The identification of genetic variation for gibberellin biosynthesis and signalling among *Arabidopsis thaliana* accessions

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Luis Orlando Barboza Barquero

aus San José; Costa Rica

Köln, 2014

Diese Arbeit wurde am Max-Planck-Institut für Züchtungsforschung in Köln in der Abteilung für Pflanzenzüchtung und Genetik (Prof. Dr. M. Koornneef) durchgeführt.

Max Planck Institute for
Plant Breeding Research

MAX-PLANCK-GESELLSCHAFT

Berichterstatter:Prof. Dr. Prof. Dr. Maarten Koornneef

Prof. Dr. Ute Höcker

Prüfungsvorsitzender: Prof. Dr. Thomas Wiehe

Tag der mündlichen Prüfung: 21.01.2014.

Contents

Chapter 1

General Introduction: About gibberellins relevance and natural variation

The analysis of natural variation

Natural variation in species with a broad distribution range is expected to reflect adaptation to the environment in which the plants are growing. The genetic difference between naturally occurring accessions is often the consequence of allelic differences at multiple loci with a quantitative effect and modification by environmental factors. These aspects require a quantitative genetic analysis of the genetic differences (Bergelson & Roux, 2010).

To dissect the genetic basis of this type of genetic variation, Quantitative Trait Loci (QTL) analysis either using the progeny of crosses among accessions or Genome Wide Association Studies (GWAS) allow the mapping of the genes underlying this variation. To identify the genes underlying allelic variation at the QTL (Alonso-Blanco et al., 2009; Weigel, 2011) additional experiments are required which involve fine-mapping, sequence and gene expression comparisons (Weigel, 2011). Both mapping approaches have been successful and revealed that the identified genes can belong to all types of ontology classes with differences in the structural part of the genes or in the promoters (Alonso-Blanco et al., 2009; Bergelson & Roux, 2010; Weigel, 2011). How this allelic variation relates to adaptation is often not known and requires insight in the pleiotropic effects that may account for trait–offs as well as field experiments, preferentially at multiple sites, to test fitness differences in nature. Additionally, molecular population genetics can provide indications for the selection of specific alleles (Nielsen, 2005) e.g. by detecting selective sweeps of specific genes (Alcázar et al., 2010) and or correlations between the presence of specific alleles and climate or other geography related parameters (Baxter et al., 2010; Gujas et al., 2012).

QTL mapping using segregating populations, based on bi-parental crosses, have allowed the identification of main effect loci and their cloning of the genes controlling this effect. This mapping procedure shows as main feature high mapping power but low resolution, which means that regions where QTL are located may contain several-hundreds of genes as possible candidates for the gene underlying the QTL. Recombinant Inbred Lines (RILs) are genotypes used for mapping derived from a cross of two parents and followed by several generations of selfing (>8) trying to achieve high homozygosity. To bypass the problem that repeated selfing for many generations takes time, another procedure to generate homozygous lines, which is the production of Doubled Haploids (DHs) (Seymour et al., 2012) can be applied. Instead of using anther culture, which is not working effectively for Arabidopsis (Scholl & Amos, 1980) recently a technology using centromere elimination was described (Ravi & Chan 2010). For this a wild type genotype, which can be a hybrid is crossed with the specific '*cenh3'* variant (a transgenic line expressing the centromere-specific histone CENH3), which then will eliminate the 'cenh3 variant' chromosomes in the zygote, thus producing a haploid progeny that will there generate the doubled haploid lines by spontaneous chromosome doubling (Ravi & Chan, 2010; Seymour et al., 2012). QTL analysis requires validation to confirm the involvement of allelic variation of the region underlying the QTL as well as further fine-mapping. This can be done by employing Near Isogenic Lines (NILs) (Keurentjes et al., 2007) that can be obtained by backcrossing specific genotypes with one or both parents. NILs are introgressions of the QTLs of interest in a contrasting parental background, or even one QTL can be introgressed in different backgrounds (Bentsink et al., 2010). Another way to select NILs make use of so-called Heterogeneous Inbred Families HIFs (Tuinstra et al., 1997). HIFs are selected based on the identification of genotypes with residual heterozygosity in the genomic region where a QTL was mapped. When such heterozygous genotypes are selfed, thereafter one can select NILs contrasting in the alleles of the parental lines. An advantage of using these lines is that they may have a genetic background that increases the differences between the alleles of the QTL under study and thereby helps the validation of QTLs, especially when the effect of the allelic variation depends on other loci (epistasis).

GWAS try to map loci at high (gene) resolution without the need to generate mapping populations and is therefore a promising method to identify genes underlying natural variation. Instead of using experimental crosses as done with QTL analysis,

GWAS employ a group of accessions to do mapping. Although the power is not as high as with bi-parental mapping populations, several genes have been identified using this procedure (Korte & Farlow, 2013). GWAS complications are the (confounding) population structure, which means some genotypes can be in linkage disequilibrium with each other e.g. due to common origin (Korte & Farlow, 2013). Additional GWAS complications are the presence of epistatic interactions (current mapping methods do not include interaction effects), and the difficulty to assign effects to rare alleles (Ingvarsson & Street, 2011).

The combination of GWAS and bi-parental QTL mapping studies may provide an efficient scenario for gene discovery as they have complementary strengths. However this is not an easy and straight forward job as shown for flowering time a trait for which several major effect QTL have been molecularly studied (Brachi et al., 2010). Recently multi-parents RILs have been generated (Kover et al., 2009; Huang et al., 2011) in an another attempt to combine the positive attributes of bi-parental and GWAS mapping. The use of multiple parents provide a higher number of segregation alleles thus exploiting more of the variation and combinations of specific alleles at different loci that can result in novel phenotypes (Huang et al 2013). When additional intercrossing is performed more recombination can take place, improving the resolution especially when combined with a large progeny (Kover & Mott, 2012). An important additional tool for the analysis of natural variation is obtained by the re-sequencing of a large number of accessions in the 1001 genomes project (http://www.1001genomes.org), which provides genomic data allowing quick identification of polymorphisms and candidate allelic variants for natural variation in *A. thaliana*.

Currently plant biologists / geneticists are not only exposed to genomic but also phenotyping data obtained by high-throughput methods. Methodologies to automatically phenotype *A. thaliana* (and crops) have been developed for seed size (Moore et al., 2013), germination (Joosen et al., 2010), hypocotyls (Cole et al., 2011; Wang et al., 2009), shoots (Granier et al., 2006; Jansen et al., 2009) and roots (Nagel et al., 2009; Nagel et al., 2012). Most of these tools provide automated evaluation platforms and semi - / automatic image processing pipelines. Still additional methods need to be developed to avoid phenotyping being the bottleneck in genetic / genomic analysis.

Does variation in hormonal biosynthesis and signalling contribute to phenotypic variation in nature?

Variation in plant hormone levels and hormone signalling might be a basis for the phenotypic variation in developmental and stress-related traits in nature. Plant hormones are small molecules, present at low concentrations, that regulate plant growth and responses to biotic and abiotic stresses (Santner et al., 2009). Well known plant growth regulators are Abscisic Acid (ABA), Auxin, Brassinosteroids, Ethylene, Gibberellins, (GAs), Jasmonates and Salicylic acid to which recently Strigolactones (Santner et al., 2009) are added. Despite all the existent knowledge about metabolism and signalling of these compounds, few examples of natural variation of plant hormone related traits have been described and even fewer about the evolutionary role of these variants. An example of natural variation related with plant hormones in Arabidopsis is the gene *BRX* that is a positive regulator of auxin signalling (Scacchi et al., 2010) and that has been implicated with brassinosteroids homeostasis (Mouchel et al., 2006). *BRX* modifies root length under acidic conditions (Gujas et al., 2012). Natural variation in the GA biosynthesis pathway has also been found and will be addressed in detail in the following paragraphs.

Gibberellin biosynthesis and signalling

Bioactive gibberellins (GAs) are plant growth regulators involved in several traits such as seed germination, flowering time, anther and petal development, and cell elongation (Sun, 2008). Their biosynthesis and signalling is well understood (Hedden & Thomas, 2012; Sun, 2008; Yamaguchi, 2008). The GA precursor is the geranylgeranyl diphosphate (GGPD), which is also precursor of other diterpenoids (Figure 1.1) (Yamaguchi, 2008). By the action of two enzymes encoded by the ent-copalyl diphosphate synthase (*CPS*) and *ent*-Kaurene synthase (*KS*) genes, GGPD is converted into *ent*-Kaurene (Yamaguchi, 2008). *Ent*-Kaurene is thereafter converted into GA_{12} by the action of two P450 enzymes: *ent*-Kaurene oxidase (KO) and *ent*-kaurenoic acid oxidase (KAO) (Yamaguchi, 2008). GA_{12} is converted into the bioactive form GA_4 by GA 20-oxidases and GA 3-oxidases. These genes have several paralogs which perform redundant functions (Mitchum et al., 2006; Rieu et al., 2008), in contrast with the early

steps of GA biosynthesis catalyzed by enzymes encoded by single or two copies genes (Figure 1.1). Gibberellins can be inactivated by GA 2-oxidases or by gibberellin methyltransferases (*GAMT1* and *GAMT2*) (Yamaguchi, 2008). The active GAs (e.g. GA_4 for Arabidopsis) by interacting with the receptor GA INSENSITIVE DWARF1 (GID1), target the growth inhibition proteins DELLAs. This allows the recognition of this complex by *SLY1* (*SLEEPY1*). Thereafter DELLAs are targeted by the 26S proteasome machinery and degraded (Sun, 2010), releasing their inhibition effect. Up to date five *A. thaliana* DELLA proteins have been described: REPRESSOR OF GA1-3 (RGA), GIBBERELLIC ACID INSENSITIVE (GAI), RGA-LIKE1 (RGL1), RGL2 and RGL3 (Sun, 2011). These DELLA differ mainly in their expression patterns.

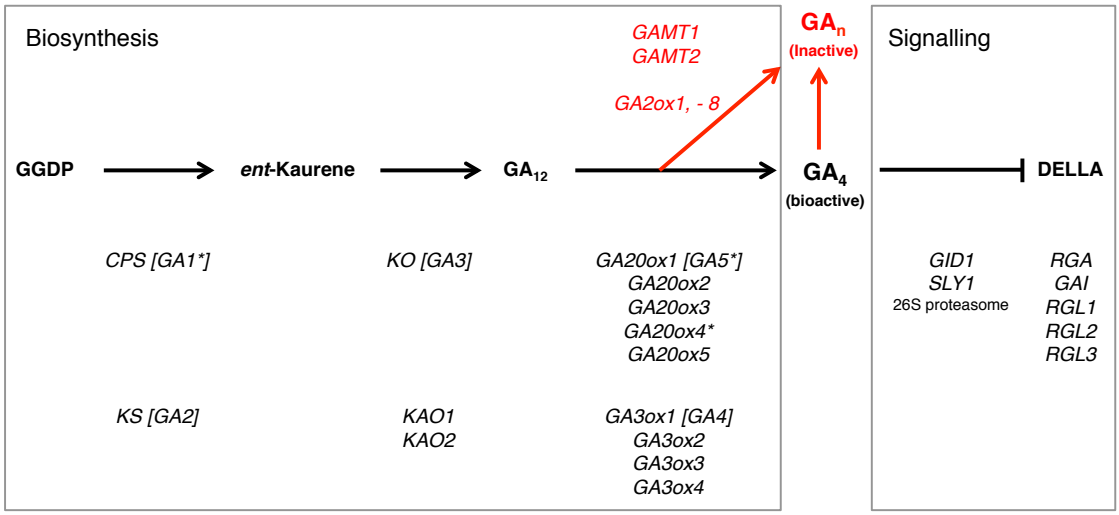

Figure 1.1. Summary of the GA biosynthesis and Signalling pathway in *Arabidopsis thaliana*. *Indicate genes with identified natural variation. The figure was summarized from previous reviews (Sun, 2010; Yamaguchi, 2008). Between brackets the previous names of the genes based on the original mutants (Koornneef and van der Veen, 1980) are mentioned.

GAs affect the expression of many genes and their levels are regulated by environmental factors as well as by other plant hormones in order to promote changes in growth or development. Seed germination is a process that can serve as example of the mode of action of GAs. In germination the antagonistic role of GA – ABA is well established. The former promotes germination while the latter inhibits germination. The hormones act both during seed maturation when dormancy is induced for which process ABA is essential and during seed imbibition when GAs levels increase and will overcome the dormancy induced by ABA among others by affecting the ABA levels (Holdsworth et al., 2008). During seed maturation GAs are regulated by the *LEAFY COTYLEDON2* (*LEC2*) and *FUSCA3* (*FUS3*) pathways (Curaba et al., 2004).

These genes are required to inhibit precocious germination and anthocyanin accumulation (Curaba et al., 2004). *FUS3* represses the *AtGA3ox2* thus reducing GA biosynthesis and consequently promoting dormancy (Holdsworth et al., 2008). During seed imbibition de-novo GA biosynthesis is induced by cold thus promoting germination (Yamaguchi, 2008). Another environmental factor affecting germination is light. When phytochromes are activated through light treatments they promote the expression of *AtGA3ox1* and *AtGA3ox2* while *AtGA2ox2* expression decreases (Yamaguchi, 2008). Besides the antagonistic GA – ABA role, GA requirement for germination is mediated by the testa characteristics, with as main evidence for this the observation that when embryos of *ga1* mutants are excised from the envelopes, germination is restored (Debeaujon & Koornneef, 2000).

GAs play an important role in promoting cell elongation and examples of this effect during hypocotyl elongation and inflorescence stem have been described. Hypocotyls grown in the dark undergo a process called skotomorphogenesis thus forming long hypocotyls with an apical hook and closed-small-chlorophyll lacking cotyledons (Gendreau et al., 1997; Lau & Deng, 2010). GA promotes hypocotyls elongation in seedlings grown in the dark by reducing the DELLA proteins levels, thus allowing the repressive transcription factors PIF3 and PIF4 to bind to their targets and promote skotomorphogenesis and repress photomorphogenesis (Lau & Deng, 2010). Another example is the interaction between GA and other hormones to promote inflorescence stem growth (Strabala & MacMillan, 2013).

GA mutations role in genetics and plant breeding

Gibberellins have played an important role in the history of genetics and in crops breeding. One of the traits studied by Mendel was stem length and the gene controlling this trait was termed *Le* (Mendel, 1865; Lester et al., 1997), which was found to encode a gibberellin 3β – hydroxylase (ortholog of *GA4* in *A. thaliana*) (Lester et al., 1997). Modification of the GA pathways was crucial in the green revolution, since it conferred semi-dwarfism and increased crop yields (Hedden, 2003; Salamini, 2003) mainly by avoiding lodging and thus allowing farmers to add more fertilizers and increasing the harvest index (Sasaki et al., 2002; Spielmeyer et al., 2002). In the case of rice the semi-dwarf locus *Semi-Dwarf-1* (*SD1*), which codes for a GA 20-oxidase-2 was targeted (Sasaki et al., 2002; Spielmeyer et al., 2002). Modern barley varieties in

which the mutated gene was called *Denso* or *Sdw1* (Jia et al., 2009) also carry mutations in the *GA 20ox2* encoding gene. In the case of wheat, the GA signalling pathway was targeted in the *Reduced height-1* (*Rht-B1* and *Rht-D1*) green revolution loci (Peng et al., 1999). This semi-dwarf genotypes have a low responsiveness to GA applications and it has been shown that *Rht-B1* / *Rht-D1* and the maize *dwarf-8* (*d8*) (Harberd & Freeling, 1989; Winkler & Freeling, 1994) are orthologs of the *A. thaliana GAI* (Koorneef et al., 1985) signalling pathway (DELLA) gene (Peng et al., 1999). As in the *A. thaliana gai* mutants, *D8* and *Rht* genes carry gain of function mutations, which conferred reduced GA responses (Peng et al., 1997). These mutations are dominant, contrasting the recessive loss of function mutations in the *GA 20-oxidase* genes. The mutations in the *gai* mutants cause semi-dwarfism because it affects the interaction with GA or GA-signals (Peng et al., 1997). Using transgenic approaches semi-dwarfism was also introduced into Poplar, where the *AtGAI* mutant variant and independently the *GA 2-oxidase* (using *Phaseolus coccineum* as gene source) were overexpressed and induced semi-dwarfism (Elias et al., 2012).

Genetic variation for GA biosynthesis and signalling in *Arabidopsis thaliana*.

Many GA biosynthesis and signalling mutants have been isolated in *A. thaliana* providing valuable tools to study this plant hormone and clone the genes controlling the pathway (Koornneef & Veen, 1980; Koorneef et al., 1985; Mitchum et al., 2006; Rieu et al., 2008; Plackett et al., 2012). GA related traits such as plant height and seed germination might help identifying natural variants with mutations in this pathway. In addition to this forward approach also the emerging genomic resources such as the 1001 Genome projects allows the identification of mutations in the now known genes of the pathway. Additional tools include the use of several GA inhibiting compounds, among them paclobutrazol (PAC) (Hedden & Graebe, 1985), ancymidol (Nambara et al., 1992), and tetcyclasis (Debeaujon & Koornneef, 2000), which allowed to the selection of mutants that do not require GAs in specific pathways such as seed germination .

How much natural variation in *A. thaliana* is due to modifications of the well described GA pathway is still unknown. Indications that variation for GA responses is present among natural *A. thaliana* accessions have been reported for germination and

hypocotyl length (Borevitz et al., 2002; van der Schaar et al., 1997). It has been described that the *A. thaliana* accession Bur-0 carries a loss-of-function allele at *GA20ox4* (Plackett et al., 2012) (Figure 1.1), which does not result in a semi-dwarf phenotype. Genetic variation in *GA1* (Figure 1.1) has been related with variation in floral morphology (Brock et al., 2012). El-Lithy et al., (2006) described that semi-dwarfism in the Kas-2 accession is due to a recessive allele at the *GA5* locus (Figure 1.1). The *GA5* gene encodes for a GA 20-oxidase causing semi-dwarfism when mutated (Xu et al., 1995). The *ga5* mutants (Koornneef & Veen, 1980) displays a semi-dwarf phenotype that do not confer obvious detrimental pleiotropic effects, and this phenotype is stable until harvesting time. It is relevant to mention, of all five GA 20-oxidases paralogs, only mutations in *GA20ox1* induce semi-dwarfism (Rieu et al., 2008).

Semi-dwarfism and water stress tolerance

Studies in rice showed that genotypes displaying drought tolerance can be found both in wild type and semi-dwarf background (Lafitte et al., 2007). In Arabidopsis, a lack of GA leading to semi-dwarfism might confer an advantage under water stress conditions as pointed by Vartanian et al. (1994) who showed *ga5* makes a drought (water withholding) stress adapted root system in a much more effective way than its wild type. In this study a higher number of short roots (present in lateral roots as a drought induced response) was observed in the *ga5* mutant compared to wild type and other evaluated genotypes. A link of gibberellin and stress tolerance makes sense in view that when plants are exposed to limited water conditions have to restrict their growth. This regulation is achieved by many environmental factors and also by other plant hormone pathways, e.g. via increasing the level of growth repressing effect of the DELLA proteins, which repressors are suppressed by gibberellins (Achard & Genschik, 2008).

Thesis objective

The main aim of this thesis was to investigate how much genetic variation in the gibberellin biosynthesis and signalling pathways contribute to natural variation in *Arabidopsis thaliana*. One goal was to study the genetic variation of the GA biosynthesis gene (*GA5*) and to obtain indications for selection in specific populations (chapter two). To understand a potential genetic advantage the possible trait-offs that

affect fitness were quantified and related to semi-dwarfism (chapter two). In addition it was aimed to understand the possible physiological consequences of *ga5* natural variants and if pleiotropic effects on other traits including the phenotype of the root system and the response to water limitation experiments are present (chapter three). Finally it was aimed to study the genetic control of a GA inhibitor application for germination (chapter four) and hypocotyl length (chapter five), for which a high throughput semi / -automated Phenotyping was developed (chapter five).

Chapter 2

Arabidopsis semi-dwarfs evolved from independent mutations in *GA20ox1*, orthologue to green revolution dwarf alleles in rice and barley

Abstract

Understanding the genetic bases of natural variation for developmental and stress-related traits is a major goal of current plant biology. Variation in plant hormone levels and signalling might underlie such phenotypic variation occurring even within the same species. Here it is reported the genetic and molecular basis of semi-dwarf individuals found in natural *Arabidopsis thaliana* populations. Allelism tests demonstrate that independent loss-of-function mutations at *GA5*, which encodes a GA 20-oxidase involved in the last steps of gibberellin (GA) biosynthesis, are found in different populations from Southern, Western and Northern Europe, Central Asia and Japan. Sequencing of *GA5* identified 21 different loss-of-function alleles causing semi-dwarfism without any obvious general trade-off affecting plant performance traits. *GA5* shows signatures of purifying selection, while *GA5* loss-of–function alleles can also exhibit patterns of positive selection in specific populations as shown by Fay and Wu's *H* statistics. These results suggest that antagonistic pleiotropy might underlie the occurrence of *GA5* loss-of-function mutations in nature. Furthermore, since *GA5* is the orthologue of rice *SD1* and barley *Sdw1/Denso* green revolution genes, this study illustrates the occurrence of conserved adaptive evolution between wild Arabidopsis and domesticated plants.

Introduction

Bioactive gibberellins (GA) are plant growth regulators involved in important traits such as seed germination, flowering time, anther and petal development, fertility and elongation growth (Hedden & Thomas, 2012). GA biosynthesis and signalling pathways are well defined (Yamaguchi, 2008; Hedden and Thomas, 2012) and have been targeted in crop breeding. Modification of GA pathways was crucial in the green revolution since it conferred semi-dwarfism thus reducing lodging and increasing crop yields (Hedden, 2003; Salamini, 2003). In addition, it allows higher fertilizer applications without detrimental effects on yield, which increased harvest index (Sasaki et al., 2002; Spielmeyer et al., 2002). Green revolution semi-dwarf varieties in wheat are due to mutations in *DELLA* genes while many short straw rice varieties carry a mutation in the *SD1* (*Semi-Dwarf-1*) locus. This locus codes for *GA 20-oxidase-2,* a GA biosynthesis gene that is also mutated in most modern barley varieties in which the gene was called *Denso* or *Sdw1* (Jia et al., 2009).

GA 20-oxidases are involved in the later steps of GA biosynthesis and belong to the group of 2-oxoglutarate-dependent dioxygenases that, together with GA 3-oxidases, form biologically active GA (Rieu et al., 2008). *Arabidopsis thaliana* (hereafter referred to as Arabidopsis) has five *GA20ox* paralogous genes. *AtGA20ox-1, -2, -3 and -4* can catalyze the *in vitro* conversion of GA_{12} to GA_9. Therefore, *GA20ox* paralogs might have partial redundant functions (Plackett et al., 2012). However, among paralog genes, only *AtGA20ox-1* (*GA5*), which was cloned on the basis of the *ga5* mutant (Xu et al., 1995), affected plant height (Rieu et al., 2008).

Natural variation for GA biosynthesis has been previously described in Arabidopsis since the Bur-0 accession carries a loss-of-function allele at *GA20ox4* (Plackett et al., 2012), which does not result in a semi-dwarf phenotype. In addition, genetic variation in *GA1* has been associated with variation in floral morphology (Brock et al., 2012). Furthermore, the semi-dwarf phenotype (here defined as a plant height shorter than half the size of genetically related individuals) observed in the Kas-2 accession, is due to a recessive allele at the *GA5* locus (El-Lithy et al., 2006). This latter finding led to the question whether green revolution alleles artificially selected in cereals could also occur in natural populations of the wild species Arabidopsis; and, if so, how many different *GA5* loss-of-function alleles exist, how are they distributed and why do they occur in some populations.

Results

Identification, characterization and geographic distribution of natural *ga5* alleles.

Phenotypic surveys for plant height in world-wide collections of Arabidopsis accessions detected 97 individuals collected in 23 different locations showing semi-dwarf phenotypes. To determine the genetic basis of semi-dwarfism, allelism tests were carried out by crossing at least one semi-dwarf from each population to the recessive *ga5* (L*er*) mutant (Koornneef & van der Veen, 1980), and to L*er* 'wild type' as comparative control (Figure 2.1A and 2.1B, Table 2.1). In addition, to discard that GA-biosynthesis mutations other than *GA5* could account for the semi-dwarf phenotypes, the complementation of the *ga5* (L*er*) mutant by two other semi-dwarf mutant alleles also affecting GA biosynthesis was tested: *ga4* (L*er*), a mutant in the *GA3ox1* gene and *ga3ox1-3* (Col-0) (Mitchum et al., 2006) (Figure 2.1B, Table 2.1). L*er* and Col mutants were used to test background effects. Control F_1 plants derived from crosses between non-dwarf accessions and *ga5* mutant, as well as F_1 plants grown from crosses with other GA mutants were all taller than their corresponding parents. The crosses *ga5* × *ga4* and L*er* × *ga4* yielded a low height due to the *erecta* mutation which remained recessive in the F_1. In addition, three accessions showing a weaker semi-dwarf phenotype (Nfro -Scandinavia-, Kar -Central Asia- and Vel -Iberian Peninsula-) were not allelic to *ga5*, which indicated that other loci accounted for their plant height phenotype. However, for all the remaining semi-dwarf accessions tested, the F_1 obtained from their cross to *ga5* exhibited the small size of the parents, whereas the semi-dwarfism was lost in the cross with L*er*. This finding confirmed the recessiveness of the semi-dwarf alleles. Therefore, most semi-dwarf accessions were allelic to *ga5*.

To evaluate if there is any general negative pleiotropic effect on plant performance associated with natural *ga5* alleles, several presumably adaptive traits in six wild *ga5* semi-dwarf accessions, as well as in the *ga5* mutants in L*er* and Col genetic backgrounds were measured (Figure 2.2). As expected from the large effect of *ga5* loss-of-function alleles, semi-dwarf accessions showed rather similar plant height. Furthermore, consistent with previous studies (Rieu et., 2008), *ga5* mutants did not differ significantly from their wild-types in the evaluated traits (Figure 2.2). However, natural *ga5* accessions strongly differed in flowering time, branch and silique number, indicating the absence of strong *ga5* effects on these traits but the substantial

contribution from other genes. Therefore, no major trade-off on silique number, assumed to be a proxy for fitness, was found for these naturally occurring *ga5* alleles.

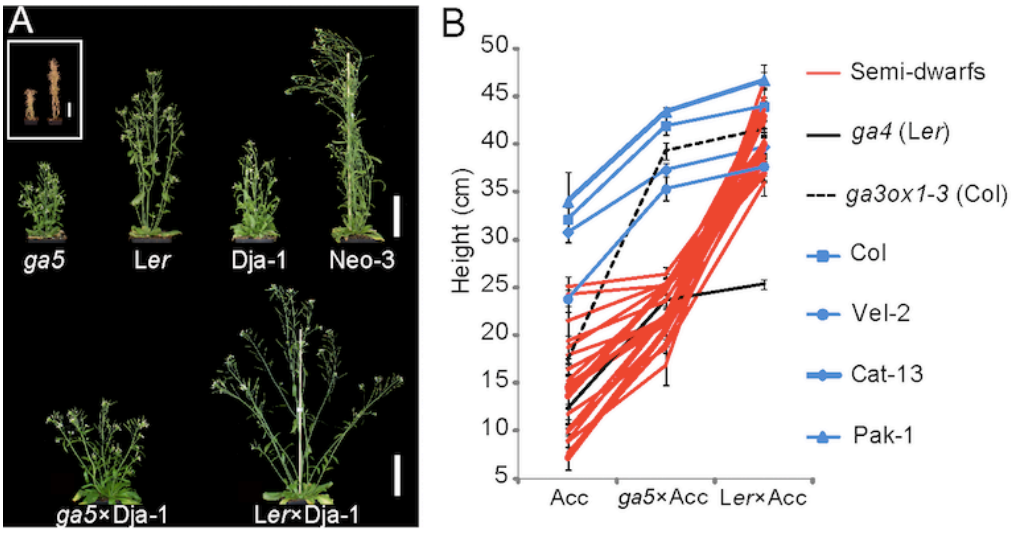

Figure 2.1. Semi-dwarf genotypes allelic to *ga5* are present in nature. (A) Allelism test between the semi-dwarf mutant *ga5* (Koornneef & van der Veen, 1980) and the semi-dwarf central Asian accession Dja-1. Neo-3 (central Asia) shows the phenotype of a functional *GA5*. Pictures were taken two weeks after flowering. On the upper left panel is shown the phenotype of *ga5* and L*er* at harvesting time. Scale bars, 7 cm. (B) Mean values of stem height ± standard errors in F_1 plants derived from crosses between *ga5* or L*er* and twenty accessions (Acc) allelic to *ga5* (red), three non-dwarf accessions (Col-0, Pak-1, and Cat-13), two semi-dwarf mutants (*ga4* and *ga3ox1-3*) and one semi-dwarf accession non-allelic to *ga5* (Vel-2).

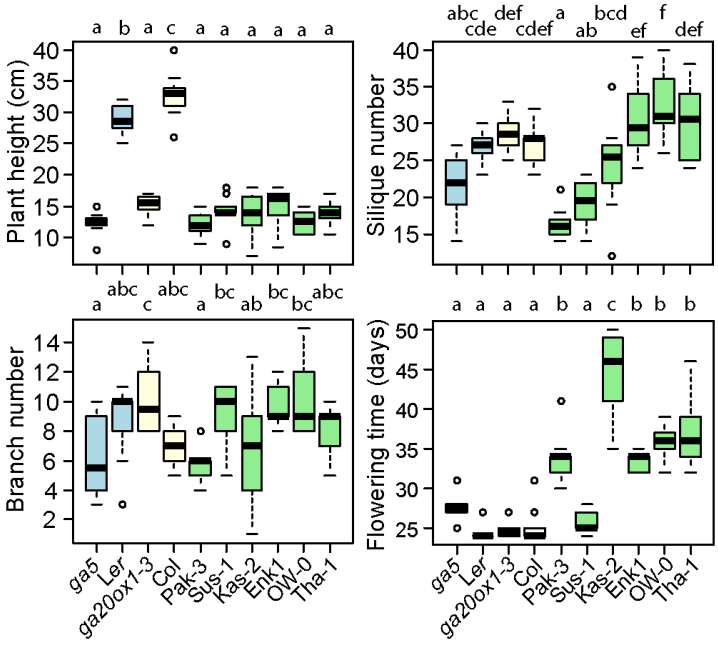

Figure 2.2. Phenotypic characterization of *ga5* semi-dwarf accessions. Figure includes six semi-dwarf accessions, the *ga5* (L*er* background) and *ga20ox1-3* (Col background) mutants and their corresponding wild-type controls. All traits show significant differences across the different genotypes at a P-value < 0.05. The thick horizontal line represents the median, boxes represent the 25th and 75th percentile (lower and upper hinges respectively), vertical lines represent whiskers (0.05th, 0.95th percentile) and open circles, extreme values. The letters above each subfigure indicate the results of a Tukey`s HSD test where means with different letters are significantly different (at P<0.05).

Table 2.1. Populations containing semi-dwarfs allelic to *ga5*. *Hapmap lines (Li et ., 2010). ** Accessions obtained from: CS and N: TAIR/ABRC stock center: http://www.arabidopsis.org/abrc/catalog/natural_accession_4.html. JW: Japanese stock center http://sassc.epd.brc.riken.jp/. NL: Dutch in house collection. KL: Cologne in house collection. NA: no data available

Population	Stock number**	Individuals number./ semi-dwarfs number	Code semi-dwarfs	Full name	Country	Latitude	Longitude	Altitude	Collector
Pak	JW107	3 / 1	Pak-3	Pakistan	Pakistan	33.9	73.4	1100	J.Mirza
Kas	N1264	2 / 1	Kas-2	Kashmir	India	35	77	1580	Sharma
YGU	JW115	1 / 1	YGU	Yamaguchi	Japan	34.18	131.48	50	Todokoro
Kyr	CS76536	8 / 7	Kyr-1, 2, 3, 4, 5, 7, 8	Kyrgyz	Kyrgyzstan	40.01	72.46	2300	O. Loudet
Sus	NA	6 / 6	Sus-1, 2, 3, 5, 6, 7	Susamyr	Kyrgyzstan	42.19	73.41	2400	O. Loudet
Dja	CS76473	9 / 9	Dja-1, 2, 3, 4, 5, 6, 7, 8, 10	Djarly	Kyrgyzstan	42.49	73.63	2550	O. Loudet
Mar	N799750, N799752, N799759	5 / 3	Mar-1, 3, 11	Marjaliza	Spain	39.58	-3.93	896	C. Alonso-Blanco
Cat	NA	25 / 13	Cat-0, 1, 5, 8, 10, 15, 17, 19, 20, 22, 23, 43, 44, 45	Cantoblanco	Spain	40.54	-3.39	713	C. Alonso-Blanco
Mdc	NA	22 / 8	Mdc-10, 53, 60, 87, 100, 113, 122, 130	Moral de Calatrava	Spain	38.52	-3.32	725	C. Alonso-Blanco
Kl	CS1278	2 / 1	Kl-2	Köln	Germany	50.93	6.98	48	Hülsbruch
Ooij	NL617	1 / 1	Ooij-1	Ooijpolder	Netherlands	51.87	5.91	11	M. Koornneef
Tha	CS28758	1 / 1	Tha-1*	The Hague	Netherlands	52.15	4.3	1	M. Koornneef
Haarl-1	NL1167	1 / 1	Haarl-1	Haarlem	Netherlands	52.38	4.64	4	N. Buiten
OW	KL40567	49 / 29	OW-0, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 23, 24, 25, 26, 28, 30, 31, 41, 42, 43, 44, 45, 46, 48, 49	Oude Wetering	Netherlands	52.22	4.64	1	M. Koornneef
Oosthuizen	NL1220	1 / 1	Oosth-1	Oosthuizen	Netherlands	52.57	4.99	1	B. Dekkers
Schar	NL1219	1 / 1	Schar-1	Scharwoude	Netherlands	52.62	5.01	-1	B. Dekkers
Hoorn	NL1237	1 / 1	Hoorn-1	Hoorn	Netherlands	52.64	5.05	1	B. Dekkers
Enk	NL1238	4 / 2	Enk-1, 3	Enkhuizen	Netherlands	52.71	5.28	2	B. Dekkers
Sch	NL183	73 / 6	Sch-183, 184, 185, 186, 187, 188, 193	Schiermonnikoog	Netherlands	53.49	6.24	6	M. Koornneef
Var	CS76298	2 / 2	Var 2-1, 2-6*	Varhallarna	Sweden	55.58	14.33	13	M. Nordborg
T1080	CS76235	1 / 1	T1080*	NA	Sweden	55.66	13.22	60	M. Nordborg
Sparta	CS76229	1 / 1	Sparta-1*	Sparta	Sweden	55.71	13.05	40	M. Nordborg
Veg	NA	1 / 1	Veg 1-2	Vega	Norway	65.71	11.93	3	Anders Bryn

Semi-dwarf *ga5* accessions were found in 23 different populations distributed in Western Europe, the Iberian Peninsula, Scandinavia, Central Asia and Japan (Figure 2.3, Table 2.1). It can be roughly estimated that, at world-wide scale, the frequency of wild populations containing semi-dwarf accessions allelic to *ga5* was at least 1%. However, since most populations segregate for *GA5* loss-of-function alleles, one cannot discard that some populations with a limited number of individuals may contain semi-dwarfs at low frequency not represented in the individuals studied. Therefore, it cannot be excluded that the frequency of wild populations containing semi-dwarfs may be higher than 1%. It was also found a semi-dwarf frequency of 1% in the Hapmap experimental population consisting of 360 world-wide accessions with empirically reduced population structure (Li et al., 2010). However, the frequency of *ga5* semi-dwarf containing populations was not homogeneous throughout the Arabidopsis geographic range since semi-dwarfs were not found among the many Central and East European accessions studied. By contrast, semi-dwarfism appeared most frequent in Central Asia than elsewhere, since 5 out of the 24 central Asian populations monitored in this and another study (Alcázar et al., 2010) carried semi-dwarf individuals (Table 2.1). A ~2% frequency was estimated for the Iberian Peninsula from the qualitative analysis of the intensive collection (Méndez-Vigo et al., 2011) used to select the Iberian accessions included in this study. In addition, detailed sampling and analysis of *ga5* semi-dwarfs in The Netherlands indicated a ~5% frequency in this region. Interestingly, Dutch semi-dwarfs seemed to have spread mainly in the west of the country, although one population was found inland at more than 100 km distance from accessions in the west (Figure 2.3).

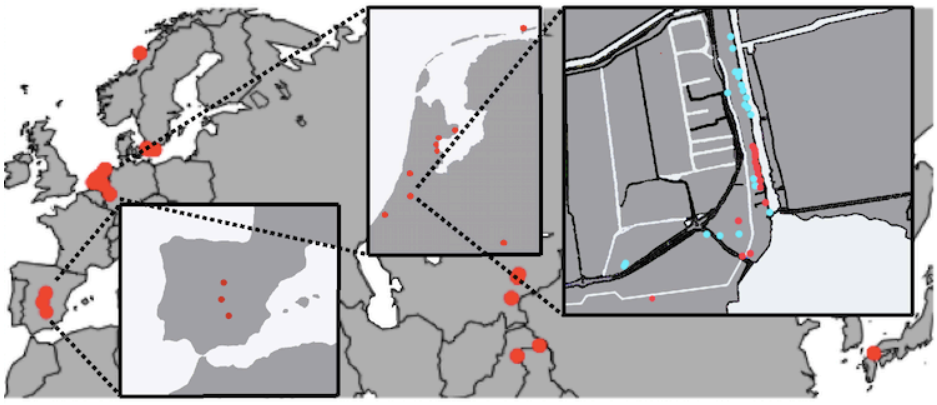

Figure 2.3. Geographical distribution of semi-dwarf accessions in Europe, Scandinavia and Central Asia. Red marks indicate the location of populations containing semi-dwarf accesions allelic to *ga5*. On the right panel it is shown the detailed local distribution of semi-dwarf (red) and wild-type (blue) individuals found in one populations from The Netherlands in 2011 (OW population).

Descriptions of the habitat of populations containing *ga5* semi-dwarf individuals show that they occur in multiple diverse environments where the species occurs. For instance, Dutch dwarf accessions were found in the anthropoid environments where Arabidopsis grows including urban (street populations) and rural (road and field sides, Figure 2.4) habitats. However, in the Iberian Peninsula and Central Asia, semi-dwarfs occurred in more natural environments where samples were collected including Mediterranean forests and mountain wet grasslands (Figure 2.4). This wide geographic and ecological distribution indicates that *ga5* semi-dwarfism does not show a strong geographic structure and is not associated with a single and common climatic factor across its distribution range.

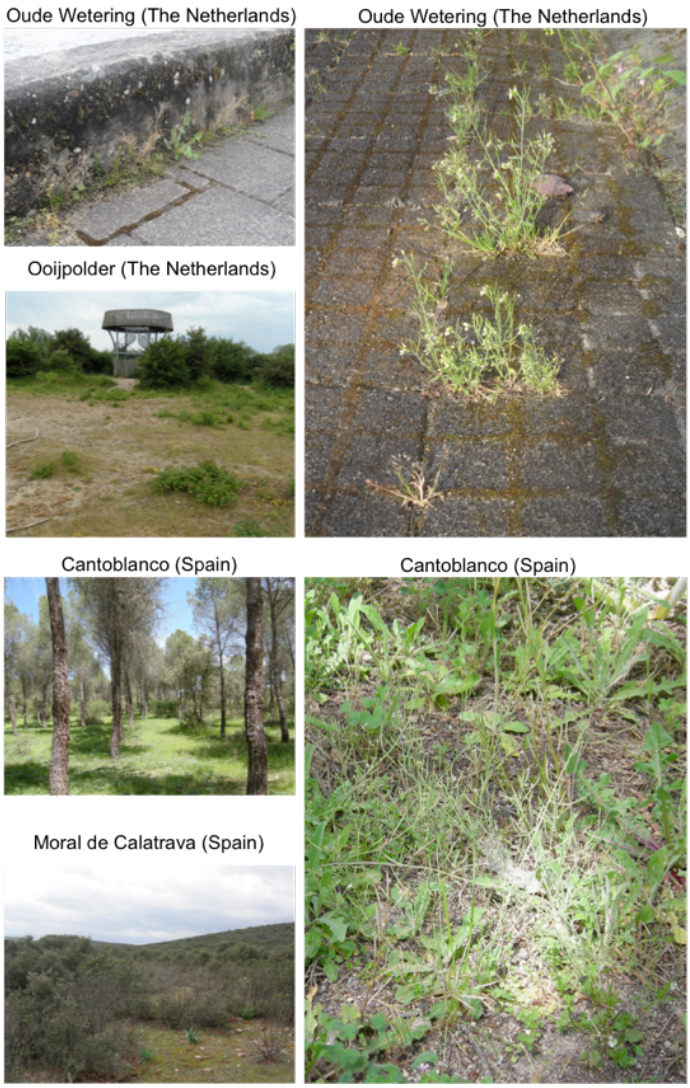

Figure 2.4. Populations of Arabidopsis containing semi-dwarf individuals. Images of semi-dwarf individuals naturally occurring in two locations from The Netherlands (top), and the habitat of Cat and Mdc populations in the Iberian Peninsula (bottom).

Identification of multiple *GA5* loss-of-function alleles.

