the year of the experiment. Red on Season line is date of experiment.

The difference between the two parental rotations for this experiment is not significant ($p>0.1$). In both rotations, the germination frequency of the CvN F_1 fell more than two standard deviations below the mean frequency of the reciprocal hybrids (Figure 3-10B), suggesting that germination frequency in both rotations is dependent upon the maternal parent.

Interestingly, the interaction between season, age and germination frequency also appears to be dependent upon the maternal parent. The CvN F_1 germination frequency correlated positively with age, while the NCv F_1 germination frequency correlated positively with the Julian harvest date (an alternative measure of "season"). Strikingly, the difference between the 10-week and the 24-week NCv F_1 populations is significant ($p=0.007$, Figure 3-10B), with the higher value being associated with the younger seeds, suggesting that in No-O borne F_1s, germination frequency correlates positively with harvest season. As with Cvi, the

CvN F1 germination frequency correlates positively with age. Though very preliminary, these data raise the intriguing possibility that harvest dates in August "prime" germination in No-O seeds by means of maternal factors [e.g. Dof zinc-finger transcription factors such as Dag1 (Papi et al. 2000)]. The fact that CvN F1's have an age correlation that matches the Cvi parent is suggestive of embryo-based control for this phenotype which is suppressed by seed-coat mediated factors in NCv F1s.

Figure 3-10: Heritability of "Season" Effects



A: Effect of harvest season on Col-O and No-O on seed germination. P-value for t-test involving unequal variance = 0.015. This time point is 24 hours post-induction; induction was preceded by a 24-hour stratification period. B: Maternal effect phenotype in reciprocal F1 progeny of a CviXNo-O cross. Time point shown is 24 hours of incubation. There was no stratification treatment.

DISCUSSION

In this chapter I examined accession specific variation in three traits predicted to confer an advantage in the maintenance of persistent soil seed banks. These traits were a) long-term germination delay, b) resistance to storage conditions relevant to growth in the unfavorable season, and c) positive correlation with germination frequency for seeds produced in the "off" season.

The experimental result that is the fundamental hallmark of heritability is the finding that accession-specific variation occurs even when environmental variables are tightly controlled. Statistical analysis of data drawn from multiple independent experiments revealed that accession-specific germination delay was most evident at 48 hours of incubation. This finding was consistent for all but two of the populations examined. One Col-O/No-O population reached completion at 48 hours of incubation, but had been stratified for 24 hours previous to the start of incubation. The Cvi X No-O F_1 populations also reached completion at 48 hours; however, the decision to examine the data from the 24-hour time point is consistent with the general strategy shown in Figure 3-1.

Long-term germination delay: The role of seed age in determining germination response has been extensively addressed by other investigators. Seeds lose primary dormancy in a mathematically predictable manner (Bauer et al. 1998) that is functionally related to incubation time, water relations (Bradford 1997) and temperature (Christensen et al. 1996). In *Arabidopsis* primary dormancy loss has a substantial genetic component in Col-O (Malmberg et al. 2005), L*er* (van der Schaar et al. 1997) and Cvi (Alonso-Blanco et al. 2003). My results provide new insight to the seed age vs. germination response question by including observations of germination patterns over periods that exceed the

literature value for *Arabidopsis* after-ripening requirements (Koornneef and Karssen 1994). Long-term delay is particularly evident in Cvi, which had not reached maximum germination efficiency at 60 week. L*er* also shows a modest long-term delay, lagging slightly behind Col-O in a majority of experiments.

The extremely long-term observations of germination response using No-O and Col-O with ages greater than 150 weeks (see Appendix 1-1) show that there are limits to the length of time for which these seeds remain viable (e.g., in No-O, the strong negative correlation between age and germination frequency is particularly noticeable). L*er* and Cvi exhibit two distinctive age-related responses. At nearly 30 weeks post-harvest, L*er* germination reaches maximum efficiency (approximately 95%), while Cvi germination has not reached maximum efficiency even at 80 weeks. The best fit line shown in Figure 3-4 for L*er* suggests that improvement in germination frequency levels off at 30 weeks at a value somewhere between the actual 30-week high value and the somewhat reduced 60-week value. It is conceivable that the decline in L*er* germination at 60 weeks is due to L*er*-specific delay alleles acting at this time. This hypothesis will need to be tested in a QTL study.

Stratification response: A short moist-chilling (stratification) requirement is the defining characteristic of winter annual plants (Allen and Meyer 1998). It was not possible to measure stratification response in Col-O, L*er,* or No-O under the incubation/induction conditions used in this work because their germination curves went to completion too quickly. However, Cvi shows a strikingly linear stratification response with r^2 values greater than 0.85 in all rotations tested (N=3 rotations). The predictability of this response made it possible to use long stratification treatments in young Cvi seeds to alleviate some of the delay that would otherwise confound the physiology of the effect being observed. Cvi also

showed the ability to improve its germination when stratification was applied after 96 hours of incubation. This phenomenon makes it possible to distinguish between delay due to dormancy and germination failure due to seed death. Application of the chilling technique would be perfectly feasible in the context of a QTL study, and its use would limit the loci discovered by QTL analysis to those responsible for short-term, reversible germination delay.

Resistance to high-temperature, high-humidity storage conditions: The seed storage assay employed in these studies is based on a combination of standard tests of seed vigor (Hampton and TeKrony 1995). In my experiments, I used temperatures and humidities at the high end of what might be considered ecologically relevant. The choice of 37°C for the temperature and 99% for the humidity was made for three reasons. First, the use of 37°C (as opposed to using the International Seed Testing Association's 45°C standard) (Hampton and TeKrony 1995), allows for a greater variety of treatment lengths and assures that the seeds will not die during treatment [e.g., *Arabidopsis* seeds lose viability at temperatures above 45°C (Grinikh et al. 1983)]. Second, the use of 99% humidity eliminates the need for a controlled humidity chamber. The results presented here document accession-specific variation in this capability that is potentially adaptive. The pattern of this variation (e.g., that Col-O and L*er* are sensitive, and No-O is resistant) follows the known genetic relationship among the European accessions (the relationship between No-O and Cvi is not known). The finding that No-O is thermal tolerant is in accordance with the prediction made in Chapter 2 regarding the necessity for frequent successful off-season reproduction in high-germination/low-establishment scenarios. If the negative correlation between seed age and germination frequency is the result of decreased viability in No-O (a hypothesis that could be confirmed by a series of

timed-viability tests), this accession exhibits two traits predicted to be associated with accessions that employ an "off-season" strategy to compensate for frequent reproductive failure.

Positive correlation between germination response and summer harvest dates: My observation of a large 24-hour "harvest season" effect on germination in L*er* must be regarded with caution because only a small number of "summer" observations (N = 2) were made and because the effect becomes undetectable at 48 hours; however, this observation is consistent with the finding of Derkx and Karssen (Derkx and Karssen 1993) that summer harvest dates were correlated positively with germination frequency. In contrast to L*er* , the "Harvest Season" effect in No-O is detectable at both 24 and 48 hours. This effect was observed when two independent "summer" rotations and three independent "winter" rotations were analyzed.

The presence of a reciprocal effect involving the No-O accession is consistent with the phenotype seen in mutants defective in seed-coat development genes which all show a positive maternal-effect correlation between germination frequency and seed-coat structural defects. The Cvi-borne F_1's for both May and August had germination frequency values greater than two standard deviations below the values for their No-O-borne counterparts. Moreover, the correlation between germination frequency and the covariates "age" and "season" was dependent on the maternal parent. If Cvi was the maternal parent, age was the better predictor of germination efficiency. If No-O was the maternal parent, "Season" was more predictive of germination efficiency. Interestingly, the difference between the dates of harvest (13 weeks) is much greater than the two-week difference in the "Season" value (Figure 3-9)

suggesting that long days early in development contribute more to the No-O rapid germination phenotype than long days later in development.

Although technical considerations prevented me from repeating this experiment, I consistently encountered positive reciprocal effects in No-O borne F_1's, whether the reciprocal parent was Col-O or Cvi. Most strikingly, CvN F_1 seeds sown on soil exhibited an absolute 7-day stratification requirement before germination occurred (N = 14). Preliminary work on tetrazolium-permeability assays on August and December rotations of No-O suggested that August-harvested seeds were more permeable to tetrazolium, though a quantitative assay would need to be developed before QTL analysis of these traits could be undertaken. Increased seed-coat permeability to tetrazolium salts is associated with defects in seed coat integrity, and in *dag-1* mutants is associated with reduced dormancy and rapid germination (Papi et al. 2002). The existence of a number of independent observations involving positive maternal effects in No-O borne F1 seeds together with the observation of season-specific seed-coat permeability effect in No-O point to several possible lines of investigation involving seasonal regulation of No-O seed coat integrity.

Conclusions: For QTL studies involving long-term germination delay, a 48-hour time point assay after imbibition is maximally informative. However, the Cvi values are low enough at 48 hours that it may be necessary to use a later mapping point in order to tease out the presence of Cvi-derived germination delay alleles. The 84-96 hour time frame may well be informative for this accession, since at 60 weeks many Cvi rotations steadily increase in germination capability after 72 hours (Figure 3-2C). QTL maps produced on 15- and 60-week old seeds, using the 48- and 96-hour time points, are likely to reveal many of the critical players in germination control and to shed light on whether short-term and

long-term germination delay are orchestrated by the same or different genetic mechanisms.

I have shown that there are accession-specific components in three traits likely to contribute to the successful persistence of *Arabidopsis thaliana* seeds in soil seed banks: seed age, storage stress response and harvest season. QTL mapping using the 48-hour germination percentage should reveal the maximum genetic effect, especially if the data are appropriately transformed to minimize the variance due to the scaling issues involved with binary data. My data suggest several feasible strategies for conducting QTL studies on seed storage and maternal growth variables. For harvest season, a cross between Col-O and No-O using deep-winter rotations (associated with the lowest No-O values in my work) may yield insight into genes which are responsible for maternal regulation of germination. The presence of seasonal dormancy patterns in L*er*, along with the large effect of harvest season at 24 hours, suggest that a cross between L*er* and Col-O could yield evidence for genetic control of these secondary dormancy patterns without the confounding profound germination delay present in Cvi. Thus, this work represents a significant step towards identifying genetic loci associated with traits that facilitate the persistence of seeds in soil seed banks.

CHAPTER 4

QTL Analysis of Long-Term Seed Germination Delay in *Arabidopsis* Accessions L*er* and Cvi

ABSTRACT

Germination delay in both dormant and non-dormant seeds is an essential component of seed survival strategy. Here I used a categorical trait approach to detect Quantitative Trait Loci (QTLs) responsible for germination delay at the 48-hour time point in 4-60 week old seeds of *Arabidopsis* accessions L*er* and Cvi. Segregation analysis of F_1 and F_2 progeny of a L*er*/Cvi reciprocal cross reveals a complex, time-dependent mode of inheritance of the delay phenotype. "Early alleles" appear to be dominant in a 22-week old F_1 population, whereas "delay associated alleles" appear to be dominant in a 4-week old F_2 population. Phenotype distributions among 12- and 60- week post-harvest L*er*/Cvi recombinant inbred lines (RILs) reveals evidence of bi-modality in both populations. The lower mean germination frequency of the 60-week old populations suggests a time-specific allele conferring delay in that population. The distribution of differences between trait values for each line at 12 and 60 weeks reveals that fully half the lines were less delayed at 12 weeks than 60 weeks when the 48-hour germination time point values were used. This negative difference disappears at 84 hours. Composite-interval (CI) mapping using values obtained at 48 hours reveals a single major locus between 75 and 83 cM on chromosome 5 (named Germination Delay 1, *GD*1) that is responsible for >70% of the phenotypic variance in 12-week old seeds. In 60-week old seeds, an additional locus on chromosome 3 (*GD*2) was required for maximum germination delay. L*er* -derived alleles within the *GD*3 locus just proximal to the *GD*1 locus

conferred the time-specific germination delay in 60-week old seeds predicted by the lower overall mean in the 60-week populations. Computational analysis of DNA sequences from this region, aided by the TIGR and PLANTGDB annotation portals, reveals that two members of the gibberellin 20-oxidase gene family reside inside and slightly toward the telomeric end of the interval, while just centromeric to the interval is an intriguing tandem cluster of germination-specific genes. Within the chromosome 3 locus required for germination delay at 60 weeks, three gene candidates were found: a member of the expansin family, a gene in the flavonol biosynthesis pathway, and an auxin-independent growth gene. Additionally, sequence-anchored markers (AFLP markers with known sequences assignable to an unambiguous genomic location) within the *GD*1/*DOG*1 region on Chromosome 5 suggest that the peak of my *GD*1 QTL lies nearly a megabase distal to the cloned *DOG*1 gene (shown to be important for after-ripening), and that *DOG*1 is outside of *GD*1. Therefore, multiple genes on chromosome 5 may underlie germination delay differences between ecotypes. These data suggest that the mechanisms underlying both the after-ripening requirement and asynchronous stratification-responsive germination are encoded by a single, structurally complex region containing multiple tightly-linked, and potentially functionally related, genes. This locus appears to epistatically interact with distinct, secondary auxiliary loci in a time-dependent manner.

INTRODUCTION

The nature of major phenotypic differences in closely related lineages is a function of both the time period over which the lineages have been isolated and the environmental conditions experienced by each lineage. Understanding the genomic basis of phenotypic variation among closely related lineages of

prokaryotes and eukaryotes remains a key challenge of the post-genomic era. *Arabidopsis thaliana*, the plant model system in which I am addressing this challenge, occurs in diverse environments from Stockholm, Sweden to the equatorial Cape Verde Isles. This broad range of habitats, combined with the presence of genetic polymorphisms in the studied accessions (Koornneef and Karssen 1994; Koornneef et al. 2000), suggests the potential for natural selection on adaptive traits. There is general agreement that *Arabidopsis* has been distributed across this broad geographic and climactic range by means of human migration (Weining et al. 2003). This fact may limit the possible time scale for mutation. Since the background mutation rate is low relative to the implied time scale for *Arabidopsis* distribution, major adaptive change would of necessity be the result of mutations in relatively few genes. Quantitative-trait locus (QTL) analysis of flowering time (Alonso-Blanco et al. 1998), nitrate storage (Harada et al. 2004), and seed storage stress response (Groot et al. 2000) suggest that a single QTL is responsible for the major differences observed in these potentially adaptive traits.

QTL analysis is conventionally conducted under the assumption of complete divergent selection (Lynch and Walsh 1998). Under this assumption, alleles with a negative effect are assumed to transgress from the "low" parent, while all positive effect alleles are assumed to transgress from the "high" parent. In such a study, the F_1 progeny are assumed to have intermediate phenotypic values. In a diverging recombinant-inbred line (RIL) population, the RILs are expected to have values arranged in a normal distribution, about a mean coinciding with the F_1 phenotypic values and extremes corresponding with the two parental values. This expectation presumes that all contributing alleles are additive, independent, and of equal effect (for review, see Lynch and Walsh 2000).

