Genetic analysis of seed dormancy and seed composition in *Arabidopsis thaliana* using natural variation

Genetische analyse van kiemrust en de samenstelling van zaden in *Arabidopsis thaliana* gebruikmakend van natuurlijke variatie

Promotor: Prof. Dr. Ir. M. Koornneef
Persoonlijk hoogleraar bij het Laboratorium voor Erfelijkheidsleer, Wageningen Universiteit

Promotiecommissie: Prof. Dr. L.H.W. van der Plas, Wageningen Universiteit
Prof. Dr. R.G.F. Visser, Wageningen Universiteit
Prof. Dr. J.C.M. Smeekens, Universiteit Utrecht
Dr. D. Vreugdenhil, Wageningen Universiteit

Leónie Bentsink

Genetic analysis of seed dormancy and seed composition in *Arabidopsis thaliana* using natural variation.

Proefschrift
Ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. Dr.Ir. L. Speelman,
in het openbaar te verdedigen
op woensdag 25 september 2002
des namiddags te vier uur in de Aula.

ISBN 90-5808-705-0

Contents

Chapter 1

General Introduction

Arabidopsis as a model plant

Arabidopsis thaliana is a small weed that belongs to the mustard family (*Brassicaceae*). Arabidopsis has become a major model system for plant biology because it allows combining genetics with molecular biology. It is not only a convenient model for plant biology but also for addressing fundamental questions of biological structure and function common to all plants and often eukaryotes in general (Meinke et al., 1998; Meyerowitz, 2002). Arabidopsis has many advantages for genetic research; it is a small self-fertilising species with a short life cycle, which is completed within 6 weeks for the commonly used laboratory strains. Furthermore, it has one of the smallest genomes among higher plants, approximately 130 megabase in size divided over 5 chromosomes. Its genome has been completely sequenced (Arabidopsis Genome Initiative, 2000) and contains approximately 26.000 genes.

Arabidopsis has a broad natural distribution throughout the Northern hemisphere (Hoffman, 2002) from northern Scandinavia to Africa, including the Cape Verde Islands at 16° latitude. It has been found from sea level up to high in the Himalayas. Many different accessions have been collected from natural populations and are available for experimental analysis. The accessions Columbia (Col) and Landsberg *erecta* (L*er*) are accepted standards for genetics and molecular studies.

Natural variation

For a long time, the functional analysis of Arabidopsis genes and the genetic dissection of complex traits was based largely on the phenotypic characterisation of mutants selected by forward and reverse genetics in a few laboratory ‘wild-type’ genotypes. However, currently naturally occurring populations of Arabidopsis are more and more used as an alternative or supplementary source to generating laboratory-induced mutants (Alonso-Blanco and Koornneef, 2000). In the earliest stages of Arabidopsis research, the phenotypic characterisation of plants collected from different geographical regions revealed considerable genetic variation (Laibach, 1943, 1951; Kugler 1951; Langridge and Griffing, 1959, reviewed by Rédei, 1970). Arabidopsis is predominantly a self-fertilising species and therefore most collected plants represent inbred lines that are practically homozygous. These wild homozygous lines are referred to as accessions but also have been called ecotypes.

Quantitative trait loci mapping

Most of the variation among accessions is of a quantitative nature, which is the result of allelic variation at several loci (multi-genic) and the effect of environmental factors on the expression of the traits. This quantitative variation can be exploited using quantitative trait loci (QTL) mapping (for review see Doerge, 2002). QTL mapping requires the generation of a segregating population and its characterisation with molecular markers (i.e. the obtainment of its genome-wide genetic map). After scoring the trait of interest, associations between the genotypes of the marker loci and the phenotypes of the trait are searched for by means of specific statistical methods. These methods use the information from flanking markers, resulting in the indirect estimation of the position and effect of QTL along the different linkage groups.

The quality of the QTL mapping is affected by a few manipulable experimental parameters, that are the size and type of mapping population, the number of markers and the statistical method that is used. The heritability of individual QTL, which is another important factor for the detection of QTL can be enhanced by minimising the environmental variation and by analysing several replications of each individual genotype. The statistical methods currently used such as composite interval mapping (Zeng, 1993) and similarly multiple QTL mapping (Jansen, 1993) allow more accurate mapping than interval mapping. Both methods extend the ideas of interval mapping by including additional markers as cofactors (outside a defined window of analysis) for the purpose of removing the variation that is associated with other (linked and independent) QTL in the genome.

Recombinant inbred lines

In Arabidopsis, relatively large populations can be grown in a small area, under controlled environments, and mapping populations can be obtained in a relatively short period of time. Although any type of mapping population can be used, RILs offer unique advantages (Burr and Burr, 1991), mainly because they are homozygous and thereby allow the analysis of replications per genotype. Well-characterized RIL populations are genetically permanent or immortal and can be used indefinitely without further genotyping, allowing their simultaneous analysis by any laboratory. Furthermore, the availability of many molecular markers makes it relatively easy to generate genome-wide genetic maps. QTL mapping experiments have been performed for various traits in Arabidopsis, mainly using three RIL populations, which are L*er*/Col, L*er*/Cape Verde Islands (Cvi) and Col-5/Niederzenz 1. During the study described in this thesis we have used the L*er*/Cvi RIL population. This population, which has been characterized for amplified fragment length polymorphism (AFLP™) and cleaved amplified polymorphic markers (CAPS) by Alonso-Blanco et al. (1998b) was already analysed for several traits (Alonso-Blanco and Koornneef, 2000). QTL analysis for seed dormancy/germination, seed oligosaccharide content, seed storability and phytate and phosphate content in seeds and leaves is described in this thesis (Chapter 3, 6 and 7).

Near isogenic Lines

To characterize an individual QTL it preferentially must be separated from the rest of the segregating loci. Commonly, this process is referred to as 'Mendelisation' of a QTL. The 'Mendelisation' of a QTL is best accomplished by constructing introgression or near isogenic lines (NILs), ideally differing only for the alleles in a small genomic region, spanning a few cM around the QTL of interest. NILs can be developed by repeated backcrossing of, either an F1 plant or a genotype derived from this F1 to one of the parents. NILs that show a monogenic segregation after being crossed with the recurrent parent can be compared to this recurrent (wild type) parent, which enables the phenotypic and genetic characterisation of a QTL in a similar way to that performed with mutants. Such a genetic characterisation includes the analysis of the dominance relationships between the two alleles under study and, in some cases, complementation tests between QTL alleles and known mutant alleles at candidate loci. In addition, NILs allow further fine mapping and map-based cloning of the locus (Alonso-Blanco and Koornneef, 2000; Remington et al., 2001).

Seed development in Arabidopsis

The seeds of Arabidopsis are produced in fruits known as siliques. A typical silique produced by a plant grown under optimal conditions contains 40-60 seeds. Seed development in Arabidopsis is a process that takes 2-3 weeks, dependent on growth conditions. One of the most striking features of Arabidopsis seeds is their small size, wild type seeds being approximately 0.5 mm long at maturity and having a dry weight of about 20-30 µg (Alonso-Blanco et al., 1999). Seed expansion occurs during the first 3-4 days after pollination. This implies that a seed with an embryo at the heart stage has the same size as mature seeds prior to desiccation (reviewed by Meinke, 1994).

Seed development in Arabidopsis can be divided into two major phases: an initial phase characterized by cell division and morphogenesis of the embryo, endosperm and seed coat layers and is referred to as embryogenesis; a subsequent phase referred to as maturation phase which involves the accumulation of storage materials in the fully grown embryo and its preparation for dormancy and germination

(Meinke, 1994). This introduction will mainly focus on the latter since most characters studied in this thesis are established during this phase. A schematic overview of the processes that take place during seed development is presented in Figure 1.1.

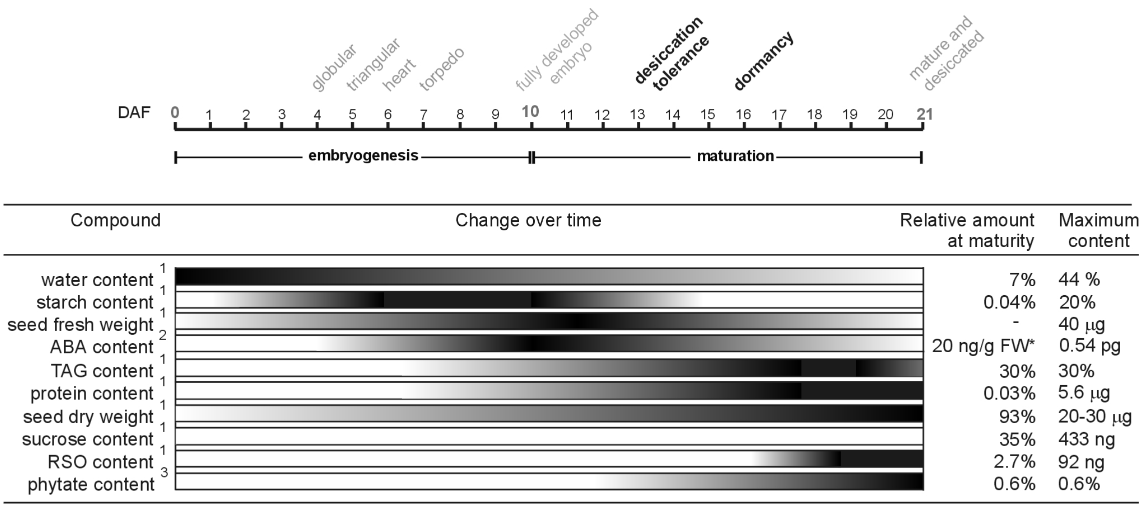


Figure 1.1. Overview of seed development in Arabidopsis. The bars indicate the relative amounts of seed products during seed development. Black indicates the highest amount and white indicates the lowest amount. The maximum amount and the relative levels present in the mature seed are indicated as well. Values are amounts present in a single seed.

DAF; days after flowering, ABA; abscisic acid, TAG; triacylglycerol, RSO, raffinose series oligosaccharides

[1] Data collected from Baud et al., 2002; accession Ws

[2] Data collected from Ooms et al., 1993; accession L*er*

[3] Data collected from Chapter 7; accession L*er*

* measurement performed on the intact silique

The seed maturation phase

An important aspect of seed maturation is the accumulation of food reserves that can be mobilized upon germination (for overview see Baud et al., 2002). In Arabidopsis, the developing embryo accumulates lipids in the form of triacylglycerols (TAG), esters of glycerol and fatty acids. The TAG will be later broken down for germination and growth of the young seedling. They are stored in cytosolic oil bodies, which occupy 30% of the mature seed (Browse and Sommerville, 1994). The most abundant mRNAs in embryos are those encoding the seed storage proteins, which are thought to serve as a source of nitrogen for the young seedling (Pang et al., 1988). Two major seed storage proteins are 12S (globulin) and 2S (albumin) (Heath et al., 1986). Starch accumulates during the first half of development and very low amounts remain in the dry seed (Focks and Benning, 1998). This is the same for glucose and fructose. However, sucrose accumulates at the end of seed development, together with raffinose series oligosacchararides (RSOs), which are derivates of sucrose to which galactose units are added (Ooms et al., 1993). Phytate (myo-inositol-1,2,3,4,5,6- hexakisphosphate) is the major form of phosphorus in seeds and accumulates also during the maturation phase (Raboy, 2001), whereas inorganic phosphate decreases (Quick et al., 1997).

Two other important aspects of seed maturation are the induction of seed dormancy and desiccation tolerance, the latter allows the seeds to be stored in dry conditions (dry seeds have a water content of approximately 7%). Seeds that are not desiccation tolerant are called recalcitrant seed. These seeds have a defective or incomplete seed maturation phase, do not enter the quiescence stage and proceed directly to the germination phase (Farnsworth, 2000). Several compounds have been thought to be involved in the acquisition of desiccation tolerance. The phyto-hormone abscisic acid (ABA) appears to play a central role in preparing embryos for maturation drying before dormancy is induced. Several lines of evidence suggest that ABA is necessary (but not sufficient) for the acquisition of desiccation tolerance (reviewed in Farnsworth, 2000). Desiccation tolerance might also be associated with a decreased sensitivity to ABA, as extreme ABA-insensitive mutants at the *ABI3* locus, are desiccation intolerant (Ooms et al., 1993). Several proteins are implicated in the acquisition of desiccation tolerance as well (Ingram and Bartels, 1996), such as a specific class of dehydration proteins that becomes prevalent in desiccation tolerant seeds when maturation drying begins. These proteins are the Em proteins, the late embryogenesis abundant (LEA) proteins, and the ABA responsiveness (Rab) proteins, which are referred to as “dehydrins”. ABA deficient mutants show reduced levels of these proteins throughout seed development, indicating that ABA signalling may be involved in the production of dehydrins. Furthermore, the induction of desiccation tolerance correlates with the appearance of seed RSOs and therefore this has been cited as an indication of the involvement of sugars in desiccation tolerance (reviewed by Horbowicz and Obendorf, 1994). However, Black et al, (1999), Oliver and Bewley (1997), Sebastian et al. (2000) and ourselves (Chapter 6) did not find this correlation.

Seed dormancy has been defined as a temporary failure of a viable seed to germinate in conditions that favour germination (Bewley, 1997) and is extensively discussed in Chapter 2.

Scope of the thesis

This thesis deals with the genetic characterisation of different seed maturation specific characters that are analysed using the natural variation present between the Arabidopsis accessions L*er* and Cvi. For these analyses a RIL population derived from a cross between these two accessions has been used. The first part of the thesis (Chapter 2-5) focuses on the characterisation of seed dormancy. Chapter 2 gives an overview of current knowledge of factors and genes controlling seed dormancy and germination in Arabidopsis. Chapter 3 describes the genetic analysis of seed dormancy/germination in the L*er*/Cvi RIL population, which resulted in the identification of 7 QTL. The QTL effects are confirmed in several NILs carrying single and double Cvi introgression fragments in a L*er* genetic background. In Chapter 4 the genetic fine mapping of *Delay of Germination 1* (*DOG1*), which is the strongest effect QTL identified in Chapter 3, is described. Chapter 5 deals with the genetic and physiological characterisation of *DOG1*. In Chapter 6 the genetic analysis of seed soluble oligosaccharides in relation to seed storability are described. In Chapter 7 the natural variation of phytate and phosphate accumulation in seeds and leaves has been analysed and the major QTL affecting phytate/phosphate in seeds and in leaves has been fine mapped. Finally in Chapter 8 the work presented in this thesis is summarised and discussed.

Chapter 2

Seed dormancy and germination

Leónie Bentsink and Maarten Koornneef

This chapter has been published in The Arabidopsis Book, eds. C.R. Somerville and E.M. Meyerowitz, American Society of Plant Biologists, Rockville, MD, DOI 10.1199 tab .0050
http://www.aspb.org/publications/arabidopsis/

Introduction

The seed is an important stage in the higher plant life cycle with respect to its survival as a species. It is the dispersal unit of the plant, which is able to survive the period between seed maturation and the establishment of the next generation as a seedling after it has germinated. For this survival, the seed, mainly in a dry state, is well equipped to sustain extended periods of unfavourable conditions. To optimise germination over time, the seed enters a dormant state. Dormancy prevents pre-harvest germination as well. Numerous studies have been performed to better understand how germination is controlled by various environmental factors and applied chemicals. However, still very little is known about the process by which the embryo emerges from the seed to complete germination and how embryo emergence is blocked in dormant seeds (Bewley, 1997).

Arabidopsis possesses dormancy, as is the case for many other plant species, which is controlled by environmental factors such as light, temperature and time of dry storage as well as by genetic factors (Koornneef and Karssen, 1994). The use of genetics and molecular genetics in Arabidopsis is starting to shed light on some aspects of the mechanism of dormancy and germination by the identification of mutants and genes that control these processes. This review provides an overview of current knowledge of factors and genes controlling seed dormancy and germination in Arabidopsis.

Seed development

Seed development comprises two major phases: embryo development and seed maturation. Embryogenesis, which is a morphogenesis phase, starts with the formation of a single-cell zygote and ends in the heart stage when all embryo structures have been formed (Mayer et al., 1991). It is followed by a growth phase during which the embryo fills the seed sac (Goldberg et al., 1994). At the end of the embryo growth phase, cell division in the embryo arrests (Raz et al., 2001). Hereafter, the seed, containing a full sized embryo, undergoes maturation during which food reserves accumulate and dormancy and desiccation tolerance develops (Goldberg et al., 1994). During normal seed development, embryo arrest and dormancy are reversed upon germination, which occurs when proper environmental conditions are provided and the dry seed imbibes water.

Seed development has been extensively studied using mutants defective in various aspects of the process. Mutants affected in the morphogenesis phase result in lethality or have a seedling phenotype (Mayer et al., 1991; Meinke, 1995). In seed germination mutants, properties of germination and dormancy are affected which sometimes are accompanied by pleiotropic effects that are specific for maturation, such as desiccation intolerance (Goldberg et al., 1994; Koornneef and Karssen, 1994).

Seed dormancy and germination

Seed dormancy has been defined as the failure of an intact, viable seed to complete germination under favorable conditions (Bewley, 1997). Since in Arabidopsis removal of the seed coat allows germination of non-germinating and strongly dormant genotypes, dormancy in Arabidopsis should be described as coat-enhanced dormancy (Bewley, 1997). However, in addition to the structures surrounding the embryo, the growth potential of the embryo is also important to overcome the constraint of these structures and thereby affects the dormancy state of a seed.

Since dormancy is regulated at different developmental phases, in interaction with environmental factors, it is difficult to detect when the genetic and physiological differences are established. This difficulty arises because all assays are based on seed germination, which is the result of the balance between the degree of dormancy and the capacity of the embryo to overcome dormancy. Mechanistically one can distinguish factors that influence dormancy and germination on the basis of their effect on germination, being either inhibiting or promoting. Mutants that germinate better or faster can represent genes that promote dormancy or those that repress germination. A further distinction can be made by defining the timing and site of action of these factors (during maturation or during imbibition of the seeds, in the embryo or in the testa). The interaction between these factors and the large effect of the environment, both during seed development and during imbibition, make seed dormancy a very complex trait.

By definition, germination incorporates those events that commence with the uptake of water by the quiescent dry seed and terminates with the elongation of the embryonic axis (Bewley and Black, 1994). The visible sign that germination is complete is usually the penetration of the structures surrounding the embryo by the radicle. Germination assays in Arabidopsis are often performed in light on seeds freshly harvested or stored for a limited time (Léon-Kloosterziel et al., 1996a). Other parameters are the germination rate after different periods of cold treatment (Cutler et al., 1996) and germination in darkness. In addition to testing mature seeds, germination of immature seeds either excised from the silique or within fruits detached from the plant, can be used to investigate genetic variation during the early stages of seed development (Raz et al., 2001).

Since tissues from both maternal (testa) and zygotic origin (embryo and endosperm) contribute to seed germination, genetic analyses of seed dormancy have to take into account these different tissue origins. Maternal effects, in contrast to zygotic factors are maternally inherited and might be due to the genetic make up of the testa surrounding the embryo, but can also be due to genetic differences related to factors that are transported into the seed from the mother plant. Maternal inheritance can be deduced from the germination of seeds obtained after reciprocal crosses, where the parental genotypes are used both as female and as male parent. Although reciprocal differences might also be attributed to cytoplasmic inheritance of the trait, genetic segregation in the subsequent generations allows the distinction between cytoplasmic and maternal inheritance (Léon-Kloosterziel et al., 1994).

Seed dormancy in Arabidopsis can be overcome by, or seed germination is activated by, the common germination promoting factors such as after-ripening, light, cold treatment (also called stratification) and chemicals such as gibberellins and KNO_3 (Derkx and Karssen, 1993b). None of these exogenous factors are an absolute requirement for germination because the need for one factor depends on the other factors, as shown for the interaction between light and temperature by Cone and Spruit (1983). This requirement for exogenous factors depends very much on the genotype.

Non-dormant seeds that are exposed for some time to unfavorable germination conditions (imbibed seeds kept at relatively high temperature in darkness for example) may enter a state of dormancy again, which is called secondary dormancy (Cone and Spruit, 1983; Derkx and Karssen 1993a).

The challenge in dormancy and germination research is to identify the nature of the crucial regulator(s) that prevent(s) the onset of germination (dormancy) and that trigger(s) the germination process and their mutual interaction.

Genetic variation for germination characters

Genetic variation can be induced by mutagenesis, but is also present among the many natural accessions (ecotypes) of Arabidopsis. Genetic variation for germination can be detected when genotypes are compared in identical environments. This implies that not only the conditions of the germination test must be identical, but also growth conditions during seed development and storage conditions, including the time that the seeds are stored, must be the same. Furthermore, the test conditions must be discriminative between genotypes.

Germination tests can be used efficiently for mutant screens because of the large numbers that can be assayed. However, variability of the germination trait may lead to genetic misclassification of individual seeds and therefore to false-positives in mutant screens. The Arabidopsis genotypes Landsberg *erecta* (L*er*) and Columbia (Col), which are mostly used in Arabidopsis research, show only a low level of dormancy. This dormancy disappears after approximately one month of after-ripening (van der Schaar et al., 1997). Because of this relatively low level of dormancy it is impossible to saturate mutations in dormancy genes. However this problem can be overcome by the use of more dormant accessions (Koornneef et al., 2000).

The genetic variants described in this review have in common that the seeds are viable, as shown by their ability to germinate after special treatments, such as disruption of the seed coat or the application of specific chemicals. Mutants that do not germinate because they are lethal, including early ovule mutants (Schneitz, 1999) and many of the so-called embryo lethals (Meinke, 1995) are not described in this review because they mainly have developmental defects that do not control specifically dormancy and germination.

The analysis of natural variation

Arabidopsis is an annual plant for which large differences in dormancy can be found between accessions collected from nature (Ratcliffe, 1976). Kugler (1951) performed the first genetic study of differences in germination between Arabidopsis accessions. That even small differences in dormancy are also amenable to genetic analysis in Arabidopsis was shown by Quantitative Trait Loci (QTL) analysis of the differences in seed germination between two commonly used Arabidopsis accessions L*er* and Col (van der Schaar et al., 1997). Germination tests were performed with seeds grown from a set of 98 recombinant inbred lines (RILs), in three different maternal environments, and each seed batch was tested in three different germination environments: in light, in darkness and in the presence of the gibberellin biosynthesis and germination inhibitor, paclobutrazol. For nine out of the fourteen loci detected, no significant interaction between the locus and environmental factors could be detected. However, three distinct loci controlling germination behavior in the presence of paclobutrazol had a much lower or no effect when germination was tested in water. Two other loci affecting germination in darkness and/or light had practically no effect on germination in paclobutrazol. The effects of the individual loci were small in all cases, which makes a thorough molecular analysis very difficult. However, the presence of QTL in specific regions, which may represent quantitative effects of specific alleles, can be studied when genes affecting seed dormancy are later identified in those regions. QTL analysis performed on RILs between accessions with a larger difference in dormancy; L*er* (low level of dormancy) and Cvi (strong dormancy), revealed individual QTL with much stronger effects. Seven chromosomal regions affecting this trait have been identified and the effects of the individual loci could be confirmed by analyzing Near Isogenic Lines (NILs) (L. Bentsink, unpublished results). These NILs are being used to clone the

QTL and also for further mutagenesis experiments that may lead to mutations in the respective QTL or in genes acting elsewhere in the dormancy/germination pathways.

It is expected that the analysis of more accessions, of which many are more dormant than the commonly used laboratory accessions (Ratcliffe, 1976; Koornneef et al., 2000), will reveal more loci affecting seed dormancy and germination.

Mutants affected in seed dormancy/germination

There are two groups of seed dormancy/germination mutants that can be recognised. In one group seed dormancy is reduced (germination is promoted) and in the other seed dormancy is increased (germination is inhibited).

Mutants that reduce and/or prevent seed dormancy or promote germination

The first important stage for dormancy induction is probably the end of the morphogenetic program, when all tissues present in a mature embryo have been formed and the embryo enters a phase of growth arrest. *ABA-INSENSITIVE3* (*ABI3*), *FUSCA3* (*FUS3*) and *LEAFY COTYLEDON* (*LEC1* and *LEC2*) play prominent roles in controlling mid- and late seed development (Meinke et al., 1994). Mutations at any of these loci affect multiple processes, including accumulation of storage proteins, as well as acquisition of dormancy and desiccation tolerance. It appears that these four genes have partially overlapping functions in the overall control of seed maturation (Parcy et al., 1997). It has been suggested that the non-dormant phenotype of *lec1, fus3* and *abi3* mutants is due to defective seed maturation and that mutant seeds germinate because this is the default state (Nambara et al., 2000). However, detailed analysis of *lec1, fus3* and *abi3* mutants showed that they differ in the time at which premature germination can occur. *LEC1 (*probably *LEC2* as well) and *FUS3* loci regulate developmental arrest, as mutations in these genes cause a continuation of growth in immature embryos. Abscisic acid (ABA) controlled dormancy (via *ABA* and *ABI*) occurs later and is additive to the developmental arrest controlled by *LEC1* and *FUS3* (Raz et al., 2001). The dependency of germination on GA is maintained in *fus3* mutants but not in *lec1*, which suggests that these mutants affect the germination potential of seeds in different ways.

Apart from mutants that influence general seed maturation, other mutants more specifically influence seed dormancy, i.e. mutants, which are altered in ABA biosynthesis or its mode of action. ABA regulates various aspects of plant growth and development, including seed dormancy. The absence of ABA-induced dormancy allows seeds to germinate without gibberellins. Therefore, the selection of mutants that germinate in the presence of GA biosynthesis inhibitors, such as paclobutrazol and tetcyclacis, is an effective way to isolate ABA biosynthesis mutants (Léon-Kloosterziel et al., 1996a). Reciprocal crosses between wild type and the ABA deficient *aba1* mutants showed that dormancy is controlled by the ABA genotype of the embryo and not by that of the mother plant. The latter is responsible for the relatively high ABA levels found in seeds halfway through seed development (Karssen et al., 1983). At this phase ABA may prevent precocious germination as shown by the maternal ABA effects in the extreme *aba abi3-1* double mutants (Koornneef et al., 1989). Since in wild type ABA levels decrease at the end of seed maturation it was proposed that after the onset of dormancy endogenous ABA is not required for its maintenance (Karssen et al., 1983). However, the observation that inhibitors of ABA biosynthesis, such as norfluorazon, promote germination (Debeaujon and Koornneef, 2000) indicates that the maintenance of dormancy in imbibed seeds is an active process involving de novo ABA synthesis as was found also for *Nicotiana plumbaginifolia* (Grappin et al., 2000). Because ABA has such a major effect on seed dormancy it can be expected that defective ABA signalling also leads to changes in germination characteristics. Substantial progress has been made in the characterization of such ABA signal transduction pathways (Bonetta and McCourt, 1998; Leung and Giraudat, 1998). Genetic screens to identify ABA signalling mutants were based primarily on the inhibition of seed germination by applied ABA. The ABA-insensitive (*abi*) mutants *abi1* to *abi5* are able to germinate in the presence of ABA concentrations that are inhibitory to the wild type. In contrast, germination of the *era1* (enhanced response to ABA) to *era3* mutant seed is prevented by low concentrations of ABA that ordinarily permit germination of the wild-type seed (Cutler et al., 1996). As judged from their impact on seed dormancy, these two sets of mutations also alter the regulation of seed germination by endogenous ABA. Like ABA-deficient mutants, the ABA-insensitive mutants *abi1* to *abi3* display marked reductions in seed dormancy. Conversely, the ABA-supersensitive mutant *era1* confers enhanced seed dormancy (Cutler et al., 1996). The *abi3, abi4* and *abi5* mutants exhibit reduced expression of various seed maturation genes but only *abi3* mutants are non-dormant, which coincides with desiccation intolerance (Nambara et al., 1992, Ooms et al., 1993, Bies et al., 1999) in strong alleles. Surprisingly no dormancy or other seed maturation phenotype was observed in *abi4* and *abi5* mutants (Finkelstein, 1994), except reduction of some seed maturation specific mRNAs (Finkelstein and Lynch, 2000; Söderman, et al., 2000). This may indicate that other genes are redundant in function to these seed specific transcription factors, which are members of the APETALA2 domain (*ABI4*, Finkelstein et al., 1998; Söderman, et al., 2000) and basic leucine zipper factor family (*ABI5,* Finkelstein and Lynch, 2000; Lopez-Molina et al., 2001).

Genetic screens based on ABA-regulated reporter genes in vegetative tissues revealed two additional ABA related mutants. The *ade1* (ABA-deregulated gene expression) and *hos5* (high expression of osmotically responsive genes) mutations enhance gene expression in response to ABA, (Foster and Chua, 1999; Xiong et al., 1999) but have little effect on seed germination. This is also observed for the *aao3* mutant, which affects a step in ABA biosynthesis, different from those in the *aba1* – *aba3* mutants (Seo et al., 2000), probably because other redundant genes with seed-specific expression compensate for the function of the mutated genes.

It has been found that sugars such as sucrose and various hexoses inhibit seed germination independently of their osmotic effects (Pego et al., 2000). Mutants that were insensitive to the inhibiting effect of glucose and sucrose were isolated by several groups and appeared to be defective in ABA biosynthesis or are among the ABA insensitive mutants. The *sugar insensitive, sis4* and *sis5* mutants are allelic to respectively *aba2* and *abi4* (Laby et al., 2000) *and sucrose uncoupled* (*sun6*) is also an *abi4* allele (Huijser et al., 2000). These results indicate that germination is controlled in an ABA and sugar dependent way. In respect to this, Garciarrubio et al (1997) published that the addition of sugars and amino acids allowed seeds to germinate in otherwise inhibitory concentrations of ABA and suggested that ABA inhibits the mobilization of food reserves. However, it cannot be excluded that these sugar effects are mediated by sugar signalling effects (Smeekens, 1998; Gibson and Graham, 1999).

A class of mutants that was directly selected on the basis of reduced dormancy are the *rdo1*-*rdo4* mutants (Léon-Kloosterziel et al., 1996b; A.J.M. Peeters unpublished results). The fact that all four mutants show some mild pleiotropic effects in adult plants indicates that the genes are not specific for dormancy/germination but affect other processes as well. The mutants differ from each other in their pleiotropic and epistatic effects, indicating that not one single pathway is involved. Recent data suggest that *RDO2*, *RDO3* and probably *RDO1* affect the GA requirement, although in a less severe way than ABA, probably due to redundancy of the function in the dormancy control of the various genes. However, *RDO4* must act on a target different from GA requirement (A.J.M. Peeters unpublished results).

The *dag1-1* mutant displays reduced dormancy, but in contrast to the *rdo* mutants the effect is determined by the maternal genotype. This is in agreement with the expression pattern of the *DAG1* gene in the vascular tissue of the developing seed. *DAG1,* which encodes a DOF transcription factor, may influence the import of compounds from the mother plant into the seed (Papi et al., 2000). It is the first gene identified as being specifically involved in maternal control of seed germination. Recently, a related gene, named *DAG2,* with a similar expression pattern as *DAG1* was isolated. However, the germination phenotype of *dag2* mutant seeds is opposite to that of *dag1* seeds (Vittorioso et al., 2001) as it shows increased dormancy. Additional mutants with a reduced dormancy phenotype at other loci, including mutants with no obvious pleiotropic effect have been isolated (unpublished results) indicating the complexity of the genetic regulation of seed dormancy.

Mutants with an altered seed coat or testa also show reduced seed dormancy (Debeaujon et al., 2000). The seed coat is a multifunctional organ that plays an important role in embryo nutrition during seed development and in protection against detrimental agents from the environment afterwards (Mohamed-Yasseen et al., 1994; Weber and Wobus, 1999). The seed coat is formed from two integuments of epidermal origin that surround the mature ovule. The development of the seed coat from the ovule has been described by Beeckman et al. (2000). The morphological differentiation of the outer integument, which excretes mucilage upon imbibition (Willats et al., 2001) has been reported by Western et al. (2000) and Windsor et al. (2000).

The seed coat exerts a germination-restrictive action, either by being impermeable to water and /or oxygen, by producing germination inhibiting compounds or by its mechanical resistance to radicle protrusion. In Arabidopsis, phenolic compounds and their derivatives present in the inner layer of the testa, called endothelium, affect seed coat properties that influence germination as can be concluded from the reduced dormancy phenotype of many testa mutants.

Seed coat mutants consist of two major groups. One group, affected in flavonoid pigmentation is represented by the *transparent testa* (*tt*) and *transparent testa glabra* (*ttg*) mutants. Mutants identified are *tt1* to *tt15*, *ttg1 and ttg2* and *banyuls* (*ban*). The color of the *tt* mutants ranges from yellow to pale brown (Debeaujon et al., 2000). *Ban* mutants accumulate pink flavonoid pigments in the endothelium of immature seeds, but have reduced brown pigments, resulting in grayish-green, spotted mature seeds (Albert et al, 1997; Devic et al, 1999). The *ttg1* mutant lacks mucilage and trichomes and is affected in the morphology of the outer layer of the seed coat as well as in pigment production.

The second group is represented by mutants affected in testa structure. The aberrant testa shape (*ats*) mutant that controls the differentiation between inner and outer integuments has only one integument with 3 cell layers instead of 2 with 2 + 3 cell layers (Léon-Kloosterziel et al., 1994). The fact that this mutant has no adult phenotype suggests that its function is restricted to this phase of development.

Mutants that increase seed dormancy or reduce the germination potential of seeds

Many mutants that increase seed dormancy or reduce the germination capacity have mutations in the biosynthesis or signaling pathways of plant hormones that stimulate seed germination such as gibberellins, brassinosteroids and ethylene.

The plant hormone gibberellin plays an important role in promoting seed germination. GA-deficient mutants are unable to germinate without exogenous GAs

(Koornneef and van der Veen, 1980). De novo biosynthesis of GAs is required during imbibition, as was concluded from the observation that inhibitors of GA biosynthesis, such as paclobutrazol and tetcyclacis prevent germination (Karssen et al., 1989) unless ABA is absent. GAs can promote germination by their ability to overcome germination constraints that exist in seeds requiring after-ripening, light and cold. This led to the suggestion that such environmental factors may induce GA biosynthesis during the early phases of germination. This light effect was supported by Yamaguchi et al., (1998), who showed that one of two 3-β hydroxylase enzymes, encoded by the *GA4H* gene is induced in germinating seeds by phytochrome. Cold treatments do not stimulate GA biosynthesis, but rather increase their sensitivity to GAs (Debeaujon and Koornneef, 2000). Two different mechanisms of action have been proposed to explain the role of endogenous GAs in the control of germination. The first one is the induction of genes encoding enzymes that reduce the mechanical resistance to radicle protrusion. This has not been proven for Arabidopsis. The second mechanism consists of a direct effect on the growth potential of the embryo (Karssen and Laçka, 1986). The latter is assumed to be restricted by the plant hormone ABA, which is produced in the embryo. GA is required to overcome the ABA-induced dormant state. Severe mutations in genes acting early in GA biosynthesis (*GA1*, *GA2* and *GA3*) display a number of GA-rescuable phenotypes. These include failure to germinate, growth of the plant as a dark green dwarf, underdeveloped petals and stamen accompanied by reduced fertility, delayed flowering in short days, reduced apical dominance, and delayed senescence (Koornneef and van der Veen, 1980). A screen aimed specifically at isolating T-DNA tagged mutants that do not germinate was reported by Dubreucq et al., (1996). This screen yielded two mutants in the first 3500 screened T-DNA lines, among which was one *ga1* mutant.

A single dominant mutant with decreased GA signal transduction has been described, *gai1* (GA-insensitive) (Koornneef et al., 1985). *Gai1* has a reduced sensitivity to GA but does not exhibit strongly reduced germination. However, loss of function alleles of this locus require slightly less GA for growth than wild type (Peng et al., 1997). This GA hypersensitive phenotype resembles the *rga* (repressor of *ga1-3*) mutant. The *RGA* gene shares 82% sequence identity with *GAI.* The *GAI* and *RGA* genes belong to the GRAS family and have overlapping but not redundant functions in transcriptional regulation (Sun, 2000) acting as repressors of GA action.

The *sleepy1 (sly1)* mutant was selected in a screen for suppressors of the ABA insensitive mutant *abi1-1*. This mutant strongly resembles the GA auxotrophs described above. However, the lack of germination of *sly1* cannot be rescued by GA and therefore *sly1* is the first GA-insensitive mutant that reflects the full spectrum of GA-associated phenotypes. Therefore, *SLY1* has postulated to be a key factor in GA reception (Steber et al., 1998).

Mutations in the *COMATOSE* (*CTS*) gene also results in a marked reduction in germination potential. Whilst the morphology of *comatose* embryos is not altered, physiological analysis reveals that mature *cts* seeds do not respond to gibberellin. Prolonged chilling of imbibed seeds only partially restores germination potential. It is suggested that *CTS* promotes increased germination potential, represses embryo dormancy and might be involved in GA signaling specific for seeds (Russel et al., 2000).

Brassinosteroids (BRs), a group of over 40 naturally occurring plant steroid hormones found in a wide variety of plant species (Clouse and Sasse, 1998; Schumacher and Chory, 2000) are also involved in the control of germination in Arabidopsis. It is suggested that the BR signal is required to reverse ABA-induced dormancy and to stimulate germination (Steber and McCourt, 2001). BRs could overcome the lack of germination of the *sleepy1* mutant, probably by bypassing its GA requirement. Two BR mutants *DET2* and *BRI1* show reduced germination but eventually germinate without BR, indicating that in contrast to GAs BRs are not absolutely required for germination (Steber and McCourt, 2001).

Mutants in ethylene signalling are also affected in their germination response. Ethylene is produced in trace amounts by almost all higher plants and is involved in the control of growth and developmental processes that range from germination to senescence. Often seeds that respond to ethylene are light sensitive for germination (Kepczynski and Kepczynska, 1997). Ethylene mimics the action of gibberellins as applied ethylene allows germination of seeds of the GA-deficient mutant (which normally only germinate in the presence of GA_{4+7}) (Karssen et al., 1989). Ethylene insensitive mutants such as *etr* and *ein2* germinate less well or after a longer period of after-ripening than wild type (Bleecker et al., 1988; Beaudoin et al., 2000). Ethylene mutants show phenotypes that resemble ABA and sugar signaling mutants. The *ein2* and *etr* mutants are hypersensitive to ABA (Beaudoin et al., 2000; Ghassemian et al., 2000), which agrees with the observation that *ein2* mutants were isolated as *abi1-1* suppressors. The *ctr1* mutant, which is characterised by a constitutive ethylene response, appeared among mutants selected as enhancers of the ABA insensitive mutant *abi1-1* and *ctr1* monogenic mutants are also slightly ABA resistant (Beaudoin et al., 2000). These observations, in combination with the non-dormant phenotype of the *ein2 abi3-4* double mutant indicated that ethylene negatively regulates seed dormancy by inhibiting ABA action (Beaudoin et al., 2000). However, Ghassemian et al (2000) suggested that, in addition to signaling, the effect of *EIN2* might also involve effects on ABA biosynthesis. The presence of cross talk between sugar signaling and ethylene was suggested by the sugar insensitive phenotype of *ctr1* (Gibson et al., 2001) and the sugar hypersensitive phenotype of *etr* (Zhou et al., 1998). Apparently ABA, ethylene and sugar signaling strongly interact at the level of germination and early seedling growth.

The germination of Arabidopsis seeds is under phytochrome-mediated photocontrol. It is therefore to be expected that phytochrome deficient mutants are affected in seed germination. The complexity of the phytochrome system comes from the presence of distinct types of phytochromes, for which five genes in the Arabidopsis genome encode different, but related, apoproteins (Sharrock and Quail, 1989). In addition different modes of action of phytochrome, described as very-low-fluence response (VLFR), low-fluence response (LFR) and high-irradiance response (HIR), which have their own fluence dependency, can affect germination (reviewed by Casal and Sánchez, 1998). Mutants lacking phytochrome B show a reduced sensitivity to red light, indicating that phyB has a primary role in seed germination. PhyA can only induce germination after a prolonged imbibition of seeds (Shinomura et al., 1994). Detailed action spectra for seed germination performed in wild type, *phyA* and *phyB* mutants revealed a typical red/ far-red (R/FR) -reversible LFR mediated by phyB, whereas the germination response mediated by phyA turned out to be a VLFR with a 10^4 -fold higher sensitivity to light (Shinomura et al, 1996). The observation that also *phyA phyB* double mutants show some light- dependent germination indicates the involvement of another R/FR-reversible photoreceptor system (Yang et al., 1995; Poppe and Schäfer, 1997) probably mediated by phyC, D, and/or E.

Although the main role of phytochrome is in light-induced stimulation of seed germination, a role in the onset of dormancy or the setting of the light requirement is suggested by the experiments of McCollough and Shropshire (1970) and Hayes and Klein (1974). These authors showed that the R/FR ratio experienced by the mother plant and therefore during seed maturation, affects the subsequent germination behavior of mature seeds. Recently Munir et al. (2001) showed that photoperiod conditions during seed formation may also influence seed germination. However, this effect was strongly genotype dependent.

The genetic control of seed storage compounds, seed deterioration and early seedling development

Seeds can survive for a long time without germinating, either when stored in dry conditions or when buried in the soil. These seeds form the seed bank waiting until environmental conditions become favorable for germination. Despite the strong desiccation tolerance of many seeds, storage also under dry conditions ultimately leads to a loss of viability. It has been suggested that compounds such as certain sugars and proteins, such as LEA (Late embryogenesis abundant) (Skriver and Mundy, 1990) and Heat shock proteins (Hsps) (Wehmeyer and Vierling, 2000; Hong and Vierling, 2001), that accumulate during the later stages of seed development have a desiccation protective role. However, mutants such as *abi5*, which show a strong reduction in some LEA proteins (Finkelstein and Lynch, 2000) do not have obvious defects in seed storability. A knock-out mutant of Hsp101 shows a loss of thermotolerance during seed germination (Hong and Vierling, 2001), but under optimal conditions germination seems unaffected, suggesting that at least Hsp101 does not have a general function in the survival of desiccated seeds.

The presence of genetic variation for storability is clear in desiccation intolerant genotypes such as *abi3, lec1* and *fus3* mutants (Meinke et al., 1994; Weber and Wobus, 1999) and effects on testa mutants are reported as well (Debeaujon et al., 2000). QTL analysis in the Cvi/Ler RIL lines (Bentsink et al., 2000) also revealed genetic variation for this trait, which was however not related to the levels of raffinose series oligosaccharides segregating in the same population.

For a seed it is important that after surviving a period in which it could not germinate, it can grow into a vigorous seedling that can compete with other seedlings. Factors that are important for seedling establishment probably mainly have to do with the availability and mobilization of storage materials. The accumulation of storage proteins and lipids is defective in seed maturation mutants such as *lec1, lec2* and *fus3*, which although they germinate on filter paper they cannot be transferred efficiently to soil from this substrate, in contrast to other seeds, and require establishment as seedling on sugar supplemented media before they can be transplanted to soil. Another mutant that affects specifically seed storage compounds, but is without a reported effect on seed germination is *wrinkled1* (*wri1*; Focks and Benning, 1998), which has an 80% reduction in seed oil content. It is suggested that *wri1* is not directly affected in fatty acid triacylglycerol biosynthesis, but controls the conversion of glucose into fatty acid precursors. It can be expected that these mutants have problems in seedling establishment as well.

The quantity of storage material is related to the size of seeds and also for this character considerable variation exists among Arabidopsis accessions (Krannitz et al., 1991), where the larger seeds also produce larger seedlings. This natural variation for seed size was genetically analysed in the L*er* x Cvi RIL population by Alonso-Blanco et al. (1999) who found that for this trait very large reciprocal differences exist that indicate the presence of a maternal inherited factor in combination with specific paternal effects that might be related to imprinting phenomena. The number of loci that are segregating for seed size in this population is 11, some of which seem to control the number of ovules formed per unit silique length as a pleiotropic effect.

For the mobilization of food reserves during germination and early seedling growth, the activity of lipid and carbohydrate degrading enzymes such as malate synthase (MLS) and iso-citrate lyase (ICL) is required. These enzymes, which are unique to the glyoxylate cycle, are expressed specifically upon imbibition of oilseeds such as Arabidopsis. Knock-out mutants in these genes (*icl* and *mls*) confirmed that in these mutants the defect is not in germination per sé, but specifically in seedling establishment. The *icl* and *mls* mutant seeds germinate and become photoautotropic when grown on agar plates or on soil in the greenhouse. Therefore, the glyoxylate cycle can be described as being non-essential in Arabidopsis, although it does play an important physiological role because the survival rate of *icl* mutant seedlings declines when they are grown in sub-optimal light conditions (Eastmond and Graham, 2001). The crucial role of GA in mobilization of starch degrading enzymes in the cereal aleuron layer has not been shown in Arabidopsis, where the aleuron layer comprises only one cell layer at maturity. However, the granule structures in this cell layer disappear upon imbibition (Debeaujon and Koornneef, 2000), which also occurs when germination takes place in a GA deficient mutant in which the testa has been removed.

Genes with an expression pattern correlated with dormancy and germination

In addition to searching for mutants and the subsequent cloning of genes, genes controlling seed dormancy and germination can also be identified on the basis of their expression pattern. However, even abolishment of gene function may not always result in an altered phenotype due to redundancy of other genes.

Various differential screening procedures using subtractive probes have proven successful in identifying tissue-specific genes. These techniques, however, are limited in application in seeds due to their requirement of large amounts of poly(A)$^+$ RNA, and also due to the relatively high expression of seed-storage protein genes which tend to mask less abundant messengers.

Many experiments have shown that during seed maturation the expression of many genes is altered and specific classes of mRNAs such as LEA appear. However, for none of these genes has a specific function in seed dormancy/germination been proven. Although it appears that seed maturation and post-germinative growth have a distinct gene expression profile, some genes that are highly expressed after germination are also expressed during the later stages of seed development (reviewed by Harada, 1997), suggesting that some aspects of post-germinative growth are initiated during maturation. The onset of early germination is also obvious in some of the maturation mutants.

To study genes which are activated during late embryo development and germination a differential display has been performed on immature siliques of the *abi3 fus3* double mutant in comparison with wild-type (Nambara et al., 2000). In addition the same authors identified 16 clones that were activated during wild-type germination of which 11 are de-repressed in the *abi3 fus3* double mutant that apparently have an activated germination process. The genes that were identified represent a variety of metabolic enzymes, regulatory proteins and a number of ribosomal proteins. Cellular processes aiming at growth and activation of protection mechanisms, such as those involved in protection against oxidative stress, and storage compound metabolism are expected to be related to germination. Such proteins were also found in a proteomic analysis of seed germination (Gallardo et al., 2001).

mRNAs that are differentially accumulated in dormant versus non-dormant seed have been identified in several species (Dyer, 1993; Johnson et al., 1995). However, here again the essential role of such genes has not been established, and even the expression level of only a surprising low number of genes has been studied into some detail. The peroxiredoxin gene was an orthologue of a barley gene associated with dormancy, but since expression in freshly harvested (dormant) seeds is much higher than in after-ripened seeds, it is also in Arabidopsis associated with dormancy (Haslekas et al., 1998).To prove that genes identified by these criteria have essential functions in seed germination, knock out mutants of such genes need to be studied.

Gene or promoter trapping with reporter genes such as GUS may identify genes with a specific expression. Dubreucq et al. (2000) isolated gene traps with an expression during seed germination and identified an insertion close to the *AtEPR1*. This gene encodes an extensin-like protein and is specifically expressed in the endosperm during seed germination, under the control of GAs (Dubreucq et al., 2000). A novel medium-chain acyl-CoA oxidase, which is transcriptionally induced during seed germination has been identified by promoter trapping (Eastmond et al., 2000).

Several novel high throughput techniques for a large scale monitoring of gene expression, such as serial analysis of gene expression (SAGE) and gel-based RNA fingerprinting techniques (cDNA-AFLP) have been added to the available techniques of gene expression analysis. Recent progress in proteome and whole transcriptome analysis using microarrays and DNA chips may provide a complete picture of genes involved in the seed germination process.

Microarrays containing 2600 seed-expressed genes for developing Arabidopsis seeds were described by Girke et al., (2000) and revealed many genes of unknown function that are highly expressed in seeds.