To determine the putative mutations causing semi-dwarf phenotypes, the *GA5* gene was sequenced (~1.5 kb) in 59 semi-dwarf accessions collected world-wide and 135 non-dwarf individuals, which were collected from the same population or geographic region as the semi-dwarfs identified. For the Dutch OW and Sch populations, the ~1 kb *GA5* region spanning semi-dwarf causal mutations was sequenced in 16 semi-dwarfs and 77 wild-type individuals. Collectively, sequencing data identified 21 different mutations, which were predicted to cause *GA5* loss-of-function alleles in semi-dwarf accessions (Figure 2.5). These mutations were classified in six loss-of-function classes according to their nucleotidic nature. First, non-sense mutations causing premature stops codons were found in Kas-2 and Sparta. Second, missense mutations were found close to the conserved metal binding sites of *GA5* in an Iberian (Mar-1, Mar-3 and Mar-11) and a Scandinavian (Var 2-1 and Var 2-6) population, which might underlie their *ga5* phenotype. Besides, the Mdc-10 and Mdc-53 semi-dwarf accessions also carried missense mutations in GA5 conserved domains. Third, a single substitution in the donor splice site of the first intron was found in all Dutch semi-dwarf accessions. This affects normal *GA5* splicing as confirmed by cDNA sequencing (data not shown), and generates a truncated GA5 protein. Forth, seven small insertions (Cat-0, Dja-1 and Pak-3) or deletions (Cat-17, Cat-23, Cat-43 and Sus-1) were predicted to cause frame-shifts and truncated GA5 proteins. Fifth, a transposon insertion, with high similarity to *At4g04410,* was identified in the MdcA-60 accession. Finally, several large deletions (> 20 bp) were found in some accessions. These included a 29 bp deletion in the first exon of Kl-2 (Germany) and a 444 bp deletion spanning part of the second exon and the complete third exon of accession T1080 (Sweden) (Figure 2.6 and 2.7, Appendix 1). This deletion was first detected by the absence of sequence coverage in the 1001 genomes data (www.1001genomes.org) and further confirmed by extensive PCR amplifications (Figure 2.6 and 2.7, Appendix 1). In addition, large *GA5* deletions of several kb were found in the Veg 1-1, Kyr-2, and YGU accessions. These deletions included not only the coding region but also the promoter (Figure 2.6 and 2.7, Appendix 1) and were associated with absence of *GA5* expression by quantitative RT-PCR in Kyr-2, Veg 1-2 and YGU (data not shown).

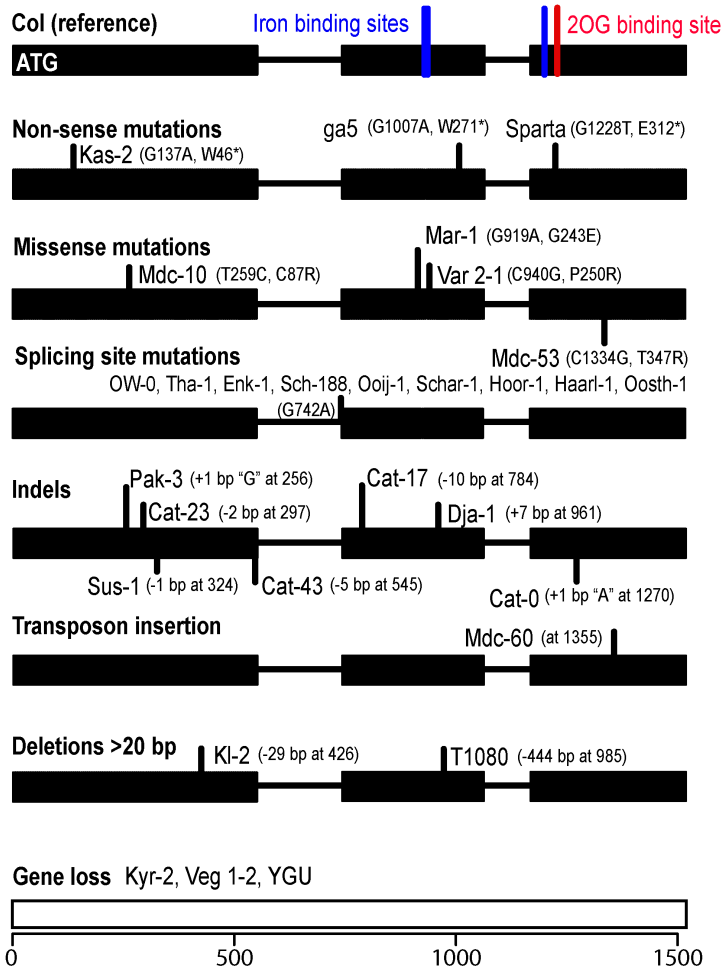

Figure 2.5. Natural loss-of-function mutations in the *AtGA20ox1* (*GA5*) gene. The different nature and position of mutations causing *GA5* loss-of-function alleles are shown in each panel. Exons (black boxes) are connected with horizontal lines representing intronic regions of *GA5*. Iron and 2-oxoglutarate binding sites Wilmouth et al., (2002) are indicated on top.

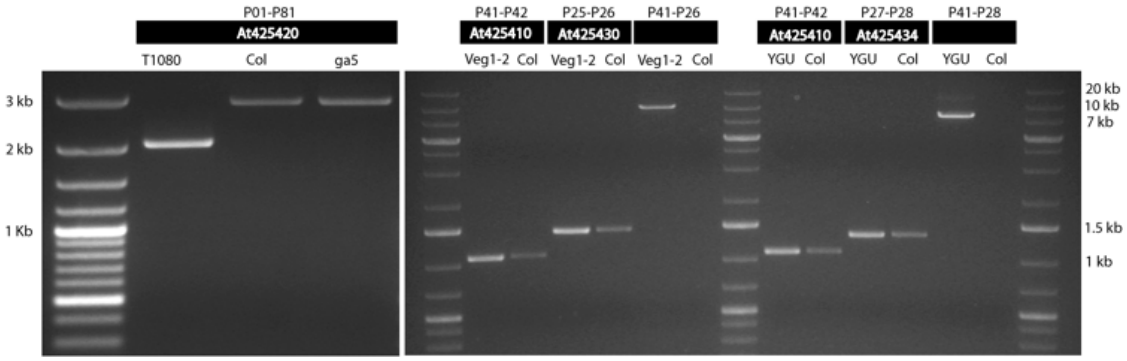

Figure 2.6. PCR amplification analyses spanning the *GA5* locus and neighbour genes identified large *GA5* deletions in the semi-dwarf accessions T1080, YGU and Veg1-2. The primer combinations used in each PCR reaction are indicated on top. The expected amplicon sizes for the primer combinations P41-P26 (middle panel) and P41-P28 (right panel) are 13.9 kb and 19.5 kb, respectively.

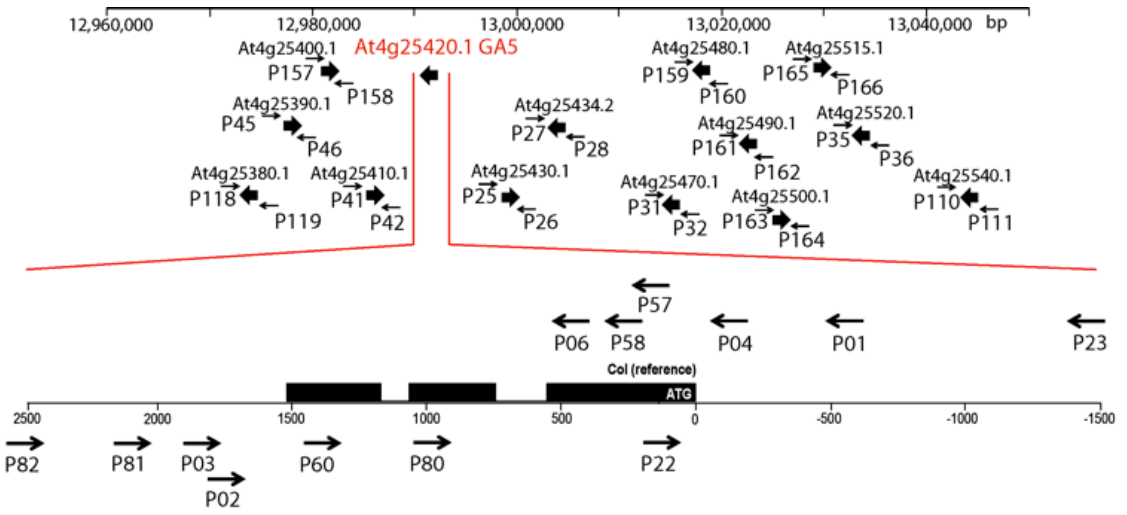

Figure 2.7. Primer positions used to identify *GA5* deletions in T1080, YGU, Kyr-2 and Veg 1-2. Thin arrows represent primers. Thicker arrows indicate different genes upstream and downstream *GA5* that have been tested for amplification.

Sequencing analyses indicated that most populations containing semi-dwarf individuals carry a single loss-of-function mutation in all dwarf plants (e.g. OW-0 in Figure 2.5). However, two Iberian populations (Cat and Mdc) segregated for four independent *GA5* loss-of-function mutations (Figure 2.5). One allele appearing more frequently as it was present in eight Cat individuals out of 22 sequenced samples. On the other hand, most *GA5* loss-of-function alleles appeared distributed in a single wild population, with the exception of the splicing site mutation widely distributed across The Netherlands. Analysis of the sequence data from the 1001 genomes project (www.1001genomes.org) detected four other putative semi-dwarf accessions from South Sweden (Sim-1, TV-22, TV-30 and TV-7), as they carry the Var-2-1 missense mutation. This result suggests that Var-2 missense loss-of function allele might be widely distributed at a local scale since Var, Sim and TV accessions originate from the same South-Swedish coastal area (data not shown).

Genome Wide Association Study (GWAS) for plant height.

Since several of the *ga5* semi-dwarf accessions identified in this study (Tha-1, Sparta, Var 2-1 and T1080) were included in the Arabidopsis Hapmap experimental population (Li et al., 2010), it was tested if the *GA5* locus could be detected by GWA mapping. Measurements of plant height in 345 accessions of this collection showed a large amount of natural variation and high broad sense heritability (h_b^2=0.80) (Figure 2.8A). However, no marker was significantly associated (P>0.05 with Bonferroni

correction for 214,000 markers) with plant height, the largest association was detected on chromosome 4, ~0.3 Mb away from *GA5* (P=3x10^{-5}; Figure 2.8B). Analysis of Linkage Disequilibrium (LD) showed a complete LD decay 10 kb upstream and downstream of *GA5* (Figure 2.3C), thus excluding the linkage of the observed association with *GA5*. By contrast, a significant association was detected when all four *GA5* loss-of-function alleles were combined as a single non-functional haplotype (P= 2.7×10^{-14}). Therefore, despite the strong effect of natural *GA5* loss-of-function alleles on plant height, GWAS was unable to detect this locus, due to the low frequency of semi-dwarf accessions and their multiple independent causal mutations.

***GA5* phylogeny and population structure.**

The genetic relationships among the semi-dwarf accessions was determined using a structure analysis with 117 genome-wide SNP markers already available (Lewandowska-Sabat et al., 2010; Platt et al., 2010) or developed in this work. Structure analysis of these accessions found five distinct genetic groups that closely corresponded to the geographic regions of origin of the semi-dwarf accessions (Figure 2.8D) in agreement with the strong global geographic structure described in Arabidopsis (Platt et al., 2010). In all cases, semi-dwarf accessions were genetically more related to the non-dwarf individuals from the same population and region than to any other accession, indicating the independent origin and expansion of semi-dwarfs in these regions. In most populations containing *ga5* semi-dwarfs where five or more individuals were collected, wild-type *GA5* alleles were found within the population except for the Central Asian populations Dja and Sus, in which all individuals were semi-dwarf. Interestingly, Dja-1 and Sus-1 accessions carried different *GA5* loss-of-function alleles (Figure 2.8D) regardless of the overall low genetic variation present in Central Asia (Cao et al., 2011). It is also remarkable that different *GA5* loss-of-function alleles were found in the Iberian Cat and Mdc populations together with wild-type alleles (Figure 2.8D). In contrast, semi-dwarf genotypes in Dutch populations were very similar and carried the same loss-of-function mutation.

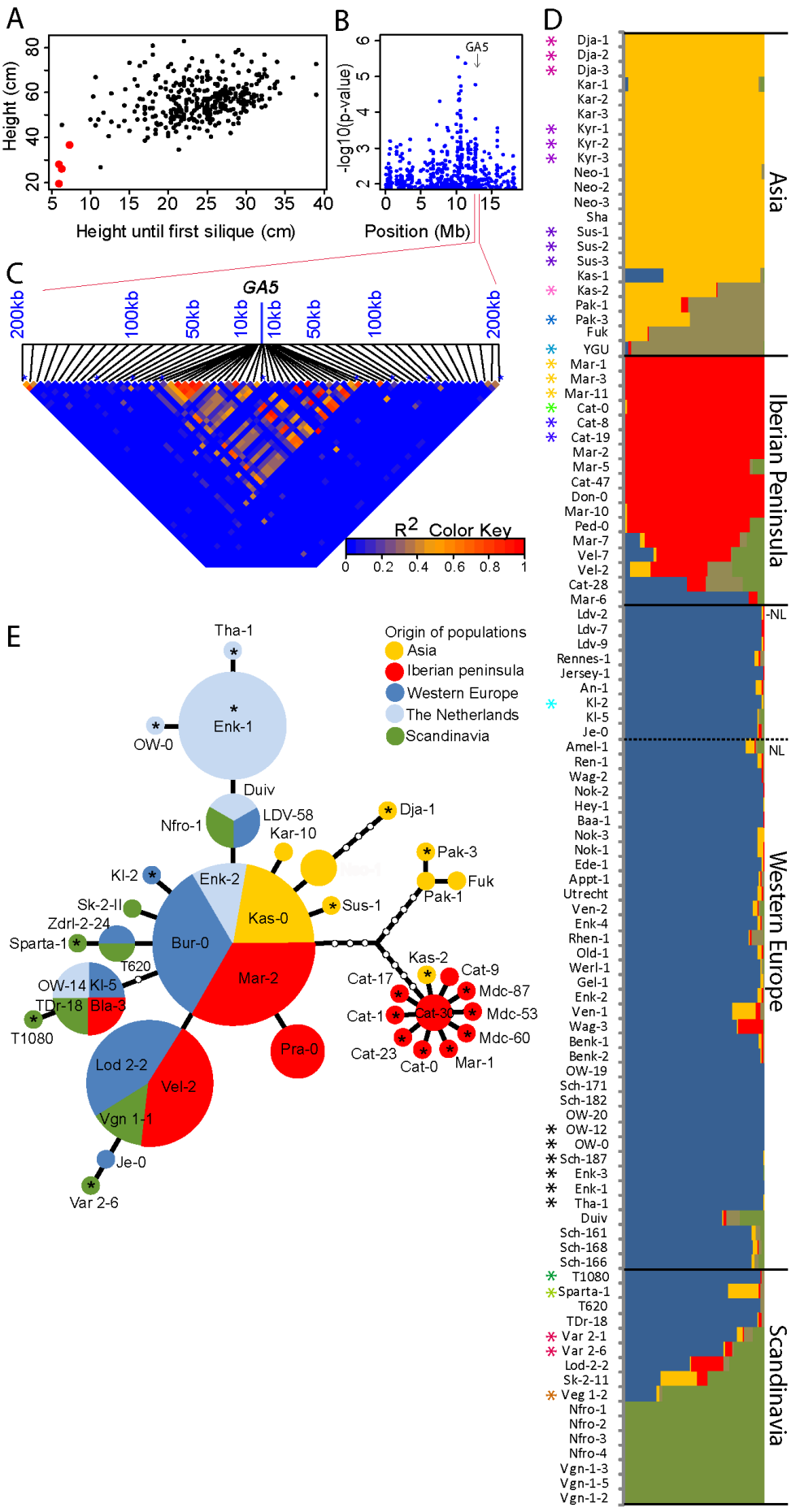
A
Height (cm)
Height until first silique (cm)
B
-log10(p-value)
GA5
Position (Mb)
C
GA5
200kb
100kb
50kb
10kb
10kb
50kb
100kb
200kb
R2 Color Key
0 0.2 0.4 0.6 0.8 1
D
Asia
Iberian Peninsula
Western Europe
Scandinavia
NL
NL
Dja-1
Dja-2
Dja-3
Kar-1
Kar-2
Kar-3
Kyr-1
Kyr-2
Kyr-3
Neo-1
Neo-2
Neo-3
Sha
Sus-1
Sus-2
Sus-3
Kas-1
Kas-2
Pak-1
Pak-3
Fuk
YGU
Mar-1
Mar-3
Mar-11
Cat-0
Cat-8
Cat-19
Mar-2
Mar-5
Cat-47
Don-0
Mar-10
Ped-0
Mar-7
Vel-7
Vel-2
Cat-28
Mar-6
Ldv-2
Ldv-7
Ldv-9
Rennes-1
Jersey-1
An-1
Kl-2
Kl-5
Je-0
Amel-1
Ren-1
Wag-2
Nok-2
Hey-1
Baa-1
Nok-3
Nok-1
Ede-1
Appt-1
Utrecht
Ven-2
Enk-4
Rhen-1
Old-1
Werl-1
Gel-1
Enk-2
Ven-1
Wag-3
Benk-1
Benk-2
OW-19
Sch-171
Sch-182
OW-20
OW-12
OW-0
Sch-187
Enk-3
Enk-1
Tha-1
Duiv
Sch-161
Sch-168
Sch-166
T1080
Sparta-1
T620
TDr-18
Var 2-1
Var 2-6
Lod-2-2
Sk-2-11
Veg 1-2
Nfro-1
Nfro-2
Nfro-3
Nfro-4
Vgn-1-3
Vgn-1-5
Vgn-1-2
E
Origin of populations
Asia
Iberian peninsula
Western Europe
The Netherlands
Scandinavia
Tha-1
Enk-1
OW-0
Duiv
Nfro-1
LDV-58
Kar-10
Dja-1
Pak-3
Fuk
Pak-1
Sus-1
Kas-0
Kl-2
Enk-2
Sk-2-II
Zdrl-2-24
Sparta-1
Bur-0
T620
OW-14
Kl-5
TDr-18
Bla-3
T1080
Mar-2
Pra-0
Kas-2
Cat-9
Cat-17
Mdc-87
Cat-1
Cat-30
Mdc-53
Mdc-60
Cat-23
Cat-0
Mar-1
Lod 2-2
Vel-2
Vgn 1-1
Je-0
Var 2-6

Figure 2.8. GWAS analyses, population structure and *GA5* diversity. (A) Correlations between height and height up to first silique. Red dots indicate the values from semi-dwarf accessions. (B) Genome wide association mapping profile for plant height on chromosome 4. The *GA5* position is indicated by an arrow. (C) Linkage disequilibrium 200 kb up and downstream of the *GA5* locus. The heat colour scale represents squared correlation (R^2) between pairs of SNPs. (D) Population structure of 100 accessions including non-dwarf and *GA5* semi-dwarfs collected in different world regions at K=5. Colored asterisks indicate accessions carrying different *GA5* loss-of-function alleles. (E) *GA5* haplotype network. Haplotypes are represented by circles with size proportional to the number of populations containing that haplotype. Each node represents a single mutation.

Network analysis of the 33 different *GA5* haplotypes detected within the genomic *GA5* sequence identified a common *GA5* functional haplotype which showed a world-wide distribution (Figure 2.8E and Appendix 2). Twenty other *GA5* haplotypes were connected to this frequent haplotype by fewer than five mutational steps and were distributed in all geographic regions. The central network position of the most frequent haplotype suggests that this is the oldest *GA5* allele, from which most other haplotypes may have derived by a small number of mutations (Figure 2.8E). Furthermore, 14 additional low frequency haplotypes, which include only Iberian and Asian haplotypes (Cat, Mdc, Mar, Kas, Pak and Fuk), were separated from the main node of the network by two long related branches. Loss-of-function *GA5* haplotypes appeared evenly distributed within this network, and all but one of these alleles was connected by a single mutational step to their presumably ancestral haplotype. In addition, all loss-of-function haplotypes occupied branch-end positions in this network but the Dutch containing semi-dwarfs haplotypes. Therefore, most of independent *GA5* loss-of-function alleles seem to be generated in multiple genetic backgrounds but they have not produced derived haplotypes (Figure 2.8E).

Signatures of selection at the *GA5* locus.

To estimate the amount and pattern of nucleotide diversity in the *GA5* gene it was analyzed the SNP data from 512 accessions available from the 1001 genomes project (www.1001genomes.org). *GA5* shows lower nucleotide diversity within coding regions than introns (Figure 2.9A). Total nucleotide diversity (π = 0.0017, Table 2.2) was lower than the average nucleotide diversity reported in previous studies (0.0081 for centromeric and 0.0059 for non centromeric regions, Schmid et al., 2005). *GA5* also presents a low ratio of non-silent to silent polymorphism ($\pi(ns)/\pi(s)$=0.132), which is consistent with a signature of purifying selection, as previously suggested for rice GA biosynthesis genes (Yang et al., 2009). In addition, significant negative values for Tajima's D at non-synonymous sites (D_n) were detected in both the aforementioned

512 accessions (D_n=-2.289 p<0.01), as well as in the more than 100 accessions used in the present study (D_n=-1.987 p<0.05) including semi-dwarf haplotypes (Table 2.2). Overall, this pattern is compatible with the occurrence of purifying selection, in which polymorphisms leading to amino acid substitutions are maintained at low frequencies.

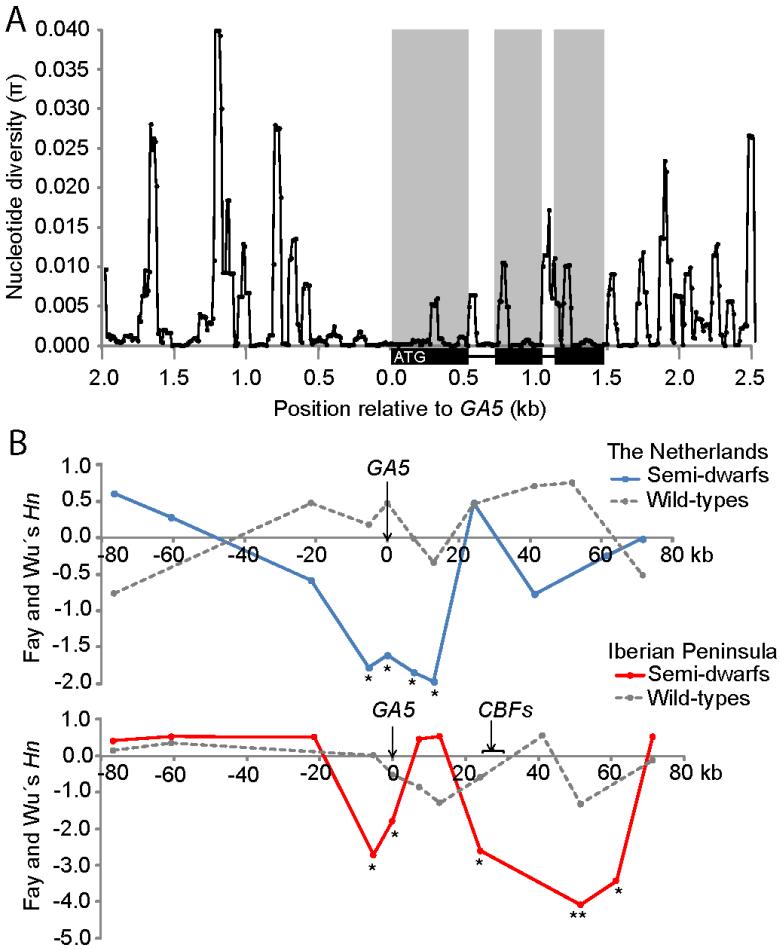

Figure 2.9. The *GA5* locus shows signatures of natural selection. (A) Nucleotide-sliding window analysis of nucleotide diversity (π) across the *GA5* locus in 505 *Arabidopsis* wild accessions. (B) Fay and Wu's *Hn* analysis across the *GA5* genomic region in populations containing semi-dwarfs from The Netherlands (blue), Iberian Peninsula (red), and populations of normal size (grey). Asterisks denote statistical significance *P<0.05, ** P<0.01.

Table 2.2. Neutrality tests conducted for the *GA5* locus. Tests were conducted in 505 accessions (data from 1001 genomes project) and in 100 additional accessions including semi-dwarf haplotypes studied in this work.

Source	Tajima's D	D Nonsyn	D Syn	D Silent	Fu and Li's D	Fu and Li's F	π	Hn
1001 genomes (512 acc)	-1.891*	-2.289**	-1,316	-1.269	-4.608**	-4.026**	0.00171	-1.559
Eurasian populations (100 acc)	-0.997	-1.978*	-0.11	-0.127	-0.429	-0.79	0.0037	-0.86

* P <0.05

** P <0.01

To test if positive selection may have contributed to an increase of *GA5* loss-of-function alleles, molecular fingerprints of recent selective sweeps were searched over a region of 80 kb upstream and downstream of *GA5* in two populations from two different regions. These Cat (Iberian Peninsula) and Ow/Sch (The Netherlands) populations were selected because they contain a moderate frequency of *GA5* loss-of-function alleles. One additional population that does not contain semi-dwarf individuals from each of the regions was analyzed as control. Significant negative values of the normalized Fay and Wu's *H* statistics were found around *GA5* in Cat and Ow/Sch populations containing semi-dwarfs ($0.019<P<0.05$) (Figure 2.9B), which is consistent with an excess of derived high-frequency mutations that commonly accompanies selective sweeps. Negative values for the Fay and Wu's H_n statistics in the semi-dwarf Iberian Peninsula population were detected around the *CBF* cluster involved in cold acclimation, for which natural variation has been reported (Figure 2.9B) (Alonso-Blanco et al., 2005). This pattern was absent in populations without semi-dwarfs from the same regions (Figure 2.9B). These results suggest that positive selection might contribute to increase the frequency of *GA5* loss-of-function mutations under particular environments, although drift and relaxed purifying selection could also contribute to a high frequency of *GA5* loss-of-function alleles in some other populations.

Discussion

In this study it was shown that Arabidopsis semi-dwarf genotypes are relatively frequent in natural populations of different regions in the world. Allelism tests demonstrate that this extreme plant height phenotype is mainly determined by multiple independent loss-of-function mutations of large effect on the GA biosynthesis gene *GA5*/*GA20ox1*. The semi-dwarf phenotype of ~20% semi-dwarf accessions studied in this work cannot be explained by *GA5* loss-of-function mutations, indicating the contribution of other genes to this trait. These results evidence a rather simple genetic basis for plant height, but its multi-allelic bases hampered *GA5* detection by GWAS mapping. Interestingly, *GA5* behaves as a functional orthologue of the green revolution genes of rice *SD1* and barley *Sdw1/Denso*. This result points to a conserved evolution for this common trait in crop and wild plant species that have been artificially selected

during domestication or naturally evolved. Thus, GA 20-oxidase is identified as a hotspot for phenotypic variation in plants (Martin & Orgogozo 2013), and illustrates the usefulness of the analysis of Arabidopsis natural variation to find genes of interest for plant breeding. The observation of major phenotypic changes caused by a large number of independent mutations resembles the situation found for the *FRIGIDA* gene of Arabidopsis involved in flowering time, another adaptive trait, which indicates that this pattern is not unique but rather common (Méndez-Vigo et al., 2011). As previously reported for *FRI* and *FLC* flowering genes, most *GA5* haplotypes show a sub-regional or local distribution, but the number of independent functional alleles was significantly larger in the Iberian Peninsula than in northern and central Europe, in agreement with the overall larger Iberian diversity (Mendez-Vigo et al., 2011; Cao et al., 2011; Picó et al 2008).

Despite that it is not known how Arabidopsis *GA5* loss-of-function alleles are maintained in nature, this study supports that different evolutionary forces might contribute to it. The relatively high frequency of several *GA5* loss-function alleles in Central Asia and within some local populations of different world regions (The Netherlands, Central Asia and Iberian Peninsula) suggests an advantage or neutrality in these populations. This is especially the case in some populations where multiple mutations have occurred and are still present. The wide geographic distribution of the same *GA5* allele found in many locations of The Netherlands separated more than 100 km indicates that this allele is spreading, further indicating the absence of deleterious effects. In addition, phenotypic characterization of *GA5* semi-dwarf accessions did not detect any strong negative effect on adaptive and fitness traits, which suggests that these alleles do not display any general obvious negative pleiotropic effect or trade-off. This result is in agreement with the phenotypes described for artificially induced *GA5* loss-of-function mutants, which show similar seed yield than wild-type accessions (Rieu et al., 2008). This lack of effect on seed production is probably due to expression of GA20ox paralogues, mainly *GA20ox2* (Rieu et al., 2008). Similarly, *GA20ox2* mutations in rice and barley do not display trade-offs (Sasaki et al., 2002; Spielmeyer et al., 2002; Jia et al., 2011). By contrast, mutations in early steps of GA biosynthesis have been associated with negative pleiotropic effects, such as the absence of seed germination shown by *ga1* null mutants or the reduced fertility and altered flower development observed even in leaky *GA1* alleles (Koornneef & van der Veen, 1980).

A similar situation has been reported in rice where the effects derived from mutations on genes involved in early steps of GA biosynthesis were found less favorable for crop production compared with mutations on rice *GA20ox2* (Itoh et al., 2004).

As suggested in this study both negative and positive selection may act on *GA5* loss-of-function alleles. The conditional negative effect of these alleles is suggested by the low frequency of most loss-of-function alleles, and by the fact that they are not maintained long enough to derive new haplotypes. Hence, such alleles seem to be only transiently maintained in nature. In addition, such potential negative effect of *GA5* loss-of-function alleles is also suggested by purifying selection inferred from the low ratio of replacement to silent polymorphisms and negative Tajima's D_n values, in agreement with previous reports in rice (Yang et al., 2009). In contrast, positive selection might contribute to transient increases in the frequency of loss-of-function alleles in certain populations, as suggested by the negative values of Fay and Wu's H_n tests across the *GA5* locus for the two tested populations segregating for semi-dwarf individuals. Remarkably, this pattern is absent in populations harboring only normal size plants from the same regions. Therefore, it is reasoned that allelic variation at *GA5* locus might be maintained in nature by antagonist pleiotropy, (*i.e.* reversed fitness effects in different environments) (Anderson et al., 2013). However, one cannot discard that *GA5* variation shows conditional neutrality in other populations (*i.e.* loss-of-function alleles might be neutral in some environments but deleterious in others). Neutrality tests should be considered carefully due to the complex demographic history of Arabidopsis populations in the wild. Furthermore, the population genetic analysis is also agnostic to the local extinction or re-colonization dynamics of *A. thaliana* populations. Remarkably, the identification of signatures for selection using genome-wide screens may be hampered by the occurrence of different loss-of-function *GA5* alleles under positive selection, a situation that also affected GWAS mapping.

It remains to be determined which are the environmental cues that could contribute to an increase in the frequency of *GA5* loss-of-function alleles since these mutations appear distributed in a wide range of anthropoid and natural environments. It has been previously shown that the short plant height phenotype caused by the well-known *erecta* loss-of-function mutation provides fitness advantage in static landscapes. On the contrary, *erecta* frequency was reduced under disturbed environments (Fakheran et al., 2010). Analogously, it can be speculated that environmental stability might favor *GA5* semi-dwarf individuals. Conclusive

demonstration about positive, negative or neutral fitness effects of *GA5* loss-of-function alleles depending on the environment will require further analyses under different natural conditions to elucidate the evolutionary forces driving *GA5* variation and its ecological significance.

Conclusion

Natural variation for GA biosynthesis is reported in numerous Arabidopsis wild populations across the species distribution range. Multiple *GA5* loss-of-function alleles are found underlying most semi-dwarf accessions, while these alleles showing strong effect on plant height do not display any obvious general trade-off. Frequencies and patterns of nucleotide variation suggest that loss-of-function alleles might be under positive and purifying selection. In addition, the common genetic basis of this extended plant height variation in Arabidopsis, rice and barley indicates conserved adaptive evolution of GA mediated plant height phenotypes in wild and domesticated species.

Materials and methods

Plant Material and growth conditions

Stock numbers and detailed information of accessions used in this work are listed in Table 2.1. For allelism tests, semi-dwarf accessions were crossed with L*er* and *ga5* (Koornneef and van der Veen, 1980). To facilitate the allelism tests, male sterility based on the *ms1* mutant (van der Veen and Wirtz, 1968) was introgressed into the *ga5* background. Plants were grown under greenhouse conditions at 16 h light, 22°C/18ºC day/night cycles. For all experiments, seeds were stratified in water at 4°C for 4-6 days prior to germination. Ten repetitions per genotype (cross) were conducted. The Ooij, Schar, Hoor, Haarl, and Oosth Dutch semi-dwarf populations and the Mdc Iberian semi-dwarf population were found in the course of this study and allelism was concluded based on sequence data that correlated with the semi-dwarf phenotypes and haplotype that were tested before in allelism tests.

Phenotyping for plant height and height up to first silique was conducted two weeks after flowering because both traits did not change after that date (data not shown). In cases of extreme flowering lateness, plants were vernalized for six weeks.

Flowering time was recorded as days after germination until the first opened flower. Branch number was scored as the number of axillary stems grown from the rosette.

Sequencing of *GA5* gene and genotyping

Genomic DNA was isolated from leaf material using the BioSprint workstation (Qiagen). Primers used for *GA5* sequencing are detailed in Appendix 3. PCR reactions were performed using LA Taq DNA polymerase (Takara) following manufacturer's instructions. Sanger sequencing of purified PCR products was made by the Max Planck Genome Center Cologne. GenBank accession numbers of DNA sequences generated in this work are listed in Table 2.2.

SNP genotyping of new accessions collected in this study was done as described in previous works (Lewandowska-Sabat et al., 2010; Platt et al., 2010) using the genotyping facility service of the University of Chicago.

Statistical analysis

Descriptive statistics, t-tests, tukey test, and principal component analysis were conducted with R. The method of Emma was used for GWAS (Kang et al., 2010) using kinship matrix to correct for population structure. Linkage disequilibrium analysis was performed with the R package LD heatmap (Shin et al., 2006).

Structure analysis

Population structure was inferred using model-based clustering algorithms implemented in the software STRUCTURE , using the haploid setting and running 20 replicates with 50,000 and 20,000 MCMC iterations of burn-in and after-burning length, respectively (Pritchard et al., 2000). To determine the K number of significantly different genetic clusters, the ΔK method was applied in combination with the absolute value of ln P(X | K) (Evanno et al., 2005).

Population genetics. Fay and Wu's H statistics and haplotype network

Population genetics analyses were conducted with the software DnaSP (5.10) (Librado & Rozas, 2009). The normalized Fay and Wu's H was performed as described (Alcázar et al., 2010) in populations containing semi-dwarfs from The Netherlands and Iberian Peninsula (Table 2.3). Representative accessions of different populations from

Central Spain, with no semi-dwarfs, were used as control (Table 2.3). For the Dutch control population, accessions from a rural area northeast of Wageningen were collected with no prior knowledge on evidences for semi-dwarfism occurring in this population (Table 2.3). The sequences of *GA5* (*At4g25420*) and flanking genes (Appendix 3) were obtained after specific PCR amplification from genomic DNA and sequencing in ABI 3730XL automated sequencers (Applied Biosystems). Sequences were aligned with ClustalW (Thompson et al., 1994) and manually inspected. *Arabidopsis lyrata* sequences were obtained by BLAST search (http://www.phytozome.net/) and used as out-group to assign ancestral and derived states to SNP variants. To assess the statistical significance of Fay and Wu's H, 10.000 coalescent simulations were computed in DnaSP v.5.10 (Librado & Rozas, 2009). The haplotype network of *GA5* was constructed using TCS1.21 (Clement et al., 2000) that implements a maximum parsimony method and excluding gaps as events in the analysis. Insertions and deletions in the semi-dwarf accessions were considered as single events and added manually to the haplotype network.

Table 2.3. Populations used for the Fay and Wu's H_n statistics.

Population	Genotypes
Dutch population containing semi-dwarfs	Duiv, Enk-1, Enk-2, OW-4, OW-7, OW-12, OW-14, OW-16, OW-19, OW-45, Sch-123, Sch-166, Sch-177, Sch-187, Sch-217, Tha-1
Dutch control population	Benk-6, Benk-08, Mir-12, Oost-25, Oost-32, Panh-2-4, Roe-02, Wol-01, Wol-06, Wol-20, Wol-26, Wag-58c, Wag-59, Wag-60a, Wag-74d, Wag-74j
Iberian population containing semi-dwarfs	Cat-0, Cat-1, Cat-5, Cat-9, Cat-13, Cat-17, Cat-19, Cat-20, Cat-22, Cat-23, Cat-27, Cat-28, Cat-30, Cat-33, Cat-35, Cat-39, Cat-43, Cat-44, Cat-45, Cat-47
Iberian control population	Agu-0, Cdc-01, Cdc-03, Cdc-04, Cdc-6, Cdc-08, San-0, San-04, San-05, San-08, San-12, San-13, Pra-0, Pra-01, Pra-08, Pra-10

Chapter 3

Does semi-dwarfism have a pleiotropic effect on root related traits that may contribute to a selective advantage under reduced water availability?

Abstract

The occurrence of natural semi-dwarf *Arabidopsis thaliana* accessions displaying reduced plant height, opened the question whether semi-dwarfism has a pleiotropic effect on other traits such as rooting depth that may contribute to a selective advantage under specific growth conditions. To answer this question, different shoot and root growth related traits were studied *in vitro* and with plants grown in soil-filled rhizotrons or pots. It was used a panel of GA biosynthesis mutants, selected semi-dwarf accessions, and derived F1 populations. Mutations in early steps of the GA biosynthesis pathway led to a reduction in shoot as well as root size. Mutations at the *ga5* locus resulted in plants with decreased root length in comparison to related wild types depending on the genetic background. Interestingly, the semi-dwarf accession Pak-3 showed the longest root system depth both *in vitro* and in soil cultivation experiments during initial growth stages (12 – 26 days after sowing). However, this long root system is independent from the *ga5* loss of function allele as shown by co-segregation analysis. When the same natural semi-dwarf accessions were grown under water limiting conditions, phenotypic differences were relatively small and not associated with functional or inactive *ga5* alleles. Remarkably the semi-dwarfs Pak-3 and Kas-2 accessions showed lower growth reduction when exposed to water limiting conditions relative to control conditions than other lines.

Introduction

Bioactive gibberellins (GA) are plant growth regulators responsible for the expression of several plant traits (Sun, 2008). Their biosynthesis and signalling pathways are well understood (Hedden & Thomas, 2012; Yamaguchi, 2008). GA-related mutations have played an important role in crop breeding where mutations in its signalling and biosynthesis pathways induced semi-dwarfism. Thereby, mutants contributed to yield increases in wheat and rice leading to the green revolution (Hedden, 2003; Salamini, 2003) especially by conferring lodging resistance. Semi-dwarf mutants that were selected by plant breeders had no negative pleiotropic effects (Jia et al., 2011; Rieu et al., 2008; Sasaki et al., 2002; Spielmeyer et al., 2002) on yield-related traits and may have positively acting pleiotropic effects that might be more difficult to detect such as effects on root systems.

Root growth is regulated by plant hormones and among these the role of auxins have been studied in detail (Overvoorde, Fukaki, & Beeckman, 2010). Auxin gradients control the length of the primary root, number of lateral root primordial and response to gravity (Overvoorde et al., 2010). A recent review showed that GA can also regulate root elongation and thickening (Gou et al., 2010; Tanimoto, 2012). The author describes examples showing the sensitivity of shoots, hypocotyls and roots to GA and GA inhibitors and provides arguments supporting the view that roots have a higher GA compared to shoot sensitivity. In *A. thaliana* the application of GA to the shoot increases elongation of primary roots (Bidadi et al., 2010). Interestingly, this study describes that application of GA inhibitors to the shoot also increases root elongation, pointing to a negative feedback mechanism controlling the GA levels (Yamaguchi, 2008). Studies in *Populus* show that GA deficient mutants (by increasing the expression of a *GA2ox1*) and GA insensitive mutants (by overexpressing *RGL1*) increase lateral roots density and elongation and that crosstalk with the auxin hormone pathway may occur (Gou et al., 2010).

Several naturally occurring *A. thaliana* accessions carrying non-functional *ga5* alleles have being identified (Barboza et al., 2013). The *GA5* gene encodes a GA 20-oxidase (a GA biosynthesis gene) causing semi-dwarfism when mutated (Koornneef & van der Veen, 1980; Xu et al., 1995). The occurrence of semi-dwarfism in nature, which shows positive selection in specific populations, may indicate that semi-dwarfism confers an advantage under specific environmental conditions. *GA5* is the functional ortholog of the *SD1* gene, mutations of which confer semi-dwarfism in

modern rice cultivars and of *Sdw1/Denso* for which mutations are present in modern semi-dwarf barley varieties. There are indications that semi-dwarf barley genotypes may have a longer root system compared with non-dwarf ones (White et al., 2009). A study by Vartanian et al. (1994) using *A. thaliana* showed that the *ga5* mutant makes a drought (measured in water withholding pot experiments) stress adapted root system in a much more effective way (higher number of short lateral roots present in lateral roots) than its wild type, thus suggesting a possible role in drought tolerance. Modifications in root growth can have an impact on the performance of plants under stress conditions and genetic variation in the hormonal pathways may affect root traits (Ghanem et al., 2011). Recent studies showed the important role of the *DEEPER ROOTING 1* (*DRO1*) mutation in rice which increased rooting depth (more than twice in Dro1-NIL compared with its genetic background IR64) and conferred drought tolerance (Uga et al., 2013). A link, especially of gibberellin and stress tolerance, makes sense in view that when plants are exposed to limited water conditions they have to restrict their growth. This regulation is achieved by many environmental factors and also by other plant hormone pathways, e.g. via regulation of the level of growth repressing effect of the DELLA proteins, of which repressors are suppressed by gibberellins (Achard & Genschik, 2008).

The aim of this study was to (i) characterize the shoot and root system of GA biosynthesis mutants and selected *A. thaliana* accessions carrying functional and mutated alleles of the *GA5* gene, (ii) and to evaluate their response to reduced water availability. In case mutations at the *GA5* locus have positive pleiotropic effects on root growth, this could help explaining their selective advantage in specific environments, especially when this relates to tolerance to reduced water availability.

