Exploring Natural Variation in *Arabidopsis thaliana*

– Flowering Time and Speciation –

der Fakultät für Biologie
der EBERHARD KARLS UNIVERSITÄT TÜBINGEN

zur Erlangung des Grades eines Doktors
der Naturwissenschaften

von

Johanne Lempe

aus Heilbronn

vorgelegte

Dissertation

2007

Tag der mündlichen Prüfung: 16.Oktober 2007

Dekan: Prof. Dr. F. Schöffl

1. Berichterstatter: Prof. Dr. D. Weigel

2. Berichterstatter: Prof. Dr. G. Jürgens

TABLE OF CONTENTS

III. RESULTS

IV. DISCUSSION

V. SUMMARY

I. INTRODUCTION

When contemplating the world of creatures, there are two obvious consequences of evolution: organisms adapted to their environment and the existence of distinct species.

Compared to other biological processes, both adaptation and speciation are quite slow and therefore difficult to observe. As a consequence, scientists have tried to infer information on the process of adaptation by investigating adapted organisms and on the process of speciation by considering already distinct species. The work of Miescher (1868), Correns (1900), Griffith (1928), Watson/Crick (1953) and many others has established that for each individual, the information for its adapted phenotype with its characteristic traits is stored and transmitted in what we call today the genome of an organism. Investigating phenotypic variation and with it accompanied sequence variation that is fixed in genomes will illuminate the process of evolution.

The first part of this thesis surveys naturally occurring phenotypic variation of an adaptive trait, flowering time, and sequence changes in major effect genes associated with it. In order to study speciation, one common approach has been to uncover the genetic causes for reproductive isolation that is found between already well-established species. Complementary to this approach, the second part of this thesis investigates incipient speciation, by uncovering the genetic cause for an example of hybrid incompatibility within the same species.

Arabidopsis thaliana

Research on natural variation is particularly promising in the plant model organism *Arabidopsis thaliana (A. thaliana),* for several reasons. Laibach favored *A. thaliana* already in 1943 because of low demands for cultivating it, short generation time, high fecundity, large seed set, availability of a large number of accessions that show phenotypic variation, small chromosome number and ease of crossing (Laibach, 1943). Today, *A. thaliana* is a well-established model plant with fully sequenced genome, convenient methods for transgenic manipulations, a seed collection comprising more than 1000 accessions and a large number of diverse molecular tools (Clough and Bent, 1998, www.arabidopsis.org; Arabidopsis-Genome-Initiative, 2000). The genetic resource in form of a prominent seed collection in combination with established sequencing and mapping technologies make it now possible to link

phenotypic variation, and therefore possibly adaptation, to sequence polymorphisms. However, it must be noted that not all naturally occurring variation is necessarily adaptive.

Natural variation in *A. thaliana* is abundant and the number of investigations is growing (Alonso-Blanco and Koornneef, 2000; Koornneef et al., 2004).

Under high selection pressure is the resistance to biotic factors such as bacteria, fungi, viruses and insects. Several studies identified phenotypic and genetic variation in pathogen response (Kunkel and Brooks, 2002; Koornneef et al., 2004; Perchepied et al., 2006). The diversity of disease resistance can mostly be accounted to sequence polymorphism in plant disease resistance genes (R-genes), such as *RPM1* and *RPS2* (Mindrinos et al., 1994; Stahl et al., 1999).

Variation is observed also in the responses to abiotic stress such as freezing (Alonso-Blanco et al., 2005b; Hannah et al., 2006), drought (McKay et al., 2003), UV-light (Cooley et al., 2001), salts (Quesada et al., 2002) and metals (Kobayashi and Koyama, 2002; Hoekenga et al., 2003).

Another aspect relevant for population success is the biochemical configuration such as chemical compositions and enzyme activities. Several studies show natural variation in seed lipids and oligosaccharides (Millar and Kunst, 1999; Bentsink et al., 2000), epicuticular waxes (Rashotte et al., 1997), glucosinolates (Mithen et al., 1995; Kliebenstein et al., 2001a; Kliebenstein et al., 2001c), phosphate uptake (Bentsink et al., 2003) and in the activity of enzymes involved primary and secondary metabolism (Mitchell-Olds and Pedersen, 1998; Sergeeva et al., 2004). Natural alleles with divergent activity could be identified of several genes involved in glucosinolate metabolism (*AOP2, AOP3*, *ESP*, *MAM1* and *MAM2)* (Kliebenstein et al., 2001b; Lambrix et al., 2001; Kroymann et al., 2003).

Considerable variation is also found in physiological traits such as circadian period (Swarup et al., 1999; Michael et al., 2003), nitrogen use efficiency (Rauh et al., 2002; Loudet et al., 2003), seed dormancy (van Der Schaar et al., 1997; Alonso-Blanco et al., 2003; Bentsink et al., 2006) and responses to light and hormone treatments (King and Stimart, 1998; Borevitz et al., 2002). Bentsink et al. succeeded in cloning a major effect Quantitative Trait Locus (QTL) that controls seed dormancy, a gene of so far unknown biochemical function (Bentsink et al., 2006). Natural alleles that confer reduced light sensitivity could be identified for *PHYA* and *PHYD* (Aukerman et al., 1997; Maloof et al., 2001).

Developmental traits showing large variation include floral morphology (Juenger et al., 2000), leaf morphology (Perez-Perez et al., 2002), root morphology (Mouchel et al., 2004; Loudet et al., 2005), venation pattern (Candela et al., 1999), branching (Alonso-Blanco et al.,

1999), plant size (Alonso-Blanco et al., 1999) and flowering time, which is discussed in detail in the following paragraph. Mouchel and colleagues succeeded in cloning *BREVIS RADIX,* a gene responsible for root length (Mouchel et al., 2004).

Natural variation has also been considered for genome wide traits. Differences between accessions are evident in methylation patterns (Cervera et al., 2002; Riddle and Richards, 2005), chiasmata frequency (Sanchez-Moran et al., 2002) and gene expression levels (West et al., 2007).

Flowering time

As an evolutionarily important trait we chose flowering time, which is ideal for the study of adaptation to specific environmental conditions in plants. The onset of flowering is crucial for seed set and finishing the life cycle before the growing season closes and thus for the success of plant populations. In order to achieve the optimal flowering time, plants adapted to their geographic location and hence react differently to environmental cues. Important environmental factors influencing flowering time are light quality and quantity, ambient temperature and a winter-like prolonged period of cold temperature, called vernalization. An increase in day length, elevated ambient growth temperature and past vernalization, conditions that indicate the onset of spring, promote the onset of flowering.

These and other environmental cues influencing flowering time are perceived and integrated though different pathways of a complex genetic network. Mutant analysis identified many components of this genetic network. Grouping these genes according to their response to various physiological conditions established four main genetic pathways controlling flowering time: the photoperiodic, autonomous, vernalization and the gibberellin (GA) pathway (Mouradov et al., 2002; Simpson and Dean, 2002; Putterill et al., 2004).

The photoperiodic pathway perceives light quality and quantity, which encodes information on changing day length and therefore seasonal information. Measuring day length in plants can be split into two functional domains: (I) an endogenous circadian clock and (II) a day-length measuring signal which itself is under circadian control (Imaizumi and Kay, 2006). The circadian clock in *A. thaliana* functions via transcriptional feedback loops and is reset by activity of the light receptors *PHYTOCHROME (PHY) A* to *E* and *CRYPTOCHROME (CRY) 1* and *CRY*2. The transcriptional feedback loops of the circadian clock involve the morning factors *CIRCADIAN CLOCK-ASSOCIATED1 (CCA1)* and *LATE ELONGATED HYPOCOTYL (LHY)* and the evening factors *TIMING OF CAB EXPRESSION1 (TOC1), EARLY FLOWERING4 (ELF4), AND LUX ARRHYTHMO (LUX),*

7

(Baurle and Dean, 2006; Imaizumi and Kay, 2006; McClung, 2006). The day length measuring component is *CONSTANS (CO)* (Hayama and Coupland, 2004). *CO* expression is under circadian regulation via *FLAVIN-BINDING, KELCH REPEAT, F-BOX1 (FKF1), GIGANTEA (GI), ELF3, CYCLING DOF FACTOR1 (CDF1)* and *RED AND FAR-RED INSENSITIVE 2 (RFI2)* (Imaizumi and Kay, 2006). CO protein stability and activity is dependent on day light, which is perceived via the photoreceptors *PHYA*, *PHYB* and *CRY2* (Hayama and Coupland, 2004; Valverde et al., 2004). The CO protein is active, only when the circadian peak of *CO* expression in late afternoons coincides with daylight. The active CO protein can then induce the floral integrator *FLOWERING LOCUS T (FT)* (Yanovsky and Kay, 2002).

The autonomous pathway was identified by late flowering mutants that are sensitive to photoperiod and require vernalization (Koornneef et al., 1991). It is still not completely resolved which environmental or endogenous cues are received through this pathway. Ambient growth temperature is supposed to effect the autonomous pathway, since mutants are insensitive to changes in growth temperature (Blazquez et al., 2003; Balasubramanian et al., 2006a). Up to date, seven flowering time genes are assigned to the autonomous pathway: *FCA, FY, FLOWERING LOCUS D (FLD), FPA, FVE, LUMINIDEPENDENS (LD), FLOWERING LATE KH MOTIF (FLK)* (Baurle and Dean, 2006). The autonomous pathway consists of several sub-pathways that all target *FLC*. *FCA* and *FY* physically interact and form one sub-pathway (Simpson et al., 2003). This interaction is required for autoregulation of *FCA* expression and also mediates changes in transcript processing. Several splice variants of *FCA* exist, of which only one, *FCAγ*, makes a functional protein (Macknight et al., 2002; Quesada et al., 2003). *FPA* might form another sub-pathway together with *FVE* (Simpson et al., 2003). *FVE* and *FLD* are likely to regulate *FLC* via histone acetylation; both are implicated in histone deacetylation complexes (He et al., 2003; Ausin et al., 2004). *FCA, FPA* and *FLK* encode putative RNA-binding proteins and *LD* a homeodomain protein of unknown function (Baurle and Dean, 2006).

The vernalization pathway also acts through the floral repressor *FLC*. A prolonged period of cold is perceived by the vernalization pathway and leads to a strong reduction of *FLC* levels. The temporal separation of experiencing cold and its effect on flowering suggested an epigenetic mechanism mediating the cold experience, which is indeed the case. Several factors contribute to histone modifications that accumulate in *FLC* after vernalization (Bastow et al., 2004). *VERNALIZATION INDEPENDENT3 (VIN3)* is required for histone deacetylation at *FLC* and is only expressed after a period of cold (Sung and Amasino, 2004b).

Mutants of *VERNALIZATION1 (VRN1)* and *VERNALIZATION2 (VRN2)* are insensitive to vernalization. *VRN2* was found to be important in maintaining low levels of *FLC* after vernalization (Gendall et al., 2001). *LIKE HP1 (LHP1)* is another factor found to be required for epigenetic silencing of *FLC* (Mylne et al., 2006).

The GA pathway promotes flowering via the floral pathway integrators *SOC1* and *LEAFY* (Blazquez et al., 1998; Moon et al., 2003). Mutants in GA synthesis (*ga1-3*) or in GA signaling (*gai*) are very late flowering, especially under short day conditions (Wilson et al., 1992; Jacobsen and Olszewski, 1993)

The molecular signals of the different pathways converge and activate the pathway integrators *SOC1, FT* and *LEAFY* to promote flowering. Flowering time signals are perceived in leaves but act in the shoot apical meristem where they lead to the transition from vegetative to reproductive development. Only recently florigen was identified, the signal that is able transport the flowering signal from leaves to the shoot apical meristem. It is the pathway integrator *FT* (Abe et al., 2005; Wigge et al., 2005; Corbesier et al., 2007). *FT* is induced in leaves and FT protein likely travels from leaves to the shoot apical meristem (Corbesier et al., 2007). In the shoot apical meristem, FT forms a transcriptional complex with FD to directly activate the floral meristem identity gene *APETALA1 (AP1)* and other flowering time genes and therefore induces flower formation (Abe et al., 2005; Wigge et al., 2005).

Flowering time is a continuously variable trait in naturally occurring accessions, spanning a wide range from flowering after a few weeks to many months (Alonso-Blanco and Koornneef, 2000). The functional status of the two genes *FRIGIDA (FRI)* and *FLOWERING LOCUS C (FLC)* has a major effect on flowering time and confers an early or late flowering phenotype (Clarke and Dean, 1994; Koornneef et al., 1994; Michaels and Amasino, 1999; Johanson et al., 2000). *FRI* activates *FLC* and *FLC* represses the onset of flowering (Michaels and Amasino, 1999, 2001; Sung and Amasino, 2005). Identification of *FRI* and *FLC* as major effect genes of flowering time was only possible in natural accessions different from the commonly used laboratory strains Col-0 and L*er*, where *FRI* is inactive and therefore *FLC* is not up regulated (Johanson et al., 2000; Michaels and Amasino, 1999; Sung and Amasino, 2005). Inactive alleles of *FRI* are observed frequently in natural accessions, with two different deletions that are found in Col-0 and L*er* being prevalent in many European accessions (Johanson et al., 2000). Several rarer, non-functional alleles of *FRI* have also been identified (Le Corre et al., 2002; Gazzani et al., 2003; Stinchcombe et al., 2004; Shindo et al., 2005). No such deletions have been observed in *FLC* alleles. However several alleles of *FLC* exist with large insertions in the first intron attenuating its activity (Gazzani et al., 2003; Michaels et al.,

2003; Sung and Amasino, 2005). Variation in the histone methylation pattern of *FLC* itself largely influences *FLC* activity and therefore vernalization response (Bastow et al., 2004; Sung and Amasino, 2004a; Shindo et al., 2006). Allelic variation of both genes, *FRI* and *FLC*, can explain much of the variation observed in the adaptive trait flowering time (Gazzani et al., 2003; Michaels et al., 2003; Hagenblad et al., 2004; Stinchcombe et al., 2004; Shindo et al., 2005; Werner et al., 2005a). Moreover, there is evidence that non-functional *FRI* alleles have been favored by natural selection (Le Corre, 2005; Toomajian et al., 2006).

Apart from *FRI* and *FLC,* natural variants of genes modulating floral behavior have been identified for *PHYD, PHYC, CRY2, FLM* and *HUA2* (Aukerman et al., 1997; El-Assal et al., 2001; Poduska et al., 2003; Doyle et al., 2005; Werner et al., 2005b; Balasubramanian et al., 2006b). Alleles of *PHYD, PHYC* and *CRY2* all affect responses to photoperiod, which suggests that day length exerts a strong selection pressure involved in specific local adaptation. For *PHYC,* two functionally distinct haplogroups were identified that are distributed along a latitudinal cline, which also suggests *PHYC* as target of adaptive evolution (Balasubramanian et al., 2006b). Several more natural alleles modifying flowering time are expected be identified in the near future since over a dozen other quantitative trait loci (QTL) for flowering time have already been mapped (Koornneef et al., 2004; Alonso-Blanco et al., 2005a). Recently, new allelic interactions between *FRI* and *FLM* have been identified using laboratory growth conditions with changing light quantity and quality, temperature and humidity which are closer to natural local conditions (Li et al., 2006).

One way to localize sequence variation responsible for phenotypic differences is QTL mapping in populations of recombinant inbred lines (RILs). RILs are obtained from an F2 population through successive selfing of single plants until homozygosity is reached in the sixth to eighths generation. Each line represents a mosaic of both parental genomes. An exact chromosomal map can be obtained for each line after genotyping. Because RILs are homozygous, they can be propagated and used indefinitely while keeping their genotype constant. Another approach to identify polymorphisms responsible for phenotypic variation is direct mapping in segregating F2 populations. A prerequisite for both ways is phenotypic variation of the surveyed trait. Therefore, it is important to identify accessions with phenotypic variation, which can facilitate mapping experiments. Despite the great interest in conspicuous accessions, a large study comparing flowering time under several conditions has been lacking when our flowering time study started. Two subsequent studies analyzed flowering time of 70 accessions under three different conditions, however individual

flowering time data had not been made available (Caicedo et al., 2004; Stinchcombe et al., 2004). Concurrent with our study, a flowering time screen of 192 ecotypes with an emphasis on vernalization response was performed (Shindo et al., 2005). Our study, described in the following, presents one of the largest comparisons of the flowering behavior in several different conditions. Specifically, we examined flowering in 16°C long day (16LD), 23°C long day (23LD), 16° C LD with vernalization (16LDV), and 23°C short days (23SD). These conditions represent the most important environmental factors affecting flowering time: photoperiod, ambient temperature and vernalization. Flowering times were surveyed for more than 155 accessions along with 32 flowering time mutants. A detailed examination of the functional status of *FRI* and *FLC* and their effect on flowering time is presented.

Speciation

The evolutionary process of adaptation typically leads to continuous natural variation in phenotypes. It is still an unsolved question how distinct species arise from continuous natural variation observed in populations. Different mechanisms of speciation have been proposed. Either species arise by splitting of old ones, or new species evolve by hybridization or polyploidization of already existing species. Hybridization, in which diploid hybrids give rise to new lineages, is estimated to give rise to 50% of all angiosperm species, whereas 35-50% of angiosperms are thought to originate through polyploidization (Stebbins, 1950; Grant, 1971; Arnold, 1997; Koch et al., 2003). Also in animals, hybrid speciation seems to be more common than previously thought but is still controversial (Mallet, 2005; Arnold and Meyer, 2006; Gompert et al., 2006). Despite its proposed role in speciation, only few cases of hybrid speciation have been convincingly documented (Rieseberg, 1997; Gompert et al., 2006). By contrast, much knowledge accumulated on species evolution via splitting of already existing species. This fact can most likely be attributed to the widely accepted Biological Species Concept (BSC) of Ernst Mayr (Mayr, 1942, 1963; Coyne and Orr, 2004): "species are actually or potentially interbreeding natural populations that are isolated from other such groups". Gene flow between species might occur, but should be negligible (Coyne and Orr, 2004). This concept relies on reproductive isolation or hybrid incompatibility as criterion to define species and gives an ideal starting point to study speciation.

The BSC is by far not the only species concept and debates about such still persist (Levin, 2000; Coyne and Orr, 2004). It is not surprising that no single concept is suitable for the contingencies of species diversity. Each concept is useful for some purpose but also has its limitations (Coyne and Orr, 2004). Whereas the BSC is based on reproductive isolation,

other species concepts were proposed that emphasize other features of species. Morphological or genetic differences are emphasized in the Genotypic Cluster, Recognition and the Cohesion Species Concepts (Paterson, 1985; Tempelton, 1989; Mallet, 1995). Evolutionary tendencies and ecology are the basis of species concepts of Van Valen and Wiley and Simpson (Van Valen, 1976; Wiley, 1978). Evolutionary history of species forms the basis of the species concept of Cracraft, and Shaw (Cracraft, 1989; Shaw, 1998).

Hybrid incompatibility is the key criterion to define a species according to the BSC. When considering hybrid incompatibility, a conceptual problem arises: how can an allele evolve, that has detrimental effects when being in a heterozygous state in hybrids, without passing a "heterozygous" transition state during the course of its evolution? This heterozygous intermediate would be highly maladaptive by definition of the BSC and is unlikely to evolve.

A genetic solution to the evolution of hybrid incompatibility was put forward three times independently by Bateson, Dobzhansky and Muller and is therefore called the Bateson-Dobzhansky-Muller (BDM) model (Bateson, 1909; Dobzhansky, 1937; Muller, 1942). In this model, two populations diverge from a common ancestor by changes at complementary genetic loci that each functions well in its own genetic context, even when heterozygous. However, when two derived loci come together in a hybrid, a negative epistatic interaction between the two loci has detrimental effects (Figure 29).

Considering the immense interest on speciation, surprisingly few speciation genes following the BDM model have been cloned so far: the *Drosophila* genes *Hybrid male rescue (Hmr)* and its interacting partner *Lethal hybrid rescue (Lhr)* (Hutter and Ashburner, 1987; Hutter et al., 1990; Barbash et al., 2000; Barbash et al., 2003; Brideau et al., 2006); *Nucleoporin96 (Nup96)* (Presgraves, 2003; Presgraves et al., 2003); the *Odysseus* Homeobox gene *(OdsH)* (Ting et al., 1998; Sun et al., 2004; Wu and Ting, 2004) and *Xmrk-2* of *Xiphophorus* (Wittbrodt et al., 1989; Schartl et al., 1999). A common feature of these genes is their fast evolving nature (Coyne and Orr, 2004; Orr, 2005; Mallet, 2006).

On the basis of so far identified speciation genes, it becomes clear that speciation genes are of diverse function. The MYB related gene *Hmr* and the homeobox gene *OdsH* are involved in gene regulation (Ting et al., 1998; Barbash et al., 2003). *Hmr* is expressed at all stages during the life cycle and is speculated to play a normal role in gene regulation (Barbash et al., 2003). *OdsH* is expressed in testis and involved in male fertility. *Xmrk-2* codes for a receptor tyrosine kinase, which causes tumor formation when it is misexpressed (Malitschek et al., 1995). *Nup96* encodes a structural protein of a nuclear pore complex (Presgraves,

2003). Lhr is an interacting partner of Heterochromatin Protein 1 (HP1) with so far unknown function (Brideau et al., 2006). Moreover, these genes involved in speciation show signatures of rapid evolution as well as positive Darwinian selection (Ting et al., 1998; Barbash et al., 2003; Presgraves et al., 2003; Orr, 2005; Wen et al., 2006; Presgraves and Stephan, 2007). Positive selection indicates that reproductive isolation arises as by-product of evolution.

Negative epistatic interactions leading to hybrid incompatibilities do not only occur between nuclear loci, but also between the co-adapted cytoplasmic and nuclear genomes. In the copepod *Tigriopus californicus*, the interaction of two genes of the mitochondrial electron transport system, cytochrome c and cytochrome c oxidase, leads to incompatibility between two different populations (Willett and Burton, 2004; Ellison and Burton, 2006; Harrison and Burton, 2006). Cytonuclear incompatibility also accounts for the reproductive isolation between two species of the monkeyflower, *Mimulus gutatus* and *M. nasutus* (Fishman and Willis, 2006). The sterility between these two *Mimulus* species reflect a cytoplasmic male sterility (CMS) (Fishman and Willis, 2006). Males are sterile due to the interaction of a maternally inherited mitochondrial factor and a nuclear gene, a *restorer-of-male-fertili*ty gene, inherited by the male (Chase, 2007). Two scenarios are possible how CMS can evolve: Either both loci evolve independently in two separate populations and the interaction of both derived loci in hybrids would cause male sterility, or male sterility evolves first by acquiring a new cytoplasmic genotype. This would lead the population through a deep fitness valley that would be passed by evolving a nuclear *restorer-of-male-fertili*ty gene. The second scenario is more likely despite the fitness valley, since male sterility is theoretically favored by selection (Fishman and Willis, 2006). CMS individuals are suggested to have a slightly higher fitness through an increased seed production. Resources that are needed for pollen production in hermaphrodites can be used to produce seeds in male sterile individuals. Also CMS mutations can spread rapidly due to high seed production.