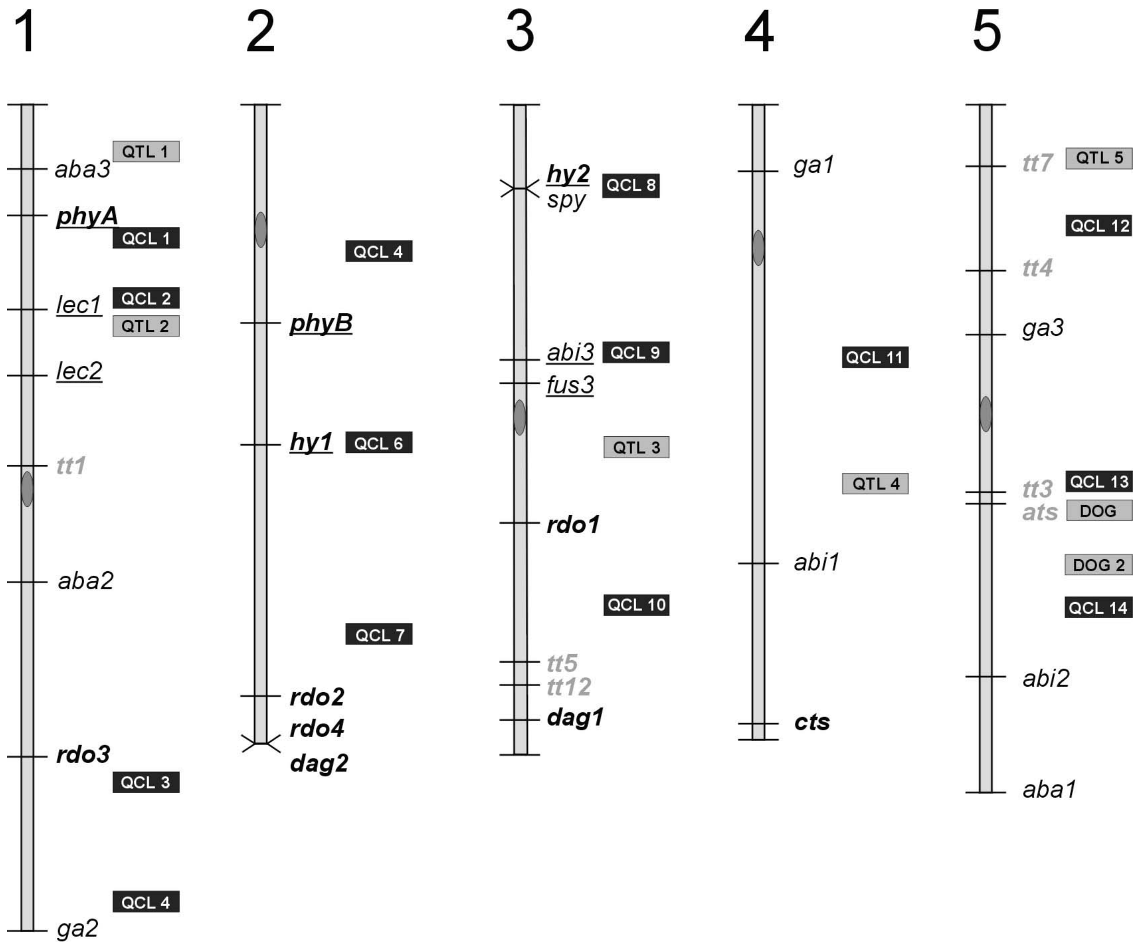


Figure 2.1. Arabidopsis genetic map showing the mutant loci and polymorphic QTL identified affecting germination/dormancy. Mutants are divided in groups, every group is represented by a different code. Germination/dormancy affected by hormones (normal), testa mutants (gray), mutants affected in light response (**bold under lined**), seed maturation mutants (under lined) and miscellaneous dormancy mutants (**bold**). Black blocks represent QTL mapped in the L*er*/Col RILs (van der Schaar et al., 1997) and the gray blocks QTL mapped in the L*er*/Cvi RILs (C. Alonso-Blanco and L. Bentsink, unpublished results).

Conclusions and perspectives

Dormancy is a very complex trait under the control of a large number of genes. Mutants and the molecular/biochemical function of the respective genes, as so far known, are listed in the appendix. Seed dormancy and germination deal with structural factors, especially in the structures surrounding the embryo and factors affecting the growth potential of the embryo. The latter may include compounds that are imported from the mother plant and also factors that are produced by the embryo itself, among which are several plant hormones. The genetic analysis of Arabidopsis identified the crucial role of ABA in seed dormancy, as well as the requirement of GAs for germination. The latter requirement only exists when ABA was and/or is present and when the envelopes surrounding the seed can prevent germination. For GAs the requirement of de novo GA synthesis upon imbibition is convincingly shown and light acting through the phytochrome system controls this step through induction of a seed specific GA 3β hydroxylase. The importance of the mechanical barrier of the maternally inherited testa was shown by the reduced dormancy of a large collection of testa mutants in Arabidopsis. However, if the single layer of endosperm tissue (the aleurone layer) provides a barrier for germination in Arabidopsis, as it does in some species such as tomato and *Datura ferox* (Bewley, 1997), which have a thicker endosperm layer is not known. The nature of other germination inhibiting or promoting factors, of which the existence is shown by mutant and natural variants in the respective genes is not known. With the exception of two DOF transcription factors none of these genes has been cloned. The complexity of the genetic control of seed dormancy and germination is shown by the large number of loci identified already (Figure 2.1) and which probably reflect only the tip of an iceberg. In what way the genes with unknown functions are downstream targets of ABA and GA or if they affect seed dormancy/germination in an independent way is currently not known. An important aspect of seed biology is also the storability and stress tolerance of seeds and factors influencing seedling establishment. For both processes genetics only very recently started to be used for their analysis. A general scheme of the factors controlling germination is shown in Figure 2.2.

The molecular identification of all these genes will be important as well as the identification of more target genes. The latter can and will be done most efficiently by using whole transcriptome and proteome approaches.

Acknowledgements

We thank Dr. Elizabeth Vierling for critical reading the manuscript.

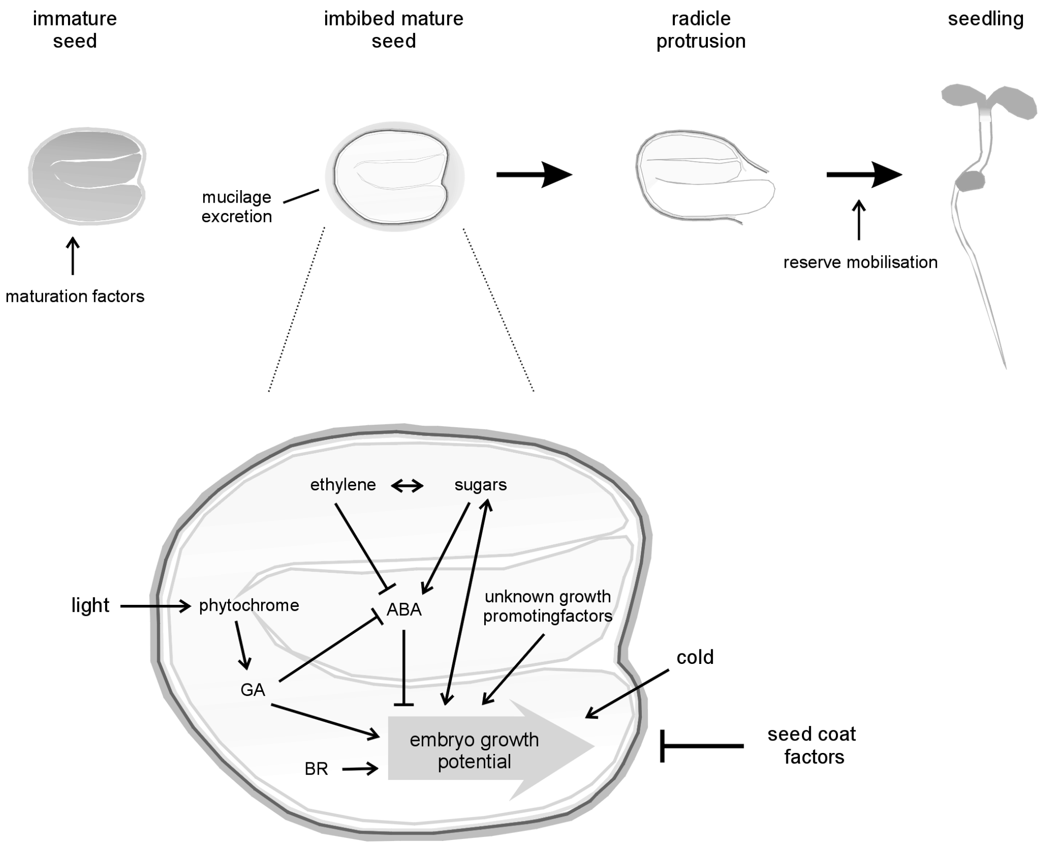


Figure 2.2. General scheme of the factors controlling germination of Arabidopsis.

Appendix

Table of mutants which are related to germination/dormancy. Germination phenotype of the mutants compared to wild type is indicated, better germination (+) and reduced or no germination (-). Proteins are divided in the following classes; transcription factor (TF), regulatory protein (RP), enzyme (E), photoreceptor (P) and transporter protein(T).

Seed maturation mutants

Mutant	Gene	Germination phenotype	Encoded protein	Protein class	Reference	Accession number
abi3	*ABI3*	+	B3 domain protein with B1 and B2 domain	TF	Giraudat et al., 1992	AJ002473
fus3	*FUS3*	+	B3 domain protein with B2 domain	TF	Luerssen et al., 1998	AF016264
lec1	*LEC1*	+	HAP3 subunit of CCAAT box binding protein	TF	Lotan et al., 1998	AF036684
lec2	*LEC2*	+	B3 domain transcription factor	TF	Stone et al., 2001	AF400124

Abscisic acid biosynthesis and signaling mutants

Mutant	Gene	Germination phenotype	Encoded protein	Protein class	Reference	Accession number
abi1	*ABI1*	+	serine/threonine phosphatase 2C	RP	Leung et al., 1994; Meyer et al., 1994	X78886
abi2	*ABI2*	+	serine/threonine phosphatase 2C	RP	Leung et al., 1997	Y11840
abi4	*ABI4*	+*	APETELA2 domain protein	TF	Finkelstein, 1994; Finkelstein et al., 1998	AF040959
abi5	*ABI5*	+*	basic leucine zipper transcription factor	TF	Finkelstein, 1994; Finkelstein and Lynch, 2000	AC006921.5
aba1	*ABA1*	+	zeaxanthin epoxidase	E	Rock and Zeevaart, 1991; Meyer et al., 1994	AF283761
aba2	*ABA2*	+	xanthoxin oxidase	E	Schwartz et al., 1997; Léon-Kloosterziel et al., 1996a	
aba3	*ABA3*	+	molybdenum cofactor sulfurase	E	Schwartz et al., 1997; Xiong et al., 2001	AF325457
era1	*ERA1*	-	farnesyl transferase	RP	Cutler et al., 1996	AF214106

* germination on ABA concentrations that inhibit wild type germination

Appendix (Continued)

Gibberellin biosynthesis and signaling mutants

Mutant	Gene	Germination phenotype	Encoding protein	Protein class	Reference	Accession number
ga1	*GA1*	-	copalyl diphosphate synthase (CPS)	E	Koornneef and van der Veen, 1980; Sun et al., 1992	U11034
ga2	*GA2*	-	ent-kaurene synthase (EKS)	E	Koornneef and van der Veen, 1980; Yamaguchi et al., 1998	AF034774
ga3	*GA3*	-	ent-kaurene oxidase	E	Koornneef and van der Veen, 1980; Helliwell et al., 1998	AF047719
spy	*SPY*	+	threonine-O-linked N-acetylglucosamine transferase	RP	Jacobsen et al., 1996	U62135
sly1		-	unknown		Steber et al., 1998	

Ethylene and Brassinosteroid mutants

Mutant	Gene	Germination phenotype	Encoding protein	Protein class	Reference	Accession number
crt1	*CTR1*	+	raf family of protein kinases	RP	Kieber et al,1993; Beaudoin et al., 2000	L08790
ein2	*EIN2*	-	bifunctional transducer	RP	Alonso et al., 1999; Beaudoin et al., 2000	AF141202
etr1	*ETR1*	-	ethylene receptor with histidine kinase activity	RP	Bleecker et al., 1988; Chang et al., 1993	L24119
det2	*DET2*	-	steroid 5 α-reductase	E	Li et al.,1996; Steber and McCourt, 2001	U53860
bri1	*BRI1*	-	transmembrane receptor kinase	RP	Li and Chory, 1997; Steber and McCourt, 2001	AF017056

Appendix (Continued)

Seed coat mutants

Mutant	Gene	Germination phenotype	Encoding protein	Protein class	Reference	Accession number
ats		+	unknown		Léon-Kloosterziel et al., 1994	
tt1		+	unknown		Debeaujon et al., 2000	
tt2	*TT2*	+	R2R3 MYB domain protein	TF	Nesi et al., 2001	AJ299452
tt3	*DFR*	+	dihydroflavonol-4-reductase	E	Shirley et al., 1992	AB033294
tt4	*CHS*	+	chalcone synthase	E	Feinbaum and Ausubel, 1988	AF112086
tt5	*CHI*	+	chalcone isomerase	E	Shirley et al., 1992	M86358
tt6	*F3H*	+	flavonol 3-hydroxylase	E	Wisman et al., 1998; Pelletier and Shirley, 1996	U33932
tt7	*F3'H*	+	flavonol 3'-hydroxylase	E	Schoenbohm et al., 2000	AF155171
tt8	*TT8*	+	basic helix-loop-helix domain protein	TF	Nesi et al., 2000	AJ277509
tt9		+	unknown		Debeaujon et al., 2000	
tt10		+	unknown		Debeaujon et al., 2000	
tt11		+	unknown		Debeaujon et al., 2000	
tt12	*TT12*	+	MATE family protein	T	Debeaujon et al., 2001	AJ294464
tt13		+	unknown		Debeaujon et al., 2000	
tt14		+	unknown		Debeaujon et al., 2000	
tt15		+	unknown		Focks et al., 1999	
ttg1		+	WD40-repeat protein	RP	Walker et al., 1999	AJ133743
ban	*LAR*	+	leucoanthocyanidin reductase	E	Devic et al., 1999	AF092912

Appendix (Continued)

Miscellaneous

Mutant	Gene	Germination phenotype	Encoding protein	Protein class	Reference	Accession number
dag1	*DAG1*	+	DOF transcription factor	TF	Papi et al., 2000	AJ224122
dag2	*DAG2*	-	DOF transcription factor	TF	Vittorioso et al., 2001	AJ237810
rdo1		+	unknown		Léon-Kloosterziel et al., 1996a	
rdo2		+	unknown		Léon-Kloosterziel et al., 1996a	
rdo3		+	unknown		A.J.M. Peeters unpublished results	
rdo4		+	unknown		A.J.M. Peeters unpublished results	
cts		-	unknown		Russel et al., 2000	
phyA	*PHYA*		phytochrome A apoprotein	P	Sharrock and Quail, 1989; Shinomura et al., 1996	X17341
phyB	*PHYB*	-	phytochrome B apoprotein	P	Sharrock and Quail, 1989; Shinomura et al., 1996	X17342
hy1	*HY1*	-	ferredoxin-dependent heme oxygenase	E	Cone and Kendrick, 1985; Muramoto et al., 1999	AB021858
hy2	*HY2*	-	phytochromobilin synthase	E	Cone and Kendrick, 1985; Kohchi et al., 2001	AB045112
hot1	*HPS101*	-**	heat shock 101/ClpB protein	RP	Hong and Vierling, 2001	AF218796
icl	*ICL*	-***	iso-citrate lyase	E	Eastmond et al., 2000; Eastmond and Graham, 2001	AB025634
mls	*MLS*	-***	malate synthase	E	Eastmond and Graham, 2001	At5g03860

** reduced germination after high temperature treatment
*** reduced seedling establishment

Chapter 3

Analysis of natural allelic variation at seed dormancy loci of *Arabidopsis thaliana*

Carlos Alonso-Blanco*, Leónie Bentsink*, Corrie J. Hanhart, Hetty Blankestijn-de Vries and Maarten Koornneef

* These authors contributed equally to the work

Abstract

Arabidopsis accessions differ largely in their seed dormancy behaviour. To understand the genetic basis of this intraspecific variation we have analysed two accessions: the laboratory strain Landsberg *erecta* (L*er*) with low dormancy and the strong dormancy accession Cape Verde Islands (Cvi). We have used a quantitative trait loci (QTL) mapping approach to identify loci affecting the after-ripening requirement measured as number of days of seed dry storage to reach 50% germination. Thus, seven QTL were identified and named *delay of germination* (*DOG*) 1 to 7. To confirm and characterise these loci, we developed 12 near isogenic lines carrying single and double Cvi introgression fragments in a L*er* genetic background. The analysis of these lines for germination in water confirmed four QTL (*DOG1*, *DOG2*, *DOG3* and *DOG6*) as showing large additive effects in L*er* background. In addition, it was found that *DOG1* and *DOG3* genetically interact, the strong dormancy determined by *DOG1*-Cvi alleles depending on *DOG3*-L*er* alleles. These genotypes were further characterised for seed dormancy/germination behaviour in five other environmental conditions, including seed coat removal, gibberellins, and an abscisic acid biosynthesis inhibitor. The role of the L*er*/Cvi allelic variation affecting dormancy is discussed in the context of the current Arabidopsis germination knowledge.

This Chapter has been submitted for publication

Introduction

To survive in a particular location, plants have developed mechanisms that regulate seed germination at the most convenient season of the year. One of such mechanisms for proper timing of seed germination is seed dormancy, which can be defined as the temporary failure of an intact viable seed to complete germination under favourable conditions (Bewley, 1997). Large variations exist for this seed characteristic among and within plant species, which are considered adaptations to particular environments (Baskin and Baskin, 1998). Therefore, seed dormancy is an important adaptive trait, that is a primary component of the different life history strategies (winter and spring habits) of annual plants. In addition, seed dormancy is also an important agronomical trait since pre-harvest sprouting, problems with uniform germination and some seed processing properties (like malting in barley) are traits greatly determined by seed dormancy characteristics (Bewley, 1997).

Seed dormancy is a very complex trait due to firstly, the complex genetic structure of the seed. Seeds consist of 3 parts with different genetic composition: the embryo and endosperm of zygotic origin, and the seed coat or testa derived from maternal tissues. The three structures together determine the germination and dormancy behaviour of seeds. Germination begins with the uptake of water by the quiescent seed and ends with the elongation of the embryonic axis leading to the protrusion of the radicle through the seed coat (Bewley and Black, 1994). In the case of dormancy, this is established during seed development and may involve any of the seed structures. Thus, seed dormancy is first classified in two overall categories so-called embryo and seed coat-imposed dormancy (Bewley and Black, 1994). Secondly, seed dormancy is influenced by environmental factors such as light and temperature during seed development on the mother plants, during seed storage and during germination. In addition, seed dormancy disappears or is released during dry storage of the seeds, the time needed for that being referred to as ‘after-ripening requirement’. These characteristics make seed dormancy a trait that is difficult to quantify because even different seeds from the same genotype may loose their dormancy at different times. The measurement of seed dormancy is best achieved by estimating the after-ripening requirement of a large number of seeds and requires germination assays at different times during seed storage to determine the ‘average’ after-ripening requirement. In this way the ‘degree’ or ‘strength’ of the seed dormancy can be precisely estimated.

The genetical and environmental complexity of seed dormancy has determined that despite its fundamental and applied importance little is known about the molecular mechanisms underlying this trait (Bewley, 1997). However, in the past decade the model annual plant *Arabidopsis thaliana* has been shown to be an ideal species to perform genetic analyses, because of the resources developed by the international community including the availability of its complete genome sequence (Meinke et al., 1998). Furthermore, it has been shown that this species is also suitable for an effective analysis of seed dormancy (for a recent review see Bentsink and Koornneef, 2002). A large number of mutations affecting seed dormancy and germination have been generated artificially and the genetic, physiological and molecular characterization of these mutations is starting to shed light in the complexity of its regulation. For instance, mutants in genes such as *ABA-insensitive 3* (*ABI3*; Ooms et al, 1993; Nambara et al., 1995), *FUSCA 3* (*FUS3*; Bäumlein et al., 1994) *LEAFY COTYLEDONS* (*LEC1* and *LEC2;* Meinke et al*.,* 1994) with defective seed maturation are non-dormant, indicating that dormancy is part of the developmental program established during the later phases of seed development. Non-germinating mutants affected in the biosynthesis of the plant hormone gibberellin (GA; Koornneef and Van Der Veen, 1980) and the non-dormant mutants deficient in abscisic acid (ABA; Koornneef et al., 1982) have shown the important and opposite roles of these two phytohormones. Embryonic ABA has been correlated with the induction of dormancy, and it has been determined that the GA requirement for dormancy release and germination is abolished in the absence of ABA, indicating that GAs are needed to counteract the ABA dormancy effects. Moreover, the characterisation of reduced seed dormancy mutants affected in the maternally inherited testa pigmentation has revealed that the GA requirement for seed germination is not only determined by the embryonic ABA but also by the testa characteristics (Debeaujon and Koornneef, 2000). Light-induced stimulation of seed germination is affected in phytochrome photoreceptor deficient mutants (Casal and Sanchez, 1998) and a phytochrome effect has also been suggested in the onset of dormancy on the mother plant (McCullough and Shropshire, 1970; Hayes and Klein, 1974). Moreover, several genes encoding transcription regulators such as *DOF affecting germination* (*DAG*; Papi et al., 2000; Gaulberti et al., 2002), *FUS3* (Luerssen et al., 1998), *LEC1* and *LEC2* (Lotan et al. 1998; Stone et al., 2001) and several genes with unknown functions such as those disrupted in the *reduced dormancy* 1 to 4 mutants (*rdo*; Léon-Kloosterziel et al., 1996; Peeters et al., 2002) have been implicated.

In addition to artificially induced mutations, genetic variation for seed dormancy and germination characteristics has been described for long time among Arabidopsis wild populations (Kugler 1951; Ratcliffe, 1976; Lawrence, 1976). Arabidopsis accessions collected at different geographical locations show a quantitative pattern of variation for light requirement (Kugler, 1951; Napp-Zinn, 1975) and for the after-ripening requirement (Ratcliffe, 1976; Lawrence, 1976). The genetic analysis of this natural variation has been attempted in some early studies. Kugler (1951) showed that the light dependency for germination of the accession Hannovrisch Münden (Hm) was recessive in crosses with the dark germinating accessions Stockholm (St) and Haarlem (Haa). Further analysis of F3 lines derived from the cross Hm x St by Napp-Zinn (1975) suggested that 3 loci determined the light requirement difference between both parents. However, the dissection of the multifactorial genetic variation into the individual loci has become feasible only recently, by using quantitative trait loci (QTL) mapping procedures. This approach has been applied in the study of seed dormancy variation in crop species, by analysing crosses between cultivated varieties like in wheat and barley (Anderson et al., 1993; Kato et al., 2001; Romagosa et al., 1999; Ullrich et al, 1993), crosses between domesticated species and their wild relatives such as in rice (Cai and Morishima, 2000), or crosses between wild relatives like in wild oat (Fennimore et al., 1999). In addition, it has been used in Arabidopsis to analyse a cross between the two most widely used laboratory accessions, Landsberg *erecta* (L*er*) and Columbia (Col), which show a low level of dormancy (van der Schaar et al., 1997). In this study, despite the small parental differences, the combination of recombinant inbred lines (RILs) and multiple QTL model (MQM) mapping methods allowed the identification of 14 loci of small effect accounting for the dormancy and germination differences between both accessions. Once the main QTL have been identified, the individual loci can be further characterised and fine mapped by developing near isogenic lines (NILs) with monogenic differences. This approach can be efficiently used in model species such as Arabidopsis since the analysis can be easily followed up to the molecular level enabling the identification of the genes underlying the genetic variation at individual QTL (Alonso-Blanco and Koornneef, 2000; Remington et al., 2001). Thus, the analysis of this source of genetic variation constitutes an important resource for the functional analysis of seed dormancy. The study of more dormant accessions might contribute to the identification of novel loci and/or alleles, since most mutant analyses have been performed in the low dormancy accessions. In addition, the identification of the genes accounting for the variation among Arabidopsis populations will contribute to the understanding of the ecological and evolutionary mechanisms involved in the development of different life history strategies of annual plants and in adaptation to different environments. This is being illustrated with the analysis of another important life history trait like flowering time for which major genetic determinants of the existing natural variation have been already identified in Arabidopsis and the respective genes have been cloned (Michaels and Amasino, 1999; Sheldon et al., 1999; Johanson et al., 2000; El-Assal et al., 2001).

In the present work we have analysed two Arabidopsis accessions differing largely in their seed dormancy behaviour: the low dormancy laboratory strain L*er*, and the very dormant strain Cape Verde Islands (Cvi). We have identified the loci accounting for the after-ripening requirement using a set of RILs derived from a cross between both accessions. Four major effect QTL were confirmed and further characterised genetically and physiologically by analysing NILs carrying specific Cvi introgression fragments in a L*er* genetic background. The possible role of this allelic variation on seed dormancy is discussed in the context of the current Arabidopsis germination knowledge.

Results

Seed dormancy behaviour of Ler, Cvi and their recombinant inbred lines

The seed dormancy behaviour of L*er*, Cvi and reciprocal F1 and F2 hybrid seeds was analysed by characterising their germination phenotypes along the time of seed dry storage (time of after-ripening). Germination was tested at six different times of storage from the harvest date until all the seeds germinated, and curves of germination percentage over time of storage were obtained (Figure 3.1). From these, the number of days of seed dry storage (after-ripening) required to reach 50% germination ($DSDS_{50}$) was estimated for each genotype, as a single measurement of seed dormancy. Although $DSDS_{50}$ values vary among experiments due to mainly environmental effects on the mother plants (Figure 3.1A and B) L*er* seeds in general germinated 100% after 6 weeks of storage, while Cvi seeds required at least 15 weeks to do so. When comparing plants grown together, the $DSDS_{50}$ values of Cvi were at least five times higher than L*er* values, L*er* varying from 12 to 17 days in different experiments while Cvi $DSDS_{50}$ values ranging from 74 to 185 days. Therefore, L*er* and Cvi differ largely in their seed dormancy, Cvi seeds being much more dormant than L*er* seeds. Maternal genetic effects on the dormancy variation were not detected as deduced from

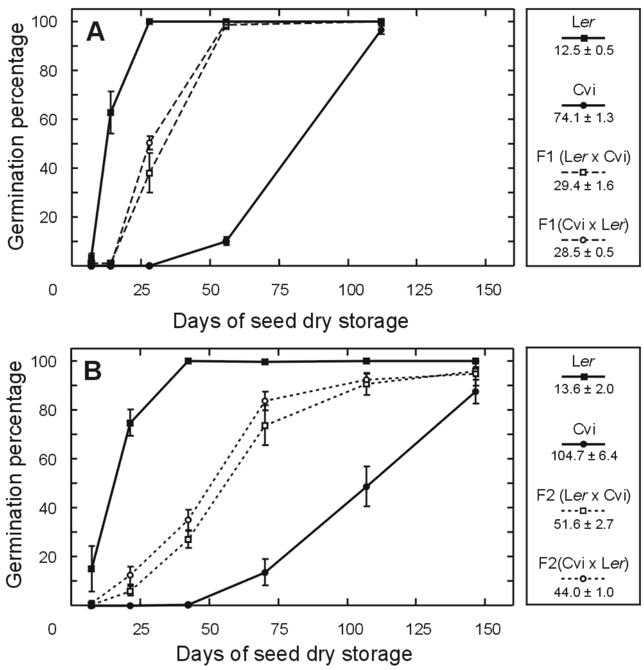


Figure 3.1. Dormancy/germination behaviour of L*er*, Cvi, and their reciprocal F1 and F2 hybrid seeds along the time of dry storage. Percentages of germination in water under white light were estimated at different times of seed dry storage, and germination curves for each genotype are presented. Plants grown in two different experiments (A and B) are shown to indicate the environmental effects on germination and to illustrate that genotypes from different experiments cannot be directly compared. The after-ripening requirement was estimated for each genotype in each experiment as the $DSDS_{50}$ (± SE) value given in the legends. In A, mean (± SE) germination percentage of 4 different seed bulks of 2 plants each is shown for the parental lines, and of 4 seed bulks of 3 different crosses for the F1 hybrid seeds. In B, each data point corresponds to the mean (± SE) of germination percentage of 4 seed bulks of 6 plants.

phenotypic comparisons of reciprocal F1 and F2 seeds (Figure 3.1). F1 hybrid seeds obtained using L*er* as mother plants (or F2 derived seeds) did not differ significantly from F1 hybrid seeds obtained using Cvi as mother plants (or respectively, from F2 seeds) either in the $DSDS_{50}$ or in the germination percentage at any tested time. In addition, the reciprocal F1 and F2 seeds showed $DSDS_{50}$ values intermediate between the parental values, indicating an overall co-dominance of L*er* and Cvi alleles. However, the germination curves consistently indicated a partial dominance of the strong Cvi dormancy during the first 3 weeks of seed storage, whereas L*er* germination phenotype appeared partly dominant after longer periods than 10 weeks of storage, which might be a consequence of the limited quantitative scale of measurement of germination percentages.

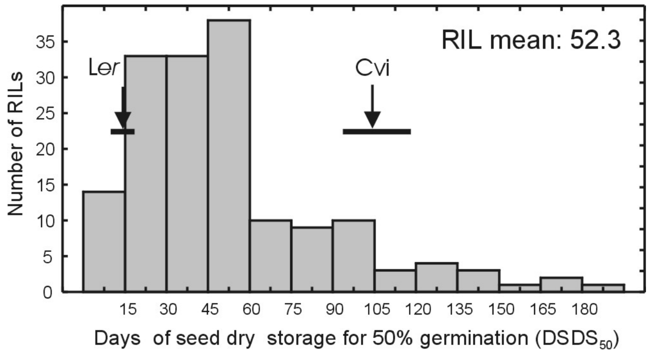


Figure 3.2. Frequency distribution of after-ripening requirements ($DSDS_{50}$ values) of the L*er*/Cvi RILs. Arrows depict the mean $DSDS_{50}$ of the parental lines estimated in 4 different seed bulks of 6 plants (3 germination assays per bulk and per time of seed storage), and horizontal bars represent their range of variation (same parental lines shown in Figure 3.1B). The $DSDS_{50}$ value of the RILs were estimated as a single value from a single seed bulk of 12 plants (3 germination assays per time of seed storage). RIL mean: average $DSDS_{50}$ of the RIL population.

The genetic variation for seed dormancy of L*er* and Cvi was further analysed by studying the germination along the time of seed storage of 161 RILs derived from crosses between both parental lines (Alonso-Blanco et al., 1998b). As shown in Figure 3.2, some transgression in both directions was detected for the $DSDS_{50}$ value in the RIL population, indicating that both parental lines carry alleles increasing and decreasing seed dormancy. Analysis of the frequency distributions of germination percentages after 1, 3, 6, 10, 15 and 21 weeks of seed storage (Figure 3.3) shows the kinetics of dormancy of the RILs. The mean germination of the RIL population increased gradually from 1.9% in one-week-old seeds, up to 94.7% in 21-week-old seeds. However, considerable phenotypic variation is present in the RIL population at each time point of seed storage, and transgression in both parental directions is observed at different times of storage; transgression towards reduced dormancy could be detected during the first three weeks of storage, whereas transgression over the Cvi parent appeared detectable after 15 weeks of seed storage (Figure 3.3). Genotype x environment (time) interactions were significant (*P* < 0.001) for any comparison of RIL germination percentages among times of seed storage, showing that RILs respond differently to seed storage. Therefore, the phenotypic effects of the dormancy allelic variation are expressed differently along the time of seed storage.

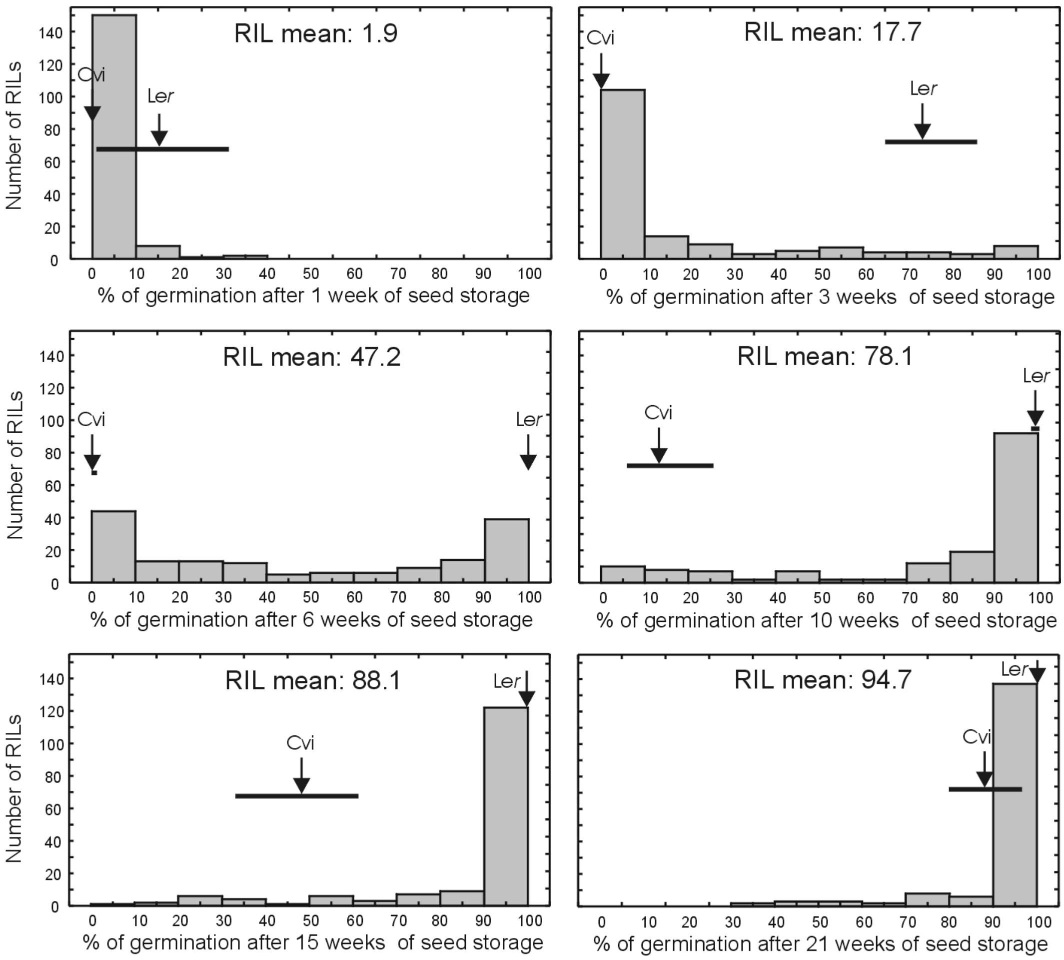


Figure 3. 3. Frequency distributions of mean germination percentage of the L*er*/Cvi RILs after 1, 3, 6, 10, 15 and 21 weeks of seed dry storage. Arrows depict the means of the parental lines estimated from 4 different seed bulks of 6 plants (3 germination assays per bulk), and horizontal bars represent their range of variation (same parental lines shown in Figure 3. 1B and 2). Mean germination percentages of RILs are estimated from 3 germination assays of a single seed bulk of 12 plants. RIL mean: average germination percentage of the RIL population.

Mapping seed dormancy loci in the Ler/Cvi RIL population

To identify the loci that control the L*er*/Cvi seed dormancy variation, QTL analysis was performed using the RIL phenotypic values of the time of seed storage required for 50% germination ($DSDS_{50}$). Conservatively, seven QTL were identified located on all chromosomes except chromosome 2, their total additive effects accounting for 61.4% of the after-ripening requirement phenotypic variation (Figure 3.4). These loci have been named as *delay of germination* (*DOG*) 1 to 7, the locus number being given according to their relative effect from larger to smaller effect. Cvi alleles at six loci increased the time of seed storage required for seed germination (increased dormancy) and only Cvi alleles at *DOG2*, located on the top of chromosome 1, reduced dormancy as compared with the L*er* allele. Three of the loci, *DOG1* to *DOG3*, showed large additive effects (each explaining more than 10% of the phenotypic variation) and together accounted for about 60% of the additive genetic variance (38.2% of the total variance). Specially complex appeared the region of chromosome 5 containing *DOG1* and *DOG7*, which using the MQM module of MapQTL could be located 20cM apart, and therefore their single QTL additive effects might be underestimated. In addition, this region might contain more than two segregating loci affecting seed dormancy, which could not be genetically distinguished.

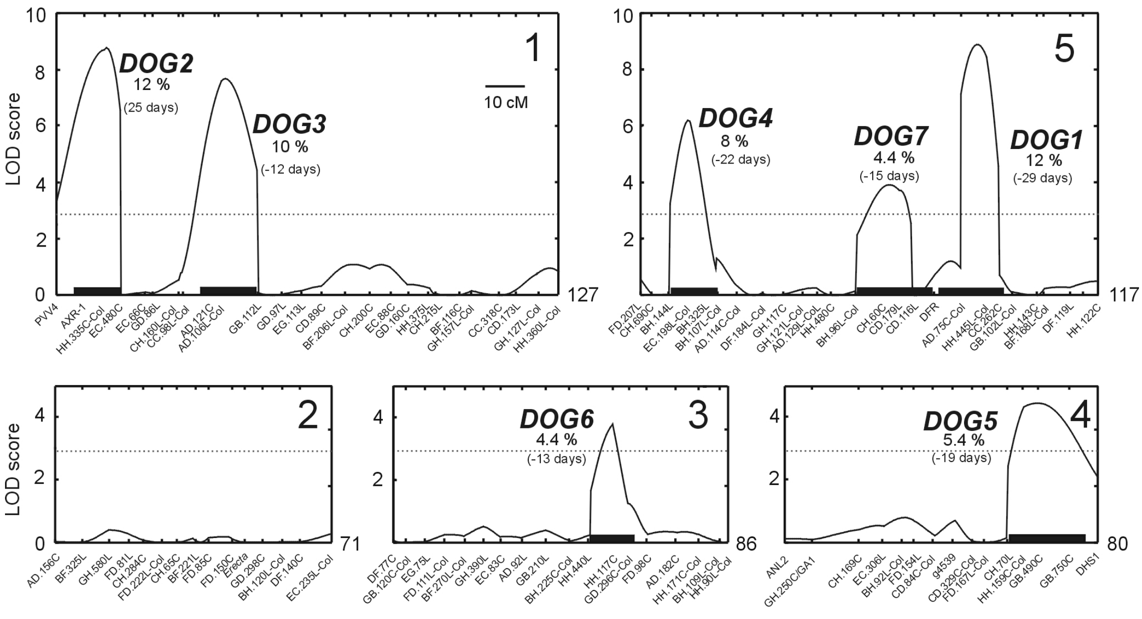


Figure 3.4. QTL likelihood maps for the after-ripening requirement ($DSDS_{50}$ values) of the L*er*/Cvi RILs. The abscissas corresponds to the genetic maps in cM, the linkage group number being indicated on the upper right corner of each map. Horizontal dashed line indicates the LOD score threshold of 2.7. Two-LOD support intervals are shown as solid bars along abscissas. For each QTL, its name, additive genetic effect and the percentage of the total variance explained by its additive effect are given. Negative additive effects indicate that Cvi genotypes have higher $DSDS_{50}$ mean than L*er* (higher dormancy), while positive effects imply that Cvi genotypes have the lower $DSDS_{50}$ mean.

Table 3.1. QTL for seed germination/dormancy detected at different times of seed dry storage. The closest marker to each QTL is shown and its location is indicated by the linkage group followed with the map position. The significant QTL are shown at the 6 times of seed storage, and for each QTL the percentage of variance explained and the additive effect is given. The total percentage of explained variance by the additive effects of all the significant QTL detected at each time of seed storage is given in the first data row. QTL detected with LOD scores between 2 and 2.7 are included and indicated with [1]. Additive effects are given as the difference between the means of the two RIL genotypic groups (a negative value implies that Cvi alleles reduce germination -increase dormancy- as compared to L*er* alleles; a positive value indicates that Cvi alleles increase germination). Additive effects are shown in the untransformed scale of percentages of germination as merely orientative. The genomic region containing *DOG1* and *DOG7* appeared genetically complex due to the linkage between these two QTL and to their different relative effects at the various times of seed storage. For this reason, cofactor markers for these two QTL were selected at slightly different positions in the different analyses, and instead of a single marker, an interval spanning the position of the cofactors is given. In addition, a single cofactor for both QTL was selected (DFR marker) in the 21 weeks analysis, due to their low individual effect. QTL interacting with environment (time of seed storage) are indicated by * ($P < 0.005$). QTL x environment interactions were tested using only the four data sets collected between 3 to 15 weeks of seed storage, at which considerable variation was detected. NS, not significant.

QTL			Time of Seed Storage												QTL x E interaction
			1 week		3 weeks		6 weeks		10 weeks		15 weeks		21 weeks		
Closest marker	QTL name	Map position	% Expl. variance	Additive effect	% Expl. variance	Additive effect	% Expl. variance	Additive effect	% Expl. variance	Additive effect	% Expl. variance	Additive effect	% Expl. variance	Additive effect	
			27.8		52.4		64.4		58		46.2		37.6		*
HH.335C-Col	*DOG2*	1-12.1	NS	NS	6.8	14.7	13.0	28.9	11.6	21.5	13.7	15.3	10.2	6.4	*
AD.106L-Col	*DOG3*	1-41.7	14.4	-3.3	21.8	-26.8	10.3	-28.1	NS	NS	NS	NS	NS	NS	*
HH.117C	*DOG6*	3-56.3	5.3 [1]	-2.1	7.1	-15.0	4.7	-17.3	2.6 [1]	-10.1	NS	NS	NS	NS	*
HH.159C-Col	*DOG5*	4-61.0	NS	NS	3.8 [1]	-11.7	3.2	-12.5	5.4	-13.9	6.9	-10.6	10.8	-7.5	NS
EC.198L-Col	*DOG4*	5-12.8	4.7 [1]	-2.4	6.7	-16.7	5.4	-18.9	6.6	-16.9	8.9	-14.5	9.1	-8.8	NS
BH.96L-Col to CD.179L	*DOG7*	5-55.5 to 66.3	NS	NS	NS	NS	4.1	-17.4	2.6 [1]	-11.1	5.2	-10.3	NS	NS	*
DFR		5-75.5	-	-	-	-	-	-	-	-	-	-	8.0	-7.1	
AD.75C-Col to GB.102L-Col	*DOG1*	5-82.1 to 94.9	10.1	-3	19.1	-22.6	16.8	-38.8	18.7	-32.3	7.3	-12.9	NS	NS	*

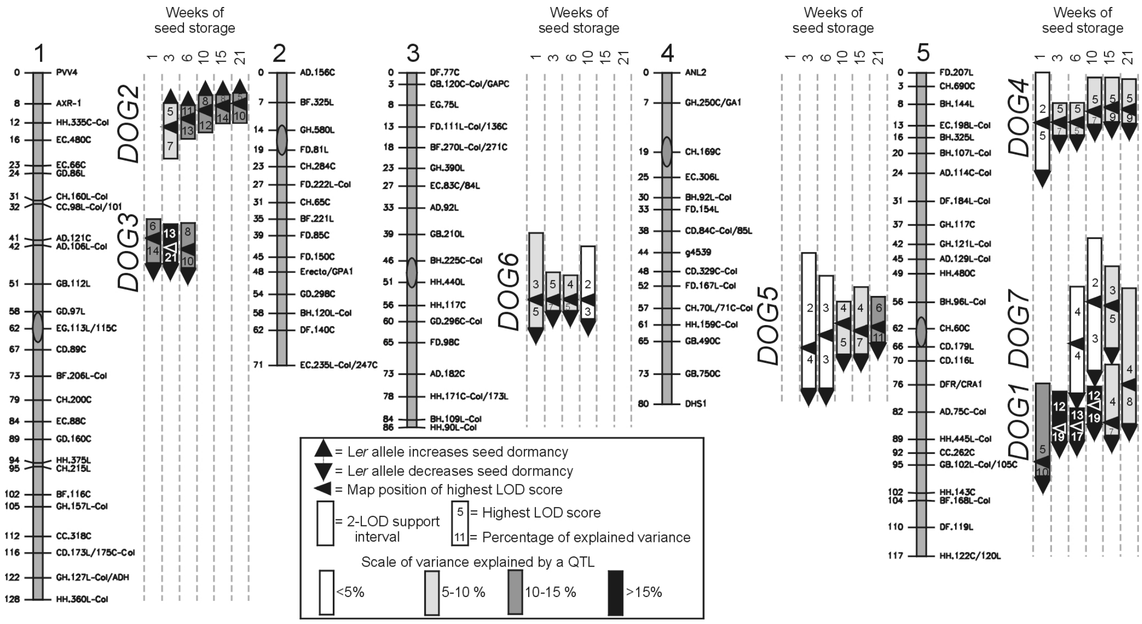


Figure 3.5. QTL mapping for germination of the L*er*/Cvi RILs at different times of seed storage. The significant QTL detected in the 6 germination assays performed at different times of seed storage are depicted close to the L*er*/Cvi genetic maps of the five linkage groups.

To further characterise genetically the *DOG* loci, we used the RIL germination percentages at the 6 different times of seed storage (1, 3, 6, 10, 15 and 21 weeks) for QTL analyses. Thus, the QTL responsible for the germination variation at each time of seed storage could be identified and their additive genetic effects followed along the time of seed storage (Table 3.1 and Figure 3.5). A conservative total number of seven QTL were identified along the six germination assays, corresponding to the same seven genomic regions previously identified using the $DSDS_{50}$ values. No other significant QTL could be detected consistently in more than one germination assay. Therefore, the detected loci affecting germination are the same *DOG* loci affecting the after-ripening requirement variation. As shown in Figure 3.5 and Table 3.1, the germination percentages at each time of seed storage detected between 2 to 7 significant QTL, their combined additive effects accounting for between 27.8% and 64.4% of the total variance. Consistently with the effects of the *DOG* loci on the after-ripening requirement, Cvi alleles at six of the seven QTL decreased the percentage of germination (increased dormancy) while only Cvi alleles at *DOG2* increased the percentage of germination, as compared with L*er* alleles. Similar to the $DSDS_{50}$ analysis the region on chromosome 5 between map positions 56 (BH.96L-Col) and 95 (GB.102L-Col) appeared as a complex region containing at least two genetically linked loci, *DOG1* and *DOG7*, with phenotypic effects in the same direction.

Table 3.2. Two-way QTL interactions. The transformed germination percentages at the 6 different times of seed storage were tested for interactions between the pairs of markers closest to the identified QTL and used as cofactors in the mapping analyses shown in Table 1. Statistical significance and R-square values are shown for the significant interactions. Significant genetic interactions were classified in three classes: A, the two alleles reducing dormancy interact synergistically; B, the two alleles increasing dormancy interact synergistically; C, L*er* alleles at *DOG3* reduce germination (increase dormancy, which is opposite to its non-dormant additive effect detected in the QTL mapping) in the presence of Cvi alleles at *DOG1*, *DOG4* or *DOG6*. The percentage of total variance explained by the additive effects of the various significant QTL and the significant two way interactions detected at each time of seed storage is given. NS, not significant ($P > 0.01$).

	Time of seed storage																	
	1 week			3 weeks			6 weeks			10 weeks			15 weeks			21 weeks		
Two-way Interaction	*p*	class	R^2	*p*	class	R^2	*p*	class	R^2	*p*	class	R^2	*p*	class	R^2	*p*	class	R^2
DOG1 x *DOG2*	NS	-	-	NS	-	-	NS	-	-	0.001	B	6.9	0.001	B	6.7	0.009	B	4.3
DOG1 x *DOG3*	0.012	A	4.1	0.002	A	6.3	0.003	A	5.5	0.001	C	6.9	0.001	C	7.3	0.002	C	5.9
DOG1 x *DOG4*	NS	-	-	NS	-	-	NS	-	-	NS	-	-	0.006	B	4.8	0.003	B	5.6
DOG1 x *DOG6*	0.01	A	4.3	0.001	A	7.0	NS	-	-	NS	-	-	NS	-	-	NS	-	-
DOG1 x *DOG7*	NS	-	-	NS	-	-	NS	-	-	0.008	B	4.4	0.004	B	5.2	NS	-	-
DOG2 x *DOG7*	NS	-	-	NS	-	-	NS	-	-	0.002	B	6.3	0.009	B	4.4	NS	-	-
DOG3 x *DOG4*	NS	-	-	NS	-	-	NS	-	-	NS	-	-	NS	-	-	0.007	C	4.6
DOG3 x *DOG7*	0.002	A	6.2	0.001	A	6.7	0.003	A	5.8	0.005	C	5	0.009	C	4.4	NS	-	-
Percentage of Total Explained Variance			35			57			67			63.5			51			39

Since the RILs were obtained from reciprocal crosses (Alonso-Blanco et al., 1998b) maternal cytoplasmic effects on the seed dormancy parameters could be analysed but no significant effect was detected either additive or as interacting with any of the nuclear markers.