Results

The root phenotype of GA biosynthesis mutants *in vitro*

To evaluate the role of mutations in the GA biosynthesis pathway in the modulation of root growth, different GA deficient mutants were phenotyped *in vitro* using the GROWSCREEN-ROOT (Nagel et al., 2009) system. Semi-dwarf mutants in the *GA20ox1 (GA5)* locus were included in the two backgrounds, L*er* (*ga5*) and Col (*ga20ox1-3*, *ga20ox1 ga20ox2*). Additional mutants with semi-dwarf (*ga3ox1-3*) and dwarf (*ga1-13*, *ga1-3 6xbxcol*) phenotypes were included. Principal Component Analysis (PCA) was performed to observe the structure of the data set. PCA1 and 2

together account for more than 80% of the information present in the different traits that were evaluated (Figure 3.1A). The *ga20ox2-1*, *ga20ox1-3,* and *ga20ox1 / ga20ox2* tend to cluster near their background accession Col, suggesting they do not differ much in root traits (Figure 3.1A). However the *ga5* mutant in L*er* background is far from its background accession L*er*. The semi-dwarf mutant *ga3ox1-3* does not cluster with any other genotype, as is the case for L*er* and the *ga1-13* mutant. When the root and shoot traits are compared (Figure 3.1B), it can be observed what are the most relevant traits for PCA1. Only the ratio rooting depth / FW (shoot fresh weight) is mainly related with PCA2 (Figure 3.1B). Root system traits are highly correlated. In addition to root traits, projected leaf area was quantified, which is highly correlated with shoot FW, DW (shoot dry weight) and root system width (Figure 3.1B). Root related traits are negatively correlated with FW / leaf area and DW / leaf area (traits used as a proxy for leaf thickness).

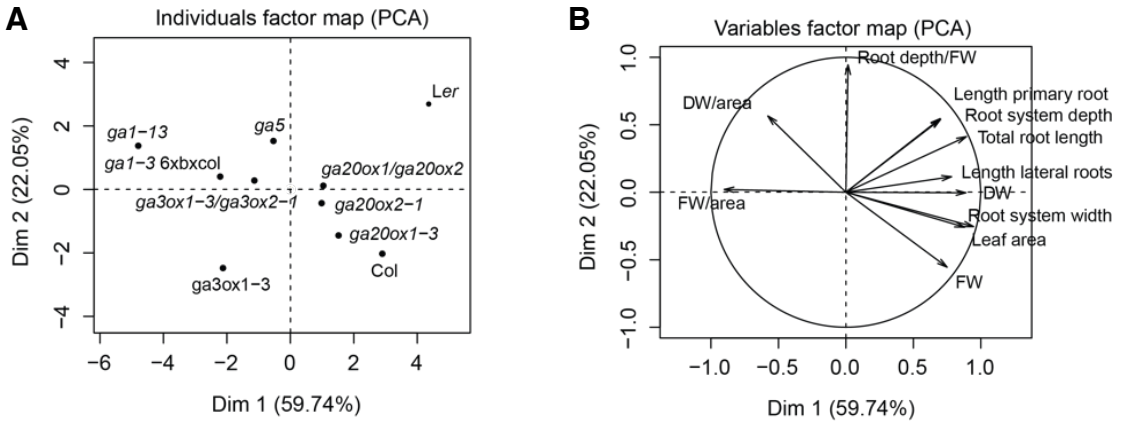

Figure 3.1. Principal Component Analysis (PCA) for shoot and root related traits. (A) PCA of GA biosynthesis *Arabidopsis thaliana* mutants and their corresponding wild type controls, based on the combined data of different shoot and root phenotypes of plants grown *in vitro* (1 % agar with 1 / 3 Hoagland solution). (B) Variables factor map. FW, shoot fresh weight; DW, shoot dry weight. All mutants are in Col background except for *ga5*, which has the L*er* background.

The parental accessions L*er* and Col exhibited the highest values for shoot and root system size compared to their semi-dwarf mutants, while *ga1-13* and *ga1-3* 6xbxcol mutant produced the shortest root systems (Figure 3.2A), as well as the smallest leaf area (Figure 3.2B). The *ga1* mutants also have thicker leaves compared with Col (Figure 3.2C). From the PCA analysis one can reduce the data to three main traits: total root length, projected shoot area, and the ratio FW / leaf area. These traits indicate that mutations in early steps of the GA biosynthesis pathway represented by the *ga1* mutants decrease root length and shoot area and increase leaf thickness

compared to their wild types (Figure 3.2A-C). The less extreme semi-dwarf mutants (*ga20ox1-3* and *ga5*) do not significantly decrease total root length, with the exception of *ga3ox1-3* (Figure 3.2A). The *ga5* mutant only showed a mild decrease compared to wild type. Semi-dwarf mutations mainly decrease the projected leaf area, for example the *ga5* mutant reached approximately half the size compared to its wild type. It seems, therefore, that the *GA 20-oxidase 1* (*ga5*) mutation has a stronger effect in the L*er* background than in the Col background because differences among *ga5* and L*er* are stronger than those observed between *ga20ox1-3* and Col.

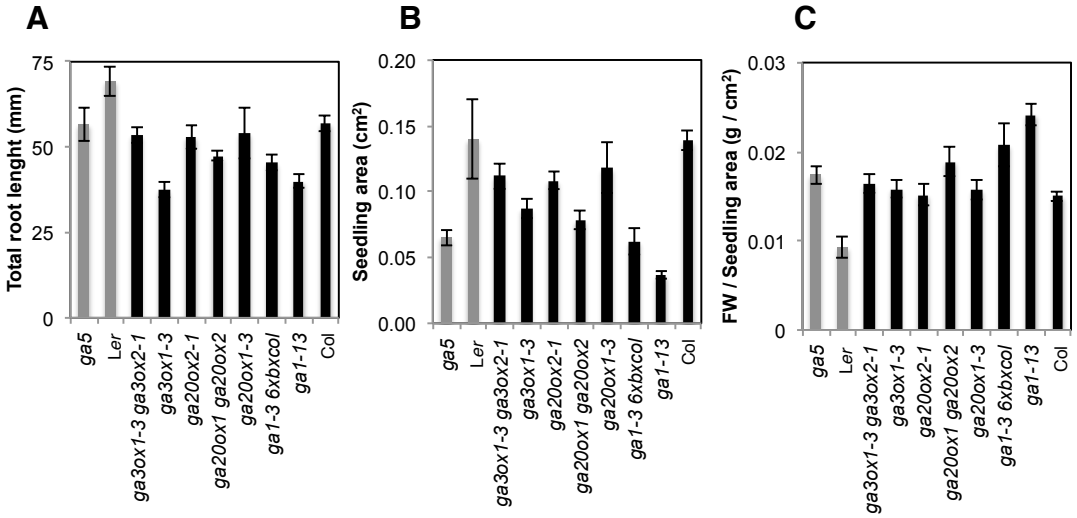

Figure 3.2. Phenotypes of GA related mutants. (A, B, C) Means (± standard errors) of different shoot and root traits in different GA biosynthesis *Arabidopsis thaliana* mutants and their corresponding wild type controls, grown *in vitro* (1 % agar with 1/3 Hoagland solution, evaluated 18 days after germination). FW, shoot fresh weight. All mutants are in Col background except *ga5,* which has the L*er* background. Different colors indicate near isogenic comparisons.

The root phenotype of semi-dwarfs accessions *in vitro*

The root system architecture of several semi-dwarf accessions allelic to *ga5* (Barboza et al., 2013) was characterized. Semi-dwarfs were compared with related, non-dwarf genotypes (Barboza et al., 2013). Both dimensions of the PCA analysis explained more than 75% of the variance observed in the data set (Figure 3.3A). PCA analysis shows no clustering of semi-dwarf accessions, thus suggesting that semi-dwarfism does not contribute to the variation of the evaluated traits. Semi-dwarf genotypes tend to cluster together with their wild, related accessions counterparts. The variable factor map shows a similar pattern to that of the mutants PCA (Figure 3.3B). The PCA indicate that Pak-3 is an outlier mainly because of its deeper root system and the length of its primary roots. The *ga5* mutant seems to have a high rooting depth / shoot FW ratio.

Mainly occurring because of the reduced leaf area compared to its background accession. It was possible to test the semi-dwarf effect among the traits because a balanced number of accessions was included carrying active and inactive *ga5* alleles. Remarkably, no semi-dwarf effect was significant. Only root system width was significantly different between dwarf and wild type genotypes, but when Neo-3 is excluded from the analysis this effect is not significant (P-value > 0.05). Neo-3 shows a much higher root system width than the other accessions (mean value ± standard error under control conditions 23.8 ± 1.6 mm vs. 7.3 ± 1.0 for Dja-1, a genetically related accession).

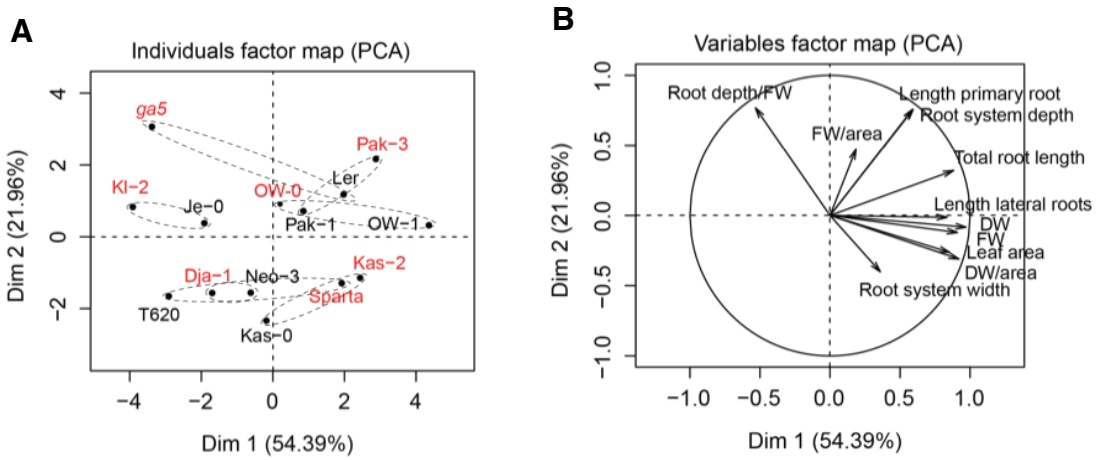

Figure 3.3. Principal Component Analysis (PCA) for shoot and root related traits. (A) PCA of 13 *Arabidopsis thaliana* accessions and the *ga5* mutant (L*er* background), based on the combined data of different shoot and root phenotypes of plants grown *in vitro* (1 % agar with 1/3 Hoagland solution, evaluated 18 days after germination). (B) Variables factor map. FW, shoot fresh weight; DW, shoot dry weight. Accessions colored in red means they have *ga5* mutant alleles. Accessions surrounded by dotted lines indicate genetically related pairs, e.g. *ga5* vs L*er* (*GA5*).

Because the observed effects were not attributed to the *ga5* mutations, possibly due to differences in the accessions background, F1 crosses of accessions with the *ga5* mutant and with its wild type L*er* provided an identical hybrid background genotype. These were evaluated in order to have a near isogenic background differing only in the *ga5* genotype. When semi-dwarf related traits such as leaf thickness (using FW / area as a proxy) were studied, differences among inactive vs. active *ga5* alleles were detected (Figure 3.4A). The presence of inactive *ga5* alleles increases the expression of this trait compared with wild type. However when total root length and leaf area were evaluated, inactive *ga5* alleles tend to decrease these traits but differences were observed among the different crosses (Figure 3.4B, C). The most remarkable difference was between the crosses *ga5* × OW-0 and L*er* × OW-0, in which the phenotype followed a trend opposite from that of the other compared pairs. This lead

to the conclusion that shoot and root modifications depend on the genetic background, which might interact to some extent with the *ga5* genotype. Apparently variation at additional loci contributes to genotype differences for the evaluated traits.

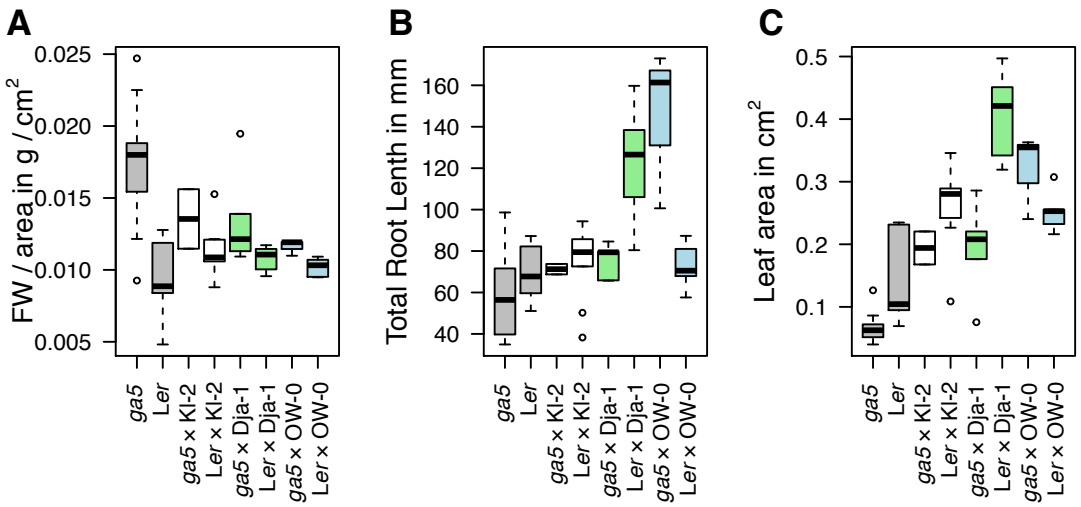

Figure 3.4. Shoot and root phenotypes in F1 populations. (A) Boxplots of the derived trait shoot fresh weight (FW) / seedling leaf area, (B) total root system length and (C) leaf area in different F1 populations derived from the crosses *ga5* × accessions and L*er* × accessions grown *in vitro* (medium 1 % Agar with 1/3 Hoagland, evaluated 18 days after germination). The thick horizontal line represents the median, boxes represent the 25th and 75th percentile (lower and upper hinges respectively), vertical lines represent whiskers (0.05th, 0.95th percentile), and open circles extreme values. Different colors indicate near isogenic comparisons.

Phenotypes of the semi-dwarfs Kas-2 and Pak-3 occur independently from the *ga5* inactive allele

The semi-dwarf accession Kas-2 showed the highest shoot biomass *in vitro* (Figure 3.5A). The *ga5* mutant was decreased in shoot biomass compared with its wild type background L*er*. However Pak-3 is a semi-dwarf accession that showed the longest primary root among all the other tested ones (Figure 3.5B). All observed phenotypes lead to the conclusion that semi-dwarfism affects neither these traits nor the relation shoot to root ratio (data not shown). Nevertheless indications that barley semi-dwarf accessions may show a higher root length (White et al., 2009) raised the question whether or not the long Pak-3 root system depth is *ga5* dependent. To test this hypothesis, the crosses *ga5* × Pak-3 and L*er* × Pak-3 were generated and their root system depth and shoot height was phenotyped in the F2 generations to quantify if root length and shoot height variation due to *ga5* genotype were correlated. Clear differences both for the plant height phenotype as well as for root system depth were

observed, but these traits segregate independently from each other (Figure 3.6A, B, C). Hence, it was concluded that *ga5* does not play a role modifying the root system depth in the Pak-3 accession.

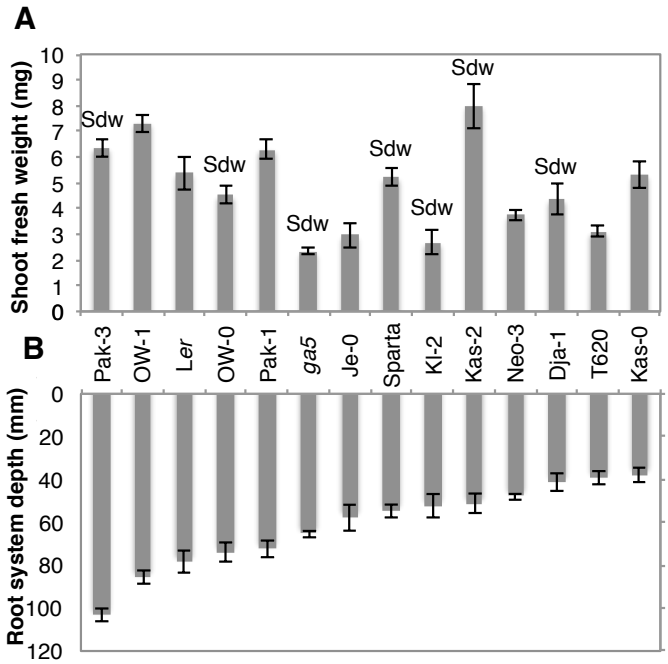

Figure 3.5. Contrasting phenotypes in the semi-dwarfs Pak-3 and Kas-2. (A) Shoot fresh weight and (B) root system depth (means ± standard error) in 13 *Arabidopsis thaliana* accessions and the *ga5* mutant (L*er* background) grown *in vitro* (medium 1 % Agar with 1/3 Hoagland, evaluated 18 days after germination). "Sdw" indicates semi-dwarf accessions.

QTL analysis was used as an independent approach to test the possible role of *ga5* in shoot biomass and root system depth *in vitro* using the L*er* × Kas-2 population (El-Lithy, 2006). The advantage of using this population is that semi-dwarfism due to the Kas-2 loss of function allele at the *GA5* locus is segregating, thus allowing to test the effect of *ga5* on root depth. Transgressive segregation was observed for seedling weight and root system depth (Figure 3.7A, B). When tested separately for each trait, *ga5* reduces both root system depth and seedling fresh weight but this effect is relatively small and the LOD score is not significant. A locus at or near *ERECTA* (chr 2) and a locus located on chr 5 (SNP304) controls the FW (LOD scores 5.0 and 4.1, explained variances are 14.3 and 10.6%, respectively, Figure 3.7C). *ERECTA* plays a minor role in root system depth together with a QTL on chromosome 4 (M4-3) but no QTL position was detected in the vicinity of the *ga5* locus (El-Lithy, 2006) (Figure 3.7D).

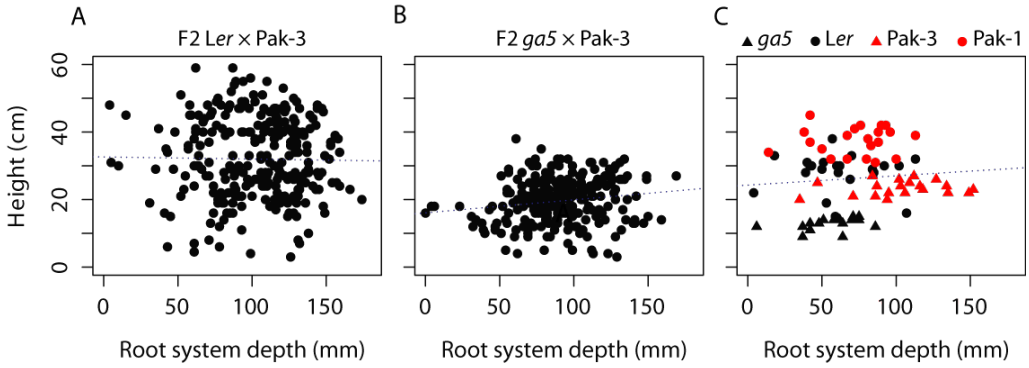

Figure 3.6. Long Pak-3 root system depth occurs independently from the *ga5* inactive allele. Correlations between shoot height (scored two weeks after flowering in greenhouse conditions) and root system depth (scored 28 days after sowing *in vitro,* medium 0.8 % Agar, 1/2 MS, pH 5.8) in (A, B) different F2 populations and (C) selected genotypes. Linear regression trend line is shown with the dotted line.

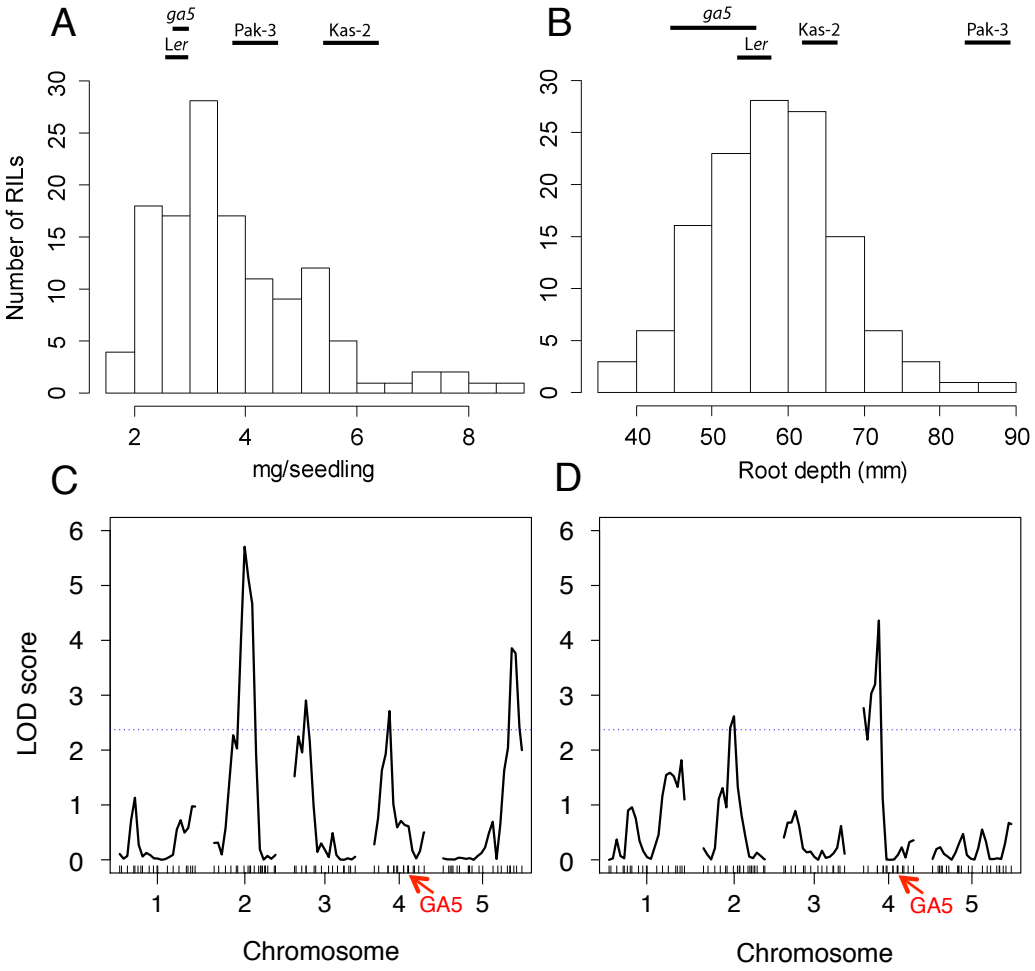

Figure 3.7. QTL mapping for shoot and root phenotypes. Frequency distributions of the traits (A) seedling weight and (B) root system depth in the L*er* × Kas-2 mapping population (125 recombinant inbred lines were analyzed; six seedlings were sown per genotype) grown *in vitro* (medium 0.8 % Agar, 1/2 MS, pH 5.8, evaluated 21 days after germination). The horizontal black bars denote the phenotype of the parental lines and two additional genotypes (mean values ± standard error). (C, D) QTL maps for the traits mentioned above. The horizontal dotted line shows the significance threshold by running 1000 permutations. Position of the GA5 marker is indicated with an arrow.

Root depth of semi-dwarf accessions grown in soil

In order to further characterize the evaluated semi-dwarf accessions, a selected group of genotypes was phenotyped in rhizotrons using soil as substrate. As observed in the *in vitro* experiment, the Pak-3 accession had the longest root system depth after the first four weeks after sowing (Figure 3.8A, B). In contrast, another Central Asian semi-dwarf accession Dja-1 presented the shortest root length that did not significantly differ from the related wild type Neo-3 during the first three weeks of growth. The semi-dwarf mutant *ga5* showed a shorter root system compared with its wild type L*er*. A final evaluation was performed two weeks after flowering in order to phenotype the plant height. As previously reported (Barboza et al., 2013), semi-dwarfism was the only trait significantly affecting plant height (P-value 4.67×10^{-11}, R^2adj = 0.68) (Figure 3.9A). Concerning the other traits, such as shoot fresh weight, flowering time, and root system depth (Figure 3.9B-D) only the genotype effect was significant (P-value < 0.0001). The late flowering west-European accession OW showed the highest shoot biomass and root system depth, whereas the opposite was observed for the early flowering accession Pak-1 (Figure 3.9B-E). The *ga5* mutant flowered slightly later than L*er*. The *ga5* root system depth was shorter than L*er* but not significantly different when applying post-hoc statistical tests (Figure 3.9D). Flowering time was positively correlated with fresh weight and rooting depth (Figure 3.10A, B).

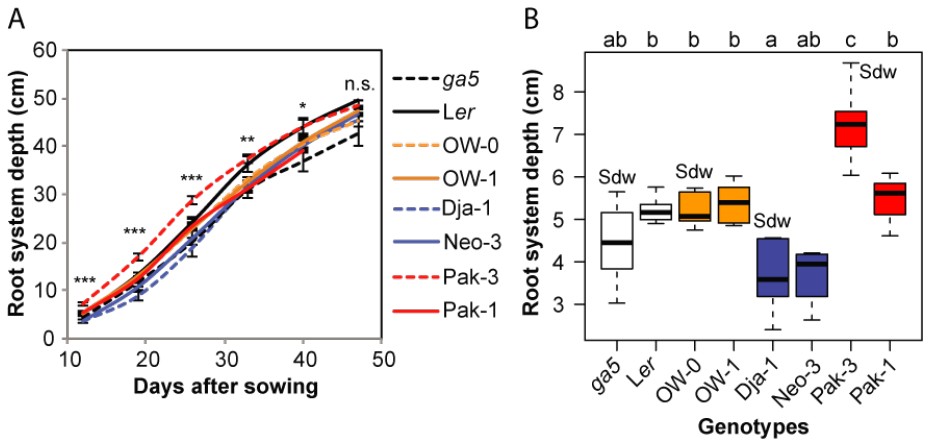

Figure 3.8. Root system depth of soil grown accessions. (A) Development of root system depth across time in seven *Arabidopsis thaliana* accessions and the *ga5* mutant (L*er* background) grown in soil-filled rhizotrons. Means ± standard errors are shown. Asterisks show P-value significance from an anova to test significant differences among the genotypes: ***0.001, **0.01, *0.05, n.s. not significant at the P-value > 0.05. (B) Boxplot shows the root system depth at day 12 after sowing. The thick horizontal line represents the median, boxes represent the 25th and 75th percentile (lower and upper hinges respectively), vertical lines represent whiskers (0.05th, 0.95th percentile) and open circles extreme values. The letters above the box plots indicate the results of a Tukey`s HSD test where means with different letters are significantly different (at P-value < 0.05). "Sdw" on the top of some boxes indicates the semi-dwarfs. Different colors indicate semi-near isogenic comparisons.

To further validate the root system phenotype of the Pak-3 accessions, several F3 lines derived from the cross L*er* × Pak-3 with differential root system depth were phenotyped *in vitro* and in rhizotrons for root system depth. It was possible to isolate F3 lines differing in their root system depth and carrying different backgrounds, e.g. segregation of *erecta* and *ga5* loss of function alleles (Figure 3.11A, B).

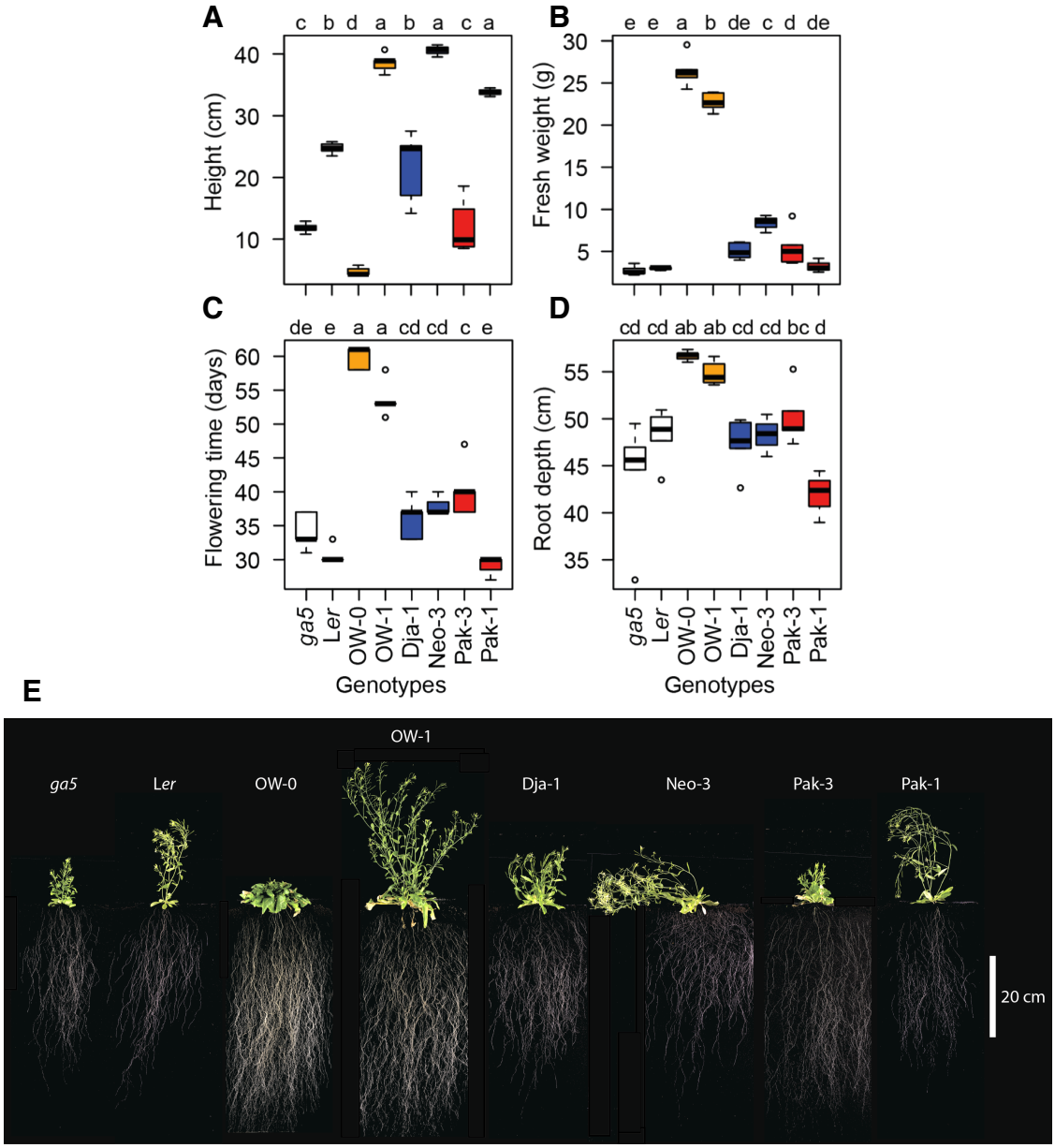

Figure 3.9. Contrasting phenotypes in selected Arabidopsis accessions. (A-D) Boxplots showing shoot height (A), rosette fresh weight (B), flowering time after sowing (C) and root system depth (D) in seven *Arabidopsis thaliana* accessions and the *ga5* mutant (L*er* background) grown in soil-filled rhizotrons. Evaluations and pictures were conducted two weeks after flowering time. The thick horizontal line represents the median, boxes represent the 25th and 75th percentile (lower and upper hinges respectively), vertical lines represent whiskers (0.05th, 0.95th percentile), and open circles extreme values. The letters above each panel indicate the results of a Tukey`s HSD test where means with different letters are significantly different (at $P<0.05$). (E) Representative images of root and shoot phenotypes of the seven accessions two weeks after flowering time. Different colors indicate semi-near isogenic comparisons.

Final evaluations show a major control of *GA5* regulating plant height (as previously known) and strong effects of the different accessions (backgrounds) regulating both root and shoot traits independent of having active or inactive *ga5* alleles. Thus it seems that *GA5* mainly controls plant height and further significant modifications in the root system are controlled by additional loci.

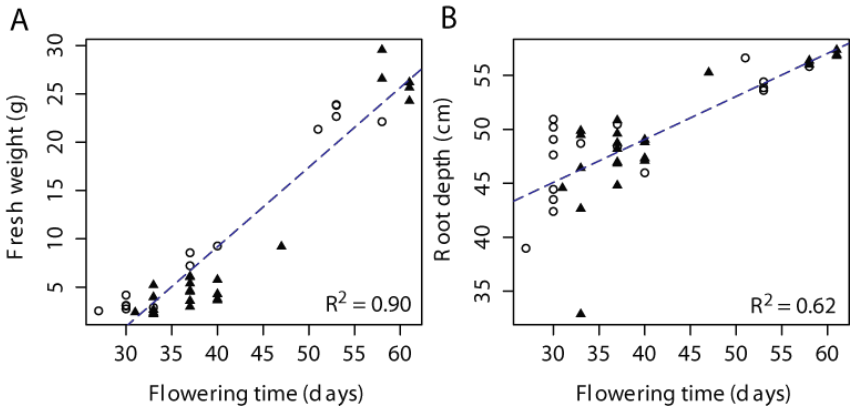

Figure 3.10. Effect of flowering time on shoot and root related traits. Correlations between flowering time and rosette fresh weight (A), and root system depth (B) in seven *Arabidopsis thaliana* accessions and the *ga5* mutant (L*er* background) grown in soil-filled rhizotrons. Triangles represent semi-dwarf accessions, open circles wild type accessions. Linear regression trend line is shown with the dotted lines. Blue dotted lines show the correlation among all variables, red dotted lanes show the correlation for semi-dwarf accessions and the red line for wild type accessions. Coefficient of determination (R^2) is shown on the lower right corner.

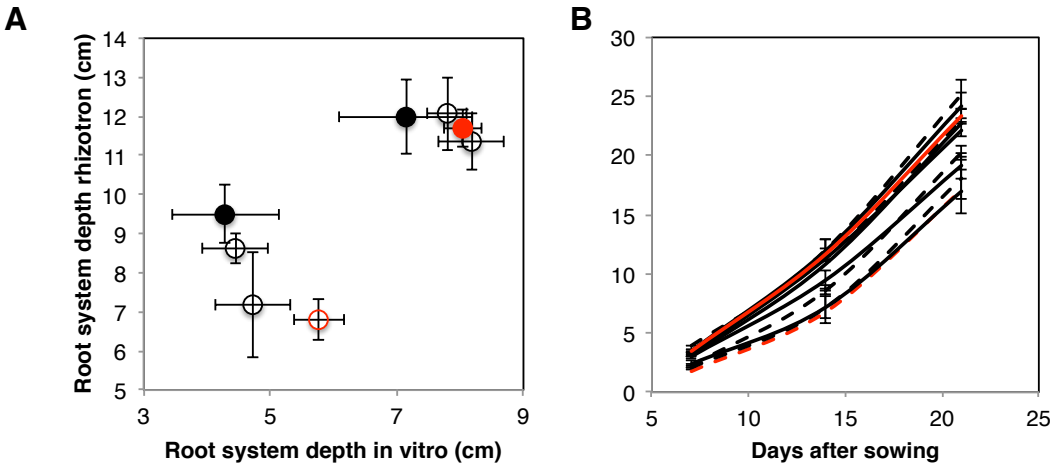

Figure 3.11. Segregation for root system depth in selected F3 lines. (A) Root system depth correlation between F3 genotypes grown *in vitro* and in rhizotrons. The F3 lines were derived from the cross L*er* × Pak-3. Plants *in vitro* were analyzed at day 15, in rhizotrons at day 14. Dashes represent standard errors of the mean. Filled squares represent functional *ERECTA* alleles (as phenotyped in parental F2 lines). Red filled square represent the parental genotype Pak-3; empty red square represents Landsberg *erecta*. Both parental lines were replicated using two biological repetitions. (B) Root system depth in rhizotron grown F3 plants across time, genotypes carrying *erecta* non functional alleles are indicated with dotted lines, parental lines are indicated with red lines (L*er* = dotted line).

Response to water limiting conditions

The presence of semi-dwarfs in nature indicates that this trait might confer a selective advantage under specific conditions. Previous studies suggest that *ga5* might be more drought tolerant due to modifications in the root system (Vartanian et al., 1994). Thus it was aimed quantifying the performance of GA biosynthesis mutants and natural semi-dwarf accessions under water limiting conditions. To mimic low water availability, a screen using osmotic stress by applying sorbitol (100 mM) was conducted *in vitro*. No strong differences were observed when comparing the ratio control / sorbitol for the traits FW and root system length for the different evaluated GA mutants. The ratios oscillated between ~ 0.58 - 0.75 for both traits. Only for the trait root system length the mutant *ga1-13* showed a ratio of ~ 0.83, and for FW the mutants *ga20ox2-1* and *ga3ox1-3* showed values of ~ 0.85 and ~ 1.0 respectively. A higher ratio, or a ratio near to one would indicate that there were no differences between the osmotic and control treatments. Semi-dwarfism does not seem to affect the response to osmotic stress. By testing the different F1 crosses no link was observed for water limiting conditions between the crosses *ga5* × accessions (mean ± standard error for root length 0.57 ± 0.02 and for FW 0.68 ± 0.02) and L*er* × accessions (root length ratio 0.60 ± 0.02; FW ratio 0.58 ± 0.05). Major differences were observed among the different genotypes were ratios ranged from 0.2 to ~ 1 (Figure 3.12). Again semi-dwarfs did not show a uniform response, e.g. the semi-dwarf Sparta showed the lowest ratios (~ 0.3 for FW and ~ 0.4 for root system length) while its semi-dwarf counterpart Dja-1 and Kas-2 were the most tolerant (ratios = > 1).

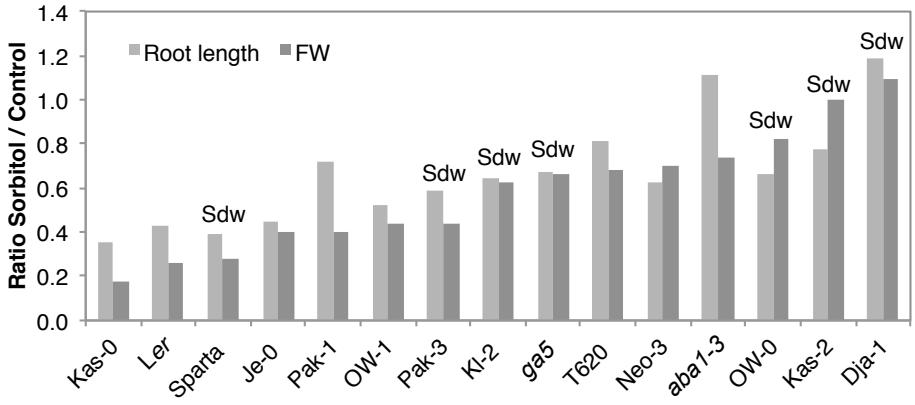

Figure 3.12. Osmotic stress response. (A) The ratio sorbitol / control was evaluated in different *A. thaliana* accessions and the *ga5* and *aba 1-3* mutants. The ratio was generated using the mean values of the Total Root Length and shoot fresh weight (FW) evaluated or harvested the final day of the experiment (day 18). "Sdw" indicates semi-dwarf accessions. Plants were grown *in vitro* (medium 1 % Agar with 1 / 3 Hoagland).

To test a different water limitation scenario, water-withholding experiments using pots filled with soil were conducted. Projected leaf area and shoot fresh / dry weight were traits used as indicators of plant performance and were evaluated using GROWSCREEN FLUORO (Jansen et al., 2009). The water-withholding experiment was divided into three stages: initial growth, then water withholding (from week four to week six after sowing watering was stopped) and recovery phase (at week six plants were re-watered and allowed to recover until the end of experiment at week seven). One group was watered to serve as control for the entire duration of the experiment (details in materials and methods). The Pak-3 accession showed no significant differences in the ratio of shoot biomass between control and water limiting conditions (Figure 3.13). Kas-2 showed an intermediate effect. Both accessions showing this high effect were semi-dwarfs. Conversely the semi-dwarf OW-0 showed the lowest ratio together with the ABA deficient mutant *aba1-3*, which was used as negative control. The use of this ratio using the final dry weight might mask possible effects because final dry weight is a result of growth during drought and recovery phases. A plant performing well in drought but poor in recovery might be very similar to one that performs poor in drought but well in recovery.

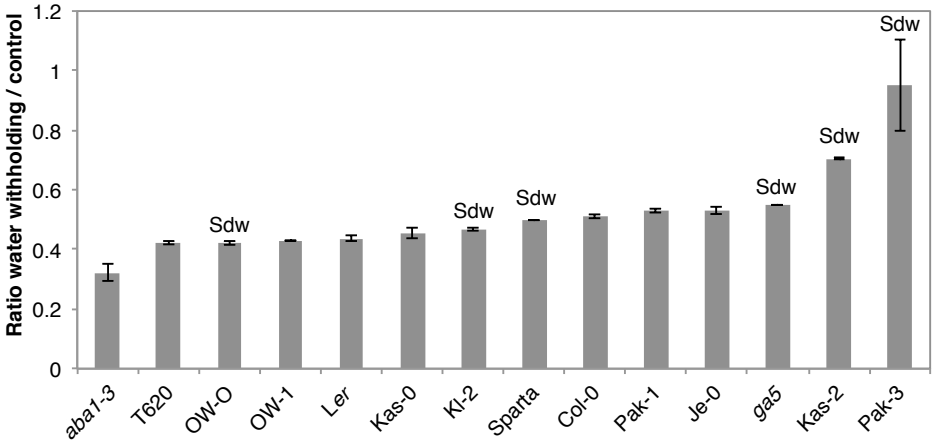

Figure 3.13. Water withholding response of different genotypes. The ratio drought/control was generated using the mean value of the dry weight of rosettes harvested the final day of the experiment. Dashes represent the standard deviation from estimating the same ratio using the fresh weight. "Sdw" indicates semi-dwarf accessions.