In a population under balancing selection, the intermediate phenotype is favored. Alleles in such populations become fixed in combinations of positive and negative effects that sum to the favored phenotypic value. In out-crossing populations, the intermediate phenotype is maintained by heterozygosity at critical loci. In self-fertilizing populations, where homozygosity predominates, the intermediate phenotype is maintained by unique sets of compensatory loci. If two distinct self-fertilizing lineages under balancing selection are crossed, and there is substantial genetic polymorphism between them, the average parental values and the F_1 values will be identical. RILs made from such a cross will follow a normal distribution, with a mean equal to the parental and F_1 values. High and low extremes represent the sum of positive and negative effects, respectively. As in the divergent-selection model, this expectation assumes that all contributing alleles are additive, independent, and of equal effect. The likelihood that unique allele sets confer the same phenotypic value—as opposed to the possibility that there is no difference between the lineages—is a function of the number of polymorphisms that are present in these lineages.

In recently separated populations under balancing selection, mutations which affect critical pathways may stabilize in a pattern of single loci which are sufficient, but not necessary to confer desirable phenotypes. Lineages produced in this manner will exhibit substantial difference from ancestral lineages, as would be expected under complete divergent selection. However, when out-crosses are made between these lineages, the parents will not represent the extreme value. RILs obtained from such crosses will segregate to either side of the parental values, and the F_1 values (assuming no dominance is present) will again be intermediate. If the mutated allele is sufficient to confer a particular phenotype, the lines carrying the allele will exhibit a range of values which fall within the

favored range. Under balancing selection, the recent, major-effect allele must be capable *on its own*, of conferring a favorable phenotypic outcome. Contrast this situation with the case where the new allele is necessary but not sufficient. Now, the presence of *at least one* other allele will be required in order to achieve a desirable outcome—meaning that more than one step would be necessary in order to attain maximum fitness.

The role of balancing selection in maintaining vigor in self-fertilizing populations cannot be underestimated. Indeed, small-effect QTLs have been observed in *Arabidopsis thaliana* RIL populations with no evident phenotypic differences between the parents (Rauh et al. 2002; Loudet et al. 2003; Harada et al. 2004). The most striking example of this stabilized pattern is seen in QTL analysis of circadian leaf movement (Swarup et al. 1999). In these analyses the parents had similar phenotypes, while phenotypic values observed between the recombinant inbred lines represented the phenotypic extremes. Each parent contributed balanced allele pairs whose effects summed to an intermediate phenotype. These data suggest that incremental balancing selection plays a major role in maintaining vigor in *Arabidopsis* self-fertilizing populations.

Arabidopsis has a number of well characterized mutants with phenotypes that reflect alterations in germination regulation which may have adaptive potential. These mutants fall into three major categories: (a) hormone production and response mutants (Nambara et al. 1998), (b) hormone-independent mutants [e.g. the reduced-dormancy mutants (Leon-Kloosterziel et al. 1996)], and (c) seed coat mutants (Debeaujon et al. 2000; Papi et al. 2000). Mutations at all of these loci represent potential candidate genes that may underlie quantitative traits. *Arabidopsis* is highly suitable for QTL-guided candidate gene inference because: (a) of the unique tractability of *Arabidopsis* QTLs with RIL analysis; (b)

the expected small number of genes underlying each QTL; (c) the extensive genomic annotation data base that is available, and (d) the compact genome of this organism.

To date, four QTL studies have been conducted on *Arabidopsis* seed germination. Two of these studies concern the after-ripening requirement (van der Schaar et al. 1997; Alonso-Blanco et al. 2003). The third study is a field study conducted in a novel set of recombinant inbred lines derived from parental stocks which differ in the seasonality of their germination behavior (Donohue 2002; Donohue et al. 2005). The final study is an evolutionary study on seed germination and other fitness traits conducted on a new set of L*er*/Col recombinant inbred lines (Malmberg et al. 2005). These studies differ both in the scope of the germination phenotype being studied and in their genetic outcomes.

In the first after-ripening study, carried out on L*er*/Col RI lines, 14 small effect loci were found to account for the relatively minor difference in after-ripening requirement between these two accessions (van der Schaar et al. 1997). By contrast, the L*er*/Cvi after-ripening study revealed two major QTLs (Alonso-Blanco et al. 2003). One major QTL proved to be a complex 30 cM locus at the bottom of chromosome 5. The second, a Ler-derived delay allele was found on chromosome 1. However, these authors could not resolve the precise location of the germination delay associated alleles. In addition, the presence of many QTLs, most of which transgress from the Cvi parent, is inconsistent with both the balancing-selection model of germination suggested by the ecological studies and the adaptive change possible given the relatively short time in which *Arabidopsis* accessions have been isolated.

QTL studies on after-ripening are likely to reveal loci responsible for maintenance of seed storage reserves (Heath et al. 1986), membrane

stabilization during desiccation (Koster and Leopold 1988; Chen and Burris 1990), and long-lived mRNAs required for the initiation of radicle extension (Obroucheva 1999). Because each of these processes is likely to have many molecular components, it is probable that the use of seed storage based measurements will reveal many loci. Since not all of these processes are involved in radicle extension, some measurement strategy must be used to confine the loci revealed solely to those governing processes affecting radicle extension. Taking measurements from within the germination time course is an effective strategy to focus specifically on germination-related events (Haigh and Barlow 1987; Bradford 1990).

There are a number of drawbacks to this strategy in species such as *Arabidopsis*. First, the rapid completion of the germination time-course necessitates frequent timed counts in order to determine the time required to reach 50% germination. Second, the likelihood that germination phenotypes are under balancing selection means that neither parent can be assumed to be carrying only "early" alleles. Thus, when the "early" or "high" parent reaches the completion of the time course, no information can be gained from the results since the effect of any minor delay alleles would be rendered undetectable. Building on the data in Chapter 3, I employ the germination frequency of a population of seeds from each of 150-155 recombinant inbred lines to determine the location of major germination-delay associated QTLs.

In this study, I show that the segregation of the F_2 progeny of a L*er*/Cvi reciprocal cross at 48 hours is consistent with a requirement for two dominant alleles responsible for germination delay. I report that the distribution of germination frequencies among the lines when assayed at 48 hours and 12 weeks is substantially bi-modal and that, at 60 weeks the distribution of

germination frequencies at 48 hours is nearly Mendelian and consistent with two unlinked, major effect delay-associated alleles. I show that there is a distinct delay-associated mechanism in 60-week old seeds detectable at 48 hours but not at 84 hours of incubation. Finally, I show that a single locus (*GD*1), located at the bottom of chromosome 5 and distinct from the published Delay Of Germination 1 gene, is substantially responsible for the large amount of germination delay observed in the Cvi accession. This locus is sufficient in 12-week seeds to prevent completion of the germination time-course by 48 hours. I show through epistasis analysis that an additional locus on chromosome 3 is required in conjunction with *GD*1 in order to confer maximum germination delay. Low densities of sequence-anchored AFLP markers in that region precluded candidate gene identification. Finally, I show that the lower overall mean trait value in 60-week old seeds is associated with a single major locus, *GD*3, whose delay phenotype transgresses from the L*er* parent. I support these findings by illustrating the reproducibility of my QTLs in multiple, independent analyses, and chart the course to individual gene identification by anchoring the QTLs to the public *Arabidopsis* genome assembly and enumerating linked candidate genes relevant to germination.

MATERIALS AND METHODS

Plant growth: Seeds representing the 162 Ler/Cvi recombinant inbred lines (RILs), along with the L*er* and Cvi parents, were initially obtained either from the *Arabidopsis* Stock Center (Columbus, Ohio, USA) or from Lehle Seeds (Round Rock, Texas, USA). Because slight variation in cultivation conditions contributes enormously to phenotypic variation (Rauh et al. 2002; Loudet et al. 2003), each experiment was conducted using seed lots that were obtained from a

simultaneously planted cohort. All seed lots were used within two years of harvest. Seeds for all experiments were produced in an air-conditioned greenhouse, which was maintained under natural lighting conditions from April 1 to October 1. From October to April natural light was supplemented with full-spectrum artificial light to produce long-day conditions (14 hours light, 10 hours darkness). Plants were randomized across benches to compensate for slight variations in light intensity. Seeds were sown on water-saturated soil (SunShine potting mixture) and thinned after 1.5-2 weeks. To isolate plants and promote uniform yield, once inflorescence stalks were well-established, individual plants were either staked, placed in perforated plastic bags, or isolated using the Arasystem™ cones and tubes (Arasystem, Inc.). Seeds were harvested when siliques began to reach maturity. Seeds were stored between two and eight weeks at room temperature before being transferred to long-term storage in 1.5 mL plastic microcentrifuge tubes at 4°C.

Reciprocal crosses: Crosses were preferentially performed on newly bolted plants using buds 1-2 days from opening. First, any open flowers on the inflorescence were removed, and then the largest 3-6 buds were dissected to remove all sepals, petals, and stamens without allowing immature stamens to contact the stigma. Dehiscing stamens from the pollen parent were promptly dabbed onto the stigma of the dissected bud. Crossed plants were labeled and spaced well apart to prevent accidental pollen transfer to the naked pistil, though experience suggests that the stigma is no longer receptive after several hours.

Handling of Recombinant Inbred Lines: Because of the intensive handling requirements inherent in the time-course based assay, a number of weeks were necessary to process all 162 RILs. Therefore, the planting dates of the two experimental runs were synchronized, and the lines were subjected to

phenotypic assay in harvest order over a period of 12 weeks ("60-week") study or five weeks ("12-week" study), according to Appendix 2-1.

Plating conditions: All seeds were sown on ½ X Murashige and Skoog (MS) (Murashige and Skoog 1962) basal salt mixture (Sigma Chemical Co) 0.8% PhytaGel (Sigma Chemical Co) in 10 cm square, 6 by 6 gridded plastic plates (Phenix Research Products). 50-200 seeds of each accession were placed in plastic 1.5 ml microcentrifuge tubes for surface sterilization. Between 750 and 900 μL of a 30% $HClO_4$ solution were added to each tube. The tubes were shaken at 20°C for 15 minutes with reciprocal agitation. At the end of the agitation period, seeds were rinsed five times with 1000 μL of sterile double-distilled water. The tubes were filled to 1.5 mL with a solution of 0.1% Phyta-Agar (Sigma Chemical Co USA) made in double distilled water. The resultant suspension ensured even spacing of seeds when sown onto plates with a pipeter. Plates were sealed with gas-permeable paper surgical tape (3M Corporation) to prevent desiccation.

Incubation conditions: For the "12-week" mapping experiments, plates were placed in a constant-temperature chamber (Conviron CMP4030) maintained at 21°C in constant white light at >45 μE. For the "60-week" mapping experiments, plates were placed in temperature-controlled growth chambers with a temperature range of 17-23.5°C in constant white light at >45 μE. At the end of 84 hours those plates containing ungerminated seeds were placed in a 4°C refrigerator for 72 hours. This stratification treatment was used to assess the degree of residual seed dormancy. Non-viable seeds, detectable visually by shriveled or discolored appearance, were not included in population totals.

Segregation analysis: 30-50 20-week-old L*er*/Cvi F_1 seeds and 80-100 5-week-old L*er*/Cvi F_2 seeds were sown as described above and placed in a 20°C

constant-temperature growth chamber for incubation in constant white light at >45 μE. Plates were scored for germination at 24, 48, 72, and 84 hours. No stratification treatment was given. T-tests for similarity between the parental, F_1, and F_2 populations were carried out using the arcsine-transformed mean germination frequencies in order to minimize over-estimates of significance (Alonso-Blanco et al. 1999).

Phenotypic Analysis: After sowing, plates were scored for germination every 12 hours, with radicle emergence the criterion for germination. Scoring was terminated at 84 hours post induction. To eliminate the possibility that low germination frequencies resulted from reduced seed viability rather than seed dormancy, all plates were placed in the dark at 4°C for 72 hours following the 84 hour incubation period. At the end of this stratification treatment, plates were returned to the growth chamber and scored for germination 72 hours later. The value obtained at the end of the stratification treatment was set as the maximum germination value, and experimental results were normalized to this value. In the "60-week" experiments, the L*er* parent was screened twice. The Cvi parent was screened six times in independent experiments. In the "12-week" experiment, the *Ler* and Cvi parents were run in conjunction with each subset of lines.

QTL Mapping: Composite-interval mapping of RIL data was conducted using QTL Cartographer Model 6 (http://statgen.ncsu.edu/). This program allows the operator to choose the number of background markers used to control for the genetic background. A window size of 10 cM and a walking distance of 2 cM was employed in this analysis. Fifteen randomly chosen markers were used to control for the genetic background in the final mapping studies. For each set of experiments, 850 permutations of the phenotypic data were performed to establish experiment-specific significance thresholds. Values for the analysis

were obtained using the arcsine transformation of germination percentages for each line at each time point past 36 hours post-induction. Maps were constructed using the marker segregation database for the 162 RIL lines provided by the *Arabidopsis* Information Resource Web site at ftp://ftp.arabidopsis.org/ (Alonso-Blanco et al. 1998).

Epistasis analysis: To determine whether epistatic interactions contributed to germination delay, genome-wide pair-wise searches of QTLs were conducted using the computer program Epistat (Chase et al. 1997). Epistat examines all possible marker pairs and checks to see whether the interaction between these pairs are non-additive (Chase et al. 1997). Output from the program was pruned by setting a significance threshold slightly more than twice that obtained from permutation testing in the composite-interval mapping process. The subset of data obtained from this process was subjected to Monte-Carlo simulation analysis to obtain a p-value. When an interaction was represented by nested marker pairs, the pair with the highest likelihood and lowest p-value in presented. Closely linked marker loci in which epistasis was detected were included if their likelihood ratio was greater than 100 and if they showed no evidence of conferring a phenotype on their own.