Gene transposition is another mechanism that can lead to reproductive isolation not mediated by the BDM model: The gene *JYAlpha* is located on the fourth chromosome in *Drosophila melanogaster* and on the third chromosome in *D. simulans*. The hybrid fraction which completely lacks this gene is male sterile (Masly et al., 2006). Masly and colleagues did find only one copy of *JYAlpha* in both *Drosophila* species. However, one can speculate that two copies of *JYAlpha* did exist at some point during evolutionary history and one copy got lost. Supporting this hypothesis is the fact that gene duplication-transposition events are quite common in *Drosophila* (Masly et al., 2006). The hypothesis of evolution of postzygotic

isolation via gene duplication followed by complementary loss of function is quite intriguing, however, there has been little evidence so far (Coyne and Orr, 2004; Lynch and Force, 2000).

Although genic incompatibilities in plants retained little attention in speciation research, several of them have been described in diverse taxa, such as *Crepis, Gossypium, Hordeum, Lycopersicon, Mimulus, Nicotiana, Phaseolus, Oryza, Solanum and Triticum,* (Hollingshead, 1930; Wiebe, 1934; Caldwell and Compton, 1943; Gerstel, 1954; Savant, 1956; Hermsen, 1963, 1967; Chu and Oka, 1972; Canvin and McVetty, 1976; Phillips, 1977; Macnair and Christie, 1983; Christie and Macnair, 1984; Singh and Gutiérrez, 1984; Valkonen and Watanabe, 1999; Yamada et al., 2001 ; Kruger et al., 2002; Kubo and Yoshimura, 2002; Mino et al., 2002; Yamada and Marubashi, 2003; Moyle and Graham, 2005; Sweigart et al., 2006). Genetic analysis revealed several two-gene incompatibilities following the BDM model (Wiebe, 1934; Caldwell and Compton, 1943; Hermsen, 1963; Chu and Oka, 1972; Christie and Macnair, 1984; Singh and Gutiérrez, 1984; Kubo and Yoshimura, 2002). However, since genetic and molecular tools are sparse in non-model plants, none of the incompatibility loci has been identified so far.

I discovered an incompatibility between the two *A. thaliana* accessions Uk-1 and Uk-3. Since many molecular tools are available for *A. thaliana,* we aimed to investigate the incompatibility on the macroscopic and molecular level as well as to map the causative loci.

II. METHODS

Standard laboratory buffers, media and cloning procedures were made and performed after "Molecular Cloning: A Laboratory Manual" by J. Sambrook (Sambrook et al., 1989). Chemicals were obtained from Roth (Karlsruhe, Germany), Sigma-Aldrich (Taufkirchen, Germany), VWR/Merck (Darmstadt, Germany) unless noted otherwise. All buffers were prepared in ddH$_2$O.

1 Non-standard protocols

1.1 Plant genomic DNA extraction with CTAB

The protocol for DNA extraction from plants protocol was modified after Doyle and Doyle (Doyle and Doyle, 1987). It was used for individual DNA isolation as well as for DNA isolation in 96-well format. If cleaner DNA was needed, RNAse treatment was performed in an extra step as mentioned in the following protocol.

CTAB- buffer (2% w/v CTAB, 1,42 M NaCl, 20 mM EDTA, 100 mM Tris, pH 8.0) was preheated to 65°C and 0,2 % v/v ß-mercaptoethanol was added; 2 μl RNAseA per sample was added if no extra RNAse treatment was done later.
Frozen plant tissue was ground in liquid nitrogen with a pistil. For 96-well format, plant material was dried under vacuum and ground as dry matter using a grinding mill and steal beads (Retsch MM300, Haan, Germany).

Tissue powder was suspended in 500 μl of the preheated CTAB-buffer with beta-mercaptoethanol and incubated at 65 °C for 1 h. 500 μl chloroform:isoamylalcohol (24:1) was added, mixed by inversion and centrifuged at maximum speed (3000-10000 G) for 20 min at room temperature. The supernatant was transferred into a new tube.

Extra RNAse steps: 10 μl RNAseT1 mix per sample was added (Fermentas, St. Leon-Rot, Germany) and incubated at 37 °C for 30 min. 500 μl chloroform:isoamylalcohol (24:1) was added and mixed by inversion

and centrifuged at full speed for 20 min. The supernatant was

transferred into a new tube.

0,7 volumes (~300-400 μl) isopropanol was added and mixed by inversion, cooled at 4 °C for 20 min and centrifuged at full speed for 30 min at room temperature. The supernatant was removed and the pellet was washed with 100 μl 70 % ethanol (EtOH) and again centrifuged at full speed for 20 min. The pellet was air dried and suspended in ddH$_2$O.

1.2 TRIzol RNA extraction

Frozen plant material was homogenized with a pistil and 1 ml TRIzol (Invitrogen, Karlsruhe, Germany) per sample was added. The suspended plant powder was incubated for 5 min at room temperature. 500 μl chloroform was added, shaken vigorously and centrifuged at 11500 rpm for 10 min at 4°C. The supernatant was transferred into a new tube, 500 μl isopropanol was added and mixed by inversion. The sample was incubated at room temperature for 10 min and centrifuged at 11500 rpm for 10 min at 4°C. The supernatant was removed and the pellet washed with 1 ml 75% EtOH. The pellet was air dried and dissolved in 30-50 μl of DEPC treated ddH$_2$O.

1.3 Tissue embedding in plastic

Solutions:

FAA: 50% EtOH, 5% acetic acid, 10% formaldehyde

Solution A: 1g Technovit Hardener I, 100 ml Technovit 7100 (Kulzer, Wehrheim,
              Germany).

Solution B: 1 ml Hardener II, 15 ml Technovit 7100

Plant material was fixed in FAA for 2 h, then dehydrated with an EtOH series (each step for 15 min: 25%, 50%, 70%, 90% and twice 100%). The tissue was infiltrated with 50 % solution A in EtOH for 30 min, followed by 100% solution A for 30 min, then left over night in 100% solution A at 4°C. For polymerization, solution A was replaced by solution B. Solution B was exchanged once before specimens were oriented properly in plastic molds within 30 min. The molds were covered with parafilm, air bubbles removed and polymerized at 4°C over night.

For staining, a droplet of 0.1% toluidine blue (0.1% toluidine blue, 0.1% sodium borate) was added onto sections, left for 1 min on a heated plate (65°C) and excess stain rinsed off with H$_2$O.

1.4 Fosmid library screening

Solutions:

2x SSPE: 0.3 M NaCl, 20mM NaH$_2$PO$_4$H$_2$O, 2mM EDTA, pH 7.4

20x SSC: 3 M NaCl, 0.3 M NaCitrate, pH 7.0

Maleic acid buffer: 100 mM maleic acid, 150 mM NaCl, pH 7.5

Washing buffer: 100 mM maleic acid, 150 mM NaCl, 0.3% (v/v) tween20, pH 7.5

Detection buffer: 100 mM Tris-HCl, 100 mM NaCl, pH 9.5

Blocking solution: 1% blocking reagent (Roche, Mannheim, Germany) in maleic acid buffer

Prehybridization solution: 5x SSC, 0.1% (w/v) N-lauiylsarcosine, 0.02% SDS, 1% blocking reagent

DIG-labeled Probe synthesis: The DIG Probe Synthesis Kit from Roche (Basel, Switzerland) was used and prepared following manufacturers descriptions. Before use, the probe was denaturized at 100°C for 10 min. The probe was reused 4 times. The primers G-5550/1 and G-5556/G- 7890 were used to generate the probes (primer sequences are listed in appendix 1).

Striping solution: 0.2N NaOH, 0.1% SDS

Making replica filter:

Bacteria containing fosmid clones (see 3.6) were plated onto a nictrocellulose membrane (Amersham Hybond™-N+, Freiburg, Germany) that was placed on a media plate and incubated over night at 37°C. A replica filter was obtained by removing the membrane from the plate, placing another nitrocellulose membrane on top of the membrane and pressing them onto each other using two glass plates. Both filters were put back onto media plates and were again incubated at 37°C for 2-4 h, until colonies are 2 to 4 mm in diameter.

Lyses of colonies, denaturing and cross-linking DNA to membrane

Membranes were placed with colony side up successively onto three different Whatmann filters, saturated with:

1) 10 % SDS (for 3 min),

2) 1.5 M NaCl, 0.5 M NaOH (for 5 min) and

3) 1.5 M NaCl, 0.5 M Tris-HCl pH 7.4 (for 5 min).

Finally, filters were floated in 2xSSPE for 5-15 min. Membranes were dried on paper towels and DNA was fixed to the membranes by using the Stratalinker® UV crosslinker (Stratagene, La Jolla, USA).

Hybridization of membranes with DIG-labeled probes

Membranes were first washed in a tray with 2x SSC for 5 min, then transferred into a rolling bottle with 6x SSC and incubated at 50 °C for 30 min. Cell debris could carefully be removed with Kim wipes soaked in 6x SSC.

For pre-hybridization, filters were transferred into a rolling bottle and incubated with pre-hybridization solution at 57°C for more than 4h. Subsequently, the denatured, labeled probes were added to the pre-hybridization solution and incubated with the membranes at 57°C in a rolling bottle over night. All filters were hybridized with both probes at the same time.

The probe hybridization was followed by a series of washing steps:

- once in 2x SSC, 0.1% SDS in a tray on a shaker for 30 min at room temperature
- four times in 2x SSC, 0.1% SDS in a rolling bottle for 30 min at 57°C
- four times in 0.5x SSC, 0.1% SDS in a rolling bottle for 30 min at 57°C
- once in 1x Maleic Acid buffer for 5 min at room temperature.

Probe detection

For detection of the DIG-labeled probe, the membrane was first blocked with 1x blocking solution for 30-60 min at room temperature under gentle agitation. The blocking solution was exchanged by antibody solution (1x blocking solution with 1:10000 Anti-Digoxygenin AP, Roche, Basel, Switzerland) and again incubated for 60 min. This step was followed by a repeated washing step in washing buffer. The membrane was then equilibrated in detection buffer. 50-100μl/100 cm^2 CDP-Star ready-to-use solution (Roche, Mannheim, Germany) was distributed on the membrane surface before the membrane was sealed between two sheets of plastic foil. A film was exposed to the sealed membrane over night.

Stripping membranes

All membranes were probed several times. Therefore, the remaining probe was removed by the following steps:

- wash plot in sterile ddH$_2$O for 2-5 min
- strip with 500 ml striping solution at 37°C for 30 min
- wash with 2 x SSC at room temperature for 1 min
- prehybridize for 5 min as described above and continue with probe hybridization.

2 Flowering time variation in *Arabidopsis thaliana*

2.1 Plant growth conditions

All experiments were carried out in growth rooms with highly controlled temperature (4°C, 16°C or 23°C), humidity (65%) and light conditions (125 to 175 μmol m^{-2} s^{-1}) with a day length of 16 hours in long day condition (LD) and 8 hours in short day (SD).

Before planting on soil, seeds were stratified in 0.1% agarose at 4°C for 7 days. For vernalization treatment, seeds were allowed to germinate at 16°C for 24 hours and then vernalized at 4°C SD with a low fluence rate (50 μmol m^{-2} s^{-1}) for 5 weeks.

2.2 Experimental design

All accessions were ordered from the European Arabidopsis Stock Centre (http://arabidopsis.info/), accession numbers can be found in appendix 2.3.

For each accession, 12 replicate plants were planted in a completely randomized design. Therefore, individually numbered positions in 60-well flats were assigned randomly to each accession and replicate. In total, four identical sets of plants were sown out and surveyed under 4 different growth conditions: 16°C long day (16LD), 16LD plus vernalization (16LDV), 23°C long day (23LD) and 23°C short day (23SD). The following traits were scored: days to flowering (DTF), juvenile leaf number (JLN), adult leaf number (ALN), cauline leaf number (CLN), rosette leaf number (RLN), total leaf number (TLN). As DTF, the time from sowing to the first macroscopic appearance of a bolting was determined. Accessions were controlled for bolting every day. The vernalization period was subtracted from DTF. Juvenile and adult leaves were distinguished by absence or presence of trichomes on the adaxial side of the leaf.

In order to minimize unwanted variation in growth conditions, flats were watered with an identical amount of water and were rotated every other day.

The experiment was carried out in two sets. The first set consisted of mostly vernalization requiring accessions and seeds from the stock centre were used. For the second set, single-seed descents from stock center seeds were used. Some accessions were included in both sets

of the experiment. Since variation between same accessions in the two sets of experiments was insignificant we merged the two datasets.

2.3 Statistical analysis of leaf numbers

For further analysis the average of all replicates of a given accession were taken. Data points with less then 5 scored replicates were removed from the analysis. Some accessions with extremely high flowering time variation were also removed from further analysis. This variation might be explained by segregation or seed contamination.

Data analysis was done with the base package of R (www.r-project.org) and Microsoft Excel. Broad sense heritability (H^2) was calculated as the between line variance (V_G) divided by the total variance. The coefficient of genetic variation (CV_G) was calculated as (100 x $\sqrt{V_G}$) / mean for each trait. Genetic correlations (r_{GE}) were calculated from covariance and variance components estimated through a restricted maximum likelihood method (REML), using those genotypes for which data across all environments was available. The significance of each genetic correlation was determined using a t-test after Z transformation of the correlation coefficient (Sokal and Rohlf, 1981). Genotype-by-environment (G x E) interactions were tested using a two-way ANOVA with strain and conditions as classifying factors. To determine day length, temperature and vernalization sensitivities of each accession, the average TLN of the two compared conditions was regressed to the individual TLN of the accession (Falconer and Mackay, 1996):

$$\frac{\log(TLN_{environmentA}) - \log(TLN_{environmentB})}{\log(meanTLN_{environmentA}) - \log(meanTLN_{environmentB})}$$

Log transformed data showed a distribution close to a normal distribution and therefore were preferred.

2.4 Analysis of *FRGIDA* and *FLOWERING LOCUS C:* sequencing and genetic complementation

All accessions were genotyped for the known L*er-* and Col-type deletions in *FRIGIDA* (*FRI*) by PCR using primers G-0360/G-0362 and G-0363/G-0365 respectively (see appendix 1 for all primer sequence). All accessions were tested for presence of the L*er*-type insertion in *FLOWERING LOCUS C* (*FLC*) (primers N-0914, N-0913, N-1124).

FRI and *FLC* functionality was tested genetically by complementation crosses to *FRI*-Sf-2/ *FLC*, *fri*-Col/*FLC* (Col-0), *FRI*-Sf-2/*flc-3, fri*-Col/*flc-3* (seeds obtained originally from R.M. Amasino). TLN of the F1 generation was scored in 23LD.

For sequence analysis, *FRI* was amplified as 3.3 kb fragment (G-0360/G-0109) and *FLC* as two overlapping fragments covering 4.2 kb of *FLC* (G-1426/G-1797, G-1799/G-1805). Fragments of the above mentioned accessions were amplified with ExTaq polymerase (*TaKaRa,* Biomedical, Shiga, Japan). Four independent reactions were pooled and sequenced from both strands (see appendix 1 for primer list). Sequences were assembled with AutoAssembler 2 and aligned with ClustalX (Thompson et al., 1997).

The cDNA sequence of *FLC* was obtained by reverse transcription using 1 μg RNA of 5 pooled plants grown for 10 days in 23LD. RNA was extracted with the Trizol method described above, treated with DNAse (Fermentas, St. Leon-Rot, Germany) and reverse transcribed using the Invitrogen Reverse Transcription Kit. For the accessions Kin-0, Lip-0 and Ll-2 RNA of 6 days old seedlings had to be used because expression levels were too low in 10 days old plants.

For transformation, 35S-constructs of FLC cDNA sequences from Cen-0, Ll-2, Lip-0, Ll-2 were made. Sf-2 served as positive control, an empty vector as negative control. PCR-fragments of FLC were amplified from cDNA with primers containing BamHI and SalI restriction sites (G-3024/G-3025). Fragments were first cloned into pGEM®-T Easy (Promega, Manheim, Germany), digested by restriction and sub cloned into the CHF3 vector. Flc-3 plants in Col-0 background (fri/flc) were transformed with the described constructs. T1 and T2 generations were scored for flowering time in 23LD.

Expression levels of *FRI* and *FLC* were tested in RNA samples by quantitative real time PCR (qRT-PCR) using primers G-0831/G-0832 and G-1981/G-1982 respectively. qRT-PCR was performed with SYBR Green (Invitrogen, Carlsbad, California, United States) and monitored with the Opticon Continuous Fluorescence Detection System (MJ Research, Reno, Nevada, United States). Two technical replicates of each accession and ubiquitin expression were measured. The relative expression ratio to the positive control Sf-2 was quantified as

$$2^{-[(C_{sample}-C_{ubiquitin})-(C_{Sf2}-C_{ubiquitin})]}$$ (Pfaffl, 2001).

2.5 Correlation of flowering time with gene expression

Levels of FLC expression were determined for all accessions surveyed in the flowering time screen using 12 days old seedlings grown at 23LD. Two biological replicates, each with two technical replicates, were analyzed. RNA samples were prepared and processed as described in the previous paragraph. The expression ratio was calculated relative to Col-0.

As part of the AtGenExpress experiment (www.arabidopsis.org/info/expression/ATGenExpress.jsp), expression profiles of 38 accessions were determined by microarray analysis. RNA of aerial parts of 5 days old seedlings was extracted with the RNeasy® Plant Mini Kit (Qiagen, Hilden, Germany). Samples were processed as described at www.weigelworld.org/resources/microarray/AtGenExpress and hybridized to GeneChip® Arabidopsis ATH1 Genome Arrays (Affymetrix, Santa Clara, USA). Triplicates of 10 accessions and singletons of 28 accessions were hybridized (for list see appendix 2.4). Normalized expression estimates were obtained using gcRMA (www.bioconductor.org).

Pearson correlation coefficients of flowering time in 23LD with expression levels of all genes represented on the ATH1 array were determined using R. Bootstrapping was performed by calculating correlation coefficients of a random set of 16 accessions for 100 times. To test significance of a correlation value, maximum correlation values were determined for 1000 random correlations.

3   Uk-1/Uk-3 hybrid incompatibility

3.1 Scanning Electron Micrographs

Samples for Scanning Electron Micrographs (SEM) were fixed in 100% EtOH for 5 minutes, washed twice and stored in 100% EtOH. Critical point drying and sputter coating was performed by Jürgen Berger.

3.2 GUS reporter essay

PR-1 expression was determined using a transgenic line containing a *PR1*::GUS construct (N6358). This line was obtained in *ndr1-1* Col-0 background. *PR1*::GUS was crossed to the Col-0/F1$_{(Uk-1xUk-3)}$ backcross line Col-12-2 that differs genotypically from Col-0 almost exclusively in the incompatibility regions on chromosome 3 and chromosome 5 intervals (see methods 3.4 and appendix 3.2). GUS staining of plants with wild type and F1 phenotype was performed as described in Weigel and Glazebrook (Weigel and Glazebrook, 2002).

3.3 Microarray experiment

Uk-1, Uk-3 and F1 plants were grown at 23SD in plant incubators for 27 days and then shifted to 16SD. Samples were collected in triplicates at the day of the temperature shift and 5 days after the shift. The experiment was carried out under SD to avoid floral induction in F1 and Uk-3 plants. Per genotype and replicate, 40 apices were collected in a random design. RNA was extracted with the Qiagen RNeasy RNA extraction kit and further processed as described at www.weigelworld.org/resources/microarray/AtGenExpress/. Samples were hybridized to Affymetrix ATH1 microarrays (see appendix 3.3 for list of micro arrays). Expression data were normalized with gcRMA (www.bioconductor.org) and significant expression differences were determined by Rank Product, implemented in R (Breitling et al., 2004), www.bioconductor.org) and LogitT (perl script by Alexis Maizel, (Lemon et al., 2003)) with p=0.05. To filter for genes that are differentially expressed in the different growth conditions we used GeneSpring 7.2. We were interested in the list of genes differentially expressed in F1 hybrids at 16°C versus 23°C from which the following genes were

subtracted: genes that change expression significantly in parents at 16°C versus 23°C, genes with significant expression differences in Uk-1 versus Uk-3 and in F1 versus parents. This resulted in a list of 232 genes that was further analyzed with GeneMerge v1.2 for Gene Ontology of Biological Processes (Castillo-Davis and Hartl, 2003).

3.4 Mapping with single nucleotide polymorphism markers

289 single nucleotide polymorphism (SNP) markers were developed by Norman Warthmann for different purposes and tested for polymorphism in 86 accessions, among them Uk-1, Uk-3 and Col-0. 64 SNP markers were polymorph between Uk-1 and Uk-3 and 136 markers showed the same genotype for Uk-1 and Uk-3 but were different to Col-0.

The two SNP marker sets were tested on several backcross lines. Therefore, the F1 generation of Uk-1 x Uk-3 was backcrossed to Uk-1, Uk-3 and Col-0 for three generations. 10, 11 and 12 lines were kept separately for Uk-1, Uk-3 and Col-0 backcrosses, respectively. The incompatibility regions were maintained in the lines by phenotypic selection at 16LD and crosses were made after shifting to 23LD where flowers developed.

The first marker set of 64 SNPs that could distinguish between Uk-1 and Uk-3 was tested on an F2 population of 181 plants (Table 1, Figure 20, appendix 3.1). Of the second marker set, 26 SNP makers were chosen in a way that one marker was located on each chromosome arm, one in the middle of each chromosome and 11 covering the incompatibility region (estimated by the results of the first mapping set). 96 individuals of backcross lines of with Col-0, including Col-0, Uk-1 and Uk-3 as controls, were genotyped with the set of 26 SNP markers (Table 1, appendix 3.2).

Table 1: *Summary table of plants genotyped with SNP marker*

Genotype	Phenotype	number of plants
Uk-1 back cross	f1-phenotype	10
Uk-3 back cross	f1-phenotype	11
F2(Uk-1 x Uk-3)	f1-phenotype	70
Uk-1 back cross	wt	10
Uk-3 back cross	wt	11
F2(Uk-1 x Uk-3)	wt	69
Total		**181**
Col back cross	f1-phenotype	66
Col back cross	wt	27
total		**93**

All SNP genotyping was performed by the Genaissance Pharmaceuticals, New Haven, CT, USA.