The use of GROWSCREEN-FLUORO allows the estimation of performance ratios in time during and after drought using projected leaf area. Thus a ratio that describes growth at water limitation was estimated dividing the relative growth rates under water limiting conditions over control. When different GA mutants were evaluated, the magnitude of the differences was modest (Figure 3.14). Relevant to

notice is that the *ga1* mutants showed a reduced performance during drought thus pointing at the fact that early mutations in the GA biosynthesis pathway may affect growth under water-limiting conditions. F1 populations using the crosses ga5 × Pak-3 and ga5 × OW-0 showed no differences under water limiting conditions compared with the crosses with the wild type L*er* (data not shown). When this effect was evaluated in different accessions, the Pak-3 accession presented a high performance during water-limiting conditions and the highest during the recovery phase (Figure 3.15). Analyzing the different growth curves of the evaluated accessions (Figure 3.16A) provide additional information. Pak-3 is more tolerant to water-withholding conditions, which means the reduction in growth is low compared to control conditions or to Pak-1 (Figure 3.16B). Interestingly, under control conditions plants stop growing after day 45, explaining why there were high performance ratios during the recovery phase. The low ratio using the shoot weight at the final date of the experiment (week 7) confirms the results. Kas-2 (Figure 3.16C) and OW accessions (Figure 3.16D) show additional examples where it seems Kas-2 is more tolerant than OW-0. Taken together, these data show that the various *ga5* alleles do not differ in their behavior from their related wild types and therefore this mutation has no detrimental effect during water limitation.

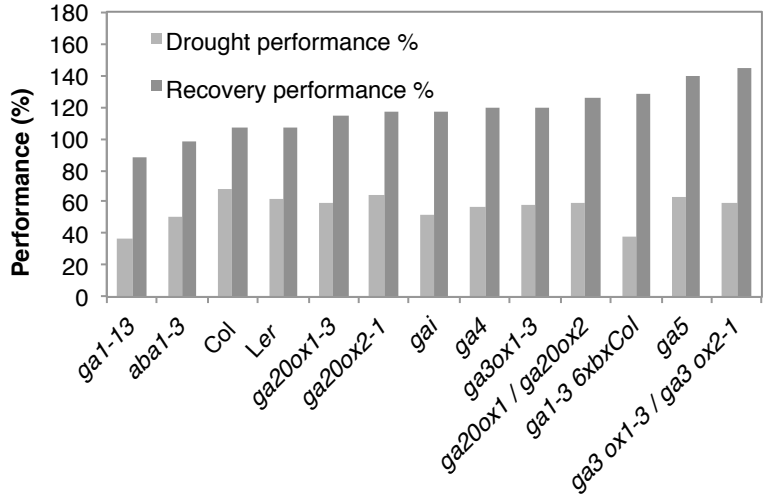

Figure 3.14. Mutations in *ga1* show a low drought performance. Relative growth performances of different *Arabidopsis thaliana* genotypes obtained during drought and after re-irrigation of the plants (recovery).

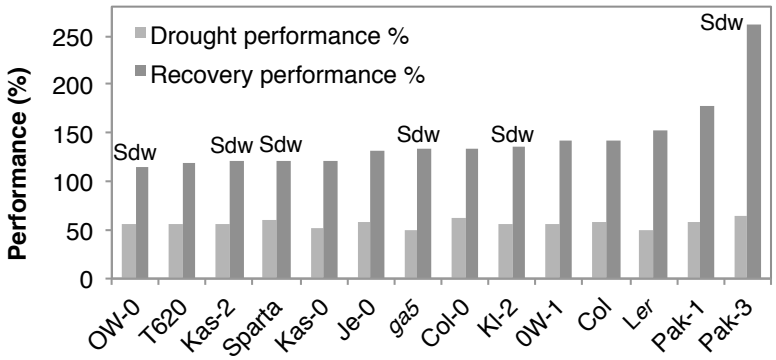

Figure 3.15. Pak-3 shows the highest drought performances. Relative growth performances of different *Arabidopsis thaliana* genotypes obtained during drought and after re-irrigation of the plants (recovery). "Sdw" indicates semi-dwarf accessions.

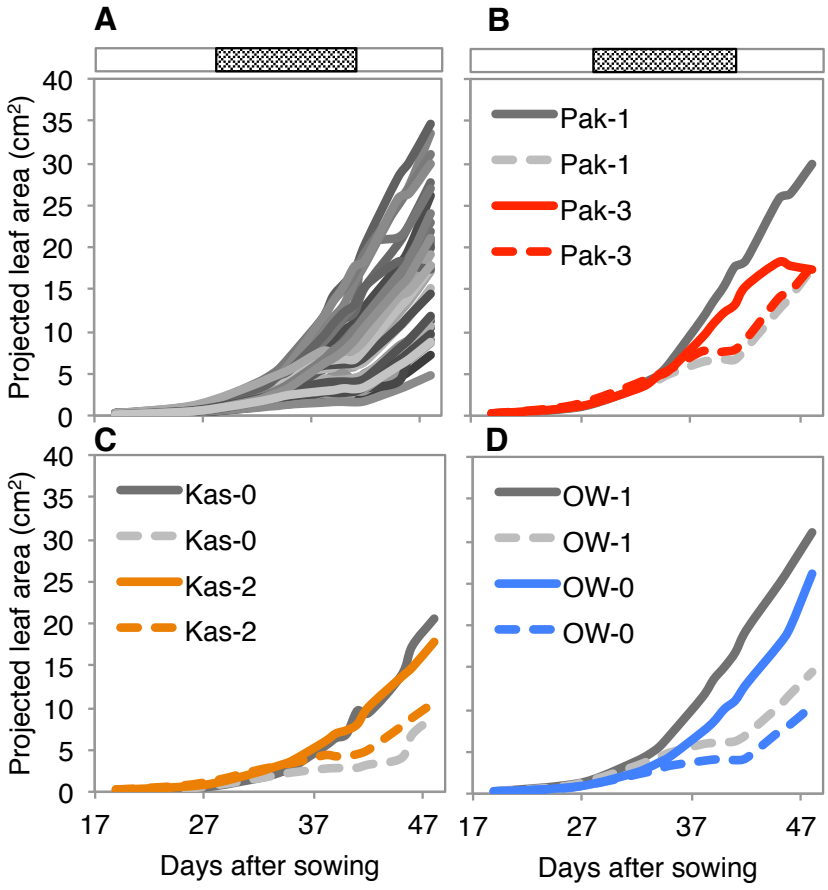

Figure 3.16. Pak accessions show a reduced leaf area. (A) Growth curves of different *A. thaliana* genotypes grown on soil under control and water-withholding conditions. Mean projected leaf area per genotype was plotted. Growth curves for the (B) Pak, (C) Kas, and (D) OW accessions. Dotted lines represent accession grown under water-withholding conditions. Vertical scale on top of the plots indicates the duration of the treatment (black rectangular bar).

Discussion

GA5 is a GA biosynthesis gene with a major effect only on plant height but without further major pleiotropic effects as are observed in mutants in early steps of the GA pathway (Koornneef & van der Veen, 1980). From all five AtGA20ox`s paralogs, *GA5* is the only one having a major effect on plant height when mutated, due to its expression pattern in comparison with that of the paralogs (Plackett et al., 2012; Rieu et al., 2008). Because the *ga5* mutants show no trade-offs, it is explained why natural variants carrying loss of function alleles can be maintained in nature in contrast to mutations in early GA biosynthesis genes such as *ga1* (Barboza et al., 2013). Redundancy in the GA20ox genes might maintain the GA homeostasis in all organs except the growing inflorescence stem and leaves thus allowing *ga5* mutants to maintain a similar root system as their corresponding wild types. For instance, *GA20ox2* and *GA20ox3* are overexpressed when GA inhibitors are applied (Bidadi et al., 2010) indicating the regulation of GA levels by a negative feedback system on biosynthesis genes. Despite this, the *ga5* mutant shows a moderate reduction of root length compared to the wild type, an effect observed in previous studies in the Landsberg *erecta* background (Vartanian et al., 1994), but not in Col (Rieu et al., 2008). The genetic background plays a role in this effect suggesting occurrence of epistasis. A good example has been shown for the *SVP* gene, for which the same mutation have a different effect on flowering time depending on the genetic background (Méndez-Vigo et al., 2013). Although no significant effect on root growth were observed in Arabidopsis, a promotion of root growth (dry weight) was observed in semi dwarfs in poplar (Elias et al., 2012) in which down regulation of a DELLA protein caused semi-dwarfism.

Natural variation for root systems architecture has been reported in *Arabidopsis*. Several studies have used QTL analysis to map loci involved in root related traits (Fitz et al., 2005; Loudet et al., 2005; Mouchel et al., 2004; Sergeeva et al., 2006). Other studies yielded mapping under different mineral concentrations showing the plasticity of the root system and the loci controlling these effects (Kellermeier et al., 2013; Prinzenberg et al., 2010). By using Genome Wide Association Studies (GWAS) the genes *PHOSPHATE 1* (*PHO1*) and *Root System Architecture 1* (*RSA1*) were associated with root system allometry (Rosas et al., 2013). However, GWAS performed for total root length, did not result in significant associations suggesting the role of many loci controlling this trait (Rosas et al., 2013).

Another example of natural variation for root length identified loss of function alleles in the positive regulator of auxin signalling *BRX,* which modifies root length under acidic conditions (Gujas et al., 2012). These examples illustrate the role of specific root system architecture traits that may be selected for depending on the environment as illustrated in a recent review (Lynch, 2013). This modulation can affect different genes from independent pathways. The possible locus (loci) controlling the root system depth in the Pak-3 accession has not been mapped and therefore it cannot be concluded if the variation is due to known pathways or genes. The different root systems found in other central Asian accessions such as Kas-2 (relatively short root system) or Neo-3 (relatively large root system width) suggests that the variation present in this region and a possible variable selective pressure may allow the evolution of specialized root systems.

The use of semi-dwarfs in crops has been a topic for discussion regarding possible trade-offs under drought conditions. The variety IR64, carrying mutations in *GA20ox2* (Sasaki et al., 2002; Spielmeyer et al., 2002), the functional ortholog of *ga5* (*GA20ox1*), has been reported as drought sensitive because it was bred for irrigated agricultural environments (Lafitte et al., 2007; Swamy et al., 2013). However some authors point to the observation of occurrence of rice semi-dwarfs carrying drought tolerance (Lafitte et al., 2007). The physiological hypothesis behind it might deal with GA to ABA antagonism, where semi-dwarfs carrying low GA levels will thus show ABA accumulation being beneficial under drought (Lafitte et al., 2007). As found in this study for *A. thaliana*, tolerant accessions can occur in semi-dwarf background. Even the isogenic comparison between *GA20ox* mutants both in the Landsberg *erecta* and Col backgrounds, show no trade-offs during drought.

Recent studies point at the relevance of the root system depth to increase drought avoidance. Natural variants in rice with deep rooting system have been isolated and the QTL behind this trait identified as being *DEEPER ROOTING 1* (*DRO1*) (Uga et al., 2013). Introgression of this gene into cultivated varieties conferred drought resistance. It is still disputable whether or not a long root system is directly translated into drought tolerance, at least as argued for rice (Lafitte et al., 2007). In this study it was addressed the possible link between water limiting conditions and root system length. For instance, previous studies have associated stomatal density with transpiration efficiency (Masle et al., 2005). Another study have related flowering time, specifically the gene *FRIGIDA* (*FRI*) as relevant to deal with drought tolerance

(Lovell et al., 2013). Yet, it remains an open question to further understand and characterize the possible link between a long root system and its plausible selective advantage in nature and if it displays possible trade-offs.

Conclusions

The *ga5* locus shows no major pleiotropic effects; here evidence is provided that this is also the case for root related traits and for drought tolerance. This indicates that semi-dwarfism is neutral under the tested drought conditions. Reduced water availability tolerance can occur in semi-dwarf backgrounds as illustrated here for the Pak-3 and Kas-2 accessions. The semi-dwarf Pak-3 accession shows a long root system depth independent from the *ga5* loss of function allele and this was confirmed using both agar- and soil-based assays.

Materials and methods

Plant material

Genotypes used in previous studies were included in the experiments (Barboza et al., 2013). For the QTL mapping experiment the L*er* × Kas-2 Recombinant Inbred line population was used (El-Lithy, 2006).

In vitro root experiments

To characterize the phenotype of the semi-dwarf accessions, an experiment was conducted with a number of natural accessions together with the *ga5* mutant and L*er*. All experiments were conducted in a complete randomized design. For the *in vitro* experiments the plates were randomized 5 times per week. To understand the role of different mutations in the GA biosynthesis and signalling pathway, different mutants were phenotyped together with selected F1 populations. In all experiments L*er* and the *ga5* mutant were included. Root system was phenotyped *in vitro* using the GROWSCREEN-ROOT system with shoots growing outside the agar plate while roots growing through the agar media (Nagel et al., 2009). The culture medium contained 1/3 Hoagland solution (as described in Nagel et al., 2009), 1% agar (Sigma). For the osmotic stress experiments 100 mM Sorbitol (Merck) was used. Plants were grown on 120 × 120 × 17 mm plates (Greiner) filled with ~166 ml medium (completely filled). Seeds were sterilized by using 70 % ethanol (3 min), thereafter, 0.5 % NaOCl (10 min)

and then seeds were rinsed three times with sterile Milli-Q H_2O. After sowing seeds were incubated at 4°C in the dark for 5 days. Subsequently, the sown plates were vertically incubated in a chamber set to 8 / 16 hours light / dark period, 22 °C day / 18 °C night temperature and 60 % air humidity. Each plate contained four plants of the same genotype; this was replicated three times. Plants were phenotyped at day 18 after germination.

Additional experiments were conducted to phenotype the root system depth in the L*er* × Kas-2 Recombinant Inbred Line mapping population (El-Lithy, 2006) and in F3 lines derived from the cross L*er* × Pak-3. Plants were grown in half MS medium (Duchefa), pH 5.8, 0.8 % agar (Plant agar Duchefa), 100 µl / liter Plant Preservative Mixture (PPM). In all experiments seeds were stratified five days at 4 °C in a dark cold room. The same plates were used, as described above. In this case plants were grown inside the plate, having six plants per plate. Plants were phenotyped at day 21 after germination.

The mutants *ga1-13*, *ga1-3* 6xbxcol, and *ga3ox1-3 / ga3ox2-1* were stratified in 100 µM GA_{4+7} solution (Duchefa) under the same temperature conditions, as described above. The GA stock solution was at a concentration of 25 mM (GA diluted in a few drops of KOH 1M).

Rhizotron experiments

To phenotype the plants in rhizotrons, an experiment was conducted with six *A. thaliana* accessions, the *ga5* mutant and L*er*. To characterize F3 lines derived from the cross L*er* × Pak-3 segregating for root length, an experiment was conducted with ten F3 lines and the parental lines, the latter with two independent biological repetitions. Plants were grown in rhizotrons (60 × 30 × 2 cm) filled with peat soil as described in previous studies (Nagel et al., 2012). Seeds were sown directly in the rhizotrons and stratified for 4 days in a cold chamber at 4 °C. After that plants were grown in greenhouse conditions located in the institute Plant Sciences (IBG-2; Forschungszentrum Jülich GmbH, Jülich, Germany). One plant per rhizotron was grown for the former rhizotron experiment; two plants per rhizotron were grown for the latter experiment. Each genotype contained six repetitions. A randomized design was used in the experiments. Root system depth was quantified in the first rhizotron experiment by image processing using ImageJ (Schneider et al., 2012) and in the latter

experiment by measuring directly with a ruler in the rhizotrons. In both experiments images were acquired once a week.

Above ground phenotyping experiments

Water-withholding experiments were conducted using GROWSCREEN-FLUORO (Jansen et al., 2009). Both control and water-withholding groups were nearly water-saturated upon transplanting and then got drier after that. They were maintained between 50 and 40 % of maximum soil water content (Table 3.1) until start of withholding (week 4). The control group then stayed between 50 and 40 % whereas the drought group looses more water until they reach 10% or even less (Table 3.1). This is the soil moisture level at which plants stopped growing. Then re-watering was done (week 6) up to the 40-50% level and plants were allowed to recover for one week. The first experiment was conducted including natural accessions together with *ga5* and L*er*. In the second experiment, a group of GA biosynthesis and signalling mutants were studied. The third experiment included the F3 lines derived from the cross L*er* × Pak-3 indicated in the root experiments. Each genotype had 10-20 repetitions (depending on experimental set up). Parental lines (L*er* and Pak-3) contained two independent biological replicates. A complete randomized design was used. Single Seeds were sown on 576 wholes trays filled with soil (Pikiererde, Balster Einheitserdewerk GmbH, Fröndenberg, Germany). Stratification was conducted for four days at 4°C. After that trays were moved to the growth chamber (8 / 16 hours light / dark period, 22 °C day / 18 °C night temperature, air humidity 50%). After cotyledon unfolding, seedlings were transplanted to pots (7 × 7 × 8 cm) filled with peat-sand-pumice substrate (SoMi 513 Dachstauden, Hawita GmbH, Vechta, Germany). Pots were randomized and arranged in trays (40 pots per tray). Tray weighing, irrigation and data acquisition with GROWSCREEN FLUORO were conducted as described in previous studies (Jansen et al., 2009). Trays were imaged and automatically randomized five times per week. The mutants *ga1-13*, *ga1-3* 6xbxcol, and *ga3ox1-3 / ga3ox2-1* were germinated as described in the *in vitro* root experiments.

Table 3.1. Gravimetric soil water content.

Experiment	Transplanting - fully saturated:	Initial growth - decreased to:	Well watered - maintained at:	Drought - decreased to:	Recovery - refilled to:
Semi-dwarf accessions	100%	40%	40%	7%	40%
Mutants	100%	40%	40%	9%	40%

Data analysis

Descriptive statistics, Anova tests, and Tukey tests were performed with R. Principal component analysis was done with the package FactoMineR (Lê et al., 2008). QTL mapping was used using the R/qtl package (Arends et al., 2010; Broman et al., 2003). The function "mqmscan" was used setting all markers as cofactors and later eliminated through backward elimination. To quantify the explained variance, the main QTLs were manually selected and fitted into a multiple QTL model using the function "fitqtl" (Haley-Knott regression was the selected method).

Chapter 4

Genetics controlling germination sensitivity to the gibberellin biosynthesis inhibitor paclobutrazol in *Arabidopsis thaliana*

Abstract

Arabidopsis thaliana natural variation for seed germination was studied to dissect the genetic control of differences in sensitivity for the gibberellin inhibitor paclobutrazol (PAC). The accession Shakdara (Sha) was found to be PAC tolerant. By using Quantitative Trait Loci (QTL) analysis and Genome Wide Association Studies (GWAS), a complex regulation of the effects of GA depletion / restoration was found. Screening different mapping populations with Sha as one of the parental lines identified different loci controlling the PAC sensitivity in the various populations. The application of higher PAC doses or the application of PAC + GA_{4+7} allowed the detection of treatment specific QTLs. A main locus for PAC sensitivity on chromosome 1 was validated and characterized for its seed dormancy and germination behavior. This QTL plays a minor role on seed dormancy. Accessions screened for GWAS showed a broad spectrum of PAC sensitivity, but despite this no major loci were identified. PAC sensitivity is partially related with dormancy. The occurrence of dormant accessions displaying PAC tolerance points to partly different mechanisms controlling these traits.

Introduction

Natural variation in species with a broad distribution range is expected to reflect adaptation to the environment in which the plants are growing. To dissect the genetic basis of this variation Quantitative Trait Loci (QTL) analysis, either using the progeny of crosses among accessions or Genome Wide Association Studies (GWAS) are the procedures to get to the genes underlying allelic variation (Alonso-Blanco et al., 2009; Weigel, 2011). Thereafter experiments need to be done to validate the QTL and to confirm the involvement of allelic variation of specific genes underlying the QTL. Validation is often done by the analysis of Near Isogenic Lines (NILs) that can be obtained by backcrossing specific genotypes with one or both parents (Bentsink et al., 2010) or by selecting NILs from so-called Heterogenous Inbred Families (Tuinstra et al., 1997). This approach has been successful in many cases and revealed that the identified genes can belong to all types of ontology classes where alleles can differ in the structural part of the genes or in the promoters (for a review see Alonso-Blanco et al., 2009), where the latter can result in variation in gene expression.

Bioactive gibberellins (GAs) are plant growth regulators affecting several traits such as seed germination, flowering time, anther and petal development, and cell elongation. The biosynthesis of GA and its signalling is well understood (Hedden & Thomas, 2012; Yamaguchi, 2008). Indications that variation for GA responses is present among natural Arabidopsis accessions have been reported for seed germination, hypocotyl length, floral morphology and plant length (van der Schaar et al., 1997; Borevitz et al., 2002; Brock et al., 2012; Barboza et al., 2013). It was found that most of the semi-dwarfs that can be found in nature have mutations in the *GA5* locus encoding the *GA20oxidase1* gene (Barboza et al. 2013, El-Lithy et al., 2006).

Apart from elongation growth and stress tolerance, gibberellins also play an important role in seed germination (Holdsworth et al., 2008). This can be shown by the inhibition of germination by gibberellin biosynthesis inhibitors such as paclobutrazol (PAC) (Hedden & Graebe, 1985), ancymidol (Coolbaugh & Hamilton, 1976), and tetcyclasis (Debeaujon & Koornneef, 2000) and the lack of germination in GA deficient mutants (Koornneef & Veen, 1980). The effect of GA on germination is linked with the inhibition of germination by the plant hormone abscisic acid (ABA), which is overcome by GA that is synthesized upon exposure of seeds to light (Yamaguchi, 2008). When ABA is absent or not functioning, germination will take

place without GA and therefore also in the presence of PAC (Koornneef et al., 1982). For this reason PAC resistance has been used to select for ABA biosynthesis mutants (Léon-Kloosterziel et al., 1996; North et al., 2007) as well as for ABA signalling mutants (Nambara et al., 1992). However accessions that differ only a little in germination behavior can also differ significantly in PAC resistance and QTLs specific for germination on paclobutrazol have been identified (van der Schaar et al., 1997). We hypothesize that these genetic differences in paclobutrazol resistance can be explained by differences in plant hormone synthesis and / or signalling (both ABA and GA). By using QTL analysis and GWAS we aim to (i) study the genetic control of PAC sensitivity in *Arabidopsis thaliana* using natural variation (ii), validate loci controlling this effect, and (iii) characterize the germination behavior of genotypes differing in PAC sensitivity.

Results

QTL analysis for paclobutrazol sensitivity in RIL populations

To analyze the natural variation for paclobutrazol sensitivity in *Arabidopsis thaliana*, different accessions were analyzed for PAC sensitivity (4 µM). Initial screens identified the Central Asian accession Sha as an accession with reduced PAC sensitivity when compared with sensitive accessions such as Bay-0 (here there after refer as Bay) (Figure 4.1A). To dissect the genetics controlling this effect, QTL mapping using a Bay × Sha Recombinant Inbred Line (RIL) (Loudet et al., 2002) population was performed. Experiments were conducted using the high throughput methodology "GERMINATOR" (Joosen et al., 2010). This method allows the quantification of several seed germination related traits. In the present study we used: germination across time described with the area under the curve (AUC); maximum germination (gMAX); time to reach 50 % of germination (t50); and uniformity of germination (u8416) which indicates the time interval between 84 % and 16 % of seeds to germinate. The PAC application reduces AUC and gMAX, and increases the t50 and u8416 mean values (Table 4.1). Broad sense heritabilities (H^2) were higher for the AUC and gMAX under PAC conditions compared to control (Table 4.1). On the other hand the heritabilities for t50 and u8416 were considerably lower than in the control conditions. The coefficient of genetic variation (CV_G) shows ample genetic variation in the different traits (Table 4.1). The germination phenotype of the RIL populations can be visualized by plotting the gMAX values of control and PAC

treatments against each other (Figure 4.1B). Few RILs germinating lower than 80 % under control conditions indicate the presence of some residual dormancy in these RILs.

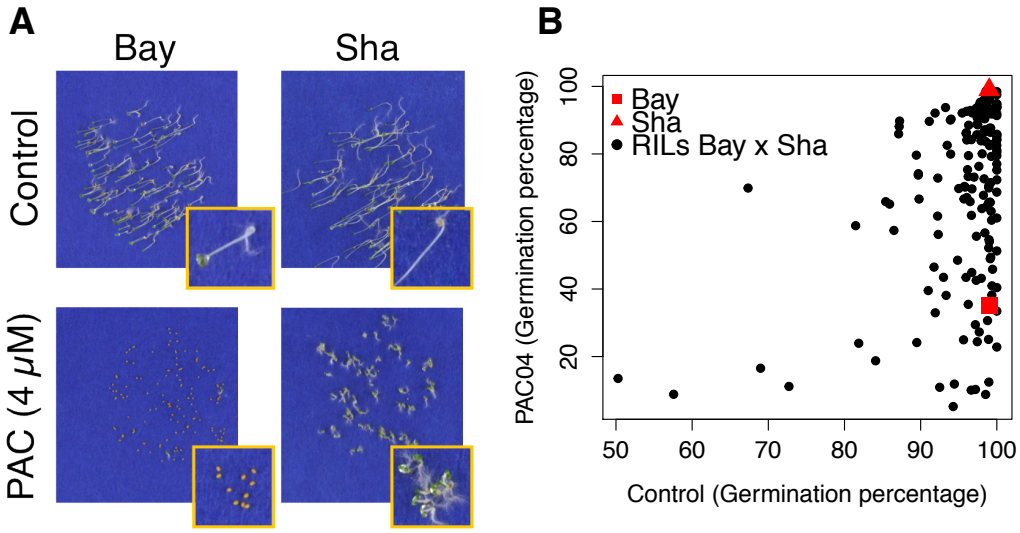

Figure 4.1. Paclobutrazol sensitivity in the Bay × Sha RIL population. (A) Germination of the Bay and Sha *A. thaliana* accessions in control and in paclobutrazol. (B) Correlation between the maximum germination of the control and paclobutrazol (PAC) treated seeds (after 120 hours) in the Bay × Sha RIL mapping population.

QTL mapping identified several loci across the genome (Figure 4.2, Appendix 4). The highest number of QTLs and the most significant were detected for the traits AUC and gMAX under PAC conditions as well as for the subtraction Control – PAC (Δ). The most significant QTL was mapped on chr 1 (Appendix 4), which together with a QTL on chr 5 confers differences in germination on PAC. At these positions no or weakly significant QTLs were detected for differences in germination under control conditions (Appendix 4). The highest QTL under control conditions was located on chr 4, and was not detected in the PAC treatment (Figure 4.2, Appendix 4). A QTL model was fitted with the most significant QTLs for gMAX under PAC conditions and showed that the QTL on chr 1 explained 16.7 % of the variance (Table 4.2). For all QTLs the Sha alleles increased the trait value. No significant interactions were present when testing pairwise interactions among QTLs.

Table 4.1. Descriptive statistics for different germination related traits in different mapping populations.

Population	Treatment	Trait[a]	$\overline{X}$	(SD)	$[V_G]$[b]	$[V_E]$[c]	$[H_2]$[d]	$[CV_G]$[e]
Bay × Sha	Control	AUC	81.59	8.75	15.64	7.88	0.80	3.43
Bay × Sha	Control	gMAX	0.96	0.07	0.13	0.08	0.74	26.55
Bay × Sha	Control	t50	33.59	4.24	7.43	4.13	0.76	5.74
Bay × Sha	Control	u8416	11.36	5.03	7.91	6.23	0.62	17.50
Bay × Sha	PAC04	AUC	55.21	23.09	44.32	13.23	0.92	8.53
Bay × Sha	PAC04	gMAX	0.68	0.26	0.50	0.17	0.90	73.22
Bay × Sha	PAC04	t50	37.71	8.01	9.00	13.26	0.32	5.62
Bay × Sha	PAC04	u8416	12.70	6.13	6.57	10.36	0.29	14.27
L*er* × Sha	Control	AUC	87.21	2.48	3.24	3.76	0.43	1.46
L*er* × Sha	PAC04	AUC	63.70	17.36	31.33	12.00	0.87	6.21
L*er* × Sha	PAC08	AUC	37.33	20.40	36.93	16.51	0.83	11.51
L*er* × Sha	PACGA	AUC	79.13	13.71	22.84	9.02	0.87	4.27
Sha × Col	Control	AUC	83.86	6.94	9.00	10.18	0.44	2.53
Sha × Col	PAC04	AUC	66.48	24.46	44.14	16.12	0.88	7.07
Sha × Col	PAC08	AUC	58.61	26.20	48.37	18.64	0.87	8.39
Sha × Col	PACGA	AUC	86.95	7.00	10.43	5.27	0.80	2.63
Col × L*er*	Control	AUC	90.30	1.84	2.01	3.10	0.30	1.11
Col × L*er*	PAC04	AUC	54.79	17.35	32.27	13.66	0.85	7.33
Col × L*er*	PAC08	AUC	38.24	21.14	34.51	24.33	0.67	10.86
Col × L*er*	PACGA	AUC	84.86	10.29	17.95	10.21	0.76	3.53
Hapmap	Control	AUC	58.08	13.49	140.35	41.03	0.77	5.93
Hapmap	Control	gMAX	0.88	0.16	0.02	0.01	0.77	43.07
Hapmap	PAC04	AUC	19.51	19.84	331.79	64.97	0.84	21.87
Hapmap	PAC04	gMAX	0.34	0.31	0.08	0.02	0.83	154.76

[a]Abbreviations: AUC, area under the curve (AUC); gMAX maximum germination (germination fraction); t50, time to reach 50 %; u8416, uniformity of germination indicating the time interval between 84 % and 16 % of seeds to germinate.
[b]Among genotype variance.
[c]Enviromental variance (estimated as total genotype variance – V_G).
[d]Broad sense heritability (V_G/V_G+V_E).
[e]Coefficient of genetic variation $\left(100 \times \sqrt{V_G}\right) / \overline{X}$

The occurrence of several loci with large effects in the Bay × Sha population raised the question whether this occurs in other populations where Sha is present. To answer this the L*er* × Sha (Clerkx et al., 2004) and Sha × Col (Simon et al., 2008) RIL populations together with the Doubled Haploid (DH) Col × L*er* population (Wijnker et al., 2012) were phenotyped for PAC sensitivity (PAC04, 4 μM). To further understand the genetic control of PAC sensitivity also a higher dose was applied (PAC08, 8 μM) as well as the treatment PAC08 + GA (PACGA, GA_{4+7}, 8 μM). The three populations

were grown in the same conditions and stored in the same way to avoid differences in possible maternal effects. The application of higher doses of PAC (PAC08) reduced germination percentage (gMAX) (Table 4.1), whereas the treatment PACGA restored germination. Broad sense heritabilities were comparable to the Bay × Sha experiment in all populations (Table 4.1). The absence of variation in the control treatment was evident as most accessions germinated nearly 100% (Figure 4.3A-C). The RILs show a transgressive segregation for the treatments PAC04 and PAC08 (Figure 4.3A, B). The Sha × Col showed the higher sensitivity range for the treatment PAC04 (Figure 4.3A, E), while the Ler × Sha population had the higher range for the treatment PACGA (Figure 4.3C, E, F). From all parental accessions, Sha is the most tolerant to PAC followed by Col and L*er*.

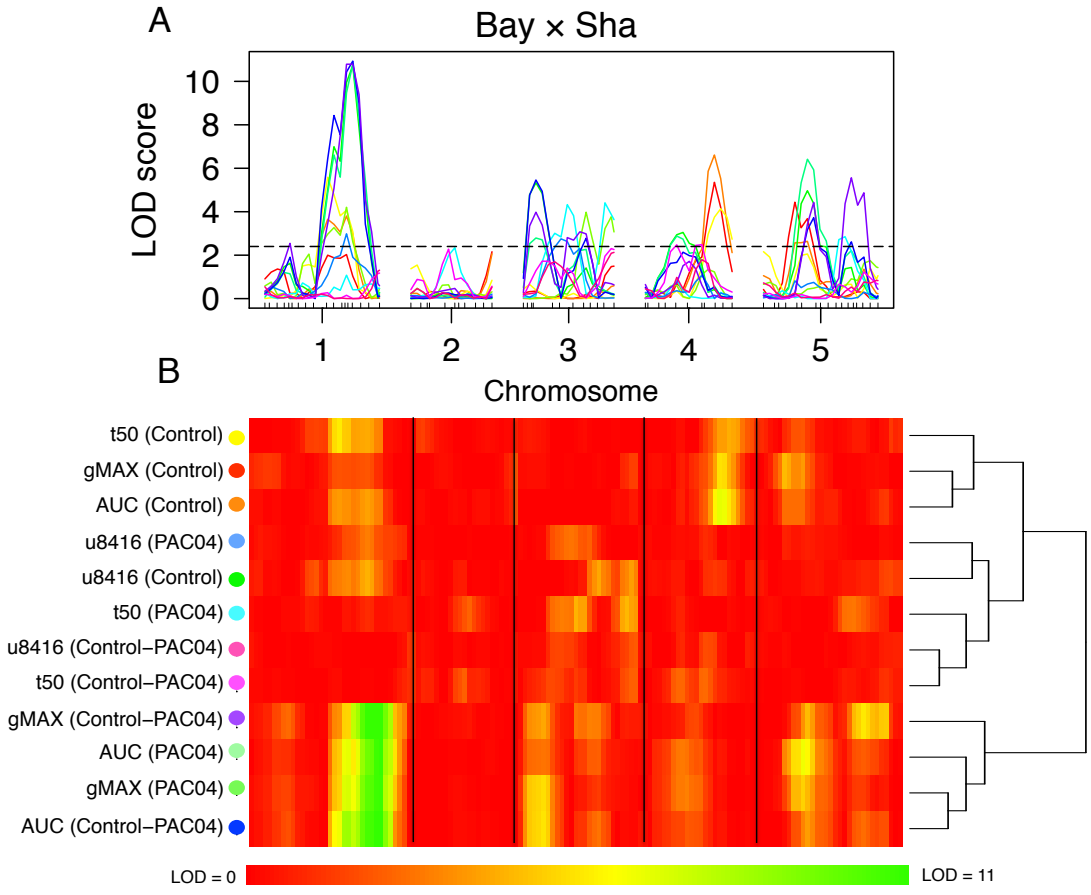

Figure 4.2. Genetics controlling pac sensitivity in the Bay × Sha RIL population. (A) QTL map of germination related traits (AUC, gMAX, t50, u8416) phenotyped under control and PAC (4 µM) conditions. The horizontal dotted line shows the significance threshold by running 1000 permutations. (B) QTL map represented as heat map organized by LOD score with the dendrogram.

Table 4.2. Summary of QTL mapping results for gMAX in the PAC treatment in the Bay × Sha RIL mapping population.

Chr	Nearest marker	Position (cM)[a]	LOD score[b]	Explained variance (%)[b]	Effect[c]
1	F5I14	69.6	7.8	16.7	+
3	MSAT305754	7.9	3.1	6.3	+
4	MSAT4.35	24.2	2.4	4.8	+
5	MSAT512110	41.8	2.2	4.4	+

[a] Position of the nearest marker in the genetic map.
[b] LOD score and explained variance estimated using the "fitqtl" model from the R/qtl package.
[c] Effect of QTL estimated μA-μB; where A refers to the Sha alleles and B to Bay alleles. Positive (+) effect means that Sha alleles at the nearest marker linked to the QTL increases the trait mean; negative (-) effect indicates that Bay alleles increase the trait mean.

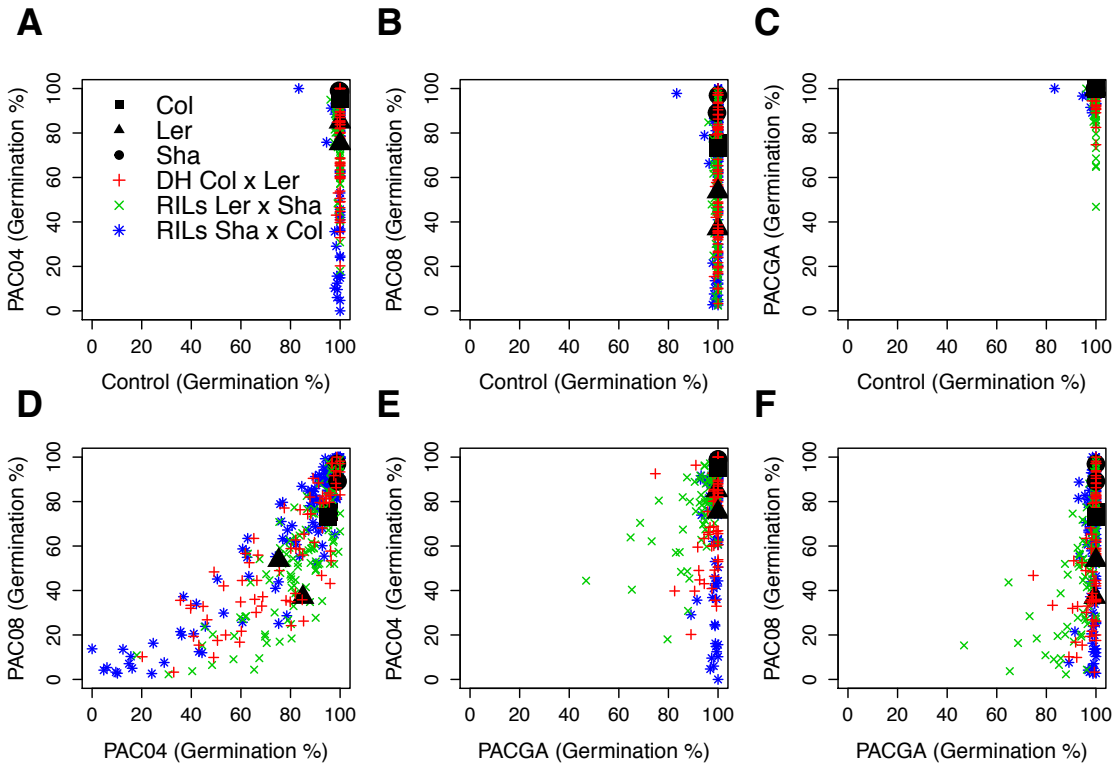

Figure 4.3. Paclobutrazol and gibberellin sensitivity in three *A. thaliana* mapping populations. (A-F) Germination of the Double Haploid (DH) Col × Ler mapping population, and the two Recombinant Inbred Lines (RILs) populations Ler × Sha and Sha × Col in different paclobutrazol (PAC) concentrations (4 μM and 8 μM) and when appling PAC (8 μM) + GA (8 μM GA_{4+7}). Mock control was performed using DMSO (same volume as PAC04). Values of the parental lines are indicated in replicates. Phenotypes were scored at maximum germination (after 120 hours).

Different QTLs affecting PAC and PACGA sensitivity were mapped in the three populations. To dissect the genetics controlling the effect of the PAC and PACGA treatments, only the trait AUC was analysed. This trait was highly correlated with gMAX (R^2 = 0.94), had a high heritability (Table 4.1) and integrates germination speed and final germination. In order to map QTL related with the response of

different PAC treatments the difference between germination in PAC04 and PAC08 was also analyzed. From the three populations the L*er* × Sha showed the highest number of QTLs (Figure 4.4A,B). The major effect QTLs are located at the lower end of chr 3, the middle of chr 4 and two regions on chr 5. The treatment PAC04 and Control – PAC04 showed the most significant QTLs located on chr 3 and chr 5 (Figure 4.4B, Appendix 5). The QTL located in the lower half chr 5 co-locates with the dormancy QTL *DOG1* (Bentsink et al., 2010). Interestingly the PAC08 treatment yields a different QTL located on chr 4 compared with PAC04. This QTL is also present for the traits PACGA − PAC08 and PACGA − PAC04. Interestingly, compared to Control not all PACGA treatments germinate 100 % and their subtraction yielded QTLs thus being more meaningful than using control. The absence of 100 % restoration by GA application indicates that GA sensitivity in these genotypes differs. This QTL co-locates with a heat germination QTL (Clerkx et al., 2004). A minor effect QTL is located in the lower half of chr 1 for the treatments PAC08 and PACGA (Figure 4.4B, Appendix 5). Dormancy and sucrose sensitivity QTLs have been mapped in this region (Clerkx et al., 2004). When the Sha × Col population was analyzed, few QTLs were mapped, mainly at top chr 3 and top chr 5, all related with PAC sensitivity (Figure 4.4C, D, Appendix 6). Sha and Col are the most PAC tolerant accessions. When the QTL profile from the Sha × Col population is compared with the other populations, many QTLs disappear, thus pointing to QTL alleles that Sha and Col share leading to higher PAC tolerance. The Col × L*er* population yielded two main QTLs located on top chr 1 for PAC08, and at the lower half chr 5 for PAC04 (Figure 4.4E, F, Appendix 7).

QTL validation in the Bay × Sha mapping population and the germination behavior of near isogenic lines

To validate the QTL mapping results we employed the procedure of selecting Near Isogenic Lines (NILs) from Heterogeneous Inbred Families (HIFs) (Tuinstra et al., 1997). We focused on the QTLs mapped on chr 1, 4 and 5 (Figure 4.5). Under control conditions (with stratification) all genotypes germinated (Figure 4.6A). Clear differences were observed between the HIF 044 derived NILs, which differ in the Bay or Sha alleles for the QTLs located on chr 1 for germination in PAC. The HIF 409 derived NILs did not show significant differences, therefore narrowing down the QTL region to 3.7 Mb (20.63 – 24.37 Mb, chr 1). The QTL located on top of chr 4 was

validated using HIF 011 derived NILs. No significant differences were observed for the NILs from HIFs 415 and 196. Both NIL pairs from HIFs 361 and 214 from chr 5 showed significant differences between the Bay and Sha alleles.