Analysis of QTL x E interactions showed that 5 of the 7 detected loci have significantly different additive effects in the various times of seed storage (environments; Table 3.1). However, the relative effects of these QTL showed different trends along the time of storage (Figure 3.5). Thus, the seven QTL could be classified in three different classes according to the behaviour of their additive genetic effects along the time of seed storage: i) *DOG1*, *DOG3* and *DOG6* show a larger effect in the germination assays carried out during the first 6 weeks of seed storage, their maximum appearing between weeks 3 to 6, thereafter their additive effects decreased. ii) *DOG2* and *DOG7* show a complementary behaviour to the previous class since they have a larger effect in the assays performed after week 6. These loci show small effects in the germination assays performed during the first 3 weeks, and increased gradually their relative additive effect to reach its maximum between weeks 10 to 21 (Figure 3.5). iii) *DOG4* and *DOG5* showed no interaction with the environments appearing as small effect loci in all assays. Therefore, the 7 QTL behave genetically different, suggesting that they might participate in different aspects of seed dormancy.

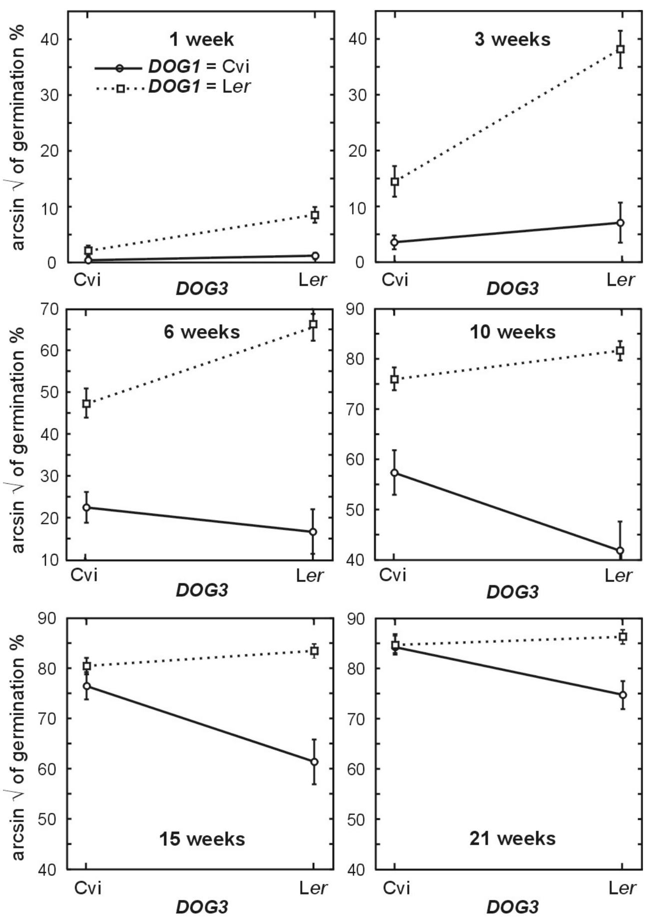


Figure 3.6. Two way interaction between *DOG1* and *DOG3*. Each panel correspond to the graphical representation of the mean (±SE) germination of the four genotypic RIL classes after 1, 3, 6, 10 15 and 21 weeks of seed storage.

Two-way interactions were analysed among the 7 dormancy QTL identified. When testing the interactions using the transformed germination percentages at each of the 6 different times of seed storage, several interactions involving all QTL except *DOG5* appeared as significant ($P < 0.005$; Table 3.2). Two-way interactions were also scanned throughout the genome by analysing all pair-wise combinations of markers, but no significant genetic interaction was consistently found in several germinations assays. The interactions detected with less than 6 weeks of storage time (when seeds are mostly dormant) corresponded to synergies between non-dormant alleles; in contrast, most interactions detected in the germination assays after long storage times (when most seeds germinate) showed synergies between dormant alleles. These interactions might be interpreted as a consequence of the limited measurement scale of percentages. In addition, epistatic effects of most loci were detected only in the same germination assays at which additive effects were previously found. However, *DOG3* showed particular interallelic interactions, also in the assays performed after 10 weeks of seed storage (Table 3.2; Figure 3.6). The interactions between *DOG3* and *DOG1*, *DOG4* and *DOG7* detected different allele effects at *DOG3* in short seed storage times than in long ones. This is illustrated in Figure 3.6 for the *DOG1* x *DOG3* interactions detected with the germination percentages in the 6 storage times. On average, *DOG3*-L*er* alleles reduced dormancy in the first two assays, similar to the *DOG3* additive effects estimated in the QTL mapping analyses. In contrast, in the later assays, *DOG3*-L*er* alleles increase dormancy in the presence of dormant alleles at the interacting QTL. This conditional *DOG3* allele effect depending on the time of storage and the genotype at several interacting loci appears not simply due to the genetic linkage of *DOG2*, whose L*er* alleles increasing dormancy have a maximum additive effect after long times of seed storage (Figure 3.5) because similar significant *DOG1* x *DOG3* interactions appear when considering only RILs with the same allele at *DOG2*. Furthermore, this *DOG1* x *DOG3* interaction was the only significant interaction detected among the 7 QTL when using the $DSDS_{50}$ values (P = 0.0002), *DOG3*-L*er* alleles showing also in this case opposite effects depending on the genotype at *DOG1*. These interactions suggests that *DOG3* alleles affect the *DOG1* effects, although we cannot discard more complex explanations such as that in the *DOG3* genomic region there are two closely linked loci with opposite effects affecting dormancy differentially along time of seed storage.

Genetical and physiological characterisation of the *DOG* loci

To characterise the various loci, 12 introgression lines carrying one or two Cvi genomic fragments around the *DOG* QTL regions into an otherwise L*er* genetic background were developed by phenotypic and genotypic selection (see Materials and Methods). These near isogenic lines (NILs) were thoroughly genotyped and the genetic position and size of the introgressions were determined (Figure 3.7). Nine of the lines carried

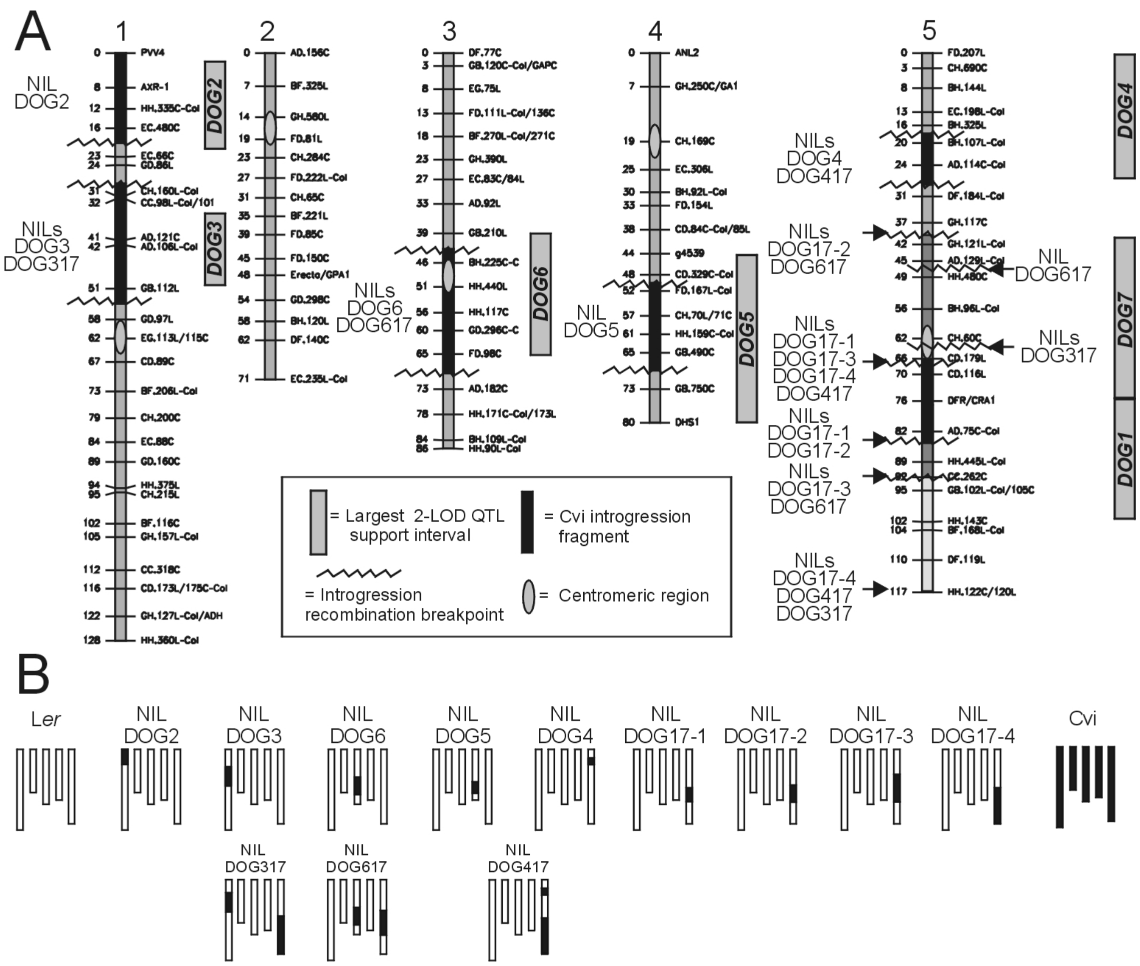


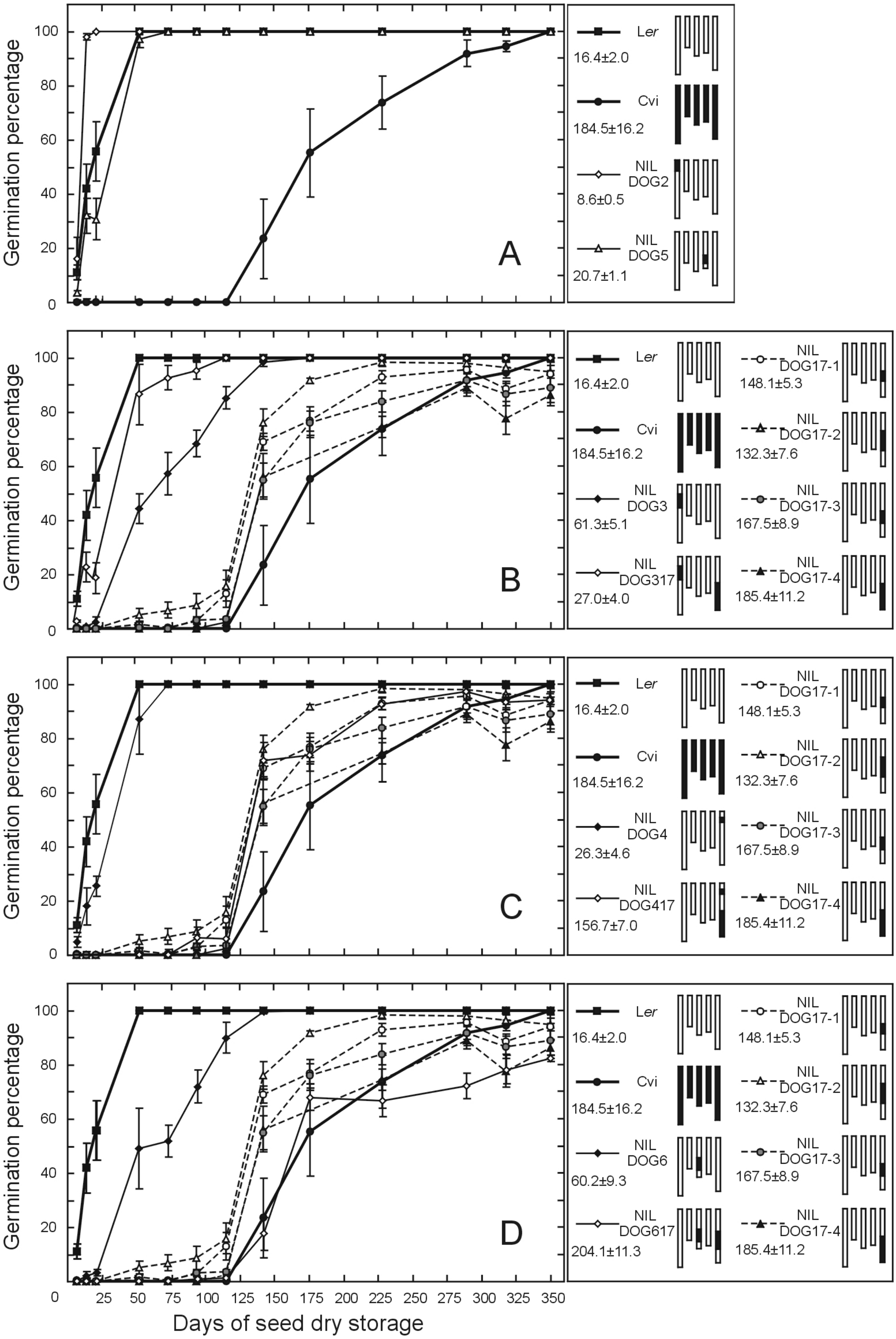
Figure 3.7. Genotype of the dormancy QTL NILs. A) Graphical representation of the Cvi introgression fragments of the 12 NILs in relation to the map position of the dormancy QTL. Support intervals for each QTL are represented as gray blocs at the right side of each chromosome, and correspond to the largest two-LOD support interval estimated in the QTL mapping experiments. Cvi introgression fragments are depicted as black or gray (when overlapping fragments were available) blocs on the chromosomes. Name(s) of the NIL(s) carrying an introgression fragment are shown at the left side of the corresponding fragment. For the *DOG1*/*DOG7* region on chromosome 5, overlapping introgression fragments were developed and in this case, the name of the NILs refer to recombination breakpoints of the introgression fragments instead to the introgression itself; this is shown with an arrowhead pointing to the corresponding breakpoint. B) Graphical genotype of the parental lines L*er* and Cvi, the 9 NILs carrying a single introgression fragment (upper row) and the 3 NILs with two introgression fragments (lower row).

single Cvi introgressions in the 6 genomic regions containing the identified *DOG* loci (NILs named as DOG2, DOG3, DOG4, DOG5, DOG6 and DOG17-1, DOG17-2, DOG17-3 and DOG17-4). In addition, 3 other NILs carried a Cvi introgression fragment of the complex region *DOG1*/*DOG7* and a second introgression in either the region of *DOG3*, *DOG4* or *DOG6* (NILs named as DOG317, DOG417 and DOG617).

The germination behaviour of these lines was analysed in water under light (Figure 3.8) aiming to confirm the existence of the QTL according to their effects in a L*er* genetic background and, in some cases, to determine the genetic interactions between the largest effect QTL *DOG1* and the remaining loci.

The dormancy behaviour of the single introgression lines measured by curves of germination percentage over time of seed dry storage and $DSDS_{50}$ values enabled to confirm several loci (Figure 3.8). Four very dormant NILs, DOG17-1 to DOG17-4, carrying introgressions of slightly different sizes around the *DOG1* and *DOG7* QTL (Figure 3.7 and 3.8) were analysed and compared. NIL DOG17-1 carried the smallest Cvi introgression of about 20cM, and was only slightly less dormant than Cvi (Figure 3.8B). Therefore in this small region between positions 65 to 85 cM of chromosome 5 we could assign the strongest Cvi dormant alleles, confirming the locus *DOG1*. However, it is not known if the strong dormancy of this line is determined by a single locus, *DOG1*, or by the two linked QTL *DOG1* and *DOG7* previously mapped in that region. Since the complexity shown in the RIL mapping experiments suggests that this region contains more than one locus closely linked, at this stage we do not claim that the dormancy difference between this line and L*er* is monogenic, since both QTL might participate. For this reason, we named this lines DOG17-1 and further mapping is needed to establish if Cvi alleles at more than one locus are introgressed. This line shared its lower recombination breakpoint with NIL DOG17-2, whose 50 cM introgression included an additional 30 cM region not present in NIL DOG17-1, which span the 2-LOD support interval of *DOG7* (Figure 3.7). These two NILs did not differ significantly in their germination behaviour (Figure 3.8B), concluding that there is no dormancy QTL detectable in the dormant background shared with NIL DOG17-1 introgression, located in the region comprised between the upper breakpoints of both lines (between positions 40 to 65 cM of chromosome 5). In addition, another line, DOG17-3, shared the upper recombination breakpoint with NIL DOG17-1 but carried

Figure 3.8. Dormancy/germination behaviour of L*er*, Cvi, and NILs along the time of seed dry storage. Percentages of germination in water under white light were estimated at different times of seed dry storage, and curves of germination for each genotype are presented. Mean (± SE) of the germination percentage of 4 different seed bulks from 3 plants is shown for each genotype. The graphical genotypes of the lines are shown in the legends together with the estimated $DSDS_{50}$ (±SE). All plants were grown in a single experiment and therefore the various genotypes are directly comparable. A) Dormancy behaviour of NILs DOG2 and DOG5 compared to L*er* and Cvi; B) dormancy behaviour of NILs DOG3, DOG17-1 to DOG17-4, DOG317 and the parental lines; C) dormancy behaviour of NILs DOG4 and DOG417 compared to NILs DOG17-1 to DOG17-4 and to the parental lines; D) dormancy behaviour of NILs DOG6 and DOG617 compared to NILs DOG17-1 to DOG17-4 and to the parental lines.


Germination percentage
Days of seed dry storage
A
B
C
D
Ler
16.4±2.0
Cvi
184.5±16.2
NIL DOG2
8.6±0.5
NIL DOG5
20.7±1.1
NIL DOG3
61.3±5.1
NIL DOG317
27.0±4.0
NIL DOG17-1
148.1±5.3
NIL DOG17-2
132.3±7.6
NIL DOG17-3
167.5±8.9
NIL DOG17-4
185.4±11.2
NIL DOG4
26.3±4.6
NIL DOG417
156.7±7.0
NIL DOG6
60.2±9.3
NIL DOG617
204.1±11.3
0 20 40 60 80 100
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350

an additional distal region of about 10 cM. These lines did not differ significantly in their dormancy behaviour (*P* < 0.05) as well. A forth line, NIL DOG17-4, had also a common upper breakpoint with NIL DOG17-1 but carried an additional 30 cM distal fragment. However, this line was slightly but significantly (*P* < 0.05) more dormant than the other NILs DOG17 (Figure 3.8B), suggesting that smaller effect Cvi alleles increasing dormancy are located on chromosome 5 between positions 85 and 117.

NILs carrying single Cvi fragments around the QTL *DOG2*, *DOG3*, and *DOG6* showed also significantly different germination behaviour than L*er* (*P* ≤ 0.005; Figure 3.8A, 3.8B and 3.8D) confirming the lower dormancy Cvi alleles at *DOG2* and the stronger dormancy ones at *DOG3* and *DOG6*. In addition, in a L*er* genetic background, the effect of *DOG6* and *DOG3* appears similar ($DSDS_{50}$ of 60.2 ± 9.3 and 61.3 ± 5.1 for NILs DOG6 and DOG3 respectively) both showing larger effect than *DOG2*. These results contrast with the relative low effect of *DOG6* and stronger relative effect of *DOG2* estimated in the RIL QTL mapping experiments (Figure 3.4 and 3.5) indicating the presence of genetic interactions.

NILs carrying single Cvi fragments around the QTL *DOG4* and *DOG5* did not differ significantly in their germination behaviour from L*er* wild type (Figure 3.8A and 3.8C), and therefore we could not confirm these loci. Both loci showed rather small additive effects in the QTL mapping experiments, and in a L*er* background they might become not easily detectable. However, it might also be possible that the introgression fragments of these lines did not include the corresponding loci. In the case of NIL DOG4, the introgression only carried one third of the region corresponding to the 2-LOD support interval, while in the case of NIL DOG5 this was about half of its region.

Furthermore, the genetic interactions between *DOG1/DOG7* and *DOG3*, *DOG4* and *DOG6* in an otherwise L*er* genetic background could be tested with the analysis of the three NILs carrying two Cvi introgressions (Figure 3.8B, 3.8C, 3.8D). NIL DOG417 showed similar germination behaviour than NIL DOG17-4 (Figure 3.8C), and therefore *DOG4* could not be confirmed either in a L*er* background or in a *DOG1/DOG7* dormant background. In contrast, NIL DOG617 was the strongest dormancy line ($DSDS_{50}$ = 204 ± 11.3), being significantly more dormant than NIL DOG17-4 and Cvi when comparing germination percentages after storage times longer than 3 months (*P* < 0.005). Therefore, the overall effects of the allelic variation at *DOG1* and *DOG6* in a L*er* background are additive, as deduced from the germination curves of these NILs (Figure 3.8D), in agreement with the behaviour observed in the RIL population. In contrast, the line carrying Cvi alleles at *DOG1/DOG7* and *DOG3* regions, interestingly, behaved as a non-dormant line, not differing significantly from Ler (*P* > 0.05). Therefore, the alleles at *DOG3* and *DOG1/DOG7* regions strongly interact, confirming the interaction observed in the epistasis analysis of the RIL population. This interaction indicates that the Cvi alleles at *DOG1/DOG7* require L*er* alleles in the *DOG3* region to produce the strong dormancy, in agreement with the larger effect of *DOG1* in a *DOG3*-L*er* than in a *DOG3*-Cvi background estimated in the RIL population analysis (Figure 3.6). In other words, Cvi alleles in the *DOG3* region increased dormancy in a *DOG1/DOG7*-L*er* background, but reduced dormancy in a *DOG1/DOG7*-Cvi genetic background confirming the epistasis analysis shown in Figure 3.6.

To characterize physiologically the 4 dormancy loci confirmed in NILs, the seed germination behaviour of these lines, L*er* and Cvi was analysed in 4 additional environmental conditions known to reduce dormancy (Bentsink and Koornneef, 2002). Germination of the 8 NILs showing a dormancy behaviour significantly different from L*er* was tested at different times of seed dry storage, after a cold treatment, or in the presence of nitrate, or with the hormone GA_{4+7}, or with the inhibitor of ABA biosynthesis norflurazon (NOR). The response of the various genotypes to these treatments was measured by comparing the $DSDS_{50}$ values obtained in water with those obtained with the corresponding treatment (Figure 3.9). Linear regression models accounted for considerable variation ($P < 0.005$; R^2 values between 0.64 for cold treatment and 0.99 for NOR) indicating overall linear responses to the treatments. The most effective condition to break the dormancy of these genotypes was a cold treatment, which

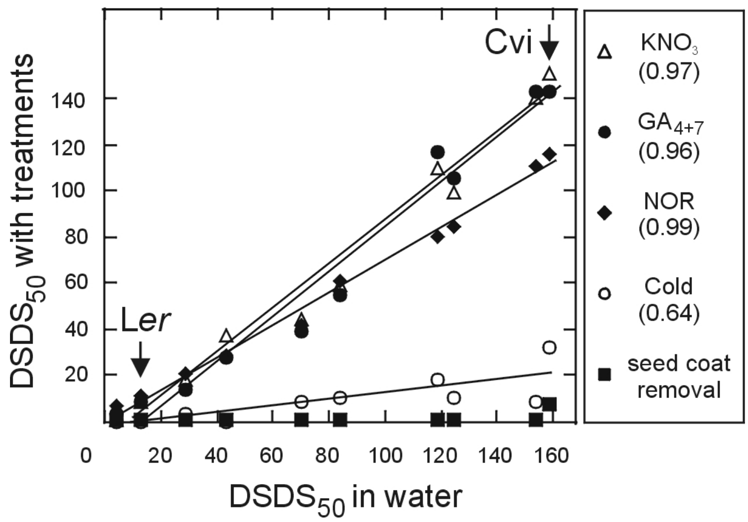


Figure 3.9. Germination responses of L*er*, Cvi and NILs to different environmental conditions. Percentages of germination were estimated at ten different times of seed dry storage in water and in five additional environmental conditions: in water after a cold treatment; with gibberellin 4 and 7 (GA_{4+7}); with norflurazon (NOR); with KNO_3; and after removal of the seed coat. For each genotype, the $DSDS_{50}$ value of the 5 treatments is plotted against the corresponding value in water, and the regression line for each environment is fitted (R^2 values of regression lines are given in the legend). Germination percentages were estimated from 4 different seed bulks of 3 plants.

reduced the $DSDS_{50}$ values of Cvi from 160 to 30 days. In addition, treatment with NOR reduced the seed dormancy of Cvi and of the high dormancy NILs (Figure 3.9) showing that the Cvi dormancy can be partly overcome by reducing ABA biosynthesis during seed imbibition. However, the fact that Cvi and NILs DOG17, DOG3 and DOG6 still retain considerable dormant after inhibition of ABA synthesis during seed imbibition, suggests that *DOG1*, *DOG3* and *DOG6* might participate in other mechanisms different or downstream to the ABA mediated seed dormancy during seed imbibition. The least effective treatments were GA_{4+7} and nitrate, which showed similar effect, both reducing rather little the dormancy of Cvi seeds ($DSDS_{50}$ values of 142 and 150.3 respectively). However, GA_{4+7} showed a similar larger effect than NOR on the seed germination of L*er* and of the low dormancy NILs. This different response of L*er* and Cvi seeds to GA_{4+7} suggest that the increased dormancy determined by Cvi alleles involves a reduction of sensitivity to GA leading to an increased effect of ABA during imbibition. However, none of the NILs showed an obvious differential response to any of the treatments, measured as deviation from the regression lines. Therefore, a clear distinct role on a specific response with respect to the parameters analysed could not be assigned to any of the 4 loci represented in these NILs.

Embryo dormancy was also analysed by testing the germination of the embryos after removal of the seed coats (Figure 3.9). L*er* parental embryos germinated 100% the first day after seed harvest. However, Cvi embryos germinate around 50% when testas are removed from seeds at day 8 after harvest, indicating that part of the Cvi dormancy, is due to the absence of the growth potential in the embryo and can be described as pure embryo dormancy. In addition, dormant NILs also showed certain embryo dormancy, although no NIL presented embryo germination as low as Cvi indicating that the Cvi embryo dormancy cannot be assigned to a single particular locus but probably requires the effects of Cvi alleles at several loci.

Discussion

Arabidopsis accessions collected from wild populations at different geographical locations differ largely in their seed dormancy (Ratcliffe, 1976; Lawrence, 1976). For instance, the laboratory strains such as Landsberg *erecta* (L*er*) and Columbia behave as almost non-dormant when germination is measured in water with white light (van der Schaar et al., 1997) while other accessions such as Cvi originally coming from Cape Verde Islands, or Enkheim 2, show much stronger seed dormancy under these same conditions (Koornneef et al. 2000). To understand the genetic basis of this intra-specific natural variation we have analysed the after-ripening requirement variation in a cross between the two accessions showing dormancy phenotypic extremes, L*er* and Cvi. L*er* seeds need between 12 and 17 days of seed dry storage for 50% germination, depending on the maternal environment, whereas Cvi seeds require between 74 to 185 days for that. This 5 to 10 fold L*er*/Cvi difference in seed dormancy measured as after-ripening requirement in $DSDS_{50}$ values is determined mainly by 7 QTL, *DOG*1 to *DOG*7. Four of these loci, *DOG1*, *DOG2*, *DOG3* and *DOG6* showed overall large phenotypic additive effects varying between 12 to 25 days of the $DSDS_{50}$ values as estimated in the mapping experiments using a RIL population. The strong additive effect of these four loci was further confirmed in NILs with a L*er* genetic background, but in addition, genetic interactions between these loci are found to participate in this variation. Genetic interactions are detected in the analysis of the RILs or by comparing the additive effects of particular loci in RILs and NILs (such as *DOG2*), which might be determined by the limited quantitative scales of germination percentages or of the $DSDS_{50}$. However, another interesting different interaction has been found between the strongest effect locus *DOG1* and *DOG3*. The strong dormancy of *DOG1*-Cvi alleles appears conditional upon the *DOG3*-L*er* alleles. Furthermore, we cannot discard that higher order and more complex interactions are involved as suggested by the difference between the additive effect of *DOG6* estimated in the RIL population and the NILs. Thus, the overall effect of the Cvi alleles at 5 of the *DOG* loci increased seed dormancy as compared with the L*er* allele; Cvi alleles at *DOG2* reduce dormancy, and Cvi alleles at *DOG3* either increase or reduce dormancy depending on the allele at *DOG1*. These additive and epistatic effect at the *DOG* loci explain the extreme parental phenotypes and the transgression in both directions observed in the RIL population. In addition, the estimated effects at the 4 major effect loci predict a more dormant than Cvi transgressive phenotype when combining Cvi alleles at *DOG1* and *DOG6* and L*er* alleles at *DOG2* and *DOG3*. This is confirmed in the NIL DOG617, which carries two Cvi introgression fragments and thus, higher dormancy alleles probably at all but the two smaller effect QTL *DOG4* and *DOG5*.

The various *DOG* loci identified in the present study behave genetically different in their additive effects along the time of seed storage and in their epistatic effects, suggesting that they might be involved in different aspects of seed dormancy. The genetic and physiological characterization of Cvi and the NILs carrying Cvi alleles at particular *DOG* regions enable several speculations on the different roles of the L*er*/Cvi dormancy allelic variation. First, the strong dormancy of Cvi is shown to involve not only seed coat imposed dormancy but also certain embryo imposed dormancy which is absent in non-dormant laboratory strains such as L*er*. This embryo dormancy is found to probably require the effects of Cvi alleles at several *DOG* loci. However, seed coat imposed dormancy appears as the major dormancy component since Cvi seeds loss their embryo dormancy one month after harvest, while retaining testa dormancy 2 months later. Maternal genetic effects on the dormancy variation were not detected by comparing reciprocal crosses or the cytoplasms of the RILs, suggesting that the L*er*/Cvi genetic variation affecting the seed coat imposed dormancy is mainly determined by the embryo genotype. Thus, it is suggested that this genetic variation is probably involved in the growth potential of the embryo required to overcome the mechanical restraints of the maternal testa. Nevertheless, preliminary analyses of the dormancy of seeds derived from reciprocal crosses between L*er* and NIL DOG2 indicate that *DOG2* has maternal effects (data not shown). Since both, NIL DOG2 and L*er*, lack embryo dormancy, it is hypothesised that this locus affects the seed coat imposed dormancy through the genetic structure of the testa, the maternal tissues surrounding the seeds during their development or by a factor imported from the mother plant. Secondly, the *DOG* loci may affect the level of either embryo or seed coat imposed dormancy in various different ways, such as influencing the induction of seed dormancy during the later phases of seed maturation, or affecting the mechanisms controlling the release of dormancy during storage, or controlling mechanisms involved in the onset of germination. The behaviour of the Cvi accession and the high dormancy NILs carrying Cvi alleles at *DOG1*, *DOG3* or *DOG6* resemble the non-germinating mutants deficient in gibberellins (*ga1, ga2* and *ga3;* Koornneef and van der Veen, 1980) or defective in GA signal transduction (*sleepy1;* Steber et al. 1998). Gibberellins are required for the onset of germination to counteract the ABA imposed dormancy (Bentsink and Koornneef, 2002). However, in contrast to those GA related mutants, the germination of Cvi and these NILs can be restored by after-ripening and cold treatment indicating that a different kind of genetic variation is leading to the increased dormancy. The observation that exogenous GA application is less effective in releasing the dormancy of Cvi seeds than the NOR inhibition of seed ABA biosynthesis during seed imbibition, suggests that part of the L*er*/Cvi dormancy allelic variation at the *DOG* loci might be involved in the mechanisms downstream to GA, controlling sensitivity to GAs. Thus, it is speculated that Cvi shows an increased ABA mediated seed dormancy not simply determined by increased seed ABA synthesis (Jullien et al., 2000) or reduction of GA biosynthesis, but by reduction of GA sensitivity. Nevertheless this function could not be assigned specifically to any of the *DOG* loci. In addition, inhibition of ABA biosynthesis during seed imbibition could overcome the strong dormancy of lines carrying Cvi alleles at *DOG1*, *DOG3* and *DOG6* only partly, the seeds of these line still retaining considerable dormancy (Figure 3.9). Therefore, it is speculated that these loci affect other mechanisms different or downstream to the ABA imposed dormancy. Finally, the strong effect of a cold treatment to reduce the seed dormancy of Cvi and of the high dormancy NILs suggests that cold does not affect GA biosynthesis, as has been proposed for the light induced germination (Yamaguchi et al., 1998). Moreover, the cold temperature mechanism must inactivate not only the mechanisms mediated by the ABA synthesised during seed imbibition but also other ABA mediated or ABA independent dormancy mechanisms that are probably involved.

Molecular interpretations of the function of the *DOG* loci require further characterisation and ultimately the gene isolation. Comparison of the map positions of the *DOG* loci with known seed dormancy and germination mutants allows the identification of primary candidate genes for all QTL except *DOG1* (Bentsink and Koornneef, 2002). Therefore, *DOG1* is likely to represent a new dormancy locus accounting for an important part of the variation for seed dormancy present in nature. *DOG2* maps close to *aba3*, *phyA* and *cry2* mutants; *DOG3* does it near to *lec1* and *lec2*; *DOG4* maps between *tt7* and *tt4*; *DOG5* around *abi1*; *DOG6* adjacent to *rdo1*, *fus3* and *abi3*; and *DOG7* close to *ats, era1* and *tt3*. The dormancy phenotypic effects of the *DOG* loci are comparable or even stronger than the currently available seed dormancy and germination mutants, which facilitates their further genetic and molecular analysis. However, the seed pleiotropic effects of several of the candidate mutants such as alterations of the pigmentation or shape of the testa, or changes in embryo pigmentation do not appear in the *DOG* NILs. In addition, other seed characteristics found in some germination mutants like changes in seed sugar composition or effects on hypocotyl elongation, are probably not affected by most of the *DOG* loci as deduced from the comparison of QTL map positions for the various traits studied in the L*er*/Cvi RIL population (Bentsink et al., 2000; Borevitz et al., 2002); conversely, other traits such as seed size, seed storability or flowering time might be influenced by some of the *DOG* loci (Alonso-Blanco et al., 1998b, 1999; Bentsink et al., 2000). We have begun the fine mapping of *DOG1*, *DOG2, DOG3* and *DOG6* by analysing crosses between the corresponding NILs and L*er*. Thus, we have discarded the *CRY2* photoreceptor and *ABA3* gene as candidates for *DOG2*, and *ABI3* for *DOG6*, further suggesting that these loci might provide new genes involved in the control of seed dormancy. A previous study of the genetic variation affecting seed dormancy and germination present between L*er* and Col showed that the small differences between both accessions were attributable to 14 loci with rather small effect (van der Schaar et al., 1997). The 7 *DOG* loci identified between L*er* and Cvi all locate in genomic regions containing the L*er*/Col QTL, suggesting that allelic series at a limited number of loci might account for the natural seed dormancy variation. However, the molecular isolation of the underlying genes and the identification of the specific allelic variants is still needed to understand the molecular basis of the genetic variation found in these works. Such endeavour will provide new components and new genetic variants of known components for the subsequent physiological and molecular understanding of seed dormancy. Ultimately, the identification of these loci will initiate the comprehension of the ecological and evolutionary significance of this quantitative natural variation and of the mechanisms involved in the development of different life history strategies for adaptation to the environment.

Acknowledgements

We are grateful to Isabelle Debeaujon and Karen M. Léon-Kloosterziel for helpful discussions at the beginning of the project. This research was supported by the European Union (E.U.) program NATURAL (contract QLG2-CT-2001-01097), by the Biotechnology TDR program of the E.U. (BIO4-CT96-5008; grant to C.A.-B) and by The Earth and Life Sciences Foundation subsidised by the Netherlands Organization for Scientific Research (grant to L.B.).

Material and Methods

Plant material

The Arabidopsis accessions Landsberg *erecta* (L*er*) originate from Northern Europe (Rédei, 1992) and Cvi, from the tropical Cape Verde Islands (Lobin, 1983) and a set of 161 recombinant inbred lines (RILs) derived from crosses between them was analysed for their seed dormancy behaviour. These lines have been previously described and characterised using AFLP and CAPS markers (Alonso-Blanco et al., 1998b).

Construction of dormancy NILs

Eleven near isogenic lines (NILs) were constructed by the introgression of increasing seed dormancy Cvi alleles into a L*er* genetic background through phenotypic and genotypic selection in three backcross generations. RILs CVL-49, CVL-122, CVL-128 and CVL-160 were used as starting material, selected on the basis of their phenotype and genotype as lines with strong seed dormancy and carrying different combinations of alleles at the six QTL genomic regions where Cvi alleles that increase dormancy were mapped. Introgression lines were derived from each RIL after 2 backcross generations and 2 further selfing generations as follows. RILs were backcrossed to L*er* and small populations of 100 to 120 F2 plants were obtained and their F3 seeds tested for germination. The two plants with highest seed dormancy from each population were backcrossed once more to L*er* and F3 seeds from 100-120 F2 plants were again assayed for dormancy. Forty-four plants with different degrees of seed dormancy (three to seven from each of the eight populations) were selected, selfed and genome wide genotyped with 182 AFLP and CAPS markers chosen from the L*er*/Cvi genetic map (Alonso-Blanco et al., 1998b). From those plants, ten introgression lines were selected as pre-NIL and used to develop, by marker assisted selection in a further backcross generation, the 11 final NILs carrying different combinations of high dormancy Cvi alleles at only one or two *DOG* QTL regions. These lines contain Cvi introgression fragments of 10 to 50 centimorgans (cM) and were named as NIL DOG followed by the number(s) of the QTL for which the Cvi allele was expected to be introgressed. When several NILs with overlapping fragments expected to carry the same Cvi dormancy alleles were obtained, these lines were named with an additional code. Thus, 8 NILs carrying a single introgression fragment were constructed and named as NIL DOG3, DOG4, DOG5, DOG6, DOG17-1, DOG17-2, DOG17-3 and DOG17-4; three other lines carrying two introgression fragments were developed and named as NIL DOG317, DOG417 and DOG617.

A NIL carrying Cvi alleles in a single genomic region of about 15-20 cM where the low dormancy *DOG2* Cvi allele has been mapped in the present work, was previously developed (NIL 45 described in Swarup et al., 1999). Following the nomenclature described above, this line is here referred to as NIL DOG2.

Growth conditions

All plants were grown in an air-conditioned greenhouse (temperature 22-25 ºC) supplemented with additional light to provide a day-length of 14 hours. Genotypes to be compared were grown together in single experiments and their ripe seeds were harvested the same day. To largely reduce the environmental effects on seed dormancy due to local greenhouse environmental differences affecting the mother plants, seeds from each genotype were harvested as a single or multiple bulks of 4 to 12 plants (as specified below). Seeds were harvested in cellophane bags and stored together in a cardboard box at room temperature.

F1 hybrid seeds from reciprocal crosses between the parental lines were obtained by emasculation of flowers and hand pollination. Three to sixteen crosses of each class were performed using different individuals plants and the seeds were harvested in 3 to 4 bulks of 2 to 4 crosses each.

RIL evaluation

The complete set of RILs, the parental lines and reciprocal F1 hybrids were grown in a single experiment. Twelve plants per RIL and 24 of the parental lines and their hybrids were grown in a two-block design. Blocks were divided in rows of 12 plants and six plants of each RIL were grown per block in half a row, lines being completely randomised. To reduce developmental and environmental effects on seed dormancy, the onset of flowering was synchronised, since the RIL population shows large variation for flowering initiation (Alonso-Blanco et al., 1998b). For that, RILs were planted at three consecutive weeks according to their flowering times. The seeds of all genotypes were harvested on the same day in a single seed bulk per RIL and 4 seed bulks from 6 plants for the parental lines and F1 hybrids.

NIL evaluation

All the NILs carrying Cvi alleles at the dormancy QTL regions and the parental lines were grown together in a similar design as that described for the RIL evaluation, but consisting of 4 blocks with 6 plants. The seeds of each genotype were harvested in four seed bulks corresponding to the different blocks, of 3 plants each.

Seed dormancy measurements and germination assays

The percentage of germinating seeds of a genotype at a particular time of seed storage was taken as a measurement of the degree of dormancy at that particular time. In each experiment, germination was tested at least at 6 different time points of dry storage from the harvest date until 100% of the seeds germinated. Curves of germination percentage on the time of storage provided the kinetics of seed dormancy of a genotype.

In addition, the seed dormancy of a genotype was estimated in a single parameter as the number of days of seed dry storage ('after-ripening') required to reach 50% germination ($DSDS_{50}$). To estimate the $DSDS_{50}$ value of each genotype, all the measurements of germination proportions at the various times during seed storage were used for probit regression on a logarithm time scale applying the regression module of the statistical package SPSS version 10.0.6.

Germination tests in water under white light were performed at each time point by incubating seeds during one week as follows. Between 50 and 100 seeds of a genotype were evenly sown on a filter paper soaked with 0.7 ml demineralised water in a 6 cm Petri dish. Petri dishes were placed in moisture chambers consisting of plastic trays containing a filter paper saturated with tap water and closed with transparent lids. Moisture chambers were stored during one week in a climate chamber at 22-25 ºC illuminated with 38 W Philips TL84 fluorescent tubes at 8 W m^{-2} with a light period of 16 hours followed by 8 hours of darkness. After that, the total number and the number of germinating seeds was scored and the percentage of germinating seeds calculated.

Germination of parental lines and NILs was also assayed in 5 different environmental conditions known to enhance germination or break seed dormancy (see Introduction). Germination was analysed after a cold treatment by placing moisture chambers in a cold room at 6°C during 7 days before to be transferred into the illuminated 25°C climate chamber. Seed germination was tested in the presence of three chemical compounds: 10 µM of gibberellins 4 and 7 (GA_{4+7}; Duchefa, The Netherlands; Koornneef and van der Veen, 1980), 10 µM of norflurazon (NOR; Chem Service, Inc.) which is an inhibitor of abscisic acid biosynthesis (Chamovitz et al., 1991) or 10 mM KNO_3 (Derkx and Karssen, 1993b) by soaking the filter paper in the corresponding solution. Concentrations for NOR, GA_{4+7} and nitrate were selected from preliminary concentration response analyses as the lowest concentration with maximum effect on the seed germination of L*er* and Cvi parental lines. GA_{4+7} were dissolved in a few drops of 1M KOH and then diluted to 10 µM with phosphate citrate buffer pH 5 containing 3.3 mM K_2HPO_4 $3H_2O$ and 1.7 mM citric acid. NOR was dissolved in a few drops of acetone and then diluted with water to 1 µM final concentration. Germination was also assayed after removal of the seed coat under a stereomicroscope by scratching carefully the seeds with two needles.

For every genotype and condition, three to four germination tests at each storage time point were performed using a single or different seed bulks. The average germination percentage at each time point of seed storage was calculated as well as the standard error to obtain an estimate of the measurement error. Since in the present study we used seed bulks from various plants, variation among plants within a genotype due to greenhouse environmental effects on the mother plants is negligible, and variation among genotype means are interpreted as the genetic variation component of the total phenotypic variation.

QTL analyses

For each RIL, the mean germination percentage at 1, 3, 6 10 15 and 21 weeks of seed storage was calculated from three replicates of the germination tests performed with a seed bulk of 12 plants. These percentages were used to estimate the $DSDS_{50}$ value of the RILs. $DSDS_{50}$ values were transformed (log_{10}) to improve the normality of the distribution and transformed data were used to perform QTL analysis. The mean germination percentages of the RILs at each time point of seed storage were transformed by the angular transformation (= arcsin√) and these data sets were used separately for QTL analyses at the six different time points of seed storage. A set of 99 markers covering most of the Arabidopsis genetic map at average intervals of 5 cM was selected from the L*er*/Cvi RIL map (Alonso-Blanco et al., 1998b). The computer software MapQTL version 4.0 (van Ooijen, 2000) was used to identify and locate QTL on the linkage map by using interval mapping and multiple-QTL-model (MQM) mapping methods as described in its reference manual (http://www.plant.wageningen-ur.nl/products/mapping/mapqtl/). In a first step, putative QTL were identified using interval mapping. Thereafter, one marker at each putative QTL (between 4 and 7 depending on the trait) was selected as a cofactor and the selected markers were used as genetic background controls in the approximate multiple QTL model of MapQTL. To refine the mapping and to identify linked QTL, cofactor markers at each QTL were moved one by one around the putative QTL position, finally selecting the closest markers to the QTL, i.e. those maximizing the logarithm-of-odds (LOD) score. LOD threshold values applied to declare the presence of a QTL were estimated by performing the permutation tests implemented in MapQTL version 4. The quantitative trait data of the RILs were permuted 1000 times over the genotypes and empirical LOD thresholds corresponding to a genome wide significance $\alpha = 0.05$ were estimated between 2.5 to 2.7 for the various data sets. Two-LOD support intervals were established as ≈ 95% QTL confidence interval (van Ooijen, 1992). The estimated additive genetic effect and the percentage of variance explained by each QTL, and the total variance explained by all the QTL affecting a trait, were obtained with MapQTL in the final multiple-QTL model in which one cofactor marker was fixed per QTL. Additive genetic effects presented correspond to the differences between the estimated means of the two homozygous RIL genotypic groups at each particular QTL. A positive additive effect implies that Cvi genotypes have higher germination (lower dormancy) than L*er*, while negative effects indicate that Cvi genotypes have the lower germination (higher dormancy). All the statistical comparisons shown were based on the transformed data, but none of the conclusions was changed when using the original data. Therefore, additive effects presented are estimated with the original scale data as merely orientative.

Since 116 of the RILs carry L*er* cytoplasm and 45 carry Cvi cytoplasm, cytoplasmic genetic effects were analysed in the RIL population using the cytoplasmic genotype as factor in one way ANOVA and in multiple factor linear models in combination with the nuclear QTL markers affecting each trait.

Two-way interactions among the QTL identified were tested by ANOVA using the corresponding two markers as random factors. In addition, two-way interactions were searched for among all pair-wise combinations of the 99 nuclear markers as well as the cytoplasmic genotype, using the computer program EPISTAT (Chase et al., 1997) with log-likelihood ratio (LLR) thresholds corresponding to a significance $P < 0.001$. Ten thousand trials were used in Monte Carlo simulations performed with EPISTAT to establish the statistical significance of the LLR values for the interactions detected (Chase et al., 1997).

The overall genotype by environment (G x E) interaction was tested for the percentage of germination by two-factor ANOVA using genotypes (RILs) and environments (time points of seed storage) as classifying factors. For each putative QTL, QTL x E interaction was tested by repeated measures ANOVA using the corresponding marker and the time of seed storage (repeated measurements of the RILs) as between and within classifying factors ($P < 0.005$). The General Linear Model module of the statistical package SPSS version 10.0.6 was used for ANOVA analyses.

Molecular markers

The introgression lines containing Cvi dormancy alleles were genotyped using AFLP marker analysis, which was performed according to Vos et al*.* (1995). Nine primer combinations chosen from the L*er*/Cvi molecular map (Alonso-Blanco et al*.*, 1998b) were used to amplify 182 polymorphic bands with known genetic location and that covered most of the genetic map at intervals of 1 to 15 cM.

CAPS and microsatellite markers previously mapped in the L*er*/Cvi RILs and/or the L*er*/Col RILs (Alonso-Blanco et al*.*, 1998b; TAIR: http://www.arabidopis.org) were used for marker assisted selection of the final NILs carrying Cvi dormancy alleles. CAPS markers were analysed according to Konieczny and Ausubel (1993) and microsatellite markers according to Bell and Ecker (1994). The following markers linked to the *DOG* QTL regions were used: DFR, MBK, nga129 and g2368 for *DOG1* and *DOG7* QTL region; PVV4, AXR1 and PhyA for *DOG2*; g2395 for locus *DOG3;* nga151 for *DOG4*; B9-1.8 for *DOG5* and TOPP5 for *DOG6*. Other CAPS and microsatellite markers were used to genotype in other genomic regions where undesired Cvi alleles had to be removed.

Chapter 4

Fine mapping of the *delay of germination 1* locus

Leónie Bentsink, Carlos Alonso-Blanco, Jemma Jowett,
Janny Peters and Maarten Koornneef

Abstract

The *Delay of Germination 1* (*DOG1*) locus was identified as the strongest effect quantitative trait locus affecting seed dormancy in the progeny of a cross between the two Arabidopsis accessions Landsberg *erecta* and Cape Verde Islands. The genetic fine mapping of *DOG1* using morphological and molecular markers allowed its assignment to a region of approximately 75 Kb. This genomic region is predicted to contain 22 open reading frames, none of them corresponding to genes previously involved in seed dormancy or germination. Therefore there are no obvious candidate genes.