To map and test for epistatic interaction the software R package qtl was used (Broman et al., 2003). LOD scores were calculated with a single QTL model using maximum likelihood via the EM algorithm. Significance of the LOD score was determined by 1000 permutations. Epistatic interaction was tested with a model predicting two QTL.

3.5 Fine Mapping

The following work on fine mapping concentrated on the chromosome 3 interval.

The mapping interval of interest could be confined by exact positioning of recombination events, which was done by sequencing fragments of recombinant lines within the mapping interval (for chromosomal position and primers used see Table 2).

A large population of F2 $_{(Uk-1x\,Uk-3)}$ plants was grown at 16°C for fine mapping. 864 individuals that showed the F1-phenotype were selected and numbered. As control, 96 wild type-like plants were included. DNA was extracted of all plants in 96-well format using the CTAB method (see 1.1). All samples were genotyped with an simple sequence length

26

polymorphism marker (SSLP) and an cleaved amplified polymorphic sequence marker (CAPS) that were designed at end positions of the mapping intervals (Table 2). After recombinant F2 plants were identified, the mapping interval was again narrowed down by sequencing small PCR fragments (Table 2). Sequences were analyzed with AutoAssembler 2.0 and the staden package pregap4 and gap4 (Staden et al., 2000).

Table 2: *markers used for screening an F2 population for recombination events (red) and to narrow down recombination events by sequencing.*

bp position chromosome3	marker type	primer pair
16052350	sequencing	G-5107/08
16106757	CAPS/MunI	G-5111/12
16108466	sequencing	G-5369/70
16119094	sequencing	G-5504/05
16121387	sequencing	G-5550/51
16126994	sequencing	G-5552/5989
16135008	sequencing	G-5554/55
16137924	sequencing	G-6065/66
16143749	sequencing	G-6036/37
16149004	sequencing	G-5371/72
16174807	sequencing	G-5564/65
16199116	sequencing	G-5288/89
16207000	SSLP	G-5115/5368
16296875	sequencing	G-5290/91
16413043	sequencing	G-5292/93
16628529	sequencing	G-5119/20
17301878	sequencing	G-5123/24

3.6 Fosmid library construction

Plant DNA of Uk-1 was extracted with the CTAB method, sheared by pipetting and size selected on an agarose gel. The fosmid library was constructed by Christa Lanz, using the

Copy Control Fosmid library Production Kit (Epicentre, Madison, USA) according to manufacturer's instructions.

About 20.000 fosmid clones were plated onto nitrocellulose membranes on media plates and screened for the insert of interest as described in 1.4.

The copy number of the fosmid was induced with the CopyControl induction solution (Epicentre, Madison, USA). DNA was isolated with buffers R1, R2, R3 of the R.E.A.L. Prep 96 Plasmid Kit (Qiagen, Hilden, Germany) without using filter columns. Each sequencing reaction was performed with 1 μl BigDye, 5 pmol primer, 1 μg DNA and ddH$_2$O to a volume of 10 μl with the following PCR cycles: 95°C for 5 min; 74 cycles of 95°C for 20 sec, 50°C for 20 sec, 60°C for 4min.

Two fosmid clones number 4 and 11 together covered the mapping interval and were sequenced. Therefore, fosmid DNA, free of genomic DNA was prepared with the Qiagen Large-Construct Kit.

Shot gun clones were produced by Christa Lanz with the Invitrogen TOPO® Shotgun subcloning kit (Invitrogen, Karlsruhe, Deutschland). All sequences of shot gun clones originating of two overlapping fosmid clones were assembled to a 74.8 kb contig with Phred/Phrap/Consed by Christa Lanz. The obtained contig was aligned to the Col-0 sequence of the BAC-clone F14L2 using the alignment software muscle (Edgar, 2004).

3.7 Testing candidate genes

Genomic complementation with candidate genes

Genomic constructs of Uk-1 alleles of both candidate genes with promoter region (at3g44530 and at3g44540) and of a large intergenic region were made. Together the three fragments cover the mapping interval. Fosmid clone 11 was restricted with ClaI to obtain at3g44530. Clone 4 was restricted with SacI & KpnI to obtain the intergenic region. To obtain the genomic fragment of at3g44540 two fragments, one from each fosmid clone, were ligated in a single step with the plasmid. Therefore, fosmid clone 4 was restricted with KpnI & BsaI and clone 11 with PstI & BsaI. All fragments were cloned into pGREEN-IIS_21000 and co transformed with pSOUP into *Agrobacterium tumefaciens* strain *ASE*.

All three constructs were transformed into Col-0 backcross lines Col-6-2 as well as Uk-3 (for SNP genotypes of Col-6-2 see appendix 3.2). Only Col-6-2 lines were used that were heterozygous at the chromosome 5 interval. Therefore, plants were genotyped with a CAPS marker (primers: G-5481 and G-5482, enzyme: Bsp1407I).

Artificial micro RNAs against two candidate genes

Artificial miRNAs (amiRNAs) were designed with the Web MicroRNA Designer (http://wmd.weigelworld.org/bin/mirnatools.pl) (Schwab et al., 2006) against at3g44530 and at3g44540. They both match to the corresponding sequence of Uk-1 as well as Col-0. Fragments containing the amiRNA and the backbone of miRNA 319a were generated using primers G-7134 – G-7141 and cloned into the plasmid CHF3 containing the 35S promoter. Uk-1 as well as Uk-3 plants were transformed with both constructs and over-expression of amiRNAs was confirmed by northern blot. T1 plants were crossed to the alternate parent and the F1s were screened at 16LD for the incompatibility phenotype.

Expression levels of both genes were tested by qRT-PCR as described in 2.4 by using the following primers for at3g44530, at3g44540 and ubiquitin as control: G-5552/G-5989, G-6061/G-9161 and G-1586/G-1587.

Ethanol inducible expression of at3g44530

To obtain an EtOH inducible construct, the Uk-1 allele of at3g44530 was cloned via BJ36-AlcA into pMLBART-AlcR. Therefore, fosmid clone 11 was restricted with SapI & Bsu36I. The construct was transformed into Col-6-2 backcross lines that were heterozygous at the chromosome 5 incompatibility region. Induction by EtOH was proofed by qRT-PCR as mentioned above.

Search for unique sequence changes in at3g44530 and at3g44540

A set of 96 DNA samples (appendix 3.6) was genotyped with G-9582/G-9583 for a 20 bp deletion in the promoter of at3g44530. The same set of DNA samples was used to sequence PCR fragments of exon 5 (G-6056/G-7883) and exon 11 (G-6051/G-6052) of at3g44530 and exon 4 (G-6061/G-7885) of at3g44540. Sequences were aligned with gap4 (Staden et al., 2000).

3.8 Other aspects of the Uk-1/Uk-3 incompatibility system

Genetic variation between Uk-1 and Uk-3

The SNP dataset of 86 accessions (data obtained from Norman Warthmann) was used to determine the relative genetic variation between Uk-1 and Uk-3. The number of polymorphic SNPs between each possible pair of accessions was determined by using R (www.r-project.org).

Crosses to other accessions

Uk-1 and Uk-3 were crossed to the following other accessions and screened at 16LD for the incompatibility phenotype. Flowering time was measured as TLN in the crosses with accessions presented here in bold (for accession numbers see appendix 2.3, Umkirch lines see below, for flowering time data see appendix 3.10).

Col-0	**Lip-0**	Umkirch 2-1-a
Ler	Lö-2	Umkirch 2-1-l
Ak-1	**M3385S**	Umkirch 2-2-i
Bd-0	**Nok-0**	Umkirch 2-3-h
BG1	**Np-0**	Umkirch 3-1-g
Bla-2	Nw-1	Umkirch 3-2-b
Bla-3	**Pog-0**	Umkirch 3-3-a
Chi-0	Te-0	Umkirch 4-1-b
Cvi-0	**Ts-1**	Umkirch 5-1-b
En-1	Uk-2	Umkirch 5-2-c
Er-0	**Uk-4**	Umkirch 5-3-g
Fr-6	**Wt-4**	Umkirch 6-1-a
Fr-7	Umkirch 1-1-e	Umkirch 7-1-d
Jl-1	Umkirch 1-2-x	Umkirch 8-1-d
Kn-0	Umkirch 1-3-f	

Natural temperature regime

In May/ June 2005, Uk-1, Uk-3, F1 and F2 seeds were placed onto soil without stratification. They grew outside on a balcony in Tübingen and were watered if required.

Average temperatures from Freiburg were obtained from the "Meteorologisches Institut Albert-Ludwigs Universität Freiburg". Monthly averages of the years 1931-1960 are shown. For Tübingen, data originate from the "Staatliche Archiverwaltung Baden-Württemberg", Landkreis Tübingen, 1967, as average of the years 1881-1930.

Umkirch collection

18 different lines were collected from 8 different locations at Umkirch (close to Freiburg, Germany) at May 11th 2005. For a map of exact collection sites see appendix 3.7.

III. RESULTS

1 Flowering time variation in *Arabidopsis thaliana*

1.1 Observed variation in flowering time

A random set of 155 accessions was chosen to screen for flowering time variation. The selection of wild strains covers most of the northern hemisphere, which represents the habitat of *A. thaliana* (Figure 1). Flowering time was also scored for 32 lines with known mutations in the flowering time signaling pathway. In total, a complete data set for all four environmental conditions was obtained for 177 strains.
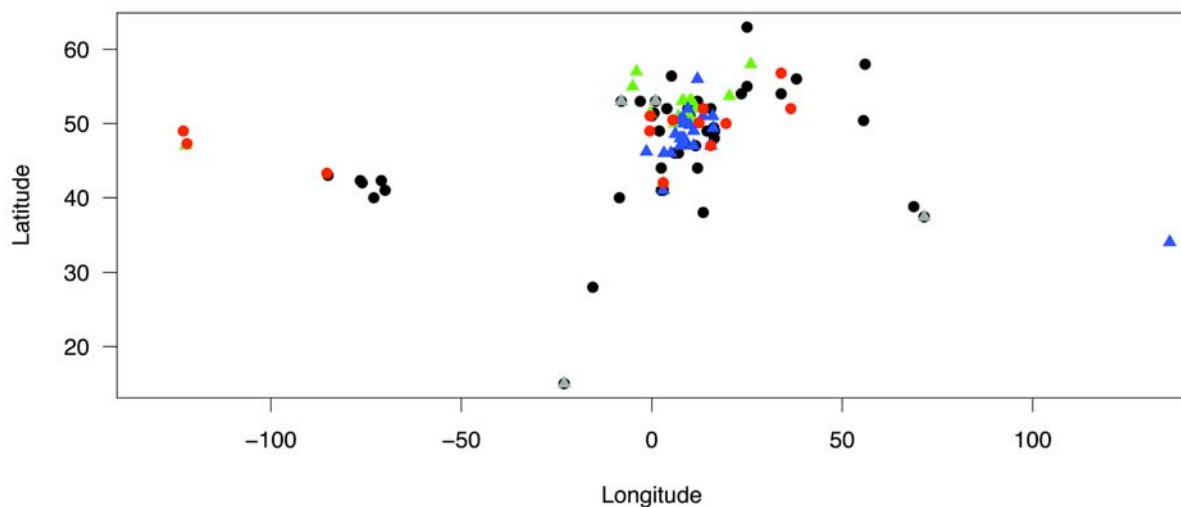
Figure 1: Geographic distribution of accessions that were surveyed for their flowering time. Black circles indicate accessions with functional FRI allele. Accessions with Col- or Ler-type deletion are represented by green and blue triangles, respectively. Accessions with rare, non-functional FRI alleles are shown in grey triangles, and red circles when identified in this study.

Flowering time of all accessions was scored as total leaf number (TLN) and days to flowering (DTF). TLN was split into juvenile, adult and cauline leaf number (all data listed in appendix 2.3). In general, TLN and DTF are highly correlated (r^2=0.6 to r^2=0.88, Figure 2).

33

Only few accessions showed variation in growth rate between the different environments. Since heritability of TLN was higher than heritability of DTF, further analysis were done with TLN (see appendix 2.1).
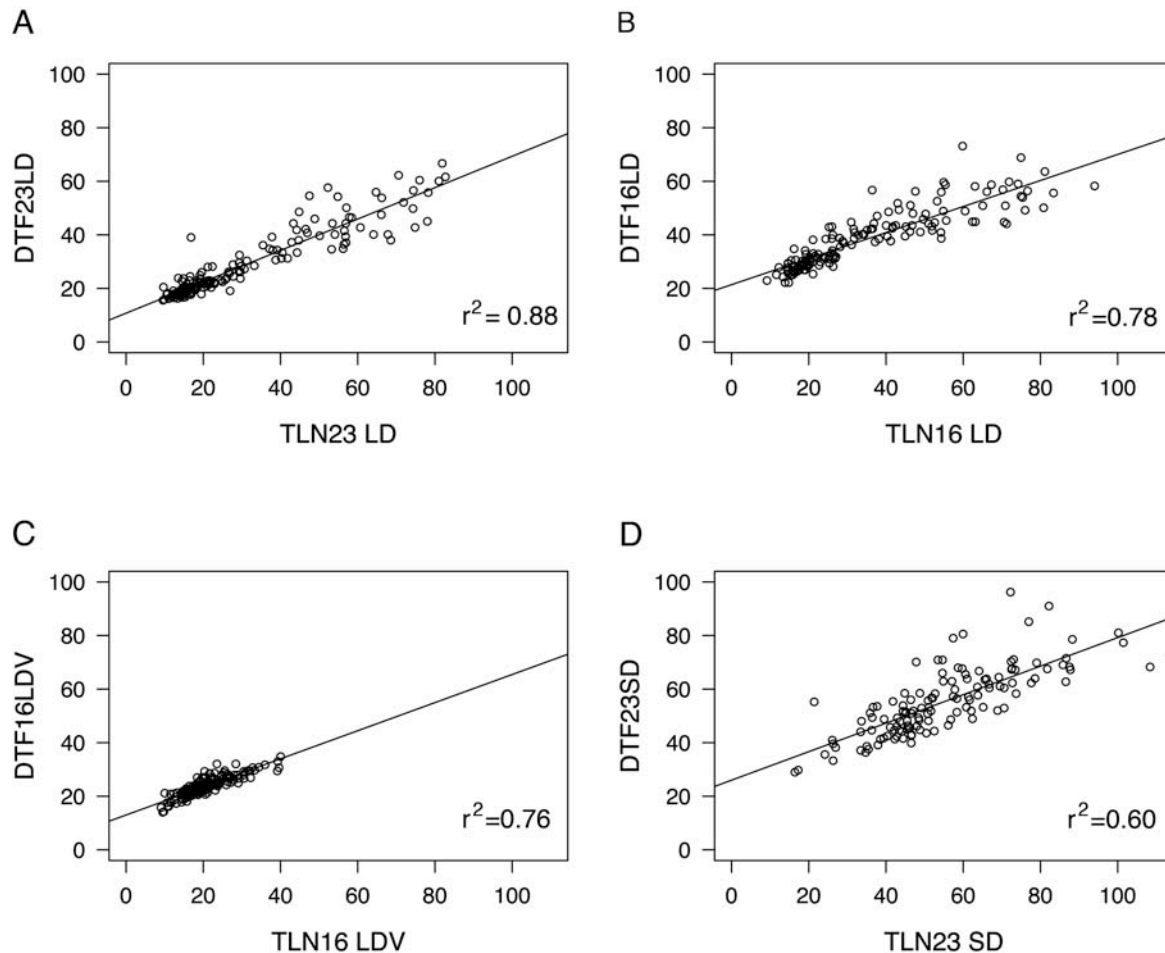
Figure 2: The regression of TLN to DTF shows high correlation between the two flowering time measurements under all examined growth conditions.

We found extensive variation in flowering time among the different accessions (Figure 3). An overall effect of the environmental cues temperature, day length and vernalization was obvious when comparing flowering behavior across different environments. At 23LD and 16LD, the majority of accessions flowered relatively early but several late-flowering accessions broaden the distribution (Figure 3 A & B). As a general trend, flowering time was slightly earlier at 23LD compared to 16LD. Under short day conditions the flowering time distribution is much broader and the onset of flowering is rather late (Figure 3 C). The long tail of late-flowering accessions in 16LD disappeared after vernalization but there is still extensive flowering time variation reaching from a TLN of 9 to 42 (Figure 3 D). Juvenile and adult phase lengths also vary in different accessions (appendix 2.3).
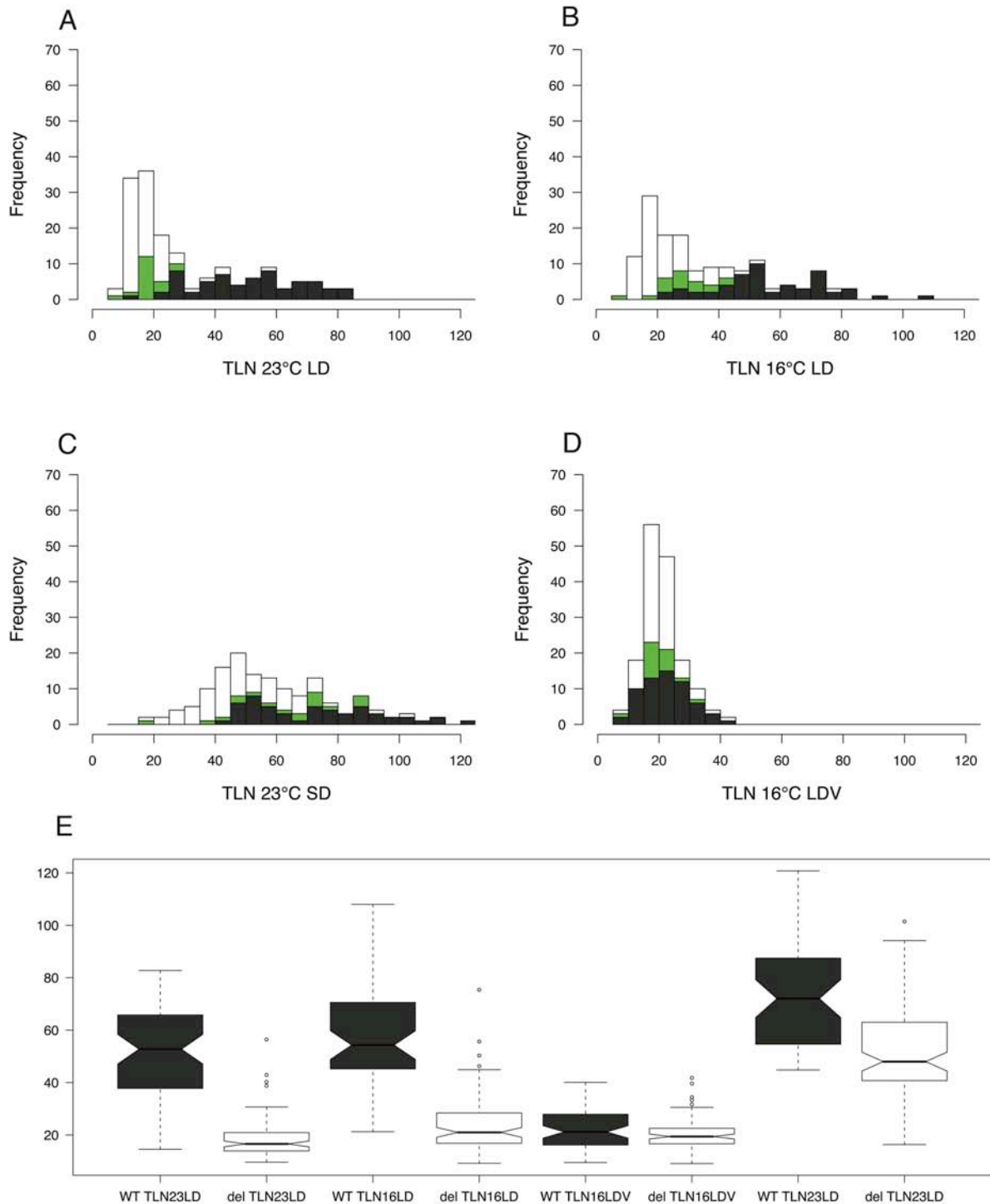
Figure 3: Flowering time of A. thaliana accessions grown in four different conditions: 23LD (A), 16LD (B), 23SD (C), 16LDV (D). White bars represent accessions with Col- or Ler-type deletions in FRI, black bars accessions without such deletions. Green bars represent accessions with defects in FRI or FLC identified in this study. (E) Summary of flowering time distributions under all growth conditions for putatively FRI/FLC functional (black) and non-functional (white) strains.

1.2 Effect of *FRI/FLC* on flowering time variation

Of 155 wild strains, 67 share the common L*er*- and Col-type deletions in *FRI*. These accessions flower early in 23LD and 16LD (Figure 3 A & B). However, none of the surveyed strains contain the insertion in the first intron in *FLC* that can be found in L*er*. Strains without the deletion in *FRI* flower mostly late and are vernalization sensitive. Few strains lacking a deletion in *FRI* are also early flowering and only weakly respond to vernalization. We used genetic complementation tests as well as sequence information to determine the functionality of *FRI/FLC* in 19 of these accessions. Crosses of 13 strains to all *FRI/FLC* wild type/ mutant combinations (*fri/flc, FRI/flc, fri/FLC, FRI/FLC*) eliminated an early flowering phenotype by complementing a weak allele of either *FRI, FLC* or both (Figure 4). A functional allele of *FRI* could complement the early flowering phenotype of Ang-1, BG1, BG9, Pog-0 and Ri-0. Bd-0, Cen-0, Gr-3, Lip-0, Ll-2 and Rsch-0 flowered later after complementation with a functional *FLC* allele (Figure 4 A & B). Jl-1 and Kin-0 were only late-flowering after complementation with functional alleles of both *FRI* and *FLC* (Figure 4 C).
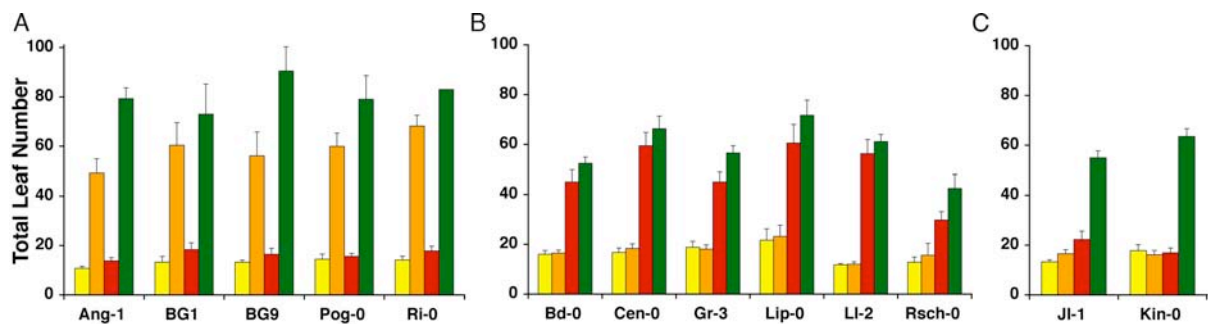
Figure 4: Flowering time of several accessions crossed to fri/flc (yellow), FRI/flc (orange), fri/FLC (red), FRI/FLC (green) in the F1 generation. Complementation of early flowering was achieved by a functional allele of either FRI (A) or FLC (B) or by both FRI and FLC (C).