Sequence analysis. To delineate the QTL region, the names of PCR-based genetic markers in the Lister-Dean RI map (Alonso-Blanco et al. 1998) were used to retrieve AFLP primer sequences from the Arabidopsis Information Resource Web Site. To localize the markers to fully sequenced BAC/PAC clones represented in the public Arabidopsis genome assembly, AFLP-derived primer sequences were used as queries in a BLASTN (Altschul et al. 1990) search of the Arabidopsis subset of the Non-Redundant GenBank database at NCBI (http://www.ncbi.nlm.nih.gov/blast). Clone names retrieved from the BLASTN

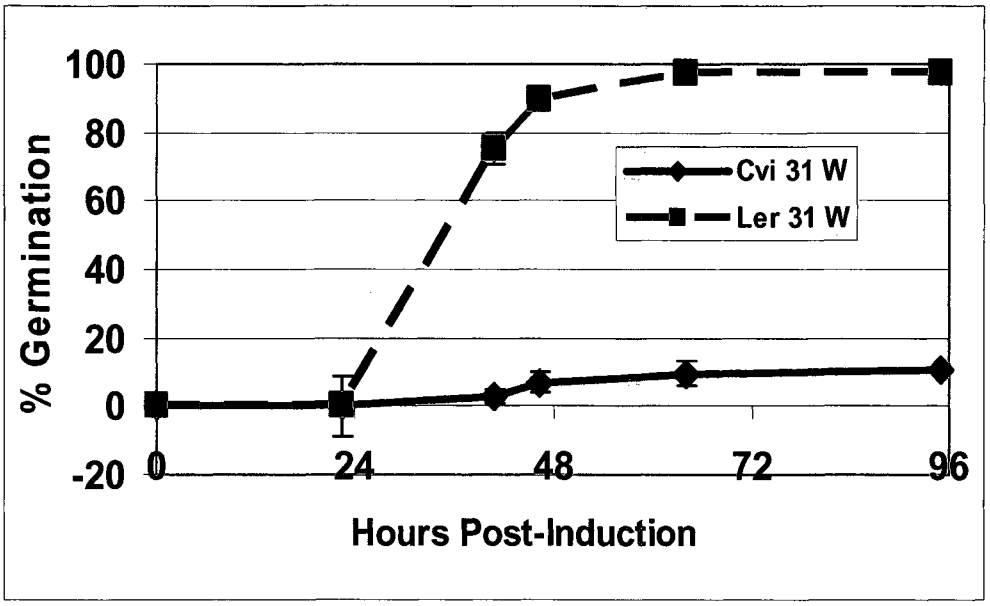
output were anchored to the Arabidopsis BAC Tiling Path feature of the TIGR (JCVI) Arabidopsis Genome Annotation Database (http://www.tigr.org/tigr-scripts/euk_manatee/listchromosomes.cgi?db=ath1&user=access&password=access). The gene list for every clone in the tiling path was retrieved, and redundancies due to clone overlaps were eliminated. A final gene list covering the region was then generated by parsing out the names of all transposon-related genes, all purely hypothetical genes lacking expression data, and all novel expressed genes of unknown function I did not include transcribed repetitive elements or genes that had no evidence of being expressed. I recognize that by also eliminating expressed genes of unknown function that I may have excluded some germination delay candidate genes simply because they lack any functional annotation.

RESULTS

The Germination Delay Phenotype

There is substantial germination delay in the Cvi accession. To examine the germination time-course behavior of the L*er* and Cvi accessions, I tested multiple plantings of these accessions ranging in age from 1-110 weeks. Since Cvi exhibits a long after-ripening requirement, this accession should show a slow increase in germination capability as the seeds age. 31-week old L*er* and Cvi seeds behaved as expected (Figure 4-1). Cvi showed less than 20% germination after 72 hours; whereas L*er* reached 85% during this same time period. The results were similar for other plantings tested during this time (data not shown).

Figure 4-1: Germination Behavior of Cvi and L*er* Parental Accessions



Representative germination time course curves for accessions L*er* and Cvi at 31 weeks post harvest. In the 31-week graph, three replicates of L*er* from a single experiment and six replicates of Cvi from two independent experiments are shown. All experiments were conducted at incubation temperatures of 18-22 degrees C, under >45 uE/s cool white light.

Analysis of F_1 and F_2 data suggests a complex inheritance mode for germination delay alleles. To determine whether the distinctive germination patterns in the L*er* and Cvi accessions exhibited a detectable mode of inheritance, I crossed L*er* and Cvi reciprocally to obtain F_1 and F_2 progeny. These seeds were subjected to a germination test under standard conditions and assayed for germination after 48 hours of incubation. Figure 4-2A shows the F_1 phenotypes. Although both L*er* and Cvi-borne F_1 offspring demonstrate a slight delay relative to the L*er* parent and the 20-week-old Col-O positive control, neither the difference between the two reciprocal F_1 populations, nor the delay between the L*er* parent and the L-Cv F_1 is statistically significant (Table 4-1). The difference between the CvL F_1 and the L*er* parent is statistically significant, but in all other respects the results match those expected if the L*er* early alleles are dominant (Table 4-1).

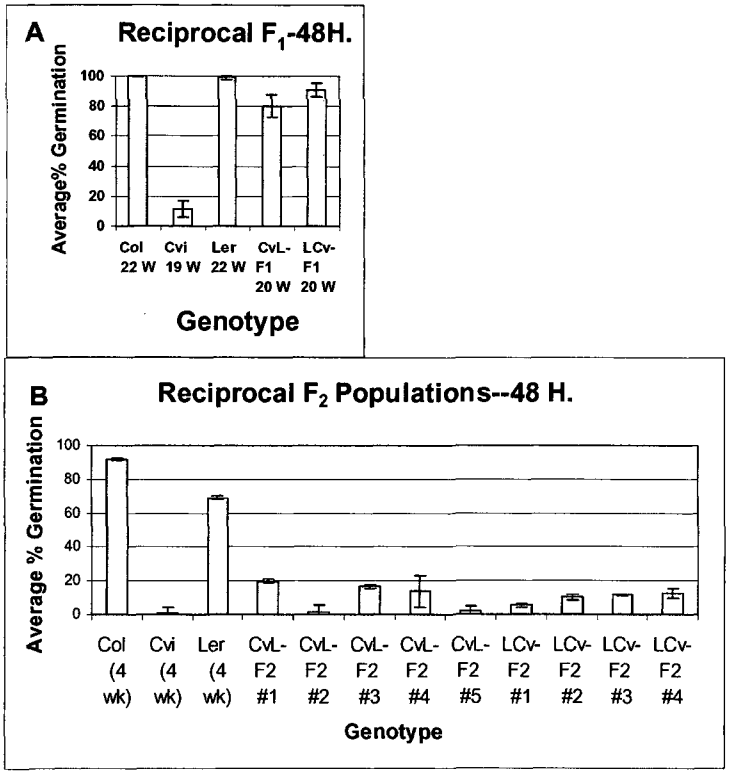
Table 4-1: Logic of L*er* Allele Dominance Hypothesis

Null Hypothesis	P Value	State	Expected (L*er* Dominant)	State (Dominance)
L*er* = Cvi	0.005	False	False	True
LCv F_1 = CvL F_1	0.1	True	True	True
LCv F_1 = L*er*	0.1	True	True	True
CvL F_1 = L*er*	0.0078	False	True	False
CvL F_1 = Cvi	0.005	False	False	True
LCv F_1 = Cvi	0.005	False	False	True

To further clarify the inheritance pattern of germination-behavior alleles, I analyzed both LCv and CvL F_2 populations in the same experiment as the F_1 populations analyzed above. The difference between the 4-week Col-O and L*er* populations is also significant (p=0.0014, Figure 4-2B), suggesting that, in

addition to environmental factors, the L*er* accession carries genetic factors associated with delay. The mean of the reciprocal F_2 populations is 10.5 for CvL and 8.2 for LCv, but the difference between the two populations is not significant ($p>0.1$), suggesting that there is no maternal inheritance.

Figure 4-2: Analysis of F_1 and F_2 Germination

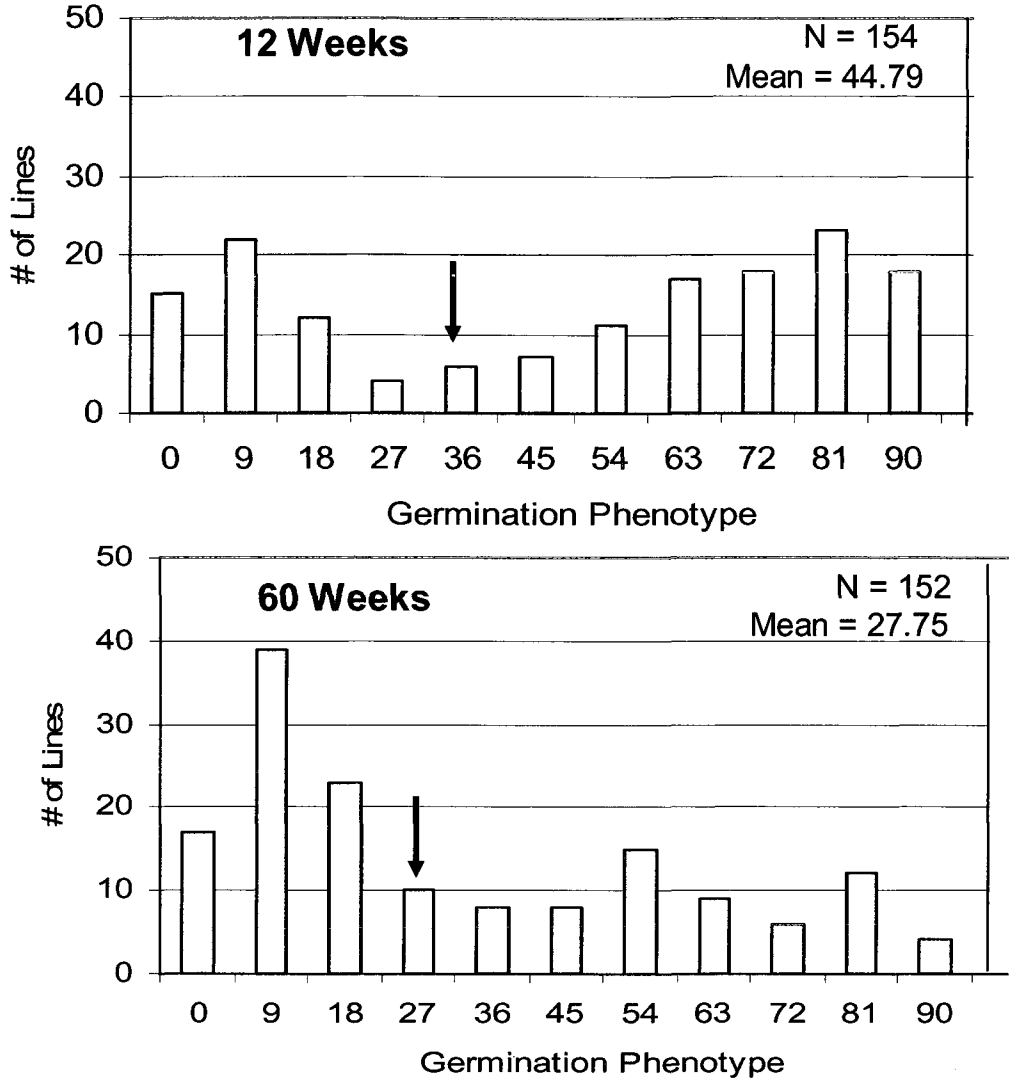


Germination response of the progeny of a L*er*/Cvi reciprocal cross. **A:** Reciprocal F_1 seeds after 48 h. incubation. **B:** Reciprocal F_2 populations at the same time point. Col-O control is added as the "non-delayed" accession. Bars represent the mean of triplicate plates; error bars represent standard deviation.

Phenotypic distributions among Ler/Cvi recombinant inbred lines are bi-modal. Recombinant inbred line populations can be used to clarify inheritance patterns in germination by substantially increasing the sample size per genotype being assayed. To determine whether the L*er*/Cvi RIL populations demonstrated the bi-modal distributions expected if a major delay allele is present and capable

of acting alone, I analyzed two independent L*er*/Cvi RIL plantings, one at approximately 12 weeks and the other at approximately 60 weeks post harvest. Both plantings were assayed at 48 hours of incubation. The 12-week distribution, at 48 hours, had a mean of 44.8 and a standard deviation of 31.6 (Figure 4-3A). Only six lines lay within the same bin as the mean, while 15 lines were at <9, and 10 were at completion. These data are suggestive of bi-modality at 48 hours. At 60 weeks post-harvest, the 48-hour mean is 27.8, while 17 lines had values <9 (Figure 4-3B). Only 4 lines had reached completion at this time, and only 10 lines lie in the same bin as the mean. The standard deviation for the 48-hour, 60-week time point is 27.5. Together these data suggest a bi-modal effect consistent with one or two major genes being responsible for the germination delay phenotype.

Figure 4-3: 48 Hour Phenotype Distribution

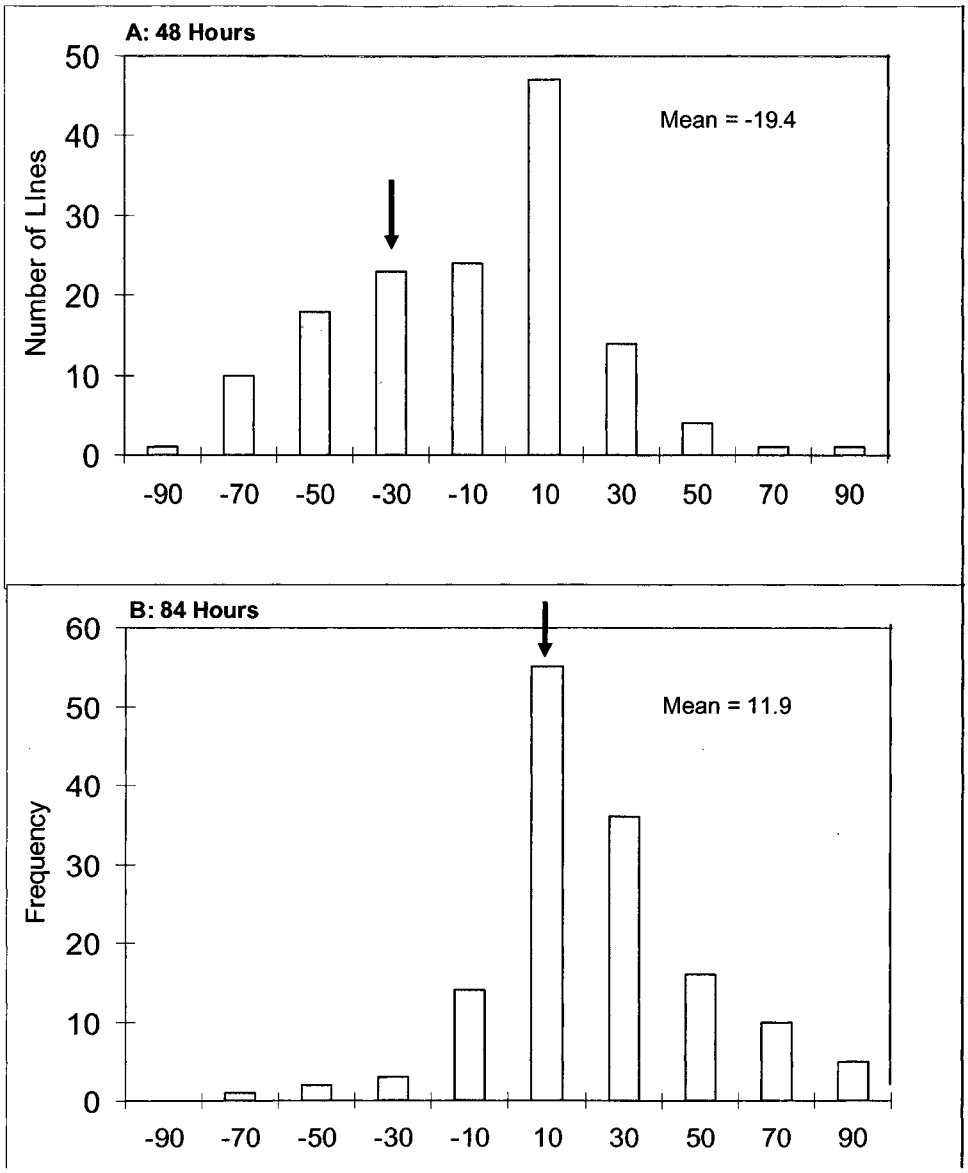


48 hour distributions at 12 and 60 weeks. Each RIL is represented by the seeds of 1-3 plants, and the number of seeds assayed per RIL was 75-200. Numbers on X-axis indicate lower bin boundaries. Arrow indicates position of the mean.