Introduction

The *Delay of Germination 1* (*DOG1*) locus has been identified as the strongest effect quantitative trait locus (QTL) affecting seed dormancy segregating in the progeny of a cross between two Arabidopsis accessions (Chapter 3). QTL mapping was performed on a recombinant inbred line (RIL) population derived from crosses between the accessions Landsberg *erecta* (L*er*) and Cape Verde Islands (Cvi). For the characterization and fine mapping of *DOG1* a near isogenic line (NIL) named DOG17-1 was developed and characterised (Chapter 3). This line contains a Cvi introgression fragment around the position of the *DOG1* QTL in a L*er* genetic background. The dormant phenotype of this NIL confirmed the effect of *DOG1* in the L*er* background, where the Cvi allele strongly increased the after-ripening requirement for seed germination (Chapter 3). To understand the molecular function and the role of *DOG1* in seed dormancy/germination, the identification of the underlying gene is pursued and the current status of this research is presented in this chapter.

To clone a gene there are several strategies depending on the molecular information available (Gibson and Somerville, 1993). Some genes, like for instance those involved in metabolism, can be isolated by their ability to complement mutations in bacteria or yeast (Minet et al., 1992). Other genes like those with a well-characterised pattern of specific expression can be isolated using transcriptional differential display techniques (Park et al., 1998). However, in case that artificially induced mutants are available for a gene and only the mutant phenotype is known, other strategies have to be used. Mutants generated by a T-DNA or transposon insertion, can be very efficiently cloned by recovering DNA fragments flanking the insertion (reviewed by Maes et al., 1999; Parinov and Sundaresan, 2000). If the mutant phenotype is caused by a deletion like is often the case when physical mutagens are used to generate the mutants, the gene can be cloned by subtractive hybridisation. However, cloning by subtractive hybridisation is complicated and has only been proven successful for two loci (Sun et al., 1992; Silverstone et al., 1998). When the above mentioned methods are not applicable, a gene can be cloned using a map-based approach (reviewed by Lukowitz et al., 2000).

The first step in map-based cloning is to locate genetically the locus of interest as accurate as possible with the help of linked markers, either morphologically or molecular. Morphological markers are based on differences in phenotype while molecular markers detect polymorphisms at the DNA level. Morphological markers flanking the locus of interest can be used to select in a simple way a large number of recombinants in the region of interest. This will limit the number of plants that have to be analysed with molecular markers, which are more expensive in terms of time and money. Closely linked molecular markers are identified sequentially and the closest ones flanking the locus of interest can be used to screen DNA libraries. Currently, large insert libraries have been developed mainly in bacterial artificial chromosomes (BACs), and clones that contain the region of the genome covering the locus are thus easily isolated. Thereafter, the gene can be identified by phenotypic complementation in transgenic plants; that is by transforming overlapping independent DNA fragments of an active allele into a genotype with a recessive (or less active) allele to determine which sequences are able to restore or complement the mutant phenotype. Ultimately, DNA sequencing of the different mutant and wild type alleles will reveal the nature of the various alleles. Further DNA sequence comparisons may indicate the putative molecular function of the gene or reveals a gene with unknown cellular function.

Map-based cloning in Arabidopsis is facilitated because it has one of the smallest genomes among higher plants and it contains very little repetitive DNA. Furthermore, there are many genetic loci and markers identified by mutations, it has a dense molecular marker map and the complete Columbia (Col) sequence is known (Arabidopsis Genome Initiative, 2000). In addition, the partial sequence of the Landsberg *erecta* genome generated by Cereon Genomics (http://www.arabidopsis.org/Cereon/index.html) is also available. This greatly facilitates the identification of DNA polymorphisms and the development of PCR based molecular markers. Currently, the most common PCR markers used for fine mapping are: Cleaved Amplified Polymorphic Sequence (CAPS; Konieczny and Ausubel, 1993), which detect a DNA polymorphism in an amplified DNA fragment on the base of the digestion with a restriction enzyme; Single Sequence Length Polymorphism (SSLP; Bell and Ecker, 1994) which detect polymorphisms based on the number of repetitions of short sequences; and insertion/deletion (INDEL) markers which are length polymorphic markers based on small insertions/deletions between accessions. In contrast to these co-dominant PCR based markers, Amplified Fragment Length Polymorphism (AFLP) markers are dominant markers that do not require previous knowledge of sequence, and have been used to genotype Arabidopsis mapping populations (Alonso-Blanco et al., 1998b). In Arabidopsis it is possible to integrate AFLP markers on the physical and sequence maps by combining in silico AFLP analysis of the available Arabidopsis genomic sequence with gel-based AFLP analysis on Arabidopsis accessions (Peters et al., 2001).

Genes underlying natural genetic variants such as those of QTL can be isolated by map-based cloning techniques, similar to artificially induced mutants for which no other information but the phenotypic effect of the genetic variants is available (Alonso-Blanco and Koornneef, 2000). However, map-based cloning has been considered problematic for isolating QTL directly from the mapping populations due to the confounding phenotypic effects of the environment and other segregating QTL. This problem can be overcome by the construction of NILs. NILs in a reference genetic background enable the phenotypic and genetic characterization of a QTL in a similar way to that performed with mutants (Alonso-Blanco and Koornneef, 2000). Monogenic crosses allow the cloning of QTL using map-based methods, which so far has been possible only in plant QTL studies (reviewed by Remington et al., 2001).

In this chapter, we used map-based techniques to fine map *DOG1*, which allowed assigning this locus to a small region of the Arabidopsis physical and sequencing maps. For that, AFLP markers previously used to map the introgression of NIL DOG17-1 were integrated with the physical map, thus defining the physical size and position of the Cvi introgression. New PCR based markers have been further developed and these have been used to genotype recombinants isolated from two new *DOG1* segregating populations specifically developed for this purpose.

Results and Discussion

Physical size and position of the Cvi introgression containing DOG1

Using the AFLP markers described by Alonso-Blanco et al. (1998b) the size of the Cvi introgression of NIL DOG17-1 has been estimated to be approximately 14 cM (Chapter 3). The upper breakpoint of this Cvi introgression was located between markers GB235C-Col and EC96L while the lower breakpoint mapped between markers BH81L-Col and GB223C. To integrate this genetic position with the physical and sequence map, we have identified the Col sequence of those AFLP markers for which either the L*er* allele or the Cvi allele amplified the same fragment as Col (Alonso-Blanco et al., 1998b). We used the in silico analysis described by Peters et al. (2001) to locate the following AFLP markers, which mapped around and within the introgression: HH225C-Col, GB235C-Col, CC277L-Col, AD75C-Col, GD350L-Col, BH81L-Col, CC540C-Col and HH445L-Col. These markers could be assigned to specific BAC clones of the Arabidopsis physical map (Table 4.1) revealing that the Cvi introgression of NIL DOG17-1 was located between BAC MWP19 and K6M13.

Table 4.1. Placement of L*er*/Cvi AFLP markers and SSLP marker nga 129, on the Col physical map. The NIL DOG17-1 allele of the AFLP markers is indicated as well.

AFLP marker	Col BAC	NIL DOG 17-1
HH.225C-Col	F15I15	L*er*
GB.235C-Col	MWP19	L*er*
CC.277L-Col	MJC20	Cvi
AD.75C-Col	MZA15	Cvi
GD.350L-Col	MQL5	Cvi
BH.81L-Col	MDN11	Cvi
CC.540C-Col	K23F3	ND*
nga129	K6M13	L*er*
HH.445L-Col	F17P19	L*er*

*ND; not determined

Mapping strategy and segregating populations

DOG1 maps very close to *DOG7*, a putative dormancy QTL of small effect identified in the same cross. This QTL might be included in the Cvi introgression fragment of NIL DOG17-1 (Chapter 3). Due to the complexity of this genomic region combined with the dominance of the high dormancy effect we decided to use a backcross population to fine map *DOG1*. This might facilitate the distinction of effects from the two putative segregating QTL. Germination analyses of F3 seeds from individual F2 lines derived from the backcrossed plants, in combination with the marker analyses described in the next section, did not detect any other locus significantly affecting seed dormancy that could interfere with the fine mapping of *DOG1*. From this we concluded that the segregation for seed dormancy in the populations derived from NIL DOG17-1 and L*er* were monogenic.

To be able to use morphological and molecular markers for fine mapping *DOG1* we derived segregating populations from crosses between a L*er* double mutant line and the NIL DOG17-1 harbouring the dormant *DOG1* allele from Cvi together with a Cvi surrounding region. The latter provides the DNA polymorphisms required for the mapping with molecular markers. The L*er* morphological marker line was homozygous for *glabrous 3* (*gl3*) and *thiamine* (*tz*). The F1 of this cross was backcrossed with this marker line and an offspring of 910 plants (910 gametes) of this cross was screened for recombinants between *gl3* and *tz,* revealing 120 recombinants. In addition a F2 population derived from the cross of NIL DOG17-1 x *gl3 tz* of 408 plants (effectively 408 gametes because only half of all recombinants between *gl3* and *tz*, can be detected; which are plants with the phenotype *GL3./tztz* and *gl3gl3*/*TZ*.) was generated and screened. This provided another 69 recombinants between *gl3* and *tz*. In total we selected 189 recombinants between *gl3* and *tz* from a population of 1318 gametes resulting in a recombination percentage of 14.3 ± 1.0 which is significantly lower than the values 20.5 ± 3.6 and 20.0 ± 5.2 published for these markers in L*er* background crosses (Koornneef et al., 1983). This might be explained by a local suppression of recombination due to possible chromosomal differences in Cvi DNA compared to L*er* in this region.

Fine mapping of DOG1 with molecular markers

The genetic fine mapping of *DOG1* was performed with molecular markers. For that, 17 new CAPS, INDEL and SSLP markers developed based on the Col sequence and 2 other CAPS markers previously described were used (Table 4.2). DNA isolated from 189 recombinants between *gl3* and *tz* was analysed with these polymorphic markers.

Table 4.2. Molecular markers polymorphic between L*er* and Cvi. Order of the markers is based on the position in the Col sequence

Marker	Primer sequence		Marker	Annealing	Restriction	Position in genome	
	Forward	Reverse	type	T (°C)	enzyme	BAC	Position in BAC
MPO12	gactaatcatcaccgactcag	cttgtcttcttccactgtcc	SSLP	50	-	MPO12	19316-19519
cMJB21b	cggtgttggatgttcataagc	aacattgttgactggtcactgg	CAPS	54	MboI/HindIII	MJB21	35118-36661
DFR[a]	tgttacatggcttcatacca	agatcctgaggtgagtttttc	CAPS	51	BsaAI	MJB21	71655-72786
cSRP40	ccgttaatccgatgttcagg	tcgttaatcctcattgcatgtg	CAPS	50	HindIII	MBD2	27883-29130
cHALFd	agaaccgagaatccgcagcgag	gcctagcgggtagcttggaaac	INDEL	54	-	MBD2	33734-35026
cMLN1	agcagatagctgccacgtcac	aggacgacatggacgacacg	CAPS	62	BsaAI	MLN1	14443-15744
K23L20	ggaacatgaaacagattaca	aggtgaacacacttcttacaa	SSLP	50	-	K23L20	50571-50661
cK21C13b	cgtttcaacgcatttggatcac	tccgtagatgagcgagcgac	CAPS	51	DraI	K21C13	38063-39144
cK21C13w	accattccaatggatgagcagg	aaaccgggatcatccatggtg	CAPS	51	AluI/DraI	K21C13	35461-37156
RPS4-ct[a]	tctctgtatccctctctcag	tcaacaggcataacgtactt	CAPS	50	Eco01091	K9E15	3562-4248
MFC19a	tcggctaattcacctcagcg	atcggtggttgttgtggctg	INDEL	52	-	MFC19	19089-19336
cK2N11a	cctttgacaatgaagaagcctg	aaccgtggctaaatacgatgtg	INDEL	52	-	K2N11	23496-24594
cMRA19z	gagaatcaatggcgggttacag	cttcgagagccgaggtttcc	CAPS	52	AluI	MRA19	63356-64513
cK15I22z	gacgagatcatgttgctgccac	gctcgtttatgctttgtgtggg	CAPS	52	HinfI	K15I22	6956-8416
K15I22	tcggtggtttactttcacttt	gaattgtagcttcttctgaacc	SSLP	52	-	K15I22	50828-50969
cMLC19b	cgaaatcgggacttacctgac	tgcaggccaaacatggtaac	CAPS	54	MboI	MCL19	51253-51535
cMPL12z	ttgctctctcatgtcctaatcg	cccactttcgtatatcagttgc	CAPS	52	HpaII/MspI	MPL12	13559-15247
MPL12a	ggacaaagagggcgttgataag	cctttcgaactctaccaccgag	INDEL	52	-	MPL12	45284-45973
K14A13	gtggaggaatactggagccgc	gcatgaattccgcctgcttct	INDEL	52	-	K14A4	31786-31887

[a] Obtained from The Arabidopsis Information Resource (www.arabidopsis.org)

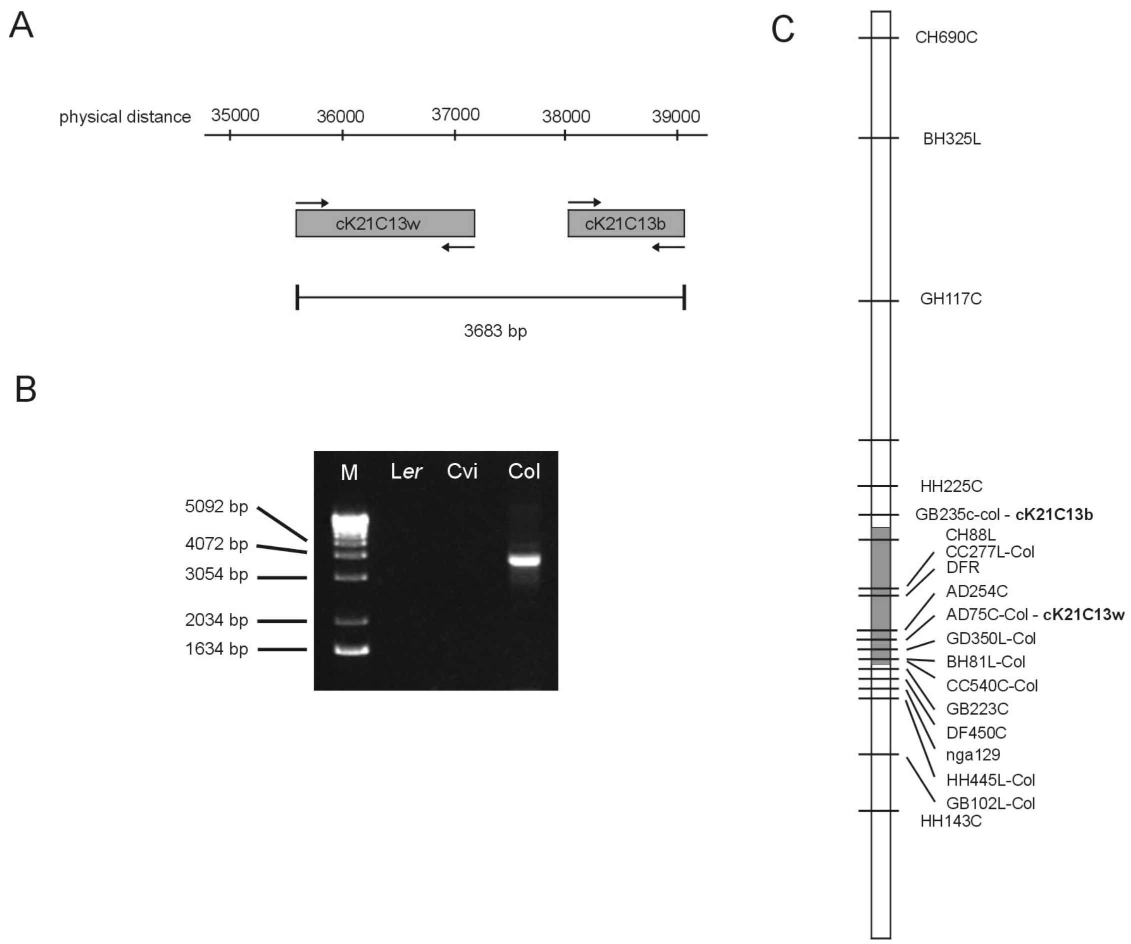


Figure 4.1. Analysis of an intrachromosomal rearrangement in the L*er*/Cvi mapping populations (A) Position of the markers cK21C13w and cK21C13b on the physical map. Arrows indicate the position of the primers; the two most outside primers were used to amplify the 3683 bp fragment. (B) PCR amplification of the 3683 bp fragment on Col genomic sequence. M: the DNA ladder marker (C) Position of the markers cK21C13w and cK21C13b on the L*er*/Cvi genetic RIL map are indicated in bold. The grey solid bar indicates the Cvi introgression of NIL DOG17-1, open bars indicate L*er* sequence in this NIL.

The phenotype of the recombinants at *DOG1* was determined by analysing the germination of: i) F3 seeds from 10-12 individual F2 plants derived from each recombinant plants selected in the backcross population; ii) or F4 seeds from 10-12 individual F3 plants derived from each recombinant selected in the F2 population.

The finding of a previously unknown chromosomal rearrangement in the *DOG1* region complicated the fine mapping of *DOG1*. During the mapping process two markers, cK21C13w and cK21C13b, located on BAC K21C13 at a physical distance smaller than 3683 bp (see Table 4.2 for marker details) according to the Col physical map (Figure 4.1A) appeared genetically distant. More than 12 recombinants were found between cK21C13w and *DOG1* (*DOG1* located south of this marker), whereas

NILD17-1 had a L*er* allele for cK21C13b, suggesting that the Cvi introgression had terminated. The two markers each amplified a PCR product with the expected Col length in L*er,* Cvi and Col. The physical distance between the two markers in Col was confirmed by PCR amplification of a 3.7 Kb fragment using the two outside primers. However, these two oligo's could not amplify DNA in L*er* and Cvi (Figure 4.1B), indicating that in these accessions the markers are further apart or present in a different orientation, making PCR amplification impossible. These two markers were genetically mapped in the original L*er*/Cvi RIL population (Alonso-Blanco et al., 1998b) and it was found that marker cK21C13b is located outside the Cvi introgression of NIL DOG17-1 (Figure 4.1C), while cK21C13w maps approximately 18 cM lower, within the introgression region of NIL DOG17-1. These results indicate that an intra-chromosomal rearrangement is present in Col genomic DNA, as compared to L*er* and Cvi DNA, in such a way that cK21C13b has been transposed to a position close to and south of K21C13w. Marker RPS4-ct, which is located physically 183 Kb away from cK21C13w in the Col sequence, seems present at the homologous position in all 3 accessions, indicating that not more than 183 Kb are likely to be translocated. Chromosomal structural changes such as the intra-chromosomal translocation found in this region might cause significant problems to fine map a locus due to suppression of recombination and wrong interpretations on the position of the markers based on the physical positions of these markers of one single sequence. In this case, the larger similarity between L*er* and Cvi in this region may determine an easier fine mapping of the *DOG1* in L*er* x Cvi crosses, while this might have been more difficult in crosses involving the Col accession. Structural regional differences between Col and other accessions have been also detected in other mapping experiments (Kowalski et al., 1994; Fransz et al., 1998; Hamburger et al., 2002) indicating that micro-linearity between the genomes of different populations of Arabidopsis is not always perfectly maintained.

The analysis of the in total 19 molecular markers in the set of 189 recombinants located *DOG1* between the markers cMRA19z and K15I22, which both had one recombinant between *DOG1* and the marker (Figure 4.2). This region is almost 75 Kb in size and contains 22 predicted open reading frames, whose encoded proteins are presented in Table 4.3. In this region we could not find any gene previously involved in seed dormancy or germination and therefore there is no obvious candidate gene. Further fine mapping of *DOG1* is underway with the development of additional markers that can be used to refine the position of the recombination events of the 2 recombinants found between markers cMRA19z and K15I22. In addition, the screening of three recombinants between cK2N11a and K15I22 that were still not analysed for the markers cMRA19z and cK15I22z will also be performed. Once fine mapping has reduced the location of *DOG1* to a smaller than 50 kb region, the final identification of the gene will eventually be pursued by plant transformation (Lukowitz et al., 2000).

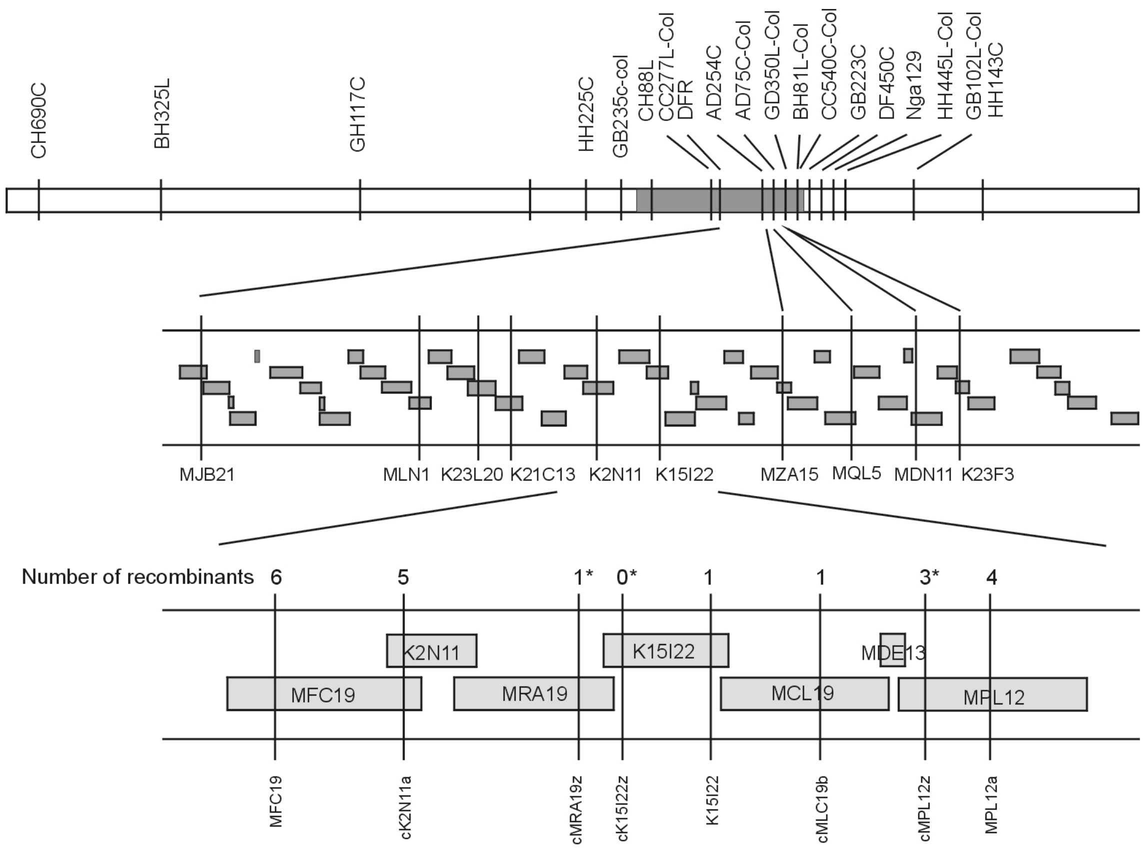


Figure 4.2. Position of *DOG1* on chromosome 5. The upper part of the figure shows the L*er*/Cvi RIL AFLP genetic map. The grey solid bar indicates the Cvi introgression fragment in NIL DOG 17-1, while open bars indicate the L*er* sequence of this NIL. The middle part of the figure depicts the physical map BAC contig of the *DOG1* region. Positions of the AFLP markers on the Col BACs are indicated as lines connecting the AFLP markers with the respective BAC. At the bottom of the figure, a close up of the BACs located in the *DOG1* region is shown. The number of recombinants between *DOG1* and the molecular markers is indicated on each marker.
* The number of recombinants is based on 6 of the in total 9 recombinants between cK2N11a and MPL12a.

Table 4.3. Open reading frames (ORFs) and encoding proteins located in the *DOG1* region (± 75 Kb). Molecular markers and their number of recombinants with *DOG1* are indicated.

ORF No.	Gene ID	Protein	Marker	No. recombinants
1	MRA19.18	Rab-type small GTP-binding protein-like	cMRA19z	1*
2	MRA19.19	unknown protein		
3	MRA19.20	unknown protein		
4	MRA19.21	ribosomal protein L11-like		
5	MRA19.22	receptor-like protein kinase		
6	MRA19.23	tyrosine-specific protein phosphatase-like		
7	MRA19.24	receptor-like protein kinase		
8	K15I22.1	serine/threonine protein kinase		
9	K15I22.2	serine/threonine protein kinase		
10	K15I22.3	tumor-related protein-like	cK15I22z	0*
11	K15I22.4	receptor protein kinase-like protein		
12	K15I22.5	unknown protein		
13	K15I22.6	unknown protein		
14	K15I22.7	unknown protein		
15	K15I22.8	Ole e I (main olive allergen)-like protein		
16	K15I22.9	senescence-specific cysteine protease		
17	K15I22.10	ubiquitin activating enzyme E1-like protein		
18	K15I22.11	GDSL-motif lipase/hydrolase-like protein		
19	K15I22.12	similarity to isoamylacetate-hydrolyzing esterase		
20	K15I22.13	magnesium chelatase subunit of protochlorophyllide reductase		
21	K15I22.14	unknown protein		
22	K15I22.15	GDSL-motif lipase/hydrolase-like protein	K15I22	1*

* The number of recombinants is based on only 6 of the in total 9 recombinants between cK2N11a and MPL12a.

Material and Methods

Plant material and growth conditions

NIL DOG17-1 has been described in Chapter 3. The mutant *gl3* has been isolated and described by Koornneef et al. (1983) and *tz* by Feenstra (1965). The L*er* marker line containing the *gl3* and *tz* mutation was previously generated in our laboratory. The mapping populations were made by crossing NIL DOG17-1 with the L*er* marker line carrying the *gl3* and *tz* mutations. The F1 hybrid of this cross was backcrossed to the marker line to construct the backcross population and selfed to retrieve the F2BC1 population. Crosses were made by hand pollination.

Plants were grown in an air-conditioned greenhouse (temp 22-25°C), supplemented with additional light (model TDL 58W/84, Philips, Eindhoven The Netherlands) from mid September till early April, providing a day length of 14 h. Plants were grown in 7-cm pots in standard soil.

Seed were harvested in cellophane bags and stored in a cardboard box at room temperature.

Seed germination/dormancy assays

Seed germination assays were performed as described in Chapters 3. Between 70 and 100 seeds were sown on water-soaked filter paper (No. 595, Schleicher & Schuell, Keene, NH) in a 6-cm Petri dish and incubated in a climate room (25°C, 16 h light/day; model TL57, Philips) during one week. Germination was scored after 7 d.

In silico AFLP analysis

In silico AFLP analysis was performed according to Peters et al. (2001).

Design of molecular markers

Based on the Columbia sequence 14 new CAPS (Konieczny and Ausubel, 1993), SSLP (Bell and Ecker, 1994) and INDEL markers were developed. Marker specifications are listed in Table 4.2.

PCR were carried out in 50 µL volumes containing, 50 ng of genomic DNA, 100 µM of each deoxynucleotide, 100 ng of both primers and 0.2 U Taq DNA polymerase. Conditions for amplification were as follows: 30 s at 94°C; annealing for 2 min at 50 to 62°C (Table 4.2); extension for 2 min at 68°C. The cycle was repeated 35 times. To detect the polymorphism 10 µL of PCR product was analysed in a 2% agarose gel (SSLP and INDEL markers) or 5 µL of PCR product was cleaved with the appropriated restriction enzyme (Table 4.2) and analysed in a 1.5% (w/v) agarose gel (CAPS markers).

Chapter 5

Genetic and physiological characterisation of *delay of germination 1*, a novel seed dormancy locus

Leónie Bentsink, Carlos Alonso-Blanco, Jemma Jowett,
Susanne Claessens and Maarten Koornneef

Abstract

The *Delay of Germination 1* (*DOG1*) locus was the strongest QTL identified by QTL mapping for seed dormancy/germination. In this work, we have characterised the natural L*er* en Cvi alleles of *DOG1* (*DOG1*-L*er* and *DOG1*-Cvi). In addition, we isolated and characterised an induced mutant allele at the *DOG1* locus, which is completely non-dormant. The *dog1* mutant germinates 100% when freshly harvested, whereas *DOG1*-L*er* and *DOG1*-Cvi require respectively 50 days and 150 days of after-ripening to reach 100% germination. *DOG1* alleles behave as semi-dominant and the dormancy of *DOG1* appears mainly controlled by the genotype of the embryo. All *DOG1* alleles require de novo GA synthesis for germination. However, the level of GA required is determined by the *DOG1* allele. Furthermore, the dormancy of *DOG1*-L*er* and *DOG1*-Cvi is dependent on ABA as the *aba1-1* and *abi3-5* mutants are completely epistatic to dormant *DOG1* alleles. We hypothesised that ABA has a dual role in the control of seed dormancy, first setting the level of dormancy, probably during seed maturation, and thereafter antagonising GA action during germination. The *DOG1* locus seems to be involved specifically in the first process.

Introduction

Seed dormancy, defined as the failure of an intact, viable seed to complete germination under favourable conditions (Bewley, 1997), is a complex trait under the control of a large number of genes. The multigenic differences between natural populations and cultivars and the large environmental effects on the expression of the trait determine the quantitative genetic nature of this trait (Koornneef et al., 2002).

Mature Arabidopsis seeds exhibit primary seed dormancy when freshly released from the mother plant. This seed dormancy can be overcome by the common germination promoting factors such as after-ripening, light, cold treatment (so-called stratification) and chemicals such as gibberellins (GAs) and nitrate (Derkx and Karssen, 1993b; reviewed by Bentsink and Koornneef, 2002; Chapter 2).

To understand seed dormancy/germination, the combination of physiology, genetics and molecular biology applied to Arabidopsis is starting to shed light on some aspects of the mechanism of dormancy and germination (Bentsink and Koornneef, 2002; Chapter 2). There are different types of seed dormancy mutants, which play their role at different moments during seed development and germination. A group of mutants that affect the induction of seed dormancy early during seed maturation are the seed maturation mutants. It has been suggested that the non-dormant phenotype of *ABA-insensitive 3* (Ooms et al, 1993; Nambara et al., 1995), *fusca 3* (Bäumlein et al., 1994) and *leafy cotyledons 1* and *2* (Meinke et al., 1994) is due to defective seed maturation and that seeds of these mutants germinate because this is the default state (Nambara et al., 2000). A group of mutants that was directly selected on the basis of reduced dormancy, are the *reduced dormancy* (*rdo 1* to *4*) mutants (Léon-Kloosterziel et al., 1996a; Peeters et al., 2002). The fact that all four mutants show some mild pleiotropic effects in adult plants indicates that the genes are not specific for dormancy/germination but affect other processes as well. Properties of the seed coat surrounding the embryo play a role in preventing outgrowth of the embryo and therefore mutations that specifically alter the seed coat affect seed germination (Debeaujon et al., 2000). The phyto-hormones abscisic acid (ABA) and GA play an important role in seed dormancy/germination, since mutants that affect GA and ABA biosynthesis or signal transduction have very obvious seed germination phenotypes. ABA deficient and ABA insensitive mutants are non-dormant, whereas GA deficient mutants are not able to germinate without applied GA (reviewed in Bentsink and Koornneef, 2002; Chapter 2). It has been suggested for Arabidopsis that GA stimulates the growth potential of the embryo allowing it to break through the surrounding envelopes (Debeaujon and Koornneef, 2000). This growth potential is restricted by ABA produced in the embryo (Karssen et al., 1983).

The differences between and within natural populations can provide material to study natural alleles of seed dormancy genes. To study natural variation, quantitative trait loci (QTL) mapping for seed dormancy was performed in a recombinant inbred line (RIL) population derived from a cross between the Arabidopsis accessions Landsberg *erecta* (L*er*) and Cape Verde Islands (Cvi). In contrast to L*er*, Cvi has a very strong dormancy, since Cvi seeds require at least 3 months of after-ripening before they are able to germinate 100%, whereas L*er* germinates fully one month after harvest. QTL mapping revealed 7 loci affecting this trait (Chapter 3), *Delay of Germination 1* (*DOG1*) being the strongest QTL identified. To characterize *DOG1,* a near isogenic line (NIL) called DOG17-1 containing a Cvi introgression fragment in a L*er* genetic background at the position of the QTL, was constructed.

In this chapter we describe the isolation and characterisation of a completely non-dormant mutant at the *DOG1* locus and the physiological and genetic interactions of the three available *DOG1* alleles in relation to different treatments and genotypes affecting seed dormancy and germination.

Results

Isolation of an induced mutant allele of DOG1

A genetic screen was carried out in a NIL DOG17-1 dormant background aiming the isolation of non-dormant mutants. M2 seeds, freshly harvested from gamma-irradiated plants were screened for germination. This screen resulted in 12 mutants whose reduced dormant phenotype could be confirmed in the progeny of the plants grown from the selected M2 seeds. Two of these mutants were allelic to the ABA-deficient mutants, *aba1* and *aba2*. In addition, the completely non-dormant mutant C119 has been identified. C119 does not show pleiotropic effects at the seedling or adult plant stage in contrast to *aba* and *rdo* mutants. F3 seeds obtained by the selfing of 1031 F2 plants derived from the cross C119 x L*er* were analysed for their germination. After 1 month of seed storage (when the two parental lines germinated 100% but NIL DOG17-1 did not start germination) no F2 progenies with a certain degree of dormancy (the dormant *DOG1*-Cvi phenotype or segregation for this phenotype) were found. Thus, the C119 mutation could not be genetically separated from the *DOG1*-Cvi allele and therefore is very likely to carry an intragenic mutation within the *DOG1* gene, from now on being considered as a novel mutant allele of this locus referred to as *dog1*. The degree of dormancy of the three *DOG1* alleles, *dog1* (C119), *DOG1*-L*er* (L*er* genotype) and *DOG1*-Cvi (NIL DOG17-1), was assessed by determining germination percentages of seed lots at different times after seed harvest until 100% germination for the most dormant line was observed. C119 is completely non-dormant when freshly harvested, L*er* seeds need approximately 50 days to reach 100% of germination, whereas NIL DOG17-1 germinates 100% only after 150 days (Figure 5.1).

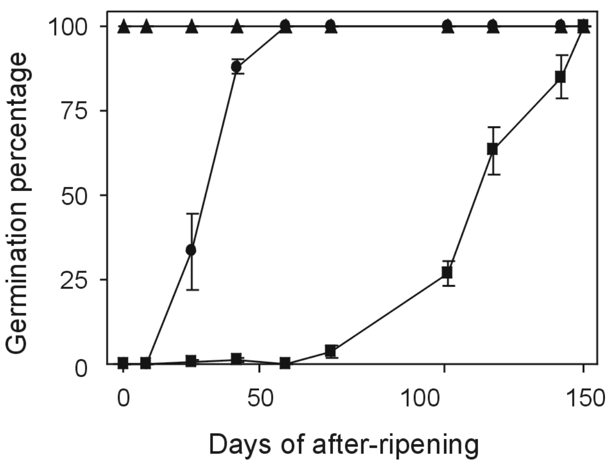


Figure 5.1. Characterisation of seed dormancy. Germination behaviour of L*er* (circles), NIL DOG17-1 (squares) and C119 (triangles) at different time points after seed harvest. The means and SE of four replicates are shown.

Genetic characterisation of DOG1 alleles

Analysis of F2 populations

To study the genetic behaviour of the *DOG1* alleles, F2 populations derived from *DOG1*-Cvi crossed with L*er* and *dog1* crossed with L*er* were analysed for their germination behaviour. F3 seeds from individual F2 plants of the cross *DOG1*-Cvi x L*er* were analysed 4 weeks after seed harvest since at this moment L*er* seeds germinated 100%. F3 seeds derived from the *dog1* x L*er* cross were tested when freshly harvested, because at that moment *dog1* germinated 100% whereas L*er* was still dormant. Both F2 populations fit monogenic segregations and suggest semi-dominance of the *DOG1* alleles (Figure 5.2), because the F3 seeds from plants that were likely to be heterozygous for the *DOG1* region exhibited a germination percentage between 25% and 75%, which is intermediate between the percentages expected when the non-dormant alleles would be respectively fully recessive or fully dominant.

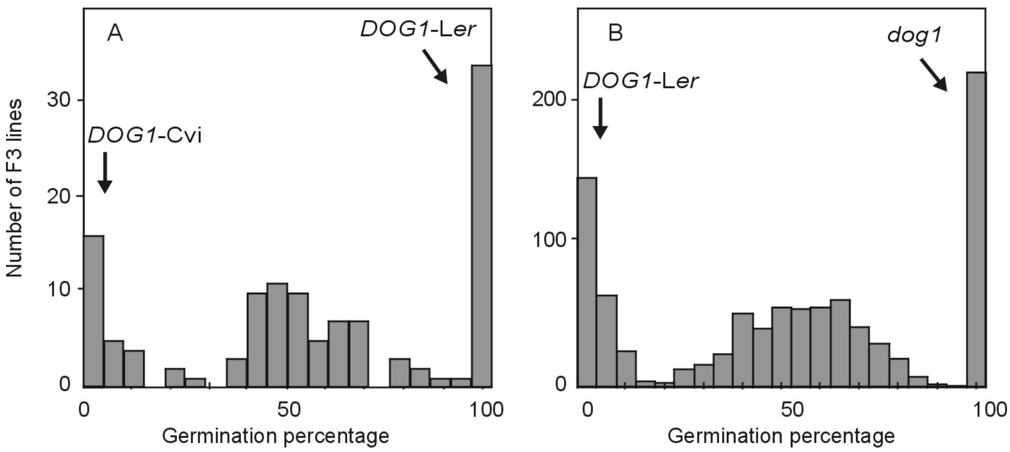


Figure 5.2. Segregation of *DOG1* alleles. Frequency distributions of germination percentages of F3 seed batches harvested on individual F2 plants derived from the crosses, *DOG1*-L*er* x *DOG1*-Cvi seeds sown 28 days after seed harvest (A) and *DOG1*-L*er* x *dog1* seeds sown directly after seed harvest (B).

Analysis of F1 reciprocal crosses

Reciprocal crosses were performed between the three *DOG1* alleles to determine the maternal and zygotic genetic components at this locus and to confirm the dominance relationships of the three *DOG1* alleles. Analysis of the after-ripening requirement measured as days of seed dry storage required to reach 50% germination ($DSDS_{50}$) of F1 seeds derived from the reciprocal crosses between *DOG1*-L*er* and *DOG1*-Cvi shows that F1 hybrid seeds derived from *DOG1*-Cvi as female parent were more dormant than the F1 seeds using *DOG1*-L*er* as mother plant (Figure 5.3). Thus, certain maternal effects of the *DOG1* locus were detected although this was partial since the level of dormancy of the F1 seeds is not identical to the maternal parent. F1 seeds derived from the crosses between *DOG1*-L*er* and *dog1* and *DOG1*-Cvi and *dog1* did not show maternal inheritance. Furthermore, the F1 seeds show intermediate dormancy phenotypes between the parental lines, confirming the semi-dominant character of the three *DOG1* alleles observed in the F2 segregation analysis.

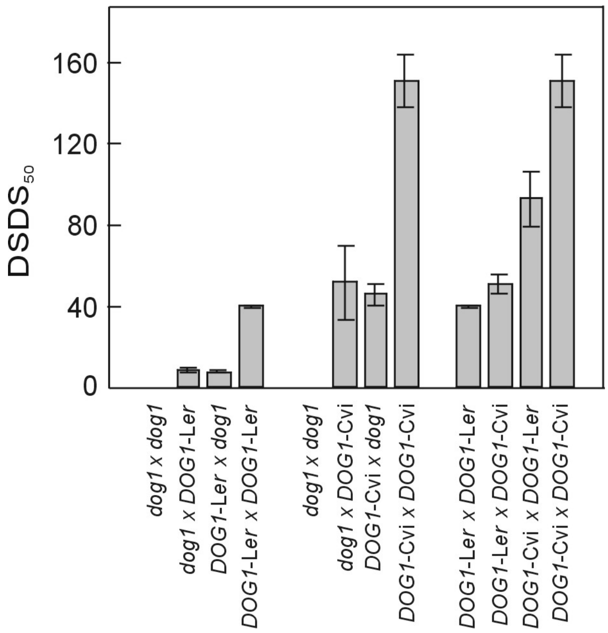


Figure 5.3. Germination potential of reciprocal crosses. Germination of F1 hybrid seeds obtained from reciprocal crosses between *DOG1*-L*er* and *DOG1*-Cvi, *dog1* and *DOG1*-L*er* and *dog1* and *DOG1*-Cvi*.* The average $DSDS_{50}$ ± SE of four replicates are shown.

Physiological characterisation of DOG1 alleles

Dormancy breaking/germination promoting treatments

To determine the effect of different dormancy breaking treatments on *DOG1*-L*er* and *DOG1*-Cvi we followed the germination at different time points after seed harvest in presence of the germination breaking compounds (nitrate, GA and norflurazon (ABA biosynthesis inhibitor)), after stratification and after removal of the seed coat. The days of seed dry storage required to reach 50% germination ($DSDS_{50}$), of *DOG1*-L*er* and *DOG1*-Cvi under the various conditions is presented in Table 5.1. The most efficient germination promoting treatment for *DOG1*-Cvi seeds is the removal of the seed coat, since in freshly harvested seeds this resulted in 100% germination. Stratification has the strongest dormancy breaking effect on the intact seed since a 7 days cold treatment on 11 days after-ripened seeds can induce 50% germination for *DOG1*-Cvi and complete germination of *DOG1*-L*er.* Nitrate, GA and norflurazon treatments break seed dormancy in a similar way but are less effective than cold. Furthermore, after-ripened seeds of the three genotypes (*dog1*, *DOG1*-L*er* and *DOG1*-Cvi) showed little or no germination in darkness indicating that all *DOG1* alleles require light for germination.

Table 5.1. Days of seed dry storage required to reach 50% of germination ($DSDS_{50}$) for *DOG*1-L*er* and *DOG1*-Cvi in the presence of the germination breaking compounds nitrate, GA and norflurazon after stratification and after removal of the seed coat.

	Water	Removal of the seed coat	Stratification	Nitrate	GA	Norflurazon
DOG1-L*er*	17 ± 3	0	0	10 ± 1	9 ± 0	12 ± 0
DOG1-Cvi	83 ± 6	0	11 ± 2	52 ± 7	52 ± 7	57 ± 9

Analysis of the GA requirement and sensitivity

To investigate the relationship between GA and *DOG1* we tested the germination of *dog1*, *DOG1*-L*er* and *DOG1*-Cvi in the presence of the hormone GA_{4+7} and on increasing concentrations of the GA biosynthesis inhibitor paclobutrazol. We followed the germination of the three genotypes at different time points after seed harvest, in the presence of GA_{4+7} and estimated the $DSDS_{50}$ values. GA promoted germination of *DOG1*-L*er* and *DOG1*-Cvi, the $DSDS_{50}$ becoming 9 ± 0 and 52 ± 7 days respectively (Table 5.1). In addition, after-ripened seeds (non-dormant seeds) of the three genotypes showed no germination anymore at 10 µM paclobutrazol (data not shown) indicating that they all require GA for germination. Furthermore, we constructed genotypes carrying the GA deficient *ga1-3* mutant allele in combination with the three *DOG1* alleles. The germination behaviour of the *ga1-3* mutant genotypes was analysed at different concentrations of GA_{4+7} (0 – 100µM) in comparison to the normal *GA1* background (Figure 5.4). After-ripened (non dormant) seeds of the three *ga1-3* genotypes were unable to germinate in the absence of exogenous GA, supporting that the three *DOG1* alleles require GA to germinate. However, the double mutant *ga1-3 dog1* germinates already more than 50% at 0.1µM of GA_{4+7}, whereas *ga1-3 DOG1*-L*er* and *ga1-3 DOG1*-Cvi need 1µM GA_{4+7} to reach the same germination percentage, indicating that *dog1* is more sensitive to GA than the other two *DOG1* alleles.

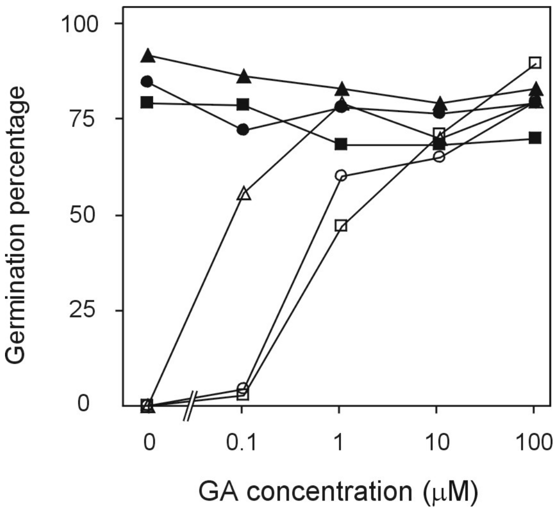


Figure 5.4. GA requirement for germination. Germination on various concentrations of GA_{4+7} of after-ripened seeds of *GA1-3 dog1* (triangles) *GA1-3 DOG1*-L*er* (circles) and *GA1-3 DOG1*-Cvi (squares), and of the GA deficient *ga1-3 dog1* (open triangles), *ga1-3 DOG1*-L*er* (open circles) and *ga1-3 DOG1*-Cvi (open squares). Seeds of *GA1-3 dog1* and *ga1-3 dog1* were 1 month after-ripened, seeds of the other combinations were after-ripened for 25 months. Percentages are averages of four measurements ± SE.

Analysis of ABA requirement and sensitivity

To study the relationship of ABA and *DOG1*, the germination behaviour of after-ripened seeds of *dog1*, *DOG1*-L*er*, and *DOG1*-Cvi was tested on increasing ABA concentrations (0-100µM), this determined the ABA sensitivity of the various genotypes. As shown in Figure 5.5 *DOG1*-Cvi required 3 times less ABA for the inhibition of germination than *DOG1*-L*er* and *dog1*. In addition, we followed the germination of these genotypes at different time points after seed harvest in the presence of the ABA biosynthesis inhibitor norflurazon. The estimated $DSDS_{50}$ of *DOG1*-L*er* and *DOG1*-Cvi were 12 ± 0 and 57 ± 9 days respectively (Table 5.1), which were comparable to the promotive effects of the exogenous application of GA. Furthermore, we constructed genotypes carrying the ABA deficient *aba1* mutant allele in combination with the *DOG1*-L*er* and *DOG1*-Cvi alleles. Their germination behaviour was analysed at different times of seed storage in comparison to the normal *ABA1* background (Figure 5.6). *DOG1*-L*er* and *DOG1*-Cvi in *aba1-1* background showed 100% germination from the time of harvest onwards thus *aba1-1* appearing as completely epistatic to *DOG1*. The same result was obtained for *DOG1*-L*er* and *DOG1*-

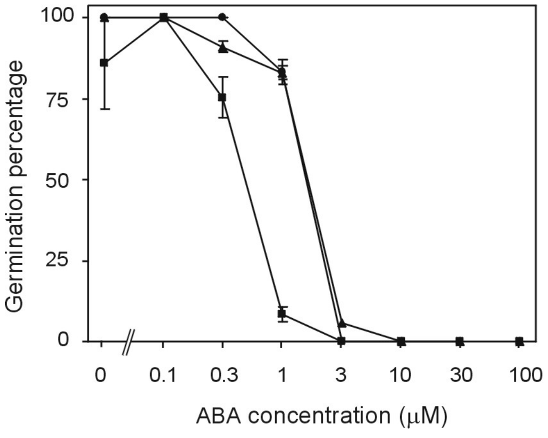


Figure 5.5. ABA sensitivity of the three *DOG1* alleles. Germination in different ABA concentrations of after-ripened seeds of *dog1* (triangles), *DOG1*-L*er* (circles) and *DOG1*-Cvi (squares). Seeds of *dog1* and *DOG1*-L*er* were after-ripened for 4 months, seeds of *DOG1*-Cvi for 25 months. The percentages are means of triplicates ± SE.

Cvi in ABA insensitive *abi3-5* background (data not shown), indicating that ABA is absolutely required for the dormancy of *DOG1*-L*er* and *DOG1*-Cvi.

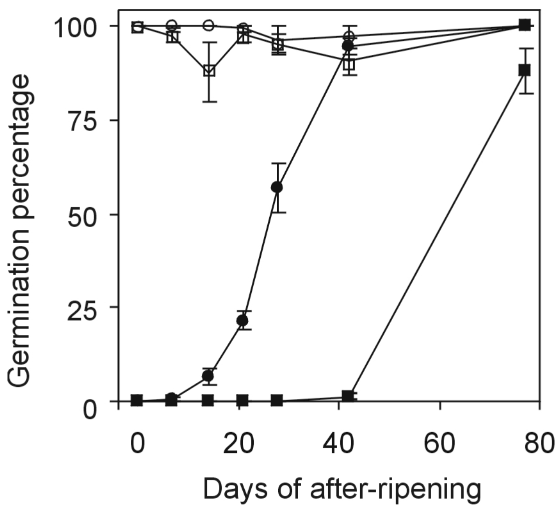


Figure 5.6. ABA dependency of *DOG1*. Germination behaviour of *aba1-1* in two different *DOG1* backgrounds (*DOG1*-L*er* and *DOG1*-Cvi) at different time points after seed harvest. *ABA1-1 DOG1*-L*er* (circles), *ABA1-1 DOG1*-Cvi (squares), aba*1-1 DOG1*-L*er* (open circles) and *aba1-1 DOG1*-Cvi (open squares).