To identify the genomic alterations underlying the non-functional alleles we performed sequence analysis of all 13 accessions. We found four different *FRI* alleles that contain polymorphisms leading to a premature stop codon (Figure 5 A). Jl-1 shows extensive sequence polymorphism in *FRI* with four indels in the coding region, all of them are in-frame. Kin-0 also lacks *FRI* activity, but does not show any obvious sequence polymorphisms. *FRI* is expressed in all surveyed accessions. Only in Ang-1 expression levels are low (Figure 6 A).
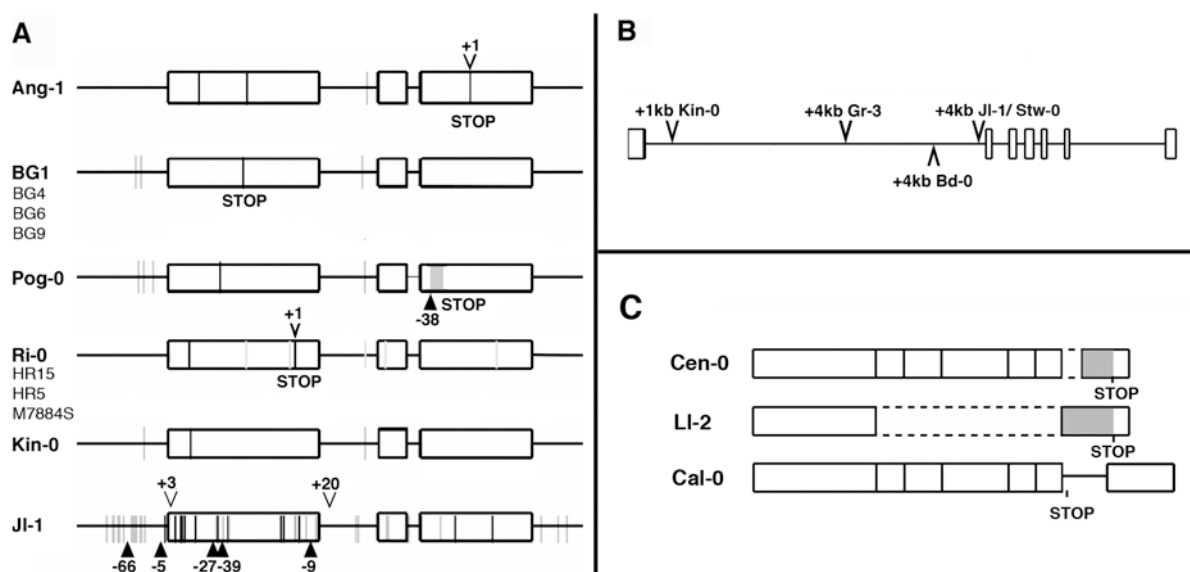
Figure 5: Newly identified alleles of FRI and FLC with exons and introns represented as boxes and lines, respectively. (A) Diagrams of FRI alleles. Grey and black lines indicate synonymous and non-synonymous polymorphisms. Insertions and deletions are depicted as triangles, numbers indicate size in base pairs. A premature stop codon (STOP) is caused by a single base insertion in Ang-1 and Ri-0, by a single base pair change in BG1 and by a deletion and inversion (grey box) event in Pog-0. The accessions BG4, BG6, BG9, HR15, HR5 and M7884S were only partially sequenced and found to have the same premature stop as BG1 and Ri-0. (B) Diagram of genomic FLC sequence with indication of big insertions in the first intron (triangles). (C) Diagram of changes in FLC transcripts. Alleles of Cen-0 and Ll-2 have deletions (doted lines) that cause a frame shift and therefore an altered amino acid sequence (grey box) and a premature stop codon (STOP). In Cal-0, alternative splicing of the last intron leads to a premature stop codon.

The genomic sequence of *FLC* revealed large insertions located in the first intron in Bd-0, Gr-3, Jl-1 and Kin-0. The insertions are likely to explain impairment of *FLC* alleles in these accessions (Figure 5B). Since the more or less unaltered genomic *FLC* sequence could not explain successful complementation by *FLC* in Cen-0, Lip-0, Ll-2 and Rsch-0, we analyzed *FLC* cDNA in these accessions. Cen-0 lacks 28 bp at the beginning of the last exon, which leads to a frame shift and a truncated protein lacking the last 72 residues (Figure 5 C). The cDNA of Ll-2 *FLC* consists of only the first and the last exon and therefore is very likely to be non-functional. Cal-0, which we also analyzed is also mis-spliced and contains 64 bp of the last intron. The cDNA sequence of Lip-0 and Rsch-0 is identical to the Col-0 sequence and cannot explain non-functionality of *FLC*. However, the level of *FLC* expression is very low in Lip-0 and Rsch-0 - even though *FRI* is expressed in these accessions (Figure 6 A & B).

37

Overexpressing *FLC* coding sequences in *flc-3* background had no significant effect in case of Ll-2 (Figure 6 C). The Cal-0 allele delayed flowering slightly and significantly. Transgenic lines overexpressing the *FLC* coding sequence of Cal-0 flowered nearly as late as a wild type allele.
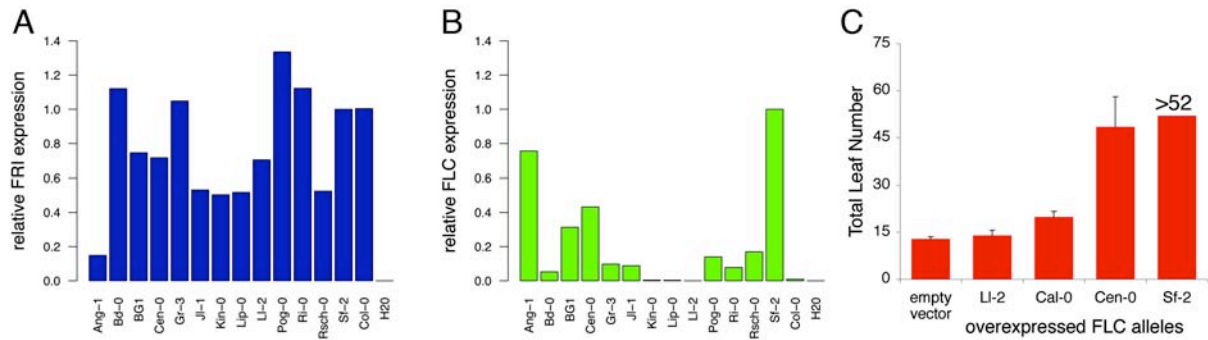
Figure 6: Expression levels of FRI (A) and FLC (B) in 10 days old seedlings by qRT-PCR relative to the level in Sf-2. (C) Total Leaf Number of transgenic lines overexpressing the coding sequence of different FLC alleles. An empty vector and the Sf-2 allele serve as negative and positive control, respectively.

All identified accessions with rare alleles of either *FRI* or *FLC* are summarized in Table 3.

Table 3: *Summary table of accessions with comprised FRI or FLC function.*

Accession	*FRI* sequence*	*FRI* compl.	*FLC* 1st intron*	*FLC* cDNA*	*FLC* compl.
Ang-1	stop (1)	no	n/d	n/d	yes
Bd-0	no change	yes	~4kb ins. 3' (3)	no change	no
BG1	stop (2)	no	n/d	n/d	yes
BG4	stop (2)	n/d	n/d	n/d	n/d
BG6	stop (2)	n/d	n/d	n/d	n/d
BG9	stop (2)	no	n/d	n/d	yes
Cen-0	no change	yes	no change	del./frame shift (5)	no
Gr-3	no change	yes	~4 kb ins. 3' (2)	no change	no
HR5	stop (4)	n/d	n/d	n/d	n/d
HR15	stop (4)	n/d	n/d	n/d	n/d
Jl-1	polymorphic (6)	no	~4 kb ins. 3' (4)	no change	no
Kin-0	no change (5)	no	~1 kb ins. 5' (1)	no change	no
Lip-0	no change	yes	no change	no change	no
Ll-2	no change	yes	no change	del./frame shift (6)	no
M7884S	stop (4)	n/d	n/d	n/d	n/d
Pog-0	stop (3)	no	n/d	n/d	yes
Ri-0	stop (4)	no	n/d	n/d	yes
Rsch-0	no change	yes	no change	no change	no
Stw-0	no change	n/d	~4 kb insertion 3' (4)	n/d	n/d
Cal-0	n/d	n/d	n/d	del./frame shift (7)	n/d

In total, we found 87 accessions with comprised *FRI* or *FLC* function; 67 accessions with either a Col- or L*er*-type deletion in *FRI* and 20 accessions with previously unknown defects in *FRI* or *FLC* (appendix 2.3). An ANOVA model of all flowering time data at 23LD with the status of *FRI*/*FLC* as a single factor and TLN as the dependent variable was able to explain 63% of the total variation in flowering time and highly significant ($p<0.001$). Under short day conditions, *FRI*/*FLC* have a smaller, but significant ($p<0.001$), effect on flowering time and account for only 23% of the observed variation.

1.3 Flowering time variation and *FLC* expression

FLC is considered a key messenger of flowering time. It is activated by *FRI* and presumably other upstream pathways. We therefore wanted to know whether *FRI*-independent variation acts upstream or downstream of *FLC*. Expression levels of *FLC* were measured in 150 strains grown at 23LD by qRT-PCR. Correlation of flowering time in all strains with measured *FLC* levels was highly significant (p<0.0001) with FLC levels explaining 42% of the observed variation in TLN and 40% in DTF (Figure 7 A). The correlation between flowering time and *FLC* levels was not significant in accessions with non-functional *FRI*. High levels of *FLC* are mostly found in accessions with functional *FRI*, but the combination of non-functional *FRI* and high levels of *FLC* also occurs.
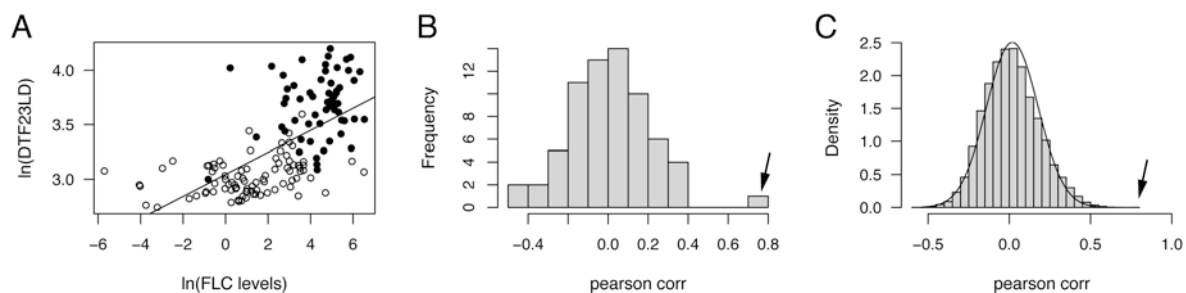
Figure 7: Correlation of FLC expression with flowering time. (A) Correlation of flowering time as DTF in 23LD with FLC levels measured by qRT-PCR. Accessions with functional and non-functional FRI/FLC are represented with filled circles and hollow circles, respectively. (B) Histogram of correlation coefficients for expression levels of 68 flowering time regulators with flowering time. Arrow indicates FLC. (C) Histogram of correlation coefficients of expression levels for all genes represented on the ATH1 microarray. The arrow indicates FLC, the line indicates the distribution recovered from 1000 permutations.

To determine whether the correlation of *FLC* levels and flowering time is significant in a genomic context we used Affymetrix microarray technology. Expression levels of all genes that are represented on the microarray were correlated with flowering time at 23LD in 34 accessions. The correlation of *FLC* levels with flowering time was very similar to the estimates from the above mentioned qRT-PCR analyses ($r^2 = 0.37$ with TLN; $r^2 = 0.42$ with DTF). Among 68 known flowering time genes (appendix 2.5), the correlation of expression

levels with flowering time was by far the highest in *FLC* (Figure 7 B). When all genes were considered, *FLC* was among the ten most highly correlated genes (Figure 7 C). This is true for both Pearson and Spearman Rank correlations. Interestingly, among the most highly correlated genes, *FLC* levels changed more than 200-fold, whereas the levels of other genes with high correlation coefficient changed less than 1.5 fold.

1.4 Variation in responses to ambient temperature, vernalization and day length

The statistical interaction of genotype and environment was tested in a two-way ANOVA model with accession and growth condition as factors. Although this model accounts for 92% of the total variation, only 19% are explained by genotype by environment interactions (appendix 2.2). This represents the action of known factors controlling flowering time, some operate in multiple environments and others only under selected conditions.

Flowering time responses were measured by comparing flowering times of the corresponding environments: 23LD with16LD for temperature response, 16LD with 16LDV for vernalization response and 23LD with 23LD for day length response. Individual flowering times were regressed onto environmental means (see methods 2.3). We used these environmental sensitivity values to identify accessions with conspicuous behavior as interesting candidates for further analysis. Accessions with extreme sensitivity value are outside the red dotted line (Figure 8).
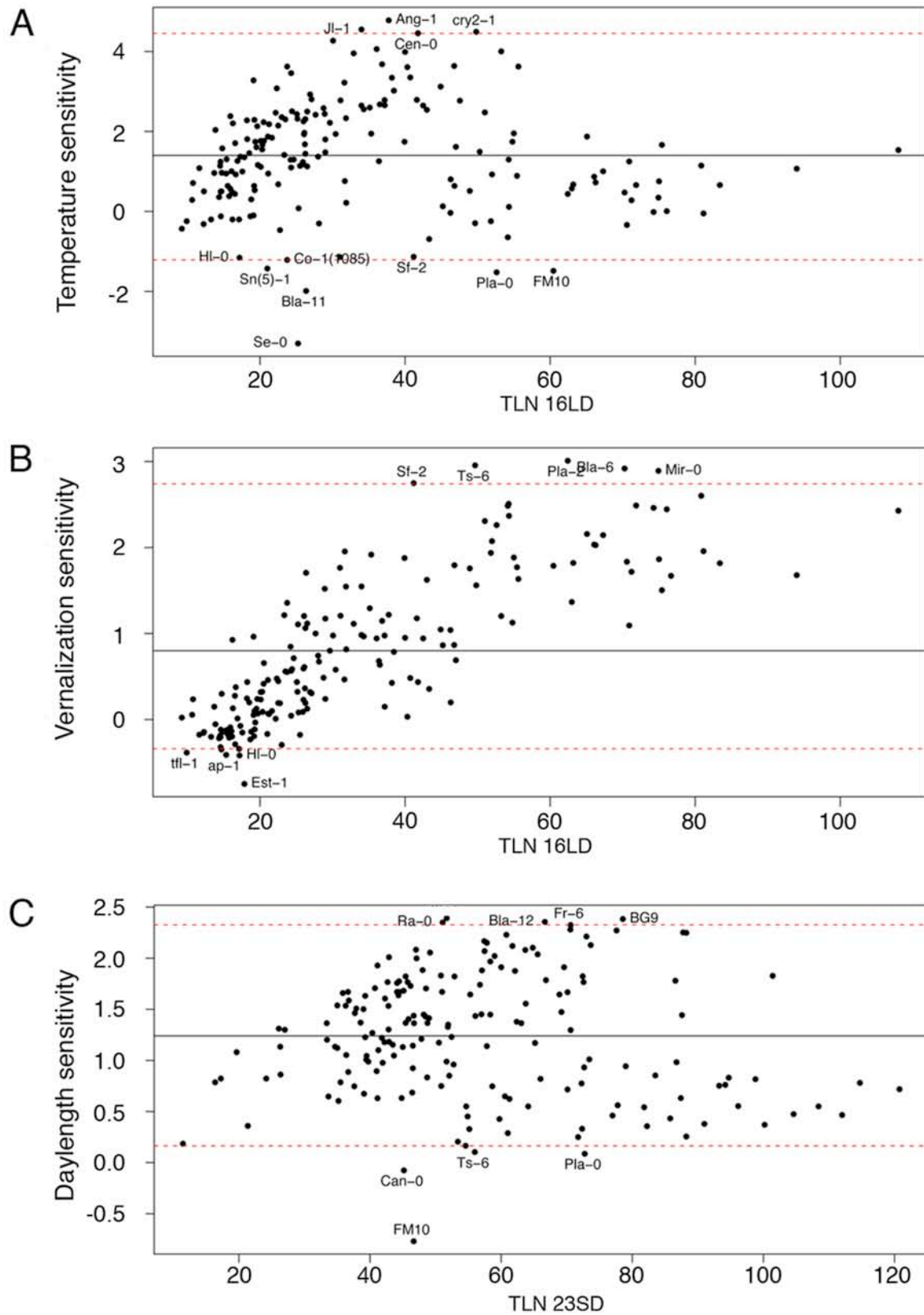
Figure 8: Environmental sensitivities to temperature (A), vernalization (B) and day length (C) plotted against TLN. 95% of data points are within the red dotted lines.

1.5 Correlation of latitude with flowering time

Without vernalization, there was no significant correlation between latitude and flowering time. Following vernalization, however, we observed a relatively high ($r^2 = 0.13$) and significant ($p < 0.0001$) correlation. This correlation was even stronger when only *FRI/FLC* functional accessions were considered ($r^2 = 0.18$) (Figure 9 A). Consistently, there was also a correlation between latitude and vernalization sensitivity (Figure 9 B; $r^2 = 0.12$; $p < 0.0001$).
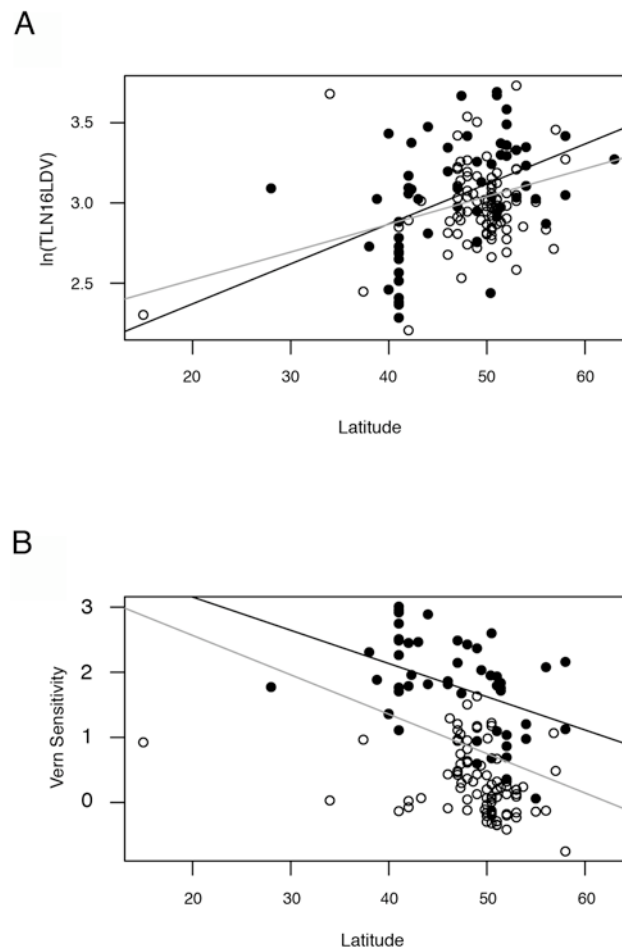
Figure 9: Latitudinal cline of flowering time after vernalization. Correlation of latitude with flowering time (A) and vernalization sensitivity (B). Accessions with putatively functional and non-functional FRI/FLC are represented by filled and open circles, respectively. Regression was performed separately on accessions with functional FRI/FLC (black line) and accessions with non-functional FRI/FLC (grey line).

2 Uk-1/Uk-3 hybrid incompatibility

At the beginning, we were interested in the flowering time of Uk-1/Uk-3 hybrids. However, the extremely unusual behavior of these hybrids shifted our interest from flowering time to speciation.

2.1 Phenotypic appearance of Uk-1/Uk-3 hybrids

Growth deficiency in F1 hybrids

Whereas the parents Uk-1 and Uk-3 appear as robust plants, F1 hybrids of these accessions are severely affected in their appearance with much reduced growth and abnormal leaf development (Figure 10).

Figure 10: The parents Uk-1 and Uk-3 and an F1 hybrid grown at 16LD. Although plants are all of the same age, they differ hugely in size. The white bar represents 1 cm.