The distribution of difference values at 48 and 84 hours post induction suggests both an after-ripening effect and long-term, constituitive germination delay. Under the conventional after-ripening scenario, the 48- and 84-hour germination frequency should be higher, on average, in 60-week old seeds. However, the mean 48-hour germination frequency in 60-week old seeds was

only approximately 2/3 that of the 12-week old seeds. To examine the nature of the differences in germination behavior between 12-week and 60 week old seeds, I constructed difference distributions by subtracting "12-week" values from "60-week" values. This method should yield a positive result if lines are more delayed in the "12-week" study, as would be expected in the conventional after-ripening scenario. Interestingly, the distribution of 60-week - 12-week values at 48 hours is skewed to the left (Figure 4-4A). The mean is –19.4, and the standard deviation is 31.4. At 84 hours, the distribution appears more suggestive of an after-ripening scenario, with a mean of 11.9 and a standard deviation of 28.0, with more lines having positive values than negative (Figure 4-4B). The lines divide themselves into four categories: 1) lines with low values in both studies, 2) lines with negative difference values, more delayed at 60 weeks, 3) lines with positive difference values, more delayed at 12 weeks, and 4) lines with high values in both studies. A number of lines switch categories between 48 and 84 hours, suggesting that some lines exhibit both a dormancy effect and a long-term delay visible early in the time-course curve. These data suggest that a variety of delay mechanisms are active in the L*er*/Cvi population in addition to the conventional, time-dependent delay associated with the after-ripening requirement.

Figure 4-4: Distribution of Difference Values



A: Histogram of difference values (60 weeks – 12 weeks) at 48 hours" incubation. Mean is –19, N-142 lines, 75-160 seeds per line. **B:** Histogram of difference values (60 weeks – 12 weeks) at 84 hours' incubation. Numbers on x-axis indicate lower bin boundaries. Arrows indicate position of the mean.

Composite Interval Mapping

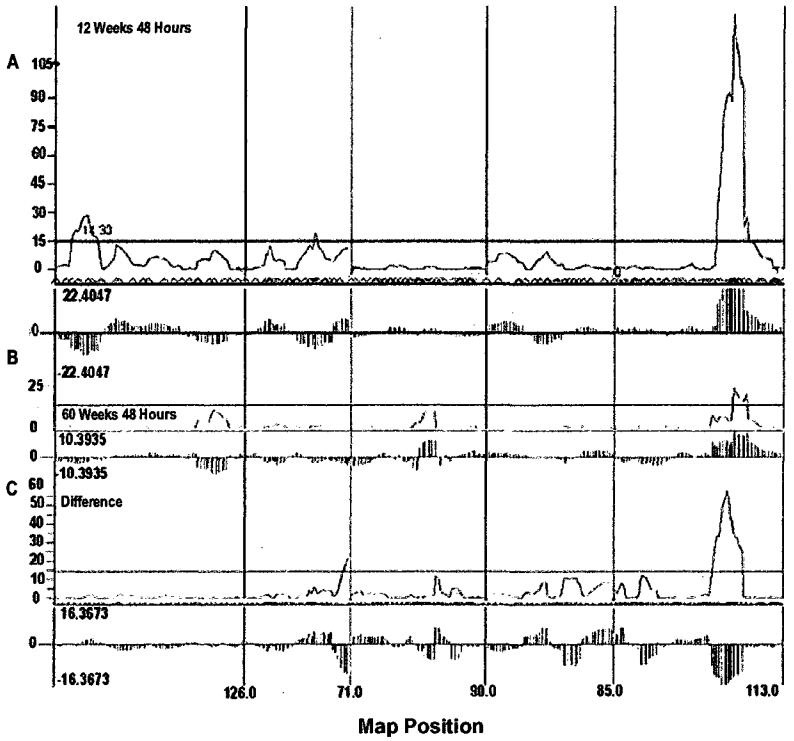
At 12 weeks, one complex locus is responsible for more than half the phenotypic variation in germination frequency. To further clarify the genetic nature of germination delay in 12-week old seeds, the germination frequency of a single planting of 162 L*er*/Cvi RILs was assayed at 36, 48, and 84 hours post-induction. Composite-interval mapping of the 48-hour arcsine-transformed germination frequency values obtained from this assay reveals a straightforward genetic picture, showing only two loci associated with germination delay. As shown in Figure 4-5A, one locus, located near the bottom of chromosome 5, exhibits a double-toothed peak. The 80-cM tooth has a likelihood ratio of 134.7 whereas the less significant tooth lies at 76 cM with a likelihood ratio of 92.00. It is theoretically possible (Fig. 4-6) that the 76-cM, 12-week peak corresponds to the cloned *DOG*1 gene. In both peaks, Cvi alleles are associated with germination delay. The combination of these two peaks, between 76 and 86 cM, accounts for 70% of the total phenotypic variation, conditional on 15 background markers. I have named this locus *GD*1 (Germination Delay 1). The second locus, whose peak is located at 21 cM on chromosome 1, has a likelihood ratio of 28.7 and is responsible for only 5% of the total phenotypic variation, conditional on 15 background markers. For this locus, the L*er* allele is associated with delay.

At 60 weeks, only the GD1 locus on chromosome 5 is independently associated with delayed germination. To examine the genetics of germination delay in seeds which are fully after-ripened, composite-interval mapping analysis was conducted on a second planting of *Ler*/Cvi RILs at approximately 60 weeks post harvest (Appendix 2). As suggested by the F_1 analysis, only one locus is associated with delayed germination (Figure 4-5B). The 60-week locus on chromosome 5, similar to the 12-week locus, exhibits a double-toothed peak. The

left-side tooth, located at 80 cM similar to the 12-week locus, had a likelihood ratio of 24.5, and explained 12% of the phenotypic variation. The right-side tooth, located at 89 cM (unique to this 60-week locus), had a likelihood ratio of 21.2, and explained 11% of the phenotypic variation. This double peak accounted for 23% of the total phenotypic variation. As in 12-week seeds, Cvi alleles at this locus were associated with germination delay.

A single major locus is associated with the negative mean "difference" value. To determine whether any loci capable of acting alone are associated with the high degree of delay in 60-week lines, I conducted composite-interval mapping on the difference values. I compared the results of this mapping study with the results of the 48-hour studies. Interestingly, only two loci are implicated in the "difference" phenotype. In both cases, the delay phenotype is associated with L*er*-derived alleles. One allele at the bottom of chromosome 2 exerts only a small (9% effect), while the other, located just proximal to the 80-cM *GD*1 locus on chromosome 5, accounts for 22% of the total phenotypic variation (Figure 4-5C). I have therefore named this locus *GD*3.

Figure 4-5: Genetic Map of Germination Response QTLs



Composite-interval maps of **A:** 12-week germination delay, **B:** 60-week germination delay, and **C:** Difference values. Y axis for the top part of each plot is log-likelihood ratio. X axis is map position, in cM. Vertical lines represent chromosome boundaries. Horizontal line represents experiment-specific significance threshold. Y axis on bottom of plot represents additive genetic variance. Positive values mean that the L*er* allele is the one associated with the early phenotype.

Epistasis Analysis

Epistasis plays a role in germination control at both 12 and 60 weeks post-harvest. The relatively small effect size and significance of the *GD*1 allele at 48 hours was surprising given the strongly bi-modal distribution of phenotypes and nearly Mendelian ratio of extremely delayed lines to mildly delayed and complete lines. This result suggests epistatic interactions may play a role in germination

delay at 60 weeks. To examine the role of epistatic interactions, I conducted a genome-wide pair-wise search for non-additive interactions using the computer program Epistat (Chase et al. 1997). To compensate for the non-normal distribution of values within single time points, I set an extremely stringent significance threshold of 30 for the Log Likelihood Ration (LLR), approximately twice the maximum experiment-wise threshold determined by permutation testing in the mapping studies. This analysis yields four interactions, two in 12-week old seeds and two in 60-week old seeds (Tables 4-2 and 4-3). One interaction, between loci on chromosome 3 and 5 confers delay only in the presence of Cvi alleles at both positions. The effect size is substantial with a genotype mean of <3% when both Cvi alleles are present as opposed to approximately 20% when the alleles are present by themselves. This reaction carries a likelihood ratio of 136, suggesting that it represents a major factor in the decline in significance of the *GD*1 locus at this time. In the second interaction between loci on chromosomes 4 and 5, the out-of-phase genotypes represent the extreme values. The lowest value is obtained with a L*er* allele in the chromosome 5 position, while the highest value is obtained with a L*er* allele in the chromosome 4 position and a Cvi allele on chromosome 5. The next highest set of values come from the in-phase genotypes, with the Cvi pair conferring a higher value together than either Cvi allele alone. When both Ler alleles are present, the value is equal to that obtained in the presence of the first L*er* allele acting alone.

The interactions present in 12-week old seeds fall into two distinct classes. The first interaction occurs between a 20-cM region on chromosome 4 and a 20 cM region on chromosome 5 which contains the *GD*1 locus. Ler alleles in the *GD*1 region confer identical phenotypic values regardless of the allele present in the chromosome 4 position. The presence of Cvi alleles in both positions confers

the most delayed value, while the remaining values confer an intermediate phenotype. This pattern is consistent with the mapping result, suggesting that *GD*1 alleles are capable of exerting substantial delay effects on their own. In the second interaction, the presence of a L*er* allele at the first locus and Cvi at the second confers a radically delayed phenotype, while all other combinations yield equivalent phenotypic values. This interaction concerns loci a scant 6 cM apart, resulting in a small number of recombinants with an appropriate breakpoint for testing. Nevertheless, the effect size and likelihood ratio are both substantial, suggesting that a L*er* allele at the first locus requires the presence of the Cvi allele at the second to confer delay at 12 weeks.

In addition to the single L*er* locus implicated in 60-week delay, two epistatic interactions were also detected. In the first, loci on either side of the *GD*1 region on chromosome 5 confer a positive value when a L*er* allele is at the first locus and a Cvi allele is at the second. In the second interaction, the highest value is also obtained when L*er* is in the first position and Cvi is in the second. (Tables 4-2 and 4-3). Interestingly, the second marker in interaction number two is also implicated, with strong significance, in delay specific to 60 weeks. L*er* alleles at the chromosome 2 locus mitigate this effect, though the difference value is still negative.

Table 4-2: Epistatic interactions

Phenotype values of significant epistatic interactions at the 48-hour time point in 12- and 60-week old seeds along with the "Difference" trait. Row headers are marker genotypes. Values in each column represent the mean value for all lines which share the genotype. Names in left-most column have the format hours-weeks-interaction number.

	Marker 1		Marker 1-Marker 2				Marker 2	
Trait	Cvi	Ler	Cvi-Cvi	Cvi-Ler	Ler-Cvi	Ler-Ler	Cvi-2	Ler-2
48-60-1	25.92	29.34	2.03	32.79	34.61	31.96	16.54	31.97
48-60-2	30.92	37.73	46.42	25.28	59.90	37.73	27.91	27.31
48-12-1	29.65	54.21	13.66	63.33	20.54	61.69	14.96	62.08
48-12-2	31.89	30.67	31.53	28.66	5.51	29.77	32.75	30.79
Diff-1	-16.65	-17.28	-13.57	-19.42	2.19	-26.26	-8.89	-27.60
Diff-2	-20.72	-13.06	-27.17	-16.69	-2.68	-15.77	-26.97	-15.14

Table 4-3: Significance and Position of Epistatic Marker Pairs

Likelihood ratios and marker positions for the six interactions. Trait name: germination time-course time point is given first, followed by the seed age, followed by a number assigned to each interaction. Positions: given as chromosome #:cM from left telomere. The likelihood ratios are between additive and null hypothesis for each marker singly, and between non-additive and additive for the marker combination.

Trait	Marker 1 Position	Marker 2 Position	Null1	Null2	Add
48-60-1	3:39 cM	5:69 cM	0.35	4.49	136.43
48-60-2	4:52 cM	5:25 cM	0.45	0.25	36.31
48-12-1	5:59 cM	5:77 cM	10.65	38.18	34.81
48-12-2	2:53 cM	2:59 cM	2.09	0.01	147.08
Diff-1	5:42 cM	5:97 cM	0.003	1.57	50.14
Diff-2	2:49 cM	3:39 cM	0.15	0.04	49.39

Genomic annotation of *GD*1 and its epistatic region

Gibberellin-pathway, auxin-responsive, and germination-specific genes are located within and near the germination-delay candidate region of chromosome 5. I surveyed all AFLP markers used in my RIL genotyping, to identify those

markers which were used in the published AFLP-based Arabidopsis genetic map (Peters et al. 2001) to sequence-anchor specific genomic clones. My resolution capacity at the sequence level was hindered by the fact that only a minority of my markers were sequence-anchorable, and to a lesser extent by the discrepancy in cM positions of identical sequences on the three available genetic maps. In spite of this challenge, I were able to localize the AFLP marker AD.75C-Col (cM 77 on my map) to genomic clone MZA15; CC.540C-Col (cM 80) to K23F3; and GB.102L-Col (cM 90) to MCO15. Unfortunately, no sequence-anchorable markers between cM 80 and cM 90 of my map were found. I ultimately integrated the published AFLP-based genetic map, my composite mapping, TIGR / JCVI genomic clone-based contigs and annotations, and MIPS tiling-path maps and annotations for a comprehensive view of the region (Figure 4-6).