The relationship of DOG1 with other seed dormancy genes

The genetic relationships between *DOG1* and other mutants known to be affected in seed dormancy were investigated by the analysis of the after-ripening requirement of *DOG1*-L*er* and *DOG1*-Cvi in different mutant backgrounds (Figure 5.7). *DOG1* alleles combined with the testa pigmentation mutant *tt4*, show a $DSDS_{50}$ value of *tt4* in *DOG1*-Cvi background being intermediate to that of *TT4 DOG1*-Cvi and that of *tt4 DOG1*-L*er*. This additive effect of *tt4* and *DOG1*-Cvi alleles indicates that different genetic pathways represented by both genes control dormancy. Additivity is also observed for *rdo1, rdo3* and *rdo4*. In contrast, *DOG1*-Cvi behaves almost completely epistatic to *rdo2.* However, the effect of the *rdo2* mutation is clearly visible in a *DOG1*-L*er* genetic background, indicating that *DOG1*-L*er* is required for *RDO2* to increase dormancy.

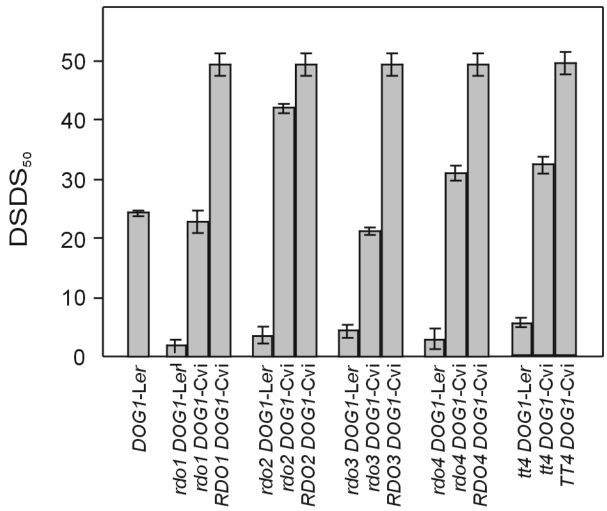


Figure 5.7. Germination of mutants known to be affected in seed dormancy in two different *DOG1* backgrounds (*DOG1*-L*er* and *DOG1*-Cvi). The average $DSDS_{50}$ ± SE of four replicates is presented.

Discussion

The genetics of seed dormancy has been studied mainly by analysing mutants that were selected for the absence of dormancy or germination. This revealed several mutants defective in hormonal pathways. The phyto-hormones ABA and GA are well known for their antagonistic functions in germination but in addition several other factors are known to play roles in germination/dormancy as shown by the genetic and physiological analysis of many mutants affected in different processes (Bentsink and Koornneef, 2002). In addition, natural variation provides another genetic resource to identify genes involved in the control of dormancy. One of these genes determining the genetic variation for seed dormancy in nature, is the *DOG1* locus identified by QTL analysis of the progeny from a cross between the accessions L*er* and Cvi (Chapter 3). In the present work we have genetically and physiologically characterised this locus. The more severe dormancy of the *DOG1*-Cvi allele compared to *DOG1*-Ler, could be confirmed in NIL DOG17-1, which contains a Cvi introgression fragment around the position of the *DOG1* locus in a furthermore L*er* genetic background. This line has been used as starting material to screen for non-dormant mutants. One of the mutants obtained was C119, which contains a mutation within the *DOG1* gene. Thus, two natural alleles and one induced mutant allele of *DOG1* are available. Since none of the three *DOG1* alleles behaves recessive, we cannot conclude unambiguously which allele represents the loss of function allele. Gamma mutagenesis usually results in loss of function alleles, which are sometimes semi-dominant as for instance the flowering-time gene *CONSTANS* (Robson et al., 2001) and several photoreceptors (Koornneef et al., 1980). Therefore, we may assume that *dog1* is the loss of function allele, suggesting that *DOG1*-L*er* and *DOG1*-Cvi alleles are the active forms promoting dormancy or inhibiting germination. In addition, the partial dominance of the various alleles suggest that *DOG1*-Cvi is a gain of function as compared to *DOG1*-L*er*, and a rate limiting mechanism on a down-stream factor on which *DOG1* acts, or on the substrate availability of the *DOG1* gene product, might be involved.

The dormancy of *DOG1* is mainly controlled by the genotype of the embryo as could be concluded by the lack of maternal effect when analysing F1 seeds derived from the reciprocal crosses between *dog1* and *DOG1*-L*er* and *dog1* and *DOG1*-Cvi. However, the partial maternal effect observed in the seeds of the reciprocal crosses between *DOG1*-L*er* and *DOG1*-Cvi suggests an additional function of the *DOG1* gene in the mother plant or in the surrounding tissues of the embryo (testa). Inhibiting factors controlled by this gene may be produced in the seed but can also be imported into the seed. Inhibiting phyto-hormones such as ABA might be candidate for this since approximately two times higher ABA levels were found in NIL DOG17-1 seeds compared to L*er* seeds (J.A.D. Zeevaart, unpublished results).

The strong dormancy of *DOG1*-Cvi can be overcome completely by removing the seed coat, indicating that the seed coat has a strong germination inhibiting effect, as has been shown before by the analysis of testa mutants (Debeaujon et al., 2000). Mechanical removal of the seed envelope can substitute for the GA requirement (Telfer et al., 1997, Debeaujon and Koornneef, 2000). However, as indicated above, the dormancy of *DOG1* is genetically not controlled by the seed coat only. After-ripening and stratification can break the dormancy of *DOG1*-Cvi. A gradual increase of GA sensitivity during dry storage and cold treatment of Arabidopsis has been reported before (Karssen and Lacka, 1986; Karssen et al., 1989). However, the fact that *ga1-1 tt4* double mutants are able to respond to cold (Debeaujon and Koornneef, 2000) without GA application, indicated that GA is not required for this response, but more likely a target of GA action. This is in agreement with the much higher response of NIL DOG17-1 (*DOG1*-Cvi) to stratification than to exogenous application of GA, indicating that other mechanisms than increase of GA biosynthesis must be targeted by the cold promotive effects on dormancy. Kraepiel et al. (1994) postulated the increase of ABA degradation as the stimulating effect on dormancy breaking or germination.

The important role of GA in promoting seed germination is well established since GA deficient mutants in Arabidopsis and tomato do not germinate (Koornneef and van der Veen, 1980; Groot and Karssen, 1987). The observation that inhibitors of GA biosynthesis, such as paclobutrazol and tetcyclacis prevent germination (Karssen et al., 1989) led to the conclusion that de novo biosynthesis of GAs is required during imbibition, unless ABA is absent. We showed that the requirement of de novo GA synthesis for germination is independent of *DOG1*, since all three alleles of *DOG1* are sensitive to paclobutrazol. However, the *DOG1* allele affects the sensitivity to GA, as concluded from the germination assay performed with *ga1-3* mutants in the different *DOG1* backgrounds (Figure 5.4). Assuming that GA is required to overcome the germination inhibition by ABA, the sensitivity to GA would be determined by the ABA sensitivity which appears also affected by the *DOG1* locus. The *dog1* mutant resembles the ABA deficient (*aba*) mutants in their complete lack of dormancy of freshly harvested seeds. This together with the observation that the dormancy of the *DOG1*-L*er* and *DOG1*-Cvi alleles completely depends on ABA might suggest that *DOG1* represents a seed germination specific downstream target of ABA. However, two substantial differences between *dog1* and *aba* mutants are observed, which suggest that *DOG1* is not a general down-stream target of ABA. First, *dog1* absolutely requires GA, this GA dependency is absent in *aba* mutants, since *ga1 aba1* double mutants in a L*er* genetic background germinate without applied GA (Koornneef et al., 1982; Debeaujon and Koornneef, 2000) and *aba* mutants are paclobutrazol resistant (Léon Kloosterziel et al.,1996b). Secondly the lack of dormancy of *aba* mutants is expressed both in light and in darkness (Koornneef et al., 1982) whereas in *dog1* it is only observed in light. The relation between GA and light requirement most likely is due to the fact that light induces GA biosynthesis during imbibition (Yamaguchi et al., 1998) and therefore the absence of dark germination may be related to this GA requirement.

In order to explain the role of *DOG1* in the dormancy control, we investigated the genetic relationships between *DOG1* and other genes that are known to affect seed dormancy/germination. A model summarising the different processes involved in the induction of seed dormancy and germination, and the possible function of *DOG1* in these processes is presented in Figure 5.8. We have shown that *DOG1*-Cvi dormancy of freshly harvested seeds can be only partly overcome with exogenous application of GA or the ABA biosynthesis inhibitor norflurazon. However, these seeds still retain a considerable degree of dormancy which is absolutely dependent on ABA since *DOG1*-Cvi in *aba1* or *abi3* backgrounds germinate 100% directly after harvest. These results suggest that there are two different dormancy/germination mechanisms dependent on ABA. We hypothesize a dual ABA role, which would explain this dependency of *DOG1* on ABA and the different behaviour of *DOG1*/GA and ABA/GA genotypes (Figure 5.8). Its first function might be setting the level of dormancy, probably during seed maturation and a second role might be inhibiting germination after imbibition which is only detected in a quiescent (non-dormant) seed. The *DOG1* locus seems not be involved in the latter process as can be concluded from the similar effects of exogenous GA and norflurazon. In addition to ABA related mutants, the only other gene that shows genetic interaction with *DOG1* is *RDO2*, *DOG1*-Cvi behaving as epistatic over *RDO2*. The increased dormancy of *RDO2* is clearly visible in *DOG1*-L*er* background, and *RDO2*, like *DOG1*, affects the GA requirement (Peeters et al., 2002). Therefore it is speculated that *DOG1* and *RDO2* may affect similar or related functions in the control of dormancy. In contrast, the *RDO1*, *RDO3*, *RDO4* genes and those affecting testa pigmentation probably affect other processes that promote dormancy or antagonise germination, since they behave additively in relation to *DOG1* alleles.
Stratification, but not GA, is effective in reducing this dormant state set by ABA, *DOG1* and probably *RDO2*. Only when this dormancy comes below a certain threshold, the seeds acquire the capacity to germinate but then, they need GA to overcome the inhibiting effect of ABA on germination of imbibed seeds. This second effect of ABA might be determined by de novo synthesised ABA during imbibition as suggested by the effect of inhibitors of ABA biosynthesis such as norflurazon (see also Grappin et al., 2000; Jullien et al. 2000), or by a limited reduction of ABA levels present in mature seeds during imbibition, as observed in dormant Cvi seeds by Jullien et al. (2000).

Thus, *DOG1* and in particular the Cvi allele, is hypothesised to be a key factor specifically involved in the determination of seed dormancy during seed development. Currently we are pursuing the molecular identification of *DOG1*, which would allow testing this hypothesis by analysing the gene function and expression during seed development and germination.

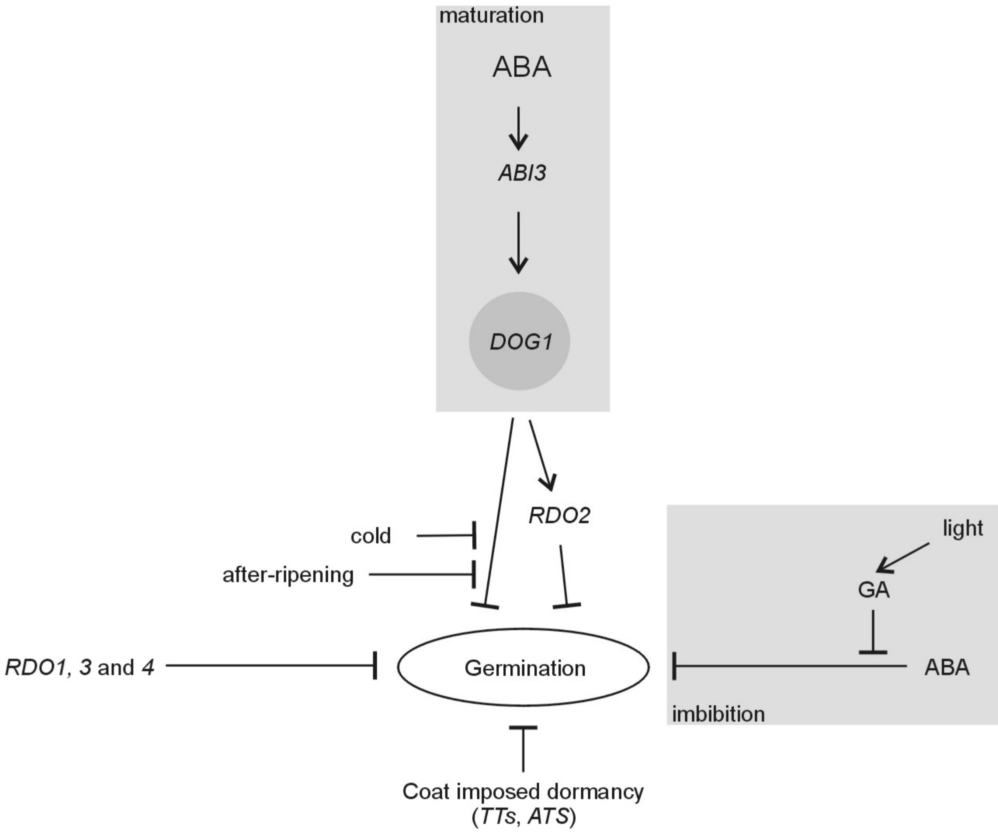


Figure 5.8. Schematic presentation of the interactions between genes that affect seed dormancy/germination. A dual function for ABA has been proposed, the first one inducing seed dormancy during the maturation phase of seed development, the second one acting antagonistic with GA during imbibition. These different phases are represented by the grey blocks. The sharp arrows and the blunt arrows stand for a promotive and an inhibitory action, respectively.

Materials and methods

Plant Material and growth conditions

Near isogenic line (NIL) DOG17-1 has a 14 cM Cvi introgression around the *DOG1* in a L*er* background, and has been previously described in Chapter 3. This line might contain Cvi alleles of another locus mapped very close in the same population and named as *DOG7* (Chapter 3). However, fine mapping using this NIL (Chapter 4) did not detect any other dormancy effects in the introgression region and it is concluded that Cvi alleles at *DOG7* are not present or they do not affect dormancy in a L*er* background. Therefore, this NIL can be considered differing in a single dormancy locus, *DOG1*, from L*er*.

The mutants *ga1-3, aba1-1*, *abi3-5* and *tt4* were isolated and described by Koornneef and van der Veen (1980), Koornneef et al. (1982), Ooms et al. (1993) and Koornneef (1981) respectively. The reduced dormancy mutants *rdo1*, *rdo2*, *rdo3* and *rdo4* were identified and described by Léon-kloosterziel et al. (1996a) and Peeters et al. (in press). All mutants are in the Landsberg *erecta* (L*er*) background.

Mutant isolation

Ten-thousand M1 seeds of NIL DOG17-1 were soaked in tubes containing 10 mM KNO_3 and 0.2% agar for 20 h at 4°C, mutagenised with 300 Gy of γ-irradiation, and dispersed on soil in an air conditioned greenhouse. M2 seeds were harvested from groups of approximately 100 M1 plants in total there were 90 bulks and a few hundred seeds of every bulk were sown on water-saturated filter paper (no. 595, Schleicher & Schuell, Keene, NH) in Petri dishes on the day of harvest. Seeds that germinated within 3 days after seed harvest were selected as putative mutants. Freshly harvested NIL DOG17-1 seeds did not germinate under these conditions. The mutant phenotype was confirmed by retesting the germination of the seeds harvested from the plants grown from these selected seedlings. To discard possible non-dormant contaminants, the selected mutants were checked for the presence of the Cvi introgression, using Cleaved Amplified Polymorphic Sequence (CAPS) marker DFR (tt3 gene; TAIR: www.arabidopsis.org).

Growth conditions

All plants were grown in an air-conditioned greenhouse supplemented with additional light (model TDL 58W/84, Philips, Eindhoven, The Netherlands) from middle September until the beginning of April, providing a day length of at least 14 h (temperature, 22–25°C).

ga1 mutants, were germinated on 10 μM GA_{4+7} and the resulted seedlings were planted in the green house, where they were sprayed once a week with 100 μM GA_{4+7} to stimulate growth, anther development and seed production.

Seed were harvested in cellophane bags and stored in a cardboard box at room temperature.

Genetic analysis

Reciprocal crosses were made by hand pollination between L*er* and NIL DOG17-1, L*er* and *C119* and NIL DOG17-1 and *C119*. Seeds of three siliques were treated as one replication. In total 3 replicates were analysed per cross. F1 hybrid plants were selfed to obtain F2 plants and the F3 seeds harvested from individual F2 plants were analysed for germination behaviour.

Construction of specific allele combinations

To combine the *DOG1*-Cvi background with *ga1-3, aba1-1*, *abi3-5*, *tt3* and *rdo1* to *rdo4*, NIL DOG17-1 was crossed with those mutants. In addition, C119 is crossed with *ga1-3* to bring *ga1-3* in the *dog1* background. To obtain the specific allele combinations, the F3 selfed progenies from F2 plants of these crosses were selected on their mutant phenotypes and these selected plants were then analysed with the CAPS marker DFR (*tt3* gene; TAIR: www.arabidopsis.org) located within the Cvi introgression on chromosome five, to determine the *DOG1* allele.

Germination Assays

All germination experiments were performed in 6-cm Petri dishes on filter paper (no. 595, Schleicher & Schuell, Keene, NH). Of each genotype three or four replicates were sown (50-100 seeds of one replicate per Petri dish). Filter papers were either soaked with water, 10 mM KNO_3, 10 µM norflurazon, 0-100 µM GA_{4+7} or 0-100 µM paclobutrazol. Norflurazon was dissolved in a few drops of acetone and thereafter diluted in water to get the appropriate concentration. GA_{4+7} was dissolved in a few drops of 1 M KOH and diluted with a phosphate citrate buffer, containing 3.3 mM K_2HPO_4 $3H_2O$ and 1.7 mM citric acid (pH 5.0). Cold treatment was performed by placing the imbibed seeds on water-soaked filter paper for 7d at 4°C in darkness. To allow germination the petri dishes with the seeds were transferred to a climate room (25°C, 16-h light/day; model TL57, Philips) and germination was scored after 7 d. The average germination percentages were calculated with SE of the triplicates.
Unless mentioned in the text, seeds compared in one experiment are grown together under the same experimental conditions and harvested on the same date.

The Days of Seed Dry Storage required to reach 50 percent of germination ($DSDS_{50}$) were calculated by probit regression analysis of the germination proportions on a logarithm time scale, using the statistical package SPSS version 10.0 (SPSS, Inc., Chicago). The average $DSDS_{50}$ of 4 replicates ± SE was calculated.

Chapter 6

Genetic analysis of seed soluble oligosaccharides in relation to seed storability of *Arabidopsis thaliana*

Leónie Bentsink, Carlos Alonso-Blanco, Dick Vreugdenhil, Karine Tesnier,
Steven P.C. Groot and Maarten Koornneef

Abstract

Seed oligosaccharides (OSs) and especially raffinose series OS (RSOs) are hypothesised to play an important role in the acquisition of desiccation tolerance and consequently in seed storability. In the present work we analysed the seed-soluble OS (sucrose, raffinose and stachyose) content of several Arabidopsis accessions and thus identified the genotype Cape Verde Islands having a very low RSO content. By performing quantitative trait loci (QTL) mapping in a recombinant inbred line population, we found one major QTL responsible for the practically monogenic segregation of seed stachyose content. This locus also affected the content of the two other OSs, sucrose and raffinose. Two candidate genes encoding respectively for galactinol synthase and raffinose synthase were located within the genomic region around this major QTL. In addition, three smaller-effect QTL were identified, each one specifically affecting the content of an individual OS. Seed storability was analysed in the same recombinant inbred line population by measuring viability (germination) under two different seed aging assays: after natural aging during 4 years of dry storage at room temperature and after artificial aging induced by a controlled deterioration test. Thus, four QTL responsible for the variation of this trait were mapped. Comparison of the QTL genetic positions showed that the genomic region containing the major OS locus did not significantly affect the seed storability. We concluded that in the studied material neither RSOs nor sucrose content had a specific effect on seed storability.

This chapter has been published in Plant Physiology (2000) **124**, 1595-1604

Introduction

In many plants species, including Arabidopsis (Ooms et al., 1993), seed maturation is accompanied by the accumulation of soluble oligosaccharides (OSs) (Horbowicz and Obendorf, 1994). These OSs, mainly Suc and raffinose series oligosaccharides (RSOs) are found in cotyledons, seed coats and hypocotyls (Obendorf, 1997 and references therein). RSOs are derivatives of Suc to which Gal units are added to the GLc moiety of Suc through alpha- (1,6) bonds. Raffinose contains one Gal unit whereas stachyose has two such units. RSOs appear during the later stages of seed development and they disappear upon germination (Obendorf, 1997). In Arabidopsis, seeds of the accession Landsberg *erecta* (L*er*) accumulate raffinose and especially stachyose at the later stages of seed development, whereas the Suc content remains constant (Ooms et al., 1993). These three OSs together represent approximately 2% of the dry weight of mature seeds.

Studies in several species such as soybean (*Glycine max*), maize (*Zea mays*) and brassica (*Brassica campestris [rapa]*) have suggested that OSs might be involved in the protection of seeds against damage during seed dehydration and aging, and therefore in seed survival and storability (reviewed by Obendorf, 1997; Sinniah et al., 1998). OSs have been speculated to be involved in the protection of membranes, proteins and nucleic acids against damage that occurs during and upon the withdrawal of water in the drying seeds (Hoekstra et al., 1991). This protective role of OSs has been explained mainly by their capacity to retain the integrity of membranes through their interaction with the phospholipid headgroups, thus replacing water during dehydration. In addition, removal of water molecules from phospholipids can lead to membrane phase transitions at physiological temperatures (Hoekstra et al, 1989). When water is available, these transitions coincide with membrane leakage and cell death. Hoekstra et al. (1991) have shown that OS can depress the temperature of membrane phase transitions and prevent leakage of cellular solutes. Furthermore, free radicals may accumulate and cause damage to cellular components and structures during seed aging because scavenging systems are not operating at low moisture contents (McDonald, 1999). It has been suggested that OSs may form a viscous glassy state (condition in which a liquid achieves such high viscosity that it resembles a solid) (Leopold et al., 1994), which prevents molecular interactions, resulting in damage to membranes and macromolecules (Crowe et al., 1987). This glassy state in seeds seems to serve as a physical stabilizer protecting against deteriorative reactions. Particularly, RSOs have been shown to have an excellent ability to form stable glasses and therefore, they have been considered essential components for the storability of seeds (Koster and Leopold, 1988).

Seed storability, defined as the longevity of seeds after storage, represents a trait important for the conservation of seed resources. However, to test this character seeds need to be stored for long times and for that reason, so-called controlled deterioration tests

(CDT) have been developed (Powell and Matthews, 1984; Hampton and Tekrony, 1995) as an alternative to analyse this property more efficiently. In such tests, seed moisture content and temperature are increased to artificially accelerate seed aging. The viability of the seeds after CDT has been shown to be a reliable measurement for determining seed storability in a number of species, including Arabidopsis, where it has been tested using mutant seeds with storability defects (K Tesnier, personal communication).

Despite substantial studies of both traits, seed OS composition and seed storability, their genetic analysis has been hampered due to their quantitative nature. It is only in the last decade with the advent of molecular marker technologies and the development of quantitative trait locus (QTL) mapping procedures that genetic analysis and the identification of genomic regions controlling quantitative traits have become feasible (Tanksley, 1993; Jansen, 1996). The use of homozygous permanent mapping populations such as recombinant inbred lines (RILs) is very efficient in this respect, because it is possible to study an indefinite number of traits on the same experimental population, enabling the detection of loci with putative pleiotropic effects (i.e. one locus affecting different traits) (Prioul et al., 1997). The localization of QTL in plant species such as Arabidopsis, where molecular analysis can be performed efficiently, may eventually lead to the molecular identification of the respective genes (for review, see Alonso-Blanco and Koornneef, 2000).

In the present work we have analysed the seed soluble OS content of several Arabidopsis accessions and identified the genotype Cape Verde Islands (Cvi) as having a very low RSO content. A QTL mapping approach has been used on a RIL population derived from a cross between Cvi and the laboratory accession Landsberg *erecta* (L*er*) (Alonso-Blanco et al., 1998b) differing in the OS composition, to identify and locate the loci responsible for OS variation. In addition, we have analysed the same RIL population to locate QTL for seed storability to investigate if there are loci with pleiotropic effects on these two traits in Arabidopsis as it has been suggested for other species.

Results

*Variation for seed-soluble OS content in accessions of Arabidopsis, RILs and reciprocal crosses of L*er *and Cvi*

To determine genetic variation in seed-soluble OS content among Arabidopsis accessions, we have analysed the seed-soluble OS composition of 10 accessions (Figure 6.1). Suc appeared as the most abundant sugar in all of them, and raffinose and stachyose were the only detectable RSOs. Monosaccharides (Fru and Glc) were not detectable. The accession Wassilewksija (Ws-1) showed the highest RSO content and, opposite to all other analysed accessions, had more raffinose than stachyose. In contrast, Cvi contained very low levels of RSOs. Nevertheless, the total OS content in Ws-1 and Cvi was not significantly different from that of other genotypes with the exception of Shahdara (Shah) and Rschew-1 (RLD-1) in which the total amount of Suc plus RSOs was significantly higher compared to the other accessions.

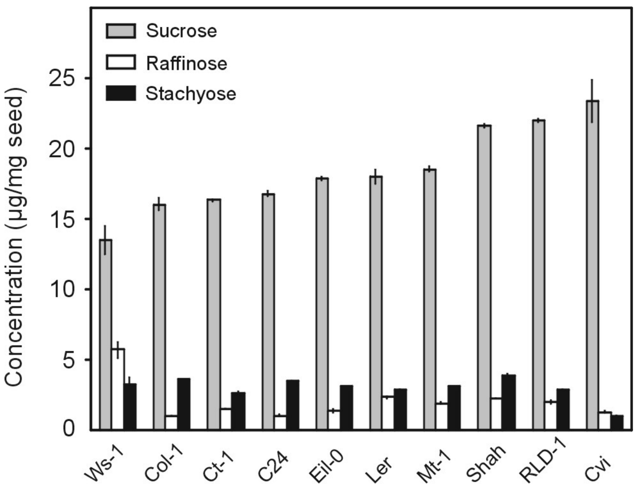
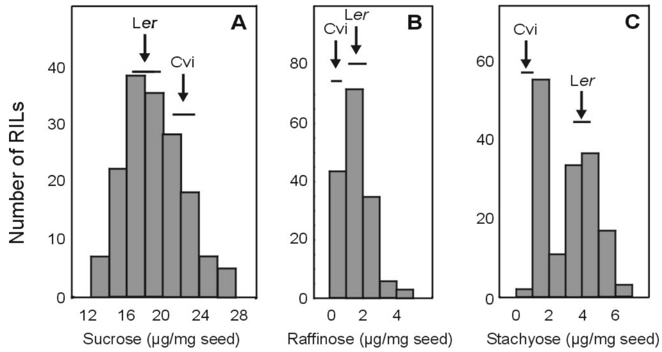


Figure 6.1. Seed-soluble OS content of 10 *Arabidopsis thaliana* accessions, Wassilewksija-1 (Ws-1), Columbia-1 (Col-1), Catania-1 (Ct-1), C24, Eilenburg-0 (Eil-0), Landsberg *erecta* (L*er*), Martuba-1 (Mt-1), Shahdara (Shah), Rschew-1 (RLD-1) and Cape Verde Islands (Cvi). Columns correspond to means of two measurements of bulked seeds from six plants; vertical bars indicate the range of variation.

To identify and locate QTL responsible for the genetic variation observed for seed OS contents we have analysed a RIL mapping population derived from the cross between Cvi, the lowest RSO content accession, and the laboratory accession L*er*. Although, the analysis of other crosses such as Ws-1 x Cvi might show larger variation, the availability of a permanent mapping population between Cvi and L*er* offers unique advantages (Alonso-Blanco and Koornneef, 2000). As shown in Figure 6.2, the RIL population showed transgression in both directions for Suc content. However, for raffinose and stachyose, transgression was only detected towards higher contents due to the low amount of RSOs in the Cvi parent. The levels of the three OSs were correlated among the RILs; raffinose content was positively correlated with stachyose content (r = 0.77), whereas Suc content correlated negatively with the raffinose and stachyose contents (r = -0.34 and r = -0.39 respectively).

Figure 6.2. Frequency distributions of Suc, raffinose and stachyose content in seeds of the L*er*/Cvi RIL population. Arrows correspond to the parental line means and horizontal bars represent their ranges of variation.



To investigate if the OS content was maternally affected we have analysed hybrid seeds obtained from reciprocal crosses between the parental lines L*er* and Cvi. As shown in Figure 6.3, the content of the three soluble OS was significantly different in the reciprocal hybrid seeds. The two sorts of hybrid seeds differed in their raffinose content, the contents being similar to those observed in seeds of the female parent, indicating maternal effects. The stachyose contents of the reciprocal hybrid seeds were different but intermediate between both parental values, suggesting both, maternal and zygotic effects and partial dominance of the higher stachyose level. In contrast, the Suc content of the hybrid seeds was comparable to that in the seeds of the male parent, suggesting paternal effects.

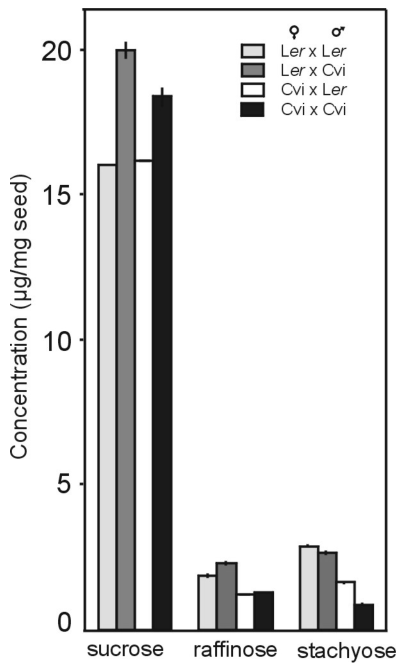


Figure 6.3. Soluble OS content in hybrid seeds obtained from reciprocal crosses between L*er* and Cvi (means of two measurements; vertical bars indicate the range of variation).

Mapping QTL controlling seed-soluble OS content

QTL mapping was performed for the quantity of the three major OSs (Suc, raffinose and stachyose) (Figure 6.4). In total, four genomic regions were detected, one affecting the content of the three OSs and an additional QTL specific for each OS. The additive effects of these QTL accounted for 72%, 45% and 32% of the total variance for stachyose, raffinose and Suc respectively. The region on the lower arm of chromosome 1 near the AFLP marker EC.88C explained 71%, 41% and 14% of the variation for stachyose, raffinose and Suc respectively. This QTL appears as the major locus responsible for the observed variation in RSO content, its L*er* allele increasing the seed content of both RSOs and decreasing the content of Suc. The L*er* allele for the additional QTL affecting raffinose (chromosome 2) and stachyose (chromosome 1) decreases the content of these OSs. The additional QTL affecting Suc (chromosome 3) is the largest QTL affecting this OS and it explained 18% of the variance of the lines. The L*er* allele at both Suc QTL decreases the content, which indicates that other QTL controlling this trait remained undetected in order to explain transgression for higher Suc (Figure 6.2).

Analysis of QTL interactions detected significant epistasis between the two loci on chromosome 1 affecting stachyose content (P = 0.003). RILs carrying a Cvi allele at marker GD.86L and a L*er* allele at marker EC.88C had a higher stachyose content than the high RSO content parent L*er*, while RILs bearing the opposite allelic combination showed a similar content as Cvi. Therefore, the Cvi allele at the QTL near the marker GD.86L increases the stachyose content only when the QTL at EC.88C carries a L*er* allele, in agreement with the observed transgression in only one direction (Figure 6.3).

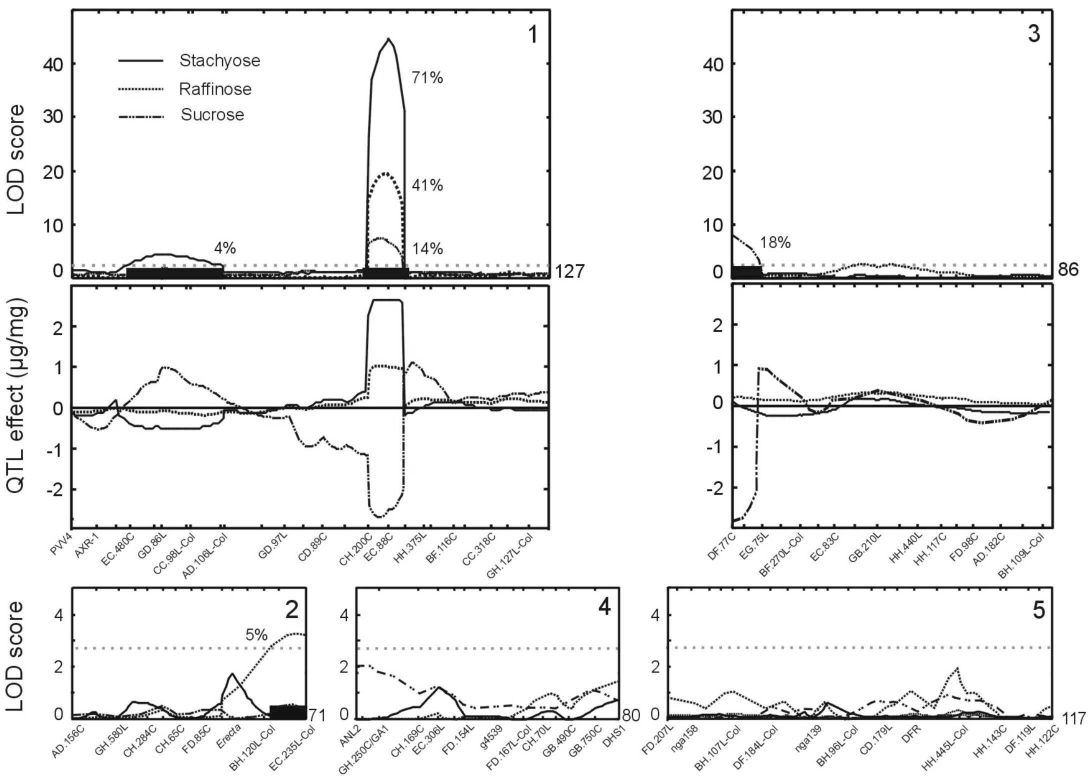


Figure 6.4. QTL likelihood maps for seed soluble OS contents of the five linkage groups of Arabidopsis. The abscissas correspond to the genetic maps in cM; 1 through 5 indicate the linkage group number. The horizontal dotted lines correspond to the LOD score threshold of 2.8 used to declare the presence of a QTL. Two-LOD support intervals for the significant QTL are shown as solid bars along the abscissas. QTL effects are shown for linkage groups 1 and 3 where the major QTL are located. These are given as twice the additive allele effects, i.e., as the mean differences between the two RIL genotypic groups carrying the L*er* and Cvi alleles. A positive QTL effect represents that the L*er* allele increases the content. The percentage of phenotypic variance explained by each QTL is reported close to the corresponding LOD score peak.

Localisation of RSOs biosynthesis genes

Since the QTL on chromosome 1 is the major locus controlling the OS variation in seeds, we searched for candidate genes in this genomic region. The genomic nucleotide sequences available in the databases as part of the International Arabidopsis Genome Project (http://www.arabidopsis.org) were analysed to look for putative genes encoding known OS biosynthesis and degradation enzymes. The four putative enzymes were: galactinol synthase (GS; EC 2.1.4.123), raffinose synthase (RS; EC 2.4.1.82), stachyose synthase (SS; EC 2.4.1.67) and α-galactosidase (EC 3.2.1.22) (Krebbers et al., 1997). GS catalyzes the first committed step in the biosynthesis of RSO (Krebbers et al., 1997). RS and SS control subsequent steps in the biosynthesis of raffinose saccharides by adding a Gal unit to Suc or raffinose, respectively. The enzyme α-galactosidase degrades the RSOs. Two genes encoding respectively, GS and RS were located on two different bacterial artificial chromosomes (BACs) from a BAC contig in the lower arm of chromosome 1. Furthermore, homologous sequences of these genes were also found in other regions of the Arabidopsis genome, which suggests that both genes belong to gene families. We designed cleaved amplified polymorphic sequence (CAPS) markers specific for both genes and mapped them in the L*er*/Cvi RIL population. The two genes appeared closely linked to the QTL around marker EC.88C on chromosome 1 and therefore, they are possible candidate genes for this major QTL.

Mapping QTL for seed storability

To determine whether there is a functional relationship between seed-soluble OS content and storability of seeds we have analysed the same L*er*/Cvi RILs to map QTL for this trait. Seed storability was measured as viability under two different seed aging assays, after natural aging following 4 years of storage and after artificial aging promoted by a controlled deterioration test (see 'Material and Methods'). The two parental accessions L*er* and Cvi were both relatively sensitive to the CDT, although they differed significantly (Figure 6.5), L*er* being more sensitive to the test conditions (lower storability) than Cvi.

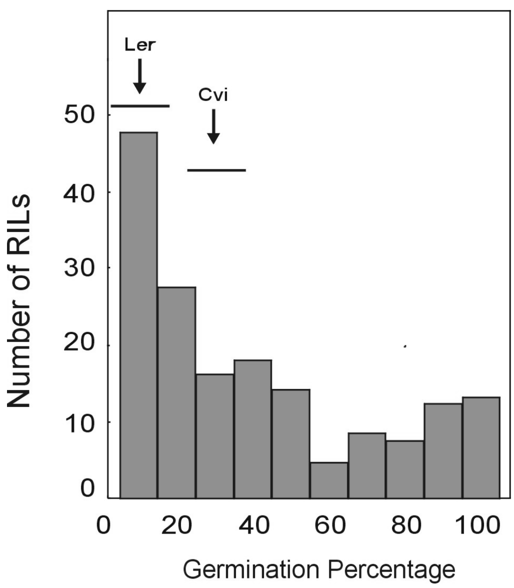


Figure 6.5. Frequency distribution of the seed viability (germination percentage) of the L*er*/Cvi RIL after the CDT. Arrows correspond to the parental line means and horizontal bars represent their ranges of variation.

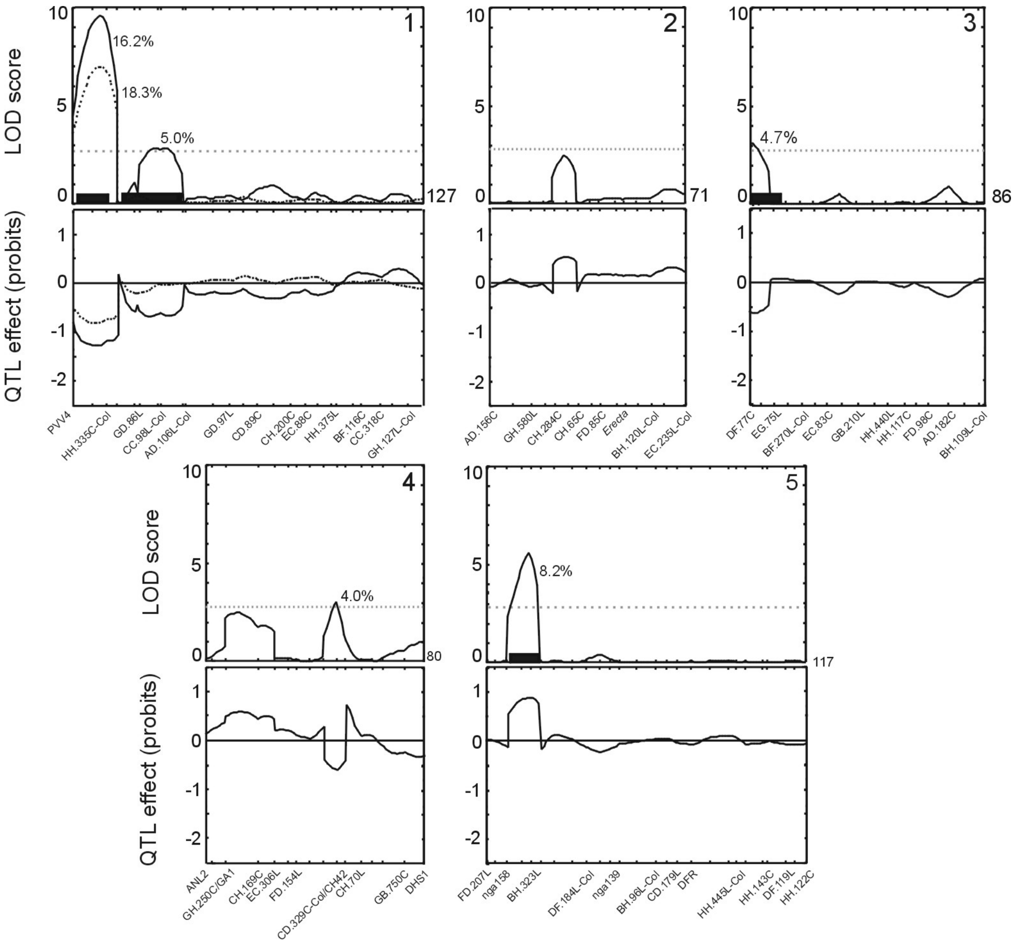


Figure 6.6. QTL likelihood maps for seed storability measured as viability after the controlled deterioration test (solid line) and four years of natural aging (dashed line). The abscissas correspond to the genetic maps in cM (the linkage group number being indicated in the right upper corner of each panel). Horizontal dotted lines corresponds to the LOD score threshold of 2.8 used to declare the presence of a QTL. Two-LOD support intervals for the significant QTLs are shown as solid bars along abscissas. The additive QTL effects are expressed as probit units of the germination percentage after the CDT. These are estimated as the mean differences between the two RIL genotypic groups carrying the L*er* and Cvi alleles (a positive value implies L*er* increases the corresponding phenotypic value). The percentage of variance explained by each QTL is reported close to the corresponding LOD score peak.

The RIL population showed transgression towards higher germination percentages, indicating the presence in the two parental lines of alleles increasing and reducing the tolerance to the given controlled stress (Figure 6.5). Four putative QTL (Figure 6.6) were identified affecting viability after controlled deterioration, their additive effects accounting for 56.5% of the phenotypic variance. These QTL are located on chromosome 1 (two closely linked QTL), chromosome 3 and chromosome 5. In addition to germination, the fraction of aberrant seedlings among germinating seeds is another parameter commonly used to measure seed storability (Coolbear, 1995) and was also analysed. This trait showed a high correlation with the germination values after the CDT (r = 0.90) and QTL at similar positions were found to be responsible for its variation (data not shown).

The germination percentages after natural aging ranged from 61% to 100% (average 97). The QTL mapping for seed viability after 4 years of storage resulted in one QTL located on chromosome 1, which accounted for 18.3% of the phenotypic variance; a QTL at similar position was also the largest effect QTL detected after the CDT. QTL mapping for natural aging did not reveal any additional loci affecting storability.

The comparison of map positions between the detected QTL for seed-soluble OS content and storability showed two genomic regions containing QTL for both traits (Figure 6.7). The top of chromosome 1 affected stachyose content and storability and the locus on top of chromosome 3 affected Suc content and storability. No significant effect on seed storability was found in the vicinity of the OS QTL region around EC.88C on chromosome 1.

Discussion

During the past decade the use of molecular marker technology and QTL mapping have contributed to a better understanding of the genetic basis of many agriculturally and biologically important quantitative traits, such as yield (Stuber et al., 1987) resistance/tolerance to biotic and abiotic stress (Koornneef and Peeters, 1999), and nutritional quality in numerous crops species (Paterson et al., 1991). In the present work we have used Arabidopsis for the analysis of seed content of the three main soluble OSs (Suc, raffinose and stachyose), and showed that considerable variation exists among accessions (Figure 6.1). By performing QTL mapping for seed OS content in a RIL population derived from a cross between the two accessions L*er* and Cvi differing in seed OS composition, we identified one major QTL responsible for the practically monogenic segregation observed for seed stachyose content. It is very likely that this QTL affected pleiotropically the contents of the three detected OSs, showing opposite effects on RSOs and Suc contents. This locus might therefore be involved in the biosynthetic pathway of RSOs. Two candidate genes encoding for GS and RS were mapped within the 2-logarithm-of-odds (LOD) support interval of this QTL, the genetic distance between both genes being about 5 cM. Nevertheless this region has not been sequenced completely and other genes not found in this analysis and involved in RSO biosynthesis and degradation (i.e. SS and α-galactosidase) might locate within this interval. Further fine mapping using recombinants in this region, as well as complementation by plant transformation can be used in the future to determine whether any of these candidate genes corresponds to this QTL.

The maternal effects for stachyose and raffinose seed content detected by the analysis of OSs in hybrid seeds obtained in reciprocal crosses between the parental lines, suggest either production of raffinose and stachyose by the testa or import of RSOs from maternal tissues. In agreement with this observation is the function that OSs and especially stachyose might have as transport sugars, as was described in cucurbits (Beebe and Turgeon, 1992). However, the apparent paternal effects observed on Suc content are rather unexpected and suggest an even more complex genetic control of the seed Suc content. Nevertheless, it should be noted that the parental and the reciprocal hybrids seeds differ considerably in size, and that paternal effects on the seed size variation of these materials could not be excluded (Alonso-Blanco et al., 1999). The two seed Suc content QTL identified in this work co-located with seed size QTL reported previously (Alonso–Blanco et al., 1999). Since Suc seems to play an important role in the metabolic control of seed size (Weber et al., 1997), it is possible that some of these co-locations are due to pleiotropic effects from the same gene on both traits.

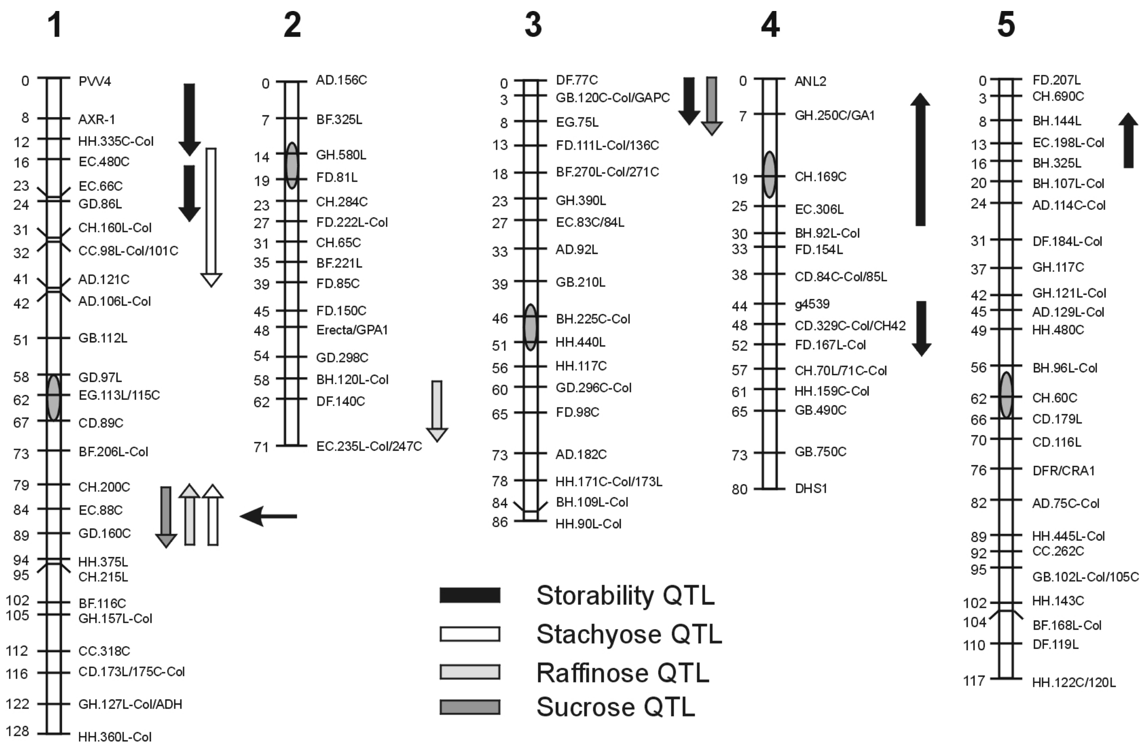


Figure 6.7. L*er*/Cvi linkage map showing the genetic location of QTL affecting seed soluble OS contents and seed storability. Arrows indicate the direction of L*er* allele phenotypic effect (up increasing, down decreasing). The length of the arrows depicts the 2-LOD support intervals. The horizontal black arrow points to the major OS QTL.