At 16°C, F1 hybrids are dwarf-like and stay very small throughout their live time. Hybrid leaves appear dark green, glossy and are often mal-shaped (Figure 11 A, B, C). In older F1 plants, many secondary meristems are active which leads to a bushy but small habit (Figure 11 B, C, E). Since hybrids do mostly not flower at 16°C, they do not complete their live cycle and stay alive for more than 12 months without producing seeds (Figure 11 C).
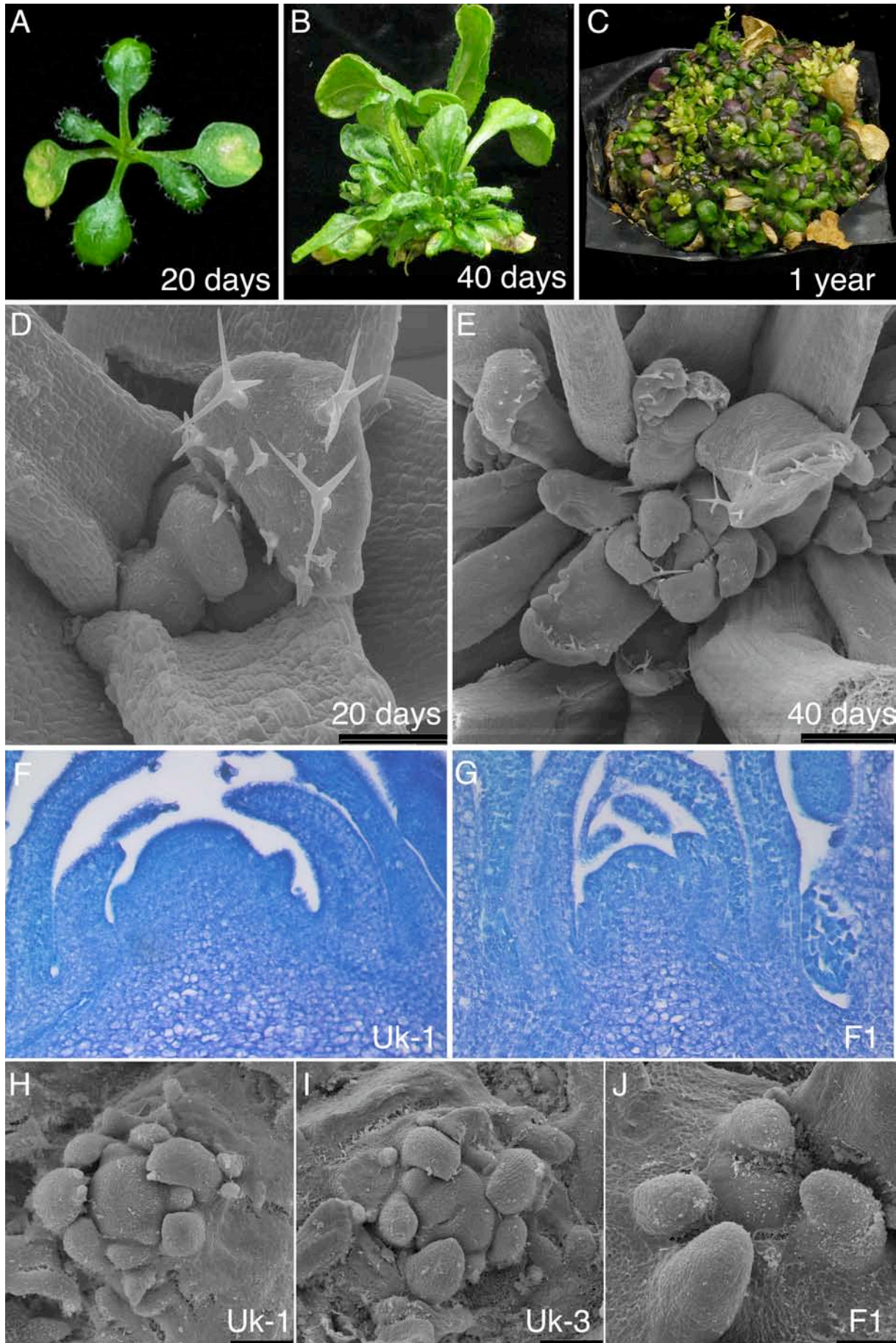
Figure 11: F1 hybrids grown at 16 °C are shown at different ages (A-E). 20 days old hybrids only have few small leaves (A & D). 40 days old plants (B & E) and 1 year old hybrids (C) are small and bushy. Scale bars in D & E represents: 500 µm. Longitudinal sections of Uk-1 (F) and F1 (G) apices grown at 16°C clearly show a size difference in the apical region in wild-type compared to hybrid plants (magnification: 20x). Scanning electron micrographs of Uk-1 (H), Uk-3 (I) and F1 hybrids (J) 8 days after a shift from 23 °C to 16 °C show the vegetative shoot apical region (scale bar: 100 µm). In F1 hybrids, the growth cone is flattened and meristematic growth is reduced after a temperature shift to 16°C.

The apical region is affected in growth and releases much less leaves per time than either of the parents (Figure 11 D-J). Being of same age, Uk-1 and Uk-3 can have 17 and 16 leaves, whereas hybrids only have 8 smaller leaves. The size of the shoot apex is reduced, although the overall structure is very similar to wild type (Figure 11 F-J).

Temperature sensitivity of the incompatibility phenotype

Interestingly, F1 hybrids only show the growth defective phenotype at 16°C. When grown at 23°C, F1 hybrids absolutely resemble wild type plants (Figure 12 A & B). The incompatibility phenotype can be induced by temperature shift from 23°C to 16°C. Already few days after shifting, a reduction in growth of new emerging leaves and leaves that are still expanding becomes obvious (Figure 12 C-D). Also the reverse shift from 16 to 23°C triggers a prompt change in phenotype (Figure 12 E & F).

Hybrids that were shifted to higher temperature also flower and reproduce. Although flowers of shifted hybrids do have some defects, they produce plenty of viable seeds (Figure 13).
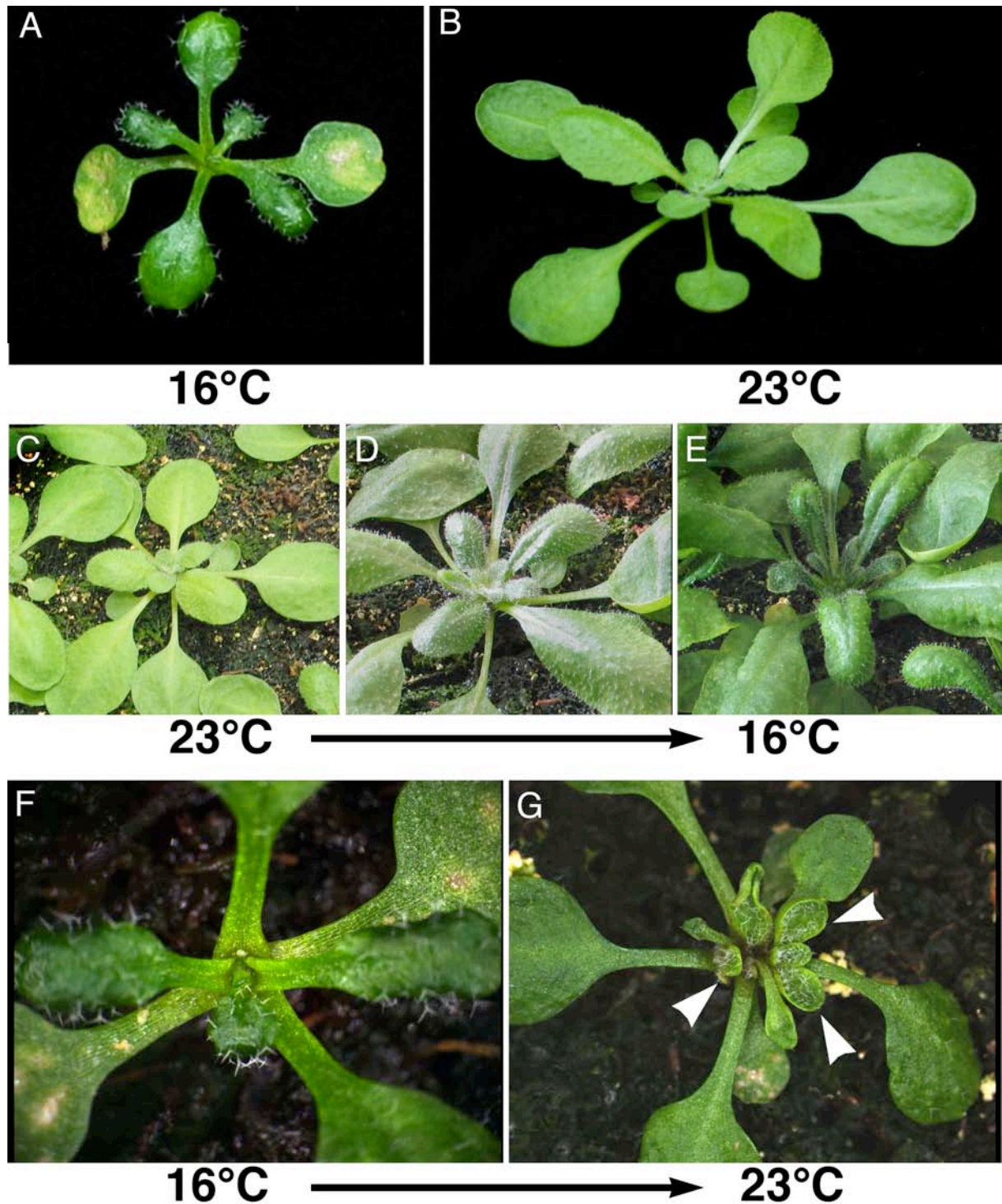
Figure 12: F1 hybrids are highly temperature sensitive and show a phenotype severely restricted in growth and development when grown at 16°C for 20 days (A), whereas they resemble wild type plants when grown at 23°C (B). When 15 days old F1 hybrids (C) are shifted from 23°C to 16°C, reduction of leaf growth is already visible after 5 days (D) and is obvious 11 days after a temperature shift (E). Shifting hybrids from 16°C to 23°C induces normal growth. In 20 days old F1 plants (F) wild type leaves (white arrowheads) emerge 6 days after shifting from 16°C to 23°C (G).
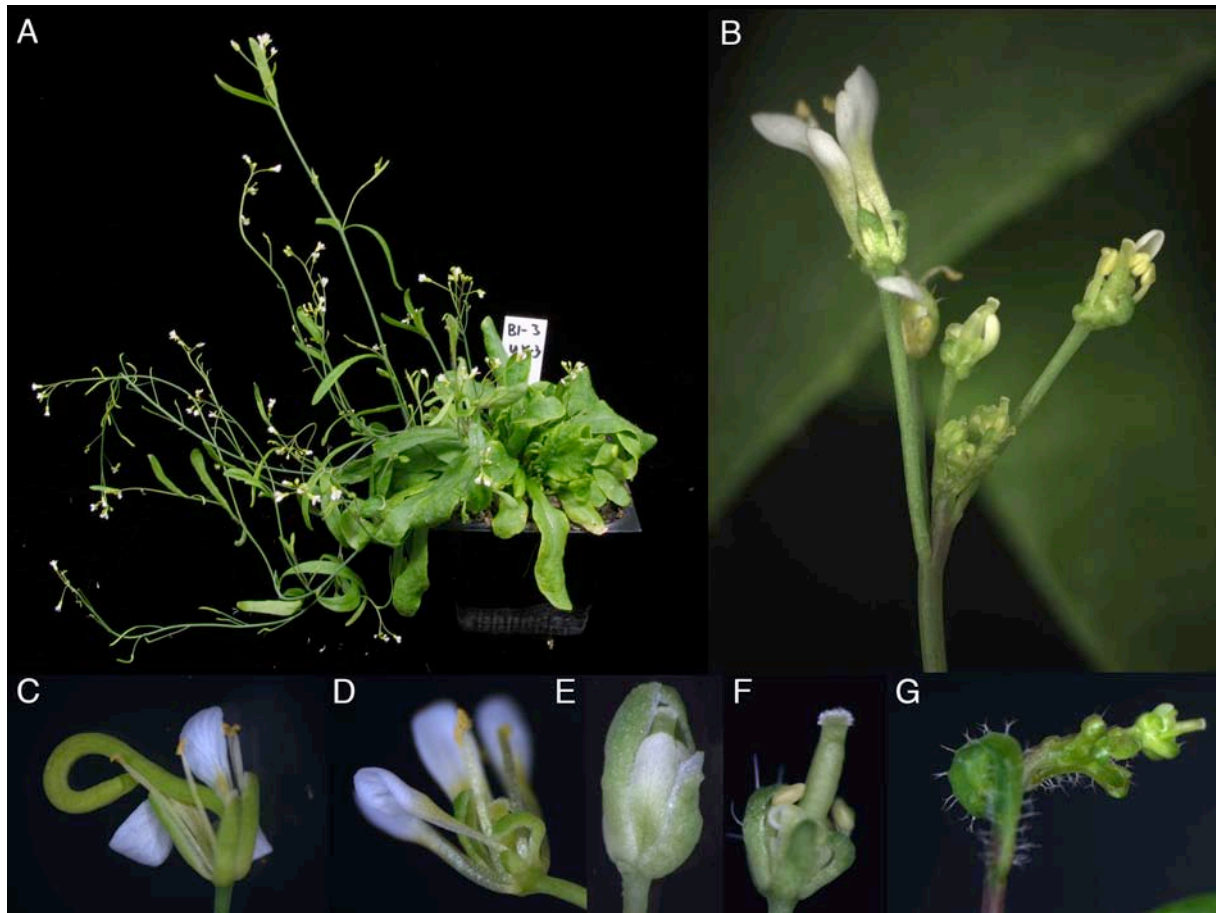
Figure 13: Plants with incompatibility phenotype are able to develop flowers after a temperature shift to 23°C. The habit (A) and inflorescences (B) show severe alterations in the phenotype compared to wild type plants. Defects are multifaceted (C-G); most frequent is a reduction in size of sepals, petals and stamens, which exposes the carpel prematurely (B, F, G).

2.2 Activated defense response is associated with the incompatibility phenotype

Microarray analysis uncovers an inappropriate pathogen response in F1 hybrids

The possibility to induce the incompatibility phenotype by temperature shift enabled us to investigate the molecular basis underlying the growth defective phenotype. We asked for differences in gene expression levels that accompany the induction of the incompatibility phenotype.

A microarray experiment with both parents and F1 hybrids was performed in which the phenotype was induced by a temperature shift from 23°C to 16°C. We obtained a gene list of 232 genes that change expression levels specifically in hybrids after temperature shift but

do not react on temperature in either of the parents, nor are they different between parents and hybrids at 23°C (rank product, p ≤ 0.05, for gene list see appendix 3.5) (Figure 14).
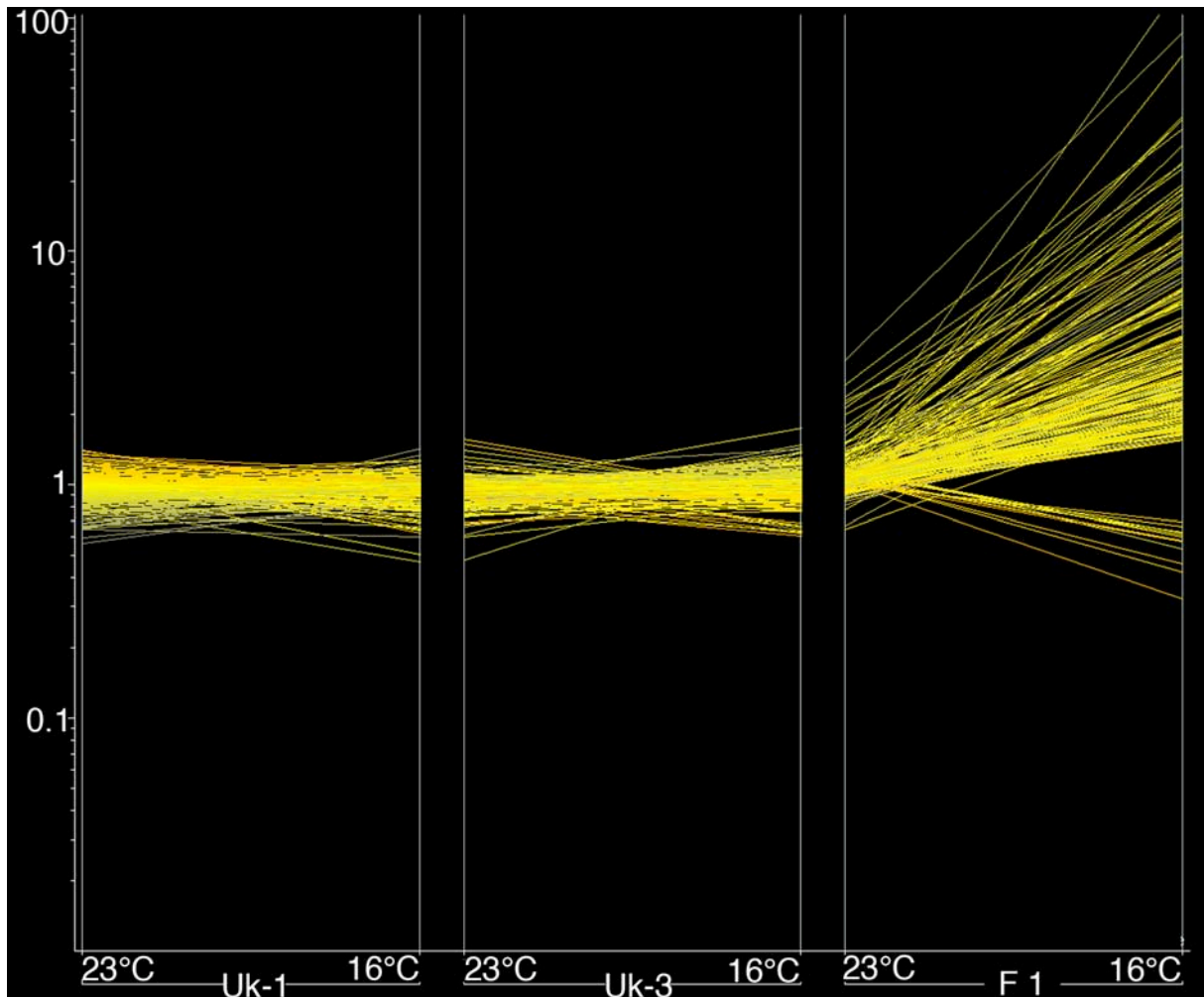
Figure 14: Expression profile of 232 genes change specifically upon temperature shift from 23 to 16°C in F1 hybrids. Normalized expression levels on logarithmic scale are plotted against temperature for each genotype Uk-1, Uk-3 and F1.

Conspicuously, the vast majority of the 232 identified genes are up regulated at 16 °C, only very few expression levels decrease upon temperature shift. Gene descriptions of the 232 genes depicted in Figure 14 clearly indicate an overrepresentation of genes involved in pathogen response. Analyses of gene ontologies using GeneMerge identify pathogen response genes as important factors involved in the F1 incompatibility (Table 4). 12 out of the 22 gene ontologies significantly represented in the 232 genes are involved in pathogen response. Even at 23°C, where hybrids do not exhibit any extraordinary morphology, we find elevated expression of several pathogen response genes in hybrids compared to both parents Uk-1 and Uk-3 (for gene list see appendix 3.4).

Table 4: *Gene ontologies that are overrepresented in the 232 genes that change expression levels in hybrids upon temperature shift frequently relate to the response of pathogens. Gene ontology descriptions derived from GeneMerge are ordered after the number of genes that fall into the corresponding ontology. The percentage of genes that fall into a gene ontology within the surveyed group of 232 genes is given as % of the 232 genes. The representation of the ontologies within the A. thaliana genome (22591 genes) is given as % of genome. Only ontologies that are significant after Bonferroni correction are listed (P-value). Gene ontologies that are obviously involved in pathogen response are marked in yellow.*

Gene Ontology Description	number of genes	% of the 232 genes	% of genome	P-value
physiological process	139	59,91	47,08	0,017
response to stimulus	42	18,10	7,56	3,1E-05
protein modification	37	15,95	6,83	0,0004
phosphorus metabolism	33	14,22	4,48	1,3E-06
phosphate metabolism	33	14,22	4,47	1,3E-06
phosphorylation	33	14,22	4,23	3,3E-07
protein amino acid phosphorylation	33	14,22	3,94	5,1E-08
response to biotic stimulus	31	13,36	2,96	7,3E-10
response to external stimulus	30	12,93	4,33	2,8E-05
response to stress	28	12,07	3,96	5,3E-05
defense response	23	9,91	1,93	5,0E-08
response to external biotic stimulus	21	9,05	1,80	4,9E-07
response to pest, pathogen or parasite	21	9,05	1,73	2,3E-07
response to pathogen	19	8,19	1,55	1,4E-06
defense response to pathogen	17	7,33	1,16	6,2E-07
defense response to pathogen, incompatible interaction	11	4,74	0,58	3,5E-05
response to wounding	8	3,45	0,42	0,0021
systemic acquired resistance	5	2,16	0,11	0,0015
cell wall catabolism	4	1,72	0,10	0,0212
salicylic acid mediated signaling pathway	4	1,72	0,08	0,0115
systemic acquired resistance, salicylic acid mediated signaling pathway	4	1,72	0,08	0,0072
salicylic acid biosynthesis	2	0,86	0,01	0,0324

In order to identify possible upstream activators, we considered 500 bp 5´ of the 232 genes as activated promoter region and looked for binding motifs that are overrepresented in this sample. The WRKY binding domain (T)TGAC is significantly overrepresented with 136 of 232 times (p < 0.001, Eulgem et al., 2000).

Induction of the salicylic acid pathway

Pathogen response pathways can be split into the salicylic acid, the jasmonic acid and the ethylene pathway (Glazebrook, 2001; Kunkel and Brooks, 2002). Key genes of the jasmonic acid and the ethylene pathway do not show differential expression in F1 hybrids compared to Uk-1 or Uk-3. However genes of the salicylic acid pathway are highly induced in F1 hybrids at permissive temperature (Figure 15). The important role of the salicylic acid pathway in F1 hybrids is also supported by overrepresentation of the gene ontologies " salicylic acid mediated signaling pathway", "systemic acquired resistance, salicylic acid mediated signaling pathway" and "salicylic acid biosynthesis" in Table 4.
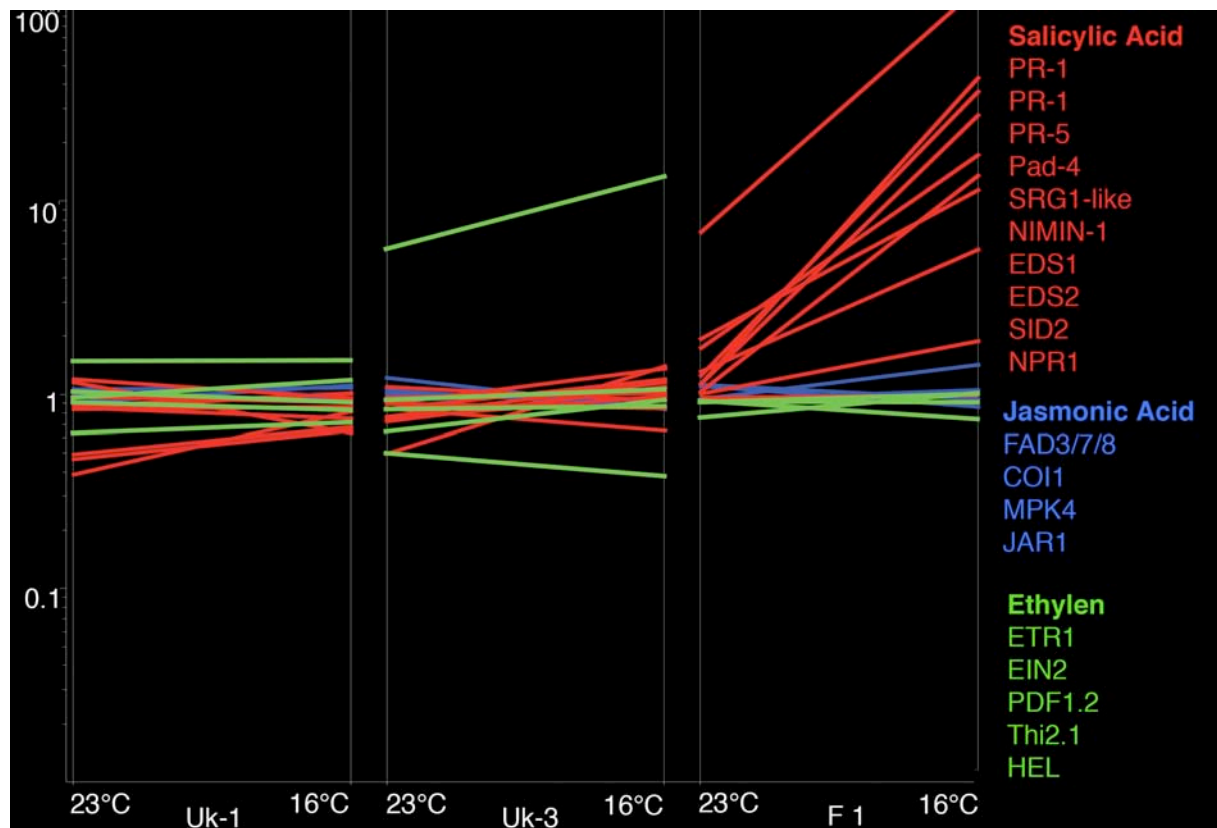
Figure 15: Expression levels of key genes of the salicylic acid, ethylene and jasmonic acid pathogen response pathways are shown. Key genes of the salicylic acid pathway change expression levels in F1 hybrids (red), whereas genes of the jasmonic acid (blue) and the ethylene pathway (green) are not induced in F1 hybrids. Normalized expression levels are plotted on logarithmic scale against genotypes. All represented genes are listed to the right.