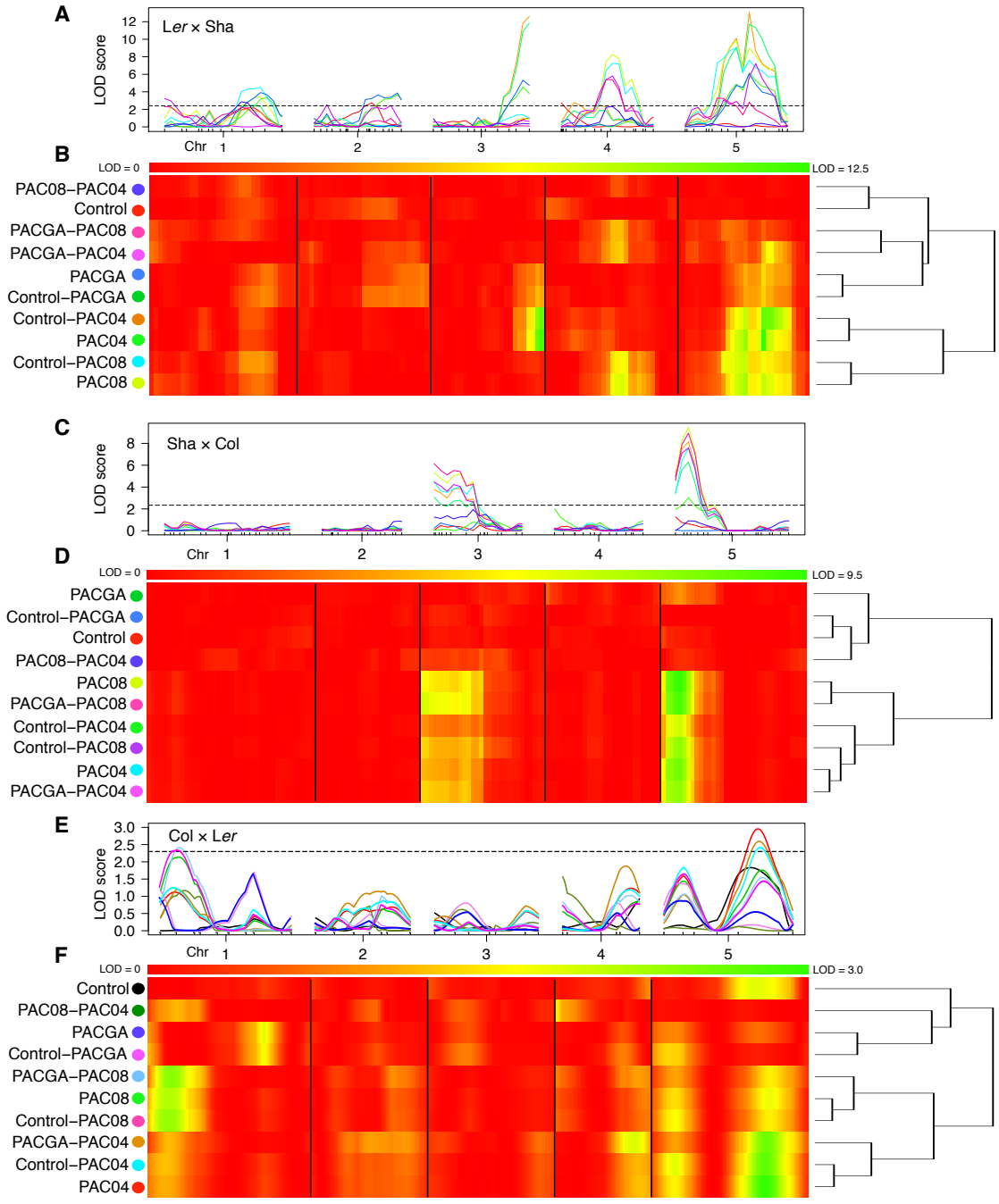

Figure 4.4. Clustered heat map of QTL controlling Pac sensitivity in three *A.thaliana* mapping populations. (A, C, E) QTL map of the germination related trait AUC (Area Under the Curve) phenotyped in the two Recombinant Inbred Lines (RILs) populations L*er* × Sha and Sha × Col and the Double Haploid (DH) Col × L*er* mapping population, respectively, in different paclobutrazol (PAC) concentrations (4 µM and 8 µM) and when appling PAC (8 µM) + GA (8 µM GA_{4+7}). The horizontal dotted line shows the significance threshold by running 1000 permutations. (B, D, F) QTL maps represented as clustered heat maps. Clustering on the right shows correlation between QTL profiles.

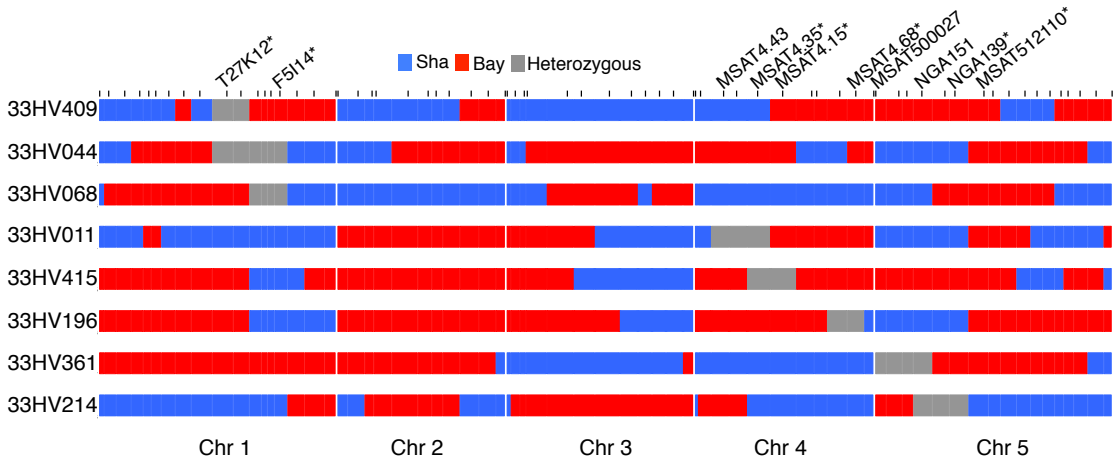

Figure 4.5. Graphical genotypes of Bay × Sha Heterogeneous Inbred Families (HIFs) used for QTL confirmation using NILs derived from these families. On top are indicated selected markers flanking the QTL regions; "*" indicate those markers with significant QTLs.

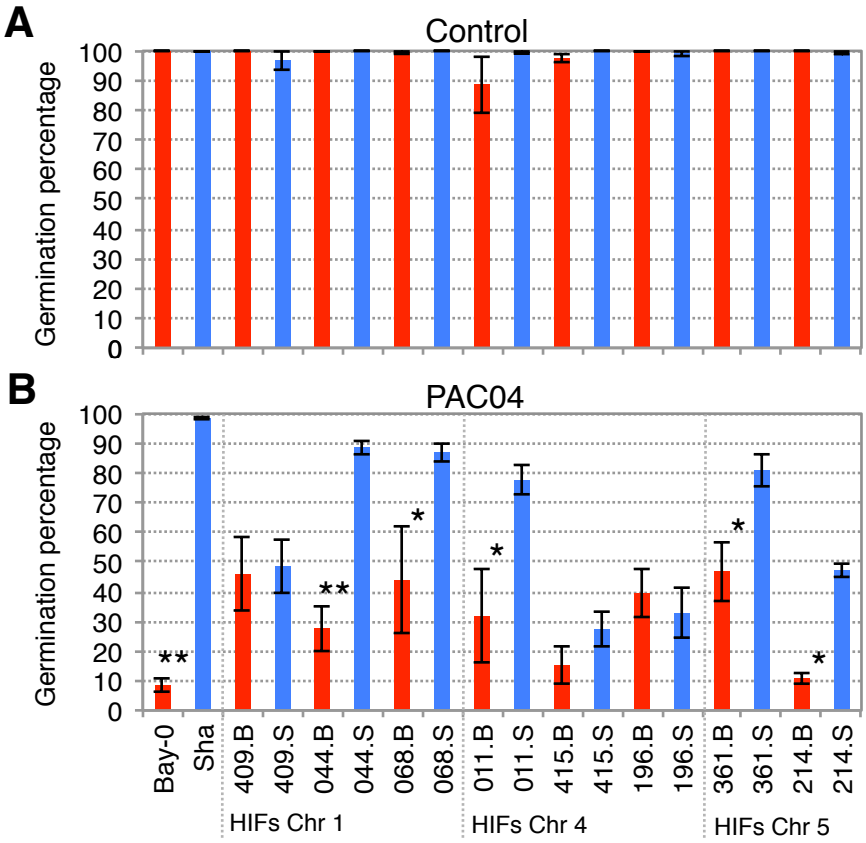

Figure 4.6. Paclobutrazol sensitivity in the Bay × Sha Heterogeneous Inbred Families (HIFs). Germination of Bay × Sha HIFs in (A) control and (B) paclobutrazol (PAC04, 4 µM) conditions. Phenotypes were scored at maximun germination (after 120 hours), after 18 weeks after ripening. All seeds were stratified for 4 days at 4 °C. Mean values ± standard error derived from three biological replicates are shown. P-values from a kruskall wallis test are indicated with asterisks at *P < 0.05 and **P < 0.01. Red colors indicate the Bay-0 alleles while red Sha alleles.

To know if differences in PAC behavior are translated into differences in seed dormancy, we characterized the germination of one HIF pair per chromosome. Germination in freshly harvested seeds under control conditions showed no strong differences for the HIF derived NILs. The stratification (cold) treatment did overcome the dormancy in most lines but significant differences were observed for this treatment between the different NILs from chr 1 and chr 4 (Figure 4.7B).

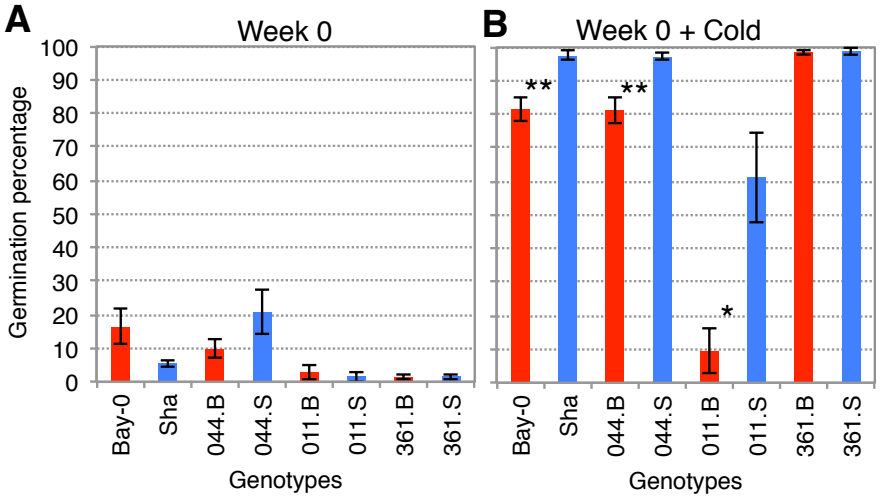

Figure 4.7. Effect of cold application on dormancy breaking in the Bay × Sha Heterogeneous Inbred Families (HIFs). Germination of Bay × Sha HIFs under (A) control, and (B) Cold conditions. Phenotypes were scored at maximum germination (after 120 hours). Experiment was conducted 1-2 days after seed harvesting, no stratification was done in A, 4 days at 4 °C stratification was done in B. Mean values ± standard error derived from three biological replicates are shown. P-values from a kruskall wallis test are indicated with asterisks at *P < 0.05 and **P < 0.01. Red colors indicate the Bay-0 alleles and blue Sha alleles.

To further study the effect of dry storage / after-ripening on dormancy release of the genotypes, germination was followed for 8 weeks (Figure 4.8). Sha is slightly dormant during the first two weeks of seed after-ripening compared with Bay, but it looses dormancy faster between weeks 4 and 8 compared to the latter. The NILs derived from HIFs 044 and 361 showed the same pattern as their parental genotypes. This observation points to a dormancy QTL, closely linked or pleiotropic with the PAC tolerance QTL. The NILs derived from HIF 011 are the most dormant genotypes. The *dog 1-2* and *aba 3-1* mutants fully germinated as expected (Figure 4.8). Col was the least dormant accession and germinated ~100 % after four weeks.

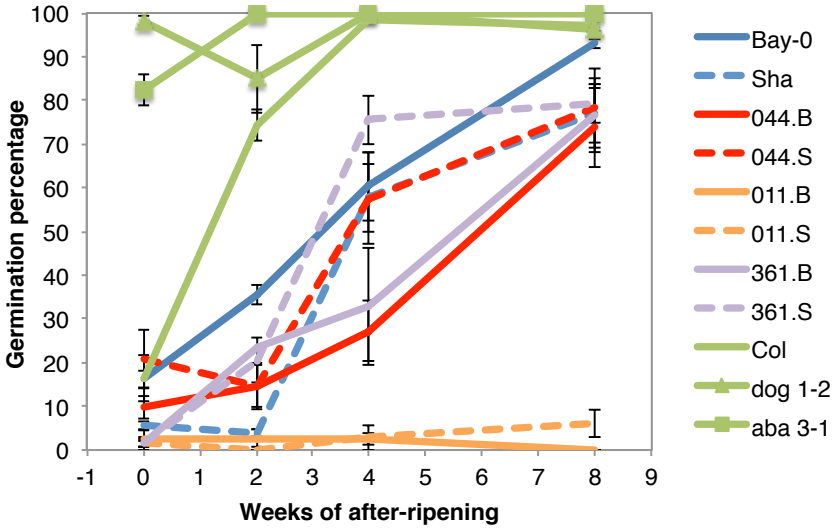

Figure 4.8. Germination profile of selected Bay × Sha Heterogeneous Inbred Families (HIFs) and selected mutants. Germination across time of selected Bay × Sha HIFs, Col and the mutants *dog 1-2* and *aba 1-3* (both Col background). Phenotypes were scored at maximun germination (after 120 hours).

The germination of partially or fully after ripened seeds from the HIF derived NILs for chr 1 and the parental genotypes (stored for 19 months) was further characterized under dark and cold conditions. In the treatment light without cold all the genotypes germinated > ~70 % (Figure 4.9A). All genotypes fully germinated under light + cold treatment (Figure 4.9B). Under the treatment dark without cold most of the genotypes did not germinate except Sha that germinated more than 80 % (Figure 4.9C). When dark + cold was tested, only the Sha accession germinated ~100 % (Figure 4.9D). No differences where observed among HIF 044 derived NILs, suggesting no specific involvement of light signalling for this QTL.

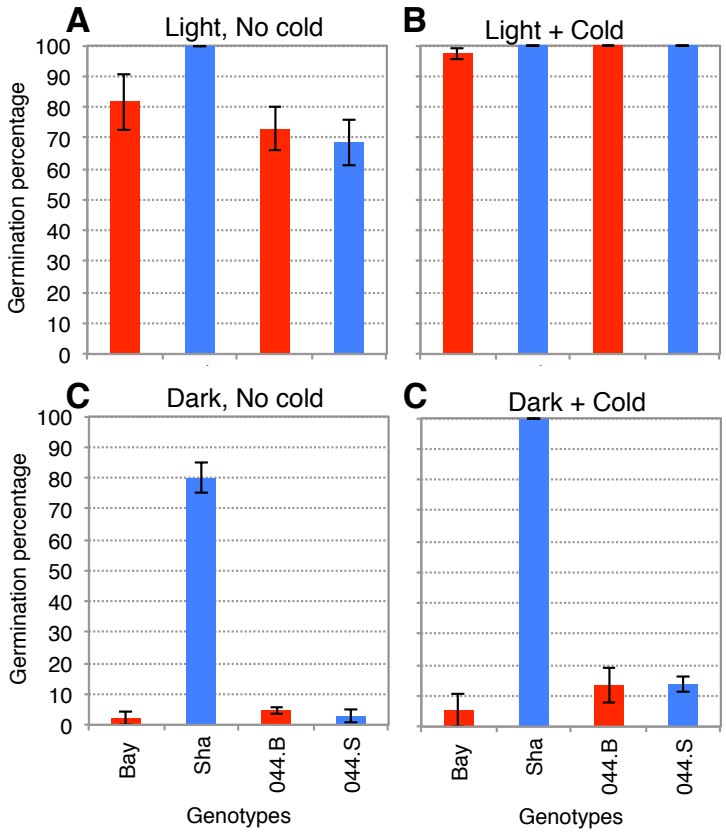

Figure 4.9. Effect of light and chilling application on germination in the Bay × Sha Heterogeneous Inbred Families (HIFs). Germination of Bay × Sha HIFs and selected ABA and GA mutants (together with their wild type accessions Col and L*er*) under (A) light, no cold (B) light + cold (4 days at 4 °C), (C) dark, no cold (D) dark + cold. Phenotypes were scored at maximun germination (after 120 hours). Experiment was conducted 19 months after seed harvest and storage at room temperature. Mean values ± standard error derived from three biological replicates are shown.

GWAS for paclobutrazol sensitivity

Association mapping (GWAS) was conducted to find potential genes affecting differential sensitivity to paclobutrazol including those causal for the detected QTLs in the study using RILs. The *A. thaliana* Hapmap population (Li et al., 2010) showed a wide response of not only PAC sensitivity but also germination under control conditions (Figure 4.10). The broad sense heritability for the evaluated traits in control and PAC conditions ranged from 0.77 to 0.84 and the coefficient of genetic variation was the highest compared to the experiments using RILs mentioned in the paragraphs above. This indicates a large genetic variation among the evaluated accessions (Table 4.1). The population contains four semi-dwarfs allelic to *ga5* (Barboza et al., 2013). Three of them were phenotyped in this experiment and showed 100 % germination under control conditions (Figure 4.10). Under PAC conditions one of them (Var 2-6), had a low germination thus indicating the effect of the background in the PAC

sensitivity. The traits gMAX and AUC were analyzed in this GWAS. When comparing the different GWAS profiles, no single PAC specific association was found (data not shown). To normalize for the differences present under control conditions, the subtraction Control – PAC04 was used for mapping PAC sensitivity (Figure 4.11A, B). Using this parameter for the trait AUC a significant association crossing the stringent Bonferroni threshold at the bottom of chr 5 was detected (Figure 4.11A). This marker is located in the gene *PSBO1* (At5g66570), which encodes for a protein involved in photosystem II (Allahverdiyeva et al., 2013). No obvious GA / ABA related genes were found in the 20 kb region flanking this gene (marker). When the GWAS profile for gMAX is analyzed, a group of markers just below the Bonferoni threshold is located in the middle part of chr 2 (Figure 4.11B). This marker is located in a gene that codes for a NAD(P)-linked oxidoreductase superfamily protein (At2g27680). It is relevant to notice that the flanking gene next to this marker (*CYP94C1*) is a cytochrome P450 (At2g27690, Appendix 8).

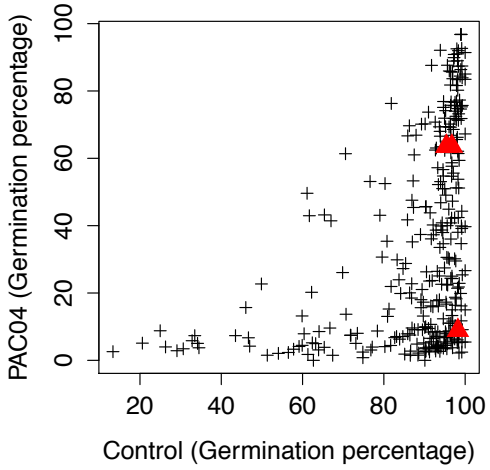

Figure 4.10. Paclobutrazol sensitivity in the Hapmap population. Correlation between germination in control and paclobutrazol (PAC 4 µM) condtions of the Hapmap population. Values of three semi-dwarfs accessions allelic to *ga5* are indicated with red triangles.

No co-localization between our main QTLs including those confirmed with NILs and the GWAS profiles was found. For the QTL located on chr 1, there is a significant marker located in a gene with unknown function (At1g51410), but similar to an *Eucalyptus gunnii* alcohol dehydrogenase and a potential candidate as an alcohol dehydrogenase in barley was reported to be involved in ABA / GA signalling (MacNicol & Jacobsen, 2001). In the vicinity of this QTL region, the *ABA4* gene is located of which a mutant was isolated on the basis of PAC tolerance (North et al., 2007). When the gMAX is analyzed for a SNP marker inside this gene (m43235, chr 1) it showed a P-value of 0.004, thus suggesting *ABA4* as possible candidate. This candidacy is also based on the PAC phenotype of the *aba4* mutants. Therefore an

allelism test (Weigel, 2011) was conducted using the mutant, the parental lines and the HIF 044 NILs. However no indications were obtained that the two natural alleles in this region complemented the *aba4* phenotype in a different way (data not shown).

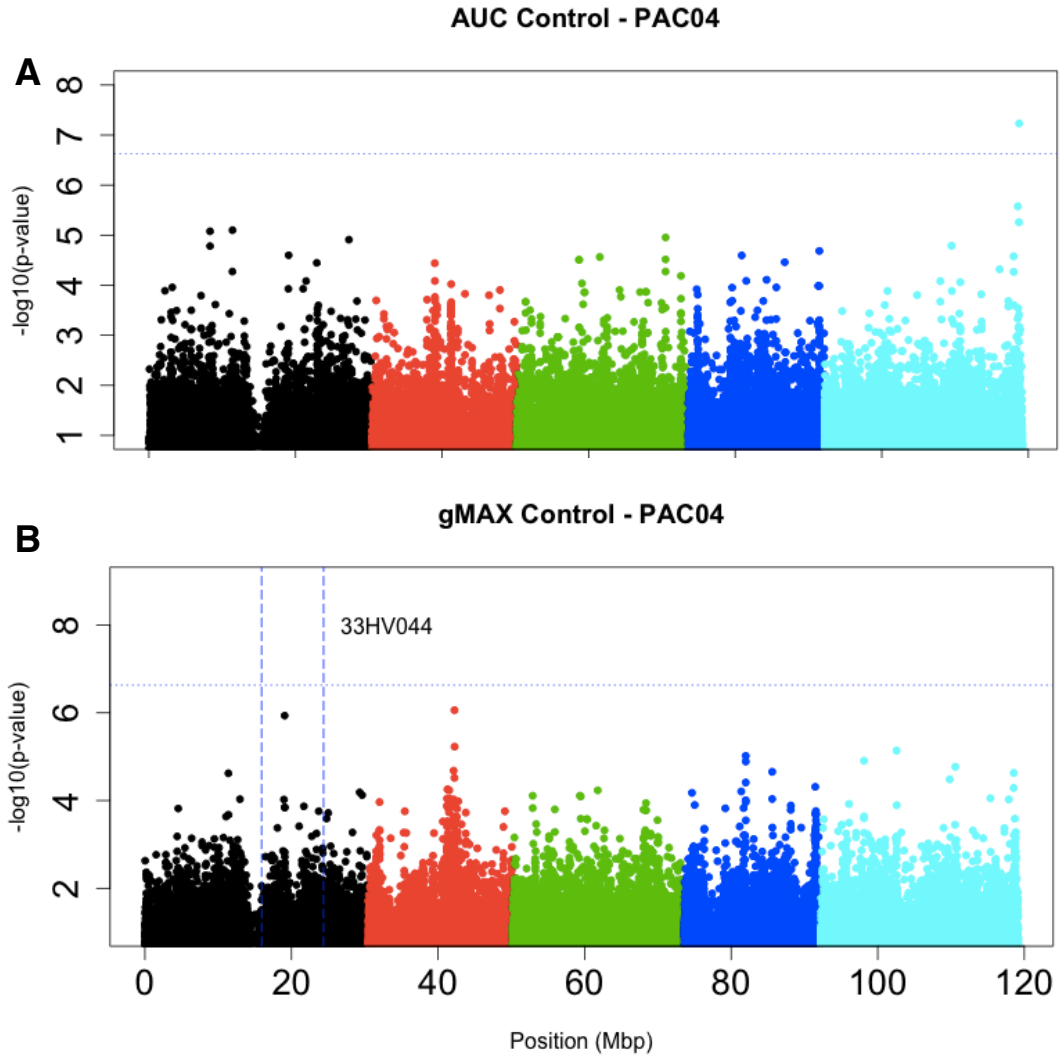

Figure 4.11. Genome wide association mapping for Paclobutrazol sensitivity in the Hapmap population. Genome wide association mapping profile for the germination related traits (A) AUC (Area Under the Curve) and (B) gMAX (maximum germination after 120 hours). The vertical line shows the position of the HIF044. The horizontal line show the 5% significance threshold with Bonferroni correction for 214,000 markers. The different colours represent different Arabidopsis chromosomes.

Germination behavior of selected Hapmap accessions

Accessions contrasting in their PAC tolerance were studied together with the NILs and mutants describe above with the aim to see if differences in PAC resistance relate to their seed germination behavior (e.g. affect dormancy and the germination response to cold and light). The PAC tolerant accessions Sha and Hovdala-2 behaved dormant in the germination test just after seed harvest contrasted by the sensitive accession Gr-1, which germinated more than 60 %. The same applies for the remaining treatments (light, dark, cold). To combine all phenotypes and genotypes in a single analysis, a principal component analysis was conducted (Figure 4.12A). The first 2 principal components explained 77 % of the variance. PCA1 mainly explains variance for all treatments except germination in Dark + Cold and Dark No cold which are explained by PCA2. Five main groups were identified. The first group, which is PAC resistant (defined as accessions that germinated > 80 % on PAC, sensitive accessions included in the PCA analysis germinated < 55 %), contains Sha, L*er*, DraIV, Zdrl, *ga5* and ABA mutants. Based on the variable map (Figure 4.12B), these accessions have a high germination under the treatments PAC04, No cold + light and No cold + dark. The second group includes the genotypes behaving as the *dog 1-2* mutant, Hovdala and Col. The third group is not highly related with the PCA1 but it is with PCA2. The genotypes Gr-1, HIF 044 NILs, Bay, and Mh-0 displayed low germination under the dark germination treatments. Group number four includes the *ga1* mutants together with the accession Alc-0, HV50 and LDV-25. This group showed a reduced germination in all treatments. Cluster number five contains those accessions that are dormant, for instance the accession Cvi. Most accessions of this cluster are sensitive to PAC, although a PAC tolerant accession (TOU-H-13) is located in this group.

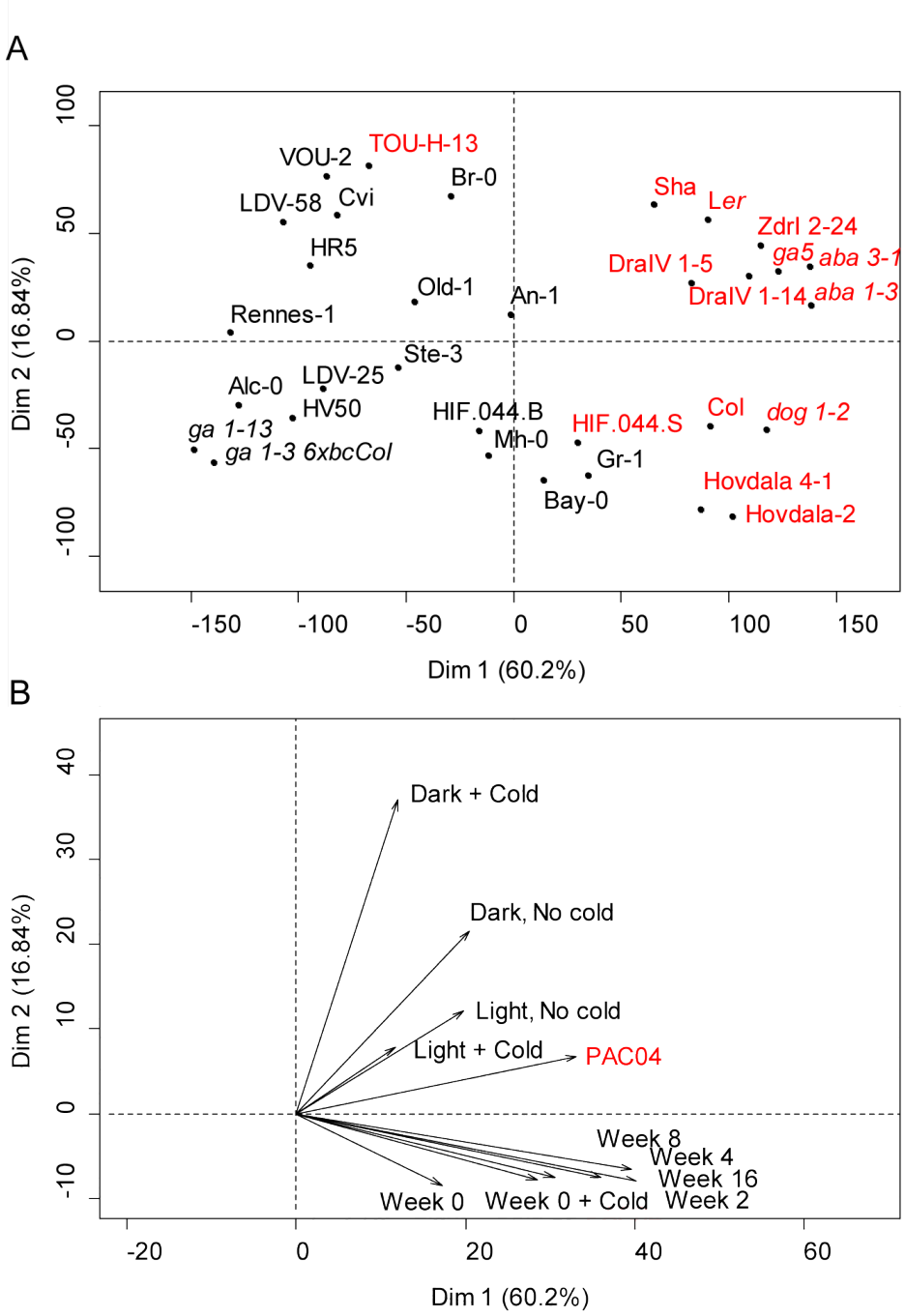

Figure 4.12. Principal component analysis (PCA) of the germination of 19 month old seeds. (A) PCA of selected *A.thaliana* accessions, ABA and GA mutants, based on the combined maximum germination data of different treatments. (B) Variable factor map. Genotypes colored in red showed a germination > 80 % under PAC conditions.

Discussion

Natural variation for PAC sensitivity is present among *A. thaliana* natural accessions but it remains less clear to what ecological relevant traits, such as seed dormancy, it relates to. Several QTLs for PAC tolerance mapped in regions for known dormancy QTLs. For instance the QTL on chr 5 in the L*er* × Sha population co-located with the

known dormancy locus *DELAY OF GERMINATION1* (*DOG1)* (Bentsink et al., 2006; Bentsink et al., 2010). *DOG1* was identified as a major QTL controlling seed dormancy in the L*er* × Cvi mapping population (Alonso-Blanco et al., 2003). Dormant accessions such as Cvi are sensitive to PAC. Thus PAC sensitivity may be seen as a pleiotropic effect of dormancy that implicates that a small reduction in GA levels in such genotypes already passes the threshold for GA needed to promote germination. Even though, *dog1* mutants can not germinate without GA as shown in the double mutant *dog1 ga1-3* (Bentsink et al., 2006) and are sensitive to the application of PAC (Nakabayashi et al., 2012). The latter indicates that such non-dormant mutants still require GA for germination probably because GA needs to overcome the inhibiting effect of ABA in these mutants. HIF 044 NILs show small differences in their germination profile across time suggesting that this locus acts as a dormancy QTL. Also as shown in previous studies (van der Schaar et al., 1997) it was demonstrated that accessions that differ little in their dormancy and in their PAC tolerance (e.g. Sha vs. Col) yield high effect QTLs in the mapping population indicating that different loci control these traits in these accessions. Using a L*er* × Col RIL mapping population, van der Schaar et al., (1997) mapped three main QTLs involved in Paclobutrazol sensitivity located on the lower half of chr 1, chr 3 and chr 5. The latter QTL seems to co-locate with the QTL mapped in this study using the DH Col × L*er* population and possibly with the DOG1 QTL (Bentsink et al., 2010). The same study indicated that in the position of the PAC sensitivity QTLs an effect of the light germination treatment was observed but that the effect of the PAC QTLs was larger. A similar pattern was also observed in this study because the effect of the QTLs on PAC treatments was higher than control.

A detailed study of the germination responses to a series of treatments in the Bay × Sha RIL population has been conducted (Joosen et al., 2011) and the identified QTLs can be compared with those identified in the present study. It appears that all QTLs identified in the Bay × Sha population in this study were also detected by Joosen et al. (2011) for other seed germination related traits, although their significance differed from those detected for PAC. The NIL with Sha alleles from HIF 044 showed a higher germination than the one with the Bay alleles under PAC conditions. This is in agreement with the validation of a QTL for ABA sensitivity (Joosen et al., 2011) and germination in NaCl (Galpaz & Reymond, 2010; Joosen et al., 2011) located near marker T27K12. The different populations yielded a plethora of QTLs not only

controlled by the application of different compounds such as PAC or GA, but the detection of some QTLs depended on the dose of PAC in specific populations. Regarding PAC dose effects, it has been described that overexpression of the ABA insensitive 5 (ABI5) transcription factor leads to hypersensitivity to low PAC doses (0.125 μM) (Piskurewicz et al., 2008). Additional examples of genes showing hypersensitivity to PAC are the calcium sensor calcineurin B-like (*cbl*) mutants (Li et al., 2013). PAC resistance has been used to select for ABA biosynthesis mutants (Léon-Kloosterziel et al., 1996; North et al., 2007) as well as ABA signalling mutants (Nambara et al., 1992). It has been observed that PAC tolerance in genotypes that are less GA dependent such as GA signalling related mutants, e.g *SPINDLY* (*SPY*) are able to germinate at 100 μM PAC. The *gai-t6* (loss of function allele of the DELLA protein GAI) was able to germinate at 1 μM PAC and showed longer stems than *GAI* plants grown on PAC (Jacobsen & Olszewski 1993; Peng et al., 1997). The occurrence of different genes / QTLs involved in PAC sensitivity indicate the presence of different mechanisms (loci / pathways) rather than a single gene affecting GA homeostasis in general and thereby directly affecting seed dormancy. This is not surprising taking as example the different DOG loci mapped in previous studies differing also across populations (Bentsink et al., 2010) or even the control of germination due to different treatments in the same population (Joosen et al., 2011). Fine mapping the QTL of the HIF 044 region and cloning the gene can provide further knowledge about the mechanisms by which this gene controls PAC tolerance.

GWAS did not identify loci co-localizing with mapped QTLs as observed in previous studies (Baxter et al., 2010). Our initial aim of combining QTL and GWAS to increase power and resolution to map the loci affecting the PAC sensitivity seemed partially fulfilled by the co-location of *ABA4* which had a GWAS QTL with a relatively low significance. However it could not be confirmed that *ABA4* underlies this QTL. Complications may rise when mapping complex traits involving different mechanisms for which a diverse set of genes may be involved. For instance it would be useful to test the effect of removing the testa in PAC sensitive accessions and see if germination is restored as shown for the *ga1-1* mutant (Debeaujon & Koornneef, 2000). The variation of mechanisms affecting PAC tolerance might also explain why it is not possible to map a single locus in the different QTL mapping populations segregating for PAC resistance. Another complication might deal with the GA biosynthesis step targeted by PAC, which is the conversion from the *ent*-kaurene to

ent-kaurenoic acid (Hedden & Graebe, 1985) performed by the P450 *ent*-kaurene oxidase (KO) (Yamaguchi, 2008). P450s are involved in different processes including the biosynthesis of other hormones such as brassinosteroids (Werck-Reichhart et al., 2002), thus additional natural variation at this step will increase the possibility to find multiple loci for PAC tolerance. A good example was the possible association found with the cytochrome P450 gene *CYP94C1* using GWAS. Yet this association remains to be confirmed, especially because it was just below the stringent bonferroni significance threshold. However, recent GWAS identified novel genes based on marker associations, which significance was near the bonferroni threshold (Meijón et al., 2013), thus highlighting the potential information behind small significance GWAS peaks.

Conclusions

Differences in Paclobutrazol sensitivity can be attributed partially to dormancy differences. However, several accessions that showed no differences in seed dormancy showed differences in PAC sensitivity. Complex genetic regulation of PAC tolerance was identified using QTL analysis. Different PAC sensitivity QTLs were validated and their seed germination behavior was characterized. These results pointed to several additive effects as observed in the high tolerance to PAC of some accessions. A wide spectrum of PAC responses was found in the Hapmap population although GWAS did not identify clear and highly significant causal loci.

Materials and methods

Plant material and experiments set up

The following populations were used to conduct our studies: Bay × Sha RIL population (n = 165) (Loudet et al., 2002); Ler × Sha RIL population (n = 114) (Clerkx et al., 2004); Sha × Col RIL population (n = 164) (Simon et al., 2008); Col × L*er* Doubled Haploid population (n = 75) (Wijnker et al., 2012); Hapmap population (n = 360) (Li et al., 2010) and from HIF derived NILs from the Bay × Sha mapping population (Joosen et al., 2011). All lines were grown together at greenhouse conditions and for the germination experiments each recombinant inbred line had two biological replicates (seed batches from two individual plants). Seeds were harvested and stored at room temperature (~ 23 °C, ~ 45 % humidity) prior to use. Seeds for the

QTL mapping experiments were stored at least two months prior to germination tests and the control treatment was conducted first to check if no residual dormancy was present. The GERMINATOR procedure (Joosen et al., 2010) was used to phenotype all experiments (except those in the dark, which were phenotyped manually). In all experiments with mapping populations seeds were stratified at 4 °C in a dark cold room. Thereafter the imbibed seeds were transferred to a germination chamber at 25 °C / 20 °C (day / night). The gibberellin inhibitor paclobutrazol (PAC, Sigma-Aldrich 46046) was applied at a dose of 4 μM (PAC04) and 8 μM (PAC08). The 4μM PAC was chosen based on preliminary experiments were this dose reduced 50% of germination in a subset of accessions (data not shown). The PAC stock solution was 25 mM and PAC was diluted using DMSO. The control solution contained equal amount of DMSO as the applied PAC04. The treatment PAC+GA was done applying 8 μM PAC + 8 μM GA_{4+7} (Duchefa, prod. No. G0938.1000). The GA stock solution was at a concentration of 25 mM (GA diluted in few drops KOH 1 M). All experiments included as control the accession Sha (PAC tolerant), Cvi (PAC sensitive) and parental accessions.

Image acquisition and data analysis

Image settings were the same as specified for the GERMINATOR package (Joosen et al., 2010). The values used for segmentation were the following ones: For S+R: Y 150 – 255, U 0 – 155, V 80 – 255 and for S-R: Y 140 – 255, 0 – 125, 80 – 255. For image processing the calculation mode “Area” and “Absolute” where use with the following settings for Variance (pixels) 25 and xy variance (mm) 1.

Heritabilities and coefficients of genetic variation were estimated as in previous studies (Keurentjes et al., 2007). For QTL mapping all traits were fitted in a linear model and coefficients were used for mapping (Trait = Genotype + Replicate + error). The trait gMAX was transformed with the arcsine function (asin command R). QTL mapping in the RIL populations was performed with the R/qtl package (Arends et al., 2010; Broman et al., 2003). The function “mqmscan” was used setting all markers as cofactors and which were later eliminated through backward elimination. To quantify the explained variance, the main QTLs were manually selected and fitted into a multiple QTL model using the function “fitqtl” (with Haley-Knott regression as selected method). In the DH population, the population type was set to “dh”. The genetic map was estimated using the est.map function (using Kosambi as map

function). Interval mapping was conducted in this population using the scanone function. The method of EMMAX was used for GWAS (Kang et al., 2010) using a kinship matrix to correct for population structure. Descriptive statistics, Anova tests and Kruskall Wallis test were done with R. Principal component analysis were performed with the package FactoMineR (Lê et al., 2008).

Chapter 5

HyPer: Hypocotyl perimeter as trait to semi-/automatically phenotype skotomorphogenic hypocotyls in *Arabidopsis thaliana*

Abstract

Hypocotyl elongation is a relevant trait in physiological and molecular studies. When hypocotyls are grown in darkness they become long, with closed small cotyledons lacking chlorophyll; a process called skotomorphogenesis. To study the factors / loci / genes controlling this process, efficient and fast phenotyping methodologies for hypocotyl length growth must be developed. A new method to phenotype skotomorphogenic hypocotyls named "HyPer" was developed and validated. The method comprises segmentation of images derived from dark grown hypocotyls and quantification of perimeter as main trait to study hypocotyl elongation. This was conducted by using ImageJ. To allow automatization of the image analysis pipeline, an application was developed on the basis of Montpellier Rio Imaging (MRI) visual scripting. To validate HyPer, a Quantitative Trait Loci (QTL) analysis was performed in the L*er* × Sha *Arabidopsis thaliana* Recombinant Inbred Line (RIL) mapping population. Different QTLs along the *A. thaliana* genome were mapped showing no differences between automatic and manual hypocotyl length measurements. Further QTLs were mapped for different treatments and it was shown an alternative for hypocotyl length measurements when dealing with treatments that inhibit germination. HyPer quantifies skotomorphogenic hypocotyls in a semi-/automatic manner. The procedure also provides an automatic report and descriptive statistics from the analyzed images. HyPer was validated using QTL analysis and it was shown its potential to phenotype hypocotyls in a wide range of sizes and different treatments.

INTRODUCTION

Recent advances in molecular technologies has led to an increase in size of mapping populations, e.g. large and multiple Recombinant Inbred Line (RIL) populations and populations that allow Genome Wide Association Studies (GWAS). In order to prevent phenotyping to become the bottleneck of genetic analysis, phenotyping methodologies need to advance in a similar fashion.

Hypocotyl length of young *A. thaliana* seedlings has been established as a suitable trait in physiological and molecular studies that is amenable to survey large quantities of samples (Borevitz et al., 2002; Gendreau et al., 1997; Lau & Deng, 2010; Sangster et al., 2008). When postgerminative seedlings grow in darkness, skotomorphogenesis takes place forming a long hypocotyl with an apical hook and closed cotyledons, which are small and lack chlorophyll (Lau & Deng, 2010). Therefore, hypocotyl growth is a well-suited trait for investigating processes controlling cell elongation and response to light. Numerous studies revealed that light and plant hormones play an essential role in the cellular and molecular basis of hypocotyl elongation (Gendreau et al., 1997; Lau & Deng, 2010). The plant hormone gibberellin (GA) is a plant growth regulator and has a major role in regulating cell elongation. Its biosynthesis and signalling is well understood (Hedden & Thomas, 2012; Yamaguchi, 2008). GA promotes hypocotyls elongation in seedlings grown in the dark by reducing the DELLA proteins levels, thus allowing the repressive transcription factors PIF3 and PIF4 to bind to their targets and promote skomorphogenesis and repress photomorphogenesis (Lau & Deng, 2010), allowing a promotive effect of GA on cell elongation. Therefore, hypocotyl growth in the dark is a potential tool to study natural variation for GA biosynthesis and signalling. Indications that natural variation in *A. thaliana* for GA responses using hypocotyls have been reported (Borevitz et al., 2002).