The accessions L*er* and Cvi were originally collected in nature and therefore, the observed variation in seed OS contents might be related to the adaptive properties of both genotypes. It has previously been shown (Liu et al., 1998) that cold increases the levels of GS mRNA, which suggests that this gene and RSOs may play a role in cold adaptation. L*er* originates from Northern Europe (Rédei 1992), whereas Cvi comes from the Cape Verde Islands (Lobin, 1983) located at 16°N and having a subtropical climate. Thus, the L*er* and Cvi RSO contents and the allele effects at the major OS QTL were in agreement with this speculation, suggesting that seed RSOs might be involved in seed survival at low temperatures.

To determine whether OSs are important for seed storability, we have mapped QTL affecting this trait in the same RIL population analysed for seed soluble OS content. We have measured seed storability as viability (germination) after natural aging and after artificial aging induced by a CDT. The major QTL affecting storability was detected in both seed aging assays. However, the effect of the CDT on the viability of the seeds was much stronger than the effect of the natural aging, resulting in a more accurate mapping. Therefore, these results indicate that the CDT is a useful method to artificially accelerate seed aging. The comparison of map positions between the QTL identified controlling seed

OS contents and the QTL affecting seed storability showed co-locations in two regions (Figure 6.7). The genomic region on top of chromosome 1 only marginally affected the stachyose content (its additive effects accounting for 4% of the variance) but considerably influenced seed storability (16.2% of the variance); the region on chromosome 3 strongly affected Suc content (18% of the variance) and had only a slight effect on seed storability (4.7% of the variance). In both cases a higher OS content co-segregates with better storability, and these co-locations might be interpreted either as the presence of two closely linked genes or as a consequence of pleiotropy, higher seed content of either stachyose or Suc leading to increased viability after CDT. However, the major locus on chromosome 1 controlling RSOs and sucrose in opposite directions did not show any significant effect on the germination ability after controlled deterioration. We conclude that in the studied material the variation observed for OS content does not evidently affect seed viability after CDT and that neither RSOs nor sucrose had an apparent specific effect on seed storability. Nevertheless, this does not imply that OSs are not involved in seed protection and storability. Both RSOs and Suc might affect seed storability but due to the effect of the major OS QTL on both RSOs and Suc contents, their roles on storability might have been compensated and masked. Given the strong effect of the major OS locus on chromosome 1 on RSOs and the relatively smaller effect on Suc seed content, it is suggested that RSOs seem not to be substantially better than Suc in protecting the seeds against controlled deterioration.

In addition, several other loci appeared to affect seed storability independently from seed OS content, which shows genetic variation for other factors involved in this complex trait. Deterioration of seeds during storage has frequently been related to free radical-mediated oxidative damage of proteins, nucleic acids and membranes (Coolbear, 1995). Factors controlling either the protection of cell structures or the recovery from damage might determine the difference between the two genotypes, L*er* and Cvi. In particular, some of the QTL that we have identified for viability after CDT co-located with a cluster of genes related to stress tolerance, on the top of chromosome 1 (Thorlby et al., 1999; Taji et al., 1999). Further fine mapping of these loci, combined with the analysis of other related traits, will allow the identification of the corresponding genes involved in the storability genetic variation.

Materials and methods

Plant material and growth conditions

The following Arabidopsis accessions (the Nottingham Arabidopsis Stock Centre nos. are indicated between parentheses) were analysed for OS content: C24 (N906), Col (N907), Ct-1 (N1094), Cvi (N8580), Eil-0 (N1133), L*er* (NW20), Mt-1 (N1380), Shah (N929), RLD-1 (N913) and Ws-1 (N2223). F10 seeds of a set of 162 RILs derived from crosses between the laboratory strain L*er* originated from Northern Europe (Rédei 1992) and the accession Cvi, from the Cape Verde Islands (Lobin, 1983) were analysed. These RILs have been previously characterised for amplified fragment length polymorphism (AFLP) and cleaved amplified polymorphic sequence (CAPS) markers (Alonso-Blanco et al., 1998b)

Plants were grown in an air-conditioned green house supplemented with additional light (model TDL 58W/84, Philips, Eindhoven, The Netherlands) from middle September until the beginning of April, providing a day-length of at least 14 hr (temperature 22–25°C). To reduce developmental and environmental effects on the various seed traits analysed, we synchronised the onset of flowering and thereby of seed development. To do so the RILs were planted at three subsequent weeks according to their respective flowering times (Alonso-Blanco et al., 1998a), and the seeds of all genotypes were harvested on the same day. Twelve plants per RIL were grown in a two-block design to avoid environmental effects, and their seeds harvested in a single bulk. For the accession analysis, six plants per genotype were grown and their seeds were bulk harvested. Seed OS content analysis and the CDT were performed on seeds of the RILs and parental lines stored in the same open box for 2 years and 2 months at room temperature. To test the natural aging the seeds have been stored for 4 years under the same conditions. The seeds of the accessions and the crosses were stored under similar conditions during 4 months before analysis of their OS content.

Sugar content measurement

One hundred seeds from bulks of six to 12 plants were weighed and homogenised in 80% (v/v) methanol, with the addition of 25 µg melezitose as internal standard. The homogenate was heated for 15 min at 75°C and centrifuged 5 min at 10,000g. The supernatant was vacuum-evaporated and its residue was resuspended in 0.5 mL pure water and injected into a Dionex HPLC system (Dionex Corporation, Sunnyvale, CA). Sugar content was determined with a high-pH-anion-exchange HPLC, using a gradient pump module (model GP40) and an ED40 pulsed electrochemical detector (Dionex, Sunnyvale, CA). Sugars were chromatographed on a CarboPac PA100 4- X 250-mm column (Dionex) preceded by a guard column (CarboPac PA100, 4 X 50 mm).

Mono-, di-, and trisaccharides were separated by elution in increasing concentration of NaOH (50-200 mM), with a flow rate of 1 mL per minute. Peaks were identified by co-elution of standards. Sugar quantity was corrected by means of the internal standard and transformed to micrograms of sugar per milligram of seed.

Seed storability measurement

Seed storability was determined as viability (germination) after natural aging during 4 years and after the CDT. The CDT was performed as follows: bulked seeds from 12 plants, stored for 2 years and 2 months at room temperature, were equilibrated at 85% relative humidity (the obtained seed moisture content was approximately 10.5%). Thereafter the seeds were artificially aged during 4 d at 40°C, because preliminary experiments showed that this deterioration is the most discriminative between L*er* and Cvi (data not shown), and dried back at 32% relative humidity and 20°C during 3 d. The seeds were stored at 5°C, and thereafter germination was tested. Germination of 100 CDT treated seeds (two replicates of 50 seeds) was tested on moist filter paper at 20°C and a 12-h-dark/12-h-light cycle, by visually inspecting root tip emergence during 2 weeks. Non-germinating seeds were not viable as shown by the absence of staining in a tetrazolium viability test (Moore, 1985). In addition, the quality of the emerging seedlings was recorded by scoring the number of morphologically aberrant seedlings.

Natural aging was tested by performing a germination test on 4-years-old seeds (stored dry at room temperature, without humidity control). Seeds were sown in triplicate (70-100 seeds per 6-cm Petri dish) on water-soaked filter paper (No. 595, Schleicher & Schuell, Keene, NH) and exposed to cold (4°C) during 3 d. Thereafter seeds were transferred to a climate room (25°C, 16 h light/day; model TL57, Philips) and germination was scored after 7 d. The average germination percentages of the three replicates were calculated.

QTL analysis

To map QTL using the RIL population, a set of 99 markers covering most of the *Arabidopsis* genetic map was selected from the previously published RIL L*er*/Cvi map (Alonso-Blanco et al., 1998a). These markers spanned 482 cM, with an average distance between consecutive markers of 5 cM and the largest genetic distance being 12 cM. Storability data (germination after 4 years storage and viability after the CDT) were transformed to probit units to achieve normality.

The computer program MapQTL™ version 4.0 (Plant Research International, Wageningen-University and Research Centre, Wageningen, The Netherlands) was used to identify and locate QTL linked to the molecular markers using both interval mapping and multiple-QTL model mapping (MQM) methods as described in its reference manual (http://www.cpro.wag-ur.nl/cbw/mapping/). The estimated additive effect and the percentage of variance explained by each QTL as well as the total variance explained by all of the QTL affecting a trait, were obtained with MapQTL in the final MQM model. For this, different cofactor markers were tested around the putative QTL positions (van Ooijen and Maliepaard, 1996), selecting as final cofactors the closest marker to each QTL, i.e. those maximizing the logarithm-of-odds (LOD) score. A LOD score threshold of 2.8 was applied to declare the presence of a QTL, which corresponds to a general genome-wide significance $P = 0.05$ for normally distributed data, as was determined by extensive simulation experiments (van Ooijen, 1999). We verified this threshold for interval mapping by applying the permutation test to each data set (10 000 repetitions) and found a $P = 0.05$ LOD threshold of 2.6 for all traits. Two-LOD support intervals were established as ≈ 95% confidence intervals (van Ooijen, 1992).

For every trait, two-way QTL interactions were analysed by analysis of variance at a significance level of $P < 0.005$, using the General Linear Model module of the statistical package SPSS version 7.5 (SPSS Inc., Chicago). For each analysis, the closest linked markers to the corresponding detected QTL were used as random factors in the ANOVA (the same markers used as cofactors in the MQM mapping with MapQTL).

Location of raffinose series oligosaccharide biosynthetic genes

CAPS markers (Konieczy and Ausubel, 1993) were developed to genetically map the genes encoding GS (EC 2.1.4.123) and RS (EC 2.4.1.82) in the L*er*/Cvi RIL population. The map locations were determined with the software package JOINMAP (http://www.cpro.wag-ur.nl/cbw/mapping/).

The primers for GS were based on the genomic nucleotide sequence of the BAC F8A5 (accession no. AC002292) located on chromosome one. The forward primer was 5'-TCG GTT ATT CTC CTT TGT TGT TTG-3'. The reverse primer was 5'-TTT CTA TGC CGT GAT GGA CTG TT -3'.

The primers for RS were based on the nucleotide sequence of BAC F20N2 (accession no. AC002328) located on chromosome one. the forward primer was 5'- GGG AGG AGT CAA ACC AGG TG-3'. The reverse primer was 5'-GGC ATC AAT GTC ACT GGT AAA G-3'.

PCR were carried out in 50-µL volumes containing, 50 ng of genomic DNA, 100 µM of each deoxynucleotide, 100 ng of both primers and 0.2 U Taq DNA polymerase. Conditions for amplification were as follows: 30 s at 94°C; annealing for 2 min at 51°C or 61°C (GS and RS respectively); extension for 2 min at 72°C. The cycle was repeated 35 times. To detect the polymorphism 10 µL of PCR product was cleaved with the restriction enzyme *Rsa*I or *Bsa*BI (for GS and RS respectively) and analysed in a 1.5% (w/v) agarose gel.

Chapter 7

The genetics of phytate and phosphate accumulation in seeds and leaves of *Arabidopsis thaliana,* using natural variation

Leónie Bentsink*, Yuan Kejun*, Maarten Koornneef and Dick Vreugdenhil

* These authors contributed equally to the work

Abstract

Phytate (myo-inositol-1,2,3,4,5,6- hexakisphosphate, InsP6) is the most abundant P-containing compound in plants, and an important anti-nutritional factor, due to its ability to complex essential micro-nutrients, e.g., iron and zinc. Analysis of natural variation for InsP6 and Pi accumulation in seeds and leaves for a large number of accessions of *Arabidopsis thaliana*, using a novel method for InsP6 detection, revealed a wide range of variation in InsP6 and Pi levels, varying from 7.0 mg to 23.1 mg InsP6 per gram seed. Quantitative trait loci (QTL) analysis of InsP6 and Pi levels in seeds and leaves, using an existing recombinant inbred line population, was performed in order to identify gene(s) that are involved in the regulation of InsP6 accumulation. Five genomic regions affecting the quantity of the InsP6 and Pi in seeds and leaves were identified. One of them, located on top of chromosome 3, affects all four traits. This QTL appears as the major locus responsible for the observed variation in InsP6 and Pi contents in the L*er*/Cvi RIL population, the L*er* allele is decreasing the content of both InsP6 and Pi in seeds and in leaves. The InsP6/Pi locus was further fine-mapped to a 99 kb region, containing 13 open reading frames. The maternal inheritance of the QTL and the positive correlation between InsP6 and total Pi levels both in seeds and in leaves indicate that the difference in InsP6 level between L*er* and Cvi is likely to be caused by a difference in transport rather than by an alteration in the biosynthesis. Therefore, we consider the vacuolar membrane ATPase subunit G, located in the region of interest, as the most likely candidate gene for the InsP6/Pi.

Introduction

The major form in which phosphorus occurs in plants is myo-inositol-1,2,3,4,5,6-hexakisphosphate, commonly referred to as phytic acid or InsP6 (Lott et al., 2000). This highly negatively charged compound forms a mixed salt with various mineral cations, e.g., potassium, magnesium, iron and zinc. It is generally assumed that the major role of InsP6 in plants is to act as a storage form for Pi and probably also for cations, and it is most abundant in seeds and fruits. The availability of Pi in soil is positively correlated with the level of InsP6 accumulated in seeds of soybean (Raboy and Dickinson, 1987) and Brassica (Lickfett et al., 1999). Since InsP6 is the major form of phosphorus in seeds, total phosphorus and InsP6 levels in seeds are usually positively correlated (Raboy, 2001).

Due to its mineral-binding capacity, InsP6 affects the nutritional quality of food; minerals, when bound to InsP6, are not or hardly absorbed in the intestine and are largely excreted, resulting in iron and zinc deficiencies, especially in developing countries, when food is mainly seed-based. Monogastric animals only very inefficiently take up phosphorus from InsP6, and hence additional Pi is added to their feed, resulting in P-pollution.

Decreasing the level of InsP6 in seeds is one of the recent strategies to improve iron and zinc availability in food (Lucca et al., 2001). In maize (*Zea mays*), barley (*Hordeum vulgare*), rice (*Oryza sativa*) and soybean (*Glycine max*), mutants with lowered InsP6 levels (*low phytic acid; lpa*) have been isolated (Larson et al., 1998, 2000, Raboy and Gerbasi, 1996, Raboy et al., 2000 and Wilcox et al., 2000). In these mutant lines, a negative correlation was observed between InsP6 and free Pi levels. The explanation is that a smaller portion of total seed phosphorus is converted into InsP6, the total amount of phosphorus in the seeds remaining unaltered. In some mutants also lower inositol-phosphates (mainly InsP3 and InsP5) were found. Mutant lines with a block in InsP6 metabolism in the developing seed appear to have a reduced seed weight, which is leading to reduced yields as well. This reduction in seed dry weight might be because of reduction in starch accumulation that results from the disturbance of phosphorus or because of disruption in inositol or inositol-phosphate metabolism found in *lpa* seeds (reviewed by Raboy, 2001).

Except its effect on seed weight it now appears that InsP6 probably serves several cellular functions as well: as a major pool in inositol phosphate pathways (Safrany et al., 1999) and as a second messenger ligand (Sasakawa et al., 1995). Furthermore, it may play a role in DNA double strand break repair (Hanakahi et al., 2000), RNA export from the nucleus (York, et al., 1999), ATP metabolism (Safrany et al., 1999), and in phosphorus and mineral storage (Raboy, 1997). A recent study indicates the role of InsP6 in the physiological response of guard cells to ABA as well (Lemtiri-Chlieh et al., 2000).

In light of these numerous functions for InsP6 in cells generally and in seeds specifically, it is somewhat surprising that the *lpa* mutants perform rather well under field conditions. Clearly the levels of InsP6 typical for normal seeds are not essential for seed functioning.

In eukaryotic cells, InsP6 is synthesised by sequential phosphorylation of *myo*-inositol and/or partly via phosphatidylinositol phosphates (Raboy, 2001; Loewus and Murthy, 2000). However, relatively little is known about plant genomic sequences encoding functions in the late part of the pathway playing a direct role in seed InsP6 accumulation and mobilisation (Raboy, 2001). Only recently, Hitz et al. (2002) reported that a defective myo-inositol-1-phosphate synthase decreased the amount of InsP6 in soybean seeds. Furthermore, it is know that during germination InsP6 is broken down by phytase enzymes, releasing *myo*-inositol and mineral contents. There probably are multiple phytases in seed and pollen, with different patterns of expression and regulation, the molecular biology and genetics of which is relatively undeveloped (Loewus and Murthy, 2000). There is significant progress in the study of microbial phytase genes (Maugenest, 1997).

In the present study we describe a genetic approach to unravel InsP6 and free phosphorus accumulation in seeds, using Arabidopsis as a model species. We have used natural variation that has been shown to be present in accessions of Arabidopsis for a large variety of traits (Alonso-Blanco and Koornneef, 2000).

Genetic variation for InsP6 and Pi accumulation in seeds and leaves has been analysed, using a novel method for InsP6 detection, for a large number of accessions. Furthermore, quantitative trait loci (QTL) analysis of InsP6 and Pi levels in seeds and leaves has been performed, using an existing recombinant inbred line (RIL) population, in order to identify gene(s) that are involved in the regulation of InsP6 accumulation.

Results

Extraction and detection of InsP6

A simple and rapid method was needed to extract, detect and quantify InsP6 in large series of small samples. Anions can easily be separated using exchange chromatography, especially at high pH, and detected by conductivity. Sensitivity of this detection method can largely be increased using post-column removal of the alkali via a suppressor (Weiss, 1991). Figure 7.1 shows chromatograms of a mix of anion standards (A), and an extract of a few mg of Arabidopsis seeds (B), extracted simply by boiling in HCl. Without any further purification, InsP6 (R_t = 15.8) and Pi (R_t = 12.3) could easily be detected. The large peak at 3.2 min in the chromatogram of the seed extract is chloride, added for the extraction. Another major peak in the chromatogram, at 10.9 min, co-eluted with sulphate. Minor peaks, co-eluting with citrate and oxalate were also present, R_t being 12.7 and 11.2 min, respectively.

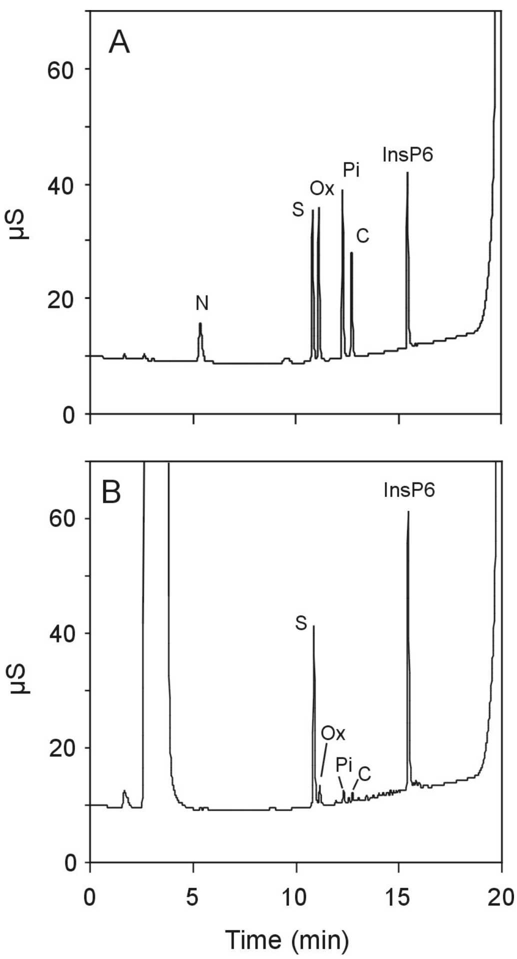


Figure 7.1. Chromatograms of anion standards (A) and a representative extract of seeds (B), analysed by high pH anion chromatography and ion-suppressed detection. Standards: N: nitrate, S: sulphate, Ox: oxalate, Pi: phosphate, C: citrate, InsP6: phytate.

Recoveries of InsP6 were determined by spiking seed samples, varying in weight from 2 – 7 mg, with authentic Na_{12}InsP6. Recoveries were found to decrease with increasing sample weight (Figure 7.2). The response of the detector was linear with the concentrations of InsP6 and P, up to at least 50 $mg.L^{-1}$ (data not shown). It is concluded that the extraction and detection method is reliable for seed samples smaller than 4 mg DW. The protocol was sensitive enough to detect InsP6 in sub-milligram samples. To avoid inaccuracy due to weighing errors or small number of seeds, routinely samples of 2-3 mg were analysed, the smallest sample being 0.8 mg, representing around 50 seeds (Alonso Blanco et al., 1999).

Natural variation for InsP6 and Pi content among Arabidopsis accessions

For a series of 101 accessions of Arabidopsis, the levels of InsP6 and Pi in mature dry seeds, and in full-grown leaves, collected from vegetative plants, were determined (data presented at www.natural-eu.org). The average InsP6 level in seeds was 16.2 ± 2.3, ranging from 7.0 to 23.1 mg/g seed. A weak but significant (r = 0.351) positive correlation between InsP6 and Pi levels in seeds was detected. Pi levels in leaves ranged from 4.6 to 30.5 mg/g fresh weight (FW), averaging at 14.6 ± 5.6 mg/g FW. The levels of InsP6 in the leaves of these accessions were more than two orders of magnitude lower than in the seeds and were not quantified. In order to test whether InsP6 levels in seeds might represent overall higher P-status in the plant, seed-InsP6 levels were plotted against leaf-Pi levels (Figure 7.3). No significant correlation was observed.

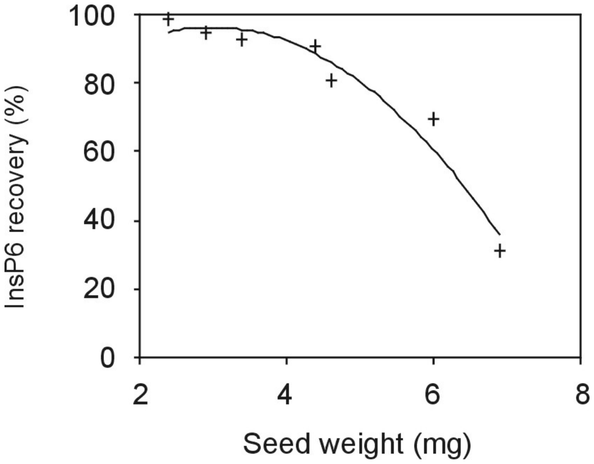


Figure 7.2. Effect of sample weight on recoveries of InsP6, extracted from Arabidopsis seeds. Recoveries were determined by spiking samples with authentic InsP6.

To identify and locate QTL responsible for the genetic variation in InsP6 and Pi between the accessions L*er* and Cvi we have analysed the RIL population derived from a cross between these accessions for InsP6 and Pi levels in seeds and leaves. Although the analysis of other crosses such as Kyoto-1 x Hey-0 (InsP6 levels 7.0 and 23.1 mg/g seed, respectively; Figure 7.3) might show larger variation, the availability of a permanent mapping population between L*er* and Cvi offers unique advantages (Alonso-Blanco and Koornneef, 2000). The L*er*/Cvi population shows considerable transgression towards both directions, for all traits analysed as is shown in Figure 7.4. This indicates that both accessions carry

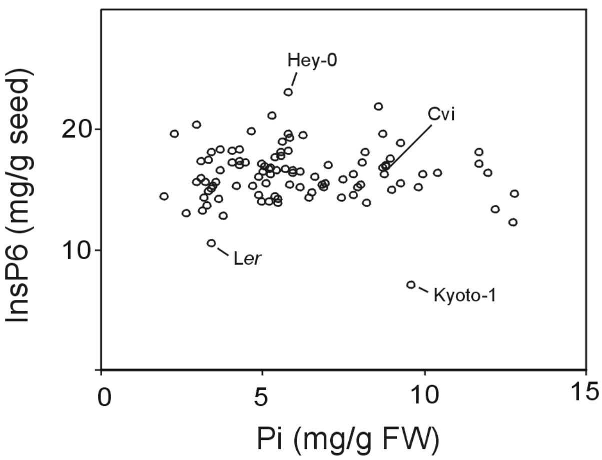


Figure 7.3. Relationship between InsP6 in seeds and free Pi levels in leaves of 101 accessions of Arabidopsis. Extremes in InsP6 levels, and the parents of the Cvi/L*er* RIL population are indicated.

genes that increase and decrease the content of the different compounds that are analysed. Furthermore, there is a positive correlation between the amount InsP6 and Pi, in seeds (r = 0.56) and in leaves (r = 0.40) as well as a positive correlation between the amount of InsP6 in seeds compared to the amount in leaves, and for the level of Pi in seeds compared to that of leaves, (r = 0.35 and 0.34, respectively).

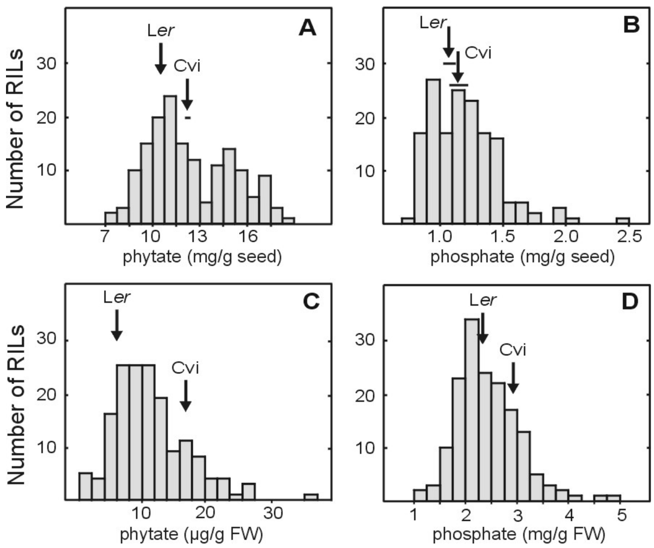


Figure 7.4. Frequency distributions of InsP6 and Pi contents in seeds (A,B) and in leaves (C,D) of plants of the L*er*/Cvi RIL population. Arrows correspond to the parental line means and horizontal bars represent their ranges of variation based on the analysis of two different plants. Parental values of InsP6 and Pi in leaves are based on measurements of a single plant.

Mapping QTL for InsP6 and Pi in seeds and leaves

QTL mapping was performed for the quantity of the InsP6 and Pi in seeds and leaves (Figure 7.5). In total five genomic regions, affecting InsP6 and Pi levels, were detected, one of them on chromosome 3 affecting all four traits. Additional QTL were detected for Pi in seeds on top chromosome 1, and bottom chromosome 2 and 4, for InsP6 in leaves on bottom of chromosome 1 and 2. The additive effects of these QTL accounted for 61.8% of the total variance for InsP6 in seeds, 55.3% for Pi in seeds 35.5% for InsP6 in leaves and 33% for and Pi in leaves. The region on the top of chromosome 3 near the AFLP marker DF.77C explained 61.8%, 36.8%, 22.5% and 28.9% of the variation for InsP6 in seeds, Pi in seeds, InsP6 in leaves and Pi in leaves respectively. Therefore, this QTL appears as the major locus responsible for the observed variation in InsP6/Pi content in the L*er*/Cvi RIL population, the L*er* allele decreasing the content of both InsP6 and Pi in seeds and in leaves. The L*er* allele for the QTL affecting Pi in seeds and InsP6 in leaves (both on chromosome 1) and for Pi in seeds (chromosome 4) increases the content. The L*er* allele for the additional QTL affecting Pi in seeds and InsP6 in leaves (chromosome 2) is decreasing the contents. Analysis of QTL interactions detected no significant epistasis between the different loci ($P > 0.005$).

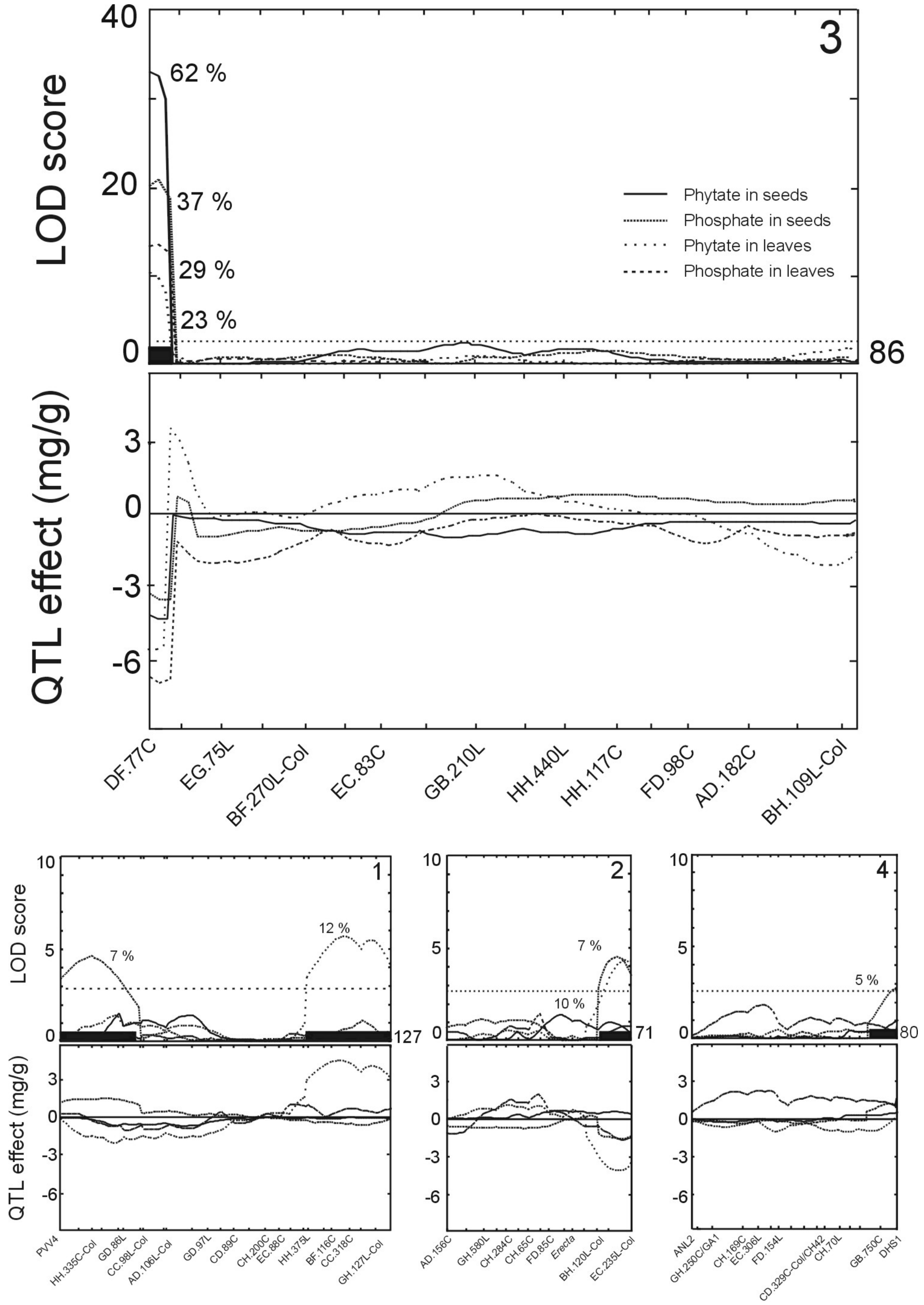


Figure 7.5. QTL likelihood maps for InsP6 and Pi contents in seeds and in leaves. Presented are four of the five linkage groups of Arabidopsis. The abscissas correspond to the genetic maps in cM; 1 through 4 indicate the linkage group number. The horizontal dotted lines correspond to the LOD score threshold of 2.8 used to declare the presence of a QTL. Two-LOD support intervals for the significant QTL are shown as black bars along the abscissa. QTL effects are shown in the second panel below the LOD graph. These are given as twice the additive allele effects, i.e. as the mean differences between the two RIL genotypic groups carrying the L*er* and Cvi alleles. The effects of Pi in seeds as well as in leaves are multiplied with 10 to allow presentation in one figure. A positive QTL effect represents that the L*er* allele increases the content. The percentage of phenotypic variance explained by each QTL is reported near the corresponding LOD score peak.

Analysis of reciprocal effects and dominance

To analyse the effect of the major QTL on the InsP6 and Pi content in more detail a Near Isogenic Line (NIL) was constructed. This NIL (NIL26) contained a Cvi introgression on top of chromosome 3 in a L*er* background. The analysis of 100 F2 plants of a cross between NIL26 and L*er* for seed InsP6 content and the genotype at marker T4P13b, showed a typical monogenic segregation and linkage of the trait to the marker. In addition it revealed the dominance effect of the Cvi allele, heterozygous plants show the higher InsP6 and Pi levels of NIL26 (Figure 7.6). The levels of InsP6 and Pi in seeds of NIL26 are similar to those in Cvi, confirming that indeed the Cvi allele(s) of the introgression on top of chromosome 3 is the major factor responsible for the higher InsP6 and Pi levels of Cvi. Figure 7.7 shows the levels of InsP6 and Pi in seeds of reciprocal crosses between L*er* and NIL26. This data indicates a clear maternal inheritance of InsP6 and Pi levels in seeds, since the InsP6 and Pi level of the hybrid with L*er* as the female parent are similar to that of L*er* itself and when NIL26 is the female parent the levels are like that of NIL26 itself.

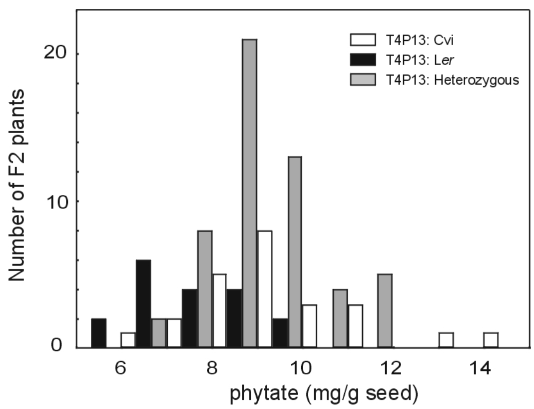


Figure 7.6. Segregation of InsP6 locus. Frequency distribution of InsP6 levels of seeds of 100 F2 plants from the cross NIL26 x L*er*. Lines are grouped based on the genotype for linked marker T4P13b.

Fine-mapping of the major InsP6 and Pi locus

Since the region containing the QTL was still approximately 3 cM in size, a further reduction of this region was needed towards the positional or candidate gene cloning. Therefore, a mapping population of 631 plants was generated from a heterozygous NIL26 sister plant. These plants were analysed with two molecular markers (T4P13b and 17D8LE, Table 7.1) one at the top of chromosome 3 and the other near the end of the Cvi introgression. This revealed 23

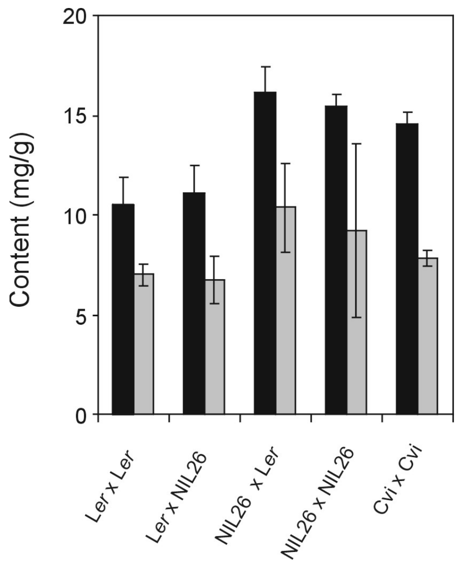


Figure 7.7. Levels of InsP6 (black bars) and free Pi (x10) (gray bars) in seeds of L*er*, Cvi, NIL26 and the reciprocal crosses between NIL26 and L*er*.

recombinants between the two markers showing a recombination frequency of 1.8 cM. Seeds of these recombinants were tested for InsP6 and Pi levels and this confirmed that the InsP6/Pi locus was located in between these two markers. For mapping nine molecular markers (Table 7.1) between T4P13b and 17D8LE (which were tested for L*er*/Cvi polymorphisms) were generated from the Columbia sequence. All 23 selected recombinants have been analysed with these markers. Hereby the QTL could be mapped to a region of 99 Kb, localizing the QTL on BAC T13O15 between marker cT4P13x and cT13O15c, which have respectively 1 and 3 recombinants with the InsP6/Pi locus as is shown in figure 7.8. This region contains 13 open reading frames. Based on the annotation of the Columbia sequence this region includes a putative peptide transporter (gene code: At3g01350), a putative protein transport protein SEC13 (At3g01340), a putative protein kinase (At3g01300), a putative phosphatidylinositolglycan class N short form (At3g01380), a vacuolar membrane ATPase subunit G1 (At3g01390), a putative RNase H (At3g01410) and 7 unknown proteins.

Table 7.1. Molecular markers polymorphic between L*er* and Cvi. The order of the markers is based on the position in the Columbia sequence

Marker	Primer sequence		Marker	Annealing	Restriction	Position in genome	
	Forward	Reverse	type	temp. (°C)	enzyme	BAC	Position in BAC
T4P13b	aacgataatcttaggctcggtcac	cccacgattacagttcctcatagac	CAPS	55	NdeI	T4P13	70213-71598
cT4P13z	gactgaatgttctgtgtgaggc	ggacggtcaagcttaacacc	CAPS	52	HgaI	T4P13	24271-25968
cT4P13x	aaccacccaaatcaaagcctac	tggattggctctcgcactc	CAPS	52	AluI	T4P13	8185-8580
						T22N4	49831-51226
cT22N4w	tgagagcgagaaaggccgagag	ggcaagctcaccattgacaggc	CAPS	52	ScrfI	T22N4	38559-37185
						T13O15	3094-4468
cT22N4a	ctggatcaacatgctcatgg	catcttgagaaacaacaatcgc	CAPS	50	DraI	T22N4	10277-12037
						T13O15	29616-31376
cT13O15b	aggagactggaaggaatccgac	aatcctcggttggttaggcac	CAPS	52	HindIII	T13O15	57449-59252
cT13O15c	tccacaggctgctttgaaacg	cggttaagttcagggagcgttg	CAPS	55	HaeIII	T13O15	84297-85714
						F4P13	3446-4863
cF4P13a	caaggacaatcccaacatcgag	ggtgctaaacgggagtccctac	CAPS	55	HaeIII	F4P13	11837-13531
CA1[a]	aaccgtctttagctcttccttctcg	ggcttctttctcacttcactttctcc	CAPS	50	DdeI	F4P13	18472-19848
nga32[a]	ccaaaacaattagctcccca	ggagactttttgagattggcc	SSLP	55	-	F1C9	35661-35401
17D8LE[a]	ccaacaacatgcatgatagttcag	ctcctttgtcatctcccgaatc	CAPS	55	RsaI	F13E7	94303-96342
nga172[a]	catccgaatgccattgttc	agctgcttccttatagcgtcc	SSLP	50	-	T21P5	73652-73817

[a] Obtained from The Arabidopsis Information Resource (www.arabidopsis.org)
CAPS (Cleaved Amplified Polymorphic Sequence)
SSLP (Single Sequence Length Polymorphism)

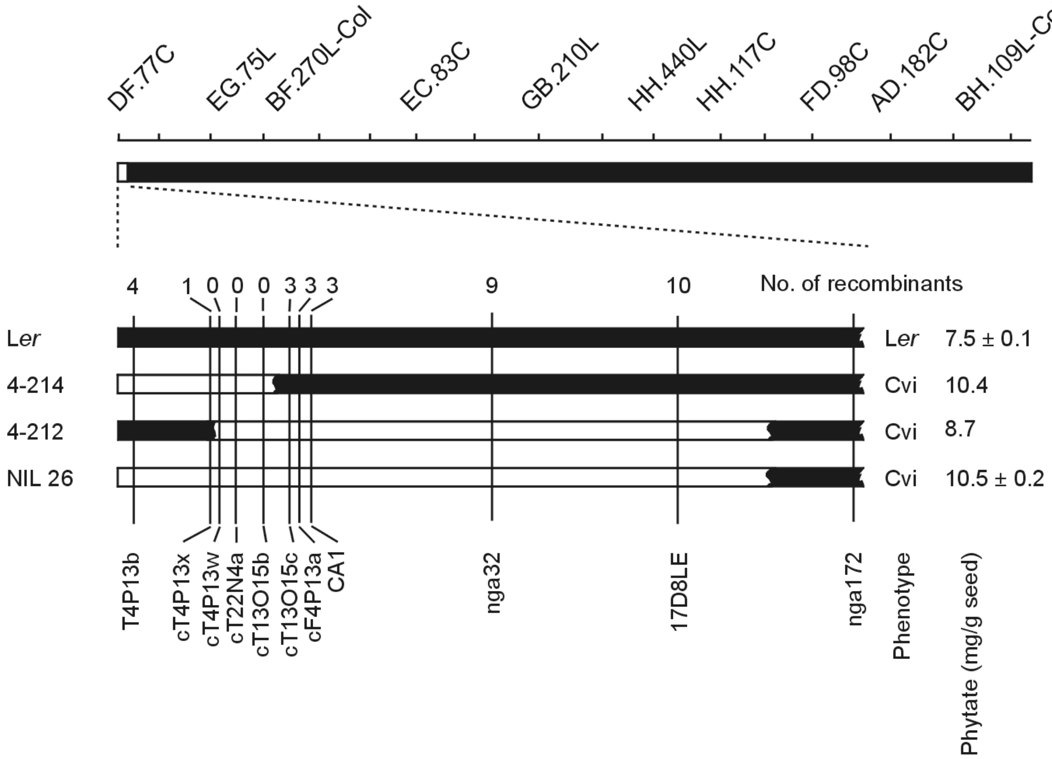


Figure 7.8. Genetic fine mapping of the major InsP6/Pi QTL. The top of the figure shows the genetic map with AFLP markers (Alonso-Blanco et al., 1998b) of chromosome 3, solid bars indicate L*er* DNA, open bars the Cvi alleles. Junctions between open and solid bars indicate crossover breakpoints. The lower part shows the QTL region in detail, essential recombinants which indicate the position of the InsP6/Pi QTL are indicated, with their average InsP6 levels (mg/g seed). Molecular markers with the number of recombinants between the marker and the InsP6/Pi locus are indicated as well. Number of recombinants with the markers nga32 and 17D8LE are underestimated, as the InsP6 levels of heterozygous plants could not be distinguished from Cvi plants. Recombinants with their crossover breakpoint between the other markers have been analysed for InsP6 by determining the seeds of 12 plants derived from these recombinants, to distinguish heterozygous plants from Cvi plants.

Discussion

We present the analysis of natural variation for InsP6 and Pi in *Arabidopsis thaliana*. To be able to analyse large series of samples we used a simple and rapid detection method, based on anion exchange chromatography and detection through ion-suppressed conductivity. This method has been used before for the analysis of steep water of corn (Hull and Montgomery, 1994), but not for plant material. The method is very sensitive, able to detect InsP6 in sub-milligram samples of seeds, with recoveries >90%, provided the samples size does not exceed 4 mg (Figure 7.2).

Total levels of P (InsP6 and Pi) in Arabidopsis seeds are higher compared with levels reported for maize, barley and rice (Raboy et al., 2001). However, Coello et al. (2001) reported around two times higher levels in seeds of accession Columbia. This might be due to differences in growth conditions, since Lickfett et al. (1999) reported that soil-P levels largely affect InsP6 levels in Brassica seeds. Besides P-fertilization, InsP6 and Pi concentrations in seeds may vary because of many factors including moisture plus climatic factors (Lott et al., 2000). We also found that genetically identical seed batches, harvested at different times, could vary in Pi and Insp6 contents. To avoid these environmental effects we have only compared plants that were grown under identical environmental conditions in same experiment.

A wide range of variation in InsP6 and Pi levels was observed amongst accessions, varying from Kyoto-1, containing 7.0 mg of InsP6 per gram seed, to Hey-0 with 23.1 mg InsP6 per gram seed. We also observed the typically high and positive correlation of InsP6 with total Pi as has been reported by Raboy (1990 and 1997), but there are accessions as well that do have altered ratios of InsP6 compared to the Pi levels, indicating that total P accumulation is affected. Altered ratios of InsP6 compared to total P were also described in mutants, isolated in maize (*Zea Mays* L.), barley (*Hordeum vulgare* L.) and rice (*Oryza sativa* L.) (Larson et al., 1998, 2000, Raboy and Gerbasi, 1996 and Raboy et al., 2000). The observation that total P in mutant seeds remained has been explained as an altered biosynthesis of InsP6, rather than differences in uptake from the environment and/or partitioning of P within the plant.

The presence of different types of natural variation among the Arabidopsis accessions provides unique genetic sources to study the genetics and molecular biology of InsP6 levels in plants. Therefore, we performed a QTL mapping procedure to study the natural variation for InsP6 and Pi between the accessions L*er* and Cvi. QTL mapping revealed five genomic regions affecting the quantity of the InsP6 and Pi in seeds and leaves. One of them, located on top of chromosome 3 is affecting all four traits. This QTL appears as the major locus responsible for the observed variation in InsP6 and Pi contents in the L*er*/Cvi RIL population, the Cvi allele is increasing the content of both InsP6 and Pi in seeds and in leaves. To Mendelize this major InsP6/Pi

QTL we generated a near-isogenic line (NIL26) carrying a 3 cM Cvi genomic region on the top of chromosome 3 in a L*er* background. Seeds of this line have higher InsP6 levels than seeds of L*er* itself, confirming that indeed the Cvi allele of the QTL is increasing the InsP6/Pi levels. The F2 of a cross between L*er* and NIL26 showed a monogenic segregation and revealed the dominance of the Cvi allele at this InsP6/Pi locus. Reciprocal crosses show maternal inheritance. This might indicate that the difference in InsP6 between L*er* and Cvi is caused by transport of P from the mother plant into the seed. This is in agreement with the finding that the InsP6 levels compared to the total P levels are not altered, indicating that there is no alteration in the biosynthesis of InsP6 controlled by this locus.

Possible candidate genes for the InsP6/Pi locus are genes, which have a function in P translocation. There are four Arabidopsis mutants known to be affected in P accumulation/mobilization. This are *pho1* which is exhibiting a reduced P-uptake (Poirier et al., 1991), *pho2* a P-accumulator (Delhaize and Randall, 1995) *pho3* a P deficient mutant probably affected in P uptake by the roots (Zakhleniuk et al., 2001) and *ppt1* which is mutated in phosphate/phosphoenolpyruvate translocator gene (formerly named *cue1-1*, and allelic to CS3156; Li et al., 1995). However, except *pho3* (which has not been mapped) none of these mutants are candidate genes for the major InsP6/Pi QTL, as they are localised at different genetic positions. *PHO1* has recently been cloned and is localised on chromosome 3, but at a lower position than the InsP6/Pi QTL (Hamburger et al., 2002). *PHO1* belongs to a gene family but none of the *PHO1* homologs map to the same chromosome region as the InsP6/Pi QTL. *PHO2* has been mapped to the bottom of chromosome 2, near marker m429 (Delhaize and Randall, 1995) and might be a candidate gene for the QTL at the bottom of chromosome 2 that affects InsP6 in leaves and Pi in seeds. *PPT1* has been cloned and is located on the middle of chromosome 5 at BAC F19N2 (Streatfield et al., 1999).

Fine mapping of the major QTL using molecular markers localised the InsP6/Pi locus to a 99 Kb region containing 13 open reading frames (ORFs). We consider the vacuolar membrane ATPase subunit G, as the most likely candidate gene for the InsP6/Pi locus, among the annotated genes in the 99 Kb region of interest. As suggested above, the InsP6/Pi locus might be involved in P-translocation in the plant, and P-transport across membranes depends on the membrane potential, which is generated by ATPases. However, we cannot exclude one of the other genes to play a role in InsP6/Pi accumulation, since only complementation analysis can proof that a candidate gene is responsible for the QTL allelic variation (El-Assal et al., 2001).

Using immortal populations has a great advantage because the same set of plants can be used to analyse different traits that might influence the same processes. Hitz et al. (2002) reported about the biochemical and molecular characterisation of a mutation that confers a decreased raffinose oligosaccharide and InsP6 phenotype on soybean seeds. This can be explained by the fact that InsP6 and raffinose oligosaccharides share at least *myo*-inositol-1-phosphate and possibly free *myo*-inositol as a common intermediate (Loewes and Murthy, 2000). Since the L*er*/Cvi RIL population has also been analysed for oligosaccharide contents and seed storability (Bentsink et al., 2000) we can check whether QTL affecting these traits co-localize with the QTL identified for InsP6/Pi. This showed that there is no co-localization between the InsP6/Pi loci and the oligosaccharide QTL.