Spatial activity of the pathogen response gene *PR-1*

PR-1, an important marker gene for pathogen response (Rogers and Ausubel, 1997), is upregulated in F1 hybrids at 16 °C. In order to visualize spatial expression of *PR-1* in these plants, we used transgenic lines containing a *PR1::*GUS reporter construct. These lines were crossed to plants with F1 phenotype. The progeny of the cross *PR1*::GUS x Col-backcross line segregates for the incompatibility phenotype in a ratio of about 3:1 (57:22). As expected there was no GUS staining observed in F1 plants exhibiting wild type phenotype. In contrast, strong staining was observed in F1 plants with incompatibility phenotype, mostly in the apical region and in still growing leaves (Figure 16).
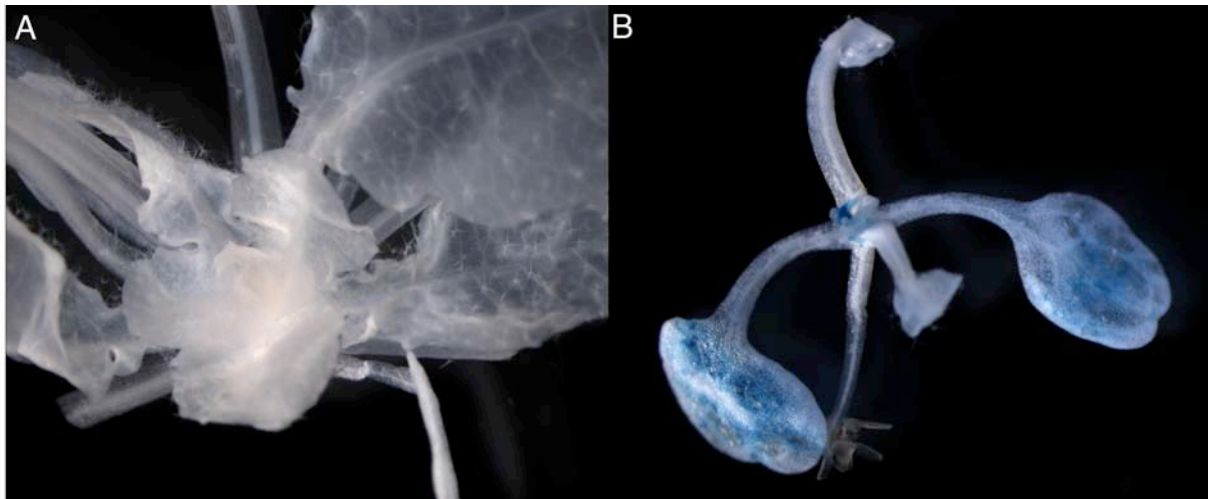
Figure 16: Spatial activity of PR1::GUS in plants with (A) or without (B) incompatibility phenotype. Blue stain in the apical region and in leaves indicates PR1 activity.

2.3 Genetic analysis suggest that 2 loci underlie the observed incompatibility

Segregation pattern

In the F2 population of the cross Uk-1 x Uk-3, there are three phenotypes: wild type (wt), F1 phenotype (f1) and a lethal phenotype (Figure 17). The different phenotypes occur in a ratio of about 7:4:5 (128 wt: 77 f1: 107 lethal).
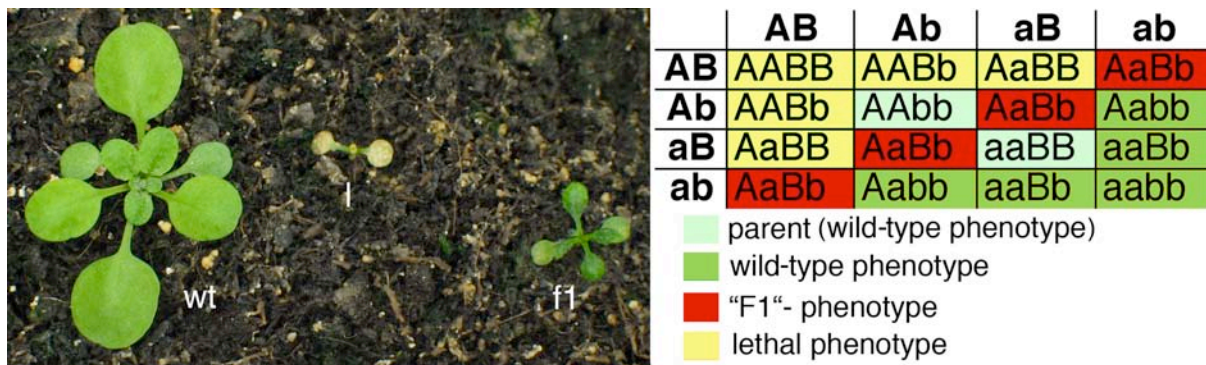
Figure 17: Observed phenotypes in the F2 population of the cross Uk-1 x Uk-3. Wild type (wt), F1-phenotype (f1) and lethal plants (l) occur in the ratio of 7:4:5. The crossing scheme with two genes with two alleles each can explain the observed segregation. Capital letters indicate derived problematic alleles.

A simple two-gene-model can explain the segregation ratio of 7:4:5 (Figure 17). In this model, each parent possesses two wild type alleles of one gene and two derived alleles of a second gene. Hybrids with the F1 phenotype (f1) would necessarily be heterozygous for both genes, as in F1 plants. Lethal plants would have at least three copies of a derived problematic allele, assuming the derived alleles produce the phenotype in a quantitative manner. Two derived alleles do not have problematic effects when homozygous (as in parents), however the epistatic interaction of derived alleles of two different genes is detrimental.

To support this model, F1 plants were also backcrossed to Col-0, Uk-1 and Uk-3. The observed phenotypic segregation patterns were also close to the ratios that are expected by the mentioned two-gene-segregation model: 3:1 in the backcross to Col-0 (30 wt:15 f1) and 2:1:1 in backcrosses to Uk-1 and Uk-3 (wt:f1:lethal, 33:10:10 and 25:19:7).

53

Mapping incompatibility regions

To verify the above-mentioned segregation model, mapping experiments were performed.

High mapping resolution was obtained with 64 SNP markers tested on a set of 181 F2 and Uk-1/-3 backcross plants (appendix 3.1). In agreement with the above-mentioned segregation model, mapping results revealed 2 loci responsible for the incompatibility phenotype. It is expected that the two causative regions are both in a heterozygous state in plants with incompatibility phenotype. In these plants, the SNP mapping results show a high percentage of heterozygosity within two chromosomal regions on chromosome 3 and 5 (Figure 18 A). QTL analysis of the SNP mapping data using the R package qtl with a single QTL model (Broman et al., 2003) clearly identifies two significant QTL in the same two regions (Figure 18 B).
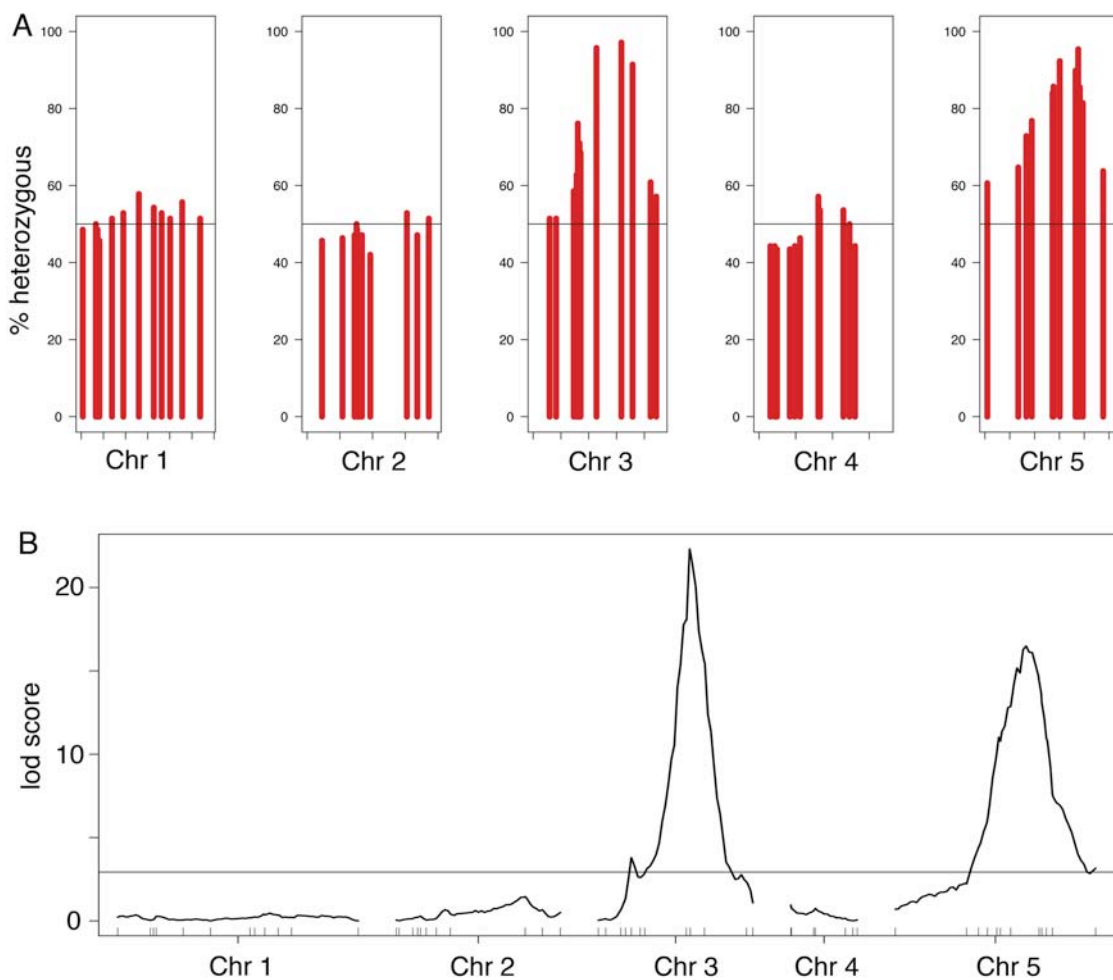
Figure 18: Mapping results of 64 SNP markers tested on 181 F2 plants. (A) Percentages of heterozygous markers in plants with F1 phenotype are plotted against chromosomal positions of SNP markers. The percentage of heterozygosity is higher than 50% (indicated by a horizontal line) at two chromosomal regions on chromosome 3 and 5. (B) QTL analysis identifies the same two loci as major effect QTL with high LOD scores. Horizontal line represents LOD score threshold after 1000 permutations.

In order to test for an epistatic interaction, a two-dimensional genome scan was performed (Broman et al., 2003). Therefore, the above-mentioned mapping results were used with genotypic and position information of each marker for each individual. There is highly significant epistatic interaction of the two genetic loci as shown in Figure 19.
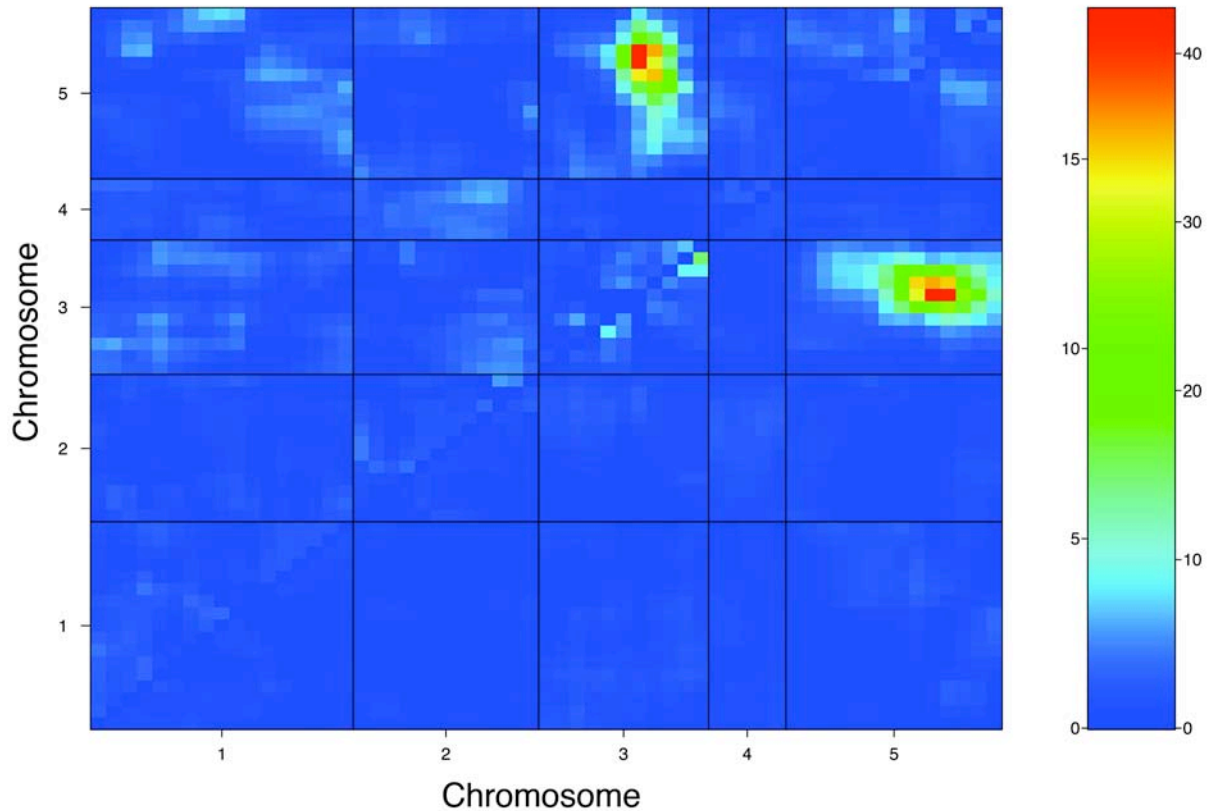
Figure 19: Two-dimensional genome scan. Upper-left triangle shows significance levels for epistatic interaction for every possible marker combination, the lower-right triangle for joint LOD-score. There is evidence for significant epistatic interaction of chromosome 3 and 5 regions (red).

Since plants with wild type phenotype were also included in the mapping sample, we can infer from the obtained results which of the parental alleles are the problematic ones. Of the 89 wild type plants included in the mapping population, 22 are heterozygous at the chromosome 3 locus and homozygous for Uk-1 at chromosome 5 region. Similarly, 20 wild type plants that are heterozygous at chromosome 5 locus are homozygous for the Uk-3 allele at chromosome 3. Following the presented scheme (Figure 17), we can conclude that the problematic derived allele on chromosome 3 originates from Uk-1 and the problematic allele on chromosome 5 from Uk-3.

F1 hybrids were also be backcrossed to Col-0 and the incompatibility regions were obtained in an nearly Col-0 background. The advantage of these lines for mapping is that more SNP markers are available for Col-0 versus Uk-1/Uk-3 than for Uk-1 versus Uk-3. Additional 24 markers concentrated in the incompatibility regions were selected for these Col-0 backcross lines (appendix 3.2). They were tested on 96 individuals of Col-0 backcross lines and both mapping intervals on chromosome 3 and 5 could be reduced to 0.68 Mbp and 0.91 Mbp respectively (Figure 20).
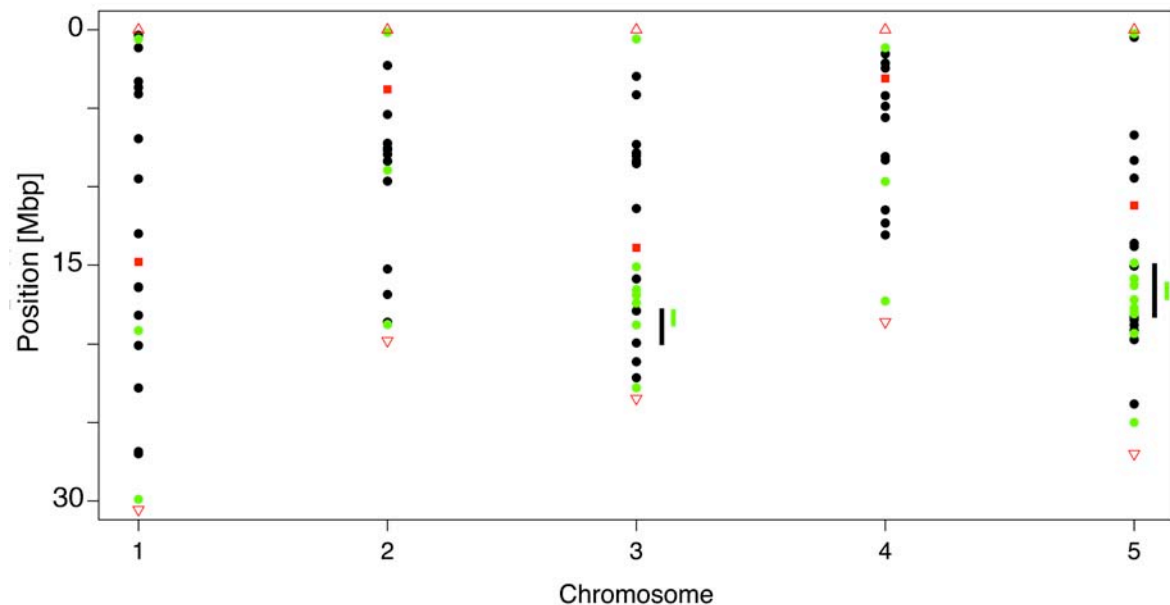
Figure 20: Chromosomal position of SNP markers used to distinguish Uk-1 and Uk-3 (black) as well as polymorph SNPs between Uk-1/Uk-3 and Col-0 (green). Red triangles indicate ends of chromosomes, red squares the position of centromeres. The mapping intervals are represented as black (from F2 and Uk backcross lines) and green bars (from Col-0 backcross lines).

Fine mapping

Further fine mapping was carried on for the mapping interval on chromosome 3.

For fine mapping, an F2 population of 864 individuals with incompatibility phenotype and 96 plants with wild type phenotype were screened with two PCR-based markers, an SSLP and a CAPS marker that were 100.2 kb apart (see methods 3.5). Only 3 out of 864 plants could be identified that contain a recombination event within the mapping interval. The recombination breakpoints were fine mapped by sequencing PCR fragments localized within the region of

interest. By doing so, the mapping interval was confined to a genomic region that corresponds to 27 kb in the Col-0 accession.

Nature of the Uk-1 sequence in the mapping interval

Since there is a chance of major genome rearrangements in strains different from Col-0, the sequence of Uk-1 within the 27 kb mapping interval was determined. To do so, a fosmid library of Uk-1 was constructed and screened for the sequence interval of interest. Two clones that together cover the mapping interval were identified and shotgun sequenced. The alignment of Uk-1 and Col-0 sequences did not show major genome rearrangements within the 27 kb interval (Figure 21).
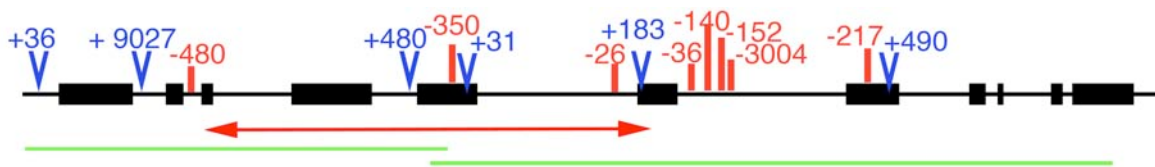
Figure 21: Schematic of Uk-1 sequence covering the mapping interval (red arrow) obtained from two fosmid clones (green lines). Genes (black boxes) and intergenic regions (black lines) are adopted from the Col-0 sequence. Blue and red arrows represent insertions and deletions present in Uk-1, numbers indicate the size of insertions/deletions. Only insertions and deletions bigger than 20 bp are shown.