The 2.1-Mb MZA15-MCO15 region of chromosome 5 was visualized as a gapless tiling path with mapped genes using the TIGR Arabidopsis Genome Annotation Database with subsequent curation (Figure 4-6, see Materials and Methods for details). No known genes directly relevant to the abscisic acid pathway or to seed coat defects were noted in the MZA15-MCO15 region. However, two members of the gibberellin 20-oxidase family reside on clones MWD22 and K17N15. Assuming that the cM / Mb ratio is constant throughout the cM 80-90 interval of my AFLP-based map, these two genes (At5g51310 and At5g51810) reside between cM 80 and cM 85 on my map. Also localized in the region are three putative auxin-responsive genes. Additionally, two ERF (ethylene-responsive transcription factor) genes reside within the *GD*1 domain at 78 cM distal to the cloned *DOG*1 gene. A functional investigation of the relationship of these genes to the Cvi germination delay phenotype thus appears warranted.

Two tandemly duplicated gene clusters with putative functional relevance to the RI-mapped germination delay associated QTLs are found on the Gramene sequence assembly (Liang et al. 2007), just proximal to *DOG*1 and to my *GD*1 region. First, 13 germin-like genes are found on the partially overlapping BAC clones K15E6, MXF12, and K3K3. According to the EST and full-length cDNA databases, 7 of the 13 are expressed. Germin gene family members have characteristic spatiotemporal expression patterns; some are upregulated during germination (Membré et al. 1997), implying a germination-related function. Second, 6 expansin-family genes are found on BAC clone MUL8, which overlaps K3K3 and is telomeric to it. In tomato, expression of an expansin transcript is associated with endosperm weakening during germination (Li et al. 2003) . Unlike the gibberellin oxidase family members, these intriguing gene clusters localize just outside *GD*1 (Fig. 4-6). However, due to the potential existence of long-range locus control regions or distant enhancers (well-established paradigms in gene regulation (Dean 2006), and to the resolution limits on my sequence anchoring of *GD*1, it is still conceivable that a loss-of-function allele in one of the germin-like or expansin-like genes, or a small deletion leading to inter-ecotype differences in gene copy number at either the germin-like or the expansin-like gene cluster, underlies part of the germination delay difference between Ler and Cvi.

The *DOG*1 gene, which is abscisic-acid-responsive, lies 1 megabase proximal to my predicted peak. This distance is substantial in terms of the sequence, and makes quite unlikely, albeit does not definitively exclude, the possibility that *DOG*1 is responsible for all or part of the phenotypes observed in my work. Nevertheless, the fact that my peak does not lie directly on top of *DOG*1 is interesting and bears further investigation.

Sparse candidates and lack of markers limit candidate gene discovery in the chromosome 3 region epistatic to GD1. Lack of sequence-supported AFLP-markers (Peters et al. 2001; Peters 2009) in the cM 23 – cM 40 region of chromosome 3 made it impossible to genomically define the distal boundary of the epistatic region (the region I will refer to as *GD2*), while the proximal boundary of the region lay just beyond the centromere (Figure 4-6B). Three candidate genes, expansin EXPA5, a putative flavonol glucosyltransferase, and an auxin-independent growth-promoting protein (At3g30300), localized to the epistatic region, although only for the latter gene could mapping within the region be confirmed. Due to the dearth of markers at the proximal end and to genomic clone contig ambiguities at the centromeric end, the region could not be annotated further.

DISCUSSION

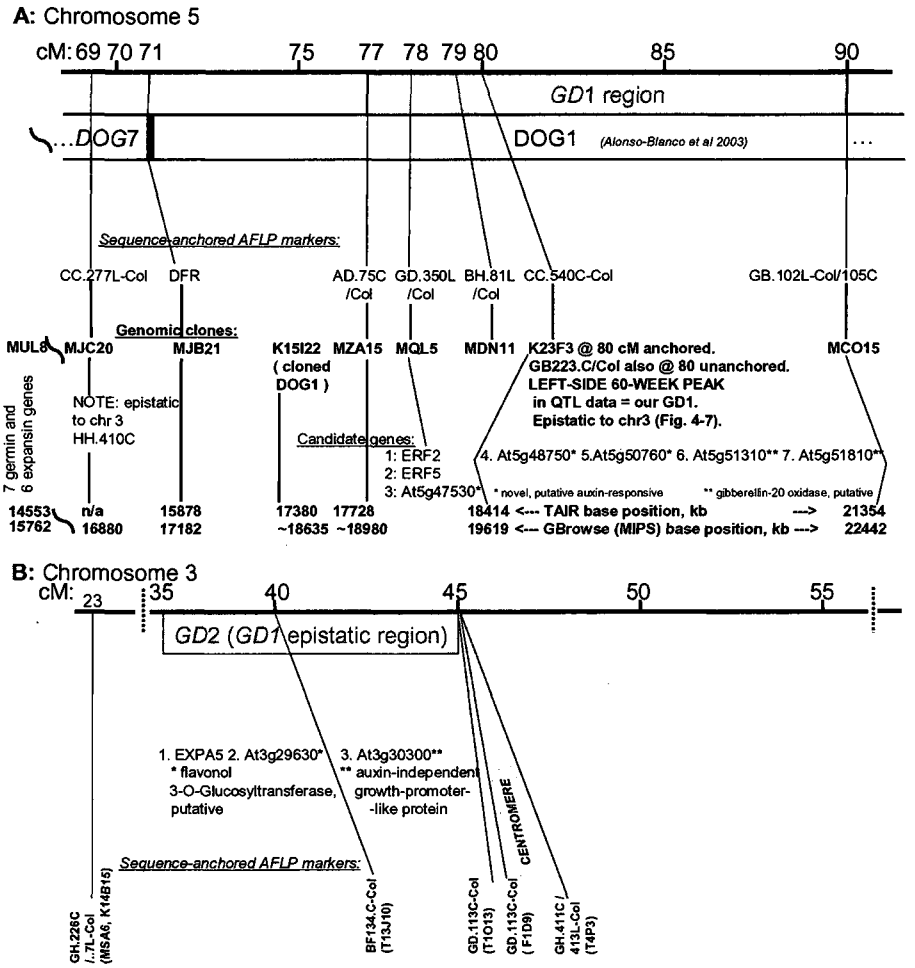
I addressed the question of pinpointing the genomic basis of Cvi germination delay by segregation analysis of L*er*-Cvi progeny as well as by mapping studies. My data show that the major accession-specific phenotypic distinction, the basis of Cvi's greater germination delay, is encoded by a single locus at the bottom of chromosome 5. In 12-week seeds, this locus is sufficient (but not necessary) to confer the germination delay phenotype. The contribution of this locus to Cvi germination delay decreases in 60-week seeds. I identified 2 specific epistatic interactions which appear to be responsible for this modulation. One interaction that occurs between loci on chromosomes 3 and 5 confers a greatly increased delay phenotype in 60-week old seeds. The observation of an epistatic interaction involving two genes in which both are required in order to achieve the

greatest delay is consistent with the observation that a quarter of the 60-week old lines lie in the lowest two bins in the distribution of un-transformed phenotypes.

Clarifying a mode of inheritance is useful for determining the direction of function for a locus. However, these experiments were complicated by increased growth rates observed in the F_1, making it unclear whether the "early" phenotype was the result of true dominance of the "early" alleles or simply the result of increased global growth rate. However, the F_2 study in which both the L*er* and Cvi parents were delayed relative to a "non-delayed" Col-O control suggested dominant inheritance of at least one delay-associated allele. This result is in agreement with data obtained from cloning of the *DOG*1 locus. A loss-of-function in this locus is associated with an early phenotype (Bentsink et al. 2006).

There are three interesting aspects of the phenotype distributions observed in this study. First is the bi-modal aspect of both 12- and 60- week distributions. This result is suggestive of a long-term major genetic effect. The second is the increased number of delayed lines in the 60-week cohort, suggestive of delay mechanisms that operate beyond the after-ripening period. The third point of interest is that when difference values are used to assess this increased delay, the phenomenon becomes nearly undetectable at 84 hours. Though by no means conclusive, this result is consistent with the idea of an after-ripening-associated gene whose expression declines as imbibition progresses (Bentsink et al. 2006) and whose effect is less well-defined beyond primary dormancy.

Figure 4-6: Annotation of Regions *GD*1 and *GD*2



A: Relationship between chr. 5 *GD*1 region and the cloned *DOG*1 gene. Black vertical bar separates *DOG*1 and *DOG*7. cM positions and corresponding lines show the relationship between *DOG*1 and *GD*1 CI maps.. Names in bold are genomic clones that anchor *DOG*1 to CI map. Wavy lines show break between genomic regions. B: Sequence analysis of chr. 3 *GD*2. cM positions and corresponding lines show the relationship between my CI map, sequence-anchored markers, and candidate genes in the *GD*2 region. Dotted line indicates break between genomic regions. Genomic clone names are in parentheses following selected AFLP marker names. Centromere of chr. 3 is between F1D9 and T4P3 on TIGR map.

The composite-interval map made on 12-week RIL germination frequencies matches well with the phenotypic distribution, and confirms that *DOG*1 is the

locus likely to be responsible for the majority of the delay effect at this time. However, the small magnitude of this germination delay-critical region in the 60-week map at first seemed contradictory. When non-transformed distributions of the 60-week lines are examined, the distribution is nearly Mendelian, with approximately ¼ of the lines being extremely delayed, and ¼ being early, along with the L*er* parent. Yet no other significant QTLs were detected. The results of the "difference" map may account for part of the 60-week delay phenomenon. This map reveals a single Ler-associated delay allele at 77cM on Chr. 5. The proximity of this locus to *GD*1 at 80 cM means that only about 8 lines would contain early alleles at both positions. This effect would tend to drive the high end of the distribution downward, and may result in an overall lower mean at 60 weeks than at 12.

The nearly Mendelian ratio of three distinct phenotype classes to each other at 60 weeks suggests that two independent loci are responsible for these classes. Thus, although the presence of the L*er* delay locus at 77 cM could account for the lower overall mean at this time, it did not explain the ratio of the phenotype classes within the distribution because of its tight linkage to the Cvi *GD*1 locus. Epistasis analysis resolves this inconsistency nicely by revealing a locus on Chr. 3 which is required for the severe Cvi delay phenotype at 60 weeks. The significance of this interaction is robust, and the effect on the phenotype intense when Cvi alleles are present at both positions. A L*er*-derived allele at the Chr. 3 position in this interaction yields the highest mean trait value, though this effect is very slight. Of the remaining interactions, the one with the highest LLR value involves tight linkage and therefore must be regarded as marginal because of the probable small number of informative lines. Of the two "Difference"-related interactions, the first between opposite ends of Chr. 5, is the

most dramatic, with a L*er* allele at 42 cM on Chr. 5 being associated with a positive value, while all other values range between –27 and –8.

The annotation data on chromosome 5 are enlightening. The detection of several potentially relevant germination loci in the *GD1*/*DOG1* region on chromosome 5 is suggestive of a "cluster" of relevant genes whose functional status combines to produce a favorable germination delay phenotype in a balancing-selection manner. The presence of linked loci in such a scenario would ensure that once the desirable phenotype was reached, loci would continue to travel together, thus ensuring the maintenance of that phenotype. However, this idea would need to be tested by examining sequences within the region in other accessions, and paralleling the germination delay phenotypes across a panel of accessions with expression, nonsynonymous substitutions, and regulatory region changes affecting the candidate genes. These suggestions for comparing ecotypes are feasible because the technology for accomplishing targeted re-sequencing of genomic regions (high-throughput Sanger sequencing, pyrosequencing, Solexa 1G), and for analyzing gene expression (microarrays, quantitative RTPCR) are widely available. In addition, to complete this work, the function of some of the closely-linked genes during germination (e. g., gibberellin-oxidase genes) would need to be verified.

Surprisingly, this project has encountered numerous limitations of *Arabidopsis* genome annotation insofar as its utility for QTL-directed candidate gene discovery is concerned. The genome sequence is only available for one or few, not all, accessions for a given locus; the majority of genes are wholly unannotated; publicly available expression data is very limited and does not cover all genes and ecotypes; and the sequence-based genomic contigs are not sufficiently saturated with genetic markers from the same panels that have

become the de-facto standard in the field for genotyping in QTL studies. Better availability of sequence-anchored markers, as well as high-throughput functional and phenotypic screens systematically profiling and prioritizing candidate genes in my regions, would have greatly enhanced this study. It is my hope that such resources will be generated by the *Arabidopsis* community by next year within the NSF-funded "Arabidopsis 2010" initiative.

Intriguingly, the characteristic up-turn of Cvi germination frequency at time points greater than 72 hours of incubation matches the expression pattern of *DOG*1, whose expression declines as imbibition progresses (Bentsink et al. 2006). The finding that this pattern persists in seeds which should theoretically be beyond the "after-ripening" requirement raises the possibility that, in addition to facilitating the acquisition of dormancy, this gene also plays a role in maintenance of dormancy. Whether the chromosome 3 epistatic region is required for *DOG*1 function in 60-week old seeds, or whether a locus closely linked to *DOG*1 interacts with the epistatic region to produce the delay phenotype is not clear from either the mapping or the annotation. These questions can only be definitively answered once the critical genes have been identified and their products placed within interaction maps of the *Arabidopsis* proteome. Clearly the availability of *Arabidopsis* lines whose *DOG*1 status is known will greatly facilitate my understanding of both primary dormancy and the maintenance of long-term germination delay, both of which appear to be adaptive across a wide range of winter annual species.

CHAPTER 5

Discussion and Conclusions

The aim of this work was to identify potential modes of adaptation in seed germination so that rapid and accessible methods for locating critical genomic regions associated with these mechanisms could be developed (Alonso-Blanco et al. 2003). I studied the adaptive nature of seed germination frequency by means of a comparative simulation analysis in which I analyzed the environmental cues that are purported to regulate seed germination (Allen and Meyer 1998). To address the physiological nature of the phenotypes predicted by the simulation analysis, I developed a phenotype assay which is both high-throughput and limited in the number of potential physiological components likely to contribute to the observed characteristic. Building on the germination responses observed in Chapter 3 and experimentally choosing the appropriate time point for monitoring seed germination, I demonstrate a nearly Mendelian mode of inheritance for the major germination frequency effect in Cvi.

The simulation analysis extends the simplest of the seed bank dynamics models proposed by Venable (Venable 1989) to include scenarios predicted to apply specifically to rapid adaptive evolution in *Arabidopsis thaliana*. The analysis is particularly focused on phenotypes that allow preservation of the seed bank during periods of sub-optimal environmental conditions, including survival in allopatric habitats (such as the Cape Verde Isles). The simulation analysis makes three predictions about the germination behavior of small-seeded colonizing annual plants. First, the central mode of adaptation required for survival in non-permissive environments is a combination of seed longevity and germination

delay. Second, it is the relationship between germination and establishment frequency that is most predictive of population success. Finally, opportunistic germination (or off-season germination) only provides protection in high-germination strategies for short-lived seed banks if successful establishment can occur every year.