Automated methods to quantify hypocotyl lengths of single seedlings grown in the light (Cole et al., 2011) and in darkness (Miller et al., 2007; Wang et al., 2009) have been described. These published procedures employ high-resolution image analysis and aim to follow individual hypocotyl elongation rates in detail and require relatively high computational power. The latter restrict their application when a high number of hypocotyls need to be analyzed. Another example is the “HypocoTool”

(http://openwetware.org/wiki/HypocoTool), which requires manual marking of hypocotyl start and end, thus limiting the throughput.

This study describes (i) a semi-/automated method to quantify hypocotyl length of skotomorphogenic seedlings in large numbers and (ii) apply and validate the method in Quantitative Trait Loci (QTL) analysis, which aims to identify loci that are responsive to GA inhibition. In order to develop such a method, the high throughput methodology 'GERMINATOR' that has been developed to study seed germination based on the color contrast of the protruding radicle and seed coat was taken as example (Joosen et al., 2010). A comparable principle and imagine processing pipeline to measure skotomorphogenic hypocotyls by adjusting the growth surface and image analysis parameters was conducted. Measuring the perimeter (length of the outside boundary of a shape) of the hypocotyls avoids complications of hypocotyls that do not grow straight and fully parallel to each other. The procedures described also shows an alternative to germinate the seedlings and quickly transfer them to treatments that would have inhibited germination. For instance, substantial natural variation for Paclobutrazol germination sensitivity is present among *A. thaliana* accessions (van der Schaar et al., 1997). This means that even at low inhibitor doses (~ 5 μM) accessions will not germinate, thus making impossible to phenotype hypocotyl length in response to such an inhibitor.

Results

Image analysis

Hypocotyl perimeters can be employed as a proxy for hypocotyl length and allows automatic quantification using ImageJ's property to provide the perimeter of binary objects (Schneider et al., 2012). To measure hypocotyl size using perimeters one must distinguish it from its background. This can be achieved by thresholding the different colors. Images containing dark-grown seedlings can be segmented into hypocotyl and root, when a contrasting background is used (Figure 5.1A, B). Both, black and blue backgrounds can be easily segmented. The use of blue filter paper (same germination paper as in the GERMINATOR (Joosen et al., 2010) allows seeds to be easily stratified (2-4 days at 4 °C). After that, seeds can be germinated under continuous light and germination can be scored. Seedlings can be transferred to new growth media when seedlings reach uniform germination, which can be quantified using t50 as a parameter

(50 % of the germination). This is important when treatments that affect seed germination are applied. When images are acquired using standard conditions (e.g. same as in GERMINATOR (Joosen et al., 2010) it is possible to crop the images automatically with software such as Adobe Photoshop. Perimeters are then measured in the cropped images using ImageJ (Schneider et al., 2012). In order to automatize seedling measurement, all operations are grouped in a single application using visual scripting for ImageJ (Baecker & Travo, 2006). The application was named 'HyPer' (Hypocotyl Perimeter). In summary, it enhances the contrast of the image and runs a macro that applies a color threshold, performs binary particle analysis, and reports results. One advantage of using imageJ is that it can save overlay masks and labels for each analyzed hypocotyl, thus having control of the quantified objects (Figure 5.1A). In a second step, an automatic R script automatically reads the results from ImageJ and generates a report showing relevant summary plots from the analyzed images. Finally, tables containing mean values and descriptive statistics per image are saved together with the report. The results obtained with Hyper were contrasted with manual measurements and found a high correlation between both methods (Figure 5.1C).

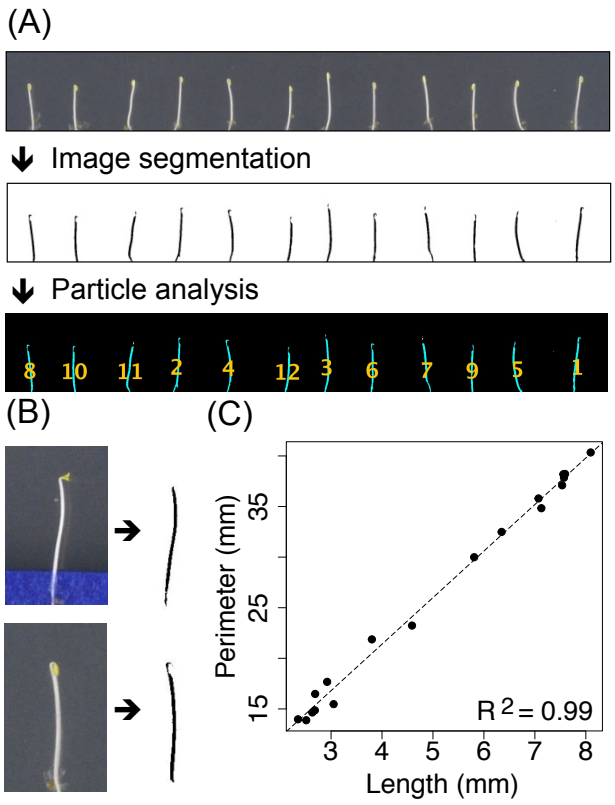

Figure 5.1. Skotomorphogenic hypocotyl image segmentation. (A, B) Image segmentation of cropped images using different backgrounds (conducted with ImageJ). (C) Correlation between mean values from images analyzed automatically (perimeter) and real length. Linear regression trend line is shown with the dotted line. Coefficient of determination (R^2) is shown on the right down corner.

HyPer validation in genetic studies.

To validate the procedure an experimental set up was designed (Figure 5.2A) and different *A. thaliana* accessions were grown in different treatments. In addition it was tested hypocotyl elongation on different growth surfaces (Figure 5.2B-D) and the effect of sucrose in three accessions. A wide range of sizes in hypocotyls from seedlings measured at 3, 5 and 7 days after germination was quantified (Figure 5.3A) in the presence and absence of sucrose and with or without blue filter paper (Figure 5.3A, B). Again, the comparison of mean values between perimeter and manually measured length revealed that values resulting from both measurement procedures are highly correlated (Figure 5.3C, D). The accession Cvi showed the highest means in all treatments, L*er* showed the lowest. Using the blue filter paper slightly decreased the length of the hypocotyls. It was compared gelrite and agar as gelling agents in the media to test which one facilitates segmentation of the images. As gelrite is clearer than agar, the correlation between perimeter and length was higher on gelrite than agar (R^2 0.93 for gelrite, 0.79 for agar).

To increase the scale of the experiments and for further validation, a QTL mapping experiment using the L*er* × Sha mapping population (Clerkx et al., 2004) was conducted. As above, the population was phenotyped using both the perimeter and length from the same images. The RIL population was grown in treatments that differed in the application of GA and the GA inhibitor paclobutrazol. For the RIL population broad sense heritability ranged from 0.59 (length in control) to 0.78 (perimeter in PAC+GA) (Table 5.1) indicating genetic variation between the lines. Furthermore, mean perimeter and length from each line were highly correlated in both biological replicates (Figure 5.4A). The coefficients of variations were similar in automatic and manual measurements (Table 5.1, Figure 5.4B). Pictures containing false positives were observed when the overlay masks were manually analyzed. In this way, artifacts (e.g. bubbles in the medium) and incorrect/partial segmentation due to excess of humidity were observed. When QTL mapping was performed, no differences were detected in explained variance and QTL number between length and perimeter (Figure 5.4C, D and Table 5.2). This indicates that the method and experimental design are suitable and accurate for QTL mapping. The major peak was located near the *ERECTA* locus that showed the highest explained variance of all mapped QTLs.

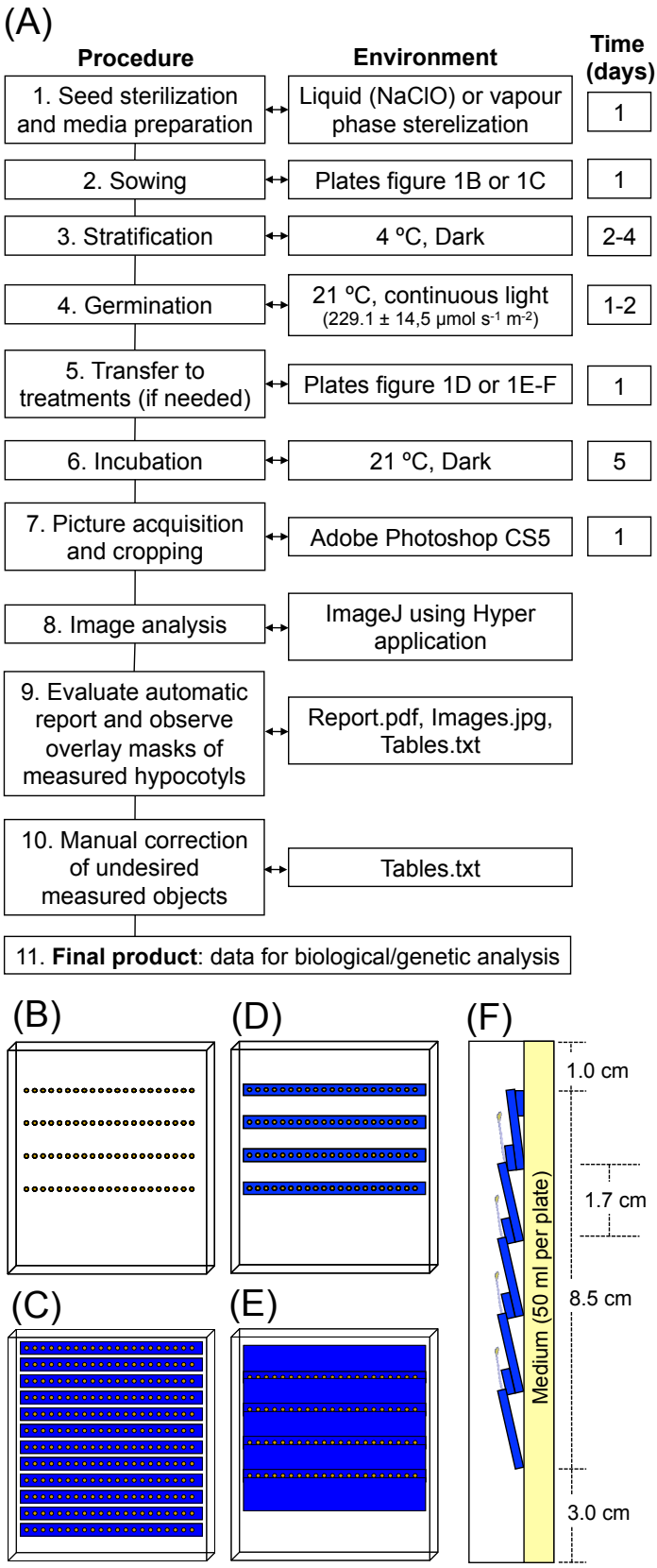

Figure 5.2. Experimental set up to quantify skotomorphogenic hypocotyl perimeters using image analysis. (A) Usual experimental set up to phenotype skotomorphogenic hypocotyls. Different plate choices for hypocotyl growth (B, standard plate; C, germination plate on blue filter paper; D, growth on blue filter paper; E-F, growth on blue filter paper all background blue).

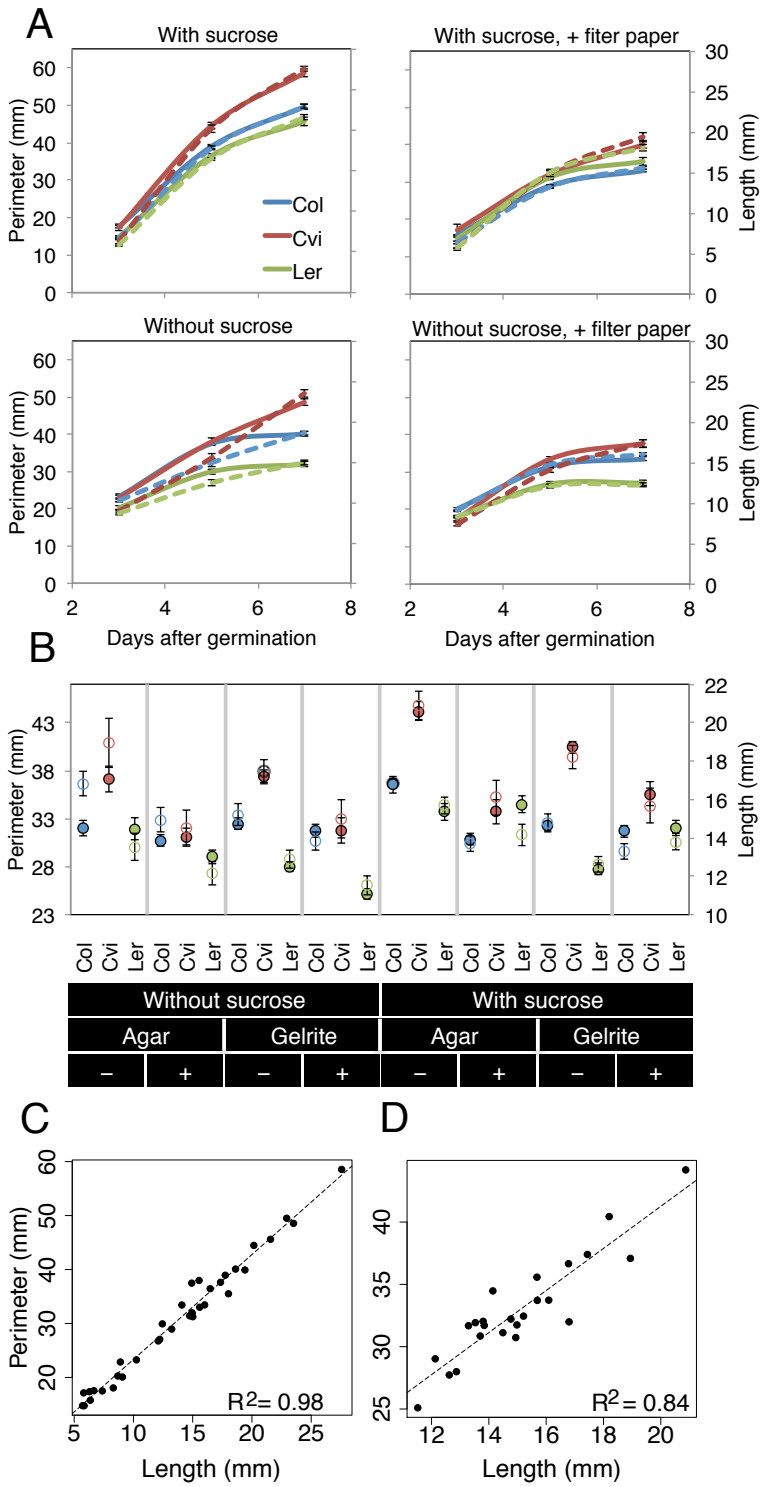

Figure 5.3. Validation of HyPer using the Col, Cvi and L*er Arabidopsis thaliana* accessions. (A, B) Hypocotyls were grown on different treatments and quantified automatically (perimeter) and manually (length, dotted lines). (B) Hypocotyls were grown for 5 days; filled dots represent perimeter, and empty dots length. Dashes show the standard error of the mean. (C) Correlation between mean values from images analyzed automatically and manually for Figure 5.4A and (D) 4B. Linear regression trend line is shown with the dotted line. Coefficient of determination (R^2) is shown on the lower right corner.

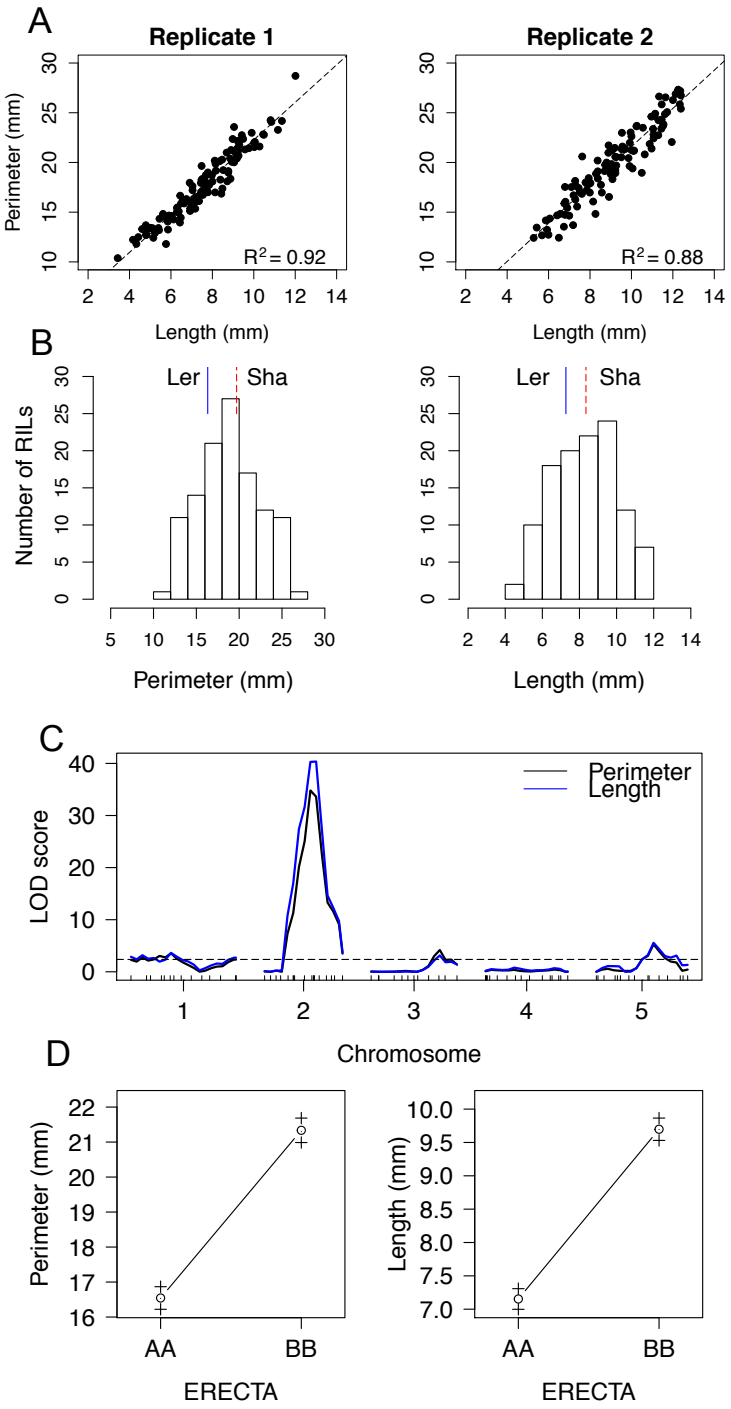

Figure 5.4. Validation of HyPer using the L*er* × Sha *Arabidopsis thaliana* mapping population. (A) Correlation are shown between mean values from images analyzed automatically (perimeter) and manually (length). Linear regression trend line is shown with the dotted line. Coefficient of determination (R^2) is shown on the lower right corner. (B) Frequency distributions for perimeter and length. Values of the parental lines are indicated. (C) QTL map of hypocotyls grown under control conditions and phenotyped automatically (perimeter) or manually (length). The horizontal dotted line shows the significance threshold by running 1000 permutations. (D) Mean values (± standard errors) of perimeter and length in the L*er* (AA) and Sha (BB) alleles at the *ERECTA* locus.

Table 5.1. Descriptive statistics for automatic (perimeter) and real length hypocotyl quantification. The different traits were evaluated in the L*er* × Sha mapping population.

Trait	Treatment	$\overline{X}$ ± (SD)		$[V_G]^a$	$[V_E]^b$	$[H_2]^c$	$[CV_G]^d$
Perimeter	Control	18.80	(3.45)	9.34	5.03	0.65	11.93
Length	Control	8.32	(1.76)	2.28	1.61	0.59	15.23
Perimeter	PAC	9.78	(1.91)	3.29	1.19	0.73	11.16
Perimeter	PAC + GA	16.91	(3.51)	10.99	3.04	0.78	10.32

[a]Among genotype variance.
[b]Enviromental variance (estimated as total genotype variance – V_G).
[c]Broad sense heritability (V_G/V_G+V_E).
[d]Coefficient of genetic variation $\left(100 \times \sqrt{V_G}\right) \big/ \overline{X}$.

Table 5.2. Summary of QTL mapping results for control in the L*er* × Sha mapping population.

Chr	Nearest marker	Position (cM)[a]	LOD score[b]	Explained variance (%)[b]	Effect[c]
			Perimeter / Length	Perimeter / Length	
1	CIW12	35	3.2 / 3.5	4.6 / 4.5	+
2	ERECTA	42.9	22.8 / 26.4	51.1 / 56.2	-
3	F8J2	60	3.3 / 3.3	4.9 / 4.2	+
5	K9D7	50	5.7 / 6.4	8.8 / 8.6	-

[a] Position of the nearest marker in the genetic map.
[b] LOD score and explained variance estimated using the "fitqtl" model from the R/qtl package.
[c] Effect of QTL estimated µA-µB; where A refers to the L*er* alleles and B to Sha alleles. Positive (+) effect means that L*er* alleles at the nearest marker linked to the QTL increases the trait mean; negative (-) effect indicates that Sha alleles increase the trait mean.

To further test the HyPer application, the effect of adding the gibberellin inhibitor paclobutrazol (PAC) and the application of this compound together with gibberellins (PAC+GA) was tested. The PAC+GA treatment restores the height of the hypocotyls in comparison with the PAC treatment and was thus highly correlated with the control (Figure 5.5A). QTL mapping by using the mean values from these treatments yields similar QTL profiles as the one obtained in control (data not shown). The same happens when all treatments are fitted in a linear model (Perimeter = replicate + treatment + genotype + error) and coefficients are used for mapping. Interestingly RILs show differential responses to these treatments. Thus to map the responses, the subtraction between treatments (e.g. Control – PAC) was tested. Again a strong QTL was located near the *ERECTA* locus. The subtraction Control – PAC, increases the significance of the QTL located on Chr 3 (Figure 5.5B, Table 5.3) thus pointing to a treatment respond QTL. A minor effect QTL response is mapped on Chr 1 using the subtraction (PAC+GA) – PAC. This suggests a possible role of these loci in GA sensitivity.

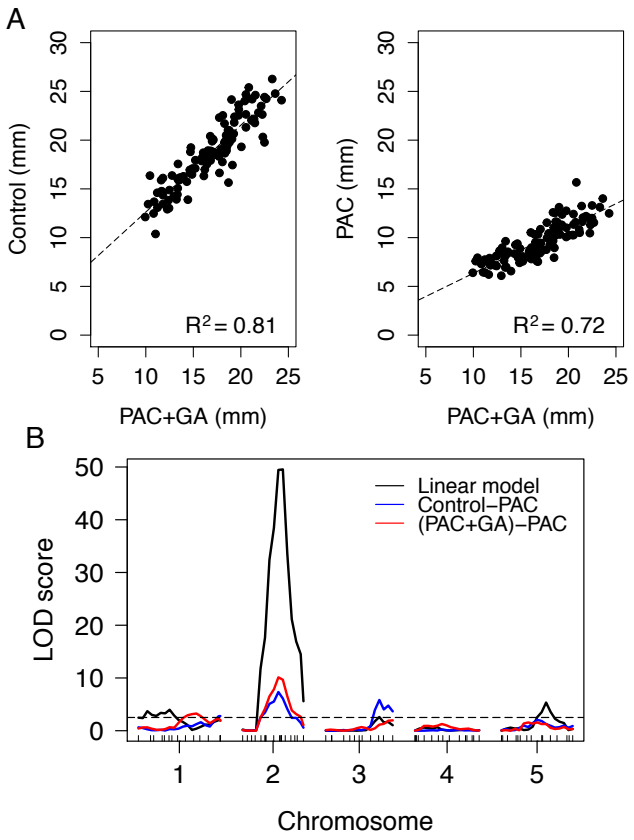

Figure 5.5. Validation of HyPer by mapping GA responses using the L*er* × Sha mapping population. (A) Correlations between the treatments: "PAC" (paclobutrazol), and PAC+GA (paclobutrazol+GA_{4+7} added at the same time). Linear regression trend line is shown with the dotted line. Coefficient of determination (R^2) is shown on the lower right corner. (B) QTL map of the response to the different treatments. Linear model shows the QTL map when the automatically phenotyped data from control, PAC and PAC+GA is analyzed together using a linear model. The horizontal dotted line shows the significance threshold by running 1000 permutations.

Table 5.3. Summary of QTL mapping results for perimeter in the L*er* × Sha mapping population.

Trait	Chr	Nearest marker	Position (cM)[a]	LOD score[b]	Explained variance (%)[b]	Effect[c]
Linear model	1	CIW12	35	3.7	4.4	+
Linear model	2	ERECTA	42.9	29	61	-
Linear model	3	F8J2	60	2.1	2.4	+
Linear model	5	K9D7	50	6.3	7.8	-
Control−PAC	2	ERECTA	42.9	5.1	16.4	-
Control−PAC	3	F8J2	60	4.1	13	+
(PAC+GA)−PAC	1	GENEA	65	2.4	6.7	+
(PAC+GA)−PAC	2	ERECTA	42.9	6.9	21.7	-

[a] Position of the nearest marker in the genetic map.
[b] LOD score and explained variance estimated using the fitqtl model from the R/qtl package.
[c] Effect of QTL estimated µA-µB; where A refers to the L*er* alleles and B to Sha alleles. Positive (+) effect means that L*er* alleles at the nearest marker linked to the QTL increases the trait mean; negative (-) effect indicates that Sha alleles increase the trait mean.

Discussion

In this chapter a method to quantify *A. thaliana* skotomorphogenic hypocotyls length based on their perimeter using ImageJ is presented. This method allows semi-/automatization of the image analysis and therefore increasing the scale of the experiments. This freely available methodology enables to conduct large experiments such as QTL mapping in a quite short time (~14 days including sowing, stratification, growing, image acquisition and analysis), with reduced costs, no complex set up, using open source software (ImageJ and R), easy installation and no special computer hardware is needed. In a single experiment 8766 hypocotyls were measured derived from 718 images. To manually phenotype the control conditions for the QTL mapping experiment only, a single user needs more than ~16 hours of work (240 images). When using HyPer in a standard personal computer (1.8 GHz), batches of 45 images can be analyzed in ~ 9 minutes (240 images will be analyzed in less than one hour). The method runs automatically but allows manual inspection of quantified hypocotyls and removal of artifacts. Furthermore a report is generated providing preliminary analysis to the users together with summary statistics per image. The simple link between ImageJ and R allows users with knowledge in R or ImageJ programming to modify and improve any of the scripts for their own requirements. The method includes a step in which seedlings are grown in blue filter paper, allowing easy application of various treatments, and thus it may help to identify novel responses. One limitation of the method is that light grown hypocotyls cannot be monitored. These contain large cotyledons that will represent a higher proportion of the measurements than the actual hypocotyls. Options to phenotype light grown hypocotyls are already available (Cole et al., 2011). Another HyPer limitation is the inability to phenotype other traits than perimeter. For instance, HYPOTrace can phenotype apical hook opening, phototropism and nutation (Wang et al., 2009). Furthermore, HyPer is sensitive to the quality of the images, which means hypocotyls need to be separate and displaying sufficient contrast with the background. The method is agnostic to the physiological effect (trade-off) of using the blue filter paper to grow the hypocotyls.

QTL analysis showed the potential to use HyPer to phenotype skotomorphogenic hypocotyls in an efficient manner. High heritabilities, QTLs with major and minor effects validate the method for usage in genetic analysis. The *ERECTA* locus was the main QTL in all treatments, even in the response to the PAC and GA treatments. A locus at or near *ERECTA* have been mapped for this trait in

previous studies (Borevitz et al., 2002; Botto et al., 2003; Wolyn et al., 2004). Van Zanten et al., (van Zanten et al., 2009) suggest to the possibility that *ERECTA* modulates elongation growth through GA but independently from the classical GA-signalling pathway. Two specific respond QTLs located in Chr 1 and Chr 3 were mapped. It remains to validate this loci and further fine mapped the region in order to identify the genes controlling PAC or PAC+GA sensitivity. HyPer is now available to be used in different populations especially with the current availability of *A. thaliana* accessions that will challenge the procedure and hopefully identify new loci/genes controlling phenotypes/responses of interest.

Conclusions

A method named "HyPer" to quantify *A. thaliana* skotomorphogenic hypocotlys was validated. The method not only allows semi-/automatization to phenotype skotomorphogenic hypocotyls, but also generates an automatic report and descriptive statistics of the analyzed images. HyPer was validated performing a series of experiments including quantitative trait loci (QTL) analysis using different treatments. It was shown that the method can also be used to phenotype hypocotyls grown on difficult treatments such as using germination inhibiting compounds.

Methods

Plant material and growth conditions

Initial experiments were conducted using the Col, Cvi and L*er A. thaliana* accessions. QTL mapping was done using the L*er* × Sha mapping population (Clerkx et al., 2004). Each recombinant inbred line had two biological replicates (seed batches from two individual plants), each containing ~15 hypocotyls. Parental lines had four repetitions. Plants were grown in half MS medium (Duchefa prod. No. M0222.0050), pH 5.8, 0.64 % agar (for germination, Plant agar Duchefa prod. No. P1001.1000), and 0.9 % agar (for seedling growth). In the experiment using gelrite (Duchefa prod no. G1101.1000), a concentration of 0.44 % was applied. The gibberellin inhibitor paclobutrazol (PAC, Sigma-Aldrich 46046) was applied at a dose of 4 μM. The PAC stock solution was 25 mM and PAC was diluted using DMSO. Mock control contained equal DMSO dose as the applied PAC. The treatment PAC+GA was done applying 4 μM PAC + 4 μM

GA_{4+7} (Duchefa, prod. No. G0938.1000). The GA stock solution was at a concentration of 25 mM (GA diluted in few drops KOH 1 M). The size of the blue filter paper (Anchor Paper Company, Seed Germination & Industrial Purchasing) was 5 mm height ×10 cm width.

In all experiments seeds were stratified at 4 °C in a dark cold room. To induce germination in the experiments conducted with Col, Cvi and L*er*, seeds were exposed to continuous light for three hours and after that plants were covered with two layers aluminum foil and placed inside a dark box. For the QTL mapping experiment blue filter paper was used to germinate the seeds under control conditions (same as in Figure 5.2C) and thus avoid the PAC germination inhibition effects. Germinated seeds were transferred to different treatments at germination t50 (~26-30 hours after incubation) thus normalizing the time of germination. Thereafter, plates were covered with two layers of aluminum foil and placed inside a box. A more complex background was used to test possible enhancements in the methodology (Figure 5.2E-F). Hypocotyls were imaged after 5 days of growth.

Image acquisition, image processing and data analysis

The same camera and settings as the ones described for GERMINATOR were used (Joosen et al., 2010). Before imaging, hypocotyls in near contact with other hypocotyls were manually arranged with forceps. Obtained images were automatically cropped using Adobe Photoshop (same action as GERMINATOR).

MRI visual scripting (Baecker & Travo, 2006) was used to create the HyPer application in ImageJ. It performs contrast enhancement (0.005) and runs a macro that makes color threshold, make binary, particle analysis (Perimeter), and report results (.txt table). The following values were used for color threshold YUV (Y is the luma-brightness component; UV the chrominance-color component) in all experiments containing blue filter paper: 150, 255; 90, 200; 85, 255. Particle analysis size was set between 300 and 8000 pixels and circularity between 0.005-0.15. Circularity is a relevant threshold since hypocotyls are not circular, however this should be carefully adjusted when very small hypocotyls are evaluated. For instance, the PAC and PAC+GA treatments considerably reduced the hypocotyl perimeter of several recombinant inbred lines, thus for these treatments the following particle size/circularity were used 150-8000 / 0.005-0.3. The application saves the overlay masks of the measured particles. Partially or incorrectly segmented hypocotyls were not included in the analysis. The application automatically runs an R (R Core Team,

2013) script from ImageJ. This script generates an automatic report showing boxplots for relevant measurement. Included measurements were the Y coordinate since hypocotyls perimeters are correlated with this measurement. Also circularity and its related traits are shown. Because hypocotyls are not circular, artifacts can be spotted with these measurements. HyPer generates descriptive statistics tables for each image (mean, standard error, standard deviation, coefficient of variation, and number of observations). Moreover a linear model fitting the effect of Perimeter ~ Image is generated. Boxplots were generated using the packages ggplot2 (Wickham, 2009) and gridExtra, and need to be installed for proper functioning of the R script. Heritabilities and coefficients of genetic variation were estimated as in previous studies (Keurentjes et al., 2007). QTL mapping was performed with the R/qtl package (Arends et al., 2010; Broman et al., 2003). The function "mqmscan" was used setting all markers as cofactors and later eliminated through backward elimination. To quantify the explained variance, the main QTLs were manually selected and fitted into a multiple QTL model using the function "fitqtl" (Haley-Knott regression was the selected method).

Chapter 6

General discussion

Plant growth regulators are small molecules that at low concentrations have an effect on the growth and development of plants. Among them, Gibberellins (GAs) are well known especially for their role on seed germination, flowering and cell elongation. The main goal of this thesis was to identify natural variation for GA biosynthesis and GA signalling in the model plant *Arabidopsis thaliana*.

The first half of the thesis (chapter 2) deals with the study of natural semi-dwarf accessions carrying mutations in a GA biosynthesis locus called *GA20ox1* or *GA5*. In addition, the possible modifications in the root system of these semi-dwarfs were studied as well as their growth performance under water limiting conditions (chapter 3). In the second half of this thesis, the natural variation of the effect of a GA inhibiting compound called paclobutrazol and the effect of adding this compound together with GAs were studied using seed germination (chapter 3) and hypocotyl elongation (chapter 4) as biological tests systems. For high – throughput analysis of hypocotyl lengths of many genotypes a semi - / automatic procedure was developed (chapter 4). In this general discussion, a summary and discussion of major findings is provided. In addition, perspectives and possible directions to continue the study of natural variation for GA biosynthesis and signalling are given.

***GA20ox1* (*GA5*) a case of convergent evolution**

GA 20-oxidases are enzymes that together with GA 3-oxidases catalyse the final steps of bioactive GA biosynthesis (Yamaguchi, 2008). *A. thaliana* contains five paralogs encoding for GA 20-oxidases (Rieu et al., 2008; Yamaguchi, 2008). The gene *GA20ox1*, or *GA5* as the original mutant was named (Koornneef & van der Veen, 1980), is together with *GA20ox2, -3*, the main paralogs, which means the absence of these three genes will result in severe dwarfism and sterility (Plackett et al., 2012). Remarkably *GA20ox1* (*GA5*) is the only paralog that will induce semi-dwarfism when

non-functional (Plackett et al., 2012; Rieu et al., 2008). Surprisingly semi-dwarfism occurs in natural *A. thaliana* accessions. A semi-dwarf (mutant) is defined in this study as a plant with half the height of a genetically related accession (wild type). The presence of natural variation of a loss of function *GA20ox1* (*GA5*) allele was first described by El-Lithy et al., (2006) in the Kas-2 semi-dwarf accession. In the present study semi-dwarf natural accessions from different parts of the distribution range of *A. thaliana* were characterized by allelism tests and sequencing (Barboza et al., 2013). In addition population molecular genetic studies were performed, which describe the pattern of variation as well as possible signatures of selection.

Natural variation for GA biosynthesis controlling plant height is mainly found for *GA20ox1* (*GA5*), based on the current knowledge. Most of the studied semi-dwarfs were alleles of *GA20ox1* (*GA5*). Only a few accessions complemented to the wild type phenotype in the cross *ga5* × accession, thus indicating that allelic variation at other loci affect the trait (Barboza et al., 2013) in these accessions. Another example of mutations in a GA 20-oxidase encoding gene has been reported for the accession Bur-0, which carries a *ga20ox4* loss of function allele. However, this does not result in semi-dwarfism (Plackett et al., 2012) or another phenotype. Mutations at different genes can lead to a reduction of plant height and thus inducing semi-dwarfism. In *A. thaliana* the *ga4* mutant, encoding the *GA3ox1* gene, shows a similar plant height phenotype as the *ga5* mutant and the total number of isolated induced mutants for this locus was significantly higher than for *ga5* (Koornneef & van der Veen, 1980). It is relevant to mention, that *GA3ox1* (*GA4*) is an ortholog of the pea *Le* gene studied by Mendel affecting stem elongation (Mendel, 1865; Lester et al., 1997). Another *A. thaliana* mutant causing semi-dwarfism is *GA INSENSITIVE* (*gai*) (Koornneef et al., 1985). The gene mutated in this semi-dominant mutant is an ortholog of the green revolution wheat genes *Reduced Height-1* (*Rht-B1* and *Rht-D1*) (Peng et al., 1999). The examples described above are semi-dwarf mutations in genes encoding for GA biosynthesis and signalling. However, also natural allelic variation at loci outside the GA pathway can lead to semi-dwarfism. For instance the *ERECTA* mutant (Torii et al., 1996), which decreases the plant height in different backgrounds, include some natural accessions (van Zanten et al., 2010). Induced mutants with a dwarf phenotype have been described many times and include mutants of the brassinosteroid (Li & Chory, 1999; Sakamoto et al., 2006; Clouse, 2011) and auxin pathways (Estelle & Somerville, 1987). Despite that allelic variants at many loci can affect plant height, semi-dwarfism

in natural populations of *A. thaliana* was shown to be caused mainly by *ga20ox1* (*ga5*) loss of function alleles, thus pointing to specific factors directing this gene as a hotspot for this trait. An indication of the frequency of these alleles in nature is the finding that four semi-dwarf accessions were present in the Hapmap, a population of 360 accessions, in which it was attempted to maximize the world wide *A. thaliana* diversity (Li et al., 2010).

Mutations in the *GA20ox1* (*GA5*) gene show no pleiotropic effects. Its mutants resemble their corresponding wild type accessions for many traits except that rosette size and plant height are reduced. Seed production is not affected in the *ga20ox1-3* mutant (Rieu et al., 2008). In the present study silique number was used as an indicator of fitness and it was found that semi-dwarfism did not affect this trait. Even during water limiting conditions, as shown in this study, the *ga20ox1-3* and *ga5* mutants show no trade-offs affecting performance when compared with their background accessions. In contrast *ga1* null mutants, encoding for a gene early in GA biosynthesis with severe dwarfism, show many pleiotropic effects that act as trade-offs, such as reduced germination and even leaky alleles will show reduced fertility and altered flower development (Koornneef & van der Veen, 1980). Recently natural variation was described for the *GA1* gene showing the presence of two distinct haplotypes (containing four non synonymous substitutions), which correlated with quantitative differences in floral morphology (Brock et al., 2012). However the effect of these mutations is small thus pointing that mutations with stronger effects at this locus will be eliminated by selection due to their trade-offs described above. Similar conclusions have been made in rice and barley (Spielmeyer et al., 2002; Jia et al., 2009; Jia et al., 2011).

GA5 is the functional ortholog of the rice green revolution semi-dwarf locus *Semi-Dwarf-1* (*SD1*), which codes for GA 20-oxidase-2 (Sasaki et al., 2002; Spielmeyer et al., 2002). Modern barley varieties carry mutations in this locus too, in which the mutated gene was called *Denso* or *Sdw1* (Jia et al., 2009). Studies have tested mutant *GAI* alleles in rice resulting in dwarf phenotypes and thus pointing to the possibility of using orthologs of this gene in crops to increase production (Peng et al., 1999; Fu et al., 2001). This suggests that *GAI* could have been selected in nature based on potential advantageous of its semi-dwarf phenotype. One possibility why this mutant has not been found yet in nature in diploid species, might be that a specific gain

of function dominant allele (Peng et al., 1997) is needed to confer this phenotype, which may be more rare than a loss of function mutation.

In this study semi-dwarfs were found together with wild type alleles indicating that semi-dwarf alleles are not fixed in nature. This observation suggests that, semi-dwarf alleles might be deleterious in some populations, especially when semi-dwarf plants are not good competitors in a mixed population. Mendel mentioned this trait-off for his short stem pea plants (Mendel, 1865). The same observation has been made for semi-dwarf *Populus* plants for which, a growth reduction was observed, when they were grown together with wild types in a competition study (Elias et al., 2012). Such observations raise the question what could be the selective advantage of semi-dwarfs in nature. It has been shown that in fragmented landscapes tall alleles will be more beneficial in terms of seed dispersal compared with short plants, represented by genotypes carrying the *erecta* allele in *A. thaliana* (Fakheran et al., 2010). Shorter plants might be beneficial under static environments (Fakheran et al., 2010), which resemble the urban environments (pedestrian side walks), where many semi-dwarfs were collected and which is common habitat of *A. thaliana* in anthropoid environments.

No obvious effect of *GA20ox1* mutations on drought tolerance as suggested by (Vartanian et al., 1994) was detected. Previous studies mentioned the hypothesis that *GA20ox2* induced semi-dwarfism in rice might confer drought tolerance because of a GA to Abscisic Acid (ABA) antagonism, where semi-dwarfs carrying low GA levels will thus show ABA accumulation (Lafitte et al., 2007). However it seems semi-dwarfism is not an advantage neither a trade-off for drought tolerance as semi-dwarfs with similar to wild types drought tolerance have been identified (Lafitte et al., 2007).

GA20ox1 (*GA5*) loss of function-induced semi-dwarfism may follow different evolutionary models based on the occurrence of different independent alleles, distributed worldwide, and present in natural and anthropoid environments. In a few populations different alleles were detected inside the same population, which might point to the presence and maintenance of these semi-dwarfs over a long time in these natural habitats in Spain. It can be expected that such populations, exposed to a variety of climatic conditions, have variable selective pressures in which some are beneficial for semi-dwarfs resulting in balancing selection. For two populations signatures of positive selection was observed, which could indicate the semi-dwarfism may become fixed due to a selective advantage. This was observed in two Asian populations.

However due to limited sampling it cannot be excluded that wild types are also present in such populations.

Cases of positive selection in specific environments have been identified in other genes, for instance, positive selection is present for the locus *SRF3* in *A. thaliana* Central Asian accessions and this might reflect an advantage under local pathogen pressure conditions (Alcázar et al., 2010). Signatures of purifying selection were also observed in this locus, which means deleterious mutations are purged and a functional *GA20ox1* (*GA5*) allele is maintained at least in the majority of the populations. This is in agreement with previous studies on the rice *GA20ox2* locus (Yang et al., 2009). Although it appears that natural *GA20ox1* (*GA5*) mutations are transiently maintained in nature, the occurrence of the same loss of function allele now in 11 locations in The Netherlands indicates that it spread across the country. Furthermore the finding of derived haplotypes of this allele provides further evidence that semi-dwarf alleles are maintained.