There are two genes present in the mapping interval: at3g44530 and at3g44540. At3g44530 encodes a WD-repeat containing protein involved in negative regulation of knox gene expression, at3g44540 codes for an oxidoreductase, similar to an acyl-CoA reductase (www.arabidopsis.org). A sequence comparison of the Uk-1 with Col-0 revealed only 18 SNPs in at3g44530 of which 7 are in the coding sequence and only 3 cause non-synonymous changes (Figure 22 A). There are no insertions or deletions. By contrast, 97 SNPs are present in at3g44540, of which 14 are silent changes within the coding region and 12 lead to amino acid changes (Figure 22 B). Additionally, there are 8 deletions and 11 insertions smaller than 20 bp and two insertions and one deletion bigger than 20 bp.
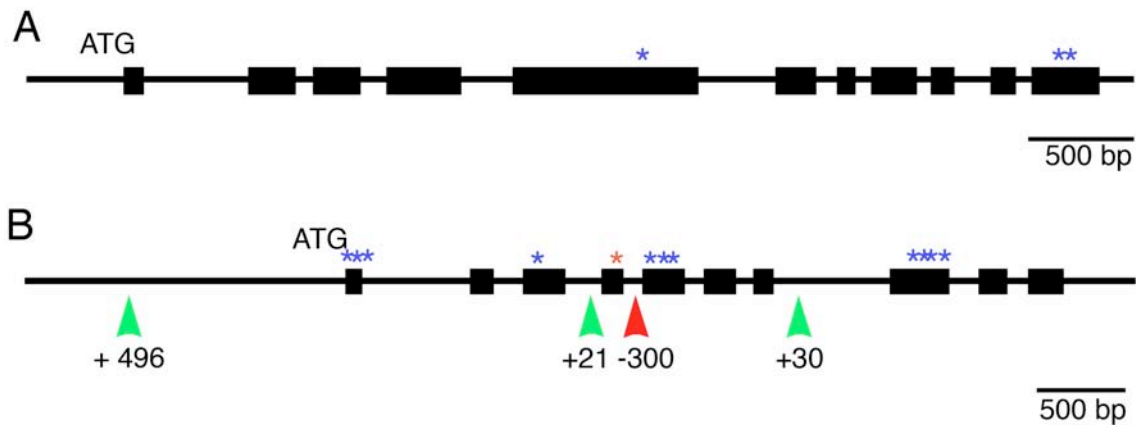
Figure 22: Schematic of Uk-1 alleles of two candidate genes at3g44530 (A) and at3g44540 (B). Boxes and lines represent exons and introns, respectively. Non-synonymous SNPs (blue stars) as well as insertions (green arrowhead) and deletions (red arrowhead) with respect to Col-0 are present in both genes. The numbers below the arrowhead indicate the size of the insertion/deletion in bp. Non-synonymous SNPs and insertions/deletions in non-coding regions that are smaller than 20 bp are not shown.

2.4 Testing two candidate genes

Three different approaches were used to identify the locus on chromosome 3 that is responsible for the incompatibility between Uk-1 and Uk-3: (I) gene knock-down by artificial micro RNAs (amiRNAs) (Schwab et al., 2006), which would identify genes required for the incompatibility, (II) genomic complementation and (III) inducible overexpression, which both might indicate whether a gene is sufficient for the incompatibility.

Gene knock-down of two candidate genes

The two genes at3g44530 and at3g44540 were knocked down by constitutive expression of amiRNAs designed to target specifically the two candidate genes. Both amiRNA constructs were transferred into Uk-1 and Uk-3 plants. The T1 generation was crossed to the reciprocal parent. We expected a reduction of the incompatibility phenotype in the F1 (T1$_{Uk-3}$ x Uk-1 or T1$_{Uk-1}$ xUk-3) in case of suppression of the causative gene by amiRNA. Since there is extensive variation in the strength of lines in the T1 and T2, a total of 62 T1s were crossed to the reciprocal parent and the F1s were surveyed at 16°C.

A reduction of the incompatibility phenotype seemed to appear in lines containing the amiRNA against at3g44530, but not in lines with at3g44540 knocked down (Figure 23 A & C). However, the normal phenotype abated with plant age (Figure 23 B).
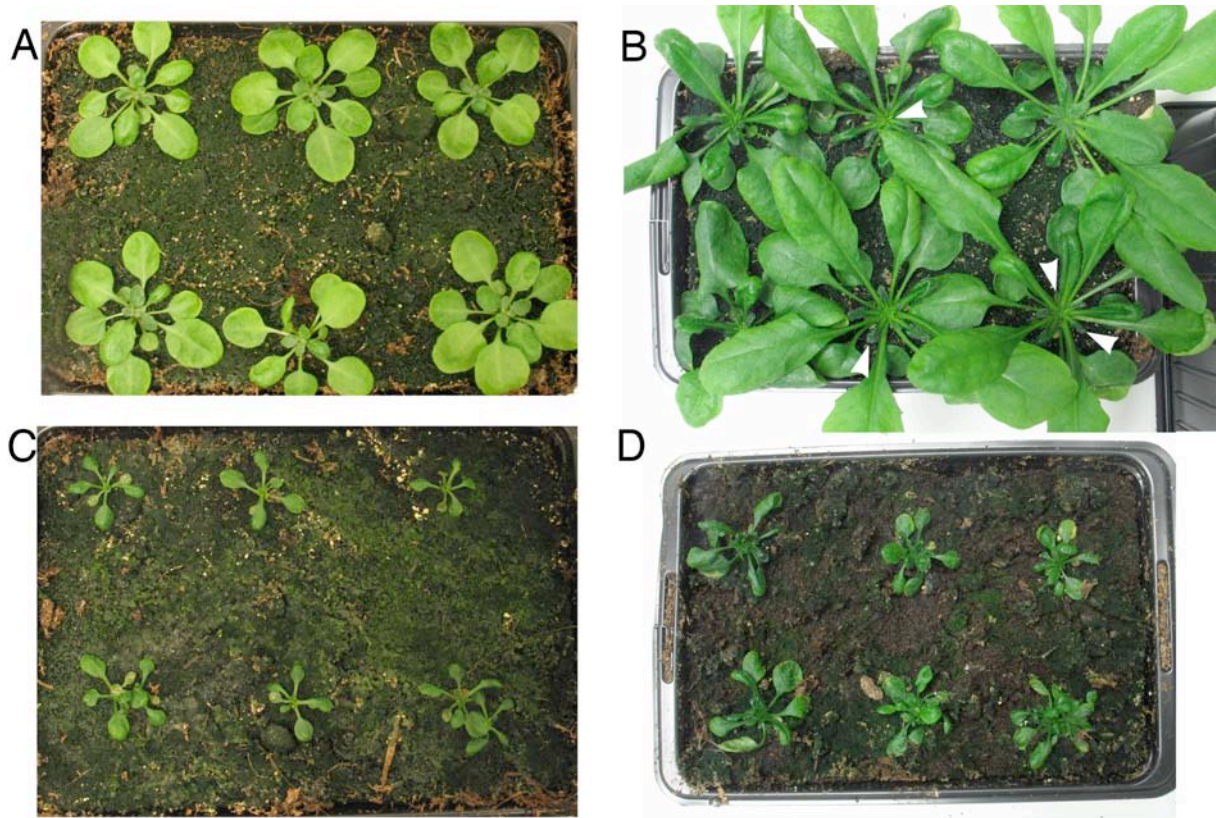
Figure 23: F1 plants (T1$_{Uk-3}$ x Uk-1 or T1$_{Uk-1}$ xUk-3) containing an amiRNA against at3g44530 exhibit a wild type phenotype (A & B), whereas the amiRNA against at3g44540 does not affect the Uk-1/Uk-3 incompatibility phenotype (C & D). The effect of the amiRNA against at3g44530 however is much stronger in 20 days old plants (A) compared to plants of 36 days age that develop incompatibility symptoms in emerging leaves (arrowhead) (B). The incompatibility phenotype is not altered by age in plants with knocked down at3g44540 (C & D).

Genomic complementation

We aimed for producing "hybrid-like" transgenic plants that should show the Uk-1/Uk-3 hybrid phenotype when the causative genomic fragment was introduced by transformation. In order to obtain a phenotype, chromosome 3 and 5 regions need to be in a heterozygous state. In the produced transgenic lines, heterozygosity for chromosome 5 region was obtained by using Col-0 backcross lines that are heterozygous in this region. For chromosome 3 region, an artificial heterozygosity was obtained by introducing a genomic fragment containing a Uk-

59

1 allele of the causative region into Col-0 backcross lines that possess two Uk-3 alleles of the chromosome 3 interval. We expected to obtain transgenic lines with a phenotype resembling at least a weak Uk-1/Uk-3 hybrid when grown at 16°C.

To do so, three genomic fragments containing (a) promoter and gene at3g44530, (b) promoter and gene at3g44540 and (c) the intergenic region between genes at3g44540 and at3g44550 (Figure 24) were transferred into Col-0 backcross lines and tested for their phenotypic effects. The Col-0 backcross line 6-2 was selected for transformation, because it is mostly of Col-0 genotype throughout the genome, except for the incompatibility region on chromosome 5 (see appendix 3.2). Only plants were transformed that were heterozygous at chromosome 5 region (Col-0/Uk-3).

Unfortunately, only few transformants were obtained for all constructs and none of them showed a phenotype different from wild type (data not shown).

Figure 24: The Uk-1 sequence covering the mapping interval (red double arrow) was spilt into three fragments containing gene with endogenous promoter (green and blue line) or an intergenic fragment (turquoise line). All three fragments were transformed into plants. Genes and intergenic regions are represented as black boxes and lines, respectively.

Inducible expression of at3g44530

The induction of the Uk-1 allele of at3g44530 confirmed that at3g44530 can alter the normal phenotype (Figure 25). The same Col-0 backcross lines used for genomic complementation were transformed with an ethanol inducible construct of the Uk-1 allele of at3g44530. If at3g44530 is causative for the incompatibility between Uk-1 and Uk-3, T1 plants that are heterozygous for chromosome 5 locus express the Uk-1 allele of at3g44530 should be similar to Uk-1/Uk-3 hybrids in their phenotype. T1 lines with strongly induced at3g44530 expression (measured by q-RT-PCR, data not shown) showed a strong aberrant phenotype (Figure 25). After one week of induction, the plants showed a phenotype similar to Uk-1/Uk-3 hybrids with slightly reduced growth and darker leaves (*Figure 25* B & C). After two weeks of induction, leaves showed a sinuate leaf shape (*Figure 25* D & E).

Unfortunately, the occurrence of the phenotype was not linked to the genotype of the chromosome 5 region. A strong phenotype was observed in both, plants that were homozygous for Uk-1 (Figure 25 B & D) as well as heterozygous at the chromosome 5 incompatibility region (Figure 25 C & E). The phenotype of induced T1 plants appears to be different from the Uk-1/Uk-3 incompatibility. In older plants the leaves get slightly dentate (Figure 25 E, compare Figure 11). Further experiments are required that assure that the Uk-1 allele is causative and that induction of a Col-0 allele would not matter to the phenotype.
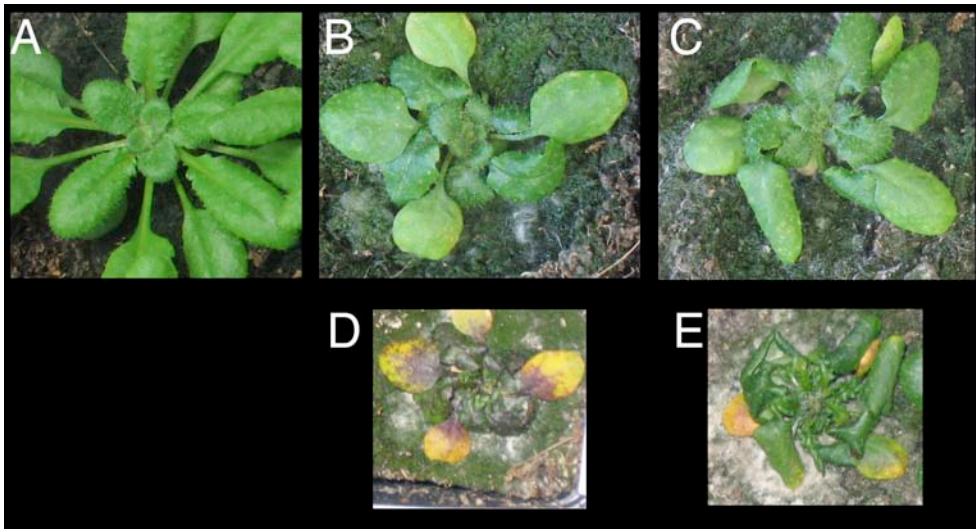
Figure 25: In contrast to non-induced plants (A), leaves start to deform when expression of the Uk-1 allele of at3g44530 is induced (B-E). However, the genotype on chromosome 5 has no phenotypic effect. T1 plants that possess either a Uk-1 allele (B & D) or are heterozygous (C & D) show the same phenotype. Leaves only start to deform a week after induction and were strongly deformed two weeks later. Pictures D & E were adopted from Kirsten Bomblies.

Search for unique sequence changes in candidate genes

Since the Uk-1/Uk-3 incompatibility is highly specific, we screened several *A. thaliana* accessions for sequence variation that is unique to Uk-1. For at3g55430, we sequenced exon 5 of 55 and exon 11 of 78 accessions. These two exons contain all three amino acid changes that are present in the Uk-1 allele with respect to Col-0. Unexpectedly, the exact same SNPs causing the 3 amino acid changes are also present in two other accessions, Ak-1 and Sav-0 (Table 5). The first SNP in exon 11 is also present in Aa-0, Pr-0 and St-0, the second SNP in Aa-0 (Table 5).

Ak-1 is among the accessions that were crossed to Uk-1 and Uk-3 (see below). Although they contain the same amino acid changes as Uk-1 in at3g44530, both F1s do not exhibit the incompatibility phenotype.

Table 5: *Nucleotide changes in at3g44530 leading to amino acid changes (yellow) in Uk-1 are also present in other accessions.*

Position (bp)	Exon5 2377	Exon 11 5016	5028	5032
consensus	G	C	G	T
Uk-1	A	A	A	G
Ak-1	A	A	A	G
Sav-0	A	A	A	G
Aa-0		A	A	
Pr-0		A		
St-0		A		

Additionally, we genotyped for a 20 bp deletion that is 1.2 kb upstream of the at3g44530 start codon. This deletion found in Uk-1 was found in 7 more accessions: Ak-1, Ang-1, Bd-0, Edi-0, Fr-6, Fr-7, Sav-0.

For at3g44540, we chose to sequence exon 4 since this exon contains an amino acid change that is not present in 20 other accessions for which we had sequence information (Clark et al., 2007). This amino acid change, from glutamine to leucine at position 1279 was not present in any of the other sequenced 66 accessions.

2.5 Other aspects of the Uk-1/Uk-3 incompatibility system

Genetic distance between Uk-1 and Uk-3

We obtained SNP marker information for 289 SNPs tested on 86 accessions, including Uk-1 and Uk-3. In order to find out about the genetic distance between Uk-1 and Uk-3, we asked if the number of polymorphic SNPs between Uk-1 and Uk-3 is particularly high. The average number of polymorphic SNPs in the given set of accessions and markers is 71.6 and 67 SNPs are polymorph between Uk-1 and Uk-3 (Figure 26). Uk-1 and Uk-3 are neither unusually similar in nor are they noticeable different in their sequence compared to other accessions.
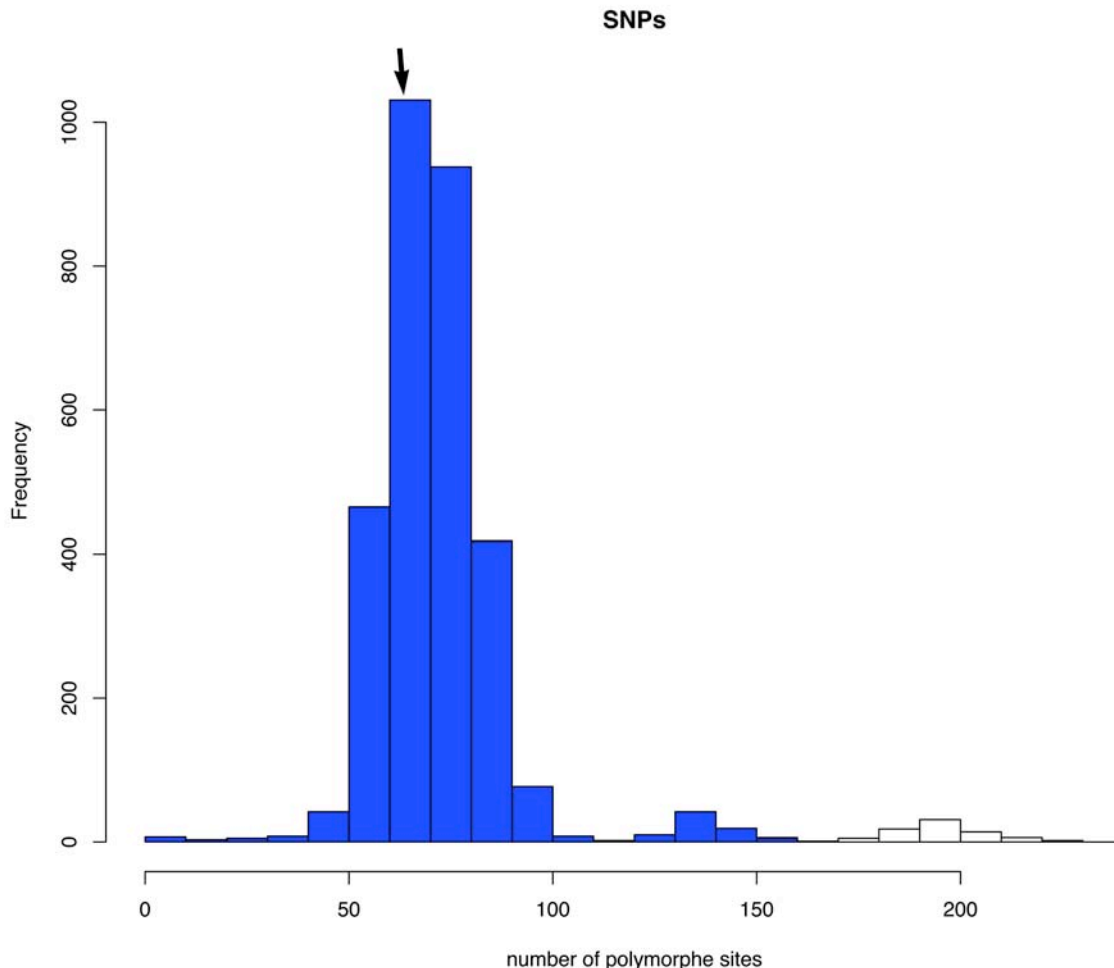
Figure 26: Histogram of number of polymorphic SNPs between each possible marker combination in a set of 86 accessions and 289 markers tested (in blue). The differences between Uk-1 and Uk-3 are about average (arrow). Col-0 was treated separately (white histogram) since a high number of SNPs is expected due to the design of the SNPs.

The Uk-1/Uk-3 incompatibility is highly specific

The incompatibility between Uk-1 and Uk-3 is rather specific to both accessions. 27 further accessions were crossed to both Uk-1 and Uk-3 (for list see methods 3.8). None of the obtained F1 plants showed any incompatibility phenotype at 16°C.

Relevant temperature resembles natural growth conditions

The standard laboratory growth conditions of *A. thaliana* are between 20-23°C. Since we can only observe the incompatibility phenotype at 16°C and not at 23°C, we were wondering if our findings would matter under natural conditions. The average temperature at Freiburg, which is close to Umkirch, the village where Uk-1 and Uk-3 were collected in 1954, is only above 16°C from mid may to mid September (Figure 27). Considering the fact that *A. thaliana* is already completing its life cycle and setting seeds in late May, 16°C much better resembles natural growth temperature than 23°C.

The relevance of the observed incompatibility under natural growth conditions also becomes evident when seeds were grown outside in Tübingen, which has a very similar temperature profile as Freiburg (see Figure 27 A & B).
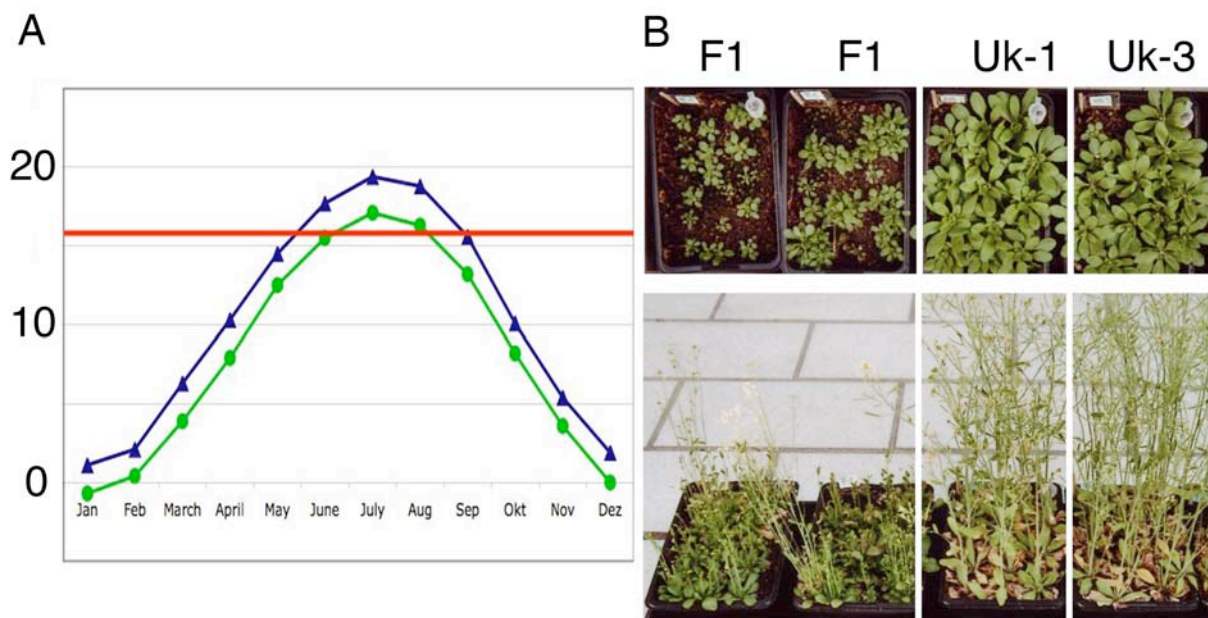
Figure 27: (A) Average monthly temperatures at Freiburg (close to Umkirch) and Tübingen. 16°C is marked by a horizontal red line. (B) F1 hybrids show noticeably reduced size and seed production when grown outside in Tübingen, compared to both parents, Uk-1 and Uk-3. Pictures were taken on June 8th (upper row) and June 24th (lower row).

Existence of incompatibility alleles in wild

The Uk-1/Uk-3 incompatibility represents a potential case of incipient speciation. However, we do not know whether the diverging populations will result in separate species any time in future. In order to place the Uk-1/Uk-3 incompatibility correctly in the evolutionary process of speciation, it is important to be able to investigate dynamics of the diverging populations. Long-term population dynamics of both incompatibility alleles would help to understand the role of this incompatibility for speciation. To do so, the first required step is to find the incompatibility alleles in wild.

In 2005, Kirsten Bomblies and I went to Umkirch, the original collection site of Uk-1 and Uk-3 in 1954, in order to sample the local *A. thaliana* population. We could find and collect 16 different *A. thaliana* accessions at Umkirch. *A. thaliana* grew mostly at ruderal sites such as dry sandy areas next to a soccer field, open grass areas, a parking lot, a side walk and as weed in a flower bed (for pictures see appendix 3.8). The different lines exhibit large variation in phenotype and flowering time (appendix 3.9).

After crossing all lines to Uk-1 and Uk-3, we could observe the Uk-1/Uk-3 incompatibility phenotype in one cross: Umkirch 2-1 x Uk-1. This suggests that the incompatibility-causing gene present in Uk-3 is also present in the Umkirch accession 2-1 which has only been under laboratory conditions for 1 generation.