In Chapter 3, I surveyed a panel of four *Arabidopsis* accessions with respect to the predictions made by the simulation analysis. In order to promote experimental efficiency and maximize genetic effects, I determined the point in time during the germination test where genetics alone exhibited the largest effect on germination frequency. Using this time point, I examined the effect of seed age, seed storage conditions, and maternal growth environment on germination frequency. As expected, Cvi (the accession collected from the allopatric habitat) exhibited a large positive correlation with seed age. As the data for this accession included seeds up to 10 years old, this result is suggestive of robust longevity in Cvi. The largest negative correlation and effect of age was observed in No-O, suggesting that seed longevity is not part of the germination control strategy of this accession. Results from experiments involving 8 days of storage at elevated temperature and combined elevated temperature and saturating humidity suggest that two accessions would be capable of an "off-season" strategy because of their tolerance to these conditions. Of these two (Cvi and No-O), only No-O exhibited germination responses that were predictable based on the time of year the seeds were produced. This effect was sizable and visible as a reciprocal effect in Cvi X No-O F1 hybrids. Together these data suggest that there are distinctions in germination strategy among the four accessions suggestive of the differences predicted in the simulation analysis.

By far the largest effect (nearly 55% of the total variation) on germination frequency was associated with the Cvi genotype. I took advantage of the availability of Cvi X L*er* recombinant inbred lines to determine the genetic location of this effect. In 12-week seeds, my detection of a single region on chromosome 5 sufficient to confer a 30% decrease in 48-hour germination frequency is in good agreement with the observations of Alonso-Blanco (Alonso-Blanco et al. 2003; Bentsink et al. 2006). In 60-week old seeds, my results add to the knowledge concerning this locus by identifying an epistatic region required for Cvi-derived decrease in 48-hour germination frequency. In addition, the overall lower mean germination frequency in 60-week seeds is associated with a novel L*er*-derived locus located only 3 cM proximal to the *DOG1*/*GD1* region on chromosome 5. Together these results suggest that long-term germination delay (such as that expected for an accession surviving in an allopatric environment) involves both after-ripening associated genes and novel loci acting specifically in 60-week old seeds. The data presented in this work tie together a theoretical framework designed to detect potentially adaptive phenotypic profiles, physiological assays aimed at detecting the "most heritable" components in a response, and mapping techniques that allow the loci discovered to be manipulated using molecular and genetic tools.

The contribution of a body of scientific research should be evaluated on the basis of whether 1) the results presented provide independent corroboration of existing results, 2) whether the new results presented follow logically from both the literature and the design of the experiments and 3) whether the new data suggest clear directions for future research.

In the case of the simulation analysis, my results match closely those of Venable (Venable 1989) with respect to the selective pressures predicted, but

add to his result in terms of the analysis of the relationship between yield and seed bank growth. In addition, the simulation program has been designed so that hypotheses concerning the effectiveness of after-ripening versus other mechanisms of timed germination may be tested. These tests will allow models of seed bank dynamics to make more accurate predictions of germination strategy based on experimental techniques and field observations.

The aim of Chapter 3 was to provide one or more effects that could be subjected to QTL analysis. For each of the responses examined, my results corroborate those of other investigators. The extensive delay phenotype of Cvi has been mapped in the context of after-ripening (Alonso-Blanco et al. 2003). The resistance to combined elevated temperature/humidity in Cvi has also been mapped in L*er*/Cvi RI populations (Groot et al. 2000), but has not been reported elsewhere for No-O. Finally my results for the "season of harvest" effect are in agreement with those of Derkx (Derkx and Karssen 1993), but also have not been reported previously for No-O. The presence of reciprocal effects between No-O and Cvi with respect to both germination frequency and "season of harvest" have also not been reported to date. The work presented in this chapter adds to the general body of knowledge about germination regulation by extending the panel of accessions for which response data are available. The effects of age, seed storage, and maternal growth environment all show promise for future mapping studies in appropriate recombinant inbred line populations.

The mapping results presented in Chapter 4 are in good agreement with those presented by Alonso-Blanco with respect to the *DOG1*/*GD1* region on Chromosome 5, although sequence-anchored markers predict that the peak of my region lies nearly a megabase to the right of cloned *DOG1*. The 60-week results extend this data by defining a novel epistatic region of sizable effect on

Chromosome 3 which is required for germination delay beyond the initial after-ripening period. In addition, the *Ler*-associated delay effect in 60-week old seeds has not been previously reported. Analysis of the mechanisms underlying these results will be greatly facilitated by the availability of the *DOG*1 gene and near-isogenic lines whose genotype with respect to all three loci is defined.

Ultimately, understanding of the complex relationship between selective pressure and the resulting adaptive phenotypes will rely on contributions from a diverse range of disciplines, including theoretical biology, physiology, and transmission genetics. Theoretical analyses help to clarify the nature of selection constraints and establish a probable direction of selection. The connection between these aspects of evolution is established in the realm of physiological mechanisms. The elements of these mechanisms actually affected by natural selection can then be identified through genome-wide analyses of quantitative traits. By combining the three disciplines in a deliberate and efficient way, I have identified a Mendelian component of the germination regulation pathway in *Arabidopsis* responsible for delay both within and beyond the after-ripening period. This information will add to the growing body of knowledge concerning the importance of germination inhibition at all stages of a seed's life history.

Appendices

Appendix 1: Supplement to Chapter 3

Tables of control variables, number of observations, and experimental conditions for Chapter 3

Table A1-1: Observations and Control Variables.

	Age Range (Weeks)	Total Plates	Days to Transfer Range	Total Number of Rotations	Total Number of Individual Lots
All	2-79, 86-524	324	0-189, 472-2873	14	63
Col-O	3-40, 105-121	93	0-189, 722	10	16
Cvi	2-79, 429-524	119	0-168, 2873	9	19
Ler	2-64, 86	44	0-185, 472	9	13
No-O	4-36, 237-245	68	1-157, 1639	7	15

Table A1-2: Number of Plates at Each Age Range

	Age Range (Weeks)	Plates--< 52W (364 days)	Plates--52-104 W. (364-728 days)	Plates-->104 W (728 days)
All	2-79, 86-524	234	33	57
Col-O	3-40, 105-121	72	0	21
Cvi	2-79, 429-524	74	27	18
Ler	2-64, 86	38	6	0
No-O	4-36, 237-245	50	0	18

Table A1-3: Number of Plates Observed in Each "Transfer" Range:

	Days to Transfer Range	Plates--Days to Transfer 0-63	Plates--Days to Transfer Range 63-126	Plates--Days to Transfer Range 126-189	Plates--Days to Transfer >400
All	0-189, 472-2873	182	34	48	60
Col-O	0-189, 722	58	8	6	21
Cvi	0-168, 2873	71	9	21	18
Ler	0-185, 472	20	6	15	3
No-O	1-157, 1639	33	11	6	18

Table A1-4: Number of Plates in Each Temperature Range:

To establish a correlation between germination frequency and the temperature ranges, an average value was used for the two variable temperature ranges: 18°C for the 17-19°C range and 23°C for the 21-25°C range. The 20°C value represents a constant temperature of 20°C.

	Plates--18C	Plates 20C Constant	Plates 23C
All	69	75	180
Col-O	6	12	75
Cvi	48	30	41
Ler	9	21	14
No-O	6	12	50

Table A1-5: Number of Rotations, Lots and Experiments

A "rotation" is a group of plants planted on the same day and grown in the same environment. A "lot" is a group of seeds from within a rotation.

	Total Plates	Total Number of Rotations	Total Number of Individual Lots	Total Exp'ts

All	324	14	63	42
Col-O	93	10	16	26
Cvi	119	9	19	26
Ler	44	9	13	13
No-O	68	7	15	19

Appendix 2. Supplement to Chapter 4

Relevant trait values and ages for lineages used in QTL analysis

Table A2-1: Ages of L*er*/Cvi RILs at Time of Experiment

Harvest numbers indicate the year (first two numbers), Julian date of harvest (middle three numbers), and unique sequence number for the seed lot (last two numbers). The Julian harvest date and the date of the experiment were used to calculate the seed age, in days, at the time of the experiment. "P*" in place of the sequence number means that several concurrently harvested lots of the same accession were pooled.

RIL #	Harvest Number "12 Weeks"	Harvest Number "60 Weeks"	Seed Age (in days) "12 Weeks"	Seed Age "60 Weeks"
1	03-329-P*	NA	97	--
3	03-339-09	01-130-01	87	467
4	03-351-21	01-127-68	83	477
5	NA	01-142-22	--	392
6	03-339-61	01-128-02	95	469
7	03-328-63	01-127-97	84	470
8	03-351-08	01-142-23	83	475
9	03-328-07	01-130-46	84	418
10	03-329-62	01-123-05	90	403
11	03-330-13	01-142-49	89	392
12	03-330-72	01-127-99	89	421
13	03-339-46	01-130-48	95	467
14	03-339-47	01-142-46	95	475
15	03-329-03	01-142-52	83	392
16	03-328-70	01-123-07	84	403
17	03-329-54	01-117-05	90	480
18	03-328-67	01-117-07	84	480
19	03-339-70	01-120-13	95	406
20	03-337-33	01-117-04	89	480
21	03-339-24	01-142-09	87	475
22	03-339-69	01-117-09	95	480
23	03-339-57	01-117-02	95	480
24	03-337-60	01-142-21	89	475
25	03-328-40	01-123-33	84	411
26	03-337-35	01-130-22	89	418
27	03-351-13	01-142-20	83	475

RIL #	Harvest Number "12 Weeks"	Harvest Number "60 Weeks"	Seed Age (in days) "12 Weeks"	Seed Age "60 Weeks"
28	03-339-43	01-127-60	95	477
29	03-339-08	01-117-14	87	480
30	03-328-53	01-120-14	84	406
31	03-339-07	01-142-11	87	**
32	03-330-64	01-120-10	89	406
33	03-328-35	01-142-14	84	475
34	03-328-57	01-127-62	84	477
35	03-339-53	01-120-16	95	406
36	03-339-60	01-120-11	95	406
37	03-328-44	01-142-24	84	475
38	03-351-04	NA	75	--
39	03-328-39	01-127-50	84	428
40	03-329-52	01-131-32	90	417
41	03-328-25	01-123-30	84	432
42	03-328-58	01-120-15	84	406
43	03-328-54	01-120-12	84	435
44	03-328-68	01-120-03	84	435
45	03-328-59	01-128-01	84	469
46	03-339-14	01-145-13	87	472
47	03-351-27	01-127-65	83	421
48	03-328-24	01-120-05	84	435
49	03-337-66	01-142-34	89	392
50	03-328-23	01-120-06	84	406
51	03-351-06	01-142-29	83	475
53	03-339-62	01-120-09	95	406
54	03-328-01	01-127-84	84	421
55	03-328-21	NA	84	--
57	03-339-50	01-127-67	95	477
58	03-333-P*	01-127-87	93	421
59	NA	01-129-P*	--	419
60	03-330-21	01-142-37	89	392
61	03-333-P*	01-127-74	79	470
62	04-012-02	01-142-38	57	392
63	03-339-63	01-130-45	95	467
64	03-328-62	01-127-71	84	477
65	NA	01-130-43	--	467
66	03-328-55	01-127-88	84	421
67	03-339-35	01-127-15	87	407

RIL #	Harvest Number "12 Weeks"	Harvest Number "60 Weeks"	Seed Age (in days) "12 Weeks"	Seed Age "60 Weeks"
68	03-339-68	01-142-41	95	475
69	03-328-33	01-127-77	84	421
70	NA	01-123-10	--	432
71	03-329-11	01-130-44	90	418
72	03-339-16	01-142-36	87	475
73	03-339-19	NA	87	--
74	03-339-48	01-116-04	95	439
101	03-330-20	01-116-05	89	439
102	03-328-22	01-123-28	84	403
103	03-329-21	01-116-11	90	410
104	03-329-19	01-142-50	90	392
105	03-328-11	01-116-09	84	410
106	03-339-66	01-130-15	95	418
107	03-328-36	01-130-09	84	467
108	03-337-59	01-131-03	89	466
109	03-339-36	01-116-10	87	410
110	03-328-32	01-116-14	106	410
111	03-339-45	01-131-10	95	473
112	03-351-28	01-116-12	83	410
113	03-339-72	01-116-08	95	410
114	03-328-10	01-131-16	84	473
115	03-339-51	01-131-17	95	486
116	03-328-12	01-117-11	84	480
117	03-339-71	01-117-03	95	480
118	03-351-22	NA	75	--
119	03-339-40	01-116-13	87	410
120	03-339-59	01-117-13	95	480
122	03-339-64	01-116-15	95	410
123	03-337-34	01-130-13	89	418
124	03-330-28	01-117-10	89	480
125	03-328-14	01-117-06	84	480
126	03-328-34	01-116-06	84	410
127	03-329-06	01-131-14	83	473
128	03-330-60	NA	96	--
129	03-329-20	01-117-12	90	480
130	03-339-33	01-142-25	87	475
131	03-329-01	01-123-29	83	411
132	03-339-41	01-130-07	87	467

RIL #	Harvest Number "12 Weeks"	Harvest Number "60 Weeks"	Seed Age (in days) "12 Weeks"	Seed Age "60 Weeks"
133	03-339-38	01-117-15	87	480
134	03-329-40	01-123-31	90	411
135	03-329-42	01-131-23	90	417
136	03-339-37	01-116-02	87	439
137	03-329-10	01-123-36	83	411
138	03-339-44	NA	95	--
139	03-337-18	NA	89	--
140	03-337-01	01-142-44	89	475
141	03-329-36	NA	97	--
142	03-329-43	01-123-35	90	411
143	03-328-13	01-127-90	84	421
144	03-339-13	01-145-06	87	472
145	NA	01-123-32	--	411
146	03-351-19	NA	83	--
147	03-329-53	01-123-09	90	403
148	03-330-15	01-123-08	89	403
149	03-329-04	01-120-07	83	406
150	03-337-17	01-127-89	89	421
151	03-329-02	01-120-04	83	435
152	03-328-29	01-127-85	84	421
153	03-329-30	01-127-93	90	421
154	03-330-06	01-127-94	89	470
155	03-351-02	01-145-09	83	472
156	03-339-11	01-145-25	87	472
157	03-329-50	01-127-81	90	421
158	03-339-30	01-127-83	87	421
159	03-329-12	01-123-12	90	432
160	03-330-29	01-127-16	89	407
161	03-328-03	01-145-24	84	472
162	03-330-34	01-127-92	89	421
163	03-337-02	01-127-13	89	407
164	03-328-15	01-127-91	84	421
165	03-339-49	01-127-86	95	421
166	03-337-63	01-127-80	89	421
167	03-339-67	01-142-27	95	475
168	03-329-29	01-123-11	90	432
169	03-330-19	01-131-15	89	473
170	NA	01-145-05	--	472

RIL #	Harvest Number "12 Weeks"	Harvest Number "60 Weeks"	Seed Age (in days) "12 Weeks"	Seed Age "60 Weeks"
171	NA	01-127-57	--	477
172	03-329-08	01-127-55	83	477
173	03-328-18	01-127-53	84	477
174	03-351-14	01-131-29	83	486
175	03-329-69	01-116-01	90	439
176	03-330-36	01-127-54	89	477
177	03-329-18	01-127-52	90	428
178	03-330-12	01-142-08	89	475
179	03-329-41	01-142-05	90	392
180	03-351-29	01-127-56	83	477
182	03-337-49	01-131-12	89	473
183	03-339-54	01-127-58	95	477
184	03-339-56	01-127-61	95	477
185	03-328-45	01-123-39	84	474
186	03-328-02	01-142-04	84	392
187	NA	NA	--	--
188	03-351-15	01-142-26	83	475
189	NA	01-117-08	--	480
190	03-329-17	01-127-63	90	477
191	03-330-25	01-127-64	89	477
192	03-329-31	01-127-51	90	428
193	03-330-16	01-127-59	89	477

NA = not available in planting -- = missing age

Table A2-2: RIL Line Comparison Table

Ranking of individual lines with respect to 48-hour germination frequency in 12- and 60-week mapping experiments. "Bin" column is placement in the distribution. Lines in the "both" column were in the same bin in both experiments.