However the InsP6/Pi locus on top of chromosome 3 co-localizes with QTL for sucrose and seed storability. The Cvi allele of this locus is increasing InsP6 levels, sucrose content and seed storability, indicating that higher InsP6 levels on their own or with higher sucrose levels might result in a better storability. A role of InsP6 in seed quality is suggested by the finding that reducing the level of InsP6 with more than 95% results in loss of viability (Raboy et al., 2001). Except for the major InsP6/Pi QTL we see co-location of the locus affecting Pi in seeds (chromosome 1) with the major seed storability QTL, decreasing Pi levels correlate with higher storability. Any functional relation between these traits has not been reported and co-location may of course also involve different linked genes.

Another advantage of the L*er*/Cvi RIL population is that we can study the link between InsP6 levels and seed size or seed weight. It has been reported that reductions in seed InsP6 have severe impacts on seed and plant growth and function, translating in more severe yield losses (reviewed by Raboy, 2001). Alonso-Blanco et al. (1999) studied allelic variation at seed size loci in relation to other life history traits using L*er*/Cvi RIL. This analysis revealed eleven loci that affected seed weight and / or seed length. One of them is located on top of chromosome 3 near the locus where the main InsP6/Pi QTL is located, however this is a QTL with a minor effect and is located below the InsP6/Pi QTL. The effects of these traits are negatively correlated, suggesting that higher InsP6 levels would result in lower seed weight. This is in contrast with results described earlier (Raboy, 2001). However, we find correlation between the major QTL influencing seed size (located on top of chromosome 1) and a QTL affecting Pi levels in seeds. This correlation is much stronger than the one mentioned before, lower Pi levels correlate positively with higher seed weights.

Concluding, there is considerable variation among natural accessions of Arabidopsis that can be used to identify genes that are involved in the regulation of InsP6 accumulation. The use of other accessions may identify additional loci affecting this trait.

Acknowledgements

The authors thank Diaan Jamar for technical assistance in the InsP6 and Pi analyses. Y.K. received a grant from the China Scholarship Council.

Material and Methods

Plant material

A set of 101 Arabidopsis accessions was analysed for InsP6 and Pi content in seeds and for Pi content in leaves. The complete list of accessions and the levels of InsP6 and Pi are presented at www.natural-eu.org.

A set of 162 RILs derived from crosses between the laboratory strain L*er* originated from Northern Europe (Rédei, 1992) and the accession Cvi, from the Cape Verde Islands (Lobin, 1983) were analysed for InsP6 and Pi in seeds and leaves as well. These RILs have been previously characterised for amplified fragment length polymorphism (AFLP) and cleaved amplified polymorphic sequence (CAPS) markers (Alonso-Blanco et al., 1998b). Results of InsP6 and Pi analysis of all individual RILs are as well presented at www.natural-eu.org.

To construct a near-isogenic line containing a Cvi introgression at the top of chromosome 3 we have used NIL35. NIL35 was constructed by two back crosses of RIL49 with L*er* for fine mapping seed dormancy QTL (Chapter 4) and contained 3cM of Cvi at the top of chromosome 3, 45 cM on the middle of chromosome 5 and a chromosome segment of 10 cM on the bottom chromosome 1, in a L*er* background, as was determined with AFLP markers. NIL35 has been backcrossed with L*er*. In the resulting F2 of this population a plant heterozygous for the top of chromosome 3 and homozygous L*er* at chromosome 5 was selected, using molecular markers. The progeny of this plant has been used as a mapping population to refine the map position of the QTL on top of chromosome 3.

A line homozygous for Cvi on the top of chromosome 3 has also been selected from this population and is called NIL26.

Plants were grown in an air-conditioned greenhouse (temp 22-25°C), supplemented with additional light (model TDL 58W/84, Philips, Eindhoven The Netherlands) from mid September till early April, providing a day length of 16 h. Plants were grown in 7-cm pots in standard soil.

Extraction and quantification of InsP6 and Pi

Dry seeds (1-4 mg) were boiled for 15-20 min in 0.5 mL 0.6 M HCl. Leaves (30-60 mg) were extracted by boiling for 15 min in 1 mL 0.6 M HCL, 10 mM EDTA. The extracts were centrifuged at 15,000 g for 5 min. The supernatants were diluted 10 times (seeds) or 3 times (leaves) with water, and 20 µL was analysed using a Dionex DX300 HPLC system (Dionex Corp., Sunnyvale, CA, USA). Anions were separated on an AS11 (4 x 250 mm) column, preceded by an AG 11 guard column and eluted with NaOH. The elution profile was: 5 min isocratic at 5 mM NaOH, followed by a 15 min linear gradient 5 – 100 mM NaOH. After each run the column was washed for 5 min with 0.5 M NaOH, followed by a 15 min equilibration at 5 mM. Flow rates were 1 $mL.min^{-1}$ throughout the run. Contaminating anions in the eluents were removed using an ion trap column (ATC-1), installed between pump and sample injection valve. Anions were determined by conductivity detection. Background conductivity was decreased using a ASRS suppressor, with water as counterflow (5 $mL.min^{-1}$), operated at 300 mA, controlled by an SRS controller (Dionex Corp., Sunnyvale, CA, USA). Peaks were identified and quantified by co-elution with known standards. Recoveries of InsP6 were determined by spiking seed samples with Na_{12}-InsP6.

QTL analysis

To map QTL using the RIL population, a set of 99 markers covering most of the Arabidopsis genetic map was selected from the previously published RIL L*er*/Cvi map (Alonso-Blanco et al., 1998b). These markers spanned 482 cM, with an average distance between consecutive markers of 5 cM and the largest genetic distance being 12 cM.

The computer program MapQTL™ version 4.0 (Plant Research International, Wageningen-University and Research Centre, Wageningen, The Netherlands) was used to identify and locate QTL linked to the molecular markers using both interval mapping and multiple-QTL model mapping (MQM) methods as described in its reference manual (http://www.plant.wageningen-ur.nl/products/mapping/mapqtl/). The estimated additive effect and the percentage of variance explained by each QTL as well as the total variance explained by all of the QTL affecting a trait, were obtained with MapQTL in the final MQM model. For this, different cofactor markers were tested around the putative QTL positions (van Ooijen and Maliepaard, 1996), selecting as final cofactors the closest marker to each QTL, i.e. those maximizing the logarithm-of-odds (LOD) score. A LOD score threshold of 2.8 was applied to declare the presence of a QTL, which corresponds to a general genome-wide significance $P = 0.05$ for normally distributed data, as was determined by extensive simulation experiments (van Ooijen, 1999). We verified this threshold for interval mapping by applying the permutation test to each data set (10 000 repetitions) and found $P = 0.05$ LOD thresholds between 2.5 and 2.6 for all traits. Two-LOD support intervals were established as ≈ 95% confidence intervals (van Ooijen, 1992).

For every trait, two-way QTL interactions were analysed by analysis of variance at a significance level of $P < 0.005$, using the General Linear Model module of the statistical package SPSS version 10.0 (SPSS Inc., Chicago). For each analysis, the closest linked markers to the corresponding detected QTL were used as random factors in the ANOVA (the same markers used as cofactors in the MQM mapping with MapQTL).

Fine mapping of the InsP6 and Pi locus

To fine map the major locus for InsP6 and Pi locus on chromosome 3, DNA of 631 plants derived from the heterozygous NIL26 sister plant was isolated and analysed using the CAPS markers T4P13b and 17D8LE (Table 7.1). The seeds of recombinants between these two markers were analysed for InsP6 and Pi content and used for further fine-mapping of the locus. To unambiguously determine the InsP6 phenotype of the recombinants, 12 F3 plants of each recombinant have been analysed sub sequentially.

Design of molecular markers

Based on the Columbia genomic sequence (TAIR; www.arabidopsis.org) new Cleaved Amplified Polymorphic Sequence (CAPS) (Konieczny and Ausubel, 1993) and Single Sequence Length Polymorpism (SSLP) (Bell and Ecker, 1994) markers were developed. Marker specifications are listed in Table 7.1.

Chapter 8

Summarising discussion

The existing genetic variation among and within naturally occurring populations of Arabidopsis, collected from different geographical regions (Rédei, 1970; Alonso-Blanco and Koornneef, 2000) provides an alternative source of genetic variation that can be used to study the function of genes. The accessions Landsberg *erecta* (L*er*) and Cape Verde Islands (Cvi) exhibited natural genetic variation for all traits described in this thesis. This variation was analysed using a recombinant inbred line (RIL) population previously developed and genotyped with amplified fragment length polymorphism (AFLP) (Alonso-Blanco et al., 1998b) and other PCR based markers.

The main aim of the work presented in this thesis is to increase the knowledge about the genetic control of seed dormancy. For that, we have analysed natural variation for this character but also for other seed traits that might be related to seed dormancy and germination. These traits are seed oligosaccharide content, seed storability and seed phytate and phosphate content.

Seed dormancy, defined as the failure of an intact, viable seed to complete germination under favourable conditions (Bewley, 1997), is a complex trait under the control of a large number of genes. The use of natural variation might reveal novel seed dormancy loci in addition to the loci that have been identified by mutagenesis and are described in Chapter 2. The accessions L*er* and Cvi provided the genetic variation that was needed for this study. Cvi is a very dormant accession and requires at least 3 months of after-ripening before it germinates 100%. L*er*, one of the standard laboratory accessions, possesses a moderate dormancy and needs approximately one month of after-ripening to fully germinate. The time of after-ripening required for reaching 100% germination is very much influenced by the environmental conditions under which the mother plant is grown and genetically identical seeds harvested on plants grown at different moments can therefore differ significantly. However, analysis on seeds harvested on plants grown within the same experiment, stored and tested under identical conditions allows the study of the genetic component of this trait.

Quantitative trait loci (QTL) mapping using the RIL population derived from a cross between L*er* and Cvi revealed the loci affecting the after-ripening requirement measured as the number of days of seed dry storage to reach 50% germination ($DSDS_{50}$) and is described in Chapter 3. A total of 7 loci were found to significantly affect this trait and these loci were named *delay of germination* (*DOG*) *1* to *7*. On average in the RIL population, Cvi alleles at six loci increased the time of seed storage required for seed germination (increased dormancy) and only Cvi alleles at *DOG2*, located on the top of chromosome 1, reduced dormancy as compared with the L*er* allele. To confirm and further characterize these loci, we developed 12 near isogenic lines (NILs) carrying single and double Cvi introgression fragments in an otherwise L*er* genetic background. The analysis of these lines for their seed germination in water confirmed four QTL, (*DOG1*, *DOG2*, *DOG3* and *DOG6*) that show large additive effects in a L*er* background. In addition, it was found that *DOG1* and *DOG3* genetically interact, the strong dormancy determined by *DOG1*-Cvi alleles depending on *DOG3*-L*er* alleles. L*er*, Cvi and these NILs were further characterised for their seed dormancy/germination behaviour in five different environments; this revealed that cold treatment (stratification) is the most effective manner to break the dormancy of the different QTL.

The *DOG1* locus was the strongest QTL identified (Chapter 3) and the genetic fine mapping of this locus is described in Chapter 4. NIL DOG17-1, which contains a Cvi introgression fragment around the position of the *DOG1* QTL, was crossed with a marker line, that is homozygous for *glabrous 3* (*gl3*) and *thiamine* (*tz*) in a L*er* background and the resulting hybrid was backcrossed with the latter line resulting in a test cross population. Recombinants between *gl3* and *tz* were selected and a subset of these recombinants has been analysed with molecular markers located in the Cvi introgression. Thus, *DOG1* could be located within a region of 75 kb, containing 22 open reading frames. This region contains no gene previously involved in seed dormancy or germination and therefore no obvious candidate gene has been identified. Additional fine mapping and subsequent cloning will reveal the function of *DOG1*.

The characterisation of the natural L*er* en Cvi alleles of *DOG1* (*DOG1*-L*er* and *DOG1*-Cvi) and the isolation and characterisation of a completely non-dormant mutant (*dog1*) in the *DOG1* locus is described in Chapter 5. The *dog1* mutant germinates 100% when freshly harvested, whereas *DOG1*-L*er* and *DOG1*-Cvi require respectively 50 days and 150 days of after-ripening to reach 100% of germination. *DOG1* alleles behave as semi-dominant and the dormancy of *DOG1* appears mainly controlled by the genotype of the embryo. All *DOG1* alleles require de novo gibberellin (GA) synthesis for germination. However, the level of GA required is determined by the *DOG1* allele. This could be investigated by analysing the GA requirement of *DOG1*-L*er* and *DOG1*-Cvi in the genetic background of the GA deficient *ga1-3* mutant. Furthermore, *DOG1*-L*er* and *DOG1*-Cvi alleles were combined in the abscisic acid (ABA) deficient and insensitive backgrounds determined by the mutants *aba1-1* and *abi3-5* respectively, revealing that the dormancy of *DOG1* alleles is dependent on ABA. Since the interaction of *DOG1* with GAs is different than the interaction of ABA and GA, we suggest that the *DOG1* locus is involved in the setting of dormancy during seed development, whereas ABA controls dormancy both during development and by inhibiting the germination promotive effect of GAs during imbibition.

Chapter 3 to 5 show how natural variation can be exploited to unravel quantitative traits, and that single major QTL for seed dormancy have strong enough effects to allow the analysis of these QTL as a single locus. Recently, we have also started the fine mapping of *DOG2*, *DOG3* and *DOG6* and the latter has already fine mapped to a 90 kb region. Like for *DOG1*, this 90 kb region containing the *DOG6* locus does not carry genes that have been previously involved in seed dormancy. In addition to the fine mapping of these *DOG* loci we are also analysing the relationship between *DOG3, DOG6* and the different dormancy pathways identified by mutants affected in seed dormancy/germination in a similar way to that described for the study of *DOG1*. For that, we have introduced the *DOG3* and *DOG6* Cvi introgressions fragments into the background of several mutants, and the seeds of these combinations have been analysed for their germination behaviour. Except for the *aba1-1* and *abi3-5* mutants that are epistatic to both the L*er* and Cvi alleles, we could not detect other interactions, meaning that *DOG3* and *DOG6* are dependent on ABA but control dormancy independent from the known pathways.

Like induced mutants, NILs containing introgressions fragments carrying strong dormancy alleles can be very useful for whole transcriptome and proteome analysis, which may allow the identification of the proteins and genes that are induced or repressed in relation to seed dormancy and germination. An example of such an approach is the proteome analysis performed by Gallardo et al. (2002) using the *ga1-1* mutant to reveal the role of GAs in germination. It is expected that we will use the NILs, generated for the work that is described in this thesis for such analysis in the near future.

We conclude that natural variation is very useful to identify novel seed dormancy loci, and the exploitation of genetic variation from other accessions might provide additional genetic variation for this trait. This has been shown before by QTL mapping for seed dormancy using the L*er*/Col RIL population (van der Schaar et al., 1997). Some of the QTL localised in this population locate at similar positions as the loci that are identified using the L*er*/Cvi RILs, but there are also unique QTL found in the Ler/Col population (Chapter 2 and 3). Currently we are analysing two new RIL populations derived from the crosses L*er* x Shahdara and L*er* x Kashmir. Furthermore, we are using another very dormant accession, Enkheim (En2) to construct NILs that contain En2 introgressions, which increase the level of dormancy in a L*er* background. To construct these dormant NILs, F3 lines of a cross between L*er* and En2 have been backcrossed with L*er* during several generations and in every generation the most dormant lines were selected. The dormant NILs obtained in this way will be characterised for seed dormancy as performed with NIL DOG17-1.

Chapter 6 describes the characterisation of raffinose series oligosaccharides (RSO) in relation to seed storability. Previous research performed by K. Léon-Kloosterziel (personal communication) indicated that apart from the large difference in dormancy, L*er* and Cvi were very different in seed oligosaccharide (OS) composition. L*er* contains higher levels of RSOs and lower levels of sucrose compared to Cvi, which was almost devoid of RSOs. RSOs, which are seed storage reserves, accumulate within seeds during the seed maturation (Chapter 1), more or less at the same moment as seed dormancy is induced. To determine whether these processes are controlled by the same loci, we performed QTL mapping for OS contents in seeds of the L*er*/Cvi RIL population. We found one major QTL responsible for the practically monogenic segregation of seed stachyose content. This locus also affected the content of the two other OSs, sucrose and raffinose, and was mapped on chromosome 1.

Although sugars are thought to be involved in dormancy/germination processes, at the map position of the major OS QTL we did not locate any seed dormancy QTL (Figure 8.1). The role of sugars in germination is supported by the observation that the *abi3-5* mutant which is non-dormant accumulates 3-6 fold more sucrose during seed development than wild type (Ooms et al., 1993) and by the fact that exogenous sugars are able to relieve the inhibition of seed germination imposed by ABA (Finkelstein and Lynch, 2000; Garciarrubio et al., 1997). Whether this effect is due to sugar acting as a

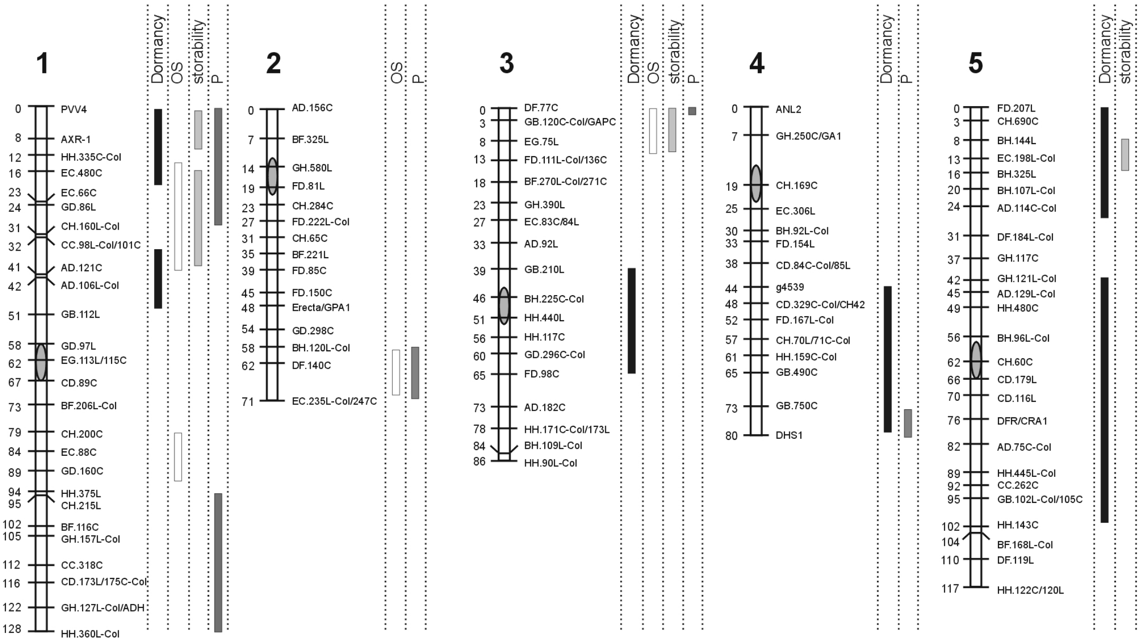


Figure 8.1. Ler/Cvi linkage map showing the genetic location of QTL affecting seed dormancy, seed OS content, seed storability and InsP6 and phosphate content in seeds and leaves (P). The length of the bars depict the QTL 2 LOD support intervals.

nutrient or as a signalling molecule remains unclear. The sugar required for germination is not provided by the breakdown of lipid reserves. Recent studies have shown that the glyoxylate cycle is not essential for germination but is important for seedling establishment and survival (Eastmond et al, 2000). In addition, mature seeds exposed to ABA express genes encoding enzymes that are required for lipid breakdown (Pritchard et al., in press) whereas they do not show visible signs of germination.

In contrast to the promotive effects of low concentrations of sugar in the presence of ABA, higher concentrations of sugar can retard germination and subsequent seedling growth. These inhibitory effects on growth and sugar related gene expression have been the basis of numerous screens for sugar-response mutants, which have identified new alleles of *ABI4* and two distinct ABA biosynthetic loci (reviewed by Finkelstein and Gibson, 2001). The observation that *abi4/glucose insensitive6/sugar insensitive5/sucrose uncoupled6/impaired sucrose induction3* mutants are sugar resistant could be consistent with either cross talk between sugar and ABA signalling or action in parallel pathways.

Seed OSs and especially RSOs are hypothesised to play an important role in the acquisition of desiccation tolerance and consequently in seed storability (Horbowicz and Obendorf, 1994; see Chapter 1). Seed storability was analysed in the same RIL population by measuring viability (germination) under two different seed aging assays: after natural aging during 4 years of dry storage at room temperature and after artificial aging induced by a controlled deterioration test (Chapter 6). Thus, four QTL responsible for the variation for this trait were mapped. Comparison of the QTL genetic positions showed that the genomic regions containing the major OS locus did not significantly affect the seed storability. We concluded that in the studied material RSO content had no specific effect on seed storability.

Seed storability and dormancy QTL co-located at chromosome 1 and at the top of chromosome 5 (Figure 8.1). It is likely that some of these QTL are the same gene because the characterisation of both traits is based on germination assays. Furthermore, there are mutants that exhibit dormancy and seed storability phenotypes, like strong alleles at the *ABI3* locus (for more details see Chapter 2). To test whether the region on chromosome 1 indeed affects these two traits we will determine if NILs that contain different Cvi introgression fragments at the position of the dormancy QTL also exhibit an altered storability phenotype.

Chapter 7 describes the genetic characterisation of phytate and phosphate (Pi) in seeds and in leaves. Phytate (myo-inositol-1,2,3,4,5,6- hexakisphosphate, InsP6) is the most abundant P-containing compound in plants, and an important anti-nutritional factor, due to its ability to complex essential micro-nutrients, e.g., iron and zinc (Raboy, 2001). QTL analysis of InsP6 and Pi levels in seeds and leaves, in the L*er*/Cvi RIL population was performed to identify gene(s) that are involved in the regulation of InsP6 accumulation. Five genomic regions affecting the quantity of the InsP6 and Pi in seeds and leaves were identified. One of them, located on top of chromosome 3, affects all four traits. This QTL appears as the major locus responsible for the observed variation in InsP6 and Pi contents in the L*er/*Cvi RIL population, the L*er* allele decreasing the content of both InsP6 and Pi in seeds and in leaves. The InsP6/Pi locus was further fine-mapped to a 99 kb region containing 13 open reading frames. The maternal inheritance of the QTL and the positive correlation between InsP6 and total Pi levels both in seeds and in leaves indicate that the difference in InsP6 level between L*er* and Cvi is likely to be caused by a difference in transport rather than by an alteration in the biosynthesis. Therefore, we consider the vacuolar membrane ATPase subunit G, located in the region of interest, as the most likely candidate gene for the InsP6/Pi locus.

Despite that InsP6 and RSO contents might be correlated because they share *myo*-inositol-1-phosphate and possibly free *myo*-inositol as a common intermediate (Hitz et al, 2002), we did not find co-location of InsP6 and RSOs. However, the InsP6/Pi locus at the top of chromosome 3 co-localised with a QTL affecting sucrose and seed storability (Figure 8.1). To determine whether these characters are controlled by the same locus or by two closely linked loci we will investigate seed OS content and seed storability of the different recombinants that were identified during the fine mapping of the InsP6/P QTL (Figure 7.8).

Possible pleiotropism of InsP6/Pi and seed dormancy QTL has been discussed in literature. Caryopses of dormant wild oat lines characteristically produce lower Pi levels than that of less dormant lines (Quick et al., 1997). InsP6 might be involved in this because phytases start to degrade the InsP6 reserves during after-ripening, which releases Pi in germinating caryopsis. The accumulating Pi might spark the ATP pump (Perl, 1986) so that enzymes associated with continued embryo development might be synthesised or activated. Furthermore, it has been shown that exogenous Pi is able to mimic the specific action of GA_3 in regulating the activity of certain enzymes in embryoless half seeds of wheat (Saluja et al., 1987). The ability of Pi to mimic GA activity may provide another mechanism whereby Pi influences the dormancy level. Foley et al. (1992; 1993) suggested that GA promoted the synthesis or the activation of enzymes involved in embryo or endosperm carbohydrate metabolism, thus providing sugar for breaking dormancy. In the L*er*/Cvi RIL population we only find co-location between dormancy loci and loci that affect seed Insp6/Pi content at two positions, top chromosome 1 and bottom chromosome 4. As described above for other QTL co-locations, whether this is due to pleiotropy or to linked loci has to be investigated by analysis of the various traits in recombinants in these regions. This can be done using for example different RILs that have crossover events within this region.

We have shown that the L*er*/Cvi RIL population segregates for many traits. Even when the parental lines do not differ for certain trait the RILs still might segregate for several loci, like has been described for flowering time (Alonso-Blanco et al., 1998a). We have illustrated how immortal populations have a great advantage because the same set of plants can be used to analyse different traits that might influence the same processes. However due to the limitated accuracy of QTL mapping, co-location of QTL for different traits does not prove pleiotropy and common regulation by a single locus as the traits can also be regulated by closely linked genes, which cannot be separated by QTL mapping. Additional fine mapping, i.e. the genetic and phenotypic characterisation of selected recombinants for the different traits, is needed to distinguish pleiotropy from close genetic linkage. In this respect, NILs are very useful because they show monogenic QTL differences as shown for seed dormancy and InsP6 (this thesis). Thus, they allow to develop monogenic fine mapping populations to test co-location of QTL for different traits and can be used to further characterise physiologically and genetically the locus of interest including genomic studies. Complete sets of NILs whose introgressions cover the whole genome of the donor parent have already proven to be very useful in several crop species for QTL mapping (Eshed and Zamir, 1995; Ramsay et al., 1996) and one of our future goals is to construct such a complete set of L*er*/Cvi NILs in a L*er* background. Thus the available resources could be further improved and the exploitation of natural variation illustrated in this thesis might be performed in a faster way.

References

Albert S, Delseny M and Devic M (1997). *BANYULS*, a novel negative regulator of flavonoid biosynthesis in the Arabidopsis seed coat. Plant J 11: 289-299.

Alonso J, Hirayama T, Roman G, Nourizadeh S and Ecker JR (1999). *EIN2*, a bifunctional transducer of ethylene and stress responses in Arabidopsis. Science 284: 2148- 2155.

Alonso-Blanco C and Koornneef M (2000). Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. Trends Plant Sci 5: 22-29.

Alonso-Blanco C, El-Assal SE-D, Coupland G, and Koornneef M (1998a). Analysis of natural allelic variation at flowering time loci in the Landsberg *erecta* and Cape Verde Islands ecotypes of *Arabidopsis thaliana*. Genetics 149: 749-764.

Alonso-Blanco C, Blankenstijn-de-Vries H, Hanhart CJ, and Koornneef M (1999). Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*. Proc Natl Acad Sci USA. 96: 4710-4717.

Alonso-Blanco C, Peeters AJM, Koornneef M, Lister C, Dean C, van den Bosch N, Pot J, and Kuiper MTR (1998b). 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. Plant J 14: 259-271.

Anderson JA, Sorrells ME and Tanksley SD (1993). RFLP analysis of genomic regions associated with resistance to preharvest sprouting in wheat. Crop Sci. 33: 453-459.

Arabidopsis Genome Initiative (2000). Analysis of the gene sequence of the flowering plant *Arabidopsis thaliana*. Nature. 408: 796-815.

Baskin CC and JM Baskin (1998). Seeds: ecology, biogeography, and evolution of dormancy and germination. (Academic Press, San Diego).

Baud S, Boutin J-P, Miquel M, Lepiniec L and Rochat C (2002). An intergrated overview of seed development in *Arabidopsis thaliana* ecotype WS. Plant Physiol Biochem 40: 151-160.

Bäumlein H, Miséra S, Luerssen H, Kölle K, Horstmann C, Wobus U and Müller, AJ, (1994) The *FUS3* gene of *Arabidopsis thaliana* is a regulator of gene expression during late embryogenesis. Plant J 6: 379-387.

Beaudoin N, Serizet C, Gosti F and Giraudat J (2000). Interactions between abscisic acid and ethylene signalling cascades. Plant Cell 12: 1103-1115.

Beebe DU and Turgeon R (1992). Localization of galactinol, raffinose, and stachyose synthesis in *Cucurbita pepo* leaves. Planta 188: 354-361.

Beeckman T, De Rycke R, Viane R and Inzé D (2000). Histological study of seed coat development in *Arabidopsis thaliana*. J Plant Res 113: 139-148.

Bell CJ and Ecker JE (1994). Assignment of 30 microsatellite loci to the linkage map of Arabidopsis. Genomics 19: 137-144.

Bentsink L and Koornneef M (2002). Seed dormancy and germination. In The Arabidopsis Book, CR Somerville and EM Meyerowitz, eds (Rockville, MD: American Society of Plant Biologists) DOI 10.1199 tab.0050.

Bentsink L, Alonso-Blanco C, Vreugdenhil D, Tesnier KJY, Groot SPC and Koornneef M (2000). Genetic analysis of seed-soluble oligosaccharides in relation to seed storability of Arabidopsis. Plant Physiol 124: 1595-1604.

Bewley JD (1997). Seed Germination and Dormancy. Plant Cell 9: 1055-1066.

Bewley JD and Black M (1994). Seeds; Physiology of development and germination. (New York: Plenum Press).

Bies N, Da Silva Conceicao A, Giraudat J, Koornneef M, Léon-Kloosterziel KM, Valon C and Delseny M (1999). Importance of the B2 domain of the Arabidopsis ABI3 protein for EM and 2S albumin gene regulation. Plant Mol Biol 40: 1045-1054.

Black M, Corbineau F, Gee H and Côme D (1999). Water content, raffinose and dehydrins in the induction of desiccation tolerance in immature wheat embryos. Plant Physiol 120, 463-471.

Bleecker AB, Estelle MA, Somerville C and Kende H (1988). Insensitivity to ethylene conferred by a dominant mutation in *Arabidopsis thaliana*. Science 241: 1086-1089.

Bonetta D and McCourt P (1998). Genetic analysis of ABA signal transduction pathways. Trends Plant Sci 3: 231-235.

Borevitz JO, Maloof JN, Lutes J, Dabi T, Redfern JL, Trainer GT, Werner JD, Asami T, Berry CC, Weigel D and Chory J (2002). Quantitiative trait loci controlling light and hormone response in two accessions of *Arabidopsis thaliana*. Genetics 160: 683-696.

Browse J and Somerville CR (1994). Glycerolipids., EM Meyerowitz and CR Somerville, eds (Cold Spring Harbor: Cold Spring Harbor Laboratory Press), pp 881-912.

Burr B and Burr FA (1991). Recombinant inbreds for molecular mapping in Maize. Trends Genet 7: 55-60.

Cai H-W, and Morishima H (2000). Genomic regions affecting seed shattering and seed dormancy in rice. Theor Appl Genet 100: 840-846.

Casal JJ and Sánchez RA (1998). Phytochromes and seed germination. Seed Sci Res 8: 317-329.

Chamovitz D, Pecker I and Hirschberg J (1991). The molecular basis of resistance to the herbicide norflurazon. Plant Mol Biol 16: 967-974.

Chang C, Kwok SF, Bleecker AB and Meyerowitz EM (1993). Arabidopsis *ethylene-response* gene *ETR1*: similarity of product to two-component regulators. Science 262: 539-544.

Chase K, Adler FR, and Lark KG (1997) Epistat: a computer program for identifying and testing interactions between pairs of quantitative trait loci. Theor Appl Genet 94: 724-730.

Clouse SD and Sasse JM (1998). Brassinosteroids: essential regulators of plant growth and development. Annu Rev Plant Physiol Plant Mol Biol 49: 427-451.

Coelle P, Maughan JP, Mendoza A, Philip R, Bollinger DW, Veum TL, Vodkin LO and Polacco JC (2001). Generation of low phytic acid Arabidopsis seeds expressing an *E. coli* phytase during embryo development. Seed Sci Res 11: 285-291

Cone JW and Kendrick RE (1985). Fluence-response curves and action spectra for promotion and inhibition of seed germination in wildtype and long-hypocotyl mutants of *Arabidopsis thaliana* L. Planta 163: 43-54.

Cone JW and Spruit CJP (1983). Imbibition conditions and seed dormancy of *Arabidopsis thaliana*. Physiol Plant 59: 416-420.

Coolbear P (1995). Mechanisms of seed deterioration. In Seed Quality; Basic mechanisms and agricultural implications. AS Basra, ed (Food Product Press, NY), pp 223-277.

Crowe JH, Crowe LM, Carpenter JF and Wistrom CA (1987). Stabilization of dry phospholipid bilayers and proteins by sugars. Biochem J 242: 1-10

Cutler S, Ghassemian M, Bonetta D, Cooney S and McCourt P (1996). A protein farnesyl transferase involved in ABA signal transduction in Arabidopsis. Science 273: 1239-1241.

Debeaujon I and Koornneef M (2000). Gibberelin requirement for *Arabidopsis thaliana* seed germination is determined both by testa characteristics and embryonic ABA. Plant Physiol 122: 415-424.

Debeaujon I, Léon-Kloosterziel KM and Koornneef M (2000). Influence of the testa on seed dormancy, germination and longevity in *Arabidopsis thaliana*. Plant Physiol 122: 403-414.

Debeaujon I, Peeters AJM, Léon-Kloosterziel KM and Koornneef M (2001). The *transparent testa 12* gene of Arabidopsis encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. Plant Cell 13: 853-871.

Delhaize E and Randall PJ (1995). Characterization of a phosphate-accumulator mutant of *Arabidopsis thaliana*. Plant Physiol 107: 207-231.

Derkx MPM and Karssen CM (1993a). Effects of light and temperature on seed dormancy and gibberellin-stimulated germination in *Arabidopsis thaliana*: Studies with gibberellin-deficient and -insensitive mutants. Physiol Plant 89: 360-368.

Derkx MPM and Karssen CM (1993b). Variability in light-gibberellin-and nitrate requirement of *Arabidopsis thaliana* seeds due to harvest time and conditions of dry storage. J Plant Physiol 141: 574-582.

Devic M, Guilleminot J, Debeaujon I, Bechtold N, Bensaude E, Koornneef M, Pelletier G and Delseny M (1999). The *BANYULS* gene encodes a DFR-like protein and is a marker of early seed coat development. Plant J 19: 387-398.

Doerge RW (2002). Mapping and analysis of quantitative trait loci in experimental populations. Nature Genet 3: 43-52.

Dubreucq B, Grappin P and Caboche M (1996). A new method for the identification and isolation of genes essential for *Arabidopsis thaliana* seed germination. Mol Gen Genet 252: 42-50

Dubreucq B, Berger N, Vincent E, Boisson M, Pelletier G, Caboche M and Lepiniec L (2000). The Arabidopsis *AtERP1* extensin-like gene is specifically expressed in endosperm during seed germination. Plant J 23: 643-652.

Dyer WE (1993). Dormancy associated embryonic mRNAs and proteins in imbibing *Avena fatua* caryopses. Plant Physiol 88: 201-211.

Eastmond PJ and Graham IA (2001). Re-examining the role of the glyoxylate cycle in oilseeds. Trends Plant Sci 6: 72-77.

Eastmond PJ, Germain V, Langa PR, Bryce JH, Smith SM and Graham IA (2000). Postgerminative growth and lipid catabolism in oilseeds lacking the glyoxylate cycle. Proc Natl Acad Sci USA 97: 5669-5674.

Eastmond PJ, Hooks MA, Williams D, Lange P, Bechtold N, Sarrobert C, Nussaume L and Graham IA (2000). Promoter trapping of a novel medium-chain acyl-coA oxidase, which in induced transcriptionally during Arabidopsis seed germination. J. Biol Biochem 275: 34375-34381.

El-Assal SE-D, Alonso-Blanco C, Peeters AJM, Raz V and Koornneef M (2001). A QTL for flowering time in Arabidopsis reveals a novel allele of *CRY2*. Nature Genet 29: 435-440.

Eshed Y and Zamir D (1995). An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141: 1147-1162.

Farnsworth E (2000). The ecology and physiology of viviparous and recalcitrant seeds. Ann. Rev. Ecol. Syst. 31: 107-138.

Feenstra WJ (1965). Production of thiamineless mutants. Arab Inf Serv 2: 24

Feinbaum RL and Ausubel FM (1988). Transcriptional regulation of the Arabidopsis chalcone synthase gene. Mol Cell Biol 8: 1985-1992.

Fennimore SA, Nyquist WE, Shaner GE, Doerge RW and Foley ME (1999). A genetic model and molecular markers for wild oat (*Avena fatua* L.) seed dormancy. Theor Appl Genet 99: 711-718.

Finkelstein RR (1994). Mutations at two new Arabidopsis ABA response loci are similar to the *abi3* mutations. Plant J 5: 765-771.

Finkelstein RR and Lynch TJ (2000). The Arabidopsis abscisic acid response gene *ABI5* encodes a basic leucine zipper transcription factor. Plant Cell 12: 599-609.

Finkelstein RR and Gibson SI (2001). ABA and sugar interactions regulating development: cross-talk or voices in a crowd? Curr Opin Plant Biol 5: 26-32.

Finkelstein RR, Wang ML, Lynch TJ, Rao S and Goodman HM (1998). The Arabidopsis Abscisic Acid response locus *ABI4* encodes an APETALA2 domain protein. Plant Cell 10: 1043-1054.

Focks N and Benning C (1998). *Wrinkled1*: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. Plant Physiol 118: 91-101.

Focks N, Sagasser M, Weisshaar B and Benning C (1999). Characterization of *tt15*, a novel transparent testa mutant of *Arabidopsis thaliana* (L.) Heynh. Planta 208: 352-357.

Foley ME, Bancal M and Nickols MB (1992). Carbohydrate status in dormant and afterripened excised wild oat embryos. Physiol Plant 85: 461-466.

Foley ME, Nichols MB and Meyers SP (1993). Carbohydrate concentrations and interactions in afterripening-responsive dormant *Avena fatua* caryopses induced to germinate by gibberellic acid. Seed Sci Res 3: 271-278.

Foster R and Chua N-H (1999). An Arabidopsis mutant with deregulated ABA gene expression: implication for negative regulator function. Plant J 17: 363-372.

Fransz P, Armstrong S, Alonso-Blanco C, Fischer TC, Torres-Ruiz RA and Jones G (1998). Cytogenetics for the model system *Arabidopsis thaliana*. Plant J 13: 867-876.

Gallardo K, Job C, Groot SPC, Puype MHD, Vandekerckhove J and Job D (2001). Proteomic analysis of Arabidopsis seed germination and priming. Plant Physiol 126: 835-848.

Gallardo K, Job C, Groot SPC, Puype M, Demol H, Vandekerckhove J and Job D (2002). Proteomics of Arabidopsis seed germination. A comparative study of wild type and gibberellin deficient seeds. Plant Physiol 129: 823-837

Garciarrubio A, Legaria JP and Covarrubias AA (1997). Abscisic acid inhibits germination of mature Arabidopsis seeds by limiting the availability of energy and nutrients. Planta 203: 182-187.

Ghassemian M, Nambara E, Cutler S, Kawaide H, Kamiya Y and McCourt P (2000). Regulation of abscisic acid signaling by the ethylene response pathway in Arabidopsis. Plant Cell 12: 1117-1126.

Gibson S and Somerville C (1993). Isolating plant genes. Trends Biotechnol 11: 306-313.

Gibson SI and Graham IA (1999). Another player joins the complex field of sugar-regulated gene expression in plants. Proc Natl Acad Sci USA 96: 4746-4748.

Gibson SI, Laby RJ and Kim D (2001). The *sugar-insensitive1 (sis1)* mutant of Arabidopsis in allelic to *ctr1*. Biochem Biophys Res Comm 280: 196-203.

Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F and Goodman HM (1992). Isolation of the Arabidopsis *ABI3* gene by positional cloning. Plant Cell 4: 1251-1261.

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

Goldberg RB, Paiva GD and Yadegari R (1994). Plant embryogenesis: zygote to seed. Science 266: 605-614.

Grappin P, Bouinot D, Sotta B, Miginiac E and Jullien M (2000). Control of seed dormancy in *Nicotiana plumbaginifolia*: post-imbibition abscisic acid synthesis imposes dormancy maintenance. Planta 210: 279-285.

Groot SPC, Bruinsma J and Karssen CM (1987). The role of endogenous gibberellin in seed and fruit development of tomato: studies with a gibberellin-deficient mutant. Physiol Plant 71: 184-190.

Gualberti G, Papi M, Bellucci L, Ricci I, Bouchez D, Camilleri C, Costantino P, Vittorioso, P (2002). Mutations in the Dof zinc finger genes *DAG2* and *DAG1* Influence with opposite effects the germination of Arabidopsis seeds. Plant Cell 14: 1253-1263.

Hamburger D, Rezzonico E, Mac-Donald-Comber Petétot J, Somerville C and Poirier Y (2002). Identification and characterization of the Arabidopsis *PHO1* gene involved in phosphate loading to the xylem. Plant Cell 14: 889-902.

Hampton JG and TeKrony DM (1995). Handbook of vigour test methods, 3rd edition. (The International Seed Testing Association, Zurich) 117 pp

Hanakahi LA, Bartlet-Jones M, Chappell C, Pappin D and West SC (2000). Binding of inositol phosphate to DNA-PK and stimulation of double-strand break repair. Cell 102: 721-729.

Harada JJ (1997). Seed maturation and control of germination. In Cellular and molecular biology of plant seed development, BA Larkin and IK Vasil, eds (Dordrecht: Kluwer) pp 545-592.

Haslekas C, Stacy RAP, Nygaard V, Culianez MFA and Aalen RB (1998). The expresion of a peroxiredoxin antioxidant gene, *AtPer1*, in *Arabidopsis thaliana* is seed specific and related to dormancy. Plant Mol Biol 36: 833-845.

Hayes GR and Klein WH (1974). Spectral quality influence of light during development of *Arabidopsis thaliana* plants in regulating seed germination. Plant Cell Physiol 15: 643-653.

Heath JD, Weldon R, Monnot C and Meinke DW (1986). Analysis of storage proteins in normal and aborted seeds from embryo lethal mutants of *Arabidopsis thaliana*. Planta 169: 304-312.

Helliwell CA, Sheldon CC, Olive MR, Walker AR, Zeevaart JAD, Peacock WJ and Dennis ES (1998). Cloning of the Arabidopsis ent-kaurene oxidase gene *GA3*. Proc Natl Acad Sci USA 95: 9019-9024.

Hitz WD, Carlson TJ, Kerr PS and Sebastian SA (2002). Biochemical and molecular characterization of a mutation that confers a decreased raffinosaccharide and phytic acid phenotype. Plant Physiol 128: 650-660.

Hoekstra FA, Crowe JH and Crowe LM (1989). Membrane behavior in drought and its physiological significance. In Recent advances in the development and germination of seeds RB Taylorson, ed (Plenum Press, NY) pp 71-88.

Hoekstra FA, Crowe JH and Crowe LM (1991). Effect of sucrose on phase behavior of membranes in intact pollen of *Typha latifolia* L., as measured with Fourier transform infrared spectroscopy. Plant Physiol 97: 1073-1079.

Hoffmann MH (2002). Biogeography of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae). J Biogeography 29: 125-134.

Hong S-H and Vierling E (2001). Hsp101 is necessary for heat tolerance but dispensable for development and germination in the absence of stress. Plant J 26: 1-12.

Horbowicz M and Obendorf RL (1994). Seed desiccation tolerance and storability: Dependence on flatulence-producing oligosaccharides and cyclitols - review and survey. Seed Sci Res 4: 385-405.

Huijser C, Kortstee AJ, Pego JV, Weisbeek P, Wisman E and Smeekens S (2000). The Arabidopsis *sucrose uncoupled-6* gene is identical to *abscisic acid insensitive-4*: involvement of abscisic acid in sugar responses. Plant J 23: 577-585.

Hull SR and Montgomery R (1994). Separation and analysis of 4'-epimeric UDP-sugars, nucleotides, and sugar phosphates by anion-exchange high-performance liquid chromatography with conductimetric detection. Anal Biochem 222: 49-54.

Ingram J and Bartels D (1996). The molecular basis of dehydration tolerance in plants. Annu Rev Plant Physiol Plant Mol Biol 47: 377-403.

Jacobsen SE, Binkowski KA and Olszewski NE (1996). *SPINDLY*, a tetratricopeptide repeat protein involved in gibberellin signal transduction in Arabidopsis. Proc Natl Acad Sci USA 93: 9292-9296.

Jansen RC (1993). Interval mapping of multiple quantitative trait loci. Genetics 135: 205-211.

Jansen RC (1996). Complex plant traits: time for polygenic analysis. Trends Plant Sci 1: 89-94.

Johanson U, West J, Lister C, Micheals SD, Amasino RM and Dean C (2000) Molecular analysis of *FRIGIDA*, a major determinant of natural varition in Arabidopsis flowering time. Science 290: 344-347.

Johnson RR, Cranston HJ, Chaverra ME and Dyer WE (1995). Characterization of cDNA clones for differentially expressed genes in embryos of dormant and nondormant *Avena fatua* L. caryopses. Plant Mol Biol 28: 113-122.

Jullien M, Bouinot D, Ali-Rachedi S, Sotta B and Grappin P (2000). Abscisic acid control of seed dormancy expression in *Nicotiana plumbaginifolia* and *Arabidopsis thaliana*. In Dormancy in Plants, J-D Viemont and J Crabbe, eds (CAB international) pp 195-210.

Karssen CM and Laçka E (1986). A revision of the hormone balance theory of seed dormancy: studies on gibberellin and/ or abscisic acid-deficient mutants of *Arabidopsis thaliana*. In Plant growth substances 1985, M Bopp, ed (Berlin Heidelberg: Springer-Verlag) pp 315-323.

Karssen CM, Brinkhorst-van der Swan DLC, Breekland AE and Koornneef M (1983). Induction of dormancy during seed development by endogenous abscisic acid: studies on abscisic acid deficient genotypes of *Arabidopsis thaliana* (L.) Heynh. Planta 157: 158-165.

Karssen CM, Zagórski S, Kepczynski J and Groot SPC (1989). Key role for endogenous gibberellins in the control of seed germination. Ann Bot 63: 71-80.

Kato K, Nakamura W, Tabiki T, Miura H and Sawada S (2001). Detection of loci controlling seed dormancy on group 4 chromosomes of wheat and comparative mapping with rice and barley genomes. Theor Appl Genet 102: 980-985.

Kepczynski J and Kepczynska E (1997). Ethylene in seed dormancy and germination. Physiol Plant 101: 720-726.

Kieber JJ, Rothenberg M, Roman G, Feldmann KA and Ecker JR (1993). *CTR1*, a negative regulator of the ethylene response pathway in Arabidopsis, encodes a member of the raf family of protein kinases. Cell 72: 427-441.

Kohchi T, Mukougawa K, Frankenberg N, Masuda M, Yokota A and Lagarias JC (2001). The Arabidopsis *HY2* gene encodes phytochromobilin synthase, a ferredoxin-dependent biliverdin reductase. Plant Cell 13: 425-436.

Konieczny A and Ausubel FM (1993). A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4: 403-410.

Koornneef M (1981). The complex syndrome of TTG mutants. Arab Inf Serv 18: 45-51.

Koornneef M and van der Veen JH (1980). Induction and analysis of gibberellin-sensitive mutants in *Arabidopsis thaliana* (L.) Heynh. Theor Appl Genet 58: 257-263.

Koornneef M and Karssen CM (1994). Seed dormancy and germination. In Arabidopsis, EM Meyerowitz and CR Somerville, eds (Cold Spring Harbor: Cold Spring Harbor Laboratory Press) pp 313-334.

Koornneef M and Peeters AJM (1999). Genetic approaches to abiotic stress responses. In Molecular responses to cold, drought, heat and salt stress in higher plants K Shinozaki and K Yamaguchi-Shinozaki, eds (RG Landes Company, Austin) pp 1-10.

Koornneef M, Rolff E and Spruit CJP (1980). Genetic control of light-inhibited hypocotyl elongation in *Arabidopsis thaliana* (L.) Heynh. Zeitschrift Fur Pflanzenphysiologie. 100: 147-160.

Koornneef M, Bentsink L and Hilhorst HWM (2002). Seed dormancy and germination. Curr Opin Plant Biol 5: 33-36.

Koornneef M, Jorna ML, Brinkhorst-van der Swan DLC and Karssen CM (1982). The isolation of abscisic acid (ABA) deficient mutants by selection of induced revertants in non-germinating gibberillin sensitive lines of *Arabidopsis thaliana* (L.) Heynh. Theor Appl Genet 61: 385-393.