The occurrence of *GA20ox1* (*GA5*) loss of function alleles illustrates a case of convergent evolution. Recent literature studies have revealed "gene re-use" which means similar traits in different lineages have involved mutations in the same gene (Martin & Orgogozo, 2013). Regarding plant height the previous authors point to the known examples of rice and wheat GA biosynthesis and signalling mutants (Hedden, 2003; Salamini, 2003) as a hotspot in domesticated plants. A review by Lenser & Theißen (2013) collected many other examples of convergent evolution among crops. Convergent evolution can be seen as a synonym of gene re-use. Lenser & Theißen (2013) point that this phenomenon is frequently observed due to the human cultivation and breeding practices, which will usually have similar and specific phenotypic demands. In this thesis a case of convergent evolution is shown between selected semi-dwarfism in domesticated plants and the occurrence of semi-dwarfs in *A. thaliana* natural accessions. In agreement with Lenser & Theißen (2013), who mention that domestication helps the understanding of evolution, evolution might provide knowledge and understanding of domestication selected traits, which indicates the potential behind natural variation found in *A. thaliana* to find useful allelic variation for plant breeding. Genes exploited in domestication convergent evolution were favored due to (i) their nodal (central) position in the giving regulatory pathway, (ii) being in simple metabolic pathways, (iii) showing few pleiotropic effects, and (iv) presence at low frequency in wild populations which will facilitate its selection

(Lenser & Theißen, 2013). One of the main features of the *ga20ox1* (*ga5*) mutants is minimal pleiotropic effects, which most likely is the main factor allowing the occurrence of these variants in nature. In this thesis trade-offs were not observed during germination, root system length, flowering time, fitness and the performance under water limiting conditions. For some traits, the genetic background showed a strong effect. Mutants of the *GA20ox1* (*GA5*) gene are mainly affected in their stem elongation and rosette size, most likely due to redundancy with the remaining GA20ox paralogs, other traits are not affected (Rieu et al., 2008; Plackett et al., 2012). A similar conclusion was drawn for in rice, where mutations in early GA biosynthesis genes (such as *OsKO*) will be less favorable due to negative pleiotropic effects (Itoh et al., 2004). The review of Lenser & Theißen (2013) shows as an additional examples, the use of *FRIGIDA* (*FRI*) and mainly the *FLOWERING LOCUS C* (*FLC*) as domestication related flowering time locus in Brassicaceae (Okazaki et al., 2007; Yuan et al., 2009; Wang et al., 2011; Wu et al., 2012). Both *FRI* and *FLC* show a high number of independent mutations in natural *A. thaliana* accessions (Méndez-Vigo et al., 2011) thus resembling the case of *GA20ox1* (*GA5*) and pointing to the occurrence of further convergent evolution cases.

Summarizing, *ga20ox1* (*ga5*) loss of function alleles can be selected and maintained in nature due to different factors (Figure 6.1). These alleles might be under positive selection, providing a selective advantage in nature to an environmental factor still unknown and which may be transient, given the co-existence of wild type alleles in the best characterized populations. The occurrence of crops displaying mutations in the same locus illustrates a case of convergent evolution between domesticated crops and natural *A. thaliana* accessions. The minimal pleiotropic effects, due to redundancy with other GA 20-oxidases, shown in *ga20ox1* (*ga5*) loss of function alleles is a main factor contributing to the selection and maintenance in nature of these alleles. To what extend the gene expression levels of the other paralogs and the *GA20ox1* (*GA5*) gene itself change in semi-dwarfs compared to wild types requires further research together with a study of the functionality of the remaining *GA20ox* genes. Functional *GA5* alleles seem to be under purifying selection, which might indicatethat this gene must be functional due to relevance of having a tall phenotype under certain environments.

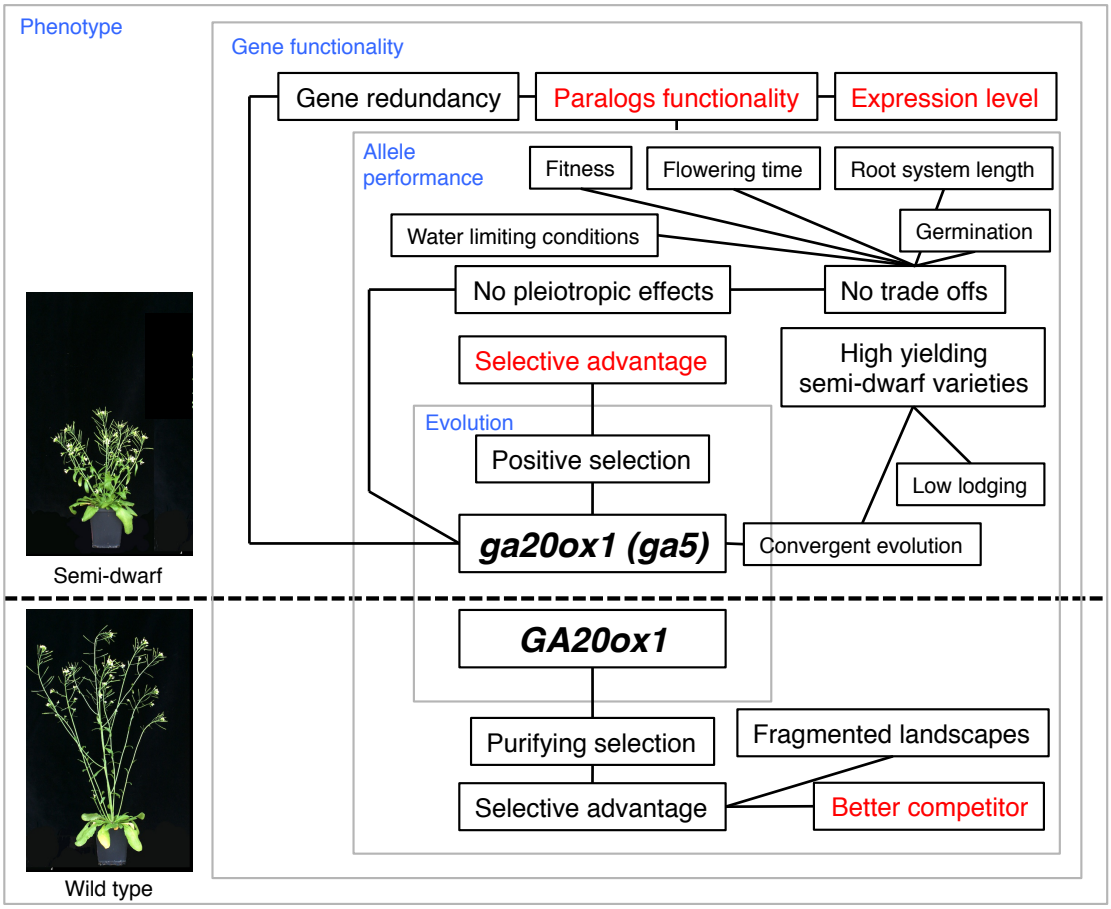

Figure 6.1. Factors affecting or related with the occurrence of loss of function *ga5* alleles vs. functional *GA5* alleles. Red colours indicate factors that need further experimental evidence.

Paclobutrazol sensitivity

The use of GA biosynthesis inhibitors such as paclobutrazol (PAC) have allowed the identification of different seed germination and seed dormancy related mutants. A dormant seed is defined as a viable seed that is unable to germinate even in a suitable environment (Holdsworth et al., 2008). For instance ABA biosynthesis (Léon-Kloosterziel et al., 1996; North et al., 2007), and signalling (Nambara et al., 1992) mutants have been isolated based on their PAC sensitivity. This selection is effective because GA is not needed when ABA and the inhibition of germination by ABA is not present. PAC hypersensitivity (germination inhibition at low ~0.1 µM doses) has been demonstrated in lines overexpressing the *ABA insensitive 5* (*ABI5*) transcription factor (Piskurewicz et al., 2008) and in *calcium sensor calcineurin B-like* (*cbl*) mutants (Li et al., 2013). In this thesis the PAC sensitivity was tested in *A. thaliana* natural accessions and substantial variation was found for the PAC effect on seed germination.

As described by (van der Schaar et al., 1997), genotypes that differ little in germination can display a significant difference in germination in PAC, controlled by different QTLs. Seeds must regulate their germination based on optimal environmental conditions, thus dormancy is present in nature. For this reason it is expected that also the regulation of ABA / GA levels in seed germination must be fine-tuned in nature and in crops. In agriculture seeds that germinate too fast might lead to a rapid loss of viability or undesired germination before harvesting (pre-harvest sprouting). A too strong dormancy is a problem as highly dormant seeds cannot be sown or are less suited for malting (e.g. in barley). Substantial dormancy variation has been found in *A. thaliana* accessions (Bentsink et al., 2010), and through natural variation relevant genes for this process have been identified and cloned such as *DELAY OF GERMINATION 1* (*DOG1*) (Bentsink et al., 2006). It is well known that ABA inhibits germination, thus by applying PAC the levels of this hormone are reduced and consequently germination is inhibited. Interestingly, although in *dog1* mutants ABA levels are slightly reduced and GA levels are increased, *dog1* mutants can not germinate without GA as shown in the double mutant *dog1 ga1-3* (Bentsink et al., 2006) and *dog1* shows PAC sensitivity (Nakabayashi et al., 2012) thus pointing to other mechanisms regulating dormancy than only the need to overcome the inhibition of germination by ABA. Due to the relatively small effect of the various loci and the fact that thus far no genes underlying the QTL could be identified, it was not elucidated what the physiological relevance is of the variation in PAC sensitivity is and if this is translated into other pleiotropic effects. PAC sensitivity is partially but not fully related with dormancy as was shown by the PAC effects on *dog1* mutants too (Nakabayashi et al., 2012). The finding that loci differ in their PAC and PAC + GA sensitivity and some loci controlling specific treatment effects, might imply novel mechanism involved with GA / ABA biosynthesis and signalling and its role in seed germination.

Hypocotyl length is a useful trait to study natural variation for GA biosynthesis and signalling because it provides an accessible experimental system and the trait is strongly affected by GAs. In this thesis a procedure to semi- / automatically quantify hypocotyls was developed. Hypocotyls grown in the dark were studied based on their known GA accumulation affecting cell elongation under this condition. (Gendreau et al., 1997; Lau & Deng, 2010). The method was proven to be successful especially when dealing with PAC as the procedure is very fast and could be adjusted to avoid the

PAC seed germination inhibition effects and thereby exclusively quantify the hypocotyl elongation inhibition. A locus at or near *erecta* was mapped as a main locus controlling hypocotyls elongation. *ERECTA* is a highly pleiotropic locus (van Zanten et al., 2009) and among the traits it regulates, is plant height (van Zanten et al., 2010). How *ERECTA*, which encodes for a receptor like kinase (Torii et al., 1996) affects plant length is unknown, but van Zanten et al., (2009) suggested the possibility of a GA mediated regulation that is independently from the classical GA-signalling way.

Perspectives about gibberellin biosynthesis and signalling natural variation studies

A. thaliana has five *GA20ox* paralogs showing the gene redundancy present in this GA biosynthesis step (Plackett et al., 2012). When the phylogeny of GA oxidases in Arabidopsis, rice and soybean was studied, the *GA20ox* genes group together with their respective homologs and most GA oxidase genes were present in paired paralogs (Han & Zhu, 2011). Transcriptional analysis showed that differences in gene expression occur when the *GA20ox1* gene is not functional (Rieu et al., 2008). The occurrence of several semi-dwarf accessions carrying loss of function *ga20ox1* alleles (Barboza et al., 2013) raises the question if gene expression of the remaining paralogs is changed and if signatures of selection are present in these genes. A hypothesis can be that accessions carrying inactive *GA20ox1* alleles have a functional *GA20ox2* and this is under selection pressure, together with a higher *GA20ox2* expression. This will provide further knowledge about the transcription regulation in this pathway. Up to date semi-dwarfs allelic to *ga20ox1* (*ga5*) have been identified in 23 populations (Barboza et al., 2013). Currently additional new semi-dwarf accessions from other populations are being analysed. The observation that not all accessions are allelic to *ga5*, immediately points to other loci / genes controlling this phenotype and will provide insight in what other allelic variation can control plant height in nature without negative pleiotropic effects.

Further germination / genetic analyses are necessary to understand the meaning of PAC sensitivity in seed germination and hypocotyl elongation. Fine mapping of the locus on chr 1, together with the cloning of the putative gene controlling this effect are remaining tasks. The availability of several *A. thaliana* bi-parental and worldwide collections of accessions will provide relevant genetic material to study the genetic control of hypocotyl elongation. All this might contribute to the understanding of the

mode of action and regulation of GAs and might help answering open questions such as GA mobility / transport in plants, novel genes and signals controlling GA biosynthesis and signalling (Yamaguchi, 2008). Besides, the occurrence of interesting traits, which might be co-selected with semi-dwarfism such as the long root system depth of the Pak-3 accession, are experiments, that will require genetics to map and clone the loci controlling it.

Summary

Variations in plant hormone levels and hormone signalling might be a basis for the phenotypic variation in developmental and stress-related traits within the same species found in nature. Understanding the genetic basis for such a variation was the major objective of this thesis. As an example gene, *GA5* was investigated. *GA5* encodes a GA 20-oxidase, an enzyme involved in the last steps of bioactive gibberellin (GA) biosynthesis. GAs are plant growth regulators involved in different traits such as seed germination, flowering and plant height. To identify mutations in the *GA20ox1 (GA5)* gene, *Arabidopsis thaliana* natural semi-dwarf accessions and the *ga5* mutant were investigated by allelism tests as well as by DNA sequencing. This approach led to the identification of a large number of independent mutations that were found to be responsible for inactive alleles in 17 different populations from various parts of the world. Population genetics was performed for the *GA20ox1* gene in the world-wide collection of natural accessions and in specific populations with many semi-dwarfs that indicate local selection. No obvious trade-offs affecting plant performance were observed among the semi-dwarf accessions. To test if the semi-dwarf GA20ox1 mutants can withstand water stress better, physiological experiments were performed under water withholding conditions. Although no significant effects could be assigned to this mutation, one semi-dwarf genotype was identified with a much longer root system that seems to correlate with tolerance to water withholding. To further screen the natural variation of GA biosynthesis and GA signalling, *Quantitative Trait Loci* (QTL) analysis and *Genome Wide Association Studies* (GWAS) were performed. Using both strategies, the influence of the gibberellin biosynthesis inhibitor paclobutrazol (PAC) and of GA_{4+7} on seed germination was tested. The QTL and GWAS studies showed a complex regulation of the effects of GA depletion / restoration. A main locus for PAC sensitivity on chromosome 1 was validated and characterized for its effect on seed germination. To extend the study, a semi-automatic method named 'HyPer' was developed for the quantification of skotomorphogenic hypocotyls. This method was used to perform QTL mapping for hypocotyl length and the effects of application of PAC and GA on this trait. These experiments revealed *ERECTA* as a locus involved in PAC sensitivity. Based on these findings, GAs play a role in *A. thaliana* natural variation. The finding of *GA5* as main GA hotspot in this study illustrates a case of convergent evolution between natural semi-dwarfs and wild domesticated plant species.

Zusammenfassung

Änderungen im Pflanzenhormongehalt und in der Hormonsignalisierung könnten die Grundlage für die phänotypische Variationen in entwicklungs- und stress-bedingten Merkmalen innerhalb der gleichen, in der Natur vorkommenden, Species darstellen.
Das Hauptziel dieser Arbeit war es, das Verständnis der genetischen Grundlagen solcher Variationen zu erweiteren. Mithilfe von Komplementationstests und DNA-Squenzierungen wurden in der Natur vorkommenden *Arabidopsis thaliana* Halbzwerg-Ökotypen mit der *ga5* Mutante verglichen. Wir konnten damit eine Vielzahl von unabhängig auftretenden Mutationen in dem *GA20ox1* (GA5)-Gen nachweisen. Diese Mutationen führen zu inaktiven Allelen in 17 verschiedenen Populationen aus verschiedenen Teilen der Welt. *GA5* kodiert für eine GA 20-oxidase, ein Enzym, das an den letzten Schritten der Gibberellin (GA)-Biosynthese beteiligt ist. Gibberelline sind Pflanzenwachstumsregulatoren, die verschiedene Merkmale wie Samenkeimung, Blühzeitpunkt und Pflanzenhöhe beeinflussen. Eine Analyse der Populationsgenetik für das *GA20ox1*-Gen wurde in der weltweiten Kollektion von natürlichen Ökotypen und in speziellen Populationen mit vielen Halbzwergen durchgeführt. Diese Untersuchungen weisen auf eine örtliche Selektion hin. Bei den GA20ox1 Halbzwerg-Mutanten wurden keine offensichtlichen Nachteile der Pflanzen bezüglich ihrer Leistungs- und Überlebensfähigkeit festgestellt. Um zu testen, ob die Mutanten Wasserstress besser vertragen, wurden physiologische Experimente unter Wasserrückhaltebedingungen durchgeführt. Obwohl den Mutationen keine eindeutigen Effekte zugewiesen werden konnten, wurde eine Halbzwerg-Variante mit einem viel größeren Wurzelsystem identifiziert, der mit einer Wasserrückhaltetoleranz zu korrelieren scheint. Um das Auftreten der natürlichen Variation hinsichtlich der GA-Biosynthese und des GA-Signalwegs zu untersuchen, wurde eine QTL (engl. Quantitative Trait Loci)-Analyse und eine genomweite Assoziationsstudie (GWAS, engl. Genome-wide association study) durchgeführt , um die Empfindlichkeit der Samenkeimung gegenüber dem Gibberellin-Biosynthese-Inhibitor Paclobutrazol (PAC) und gegenüber GA_{4+7} zu testen. Die QTL- und die GWAS-Analyse zeigten eine komplexe Regulierung der Effekte des Abbaus und Wiederaufbaus von GA. Ein Hauptlocus für die Empfindlichkeit gegenüber PAC wurde bestätigt und hinsichtlich seines Einfluses auf die Samenkeimung charakterisiert. Um unsere Studie auszudehnen, wurde eine `HyPer´genannte, halb-automatisierte Methode zur

Quantifizierung des sktotomorphogenesen Hypokotyls entwickelt,. Diese Methode wurde für eine QTL-Kartierung der Hypokotyllänge und der Effekte von PAC und GA genutzt. Diese Experimente zeigten *ERECTA* als einen Locus, der mit der Empfindlichkeit gegenüber PAC zu tun hat. Anhand dieser Ergebnisse spielen Giberelline eine Rolle in der natürlichen Variation von *Arabidopsis thaliana*. Der Fund von *GA5* als GA-"hotspot" in dieser Arbeit zeigt, dass eine konvergente Entwicklung zwischen natürlich vorkommenden Halbzwergen und wilden, domestizierten Pflanzenspecies vorliegt.

References

Achard P, Genschik P (2008) Releasing the brakes of plant growth: how GAs shutdown DELLA proteins. *J Exp Bot* 60:1085–1092.

Alcázar R, García AV, Kronholm I, de Meaux J, Koornneef M, Parker JE, Reymond M (2010) Natural variation at strubbelig receptor kinase 3 drives immune-triggered incompatibilities between *Arabidopsis thaliana* accessions. *Nat Genet* 42:1135–1139.

Allahverdiyeva Y, Suorsa M, Rossi F, Pavesi A, Kater MM, Antonacci A, Tadini L, Pribil M, Schneider A, Wanner G, Leister D, Aro EM, Barbato R, Pesaresi P (2013) Arabidopsis plants lacking PsbQ and PsbR subunits of the oxygen-evolving complex show altered PSII super-complex organization and short-term adaptive mechanisms. *Plant J Cell Mol Biol* 75:671–684.

Alonso-Blanco C, Bentsink L, Hanhart CJ, Blankestijn-de Vries H, Koornneef M (2003) Analysis of natural allelic variation at seed dormancy loci of *Arabidopsis thaliana*. *Genetics* 164:711–729.

Alonso-Blanco C, Gomez-Mena C, Llorente F, Koornneef M, Salinas J, Martínez-Zapater JM (2005) Genetic and molecular analyses of natural variation indicate *CBF2* as a candidate gene for underlying a freezing tolerance quantitative trait locus in Arabidopsis. *Plant Physiol* 139:1304–1312.

Alonso-Blanco C, Aarts MG, Bentsink L, Keurentjes JJ, Reymond M, Vreugdenhil D, Koornneef M (2009) What has natural variation taught us about plant development, physiology, and adaptation? *Plant Cell* 21:1877–1896.

Anderson JT, Lee CR, Rushworth CA, Colautti RI, Mitchell-Olds T (2013) Genetic trade-offs and conditional neutrality contribute to local adaptation. *Mol Ecol* 22:699–708.

Arends D, Prins P, Jansen RC, Broman KW (2010) R / qtl: high-throughput multiple QTL mapping. *Bioinformatics* 26:2990–2992.

Baecker V, Travo P (2006) Cell image analyzer - a visual scripting interface for ImageJ and its usage at the microscopy facility Montpellier RIO Imaging. *In Proc ImageJ User Dev Conf* :105–110.

Barboza L, Effgen S, Alonso-Blanco C, Kooke R, Keurentjes JJ, Koornneef M, Alcázar R (2013) Arabidopsis semidwarfs evolved from independent mutations in *GA20ox1*, ortholog to green revolution dwarf alleles in rice and barley. *Proc Natl Acad Sci* 110:15818–15823.

Baxter I, Brazelton JN, Yu D, Huang YS, Lahner B, Yakubova E, Li Y, Bergelson J, Borevitz JO, Nordborg M, Vitek O, Salt DE (2010) A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter AtHKT1;1. *PLoS Genet* 6:e1001193.

Bentsink L, Jowett J, Hanhart CJ, Koornneef M (2006) Cloning of *DOG1*, a quantitative trait locus controlling seed dormancy in Arabidopsis. *Proc Natl Acad Sci* 103:17042–17047.

Bentsink L, Hanson J, Hanhart CJ, Blankestijn-de Vries H, Coltrane C, Keizer P, El-Lithy M, Alonso-Blanco C, de Andrés MT, Reymond M, van Eeuwijk F, Smeekens S, Koornneef M (2010) Natural variation for seed dormancy in Arabidopsis is regulated by additive genetic and molecular pathways. *Proc Natl Acad Sci* 107:4264–4269.

Bergelson J, Roux F (2010) Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nat Rev Genet* 11:867–879.

Bidadi H, Yamaguchi S, Asahina M, Satoh S (2010) Effects of shoot-applied gibberellin / gibberellin-biosynthesis inhibitors on root growth and expression of gibberellin biosynthesis genes in *Arabidopsis thaliana*. *Plant Root* 4:4–11.

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

Botto JF, Alonso-Blanco C, Garzarón I, Sánchez RA, Casal JJ (2003) The Cape Verde Islands allele of cryptochrome 2 enhances cotyledon unfolding in the absence of blue light in Arabidopsis. *Plant Physiol* 133:1547–1556.

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F (2010) Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genet* 6:e1000940.

Brock MT, Kover PX, Weinig C (2012) Natural variation in GA1 associates with floral morphology in *Arabidopsis thaliana*. *New Phytol.* 195:58–70.

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

Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Müller J, Alonso-Blanco C, Borgwardt K, Schmid KJ, Weigel D (2011) Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nat Genet* 43:956–963.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9:1657–1659.

Clerkx EJM, El-Lithy ME, Vierling E, Ruys GJ, Blankestijn-De Vries H, Groot SPC, Vreugdenhil D, Koornneef M (2004) Analysis of natural allelic variation of Arabidopsis seed germination and seed longevity traits between the accessions Landsberg *erecta* and Shakdara, using a new recombinant inbred line population. *Plant Physiol* 135:432–443.

Clouse SD (2011) Brassinosteroid signal transduction: from receptor kinase activation to transcriptional networks regulating plant development. *Plant Cell* 23:1219–1230.

Cole B, Kay SA, Chory J (2011) Automated analysis of hypocotyl growth dynamics during shade avoidance in Arabidopsis. *Plant J* 65:991–1000.

Coolbaugh RC, Hamilton R (1976) Inhibition of ent-Kaurene oxidation and growth by -Cyclopropyl- -(p-methoxyphenyl)-5-pyrimidine methyl alcohol. *Plant Physiol* 57:245–248.

Curaba J, Moritz T, Blervaque R, Parcy F, Raz V, Herzog M, Vachon G (2004) *AtGA3ox2*, a key gene responsible for bioactive gibberellin biosynthesis, is regulated during embryogenesis by LEAFY COTYLEDON2 and FUSCA3 in Arabidopsis. *Plant Physiol* 136:3660–3669.

Debeaujon I, Koornneef M (2000) Gibberellin requirement for Arabidopsis seed germination is determined both by testa characteristics and embryonic abscisic acid. *Plant Physiol* 122:415–424.

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

Elias AA, Busov VB, Kosola KR, Ma C, Etherington E, Shevchenko O, Gandhi H, Pearce DW, Rood SB, Strauss SH (2012) Green revolution trees: semidwarfism transgenes modify gibberellins, promote root growth, enhance morphological diversity, and reduce competitiveness in hybrid Poplar. *Plant Physiol* 160:1130–1144.

Estelle MA, Somerville C (1987) Auxin-resistant mutants of *Arabidopsis thaliana* with an altered morphology. *Mol Gen Genet* 206:200–206.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620.

Fakheran S, Paul-Victor C, Heichinger C, Schmid B, Grossniklaus U, Turnbull LA (2010) Adaptation and extinction in experimentally fragmented landscapes. *Proc Natl Acad Sci USA* 107:19120–19125.

Fitz JN, Lehti-Shiu MD, Ingram PA, Deak KI, Biesiada T, Malamy JE (2005) Identification of Quantitative Trait Loci that regulate Arabidopsis root system size and plasticity. *Genetics* 172:485–498.

Fu X, Sudhakar D, Peng J, Richards DE, Christou P, Harberd NP (2001) Expression of Arabidopsis GAI in transgenic rice represses multiple gibberellin responses. *Plant Cell* 13:1791–1802.

Galpaz N, Reymond M (2010) Natural variation in *Arabidopsis thaliana* revealed a genetic network controlling germination under salt stress. *PLoS One* 5:e15198.

Gendreau E, Traas J, Desnos T, Grandjean O, Caboche M, Höfte H (1997) Cellular basis of hypocotyl growth in *Arabidopsis thaliana*. *Plant Physiol* 114:295–305.

Ghanem ME, Hichri I, Smigocki AC, Albacete A, Fauconnier ML, Diatloff E, Martinez-Andujar C, Lutts S, Dodd IC, Pérez-Alfocea F (2011) Root-targeted biotechnology to mediate hormonal signalling and improve crop stress tolerance. *Plant Cell Rep* 30:807–823.

Gou J, Strauss SH, Tsai CJ, Fang K, Chen Y, Jiang X, Busov VB (2010) Gibberellins regulate lateral root formation in *Populus* through interactions with auxin and other hormones. *Plant Cell* 22:623–639.

Granier C, Aguirrezabal L, Chenu K, Cookson SJ, Dauzat M, Hamard P, Thioux JJ, Rolland G, Bouchier-Combaud S, Lebaudy A, Muller B, Simonneau T, Tardieu F (2006) PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in *Arabidopsis thaliana* permitted the identification of an accession with low sensitivity to soil water deficit. *New Phytol* 169:623–635.

Gujas B, Alonso-Blanco C, Hardtke CS (2012) Natural Arabidopsis *brx* loss-of-function alleles confer root adaptation to acidic soil. *Curr Biol* 22:1962–1968.

Han F, Zhu B (2011) Evolutionary analysis of three gibberellin oxidase genes in rice, Arabidopsis, and soybean. Gene 473:23–35.

Harberd NP, Freeling M (1989) Genetics of dominant gibberellin-insensitive dwarfism in maize. Genetics 121:827–838.

Hedden P, Graebe JE (1985) Inhibition of gibberellin biosynthesis by paclobutrazol in cell-free homogenates of *Cucurbita maxima* endosperm and *Malus pumila* embryos. *J Plant Growth Regul* 4:111–122.

Hedden P (2003) The genes of the Green Revolution. *Trends Genet* 19:5–9.

Hedden P, Thomas SG (2012) Gibberellin biosynthesis and its regulation. *Biochem J* 444:11–25.

Holdsworth MJ, Bentsink L, Soppe WJJ (2008) Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytol* 179:33–54.

Huang X, Paulo MJ, Boer M, Effgen S, Keizer P, Koornneef M, van Eeuwijk FA (2011) Analysis of natural allelic variation in Arabidopsis using a multiparent recombinant inbred line population. *Proc Natl Acad Sci* 108:4488–4493.

Huang X, Ding J, Effgen S, Turck F, Koornneef M (2013) Multiple loci and genetic interactions involving flowering time genes regulate stem branching among natural variants of Arabidopsis. *New Phytol* 199:843–857.

Ingvarsson PK, Street NR (2011) Association genetics of complex traits in plants. *New Phytol* 189:909–922.

Itoh H, Tatsumi T, Sakamoto T, Otomo K, Toyomasu T, Kitano H, Ashikari M, Ichihara S, Matsuoka M (2004) A rice semi-dwarf gene, Tan-Ginbozu (D35), encodes the gibberellin biosynthesis enzyme, ent-kaurene oxidase. *Plant Mol Biol* 54:533–547.

Jacobsen SE, Olszewski NE (1993) Mutations at the *SPINDLY* locus of Arabidopsis alter gibberellin signal transduction. *Plant Cell* 5:887–896.

Jansen M, Gilmer F, Biskup B, Nagel KA, Rascher U, Fischbach A, Briem S, Dreissen G, Tittmann S, Braun S, De Jaeger I, Metzlaff M, Schurr U, Scharr H, Walter A (2009) Simultaneous phenotyping of leaf growth and chlorophyll fluorescence via GROWSCREEN FLUORO allows detection of stress tolerance in *Arabidopsis thaliana* and other rosette plants. *Funct Plant Biol* 36:902-914.

Jia Q, Zhang J, Westcott S, Zhang XQ, Bellgard M, Lance R, Li C (2009) GA-20 oxidase as a candidate for the semidwarf gene *sdw1/denso* in barley. *Funct Integr Genomics* 9:255–262.

Jia Q, Zhang XQ, Westcott S, Broughton S, Cakir M, Yang J, Lance R, Li C (2011) Expression level of a gibberellin 20-oxidase gene is associated with multiple agronomic and quality traits in barley. *Theor Appl Genet* 122:1451–1460.

Joosen RVL, Kodde J, Willems LA, Ligterink W, van der Plas LH, Hilhorst HW (2010) GERMINATOR: a software package for high-throughput scoring and curve fitting of Arabidopsis seed germination. *Plant J Cell Mol Biol* 62:148–159.

Joosen RVL, Arends D, Willems LA, Ligterink W, Jansen RC, Hilhorst HW (2011) Visualizing the genetic landscape of Arabidopsis seed performance. *Plant Physiol* 158:570–589.

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42:348-354.

Kellermeier F, Chardon F, Amtmann A (2013) Natural variation of Arabidopsis root architecture reveals complementing adaptive strategies to potassium starvation. *Plant Physiol* 161:1421–1432.

Keurentjes JJB, Bentsink L, Alonso-Blanco C, Hanhart CJ, Blankestijn-De Vries H, Effgen S, Vreugdenhil D, Koornneef M (2007) Development of a near-isogenic line population of Arabidopsis thaliana and comparison of mapping power with a recombinant inbred line population. *Genetics*, 175:891–905.

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

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

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

Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9:29.

Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, Durrant C, Mott R (2009) A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet* 5:e1000551.

Kover PX, Mott R (2012) Mapping the genetic basis of ecologically and evolutionary relevant traits in *Arabidopsis thaliana*. *Curr Opin Plant Biol* 15:212–217.

Lafitte HR, Yongsheng G, Yan S, Li ZK (2007) Whole plant responses, key processes, and adaptation to drought stress: the case of rice. *J Exp Bot* 58:169–175.

Lau OS, Deng XW (2010) Plant hormone signalling lightens up: integrators of light and hormones. *Curr Opin Plant Biol* 13:571–577.

Lê S, Josse J, Husson F (2008) FactoMineR: An R package for multivariate analysis. *J Stat Softw* 25:1–18.

Lenser T, Theißen G (2013) Molecular mechanisms involved in convergent crop domestication. *Trends Plant Sci*. doi: 10.1016/j.tplants.2013.08.007.

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

Lester DR, Ross JJ, Davies PJ, Reid JB (1997) Mendel's stem length gene (*Le*) encodes a gibberellin 3 beta-hydroxylase. *Plant Cell* 9:1435–1443.

Lewandowska-Sabat AM, Fjellheim S, Rognli OA (2010) Extremely low genetic variability and highly structured local populations of *Arabidopsis thaliana* at higher latitudes. *Mol Ecol* 19:4753–4764.

Li J, Chory J (1999) Brassinosteroid actions in plants. *J Exp Bot* 50:275–282.

Li Y, Huang Y, Bergelson J, Nordborg M, Borevitz JO (2010) Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 107:21199–21204.

Li ZY, Xu ZS, Chen Y, He GY, Yang GX, Chen M, Li LC, Ma YZ (2013) A novel role for Arabidopsis CBL1 in affecting plant responses to glucose and gibberellin during germination and seedling development. *PLoS One* 8:e56412.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452.

Loudet O, Chaillou S, Camilleri C, Bouchez D, Vedele D (2002) Bay-0 x Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in Arabidopsis. *Theor Appl Genet* 104:1173–1184.

Loudet O, Gaudon V, Trubuil A, Daniel-Vedele F (2005) Quantitative trait loci controlling root growth and architecture in *Arabidopsis thaliana* confirmed by heterogeneous inbred family. *Theor Appl Genet* 110:742–753.

Lovell JT, Juenger TE, Michaels SD, Lasky JR, Platt A, Richards JH, Yu X, Easlon HM, Sen S, McKay JK (2013) Pleiotropy of FRIGIDA enhances the potential for multivariate adaptation. *Proc R Soc B Biol Sci* 280:1043–20131043.

Lynch JP (2013) Steep, cheap and deep: an ideotype to optimize water and N acquisition by maize root systems. *Ann Bot* 112:347–357.

MacNicol PK, Jacobsen JV (2001) Regulation of alcohol dehydrogenase gene expression in barley aleurone by gibberellin and abscisic acid. *Physiol Plant* 111:533–539.

Martin A, Orgogozo V (2013) The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* 67:1235-1250.

Masle J, Gilmore SR, Farquhar GD (2005) The *ERECTA* gene regulates plant transpiration efficiency in Arabidopsis. *Nature* 436:866–870.

Meijón M, Satbhai SB, Tsuchimatsu T, Busch W (2013) Genome-wide association study using cellular traits identifies a new regulator of root development in Arabidopsis. *Nat Genet.* doi:10.1038/ng.2824

Mendel G (1865) Versuche über Plflanzenhybriden. Verh Natforsch Ver Brünn 4:3–47. English translation available at: http://www.mendelweb.org/Mendel.html

Méndez-Vigo B, Picó FX, Ramiro M, Martínez-Zapater JM, Alonso-Blanco C (2011) Altitudinal and climatic adaptation is mediated by flowering traits and *FRI*, *FLC*, and *PHYC* genes in Arabidopsis. *Plant Physiol* 157:1942–1955.

Méndez-Vigo B, Martínez-Zapater JM, Alonso-Blanco C (2013) The flowering repressor SVP underlies a novel *Arabidopsis thaliana* QTL interacting with the genetic background. *PLoS Genet* 9:e1003289.

Miller ND, Parks BM, Spalding EP (2007) Computer-vision analysis of seedling responses to light and gravity. *Plant J* 52:374–381.

Mitchum MG, Yamaguchi S, Hanada A, Kuwahara A, Yoshioka Y, Kato T, Tabata S, Kamiya Y, Sun TP (2006) Distinct and overlapping roles of two gibberellin 3-oxidases in Arabidopsis development. *Plant J* 45:804–818.

Moore CR, Gronwall DS, Miller ND, Spalding EP (2013) Mapping quantitative trait loci affecting *Arabidopsis thaliana* seed morphology features extracted computationally from images. *G3* 3:109–118.

Mouchel CF, Briggs GC, Hardtke CS (2004) Natural genetic variation in Arabidopsis identifies BREVIS RADIX, a novel regulator of cell proliferation and elongation in the root. *Genes Dev* 18:700–714.

Mouchel CF, Osmont KS, Hardtke CS (2006) BRX mediates feedback between brassinosteroid levels and auxin signalling in root growth. *Nature* 443:458–461.

Nagel KA, Kastenholz B, Jahnke S, van Dusschoten D, Aach T, Mühlich M, Truhn D, Scharr H, Terjung S, Walter A, Schurr U (2009) Temperature responses of roots: impact on growth, root system architecture and implications for phenotyping. *Funct Plant Biol* 36:947-959.

Nagel KA, Putz A, Gilmer F, Heinz K, Fischbach A, Pfeifer J, Faget M, Blossfeld S, Ernst M, Dimaki C, Kastenholz B, Kleinert A, Galinski A, Scharr H, Fiorani F, Schurr U (2012) GROWSCREEN-Rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons. *Funct Plant Biol* 39:891-904.

Nakabayashi K, Bartsch M, Xiang Y, Miatton E, Pellengahr S, Yano R, Seo M, Soppe WJ (2012) The time required for dormancy release in Arabidopsis is determined by DELAY OF GERMINATION1 protein levels in freshly harvested seeds. *Plant Cell* 24:2826–2838.

Nambara E, Naito S, 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.

Nielsen R (2005) Molecular signatures of natural selection. *Annu Rev Genet* 39:197–218.

North HM, De Almeida A, Boutin JP, Frey A, To A, Botran L, Sotta B, Marion-Poll A (2007) The Arabidopsis ABA-deficient mutant *aba4* demonstrates that the major route for stress-induced ABA accumulation is via neoxanthin isomers. *Plant J* 50:810–824.

Okazaki K, Sakamoto K, Kikuchi R, Saito A, Togashi E, Kuginuki Y, Matsumoto S, Hirai M (2007) Mapping and characterization of FLC homologs and QTL analysis of flowering time in *Brassica oleracea*. *Theor. Appl. Genet* 114:595–608.

Overvoorde P, Fukaki H, Beeckman T (2010) Auxin control of root development. *Cold Spring Harb Perspect Biol* 2:a001537–a001537.

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

Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, Sudhakar D, Christou P, Snape JW, Gale MD, Harberd NP (1999) "Green revolution" genes encode mutant gibberellin response modulators. *Nature* 400:256–261.

Picó FX, Méndez-Vigo B, Martínez-Zapater JM, Alonso-Blanco C (2008) Natural genetic variation of *Arabidopsis thaliana* is geographically structured in the Iberian penninsula. *Genetics* 180:1009–1021.

Piskurewicz U, Jikumaru Y, Kinoshita N, Nambara E, Kamiya Y, Lopez-Molina L (2008) The gibberellic acid signalling repressor RGL2 inhibits Arabidopsis seed germination by stimulating abscisic acid synthesis and ABI5 activity. *Plant Cell* 20:2729–2745.

Plackett AR, Powers SJ, Fernandez-Garcia N, Urbanova T, Takebayashi Y, Seo M, Jikumaru Y, Benlloch R, Nilsson O, Ruiz-Rivero O, Phillips AL, Wilson ZA, Thomas SG, Hedden P (2012) Analysis of the developmental roles of the Arabidopsis gibberellin 20-oxidases demonstrates that *GA20ox1*, *-2*, and *-3* are the dominant paralogs. *Plant Cell* 24:941–960.

Platt A, Horton M, Huang YS, Li Y, Anastasio AE, Mulyati NW, Agren J, Bossdorf O, Byers D, Donohue K, Dunning M, Holub EB, Hudson A, Le Corre V, Loudet O, Roux F, Warthmann N, Weigel D, Rivero L, Scholl R, Nordborg M, Bergelson J, Borevitz JO (2010) The scale of population structure in *Arabidopsis thaliana*. *Plos Genet* 6:e1000843.

Prinzenberg AE, Barbier H, Salt DE, Stich B, Reymond M (2010) Relationships between growth, growth response to nutrient supply, and ion content using a recombinant inbred line population in Arabidopsis. *Plant Physiol* 154:1361–1371.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.

R Core Team (2013) R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Ravi M, Chan SWL (2010) Haploid plants produced by centromere-mediated genome elimination. *Nature* 464:615–618.

Rieu I, Ruiz-Rivero O, Fernandez-Garcia N, Griffiths J, Powers SJ, Gong F, Linhartova T, Eriksson S, Nilsson O, Thomas SG, Phillips AL, Hedden P (2008) The gibberellin biosynthetic genes *AtGA20ox1* and *AtGA20ox2* act, partially redundantly, to promote growth and development throughout the Arabidopsis life cycle. *Plant J* 53:488–504.

Rosas U, Cibrian-Jaramillo A, Ristova D, Banta JA, Gifford ML, Fan AH, Zhou RW, Kim GJ, Krouk G, Birnbaum KD, Purugganan MD, Coruzzi GM (2013) Integration of responses within and across Arabidopsis natural accessions uncovers loci controlling root systems architecture. *Proc Natl Acad Sci* 110:15133–15138.

Sakamoto T, Morinaka Y, Ohnishi T, Sunohara H, Fujioka S, Ueguchi-Tanaka M, Mizutani M, Sakata K, Takatsuto S, Yoshida S, Tanaka H, Kitano H, Matsuoka M (2006) Erect leaves caused by brassinosteroid deficiency increase biomass production and grain yield in rice. *Nat Biotechnol* 24:105–109.

Salamini F (2003) Hormones and the green revolution. *Science* 302:71–72.

Sangster TA, Salathia N, Undurraga S, Milo R, Schellenberg K, Lindquist S, Queitsch C (2008) HSP90 affects the expression of genetic variation and developmental stability in quantitative traits. *Proc Natl Acad Sci* 105:2963–2968.

Santner A, Calderon-Villalobos LIA, Estelle M (2009) Plant hormones are versatile chemical regulators of plant growth. *Nat Chem Biol* 5:301–307.

Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D, Ishiyama K, Saito T, Kobayashi M, Khush GS, Kitano H, Matsuoka M (2002) Green revolution: a mutant gibberellin-synthesis gene in rice. *Nature* 416:701–702.

Scacchi E, Salinas P, Gujas B, Santuari L, Krogan N, Ragni L, Berleth T, Hardtke CS (2010) Spatio-temporal sequence of cross-regulatory events in root meristem growth. *Proc Natl Acad Sci* 107:22734–22739.

Schmid KJ, Ramos-Onsins S, Ringys-Beckstein H, Weisshaar B, Mitchell-Olds T (2005) A multilocus sequence survey in *Arabidopsis thaliana* reveals a genome-wide departure from a neutral model of DNA sequence polymorphism. *Genetics* 169:1601–1615.

Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9:671–675.

Scholl RL, Amos JA (1980) Isolation of doubled-haploid plants through anther culture in *Arabidopsis thaliana*. *Z für Pflanzenphysiol* 96:407–414.

Sergeeva LI, Keurentjes JJ, Bentsink L, Vonk J, van der Plas LH, Koornneef M, Vreugdenhil D (2006) Vacuolar invertase regulates elongation of Arabidopsis thaliana roots as revealed by QTL and mutant analysis. *Proc Natl Acad Sci* 103:2994–2999.

Seymour DK, Filiault DL, Henry IM, Monson-Miller J, Ravi M, Pang A, Comai L, Chan SW, Maloof JN (2012) Rapid creation of Arabidopsis doubled

haploid lines for quantitative trait locus mapping. *Proc Natl Acad Sci* 109:4227–4232.

Shin J-H, Blay S, McNeney B, Graham J (2006) LDheatmap: An R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J Stat Soft* 16.

Simon M, Loudet O, Durand S, Bérard A, Brunel D, Sennesal FX, Durand-Tardif M, Pelletier G, Camilleri C (2008) Quantitative trait loci mapping in five new large recombinant inbred line populations of *Arabidopsis thaliana* genotyped with consensus single-nucleotide polymorphism markers. *Genetics* 178:2253–2264.

Spielmeyer W, Ellis MH, Chandler PM (2002) Semidwarf (*sd-1*), "green revolution" rice, contains a defective gibberellin 20-oxidase gene. *Proc Natl Acad Sci USA* 99:9043–9048.

Strabala TJ, MacMillan CP (2013) The Arabidopsis wood model–the case for the inflorescence stem. *Plant Science* 210:193–205.

Sun T (2008) Gibberellin metabolism, perception and signalling pathways in Arabidopsis. *Arab Book* 6: e0103.

Sun T (2010) Gibberellin-GID1-DELLA: a pivotal regulatory module for plant growth and development. *Plant Physiol* 154:567–570.

Sun T (2011) The molecular mechanism and evolution of the GA–GID1–DELLA signalling module in plants. *Curr Biol* 21:338–345.

Swamy BPM, Ahmed HU, Henry A, Mauleon R, Dixit S, Vikram P, Tilatto R, Verulkar SB, Perraju P, Mandal NP, Variar M, Robin S, Chandrababu R, Singh ON, Dwivedi JL, Das SP, Mishra KK, Yadaw RB, Aditya TL, Karmakar B, Satoh K, Moumeni A, Kikuchi S, Leung H, Kumar A (2013) Genetic, physiological, and gene expression analyses reveal that multiple QTL enhance yield of rice mega-variety IR64 under drought. *PLoS One* 8:e62795.

Tanimoto E (2012) Tall or short? Slender or thick? A plant strategy for regulating elongation growth of roots by low concentrations of gibberellin. *Ann Bot* 110:373–381.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.

Torii KU, Mitsukawa N, Oosumi T, Matsuura Y, Yokoyama R, Whittier RF, Komeda Y (1996) The Arabidopsis ERECTA gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *Plant Cell* 8:735–746.

Tuinstra MR, Ejeta G, Goldsbrough PB (1997) Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. *Theor Appl Genet* 95:1005–1011.

Uga Y, Sugimoto K, Ogawa S, Rane J, Ishitani M, Hara N, Kitomi Y, Inukai Y, Ono K, Kanno N, Inoue H, Takehisa H, Motoyama R, Nagamura Y, Wu J, Matsumoto T, Takai T, Okuno K, Yano M (2013) Control of root

system architecture by DEEPER ROOTING 1 increases rice yield under drought conditions. *Nat Genet* 45:1097–1102.

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

van der Veen JH, Wirtz P (1968) EMS-induced genic male sterility in *Arabidopsis thaliana*: A model selection experiment. *Euphytica* 17:371–377.

van Zanten M, Snoek LB, Proveniers MCG, Peeters AJM (2009) The many functions of ERECTA. *Trends Plant Sci* 14:214–218.

van Zanten M, Basten Snoek L, van Eck-Stouten E, Proveniers MC, Torii KU, Voesenek LA, Peeters AJ, Millenaar FF (2010) Ethylene-induced hyponastic growth in *Arabidopsis thaliana* is controlled by *ERECTA*. *Plant J* 61:83–95.

Vartanian N, Marcotte L, Giraudat J (1994) Drought rhizogenesis in *Arabidopsis thaliana* (differential responses of hormonal mutants). *Plant Physiol* 104:761–767.

Wang L, Uilecan IV, Assadi AH, Kozmik CA, Spalding EP (2009) HYPOTrace: Image analysis software for measuring hypocotyl growth and shape demonstrated on Arabidopsis seedlings undergoing photomorphogenesis. *Plant Physiol* 149:1632–1637.

Wang N, Qian W, Suppanz I, Wei L, Mao B, Long Y, Meng J, Müller AE, Jung C (2011) Flowering time variation in oilseed rape (*Brassica napus* L.) is associated with allelic variation in the FRIGIDA homologue BnaA.FRI.a. *J Exp Bot* 62:5641–5658.

Weigel D (2011) Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiol* 158:2–22.

Werck-Reichhart D, Bak S, Paquette S (2002) Cytochromes P450. *Arab Book* 2:1.

White P, Bengough AG, Bingham IJ, George TS, Karley AJ, Valentine TA (2009) Induced Plant Mutations in the Genomics Era. Food and Agriculture Organization of the United Nations, Rome, pp 338–340.

Wickham H (2009) Ggplot2: elegant graphics for data analysis. New York: Springer.

Wijnker E, van Dun K, de Snoo CB, Lelivelt CL, Keurentjes JJ, Naharudin NS, Ravi M, Chan SW, de Jong H, Dirks R (2012) Reverse breeding in *Arabidopsis thaliana* generates homozygous parental lines from a heterozygous plant. *Nat Genet* 44:467–470.

Wilmouth RC, Turnbull JJ, Welford RW, Clifton IJ, Prescott AG, Schofield CJ (2002) Structure and mechanism of anthocyanidin synthase from *Arabidopsis thaliana*. *Structure* 10:93–103.

Winkler R, Freeling M (1994) Physiological genetics of the dominant gibberellin-nonresponsive maize dwarfs, *Dwarf8* and *Dwarf9*. *Planta* 193:341-348.

Wolyn DJ, Borevitz JO, Loudet O, Schwartz C, Maloof J, Ecker JR, Berry CC, Chory J (2004) Light-response quantitative trait loci identified with

composite interval and extreme array mapping in *Arabidopsis thaliana*. *Genetics* 167:907–917.

Wu J, Wei K, Cheng F, Li S, Wang Q, Zhao J, Bonnema G, Wang X (2012) A naturally occurring InDel variation in BraA.FLC.b (BrFLC2) associated with flowering time variation in *Brassica rapa*. *BMC Plant Biol* 12:151.

Xu Y, Li L, Wu K, Peeters AJ, Gage DA, Zeevaart JA (1995) The *GA5* locus of *Arabidopsis thaliana* encodes a multifunctional gibberellin 20-oxidase: molecular cloning and functional expression. *Proc Natl Acad Sci USA* 92:6640–6644.

Yamaguchi S (2008) Gibberellin metabolism and its regulation. *Annu Rev Plant Biol* 59:225–251.

Yang YH, Zhang FM, Ge S (2009) Evolutionary rate patterns of the gibberellin pathway genes. *BMC Evol Biol* 9:206.

Yuan YX, Wu J, Sun RF, Zhang XW, Xu DH, Bonnema G, Wang XW (2009) A naturally occurring splicing site mutation in the *Brassica rapa FLC1* gene is associated with variation in flowering time. *J Exp Bot* 60:1299–1308.

Appendices

Appendix 1. List of primers and PCR results used to identify deletions in *Arabidopsis thaliana* accessions T1080, YGU, Kyr-2 and Veg1-2. Col and *ga5* were used as controls. Reactions were conducted from two independent samples of genomic DNA for each accession. Product sizes were predicted using Col as reference sequence.

Primer forward	Primer reverse	Product size (kb)	Locus	T1080	Kyr-2	YGU	Veg1-2	Col	*ga5*
P01	P02	2.2	At4g25420.1	-	-	-	-	+	+
P01	P03	2.3	At4g25420.1	-	-	-	-	+	+
P01	P60	1.8	At4g25420.1	-	-	-	-	-	-
P04	P02	1.8	At4g25420.1	-	-	-	-	+	+
P06	P02	1.2	At4g25420.1	-	-	-	-	+	+
P57	P02	1.5	At4g25420.1	-	-	-	-	+	+
P57	P60	1.1	At4g25420.1	-	-	-	-	+	+
P04	P03	1.9	At4g25420.1	-	-	-	-	+	+
P01	P82	3.5	At4g25420.1	+	-	-	-	+	+
P01	P81	3.1	At4g25420.1	+	-	-	-	+	+
P01	P80	1.4	At4g25420.1	+	-	-	-	+	+
P01	P58	0.8	At4g25420.1	+	-	-	-	+	+
P22	P23	1.5	promoter At4g25420.1	+	-	-	-	+	+
P45	P46	1.0	At4g25390.1	NT	+	+	+	+	+
P118	P119	1.1	At4g25380.1	NT	+	NT	NT	+	+
P41	P42	1.0	At4g25410.1	NT	-	+	+	+	-
P25	P26	1.4	At4g25430.1	NT	-	-	+	+	+
P27	P28	1.3	At4g25434.2	NT	-	+	+	+	+
P31	P32	1.2	At4g25470.1	NT	-	NT	NT	+	+
P35	P36	1.1	At4g25520.1	NT	+	NT	NT	+	+
P110	P111	1.1	At4g25540.1	NT	+	NT	NT	+	+
P41	P26	13.9	At4g25410.1-At4g25430.1	NT	-	-	+	-	-
P41	P28	19.5	At4g25410.1-At4g25434.2	NT	NT	+	NT	-	-
P45	P32	37.9	At4g25390.1-At4g25470.1	NT	-	NT	NT	-	-
P45	P36	57.3	At4g25390.1- At4g25520.1	NT	-	NT	NT	-	-
P45	P111	67.3	At4g25390.1-At4g25540.1	NT	-	NT	NT	-	-
P118	P36	58.8	At4g25390.1-At4g25520.1	NT	-	NT	NT	-	-
P118	P111	68.8	At4g25390.1-At4g25540.1	NT	-	NT	NT	-	-
P157	P158	1.2	At4g25400.1	NT	+	NT	NT	+	+
P159	P160	1.0	At4g25480.1	NT	-	NT	NT	+	+
P161	P162	1.0	At4g25490.1	NT	+	NT	NT	+	+
P163	P164	1.1	At4g25500.1	NT	+	NT	NT	+	+
P165	P166	1.8	At4g25515	NT	+	NT	NT	+	+
P45	P162	44.6	At4g25390.1-At4g25490.1	NT	+	NT	NT	-	-
P45	P164	47.4	At4g25390.1-At4g25500.1	NT	+	NT	NT	-	-

"NT" not tested, "-" PCR product did not amplify, "+" PCR amplified.

Appendix 2. List of *GA5* haplotypes. Asterisks indicate haplotypes containing semi-dwarfs.

Haplotype	GeneBank accession numbers	Accessions sharing the same haplotype
1	KF312645	Bla-3, Kl-5, TDr-18, OW-14
2	KF312646	Bla-1, Bur-0, Enk-2, Kar-2, Kar-3, Kar-4, Kar-6, Kar-7, Kar-8, Kar-9, Kas-0, LDV-25, Rennes-1, Mar-2, Mdc-1, Mdc-29, Mdc-32, Mdc-56, Mdc-63, Mdc-95
3*	KF312647	Dja-1, Dja-10, Dja-2, Dja-3, Dja-4, Dja-5, Dja-6, Dja-7, Dja-8
4*	KF312648	Enk-1, Enk-3, Haarl-1, Ooij-1, OW-12, Sch-187, Schar-1
5	KF312649	Duiv , Nfro-1, Nfro-2, Nfro-3, Nfro-4, Nfro-5, LDV-58
6	KF312650	Fuk
7	KF312651	Je-0
8	KF312652	Kar-10, Kar-11, Kar-12
9*	KF312653	Kas-2
10*	KF312654	Kl-2
11	KF312655	Lod-2-2, San-0, Ts-5, Tou-H-13, Vel-2, Vel-7, Vgn1-1, Vgn1-2, Vgn1-3, Vgn1-4, Vgn1-5, Vou-2
12	KF312656	Neo-1, Neo-2, Neo-3, Neo-4, Neo-5, Neo-6, Neo-7, Neo-8, Sha
13	KF312657	Pak-1
14*	KF312658	Pak-3
15	KF312659	Cat-28, Cat-39, Mar-6, Pra-0
16	KF312660	Sk-2-11
17*	KF312661	Sparta
18*	KF312662	Sus-1, Sus-2, Sus-3, Sus-5, Sus-6, Sus-7
19	KF312663	T620, Zdrl-2-24
20*	KF312664	Var 2-6, Var 2-1
21	KF312665	Cat-13, Cat-27, Cat-30, Cat-33, Cat-47, MdcA-0, MdcA-6, MdcA-6-2, MdcA-11, MdcA-14, MdcA-19, MdcA-22, MdcA-44
22*	KF312666	Cat-0, Cat-10
23*	KF312667	Cat-1, Cat-5, Cat-8, Cat-19, Cat-20, Cat-22, Cat-43, Cat-44
24*	KF312668	Cat-15, Cat-17
25	KF312669	Cat-09
26*	KF312670	Cat-23, Cat-45
27*	KF312671	Mar-1, Mar-3, Mar-11
28*	KF312672	Mdc-60
29*	KF312673	Mdc-53, Mdc-122
30*	KF312674	Mdc-10, Mdc-87, Mdc-100, Mdc-113, Mdc-130
31*	KF312675	Tha-1
32*	KF312676	OW-0
33*	-	T1080

Appendix 3. List of oligonucleotides.

GA5 gene amplification

Name	Locus	Primer	Sequence (5`→3`)
P01	At4g25420.1	GA5_F_200511_LBB	TGTCCATGTTGCCACACAACA
P02	At4g25420.1	GA5_R_200511_LBB	TCCCCCATTCCCTAAACTTGCT

GA5 promoter amplification and sequencing

Name	Locus	Primer	Sequence (5`→3`)
P22	At4g25420.1	GA5_F_180712_LBB	ATTTCCAAGGCCTAGCTTCG
P23	At4g25420.1	GA5_R_180712_LBB	CCCCACCAAAAAGATCCACA

Primers used for *GA5* sequencing

Name	Locus	Primer	Sequence (5`→3`)
P02	At4g25420.1	GA5_R_200511_LBB	TCCCCCATTCCCTAAACTTGCT
P04	At4g25420.1	GA5_F_200511_LBB	TGGTTCCCGTATCTCCTCGCA
P06	At4g25420.1	GA5_F_200511_LBB	CTTCTGCGATGCGTTGGGACA

Fay & Wu's *Hn* analyses

Name	Locus	Primer	Sequence (5`→3`)
P25	At4g25430.1	unk_F_180712_LBB	TTTTCGCCTTCCACTTGTCA
P26	At4g25430.1	unk_R_180712_LBB	CTGACCAGAGCATATCGTCG
P27	At4g25434.2	NUD_F_180712_LBB	CGGGAAAGATTGCGAGAGAA
P28	At4g25434.2	NUD_R_180712_LBB	TTTGTGTTGGTGCGAGATCA
P31	At4g25470.1	CBF2_F180712_LBB	ACACGGAAATGCCAGAATCA
P32	At4g25470.1	CBF2_R180712_LBB	TATCCACGTGGCATTCACAG
P41	At4g25410.1	BHLH_F190712LBB	TTTGGTGAGATTTGGCTGCT
P42	At4g25410.1	BHLH_R190712LBB	GCACCTTGCTCTCATAACGA
P45	At4g25390.1	PKS_F_190712_LBB	TCTGTGAGTGCTTCTCCTGA
P46	At4g25390.1	PKS_R_190712_LBB	TGGTGAGAACTCACTAGGCA
P49	At4g25360.1	TBL_F_190712_LBB	GTATTGCATGCCCCTGAAGA
P50	At4g25360.1	TBL_R_190712_LBB	TTTCCCTGGCAATTCTGCAT
P35	At4g25520.1	SLK1_F180712_LBB	GTCTGGCAAGCACATCAAAC
P36	At4g25520.1	SLK1_R180712_LBB	ATGCGTATCCCAACATCACC
P53	At4g25315.1	DUF_F_190712_LBB	TTCGAGGTACGCTTCTTTGG
P54	At4g25315.1	DUF_R_190712_LBB	GCACTAAGCAGACACGATGA
P86	At4g25560.1	LAF_F_291112_LBB	AGAGAAATGGCGAAGACGAA
P87	At4g25560.1	LAF_R_291112_LBB	GGCTTTGCTACTTCTGGTGT
P90	At4g25640.2	FFT_F_291112_LBB	AGACTACTCGGTCAAGCAGA
P91	At4g25640.2	FFT_R_291112_LBB	TTCCAGACCAAAGTCCCTGA
P96	At4g25240.1	SKS_F_291112_LBB	CCCTTTCGTCTCCTACGACT
P97	At4g25240.1	SKS_R_291112_LBB	GAGACTCAACACAGGCAACT
P102	At4g25190.1	QWR_F_291112_LBB	CTCCGGTACAAGAGGAGGAT
P103	At4g25190.1	QWR_R_291112_LBB	GGCTCTCGTCCATTTCTTCA
P110	At4g25540.1	MSH3_F_060213LB	AGAAATGAAGCTGGAGGCTG
P111	At4g25540.1	MSH3_R_060213LB	GGCACATACATAAGAGGGCA
P112	At4g25610.1	C2H2_F_060213LB	CCTGAGATGTCGGTTTCCTG
P113	At4g25610.1	C2H2_R_060213LB	TCATGGACCTGTGTCCTGAT
P118	At4g25380.1	SAP_F_060213LB	GGCCATCAAAGATCACCATGT
P119	At4g25380.1	SAP_R_060213LB	TGGAAAGATGGCTTGCTTGT

Continue next page

From previous page

Primers used to identify deletions (not listed above)

Name	Locus	Primer	Sequence (5`→3`)
P03	At4g25420.1	GA5_R_200511_LBB	ACCCCAAGATGATGCATGATGAACA
P57	At4g25420.1	GA5_F_030812_LBB	GATCCATCCTCCACTTTAGA
P58	At4g25420.1	GA5_R_030812_LBB	GTGTATTCATGAGCGTCTGA
P60	At4g25420.1	GA5_R_030812_LBB	GGCTTGGAGAGTGTTCATGT
P80	At4g25420.1	GA5_R_071112_LBB	AGGTCCTGTTCCTAGTGTGA
P81	At4g25420.1	GA5_R_071112_LBB	ACCCAACTACAGAAACAAGACC
P82	At4g25420.1	GA5_R_071112_LBB	CACCCGAAAGCTAACTCACA
P157	At4g25400.1	At4g25400_F_LBB	TCCCAGAGCACAAGATCAGA
P158	At4g25400.1	At4g25400_R_LBB	GTCGATCAAAGCCAGATCGT
P159	At4g25480.1	At4g25480_F_LBB	CGCCGTCGACTTCATGATTA
P160	At4g25480.1	At4g25480_R_LBB	AGCCACACATTCATACGCAA
P161	At4g25490.1	At4g25490_F_LBB	GACACGTCACCATCTCCTTC
P162	At4g25490.1	At4g25490_R_LBB	CGATTGTAACAACAGCAGCC
P163	At4g25500.1	At4g25500_F_LBB	ACTACGCCTGCCAAAATCAT
P164	At4g25500.1	At4g25500_R_LBB	CTTCTCCATTGCAGAGGCAT
P165	At4g25515.1	At4g25515_F_LBB	GCATTAACCTGCCTGGATCA
P166	At4g25515.1	At4g25515_R_LBB	ACCACTCTCAGAAATCGTGC

Appendix 4 - Summary of QTL mapping results for control, PAC04 and the response (Δ=Control- PAC04) in the Bay × Sha mapping population. QTLs above significance threshold are highlighted. The numbers in the different traits indicate the LOD scores.

Marker	Chr	cM	gMAX	AUC	t50	u8416	gMAX PAC04	AUC PAC04	t50 PAC04	u8416 PAC04	AUC (Δ)	gMAX (Δ)	t50 (Δ)	u8416 (Δ)
T27K12	1	49.1	2	4	5	3	5	4	0	1	6	3	0	0
MSAT1.42	1	54.7	2	3	5	3	7	7	1	2	8	5	0	0
NGA128	1	61.3	2	3	4	3	7	7	1	2	8	8	0	0
IND2188	1	63.8	2	4	4	4	9	9	1	3	10	10	0	0
dCAPsAPR2	1	66.1	2	4	4	4	10	10	1	3	11	11	0	0
F5I14	1	69.6	1	3	3	3	11	11	0	2	11	11	0	0
MSAT1.13	1	76.3	0	1	0	1	7	8	0	2	8	8	0	0
MSAT127088	1	82.7	0	0	0	0	3	4	1	1	3	4	1	0
MSAT3.99	3	3.2	0	0	1	0	3	2	0	0	3	2	0	0
ATHCHIB2	3	6.6	0	0	1	0	5	3	0	0	5	4	1	0
MSAT305754	3	7.9	0	0	1	0	5	3	0	0	5	4	1	0
MSAT3.19	3	23.2	0	0	0	0	2	1	3	2	2	1	2	2
MSAT3.117	3	28.8	0	0	0	0	0	0	3	3	0	1	1	2
MSAT3.32	3	39.5	0	0	0	1	1	1	4	2	1	3	1	1
MSAT3.21	3	48.0	0	0	0	3	1	2	1	1	3	3	0	0
MSAT318406	3	53.3	0	0	0	3	1	1	0	0	2	2	0	0
MSAT3.18	3	64.1	1	0	0	3	0	0	4	0	0	1	2	1
MSAT3.70	3	72.2	1	1	0	3	0	0	4	0	0	1	2	2
MSAT4.35	4	24.2	1	1	0	1	3	3	1	0	2	1	2	1
MSAT4.15	4	33.5	1	1	0	0	3	2	0	0	2	2	1	0
CIW7	4	45.0	2	2	1	1	2	1	0	1	1	1	2	2
MSAT4.18	4	47.0	3	4	2	1	1	1	0	1	1	0	2	2
MSAT4.9	4	55.6	5	6	4	1	0	1	0	0	0	0	1	0
MSAT4.68	4	61.8	4	5	4	0	0	0	0	0	0	0	1	0
MSAT4.37	4	69.1	1	2	3	0	0	0	0	0	0	0	0	0
MSAT5.14	5	26.6	4	3	1	0	2	4	0	1	1	1	0	0
NGA139	5	30.4	3	3	2	0	4	5	0	1	2	2	0	0
MSAT512110	5	41.8	2	1	1	1	4	5	0	1	3	4	0	0
MSAT5.22	5	45.4	1	1	0	1	2	3	0	1	2	2	0	0
MSAT518662	5	62.3	0	0	0	1	1	1	3	0	1	3	1	0
MSAT520037	5	67.4	0	0	1	1	1	2	3	0	2	5	0	0
MSAT5.12	5	71.6	0	0	0	1	1	2	2	0	2	5	0	0
JV6162	5	74.2	0	1	0	1	1	2	1	0	2	5	1	0
JV7576	5	79.1	0	1	0	1	1	1	1	0	2	5	1	1

Appendix 5 - Summary of QTL mapping results for control, PAC04, PAC08, PACGA and the response in the L*er* × Sha mapping population. QTLs above significance threshold are highlighted. The numbers in the different traits indicate the LOD scores.

Marker	Chr	cM	Control	PAC04	PAC08	PACGA	Control-PAC04	Control-PAC08	Control-PACGA	PAC08-PAC04	PACGA-PAC04	PACGA-PAC04
NGA59	1	0	0	0	1	0	0	0	0	0	3	2
GENEA	1	61.9	2	2	3	1	2	4	2	3	0	2
F5I14	1	68.8	2	2	3	2	2	4	3	2	0	2
M1-13	1	78.3	1	1	3	3	2	4	4	1	0	1
M2-26	2	10	0	0	0	1	0	0	1	0	3	0
M2-17	2	41.3	2	0	0	2	0	0	3	0	2	0
ERECTA	2	42.9	3	0	0	2	0	0	3	0	2	0
NGA361	2	47.7	2	0	0	2	0	0	3	0	2	1
T2N18	2	54	1	0	0	3	0	0	3	0	2	1
T3K9	2	58.4	0	0	0	3	0	0	3	0	2	1
M2-9	2	61.1	0	0	0	3	0	0	3	0	2	1
F17A22	2	68.1	0	0	0	3	1	0	3	0	0	0
F8J2	3	59.2	0	4	0	3	4	1	3	0	0	0
M3-18	3	64.6	0	7	1	3	7	1	4	0	1	0
NGA6	3	75	1	13	1	4	12	1	5	0	1	0
M4-41	4	0	3	1	0	0	0	0	1	0	0	0
M4-39	4	0	3	1	0	0	0	0	1	0	0	0
FRI	4	0.8	3	1	0	0	0	0	1	0	0	0
M4-8	4	1.5	2	1	0	0	1	0	1	0	0	0
C6L9-78	4	10.2	1	3	0	0	2	0	0	0	1	0
M4-35	4	26.1	0	2	2	0	1	2	0	0	2	2
M4-15	4	31.1	0	2	4	0	2	4	0	1	4	4
CIW7	4	38.4	0	1	8	0	1	7	0	2	6	5
M4I22	4	49.6	0	1	5	1	1	4	1	1	2	2
M4-14	4	56.5	0	1	5	0	1	4	1	1	2	2
M4-9	4	58.1	0	1	4	0	1	3	1	1	2	1
hua2-5	5	18.4	0	1	2	0	1	2	0	0	1	1
NGA139	5	21	0	1	3	0	1	4	0	0	1	2
SO262	5	28.4	0	6	6	2	5	8	2	0	3	3
M5-1	5	33.3	0	8	9	4	7	8	3	0	3	3
M5-9	5	45.1	0	7	7	4	6	6	5	0	2	1
K9D7	5	46.1	0	8	7	5	7	7	5	0	3	2
NGA129	5	54.8	0	9	8	5	11	7	5	0	7	2
JV61/62	5	62.2	0	7	7	5	9	6	4	0	5	1
MBK5	5	73.3	0	3	4	2	3	3	2	0	2	0

Appendix 6 - Summary of QTL mapping results for control, PAC04, PAC08, PACGA and the response in the Sha × Col mapping population. QTLs above significance threshold are highlighted. The numbers in the different traits indicate the LOD scores.

Marker	Chr	cM	Control	PAC04	PAC08	PACGA	Control-PAC04	Control-PAC08	Control-PACGA	PAC08-PAC04	PACGA-PAC04	PACGA-PAC04
c3_00580	3	0.0	0	4	5	0	3	5	0	1	4	6
c3_00885	3	0.9	0	4	5	0	3	4	0	1	4	6
c3_01901	3	3.8	0	4	5	0	3	4	1	1	4	6
c3_02968	3	5.9	0	3	5	0	2	4	1	1	4	5
c3_04141	3	5.9	0	3	5	0	2	4	1	1	4	5
c3_05141	3	5.9	0	3	5	0	2	4	1	1	4	5
c3_06631	3	9.4	0	3	4	0	2	4	1	1	4	5
c3_08042	3	12.9	0	3	5	0	3	4	0	1	4	5
c3_09748	3	24.7	0	3	4	0	2	3	0	1	3	4
c5_00576	5	0.0	1	5	5	2	4	3	0	0	5	5
c5_01587	5	6.4	1	8	9	2	6	7	0	1	7	8
c5_02900	5	9.0	1	8	9	3	6	7	0	1	7	9
c5_04011	5	11.8	1	7	9	3	6	7	0	1	7	8
c5_05319	5	17.2	0	5	6	2	3	5	0	1	4	6
c5_06820	5	21.7	0	2	3	2	2	2	0	1	2	3

Appendix 7 - Summary of QTL mapping results for control, PAC04, PAC08, PACGA and the response in the Col × L*er* mapping population. *Indicates augmented markers. QTLs above significance threshold are highlighted. The numbers in the different traits indicate the LOD scores.

Marker	Chr	cM	Control	PAC04	PAC08	PACGA	Control-PAC04	Control-PAC08	Control-PACGA	PAC08-PAC04	PACGA-PAC04	PACGA-PAC04
3504562	1	16.3	0.0	1.1	2.1	0.0	1.2	2.3	0.0	1.0	1.2	2.4
FM4_2	5	76.3	1.6	1.6	0.7	0.5	1.4	0.6	0.1	0.1	1.2	0.5
c5.loc86*	5	86.0	1.8	2.4	1.2	0.7	2.2	1.0	0.2	0.1	2.0	0.9
c5.loc89*	5	89.0	1.8	2.7	1.4	0.7	2.4	1.2	0.2	0.1	2.2	1.1
c5.loc92*	5	92.0	1.8	2.9	1.5	0.7	2.6	1.3	0.2	0.1	2.4	1.3
23115566	5	117.7	1.2	1.6	1.2	0.2	1.5	1.1	0.0	0.0	1.5	1.1

Appendix 8 - Flanking 10 kb regions from top associated maker in chromosome 2 for the trait gMAX (Δ=Control- PAC04) in the Hapmap population. In bold is highlighted the position for the most significant marker.

Marker	Chr	Pos (bp)	Distance from top snp (kb)	gMAX (Δ)	Locus ID	Gene symbol
69790	2	11795352	-10.9	3.7E-01	AT2G27650	-
69791	2	11796818	-9.4	5.2E-01	AT2G27650	-
69792	2	11797139	-9.1	3.4E-01	AT2G27650	-
69793	2	11797262	-8.9	5.2E-01	AT2G27650	-
69794	2	11798019	-8.2	5.2E-01	AT2G27650	-
69795	2	11798076	-8.1	6.1E-03	AT2G27650	-
69796	2	11799504	-6.7	6.0E-02	AT2G27660	-
69797	2	11799543	-6.7	6.1E-01	AT2G27660	-
69798	2	11799786	-6.4	3.8E-02	AT2G27660	-
69799	2	11801030	-5.2	4.7E-01	AT2G27660	-
69800	2	11801330	-4.9	3.2E-01	AT2G27660	-
69801	2	11801562	-4.6	9.0E-02	AT2G27660	-
69802	2	11802817	-3.4	4.8E-01	AT2G27670	-
69803	2	11803949	-2.3	2.4E-01	AT2G27680	-
69804	2	11804774	-1.4	9.4E-04	AT2G27680	-
69805	2	11805194	-1.0	1.7E-01	AT2G27680	-
69806	2	11805633	-0.6	3.0E-05	AT2G27680	-
69807	2	11806062	-0.1	1.0E+00	AT2G27680	-
69808	**2**	**11806211**	**0.0**	**8.7E-07**	**AT2G27680**	**-**
69809	2	11806331	0.1	3.7E-01	AT2G27680	-
69810	2	11806526	0.3	5.7E-01	AT2G27680	-
69811	2	11806647	0.4	5.9E-01	AT2G27680	-
69812	2	11806766	0.6	1.6E-01	AT2G27680	-
69813	2	11806851	0.6	7.2E-01	AT2G27680	-
69814	2	11807078	0.9	4.7E-03	AT2G27680	-
69815	2	11807406	1.2	6.4E-01	AT2G27680	-
69816	2	11810505	4.3	1.1E-01	AT2G27690	CYP94C1
69817	2	11810639	4.4	6.9E-01	AT2G27690	CYP94C1
69818	2	11811188	5.0	2.8E-01	AT2G27690	CYP94C1
69819	2	11811801	5.6	2.8E-01	AT2G27690	CYP94C1
69820	2	11813421	7.2	8.1E-04	AT2G27690	CYP94C1
69821	2	11814724	8.5	7.2E-02	AT2G27700	-
69822	2	11816665	10.5	5.2E-01	AT2G27700	-

Acknowlegments

I would like to express my sincerest gratitude to Maarten Koornneef for an exceptional supervision, an awesome and interesting project to work with, all patience, teaching and critical review of preliminary results, papers manuscripts and Ph.D. thesis. Thanks for all constructive discussions, visits and *in situ* teaching in greenhouse, honesty, and for always keeping positive and supportive as well as transmitting the magnificence of science and genetics.

I would like to thank my Phd thesis committee: Ute Höcker (University of Cologne), Thomas Wiehe (University of Cologne) and Wim Soppe (MIPIZ). Besides thanks to Wim for allowing using the GERMINATOR facilities to image many of the experiments included in the thesis.

I express my gratitude to Ruben Alcázar (Universitat de Barcelona) for all teaching with molecular, population genetics, and scientific writing, and for keeping positive and sharing his enthusiasm for science.

I thank all Lab. Koornneef for support (+ lunch / coffee time): Sigi, Marianne, Valentina, Regina, Sheila, Xueqing, Corinna, Dorothea, Wilma, and Ralph. Besides, thanks to Dorothea and Manfred Wozny for help with translation of German writing parts of the thesis.

Thanks to Carlos Alonso-Blanco (Centro Nacional de Biotecnología, CSIC), Rik Kooke, and Joost Keurentjes (Wageningen University) for sharing information, genetic material and relevant comments / corrections to publish the second chapter of this thesis.

Thanks to Ulrich Schurr for the opportunity to conduct the root / water with holding experiments in IBG-2: Plant Sciences, Forschungszentrum Jülich. My sincere gratitude to Kerstin A. Nagel, Marcus Jansen, Bernd Kastenholz, Silvia Braun, Birgit Bleise, Thorsten Brehm, and Fabio Fiorani, for all help and discussions before-during-after the execution of the experiments included in chapter three.

Thanks to Ronny Joosen, Wilco Ligterink, Leo Willems (Wageningen University) for all teaching about GERMINATOR and GWAS, providing genetic material, and for help and discussions before-during-after the execution of the

experiments included in chapter four. Besides thanks to Ronny for advice and sharing ideas regarding HyPer.

Thanks to Jane Lempe and José Jiménez for all help with HyPer. Besides, thanks to José for all teaching in R and in journal clubs; I extend my gratitude to his entire group.

Thanks to Jonas Klasen for all teaching with R (+statistics+marathon). I extend my thanks to Andreas Benke for R induction course and for all discussion regarding QTL mapping.

Thanks to the Master students who joined us for six weeks term projects for the course Plant Genetics and Development (MN-B-Gen 8) University of Cologne: Katja Nagler, Helene Wiebe, and Rahere Thoma. Especial thanks for Willem van Erp who conducted a six months internship in the lab as part of his bachelor degree studies (University of Applied Sciences HAS Den Bosch).

I would like to thank Seth Davis for co-supervision during the initial phase of my Ph.D. project.

My gratitude to the International Max Planck Research School (IMPRS) for funding. Thanks to Olof Persson for managing all logistics regarding IMPRS.

Thanks to all Koornneef department, Phd Colleagues, Gardeners, Genome Center and all those who were not mentioned that provided help, and interesting discussions during my studies in the MPIPZ.

I would like to thank the directors and professors from “Centro para Investigaciones en Granos y Semillas (CIGRAS)” and “Facultad de Agronomía” from the University of Costa Rica for the authorization to conduct my Ph.D. studies in the MPIPZ. Thanks to the staff from the “Oficina de Asuntos Internacionales y Cooperación Externa (OAICE)” for dealing with all administrative issues regarding my permission to study abroad.

Thanks to my parents (Carmen and Orlando) and grandparents (Nina and Antonio) for all support, carrier orientation and life example. Thanks to my family (especially Priscilla, Laura, M.G., Gdynia, Hans, Geraldine, and Francisco C.). and friends for all support (especially Chidi, Manu and Shreyas for extra support when finishing the thesis) and for keeping the motivation and morale high.

Last but never least thanks to my wife Jhamna Magsig-Castillo for unconditional *in situ* and overseas support, for being an outstanding “manager” and for always encourage my / our dreams.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel genutzt habe. Alle wörtlich oder inhaltlich übernommenen Stellen habe ich als solche gekennzeichnet.
Ich versichere außerdem, dass ich die beigefügte Dissertation nur in diesem und keinem anderen Promotionsverfahren eingereicht habe und, dass sie abgesehen von den unten angegebenen Teilpublikationen noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde.

Die Daten über die Pflanzenhöhe, die verwendet wurden um die genomweite Assoziationsstudie (GWAS, engl. Genome-wide association study) in Kapitel 2 durchzuführen, wurden von Rik Kook (Labor für Pflanzenphysiologie, Universität Wageningen, Niederlande) erhalten. Spross-und Wurzel-Phänotypisierungsexperimente wurden in Zusammenarbeit mit dem Forschungszentrum Jülich – Pflanzenwissenschaften (IBG-2) durchgeführt. Die Phänotypisierung der Keimung für GWAS und für die Bay und Sha Populationen wurde von mir in Zusammenarbeit mit dem Labor für Pflanzenphysiologie der Universität Wageningen durchgeführt.

Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Maarten Koornneef betreut worden.

_____

Luis Orlando Barboza Barquero

Köln, November 2013

Teilpublikationen

Luis Barboza, Sigi Effgen, Carlos Alonso-Blanco, Rik Kooke, Joost Keurentjes, Maarten Koornneef, Rubén Alcázar (2013) Arabidopsis semidwarfs evolved from independent mutations in *GA20ox1*, ortholog to green revolution dwarf alleles in rice and barley. Proc Natl Acad Sci 110:15818–15823.

Lebenslauf

Name: Luis Barboza

Adresse: Carl von Linné Weg 10, 50829 Köln

Geburtsdatum: April 3rd 1982

Ausbildung

Oct, 10-Up to date. Phd International Max Planck Research School (IMPRS), Max Planck Institute for Plant Breeding Research. Under supervision Prof. Dr. Maarten Koornneef (Plant Breeding and Genetics Department). Cologne, Germany.

2006-2008 M.S.c in Plant Biotechnology (specialization in Molecular Plant Breeding), Wageningen University and Research Centrum, Wageningen, The Netherlands. Graduated.

2000-2004 B.Sc. in Agricultural Sciences, with emphasis in Plant Sciences, University of Costa Rica (UCR). Graduated.

1995-1999 María Inmaculada High School, San José, Costa Rica. Graduated.

1989-1994 Inglaterra School, San José, Costa Rica. Graduated.

Arbeitserfahrung

Oct, 08 – Sep, 10 Researcher/professor, "Centro para investigaciones en Granos y Semillas (CIGRAS)" (Seed and Grain Research Center), Agronomy Faculty, University of Costa Rica (UCR).

2004–2006 Research Assistant in the Biotechnology Laboratory of the CIGRAS, University of Costa Rica (UCR).

Stipendien

Oct, 10-Up to date Phd International Max Planck Research School (IMPRS), Max Planck Institute for Plant Breeding Research. Under supervision Prof. Dr. Maarten Koornneef (Plant Breeding and Genetics Department). Granted Fellowship: IMPRS.

2009 Internship University of California Riverside: Genetic linkage maps and QTL mapping in citrus. Under the Guidence of Mikeal Roose and Claire Federici. 4th Ausgust - 30th October. California, USA. Granted Fellowship: UCR-professors mobility.

2006-2008 M.Sc. Plant Biotechnology. Wageningen University, The Netherlands. Granted Fellowship: NFP-NUFFIC.

2006 "Molecular techniques in Aquiculture". San Pedro de Manglaralto, Ecuador. From 3rd - 14th July, 2006, 30 hours theory, 50 hours practice. Granted Fellowship: BTC/CTB Belgique.

Wissenschaftliche Publikationen

Luis Barboza, Sigi Effgen, Carlos Alonso-Blanco, Rik Kooke, Joost Keurentjes, Maarten Koornneef, Rubén Alcázar. 2013. Arabidopsis semidwarfs evolved from independent mutations in *GA20ox1*, ortholog to green revolution dwarf alleles in rice and barley. Proc Natl Acad Sci 110:15818–15823

Christiane Noronha Fernandes, Emanuelle Ferreira Melo, Antonio Chalfun Júnior, Luis Orlando Barboza Barquero, José Donizeti Alves. Modifications in the metabolism of carbohydrates in *Coffea arabica* L. cv. Siriema seedlings under drought conditions. *In press.* Coffee Science, Lavras, Brazil.

Artak Ghandilyan, Nadine Ilk, Corrie Hanhart, Malick Mbengue, Luis Barboza, Henk Schat, Maarten Koornneef, Mohamed El-Lithy, Dick Vreugdenhil, Matthieu Reymond and Mark G. M. Aarts. 2009. A strong effect of growth medium and organ type on the identification of QTLs for phytate and mineral concentrations in three *Arabidopsis thaliana* RIL populations. Journal of Experimental Botany, 60 (5): 1409–1425

Artak Ghandilyan, Luis Barboza, Sébastien Tisné, Christine Granier, Matthieu Reymond, Maarten Koornneef, Henk Schat, Mark G.M. Aarts. 2009. Genetic analysis identifies quantitative trait loci controlling rosette mineral concentrations in *Arabidopsis thaliana* under drought. New Phytologist 184: 180-192

Masterarbeit

Genetic regulation of mineral concentration in an *Arabidopsis thaliana* L*er*/An-1 mapping population grown under different environmental conditions. Chair group: Botanical Genetics. Supervisor: Artak Ghandilyan, examiner: Mark Aarts, second examiner: Maarten Koornneef. Wageningen University. M.Sc. thesis Luis Barboza, March 2008.