Bin	Both 12 and 60 Week Lines	12 Week Lines	60 Week Lines
0-9.9	10	21	43
	19	24	53
	25	31	101
	74	46	149
	122	49	157

	159	57	
	164	58	
		62	
		117	
		136	
		142	
		143	
		160	
		169	
10-19.9		22	32
		37	
		184	
20-29.9		32	102
		190	142
			151
30-39.9		27	14
		109	
		144	
		177	
40-49.9	30	3	160
	54	16	166
		34	
		114	
		132	
		133	
		161	
		167	
		188	
50-59.9		14	36
		29	58
		43	115
		63	131
		67	143
		108	169
		124	192
		173	
		178	
		180	
		193	
60-69.9	20	8	9
	106	15	22
	152	36	31
		41	45

		53	46
		101	57
		110	120
		111	132
		123	133
		130	155
		131	177
		137	
		156	
		157	
		179	
		183	
70-79.9	4	6	3
	7	9	8
	17	11	21
	23	12	24
	40	13	26
	47	18	27
	50	28	29
	69	35	33
	71	44	34
	125	68	37
	134	72	39
	165	102	48
	175	103	61
	182	104	107
	191	112	108
		113	110
		115	114
		119	116
		120	117
		126	123
		127	124
		129	136
		147	137
		148	158
		151	167
		155	168
		162	184
		163	185
		166	188
		172	190
		174	

		176	
80-89.9	42	26	6
		33	16
		45	44
		51	63
		60	72
		64	112
		107	127
		135	129
		149	130
		168	140
		186	144
		192	147
			153
			156
			162
90+	66	39	11
	105	48	12
	150	61	13
	154	116	15
		140	18
		153	28
		158	35
		185	41
			49
			51
			60
			62
			64
			67
			68
			103
			104
			109
			111
			113
			119
			126
			135
			148
			161
			163
			172

			173
			174
			176
			178
			179
			180
			183
			186
			193

Table A2-3: Table of F_1 and F_2 Populations Used in Segregation Analysis

Parent / Cross Type	Generation	Seed Age at Experiment	Seed Lot
Cvi	P	130	04-132
Cvi	P	32	04-230
Cvi	P	14	05-024-08
Ler	P	158	04-104
Ler	P	32	04-230
Ler	P	14	05-024-01
Cv X L	F_1	142	04-120
L X Cv	F_1	142	04-120
Cv X L	F_2 (#1-5)	32	04-230
L X Cv	F_2 (#1-5)	32	04-230

References

Akalehiywot, T. and B. J. D (1977). "Promotion and synchronization of cereal grain germination by osmotic pretreatment with polyethylene glycol." J. Agric. Sci. Comb **90**: 503-506.

Alexander, H. M. and A. M. Schrag (2003). "Role of soil seed banks and newly dispersed seeds in the population dynamics of the annual sunflower, *Helianthus annuus*." *Journal of Ecology* **91**: 987-998.

Allen, P. S. and S. E. Meyer (1998). "Ecological aspects of seed dormancy loss." Seed Science Research **8**: 183-191.

Allen, P. S., S. E. Meyer, et al. (2000). Hydrothermal Time as a Tool in Comparative Germination Studies. Seed Biology: Advances and Applications. M. Black, K. J. Bradford and J. Vazquez-Ramos. Oxon, UK, CABI Publishing: 401-410.

Alonso-Blanco, C., L. Bentsink, et al. (2003). "Analysis of Natural Allelic Variation at Seed Dormancy Loci of *Arabidopsis thaliana*." Genetics **164**: 711-729.

Alonso-Blanco, C., H. Blankestun-De-Vries, et al. (1999). "Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*." Proceedings of the National Academy of Science USA **96**: 4710-4717.

Alonso-Blanco, C., S. El-Din El-Assal, et al. (1998). "Analysis of Natural Allelic Variation at Flowering Time Loci in Landsberg *erecta* and Cape Verde Islands Ecotypes of *Arabidopsis thaliana*." Genetics **149**(2): 749-764.

Alonso-Blanco, C., A. J. M. Peeters, et al. (1998). "Development of an AFLP based linkage map of L*er,* Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a L*er*/Cvi recombinant inbred line population." The Plant Journal **14**(2): 259-271.

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment." Journal of Molecular Biology **215**((3)): 403-10.

Bahulikar, R. A., D. Stanculescu, et al. (2004). "ISSR and AFLP analysis of the temporal and spatial population structure of the post-fire annual, Nicotiana attenuata, in SW Utah." BMC Ecology **4**(12).

Bauer, M. C., S. E. Meyer, et al. (1998). "A simulation model to predict seed dormancy loss in the field for *Bromus tectorum* L." Journal of Experimental Botany **49**(324): 1235-1244.

Benech-Arnold, R. L., S. Enciso, et al. (2000). Involvement of ABA and GAs in the Regualtion of Dormancy in Developing Sorghum Seeds. Seed Biology: Advances and Applications. M. Black, K. J. Bradford and J. Vazquez-Ramos. Oxon, UK, CABI Publishing: 101-112.

Bentsink, L. (2002). Genetic analysis of seed dormancy and seed composition in *Arabidopsis thaliana* using natural variation, Wageningen University: 159.

Bentsink, L., J. Jowett, et al. (2006). "Cloning of DOG1 a quantitative-trait locus controlling seed dormancy in *Arabidopsis*." Proc. Nat. Acad. Science USA **103**(45): 17042-17047.

Bewley, J. D. (1997). "Seed Germination and Dormancy." The Plant Cell **9**(7): 1055-1066.
Botto, J. F., R. A. Sanchez, et al. (1998). "Burial conditions affect light responses of Datura ferox seeds." Seed Science Research **8**(4): 423-429.
Bradbeer, J. W. (1988). Seed Dormancy and Germination. New York, Chapman and Hall.
Bradford, K. J. (1990). "A Water Relations Analysis of Seed Germination Rates." Plant Physiology **94**: 840-849.
Bradford, K. J. (1997). The Hydrotime concept in Seed Germination and Dormancy. Basic and Applied Aspects of Seed Biology. R. H. Ellis, M. Black, A. J. Murdoch and T. D. Hong. Dordrecht, Kluwer Academic Publishers: 349-360.
Campalans, A., R. Messeguer, et al. (1999). "Plant responses to drought, from ABA signal transduction events to the action of the induced proteins." Plant Physiology and Biochemistry **37**(5): 327-340.
Capon, S. J. and M. A. Brock (2006). "Flooding, soil seed bank dynamics and vegetation resilience of a hydrologically variable desert floodplain." Freshwater Biology **51**(2): 206-223.
Chase, K., F. R. Adler, et al. (1997). "Epistat: a computer program for identifying and testing interactions between pairs of quantitative trait loci." Theoretical and Applied Genetics **94**: 724-730.
Chen, J.-G., S. Pandey, et al. (2004). "GCR1 Can Act Independently of Heterotrimeric G-Protein in Response to Brassinosteroids and Gibberellins in Arabidposis Seed Germination." Plant Physiol. **135**: 907-915.
Chen, J.-G., H. Ullah, et al. (2001). "ABP1 is required for organized cell elongation and division in *Arabidopsis* embryogenesis." Genes & Development **15**(7): 902-911.
Chen, Y. and J. S. Burris (1990). "Role of Carbohydrates in Dessication Tolerance and Membrane Behavior in Maturing Maize Seed." Crop Science **30**: 971-975.
Cheng, Z. and K. J. Bradford (1999). "Hydrothermal time analysis of tomato seed germination responses to priming treatments." Journal of Experimental Botany **50**(330): 89-99.
Chono, M., I. Honda, et al. (2006). "Field studies on the regulation of abscisic acid content and germinability during grain development of barley: molecular and chemical analysis of pre-harvest sprouting." Journal of Experimental Botany **57**(10): 2421-2434.
Christensen, M., S. Meyer, et al. (1996). "A hydrothermal time model of seed after-ripening in *Bromus tectorum* L." Seed Science Research **6**: 155-163.
Claus, M. J. and D. L. Venable (2000). "Seed Germination in Desert Annuals: An Empirical Test of Adaptive Bet-Hedging." American Naturalist **155**(2): 168-186.
Clerkx, E. J., H. Blankestijn-De Vries, et al. (2004). "Genetic differences in seed longevity of various *Arabidopsis* mutants." Physiologia plantarum **121**: 448-461.

Clerkx, E. J. M., M. El-Lithy, et al. (2004). "Analysis of Natural Allelic Variation of Arabidopsis Seed Germination and Seed Longevity Traits between the Accessions Landsberg *erecta* and Shakdara, Using a New Recombinant Inbred Line Population." Plant Physiology **135**: 432-443.

Comai, L. and J. J. Harada (1990). "Transcriptional activities in dry seed nuclei indicate the timing of the transition from embryogenesis to germination." Proceedings of the National Academy of Sciences of the United States of America **87**: 2671-2674.

Corbineau, F., N. Ozbingol, et al. (2000). Improvement to Tomato Seed Germination by Osmopriming as Related to Energy Metabolism. Seed Biology: Advances and Applications. M. Black, K. J. Bradford and J. Vazquez-Ramos. Oxon, UK, CABI Publishing: 467-476.

Curaba, J., T. Moritz, et al. (2004). "AtGA3ox2, a Key Gene Responsible for Bioactive Gibberllin Biosynthesis, Is Regulated during Embryogenesis by LEAFY COTYLEDON2 and FUSCA3 in Arabidopsis." Plant Physiol. **136**: 3660-3669.

Dean, A. (2006). "On a Chromosome Far, Far Away: LCRs and Gene Expression." Trends in Genetics **22**(1): 38-45.

Debeaujon, I. and M. Koornneef (2000). "Gibberellin Requirement for Arabidopsis Seed Germination Is Determined Both by Testa Characteristics and Embryonic Abscisic Acid." Plant Physiology **122**: 415-424.

Debeaujon, I., K. M. Leon-Kloosterziel, et al. (2000). "Influence of the Testa on Seed Dormancy, Germination, and Longevity in *Arabidopsis*." Plant Physiology **122**: 403-413.

Derkx, M. P. M. and C. M. Karssen (1993). "Effects of light and temperature on seed dormancy and gibberellin-stimulated germination in Arabidopsis thaliana: studies with gibberellin-deficient and -insensitive mutants." Physiologia Plantarum **89**(2): 360-368.

Derkx, M. P. M. and C. M. Karssen (1993). "Variability in light, gibberellin and nitrate requirement of Arabidopsis thaliana seeds due to harvest time and conditions of dry storage." Journal of Plant Physiology **141**(5): 574-582.

Derkx, M. P. M. and C. M. Karssen (1994). "Are seasonal dormancy patterns in Arabidopsis thaliana regulated by changes in seed sensitivity to light, nitrate and gibberellin?" Annals of Botany **73**(2): 129-136.

Derkx, M. P. M., E. Vermeer, et al. (1994). "Giberellins in seeds of *Arabidopsis thaliana*: biological activities, identification and effects of light and chilling on endogenous levels." Plant Growth Regulation **15**: 223-234.

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

Donohue, K., L. Dorn, et al. (2005). "Environmental and Genetic Influences on the Germination of *Arabidopsis thaliana* in the Field." Evolution **59**(4): 740-757.

Donohue, K., L. Dorn, et al. (2005). "The Evolutionary Ecology of Seed Germination of *Arabidopsis Thaliana*: Variable Natural Selection on Germination Timing." Evolution **59**(4): 758-770.

Dyer, W. E. (1995). "Exploiting weed seed dormancy and germination requirements through agronomic practices." Weed Science **43**(3): 498-503.

Federici, M. T., D. Vaughan, et al. (2001). "Analysis of Uruguayan weedy rice genetic diversity using AFLP molecular markers." EJB Electronic Journal of Biotechnology **4**(3): 130-145.

Fenner, M. (1985). Seed Ecology. New York, Chapman and Hall.

Ferrandis, P., J. M. Herranz, et al. (1996). "The Role of Soil Seed Bank in the Early Stages of Plant Recovery After Fire in a *Pinus pinaster* Forest in SE Spain." *Int. J. Wildland Fire* **6**(1): 31-35.

Finkelstein, R. R., M. L. Wang, et al. (1998). "The Arabidopsis abscisic acid response locus ABI4 encodes an APETALA2 domain protein." The Plant Cell **10**(6): 1043-1054.

Flores, E. M. (2002). Seed Biology. Tropical Tree Seed Manual. J. A. Vozzo. Washington, D.C., United States Department of Agriculture Forest Service: 13-42.

Garvin, S. C. and S. F. Meyer (2003). "Multiple mechanisms for seed dormancy regulation in shadescale (Atriplex confertifolia Chenopodiaceae)." Canadian Journal of Botany **81**: 601-610.

Gilmour, S. J., N. N. Artus, et al. (1992). "cDNA sequence analysis and expression of two cold-regulated genes of Arabidopsis thaliana." Plant Molecular Biology **18**(1): 13-21.

Gosti, F., N. Beaudoin, et al. (1999). "ABI1 Protein Phosphatase 2C Is a Negative Regulator of Abscisic Acid Signaling." The Plant Cell **11**(10): 1897-1909.

Grinikh, L. I., V. V. Shevchenko, et al. (1983). "Effect of high temperature and absence of light on germination of the seeds of different Arabidopsis thaliana lines." Arabidopsis Information Management Service(20): 109-114.

Groot, S. P. C., A. H. M. van der Geest, et al. (2000). Molecular Genetic Analysis of *Arabidopsis* Seed Quality. Seed Biology: advances and applications. M. Black, K. J. Bradford and J. Vazquez-Ramos. Oxon, UK, CABI Publishing: 123-132.