Koornneef M, Hanhart CJ, Hilhorst HWM and Karssen CM (1989). In vivo inhibition of seed development and reserve protein accumulation in recombinants of abscisic acid biosynthesis and responsiveness mutants in *Arabidopsis thaliana*. Plant Physiol 90: 463-469.

Koornneef M, Van Eden J, Hanhart CJ, Stam P, Braaksma FJ and Feenstra WJ (1983). Linkage map of *Arabidopsis thaliana*. J Heredity 74: 265-272.

Koornneef M, Elgersma A, Hanhart CJ, van Loenen-Martinet EP, van Rijn L and Zeevaart JAD (1985). A gibberellin insensitive mutant of *Arabidopsis thaliana*. Physiol Plant 65: 33-39.

Koornneef M, Alonso-Blanco C, Bentsink L, Blankestijn-de Vries H, Debeaujon I, Hanhart CJ, Léon-Kloosterziel KM, Peeters T and Raz V (2000). The genetics of seed dormancy in *Arabidopsis thaliana*. In Dormancy in Plants, J-D Viémont and J Crabbé, eds (Oxon:CABI publishing) pp 365-373.

Koster KL and Leopold AC (1988). Sugars and desiccation tolerance in seeds. Plant Physiol 88: 829-832.

Kowalski SP, Lan TH, Feldmann KA and Paterson AH. (1994). QTL mapping of naturally-occurring variation in flowering time of *Arabidopsis thaliana*. Mol Gen Genet. 1 245, 548-555.

Kraepiel Y, Rousselin P, Sotta B, Kerhoas L, Einhorn J, Caboche M and Miginiac E (1994). Analysis of phytochrome- and ABA-deficient mutants suggests that ABA degradation is controlled by light in *Nicotiana plumbaginifolia*. Plant J 6: 665-672.

Krannitz PG, Aarssen LW and Dow JM (1991). The effect of genetically based differences in seed size on seedling survival in *Arabidopsis thaliana* (Brassicaceae). Am J Bot 78: 446-450.

Krebbers E, Broglie R, Hitz B, Jones T, and Hubbard N (1997). Biotechnological approaches to altering seed composition. In Cellular and Molecular Biology of Plant Seed Development. BA Larkins and IK Vasil, eds (Kluwer Academic Publishers, Dordrecht) pp 595-633.

Kugler I (1951). Untersuchungen über das Keimverhalten einiger Rassen von *Arabidopsis thaliana* (L.) Heynh. Ein Beitrag zum problem der Lichtkeimung. Beitr Biol Pflanzen 28: 211-243.

Laby RJ, Kincaid S, Kim D and Gibson SI (2000). The Arabidopsis *sugar-insensitive mutants sis4* and *sis5* are defective in abscisic acid synthesis and response. Plant J 23: 587-596.

Laibach F (1943). *Arabidopsis thaliana* (L.) Heynh. als object für genetische und entwicklunsphysioloische untersuchungen. Botanisches Archiv 44: 439-455.

Laibach F (1951). Uber sommer und winterannuelle rasse von *Arabidopsis thaliana* (L.) Heynh. Ein beitrag zur atiologie der blutenbildung. Beitr Biol Pflanzen 28: 173-210.

Langridge J and Griffing B (1959). A study of high temperatures lesions in *Arabidopsis thaliana*. Aust J Biol Sci 12: 117-135.

Larson S, Young K, Blake T and Raboy V (2000). Isolation and genetic mapping of a non-lethal rice low phytic acid I mutation. Crop Sci 40: 1397-1405.

Larson S, Young K, Cook A, Blake T and Raboy V (1998). Linkage mapping of two mutations that reduce phytic acid content of barley grain. Theor Appl Genet 97: 141-146.

Lawrence MJ (1976) Variations in natural populations of *Arabidopsis thaliana* (L.) Heynh. In The biology and chemistry of the cruciferae, JG Vaughan, AJ Macleod and BMG Jones eds (Academic Press, London, New York, San Francisco) pp 167-190.

Lemtiri-Chlieh F, MacRobbie EAC and Brearley CA (2000). Inositol hexaphophate is a physiological signal regulating the K+-inward rectifying conductance in guard cells. Proc Natl Acad Sci USA 97: 8687-8692.

Léon-Kloosterziel KM, Keijzer CJ and Koornneef M (1994). A seed shape mutant of Arabidopsis that is affected in integument development. Plant Cell 6: 385-392.

Léon-Kloosterziel KM, van de Bunt GA, Zeevaart JAD and Koornneef M (1996a). Arabidopsis mutants with a reduced seed dormancy. Plant Physiol 110: 233-240.

Léon-Kloosterziel KM, Alvarez-Gil M, Ruijs GJ, Jacobsen SE, Olszewski NE, Schwartz SH, Zeevaart JAD and Koornneef M (1996b). Isolation and characterization of abscisic acid-deficient Arabidopsis mutants at two new loci. Plant J 10: 655-661.

Leopold AC, Sun WQ and Bernal-Lugo I (1994). The glassy state in seeds: analysis and function. Seed Sci Res 4: 267-274.

Leung J and Giraudat J (1998). Abscisic acid signal transduction. Ann Rev Plant Physiol Plant Mol Biol 49: 199-222.

Leung J, Merlot S and Giraudat J (1997). The Arabidopsis *ABI2* gene is a homologue of *ABI1* and implicates redundant protein serine/threonine phosphatases 2C in abscisic acid signal transduction. Plant Cell 9: 759-771.

Leung J, Bouvier-Durand M, Morris P-C, Guerrier D, Chefdor F and Giraudat J (1994). Arabidopsis ABA-response gene *ABI1*: features of a calcium-modulated protein phosphatase. Science 264: 14481-11452.

Li J and Chory J (1997). A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. Cell 90: 825-827.

Li H-m, Culligan K, Dixon RA and Chory J (1995). *CUE1*: A mesophyll cell-specific positive regulator of light-controlled gene expression in Arabidopsis. Plant Cell 7: 1599-1610.

Li J, Nagpal P, Vitart V, McMorris TC and Chory J (1996). A role for brassinosteroids in light-dependent development of Arabidopsis. Science 272: 398-401.

Lickfett T, Matthäus B, Velasco L, Möllers C (1999). Seed yield, oil and phytate concentration in the seeds of two oilseed rape cultivars as affected by different phosphorus supply. Eur J Agron 11: 292-299.

Liu J-JJ, Krenz DC, Galvez AF and de Lumen BO (1998). Galactinol synthase (GS): increased enzyme activity and levels of mRNA due to cold and desiccation. Plant Sci 134: 11-20.

Lobin W (1983). The occurrence of *Arabidopsis thaliana* in the Cape Verde Islands. Arab Inf Serv 20: 119-123.

Loewus FA and Murthy PPN (2000). Myo-inositol metabolism in plants. Plant Sci 150: 1-19.

Lopez-Molina L, Mongrand S and Chua N-H (2001). A postgermination developmental arrest checkpoint is mediated by abscisic acid and requires the *ABI5* transcription factor in Arabidopsis. Proc Natl Acad Sci USA 98: 4782-4787.

Lotan T, Ohto M, Yee KM, West MA, Lo R, Kwong RW, Yamagishi K, Fischer RL, Goldberg RB and Harada JJ (1998). Arabidopsis *LEAFY COTYLEDON1* is sufficient to induce embryo development in vegetative cells. Cell 93: 1195-1205.

Lott JNA, Ockenden I, Raboy V and Batten GD (2000). Phytic acid and phosphorus in crop seeds and fruits: a global estimate. Seed Sci Res. 10: 11-33.

Lucca P, Hurrell R and Potrykus I (2001). Genetic engeneering approaches to improve the bioavailability and the level of iron in rice grains. Theor Appl Genet 102: 392-397.

Luerssen H, Kirik V, Herrmann P and Misera S (1998). *Fusca3* encodes a protein with a conserved VP1/ABI3-like B3 domain which is of functional importance for the regulation of seed maturation in *Arabidopsis thaliana*. Plant J 15: 755-764.

Lukowitz W, Gillmor S and Scheible W-R (2000). Positional cloning in Arabidopsis. Why it feels good to have a genome initiative working for you. Plant Physiol 123: 795-805.

Maes T, De Keukeleire P and Gerats T (1999). Plant tagnology. Trends Plant Sci 4: 90-96.

Maugenest S Martinez I and Lescure AM (1997). Cloning and characterization of a cDNA encoding a maize seedling phytase. Biochem J 322: 511-517.

Mayer U, Ruiz RAT, Berleth T, Miséra S and Jürgens G (1991). Mutations affecting body organization in the Arabidopsis embryo. Nature 353: 402-407.

McCullough JM and Shropshire WJ (1970). Physiological predetermination of germination responses in *Arabidopsis thaliana* (L.) Heynh. Plant Cell Physiol. 11: 139-148.

McDonald MB (1999). Seed deterioration: physiology, repair and assessment. Seed Sci Techn 27: 177-237.

Meinke DW (1994). Seed development in *Arabidopsis thaliana*. In Arabidopsis, EM Meyerowitz and CR Somerville, eds (Cold Spring Harbor: Cold Spring Harbor Laboratory Press) pp 253-298.

Meinke DW (1995). Molecular genetics of plant embryogenesis. Annu Rev Plant Physiol Plant Mol Biol 46: 369-394.

Meinke DW, Franzmann LH, Nickle TC and Yeung EC (1994). *Leafy cotyledon* mutants of Arabidopsis. Plant Cell 6: 1049-1064.

Meinke DW, Cherry JM, Dean C, Rounsley SD and Koornneef M (1998). *Arabidopsis thaliana*: A model plant for genome analysis. Science 282: 662-682.

Meyer K, Leube MP and Grill E (1994). A protein phosphatase 2C involved in ABA signal transduction in *Arabidopsis thaliana*. Science 264: 1452-1455.

Meyerowitz EM (2002). Plants compared to animals: the broadest comparative study of development. Science 295: 1482-1485.

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

Minet M, Dufour M and Lacroute F (1992). Complementation of *Saccharomyces cerevisiae* autotrophic mutants by *Arabidopsis thaliana* cDNAs. Plant J 2: 417-422.

Mohamed-Yasseen Y, Barringer SA, Splittstoesser WE and Costanza S (1994). The role of seed coats in seed viability. Botanical Rev 60: 426-439.

Moore RP (1985). Handbook on tetrazolium testing. (International seed testing association, Zurich) 99 pp.

Munir J, Dorn LA, Donohue K and Schmitt J (2001). The effect of maternal photoperiod on seasonal dormancy in *Arabidopsis thaliana* (Brassicaceae). Amer J Bot. 88: 1240-1249.

Muramoto T, Kohchi T, Yokota A, Hwang I and Goodman HM (1999). The Arabidopsis photomorphogenic mutant *hy1* is deficient in phtyochrome chromophore biosynthesis as a result of a result of a mutation in a plastid heme oxygenase. Plant Cell 11: 335-347.

Nambara E, Naito S and McCourt P (1992). A mutant of Arabidopsis which is defective in seed development and storage protein accumulation is a new *abi3* allele. Plant J 2: 435-441.

Nambara E, Keith K, McCourt P and Naito S (1995) A regulatory role for the *ABI3* gene in the establishment of embryo maturation in *Arabidopsis thaliana*. Development 121: 629-636.

Nambara E, Hayama R, Tsuchiya Y, Nishimura M, Kawaide H, Kamiya Y and Naito S (2000). The role of *ABI3* and *FUS3* loci in *Arabidopsis thaliana* on phase transition from late embryo development to germination. Dev Biol 220: 412-423.

Napp-Zin K (1975). On the genetical basis of light requirement in seed germination of Arabidopsis. Arab Inf Serv 12: 10-10.

Nesi N, Jond C, Debeaujon I, Caboche M and Lepiniec L (2001). The Arabidopsis *TT2* gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. Plant Cell 13: 2099-2114.

Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M and Lepiniec L (2000). The *tt8* gene encodes a basic helix-loop helix domain protein required for expression of *DFR* and *BAN* genes in Arabidopsis siliques. Plant Cell 12: 1863-1878.

Obendorf RL(1997). Oligosaccharides and galactosyl cyclitols in seed desiccation tolerance. Seed Sci Res 7: 63-74.

Oliver MJ and Bewley JD (1997). Desiccation tolerance of plant tissues: a mechanistic overview. Hort Rev 18: 171-213.

Ooms JJJ, Léon-Kloosterziel KM, Bartels D, Koornneef M, and Karssen CM (1993). Acquisition of desiccation tolerance and longevity in seeds of *Arabidopsis thaliana* - A comparative study using *abscisic acid- insensitive abi3* mutants. Plant Physiol 102: 1185-1191.

Pang PP, Pruitt RE and Meyerowitz EM (1988). Molecular cloning, genomic organization, expression and evolution of 12S seed storage protein genes of *Arabidopsis thaliana*. Plant Mol Biol 11: 805-820.

Papi M, Sabatini S, Bouchez D, Camilleri C, Costantino P and Vittorioso P (2000). Identification and disruption of an Arabidopsis zinc finger gene controlling seed germination. Genes Dev 14: 28-33.

Parcy F, Valon C, Kohara A, Miséra S and Giraudat J (1997). The *ABSCISIC ACID-INSENSITIVE3, FUSCA3*, and *LEAFY COTOLEDON1* loci act in concert to control multiple aspects of Arabidopsis seed development. Plant Cell 9: 1265-1277.

Parinov S and Sundaresan V (2000). Functional genomics in Arabidopsis: large-scale insertional mutagenesis complements the genome sequencing project. Curr Opin Biotechnol 11: 157-161.

Park DH, Oh SA, Kim YH, Woo HR and Nam HG. (1998). Differential expression of scenescence-associated mRNAs during leaf senescence induced by different senescence-inducing factors in Arabidopsis. Plant Mol Biol 37: 445-454.

Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES and Tanksley SD (1991). Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. Genetics 127: 181-197.

Peeters JM, Blankestijn-de Vries H, Hanhart CJ, Léon-Kloosterziel KM, Zeevaart JAD and Koornneef M (2002). Characterization of mutants with reduced seed dormancy at two novel *rdo* loci and a further characterization of *rdo1* and *rdo2* in *Arabidopsis*. Physiol Plant 115: 604-612.

Pego JV, Kortstee AJ, Huijser C and Smeekens S (2000). Photosynthesis, sugars and the regulation of gene expression. J Exp Bot 51: 407-416.

Pelletier MK and Shirley BW (1996). Analysis of flavanone 3-hydroxylase in Arabidopsis seedlings. Coordinate regulation with chalcone synthase and chalcone isomerase. Plant Physiol 111: 339-345.

Peng J, Carol P, Richard DE, King KE, Cowling RJ, Murphy GP. and Harberd NP (1997). The Arabidopsis *GAI* gene defines a signalling pathway that negatively regulates gibberellin responses. Genes Dev 11: 3194-3205.

Perl M (1986). ATP synthesis and utilization in the early stages of seed germination in relation to seed dormancy and quality. Physiol Plant 66: 177-182.

Peters JL, Constandt H, Neyt P, Cnops G, Zethof J, Zabeau M and Gerats T (2001). A physical amplified fragment-length polymorphism map of Arabidopsis. Plant Physiol 127: 1579-1589.

Poirier Y, Thom S, Somerville C and Schiefelbein J (1991). A mutant of Arabidopsis deficient in xylem loading of phosphate. Plant Physiol 97: 1087-1093.

Poppe C and Schäfer E (1997). Seed germination of *Arabidopsis thaliana phyA/phyB* double mutants is under phytochrome control. Plant Physiol 114: 1487-1492.

Powell A and Matthews S (1984). Application of the controlled deterioration vigour test to detect seed lots of Brussels sprouts with low potential for storage under commercial conditions. Seed Sci Techn 12: 649-657.

Prioul J-L, Quarrie S, Causse M and de Vienne D (1997). Dissecting complex physiological functions through the use of molecular quantitative genetics. J Exp Bot 48: 1151-1163.

Quick WA, Hsiao AI and Hanes JA (1997). Dormancy implications of phosphorus levels in developing caryopses of wild oats (*Avena fatua* L.). J Plant Growth Regul 16: 27-34

Raboy V (1997). Accumulation and storage of phosphate and minerals. In Cellular and molecular biology of plant seed development, BA Larkins and IK Vasil, eds (Dordrecht: Kluwer Academic Publishers) pp 441-477.

Raboy V (2001). Seeds for a better future: "low phytate" grains help to overcome malnutrition and reduce pollution. Trends Plant Sci 6: 458-462.

Raboy V and Dickinson D (1987). The timing and rate of phytic acid accumulation in developing soybean seeds. Plant Physiol 85: 841-844.

Raboy V and Gerbasi P (1996). Genetics of myo-inositol phosphate synthesis and accumulation. In Myo-inositol phosphates, phosphoinositides and signal transduction, B Biswas and S Biswas, eds (New York: Plenum Press) pp 257-285.

Raboy V, Dickinson DB and Neuffer MG (1990). A Survey of maize kernel mutants for variation in phytic acid. Maydica 35: 383-390.

Raboy V, Young KA, Dorsch JA and Cook A (2001). Genetics and breeding of seed phosphorus and phytic acid. J Plant Physiol 158: 489-497.

Raboy V, Gerbasi PF, Young KA, Stoneberg SD, Pickett SG, Bauman AT, Murthy PPN, Sheridan WF and Ertl DS (2000). Origin and seed phenotype of maize *low phytic acid 1-1* and *low phytic acid 2-1*. Plant Physiol 124: 355-368.

Ramsay LD, Jennings DE, Bohuon EJR, Arthur AE, Lydiate DJ, Kearsey MJ and Marshall DF (1996). The construction of a substitution library of recombinant backcross lines in *Brassica oleracea* for the precision mapping of quantitative trait loci. Genome 39: 558-567.

Ratcliffe D (1976). Germination characteristics and their inter-and intra-population variability in Arabidopsis. Arab Inf Serv 13: 34-45.

Raz V, Bergervoet JHW and Koornneef M (2001). Sequential steps for developmental arrest in Arabidopsis seeds. Development 128: 243-252.

Rédei GP (1992). A heuristic glance to the past of Arabidopsis genetics. In Methods in Arabidopsis Research. C Koncz, N Chua and J Schell eds (World Scientific, Singapore) pp 1-15.

Rédei GP (1970). *Arabidopsis thaliana* (L.) Heynh. A review of the genetics and biology. Bibliogr Genet 20: 1-151.

Remington DL, Ungerer MC and Purugganan MD (2001). Map-based cloning of quantitative trait loci: progress and prospects. Genet Res 78: 213-218.

Robson F, Costa MMR, Hepworth SR, Vizir I, Pineiro M, Reeves PH, Putterill J and Coupland G (2001). Functional importance of conserved domains in the flowering-time gene *CONSTANCE* demonstrated by analysis of mutant alleles and transgenic plants. Plant J 28: 619-631.

Rock CD and Zeevaart JAD (1991). The *aba* mutant of *Arabidopsis thaliana* is impaired in epoxy-carotenoid biosynthesis. Proc Natl Acad Sci USA 88: 7496-7499.

Romagosa I, Han F, Clancy JA and Ullrich SE (1999) Individual locus effects on dormancy during seed development and after ripening in barley. Crop Sci 39: 74-79.

Russell L, Larner V, Kurup S, Bougourd S and Holdsworth MJ (2000). The Arabidopsis *COMATOSE* locus regulates germination potential. Development 127: 3759-3767.

Safrany ST, Caffrey JJ, Yang X and Shears SB (1999). Diphosphoinositol polyphosphates: the final frontier for inositide research? Biol Chem 380: 945-951.

Sajula D, Berry M and Sachar RC (1987). Inorganic phosphate mimics the specific action of gibberellic acid in regulating the activity of monophenolase in embryo-less half seeds of wheat. Phytochem 26: 611-614.

Sasakawa N, Sharif M and Hanley MR (1995). Metabolism and biological activities of inositol pentakisphosphate and inositol hexakisphosphate. Biochem Pharmacol 50: 137-146.

Schneitz K (1999). The molecular and genetic control of ovule development. Curr Opin Plant Biol 2: 13-17.

Schoenbohm C, Martens S, Eder C, Forkmann G and Weisshaar B (2000). Identification of the *Arabidopsis thaliana* flavonoid 3'-hydroxylase gene and functional expression of the encoded P450 enzyme. Biol Chem 381: 749-753.

Schumacher K and Chory J (2000). Brassinosteroid signal transduction: still casting the actors. Curr Opin Plant Biol 3: 79-84.

Schwartz SH, Léon-Kloosterziel KM, Koornneef M and Zeevaart JAD (1997). Biochemical characterization of the *aba2* and *aba3* mutants in *Arabidopsis thaliana.* Plant Physiol 114: 161-166.

Sebastian SA, Kerr PS, Pearlstein RW and Hitz WD (2000). Soybean germplasm with novel genes for improved digestibility. In Soy in animal nutrition, JK Drackley, ed (Savoy, IL: Federation of Animal Science Societies) pp 56-74.

Seo M, Peeters AJM, Koiwai H, Oritani T, Marion-Poll A, Zeevaart JAD, Koornneef M, Kamiya Y and Koshiba T (2000). The Arabidopsis *aldehyde oxidase 3* (*AAO3*) gene product catalyzes the final step in abscisic acid biosynthesis in leaves. Proc Natl Acad Sci USA 97: 12908-12913.

Sharrock RA and Quail PH (1989). Novel phytochrome sequences in *Arabidopsis thaliana*: structure, evolution, and differential expression of a plant regulatory photoreceptor family. Genes Dev 3: 1745-1757.

Sheldon CC, Burn JE, Perez PP, Metzger J, Edwards JA, Peacock WJ and Dennis ES (2002) The *FLF* MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. Plant Cell 11: 445-458.

Shinomura T, Nagatani A, Chory J and Furuya M (1994). The induction of seed germination in *Arabidopsis thaliana* is regulated principally by phytochrome B and secondarily by phytochrome A. Plant Physiol 104: 363-371.

Shinomura T, Nagatani A, Hanzawa H, Kubota M, Watanabe M and Furuya M (1996). Action spectra for phytochrome A- and B specific photoinduction of seed germination in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 93: 8129-8133.

Shirley BW, Henley S and Goodman HM (1992). Effects of ionizing radiation on a plant genome: analysis of two Arabidopsis transparent testa mutations. Plant Cell 4: 333-347.

Silverstone AL, Ciampaglio CN and Sun T-P (1998). The Arabidopsis *RGA* gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway. Plant Cell 10: 155-169.

Sinniah UR, Ellis RH and John P (1998). Irrigation and seed quality development in rapid-cycling Brassica: Soluble carbohydrates and heat-stable proteins. Ann Bot 82: 647-655.

Skriver K and Mundy J (1990). Gene expression in response to abscisic acid and osmotic stress. Plant Cell 2: 503-512.

Smeekens S (1998). Sugar regulation of gene expression in plants. Curr Opin Plant Biol 1: 230-234.

Söderman EM, Brocard IM, Lynch TJ and Finkelstein RR (2000). Regulation and function of the Arabidopsis ABA-insensitive gene in seed abscisic acid response signaling networks. Plant Physiol 124: 1752-1765.

Steber CM and McCourt P (2001). A role for brassinosteriods in germination in Arabidopsis. Plant Physiol 125: 763-769.

Steber CM, Cooney S and McCourt P (1998). Isolation of the GA-response mutant *sly1* as a suppressor of *ABI1-1* in *Arabidopsis thaliana*. Genetics. 149, 509-521.

Stone SL, Kwong LW, Yee KM, Pelletier J, Lepiniec L, Fischer RL, Goldberg RB and Harada JJ (2001). *LEAFY COLTYLEDON2* encodes a B3 domain transcriptionfactor that induces embryo development. Proc Natl Acad Sci USA 98: 11806-11811.

Stuber CW, Edward MD and Wendel JF (1987). Molecular marker-facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its component traits. Crop Sci 27: 639-648.

Streatfield SJ, Weber A, Kinsman EA, Hausler RE, Li J, Post-Beittenmiller D, Kaiser WM, Pyke KA, Flugge U-I and Chory J (1999). The phosphoenolpyruvate/phosphate translocator is required for phenolic metabolism, palisade cell development and plastid-dependent nuclear expression. Plant Cell 11: 1609-1621.

Sun T, Goodman HM and Ausubel FM (1992). Cloning the Arabidopsis *GA1* locus by genomic subtraction. Plant Cell 4: 119-128.

Sun T-P (2000). Gibberellin signal transduction. Curr Opin Plant Biol 3: 374-380.

Swarup K, Alonso-Blanco C, Lynn JR, Micheals SD, Amasino RM, Koornneef M and Millar AJ (1999) Natural allelic variation identifies new genes in the Arabidopsis circadian system. Plant J 20: 67-77.

Taji T, Seki M, Yamaguchi-Shinozaki K, Kamada H, Giraudat J, and Shinozaki K (1999). Mapping of 25 drought-inducible genes, *RD* and *ERD*, in Arabidopsis thaliana. Plant Cell Physiol 40: 119-123.

Tanksley SD (1993). Mapping polygenes. Annu Rev Genet. 27: 205-233.

Telfer A, Bolman KM and Poethig RS (1997). Phase change and the regulation of trichome distribution in Arabidopsis. Development 124: 645-654.

Thorlby G, Veale E, Butcher K, and Warren G (1999). Map positions of *SFR* genes in relation to other freezing-related genes of *Arabidopsis thaliana*. Plant J 17: 445-452.

Ullrich SE, Hayes PM, Dyer WE, Blake TK and Clancy JA (1993) Quantititative trait locus analysis of seed dormancy in "Steptoe"barley. In Pre-harvest Sprouting in Cereals 1992, MK Walker Simmons and JL Ried eds (Am Assoc Cereal Chem, St. Paul, MN) pp 136-145.

van der Schaar W, Alonso-Blanco C, Léon-Kloosterziel KM, Jansen RC, Van Ooijen JW and Koornneef M (1997). QTL analysis of seed dormancy in Arabidopsis using recombinant inbred lines and MQM mapping. Heredity 79: 190-200.

van Ooijen JW (1992). Accuracy of mapping quantitative trait loci in autogamous species. Theor Appl Genet 84: 803-811.

van Ooijen JW (1999) LOD significance thresholds for QTL analysis in experimental populations of diploid species. Heredity 83: 613-624.

van Ooijen JW (2000) MapQTL (r) version 4.0: Userfriendly power in QTL mapping; addendum to the manual of version 3.0. Plant research International, Wageningen, The Netherlands.

van Ooijen JW and Maliepaard C (1996). MapQTL(tm) version 4.0: Software for the calculation of QTL positions on genetic maps. Plant Research International, Wageningen, The Netherlands.

Vittorioso P, Gualberti G, Papi M, Ricci I and Costantino P (2001). *DAG1* and *DAG2*: two transcription factors involved in seed germination in Arabidopsis. Abstractbook 12th International Conference on Arabidopsis Research.

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

Walker AR, Davidson PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD and Gray JC (1999). The *transparent testa glabra1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. Plant Cell 11: 1337-1349.

Weber H and Wobus U (1999). Sugars as signal molecules in plant seed development. Biol Chem 380: 937-944.

Weber H, Borisjuk L and Wobus U (1997) Sugar import and metabolism during seed development. Trends Plant Sci 2: 169-174.

Wehmeyer N and Vierling E (2000). The expression of small heat shock proteins in seeds responds to discrete developmental signals and suggests a general protective role in desiccation tolerance. Plant Physiol 122: 1099-1108.

Weiss J (1991). Ionenchromatographie. VCH Verlag, Weinheim, New York, Basel, Cambridge. ISBN 3-527-28236.

Western TL, Skinner DJ and Haughn GW (2000). Differentiation of mucilage secretion cells of the Arabidopsis seed coat. Plant Physiol 122: 345-355.

Wilcox J, Premachandra G, Yound K and Raboy V (2000). Isolation of high seed inorganic P, low-phytate soybean mutants. Crop Sci 40: 1601-1605.

Willats WGT, McCartney L and Knox JP (2001). In-situ analysis of pectic polysaccharides in seed mucilage and at the root surface of Arabidopsis. Planta 213: 37-44.

Windsor JB, Symonds VV, Mendenhall J and Lloyd AM (2000). Arabidopsis seed coat development: morphological differentiation of the outer integument. Plant J 22: 483-493.

Wisman E, Hartmann U, Sagasser M, Baumann E, Palme K, Hahlbrock K, Saedler H and Weisshaar B (1998). Knock-out mutants from an En-1 mutagenized *Arabidopsis thaliana* population generate phenylpropanoid biosynthesis phenotypes. Proc Natl Acad Sci USA 95: 12432-12437.

Xiong L, Ishitani M, Lee H and Zhu J-K (1999). *HOS5* - a negative regulator of osmotic stress-induced gene expression in *Arabidopsis thaliana*. Plant J 19: 569-578.

Xiong L, Ishitani M, Lee H and Zhu J-K (2001). The Arabidopsis *LOS5/ABA3* locus encodes a molybdenum cofactor sulfurase and modulates cold stress- and osmotic stress-responsive gene expression. Plant Cell 13: 2063-2083.

Yamaguchi S, Smith MW, Brown RG, Kamiya Y and Sun T (1998). Phytochrome regulation and differential expression of *gibberellin 3 β-hydroxylase* genes in germinating Arabidopsis seeds. Plant Cell 10: 2115-2126.

Yang Y-Y, Nagatani A, Zhao Y-J, Kang B-J, Kendrick RE and Kamiya Y (1995). Effects of gibberellins on seed germination of phytochrome- deficient mutants of *Arabidopsis thaliana*. Plant Cell Physiol 36: 1205-1211.

York JD, Odom AR, Murphy R, Ives EB and Wente SR (1999). A phospholipase C-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export. Science 285: 96-100.

Zeng Z-B (1993). Theoretical basis of precision mapping of quantitative trait loci. Proc Natl Acad Sci USA 90: 10972-10976.

Zhou L, Jang J-C, Jones TL and Sheen J (1998). Glucose and ethylene signal transduction crosstalk revealed by an Arabidopsis glucose-insensitive mutant. Proc Natl Acad Sci USA 95: 10294-10299.

Zakhleniuk OV, Raines CA and Lloyd JC (2001). *pho3*: a phosphorus-deficient mutant of *Arabidopsis thaliana* (L.) Heynh. Planta 212: 259-534.

Samenvatting

Arabidopsis thaliana (Zandraket) komt van nature voor in een groot aantal verschillende gebieden in de wereld. Binnen de soort bestaat voor veel eigenschappen een aanzienlijke genetische variatie. Deze genetische variatie kan gebruikt worden bij het vaststellen van de functies van genen o.a. die betrokken zijn bij de aanpassing aan specifieke groeiomstandigheden. Alle eigenschappen, waarvan de genetica beschreven staat in dit proefschrift, zijn gebaseerd op de genetische variatie die aanwezig is tussen de accessies Landsberg *erecta* (L*er*) en Kaap Verdische Eilanden (Cvi). Deze variatie is geanalyseerd met behulp van “Recombinant Inbred Lines” (RILs), waarvan de genotypen al geanalyseerd waren met behulp van moleculaire merkers.
Het doel van het onderzoek dat in dit proefschrift beschreven staat is het vergroten van de kennis over de genetische regulatie van kiemrust. Verder is de genetica van andere zaad eigenschappen onderzocht die met kiemrust in verband zouden kunnen staan. Dit zijn de oligosaccharide samenstelling, de bewaarbaarheid en de hoeveelheid fytaat en fosfaat in zaden.

Kiemrust, gedefinieerd als de toestand van het zaad waarin het wel levensvatbaar is maar toch niet kiemt na het zaaien ook al zijn de omstandigheden gunstig, is een complexe eigenschap die door veel genen bepaald wordt. Door gebruik te maken van natuurlijke variatie worden mogelijk nieuwe kiemrust loci geïdentificeerd in aanvulling op de genen die al tijdens mutantenselecties gevonden zijn. De accessies L*er* and Cvi leverden de genetische variatie die nodig was voor dit onderzoek. Cvi heeft een diepe kiemrust, het zaad van deze plant moet gemiddeld drie maanden droog bewaard worden voordat het in staat is te kiemen. L*er*, een standaard laboratorium accessie, die slechts een beperkte kiemrust heeft, kiemt volledig na ongeveer een maand droog bewaren. De bewaringstijd die zaden nodig hebben om te kunnen kiemen wordt behalve door het genotype sterk bepaald door de omgevingsfactoren waaronder de moederplant groeit en de bewaar- en kiemingsomstandigheden. Genetisch identieke zaden afkomstig van planten die op verschillende momenten opgekweekt zijn kunnen daarom sterk verschillen in hun kiemingsgedrag. Echter de genetische en fysiologische analyse van zaden kan zonder problemen uitgevoerd worden, wanneer de zaden geoogst zijn van planten uit hetzelfde experiment en vervolgens onder gelijke omstandigheden bewaard en getest worden.

“Quantitative Trait Loci” (QTL) mapping voor kiemrust, gebruik makend van de RIL populatie gemaakt uit een kruising tussen L*er* and Cvi, resulteerde in de identificatie van de loci die verantwoordelijk zijn voor de kiemrust verschillen tussen deze twee accessies. Er zijn 7 loci geidentificeerd die kiemrust significant beïnvloeden,

deze loci zijn “*delay of germination 1-7*” (*DOG1-7*) genoemd. Voor de QTL die boven aan chromosoom 1 gelokaliseerd is, leidt het Cvi allel tot minder dormantie. Voor alle andere QTL levert Cvi de dormante allelen. De kiemrust van de meeste QTL kon bevestigd worden door “near isogenic lines” (NILs) te maken. De kiemrust van deze NILs is geanalyseerd onder diverse kiemrust brekende of kiemingsbevorderde omstandigheden, waaruit geconcludeerd kon worden dat een koude behandeling na het zaaien (stratificatie) de efficiëntste manier is om kiemrust van de verschillende QTL te breken. De genetische interactie tussen de meeste QTL waren additief, maar *DOG1* en *DOG3* vertonen een onverwachte interactie. De aanwezigheid van Cvi allelen voor zowel *DOG1* als *DOG3* resulteert in een bijna non-dormant fenotype. Dit is gebaseerd op zowel de QTL analyse als op de kiemrust van NIL DOG317, die in een L*er* achtergrond Cvi allelen op beide loci bevat.

De natuurlijke L*er* en Cvi allelen van *DOG1* (*DOG1*-L*er* en *DOG1*-Cvi), het locus met het grootste effect op kiemrust zijn verder gekarakteriseerd, tezamen met een compleet non-dormante mutant in het *DOG1* locus. *DOG1* allelen erven semi-dominant over en *DOG1* reguleert kiemrust voornamelijk via het embryo. Alle *DOG1* allelen hebben gibberelline (GA) nodig om te kunnen kiemen, hoewel de hoeveelheid GA nodig voor kieming wordt bepaald door het betreffende *DOG1* allel. Dit is bepaald door de analyse van GA afhankelijkheid van de *DOG1* allelen in de genetische achtergrond van de *ga1-3* mutant. Een zelfde soort experiment (*DOG1*-L*er* en *DOG1*-Cvi in *aba1-1* en *abi3-5* achtergrond) liet zien dat de kiemrust van *DOG1* volledig afhankelijk is van abscisinezuur (ABA). Omdat de interactie van *DOG1* met GA anders is dan de interactie van ABA met GA veronderstellen we dat het *DOG1* locus alleen betrokken is bij de inductie van kiemrust tijdens de zaadontwikkeling, terwijl ABA zowel tijdens de ontwikkeling als in de interactie met GA tijdens imbibitie en kieming een rol speelt.

De precieze genetische positie van *DOG1* is bepaald met behulp van een toetskruisingspopulatie. Hiervoor is NIL DOG17 gekruist met een merker lijn die homozygoot was voor *glabrous 3* (*gl3*) en *thiamine* (*tz*) in een L*er* achtergrond, waarna de verkregen hybride weer teruggekruist is met laatste genoemde merker lijn Recombinanten tussen *gl3* en *tz* zijn geselecteerd en een subset van deze recombinanten is geanalyseerd met moleculaire merkers. Op deze manier kon *DOG1* gelokaliseerd worden in een gebied van 75 kb waar 22 “open reading frames” liggen. Dit gebied bevat geen genen die voorheen met kiemrust gerelateerd zijn, daarom is er ook nog geen duidelijk kandidaat gen aan te wijzen. Het verder verkleinen van het *DOG1* gebied en het kloneren van het gen zal meer vertellen over de functie van *DOG1*.

Zaden van L*er* en Cvi verschillen behalve in kiemrust ook in oligosaccharide (OS) samenstelling. L*er* heeft meer raffinose en stachyose en minder sucrose dan Cvi. Raffinose en stachyose zijn voedingsreserves van zaden en hopen zich op in zaden, op hetzelfde moment waarop kiemrust geïnduceerd wordt. Om te onderzoeken of deze twee processen door dezelfde loci gereguleerd worden is er QTL mapping voor OS samenstelling in zaden van de L*er*/Cvi RIL populatie uitgevoerd. Eén hoofd QTL is verantwoordelijk voor de bijna monogene segregatie voor stachyose. Dit locus beïnvloedt ook de hoeveelheid raffinose en sucrose in zaden. Op de positie waar dit gen gelokaliseerd is liggen geen kiemrust loci.

Lang werd gedacht dat OSs in zaden en voornamelijk RSOs een belangrijke rol spelen bij de bewaarbaarheid van zaden. Om deze hypothese te testen is de bewaarbaarheid bestudeerd in dezelfde RIL populatie, door de levensvatbaarheid van zaden (kieming) te analyseren met behulp van twee methoden: na natuurlijke veroudering gedurende 4 jaar droge bewaring en na versnelde veroudering geïnduceerd met een zogenaamde “controlled deterioration test”. Op deze manier zijn er 4 QTL geidentificeerd, waarvan er geen op dezelfde positie ligt als het major OS locus. Hierdoor kunnen we concluderen dat raffinose en stachyose geen specifiek effect op bewaarbaarheid hebben.

Wel vinden we twee posities waarvoor kiemrust QTL en bewaarbaarheids QTL co-localiseren. Dit kan worden veroorzaakt doordat we bij de analyse van beide eigenschappen kiemproeven gebruiken. Verder bestaan er Arabidopsis mutanten die zowel een kiemingsfenotype als een bewaringsfenotype hebben, zoals sterke allelen van het *ABI3* locus.

Fytaat (InsP6) is een ander zaad bestandsdeel dat accumuleert gedurende de zaad ontwikkeling. InsP6 is het meest voorkomende P-bevattende verbinding in planten en is belangrijk doordat het micro-elementen zoals ijzer en zink kan binden. QTL analyses voor InsP6 en fosfaat (Pi) in zaden en bladeren van de L*er*/Cvi RIL populatie is gedaan om genen te identificeren die betrokken zijn bij de accumulatie van InsP6. Er zijn 5 loci geidentificeerd die de hoeveelheid InsP6 en Pi in zaden en bladeren significant beïnvloeden. Een QTL gelokaliseerd boven aan chromosoom 3 beïnvloed alle vier de eigenschappen. Deze QTL is de hoofd QTL en het L*er* allel verlaagt de hoeveelheid InsP6 en Pi zowel in zaden als in bladeren. Dit locus is gelokaliseerd tot een gebied van 99 kb, dat 13 “open reading frames” bevat. De maternale overerving van de QTL en de positieve relatie tussen de hoeveelheid InsP6 en Pi in bladeren en in zaden, suggereert dat het verschil in InsP6 tussen L*er* en Cvi wordt veroorzaak door een verschil in transport, in plaats van een verschil in biosynthese. Daarom denken wij dat de vacuolar membrane ATPase subunit G, dat in het 99 kb gebied ligt, het meest waarschijnlijke kandidaat gen is voor het InsP6/Pi locus.

InsP6 en RSO zijn mogelijk met elkaar gereguleert doordat ze *myo*-inositol-1-fosfaat en eventueel vrije *myo*-inositol als gezamenlijke precursor (voorloper) hebben. Wij vinden echter geen co-locatie tussen InsP6 en RSOs.

De L*er*/Cvi RIL populatie splitst uit voor vele eigenschappen, zelfs wanneer de ouders niet verschillen voor een bepaalde eigenschap kan het zijn dat de RILs splitsen voor deze eigenschap, zoals het geval is voor bloeitijd. De resultaten die in dit proefschrift beschreven zijn laten zien dat permanente populaties een groot voordeel hebben omdat dezelfde serie planten voor verschillende eigenschappen geanalyseerd kan worden. Co-locatie is echter geen bewijs dat eigenschappen door dezelfde loci gereguleerd worden, omdat de eigenschappen ook gereguleerd kunnen zijn door nauw gekoppelde genen die met QTL mapping niet van elkaar gescheiden kunnen worden. Fijn-mapping van QTL en de analyse van geselecteerde recombinanten voor de verschillende eigenschappen kunnen mogelijk deze vragen beantwoorden.

Publications

Published articles from this thesis:

Bentsink L, Alonso-blanco C, Vreugdenhil D, Tesnier, K, Groot SPC and Koornneef M (2000) Genetic analysis of seed soluble oligosaccharides in relation to seed storability of Arabidopsis. Plant Physiol 124: 1595-1604.

Bentsink L, Koornneef M (2002) Seed dormancy and germination. In The Arabidopsis Book. CR Somerville, EM Meyerowitz eds (American Society of Plant Biologist, Rockville, MD), pp DOI 10.1199 tab.0050

Additional articles:

Croes, AF, Schapendonk, W, Jansen, F, Stommen, P, **Bentsink, L**, Zwanenburg, B, and Wullems, GJ (1994) Germination of *Orobanche aegyptiaca* and *O. crenata* is stimulated in vitro by isolated roots: a comparative study. Biology and Management of Orobanche.Proc.of the 3rd Intern.Workshop on Orobanche and related Striga research, 157-168.

Soppe WJJ, **Bentsink L** and Koornneef M (1999). The early flowering mutant *efs* is involved in the autonomous promotion pathway of *Arabidopsis thaliana*. Development 126: 4763-4770.

Groot SPC, van der Geest AHM, Tesnier K, Alonso-Blanco C, **Bentsink L**, Donkers H, Koornneef M, Vreugdenhil D and Bino RJ (2000). Molecular genetic anlysis of *Arabidopsis* seed quality. In Seed biology: Advances and applications. M Black, KJ Bradford and J Vázques-Ramos eds (CAB international), pp 123-132.

Koornneef M, Alonso-Blanco C, **Bentsink L**, Blankenstijn-de Vries, Debeaujon I, Hanhart CJ, Léon-Kloosterziel KM, Peeters AJM and Raz V (2000). The genetics of seed dormancy in *Arabidopsis thaliana*. In Dormancy in plants. J-D Viémont and J Crabbé, eds (CAB international), pp 365-373.

Koornneef M, **Bentsink L**, Hilhorst HWM (2002) Seed dormancy and germination. Curr Opin Plant Biol 5: 33-36.

Bentsink L, Leone GOM, van Beckhoven JRCM, van Schijndel HB, van Gemen B and van der Wolf JM (2002) Amplification of RNA by NASBA allows direct detection of viable cells of *Ralstonia solanacearum* in potato. J Appl Microbiol, in press.

Curriculum vitea

Leónie Bentsink werd op 29 augustus 1973 in Eibergen geboren. Na het behalen van de MAVO diploma in 1989 en het MLO-botanisch in 1993 begon zij met een studie planten biotechnologie aan de International Agrarische Hogeschool Larenstein in Velp. Tijdens deze opleiding liep zij stage op het Instituut voor Plantenziektenkundig Onderzoek (IPO-DLO) nu onderdeel van Plant Research International. Daarna werd de afstudeeropdracht uitgevoerd aan de Leerstoelgroep voor Erfelijkheidsleer van de Wageningen Universiteit in de groep botanische genetica. Na het afstuderen in 1997 werkte zij 7 maanden bij het IPO-DLO om vervolgens in januari 1998 naar de Leerstoelgroep voor Erfelijkheidsleer terug te keren om daar als onderzoeker in opleiding te gaan werken aan de karakterisering van kiemrust en andere eigenschappen in Arabidopsis. De resultaten van dit onderzoek staan beschreven in dit proefschrift. Met ingang van 1 april 2002 werkt zij als post-doc bij dezelfde vakgroep als onderdeel van het Europese Project “Natural” aan de analyse van kiemrust in verschillende accessies van Arabidopsis.

Nawoord

Hoewel dit de laatste pagina's van mijn proefschrift zijn, zullen ze door vele als eerste gelezen worden. Toch betekent dat niet dat deze paginas extra vernieuwend zijn, want net als vele promovendi voor mij zal ik deze pagina's gebruiken om de mensen die mij geholpen hebben tijdens dit onderzoek te bedanken. Allereerst Maarten, ik wil je bedanken voor de mogelijk die ik kreeg OIO in jouw groep te worden, hoewel ik wist dat dit toegestaan was met een HBO diploma had ik dit zelf nog nooit overwogen. Ik was dan ook zeer verrast toen je aan me vroeg of ik na mijn afstuderen terug wilde komen, met als resultaat dat ik er een tijdje over moest nadenken. Ik kwam tot de conclusie dat ik zo'n kans niet kon laten schieten en tot op vandaag heb ik hier nog geen seconde spijt van gehad! Ik wil je natuurlijk ook bedanken voor je begeleiding, ik had me geen betere begeleider en promoter kunnen wensen. Verder ben ik erg blij dat ik de komende jaren in jouw groep verder kan blijven werken en ik hoop *DOG1* snel te kloneren. Next to Maarten there was Carlos who took care of my supervision. Carlos you taught me a lot not only during the daily supervision in the first year but as well as guide by mail and during occasional visits in the other three years. Further more thanks for correcting my manuscripts. Naast mijn begeleiders hebben de andere collega's uit de groep botanische genetica voor de goede sfeer gezorgd vandaar dat ik deze mensen die deel van de groep hebben uitgemaakt of nog doen één voor één zal noemen. Ik zal dit op alphabetische volgorde doen, maar eerst nog een speciaal woordje voor Corrie, behalve dat je een gezellige collega bent heb je me ook regelmatig geholpen met het tellen van kiemproeven en uiteindelijk heb je zelfs je vakantie verzet zodat ik verder kon schrijven aan dit proefschrift, bedankt. Verder Diana, Elisabeth, Emile, Isabelle, Jim, Gerda, Hetty, Mark, Mohamed, Salah, Ton, Wim en Vered BEDANKT!!!. Verder wil ik Hans bedanken voor zijn hulp en het gebruik van zijn computer. Zonder de ondersteuning van Gerrit voor het verzorgen van de planten en Aafke en Corrie E. voor het draaiende houden van de vakgroep kan er natuurlijk geen onderzoek gedaan worden en ik wil dan ook jullie van harte bedanken. Nu ontbreken nog de studenten die ik heb mogen begeleiden in volgorde van verschijnen, Jemma, Sophie, Klaas, Susanne en Daniela: I enjoyed it very much working with you all and although not all the work you did is included in this thesis most of it was very useful and will lead to more publications in the nearby future.

Mijn onderzoek is niet alleen binnen de leerstoelgroep erfelijkheidsleer uitgevoerd. Het werk dat in hoofdstuk 6 beschreven staat en in "Plant Physiology" gepubliseerd is was een samenwerking met de vakgroep plantenfysiology en de Business Unit "Plant Development and Reproduction van PRI. Ik wil Dick, Karine, Lonneke en Steven dan ook hartelijk bedanken voor de samenwerking. Furthermore, the work that has been described in Chapter 7 was performed in collaboration with Yuan, Diaan en Dick from the department Plant Physiology, thanks. Ik wil ook nog graag Janny Peters bedanken, die toen ik helemaal geen idée meer had waar *DOG1* gelokaliseerd is mij via "insilco AFLP analyse" weer de juiste weg kon wijzen.

Dan rest mij nog mijn familie en vrienden, met in het bijzonder: Pap, Mam, Sabine, Diane, Barbara en natuurlijk Mike te bedanken. Want, hoewel ik de afgelopen jaren de meeste vrije en niet vrije tijd in mijn promotie onderzoek gestoken heb betekent dat niet dat jullie niet belangrijk voor mij zijn. Bedankt voor jullie steun, vriendschap en liefde!!!

Tot slot, ik ga mijn koffers pakken en reis hiermee af naar de Kaap Verdische Eilanden. Wees gerust Mike, op deze tijd van het jaar groeit hier geen Arabidopsis.

The work presented in this thesis was performed at the Wageningen University, Laboratory of Genetics, Arboretumlaan 4, 6703 BD Wageningen, The graduate School Experimental plant science, NL.

The research was financially supported by The Earth and Life Sciences Foundation subsidised by The Netherlands Organisation for Scientific Research.

Omslag: Barbara Marinus