Gualberti, G., M. Papi, et al. (2002). "Mutations in the Dof Zinc Finger Genes DAG2 and DAG1 Influence with Opposite Effects the Germination of Arabidopsis Seeds." The Plant Cell **14**: 1253-1263.

Gutierrez, L., O. Van Wuytswinkel, et al. (2007). "Combined networks regulating seed maturation." Trends in Plant Science **12**(7): 294-300.

Haigh, A. and E. Barlow (1987). "Water Relations of Tomato Seed Germination." Australian Journal Plant Physiology **14**: 485-492.

Halpern, C. B., S. A. Evans, et al. (1999). "Soil seed banks in young, closed-canopy forests of the Olympic Peninsula, Washington: potential contributions to understory reinitiation." Canadian Journal of Botany **77**: 922-935.

Hampton, J. G. and D. M. TeKrony, Eds. (1995). Handbook of Vigour Test Methods. Zurich, The International Seed Testing Association.
Harada, H., T. Kuromori, et al. (2004). "Quantitative Trait Loci Analysis of Nitrate Storage in Arabidopsis Leading to an Investigation of the Contribution of the Anion Channel Gene AtCLC-c to Variation in Nitrate Levels." Journal of Experimental Botany **55**(405): 2005-2014.
Heath, J. D., R. Weldon, et al. (1986). "Analysis of storage proteins in normal and aborted seeds from embryo-lethal mutants of *Arabidopsis thaliana*." Planta **169**: 304-312.
Hewitt, J. (2007). Study of Arabidopsis thaliana Seed Development: Occurrence of Germinability. Dept. of Horticulture, Oregon State University: 66.
Huh, M. K. and H. W. Huh (2001). "AFLP Fingerprinting of Brassica campestis L. ssp. napus var. nippo-oleifera Makino from Korea." Korean Journal of Biological Science **5**: 101-106.
Koornneef, M., L. Bentsink, et al. (2000). "Natural variation, an unexploited source of genes in Arabidopsis."
Koornneef, M., C. J. Hanhar, et al. (1989). "*In Vivo* Inhibition of Seed Development and Reserve Protein Accumulation in Recombinants of Abscisic Acid Biosynthesis and Responsiveness mutants in *Arabidopsis thaliana*." Plant Physiology **90**: 463-469.
Koornneef, M. and C. M. Karssen (1994). Seed Dormancy and Germination. Arabidopsis. E. M. Meyerowitz and C. R. Somerville. New York, Cold Spring Harbor Laboratory Press: 313-334.
Koornneef, M., K. M. Leon-Kloosterziel, et al. (1998). "The genetic and molecular dissection of abscisic acid biosynthesis and signal transduction in *Arabidopsis*." Plant Physiology and Biochemistry **36**(1-2): 83-89.
Koster, K. L. and A. C. Leopold (1988). "Sugars and Dessication Tolerance in Seeds." Plant Physiology **88**: 829-832.
Kucera, B., M. A. Cohn, et al. (2005). "Plant hormone interactions during seed dormancy release and germination." Seed Science Research **15**: 281-307.
Lefebvre, V., H. North, et al. (2006). "Functional analysis of Arabidopsis NCED6 and NCED9 genes indicates that ABA synthesized in the endosperm is involved in the induction of seed dormancy." The Plant Journal **45**: 309-319.
Leishman, M. R., I. J. Wright, et al. (2000). The Evolutionary Ecology of Seed Size. *Seeds: The Ecology of Regeneration in Plant Communities*. M. Fenner, CABI Publishing: 31-57.
Leon-Kloosterziel, K. M., G. A.-v. d. Bunt, et al. (1996). "Arabidopsis mutants with a reduced seed dormancy." Plant Physiology **110**(1): 233-240.
Li, Y., L. Jones, et al. (2003). "Expansins and cell growth." Curr. Opin. Plant **6**((6)): 603-10.
Liang, C., P. Jaiswal, et al. (2007). "Gramene: a growing plant comparative genomics resource." Nucleic Acids Research **PMID 17984077**.

Loprz, M., J. M. Humaara, et al. (1999). "The effect of temperature and water stress on laboratory germination of Eucalyptus globulus Labill. seeds of different sizes." Annual Forum for Science **57**(2000): 245-250.
Lorthe, C. J. and R. Turkington (2002). "The small-scale spatio-temporal pattern of a seed bank in ghe Negev Desert, Israel." EcoScience **9**(3): 407-413.
Loudet, O., S. Chailleu, et al. (2003). "Quantitative Trait Loci Analysis of Nitrogen Use Efficiency in *Arabidopsis*." Plant Physiology **131**: 345-325.
Lynch, M. and B. Walsh (1998). Genetics and Analysis of Quantitative Traits. Sunderland, MA, Sinauer Associates.
Mahy, G., X. Vekemans, et al. (1999). "Patterns of allozymic variation within Calluna vulgaris populations at seed bank and adult stages doi:10.1046/j.1365-2540.1999.00499.x." Heredity **82**(4): 432-440.
Malmberg, R. L., S. Held, et al. (2005). "Epistasis for Fitness-Related Quantitative Traits in *Arabidopsis thaliana* Grown in the Field in the Greenhouse." Genetics.
Mapes, G., G. W. Rothwell, et al. (1989). "Evolution of seed dormancy." Nature **337**: 645-646.
McCue, K. and T. Holtsford (1998). "Seed bank influences on genetic diversity in the rare annual Clarkia springvillensis (Onagraceae)." Am. J. Bot. **85**(1): 30-.
McCullaugh, P. and J. A. Nelder (1989). Generalized Linear Models. New York, Chapman and Hall/CRC.
Membré, N., A. Berna, et al. (1997). "cDNA sequence, genomic organization and differential expression of three Arabidopsis genes for germin/oxalate oxidase-like proteins." Plant Molecular Biology **35**((4)): 459-69.
Mena, M., F. J. Cejudo, et al. (2002). "A Role for the DOF Transcription Factor BPBF in the Regulation of Gibberellin-Responsive Genes in Barley Aleurone." Plant Physiol. **130**: 111-119.
Morris, A., R. Baucom, et al. (2000). "Stratified Analysis of the Soil Seed Bank in the Cedar Glade Endemic, *Astragalus bibullatus:* Evidence for Historical Changes in Genetic Structure." American Journal of Botany **89**: 29-36.
Munir, J., L. A. Dorn, et al. (2001). "The Effect of Maternal Photoperiod on Seasonal Dormancy in *Arabidopsis thaliana* (Brassicaceae)." American Journal of Botany **88**(7): 1240-1249.
Murashige and Skoog (1962). "A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Culture." Physiologia plantarum **15**: 473-497.
Nambara, E., H. Kawaide, et al. (1998). "Characterization of an Arabidopsis thaliana mutant that has a defect in ABA accumulation: ABA-dependent and ABA-independent accumulation of free amino acids during dehydration." Plant Cell Physiology **39**(8): 853-858.
Ni, B.-R. and K. J. Bradford (1993). "Germination and Dormancy of Abscisic Acid- and Gibberellin-Deficient Mutant Tomato (Lycopersicon esculentum) Seeds (Sensitivity of Germination to Abscisic Acid, Gibberellin, and Water Potential)." Plant Physiology **101**(2): 607-617.

Nicolas, C., G. Nicolas, et al. (1996). "Antagonistic effects of abscisic acid and gibberellic acid on the breaking of dormancy of *Fagus sylvatica* seeds." Physiologia Plantarum **96**(2): 244-250.
Nicolas, C., D. Rodriguez, et al. (1997). "The Expression of an Abscisic Acid-Responsive Glycine-Rich Protein Coincides with the Level of Seed Dormancy in *Fagus sylvatica*." Plant Cell Physiology **38**(12): 1303-1310.
Obroucheva, N. V. (1999). Seed Germination: A guide to the early stages. Leiden, Backhuys Publishers.
Okamoto, M., A. Kuwahara, et al. (2006). "CYP707A1 and CYP707A2, Which Encode Abscisic Acid 8'-Hydroxylases, Are Indispensable for Proper Control of Seed Dormancy and Germination in Arabidopsis." Plant Physiol. **141**: 97-107.
Papi, M., S. Sabatini, et al. (2002). "Inactivation of the Phloem-Specific Dof Zinc Finger Gene *DAG1* Affects Response to Light and Integrity of the Testa of Arabidopsis Seeds
10.1104/pp.010488." Plant Physiol. **128**(2): 411-417.
Papi, M., S. Sabatini, et al. (2000). "Identification and disruption of an Arabidopsis zinc finger gene controlling seed germination." Genes & Development **14**(1): 28-33.
Parcy, F., C. Valon, et al. (1997). "The *ABSCISIC ACID-INSENSITIVE3, FUSCA3,* and *LEAFY COTYLEDON1* Loci Act in Concert to Control Multiple Aspects of Arabidopsis seed development." The Plant Cell **9**(8): 1265-1277.
Parcy, F., C. Valon, et al. (1994). "Regulation of Gene Expression Programs during Arabidopsis Seed Development: Roles of the *ABI3* Locus and of Endogenous Abscisic Acid." The Plant Cell **6**(11): 1567-1582.
Penfield, S., E.-M. Josse, et al. (2005). "Cold and Light Control Seed Germination through the bHLH Transcription Factor SPATULA." Current Biology **15**: 1998-2006.
Peters, J. L. (2009). AFLP markers. L. Lipovich. Detroit.
Peters, J. L., H. Constandt, et al. (2001). "A physical amplified fragment-length polymorphism map of Arabidopsis." Plant Physiol. **127**: 1579-1589.
Phillips, J., O. Artsaenko, et al. (1997). "Seed-specific immunomodulation of abscisic acid activity induces a developmental switch." The EMBO Journal **16**(15): 4489-4496.
Rauh, B. L., C. Bates, et al. (2002). "Quantitative trait loci analysis of growth response to varying nitrogen levels in *Arabidopsis thaliana*." Theoretical and Applied Genetics **101**: 743-750.
Reyes, O. and M. Casal (2001). "The influence of seed age on germinative response to the effects of fire in Pinus pinaster, Pinus radiata and Eucalyptus globulus." Annals of Forest Science **58**: 439-447.
Robertson, M., S. M. Swain, et al. (1998). "Identification of a negative regulator of gibberellin action HvSPY, in barley." The Plant Cell **10**(6): 995-1007.
Russell, L., V. Larner, et al. (2000). "The *Arabidopsis COMATOSE* locus regulates germination potential." Development **127**(17): 3759-3767.

Saito, S., N. Hirai, et al. (2004). "Arabidopsis CYP707As Encode (+)-Abscisic Acid 8'-Hydroxylase, a Key Enzyme in the Oxidative Catabolism of Abscisic Acid." Plant Physiol. **134**: 1439-1449.

Schmidt, K. and K. Jensen (2000). "Genetic structure and AFLP variation of remnant populations in the rare plant Pedicularis palustris (Scrophulariaceae) and its relation to population size and reproductive components." American Journal of Botany **87**(5): 678-689.

Schneider, W. L. and D. J. Gifford (1994). "Loblolly pine seed dormancy. I. The relationship between protein synthesis and the loss of dormancy." Physiologia plantarum **90**(2): 246-252.

Smith, M. T., B. S. P. Wang, et al. (2002). Dormancy and Germination. Tropical Tree Seed Manual. J. A. Vozzo. Washington, D.C., United States Department of Agriculture Forest Service: 149-176.

Swarup, K., C. Alonso-Blanco, et al. (1999). "Natural allelic variation identifies new genes in the *Arabidopsis* circadian system." The Plant Journal **20**(1): 67-77.

To, A., C. Valon, et al. (2006). "A Network of Local and Redundant Gene Regulation Governs Arabidopsis Seed Maturation." The Plant Cell **18**: 1642-1651.

Toorop, P. T., R. M. Barroco, et al. (2005). "Differentially expressed genes associated with dormancy or germination of *Arabidopsis thaliana* seeds." Planta **221**: 637-647.

van der Schaar, W., C. Alonso-Blanco, et al. (1997). "QTL analysis of seed dormancy in *Arabidopsis* using recombinant inbred lines and MQM mapping." Heredity **79**: 190-200.

Venable, D. L. (1989). Modeling the Evolutionary Ecology of Seed Banks. Ecology of Soil Seed Banks, Academic Press, Inc.: 67-87.

Venable, D. L., E. Dyerson, et al. (1995). "Population dynamic consequences and evolution of seed traits in Heterosperma pinatum (Asteraceae)." American Journal of Botany **82**(3): 410-420.

Venable, D. L., E. Dyerson, et al. (1998). "Seed Morphometrics and Adaptive Geometric Differentiation." Evolution **52**(2344-354).

Wang, H., L. V. Caruso, et al. (2004). "The Embryo MADS Domain Protein AGAMOUS-Like 15 Directly Regulates Expression of a Gene Encoding an Enzyme Involved in Gibberellin Metabolism." The Plant Cell **16**: 1206-1219.

Warren, J. E. and M. A. Bennett (2000). Bio-osmopriming Tomato (*Lycopersicon esculentum* Mill.) Seeds for Improved Seedling Establishment. Seed Biology: Advances and Applications. M. Black, K. J. Bradford and J. Vazquez-Ramos. Oxon, UK, CABI Publishing: 477-488.

Weining, C., L. A. Dorn, et al. (2003). "Heterogeneous Selection at Specific Loci in Natural Environments in *Arabidopsis thaliana*." Genetics **165**: 321-329.

Weining, C., M. c. Ungerer, et al. (2002). "Novel Loci Control Variation in Reproductive Timing in *Arabidopsis thaliana* in Natural Environments." Genetics **162**: 1875-1884.

West, M. A. L., K. M. Yee, et al. (1994). "*LEAFY COTYLEDON1* is an essential regulator of late embryogensis and cotyledon identity in *Arabidopsis*." The Plant Cell **6**(12): 1731-1745.

Westerman, J. M. and M. J. Lawrence (1970). "Genotype-environment interaction and developmental regulation in *Arabidopsis thaliana*." Heredity **25**: 609-627.

BIOGRAPHY

Anne Harlan Prather was the eldest of five children and the only girl. As a result, she learned to teach, lead, and survive at the hands of her younger brothers--who never let her take the easy way out. She holds a Bachelor of Music Performance degree in classical guitar from Wheaton College. She has been, at various times, a musician, artist, storyteller, teacher, and scientist. She has recorded four albums, published a classical song cycle, written a workbook for college-level general genetics courses, and published magazine articles on the subject of telecommunications. During her research career, she received the Mary Gates Undergraduate Research Scholarship, the NASA Space Grant research fellowship, and the Howard Hughes research fellowship. As a member of the GK-12 Fellowship in Mathematics, she began work on developing a method for keeping math journals as an enhancement to upper primary students' daily practice in mathematics.