IV. DISCUSSION

1 Flowering time variation in *Arabidopsis thaliana*

1.1 Flowering time variation under different environmental conditions

In this study, flowering time variation was surveyed among 155 accessions of *A. thaliana* under four different environmental conditions. Under LD conditions the distribution of flowering time was bimodal. This can be attributed to the activity of *FRI* and *FLC,* which leads to early or late flowering behavior, noted in several previous studies (Koornneef et al., 2004; Sung and Amasino, 2004a). After vernalization treatment, the long tail of late flowering accessions vanished, however extensive variation reaching from TLN 9 to TLN 42 remained. Under SD conditions, the flowering time distribution was much broader and only consisted of a single peak.

Extensive variation was found in flowering time responses to the environmental cues day-length, ambient temperature and vernalization. Nonetheless, the overall distribution of flowering times in the four different environmental conditions reflected the action of known flowering time pathways: prolonged photoperiod, vernalization and elevated ambient temperature promote early flowering. Several accessions with conspicuous behavior could be identified. They represent an important resource for further investigations aiming to map imprints of adaptation in known or unknown flowering time genes.

Considering environmental sensitivities as well as hierarchical clustering (not presented, see Lempe et al., 2005), Fr-2 attracted our attention since it showed a very weak photoperiodic response. In a subsequent study, an allele of *PHYC* with a premature stop codon could be identified as cause for the reduced day length sensitivity of the accession Fr-2 (Balasubramanian et al., 2006b).

1.2 Contribution of *FRI/FLC* to flowering time variation

FRI and *FLC* are two genes with major effect on flowering time. That sequence variation plays an important role in selection could be shown for *FRI* (Le Corre, 2005; Toomajian et al., 2006). Haplotype comparisons demonstrated that selection had acted on *FRI* favoring non-functional *FRI* alleles and therefore an early flowering phenotype (Toomajian et al., 2006). The widely spaced geographic distribution of accessions with non-functional *FRI* alleles suggests that non-functional alleles evolved multiple times (Le Corre et al., 2002; Le Corre, 2005; Shindo et al., 2005).

Because of the evolutionary importance and the large effect of *FRI* and *FLC* on flowering time, we determined the functional status of both genes by genotyping, sequencing and complementation analysis in all accessions. The two frequent and well known Col-0- and L*er* –type deletions were found in 43% of all surveyed accessions. All early flowering accessions lacking these *FRI* deletions harbor non-functional alleles of either *FRI* or *FLC*. We could identify six new, non-functional alleles of *FRI* that are present in 12 different accessions. Diverse non-functional alleles of *FRI* are known, but only two of the alleles observed in our study, the *FRI* allele of HR5 and Kin-0, are part of the concurrent study of Shindo et al. (Shindo et al., 2005). Interestingly, the non-functional *FRI* allele of Jl-1 exhibits extensive nucleotide variation without a premature stop codon. The level of polymorphism in the *FRI* allele of Jl-1 is much higher compared to other alleles. One could speculate that this *FRI* allele accumulates sequence polymorphisms and is hence evolving to become a pseudogene.

Apart from non-functional *FRI* alleles, we found impaired *FLC* function in early flowering accessions. Few accessions have been identified previously with reduced *FLC* function and all of them but one have transposon insertions in the first intron (Michaels and Amasino, 1999; Gazzani et al., 2003; Caicedo et al., 2004). The first intron is an important site for regulation of transcript levels via histone methylation (Bastow et al., 2004; Sung and Amasino, 2005). Only one *FLC* allele has been described previously with premature stop codon (Werner et al., 2005a). We found reduced *FLC* function to be caused by large insertions within first intron in five accessions and altered transcripts in three accessions. The altered transcripts are likely to be null alleles, considering known null alleles of other MADS box genes (Jack et al., 1992; Mandel et al., 1992). Like *FRI*, *FLC* has also been proposed to be a target of evolution for early flowering accessions, although attenuated or null alleles are much less frequent than in *FRI* (Michaels et al., 2003). Shindo et al. could show that variation

in vernalization response and stable repression of *FLC* is correlated with the rate of accumulation of histone methylation in *FLC* itself (Shindo et al., 2006). *FLC* alleles are geographically dispersed, suggesting that *FLC* alleles evolved multiple times, similar to *FRI*.

The functional status of *FRI/FLC* caused flowering behavior to split accessions into early and late flowering groups, leading to a bimodal distribution. Within both groups however, there is still extensive variation. This suggests that in addition to *FRI* and *FLC*, there is substantial naturally occurring variation in other genes of the complex signaling network regulating flowering time. 63% of flowering time variation was explained by *FRI/FLC* under LD conditions and 23% in SD, indicating a particularly important effect of *FRI/FLC* in LD. This is consistent with the notion that the antagonistic effect of *FRI/FLC* on the LD induced flowering pathway is less important in SD, when flowering is not induced. A slightly higher correlation (70%) was estimated in LD by Shindo et al. by only considering sequence information of *FRI* and not *FLC* (Shindo et al., 2005). Other studies did not find a significant effect of *FRI* under SD conditions, which might be due to a sampling bias (Caicedo et al., 2004; Stinchcombe et al., 2004).

1.3 Correlation of *FLC* expression levels with flowering time

It is well known that *FLC* expression levels are high in late flowering and low in early flowering accessions (Michaels and Amasino, 1999; Sheldon et al., 1999; Sheldon et al., 2000; Gazzani et al., 2003; Michaels et al., 2003). However, it is not known how *FLC* levels explain variation within the two classes of accessions. We found a high linear correlation of *FLC* expression with flowering time under LD conditions in all accessions (r^2=0.42). The correlation of vernalization response with *FLC* levels showed an even stronger correlation (r^2=0.47).

Shindo et al. found expression levels of *FLC* being less correlated with flowering time (r^2=0.33) and vernalization response (r^2=0.38) than in our study, which might be due to differences in methodology or sampling differences. QRT-PCR with two technical replicates for each of two biological replicates were used in this study, northern plots without replicates were used by Shindo et al. (Shindo et al., 2005). In both studies, however, there was considerable variation of *FLC* levels in early flowering and late flowering accessions.

We observed a stronger correlation of flowering time with *FRI/FLC* functional status (r^2=0.63) compared to *FLC* levels (r^2=0.42). This could be explained by *FRI* affecting other

pathways controlling flowering time. However no other loci but *FLC* are known to be controlled by *FRI* (Schmid et al., 2003) (our own microarray data, not shown). It is also possible that other flowering time pathways are genetically linked to *FRI*. This hypothesis is in line with the finding that there is linkage disequilibrium between *FRI* and *FLC* alleles (Caicedo et al., 2004).

The dominant role of *FLC* in flowering time control was confirmed by a microarray experiment approving that *FLC* levels are most highly correlated with flowering time, compared to all 24,000 genes represented on the ATH1 array. It is essential to also consider the genomic context of *FLC* expression levels since observed patterns of polymorphisms in *A. thaliana* deviate from standard neutral models of evolution (Nordborg et al., 2005).

1.4 Latitudinal cline and flowering time

In the light of considerable phenotypic variation, it would be highly interesting to proof that an adaptive trait is under selection. Evidence for a trait, such as flowering time, to be under selection can be covariation with environmental factors. An obvious environmental factor to test for covariation with flowering time is latitude, since latitudinal clines are not uncommon for various phenotypes (Endler, 1977).

For *A. thaliana,* latitudinal clines have been shown for growth rate, light sensitivity, germination and establishment as well as freezing tolerance (Maloof et al., 2001; Stenoien et al., 2002; Hannah et al., 2006; Schmuths et al., 2006). Germination and establishment was only significant at 14°C but not at 22°C (Schmuths et al., 2006). A strong correlation between flowering time and latitude has been demonstrated in accessions that over-wintered in a common garden in Rhode Island (Stinchcombe et al., 2004). The correlation with latitude and flowering time was particularly high in accessions with functional *FRI* allele (r^2=0.38).

We found the strongest correlation with latitude in flowering time of vernalized plants and in vernalization sensitivity (r^2=0.18 and r^2=0.12). Our findings are in line with the over-wintered plants of the aforementioned common garden experiment, although the correlation coefficient is much smaller. This difference might be a sampling bias since we also found stronger correlation (r^2=0.38) in nine accessions that were part of both studies.

However, further testing is required to assure a biological relevant meaning of the observed latitudinal cline since population structure was not taken into consideration in both studies. When testing for latitudinal cline, population structure has to be taken into

consideration to assure that an observed latitudinal cline is really due to adaptation and not only to population structure. For several years, it was thought that there was little population structure in *A. thaliana* (Sharbel et al., 2000; Nordborg et al., 2002). Only an increased amount of sequence information has recently revealed the existence of large-scale population structure in *A. thaliana* (Nordborg et al., 2005; Schmid et al., 2006).

2 Uk-1/Uk-3 hybrid incompatibility

2.1 Genetic architecture of the Uk-1/Uk-3 incompatibility

F2 segregation and QTL mapping revealed teo genetic loci being responsible for the Uk-1/Uk-3 incompatibility, consistent with a two-gene Bateson-Dobzhansky-Muller (BDM)-type interaction. In the BDM model of speciation, divergence of two genes is responsible for reproductive isolation between two populations (Figure 28). In each population at least one gene has diverged from a common ancestor. Each derived allele functions well in its own genomic context, but when the two derived alleles come together in hybrids, their interaction leads to a phenotype that is either lethal or sterile and thus underlies reproductive isolation.

As described above (results 2.3), the F2 segregation pattern and mapping results in the Uk-1/Uk-3 incompatibility suggest the occurrence of the incompatibility phenotype if both derived alleles are present in a heterozygous state. If three derived alleles are present, the plant phenotype is lethal.

The two genomic regions that make Uk-1 and Uk-3 incompatible could be localized to small regions on chromosome 3 and chromosome 5, and candidate genes were tested (see below). Genotypes of F2 plants revealed that the Uk-1 allele on chromosome 3 and the Uk-3 allele on chromosome 5 are the derived problematic alleles. Statistical analysis suggested significant epistatic interaction of both loci to cause the observed maladaptive phenotype. This is consistent with the BDM model where deleterious epistatic interactions are the primary genetic basis of incompatibilities.
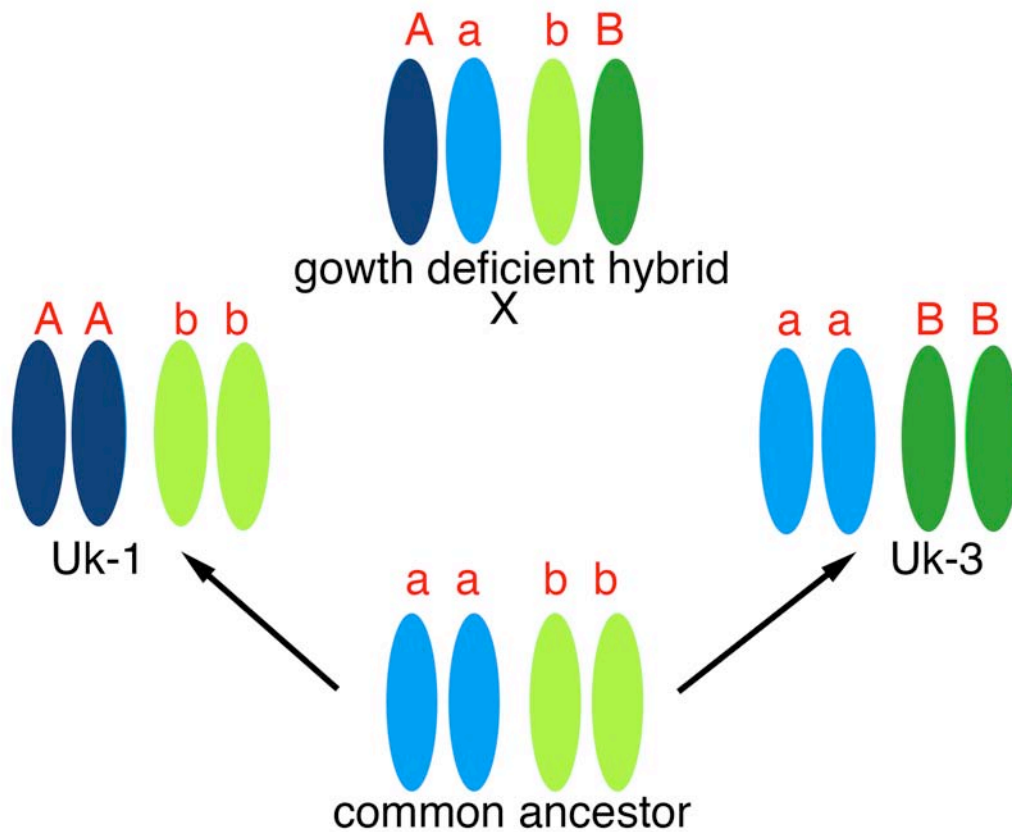
Figure 28: Schematic of the BDM model (red letters) with two genes, a and b (blue and green symbols) of which "a" diverged in one population to "A" (dark blue symbol) and "b" diverged in another population to "B" (dark green symbol). Both alleles "A" and "B" function well in their own genetic context, but in hybrids the interaction of two derived alleles is detrimental. This model fits to the observed incompatibility between Uk-1 and Uk-3 (black letters). Blue symbols correspond to the locus on chromosome 3 and green symbols to the chromosome 5 locus. Darker colors represent derived problematic alleles.

Since the causal alleles for the Uk-1/Uk-3 incompatibility are dominant, F1 hybrids show severe growth deficits and do not produce seeds under ecologically relevant temperature conditions (16°C), suggesting an extreme loss of fitness. Therefore, this incompatibility constitutes a functional reproductive isolation. Fortunately, F1 non-propagation could be bypassed in the laboratory using the temperature-dependence of the system: F1 hybrids recover a wild type-like phenotype and produce seeds after shifting to higher, rather non-natural, ambient temperature (23°C). Thus, the F2 segregation ratio could be determined and mapping experiments performed.

73

2.2 Physiological basis of the Uk-1/Uk-3 incompatibility

Active pathogen defense response in F1 hybrids

Our microarray analysis revealed a clear association of the growth defective incompatibility phenotype observed in F1 hybrids with an activated pathogen defense response. Gene classes involved in pathogen response are overrepresented in the set of genes induced in hybrids as is the number of genes with a W-box motif in the promoter region. WRKY transcription factors can bind to W-box motifs and are implicated in defense response regulation (Eulgem et al., 2000; Eulgem, 2005).

When a plant is exposed to a pathogen attack, the plant can defend itself both by an unspecific basal defense response and by pathogen-specific, highly effective immune responses, the R-mediated response. This immune response is governed by specific interaction of a pathogen derived avirulence protein (avr) and a corresponding disease resistance (R) gene of the plant, also known as the gene-for-gene resistance (Flor, 1971; Eulgem, 2005). The interaction of pathogen avr protein and host R-gene product mediates the fast activation of pathogen defense, leading to a hypersensitive cell death response (Dangl and Jones, 2001; Glazebrook, 2001; Chisholm et al., 2006).

Defense signaling pathways can be distinguished genetically based on their activation by diverse R genes and by the requirement of distinct pathway components. However, the different components of this complex signaling network are highly interconnected and activate a common set of target genes (Eulgem, 2005). Different induction pathways merge at several positions and expression levels of induced pathways differ rather in timing and strength than in components involved (Eulgem, 2005). Nevertheless, several distinctions of pathways have been made and can be connected to the induced signaling pathway in F1 hybrids.

The signaling molecules salicylic acid, jasmonic acid and ethylene play important roles in pathogen defense and define separate signaling pathways (Glazebrook, 2001; Kunkel and Brooks, 2002). In dwarf-like F1 hybrids expression levels of genes acting in the salicylic acid pathway were elevated, whereas genes of the jasmonic acid or ethylene pathway were not.

Genetic analysis of pathogen response mutants lead to the general distinction of at least two routes of activation, one requiring EDS1 and PAD4, the other NDR1 (Falk et al.,

1999; Feys et al., 2001; Glazebrook, 2001). The high levels of *EDS1* and *PAD4* transcripts we found in F1 hybrids suggest that the EDS1 and PAD4 dependent pathway is activated in the Uk-1/Uk-3 incompatibility. This pathway is associated with R-proteins of the class Toll-interleukin-nucleotide-binding-leucine-rich-repeat (TIR-NB-LRR) (Dangl and Jones, 2001; Glazebrook, 2001). This is in line with a TIR-NB-LRR gene that has been identified as the chromosome 5 locus responsible for the Uk-1/Uk-3 incompatibility (Bomblies et al., 2007).

Genes responsible for Uk-1/Uk-3 hybrid incompatibility

As mentioned above, the incompatibility gene on chromosome 5 was identified to be the R-gene at5g41740/50 of the class TIR-NB-LRR (Bomblies et al., 2007). This is consistent with a constitutively active pathogen response in F1 hybrids.

For the incompatibility locus on chromosome 3, candidate genes located within the mapping interval were tested by genomic complementation, gene knock-down using amiRNA and gene induction experiments. Surprisingly, introducing a genomic construct of any candidate gene located within the mapped interval did not alter the phenotype at all. However, knock-down experiments point to at3g44530 (*HIRA)* as the candidate gene on chromosome 3. F1 hybrids in which at3g44530 expression was artificially reduced exhibit a more wild type-like phenotype compared to Uk-1/Uk-3 hybrids. And specific induction of the Uk-1 allele revealed an F1-like phenotype, which was slightly different from Uk-1/Uk-3 hybrids.

Phelps-Durr and colleagues identified *HIRA* as an interacting partner of ASYMMETRIC LEAVES1 that is required for *knox* gene silencing (Phelps-Durr et al., 2005). Yeast and mammalian orthologues of *HIRA* are histone chaperones that control expression patterns of specific euchromatic genes by modulating chromatin structure (Spector et al., 1997; Kaufman et al., 1998; Magnaghi et al., 1998; Roberts et al., 2002). Interestingly, the maize orthologue of *HIRA* interacts with WRKY proteins, but not the *A. thaliana HIRA* protein (Phelps-Durr et al., 2005). In *A. thaliana, HIRA* function is essential for plant survival and expression is tightly regulated. Homozygous mutants were found to be lethal and in the few viable transformants containing a 35S-construct, *HIRA* turned out to be co-suppressed (Phelps-Durr et al., 2005). Remarkably, the phenotype of these co-suppression lines has significant similarities to Uk-1/Uk-3 hybrids at 23°C. In both, sepals, petals and stamens of flowers at 23°C are shortened and leaves are asymmetric and develop lobes.

Taken together, knock-down of *HIRA* as well as the phenotype of *HIRA* mutants suggested that the UK-1 allele of this gene might be the chromosome 3 locus responsible for

the incompatibility phenotype. Against *HIRA* argues that neither genomic complementation experiments with *HIRA* were successful nor could be find any sequence changes unique to the Uk-1 allele of *HIRA*. However, it is puzzling that none of the genomic constructs that entirely cover the mapping interval affected the phenotype. Epigenetic changes can also impair gene function (Stokes et al., 2002; Yi et al., 2004), but the methylation pattern of *HIRA* was not examined in this study.

In conclusion, it remains unclear whether any of the genes defined in this interval is causative for the hybrid incompatibility. It also needs to be considered that the phenotypic classification of plants in the mapping populations as being either "F1"- or "F2"-like was sometimes difficult. If indeed some hybrids were mis-classified, the mapping interval would have to be revisited.

Activated pathogen response brings about the incompatibility phenotype

A tight control of the pathogen defense response is crucial for the survival of plants as well as other organisms. This control has to allow both, prompt reaction to a pathogen attack as well as effective shutdown of the pathogen response machinery under non-pathogenic circumstances. Constitutive expression of pathogen response genes has detrimental effects on plant stature and development. A phenotype similar to the described F1 hybrid has been found in several mutants that constitutively express pathogen response genes: *cpr1* (Bowling et al., 1994), *mpk4* (Petersen et al., 2000), *ssi1* (Shah et al., 1999) and *bal* (Stokes et al., 2002). It includes twisted leaves, dwarfed stature and reduced fecundity. Further examples are the *bon1/cop1* and *snc1* mutants. Both proteins act in the same pathway and constitutive expression of pathogen response genes results in a dwarf phenotype that is temperature-sensitive - like Uk-1/Uk-3 hybrids (Hua et al., 2001; Zhang et al., 2003; Yang and Hua, 2004). Temperature sensitivity of pathogen response reaction has been noted before (Malamy et al., 1992; Wang et al., 2001). Similar to Uk-1/Uk-3 hybrids, PR-genes and salicylic acid accumulate at lower temperature but not at higher temperature (Malamy et al., 1992). Interestingly, R-gene accumulation increases with temperature (Wang et al., 2001).

These findings strongly suggest that an up-regulated pathogen response is causative for the growth-deficient and temperature-sensitive phenotype observed in F1 hybrids of Uk-1/Uk-3. Assuming that the strength of the pathogen response – and thus the severity of the phenotype – may depend on gene dosage could explain the lethality observed in F2

76

generations. Consistent with this hypothesis, constitutively active pathogen response cannot only result in growth-deficiency but also in lethality (Mackey et al., 2003).

Is ectopic defense response a common mechanism of reproductive isolation?

Speciation genes identified so far are of diverse function without common theme. Although no gene that causes reproductive isolation in plants is known yet, several hybrid incompatibilities within and between plant species have been described (Rieseberg and Carney, 1998). Interestingly, necrosis, extensive cell death and dwarf-like habitus characterize many incompatibilities observed in diverse taxa (Caldwell and Compton, 1943; Gerstel, 1954; Hermsen, 1963, 1967; Phillips, 1977; Singh and Gutiérrez, 1984; Mino et al., 2002; Yamada and Marubashi, 2003). Also temperature sensitive incompatibilities are not uncommon (Phillips, 1977; Valkonen and Watanabe, 1999; Mino et al., 2002).

High similarities of the incompatibility phenotypes found in different plants suggest also similarities in the physiology underlying the phenotype. Therefore, one can propose that the ectopic defense response observed in Uk-1/Uk-3 hybrids might be a common physiological basis of reproductive isolation found in diverse plant taxa.

2.3   Proposed Model

Based on the BDM model, we suggest that Uk-1 and Uk-3 alleles have evolved from a common ancestor (Figure 28 & Figure 29). Both alleles function well in their own genetic background. However, the interaction of the Uk-1 allele of the chromosome 3 locus with the Uk-3 allele of the R-gene on chromosome 5 is detrimental. This interaction leads to a constitutive activation of the pathogen response and is responsible for the growth deficiency observed in F1 hybrids.
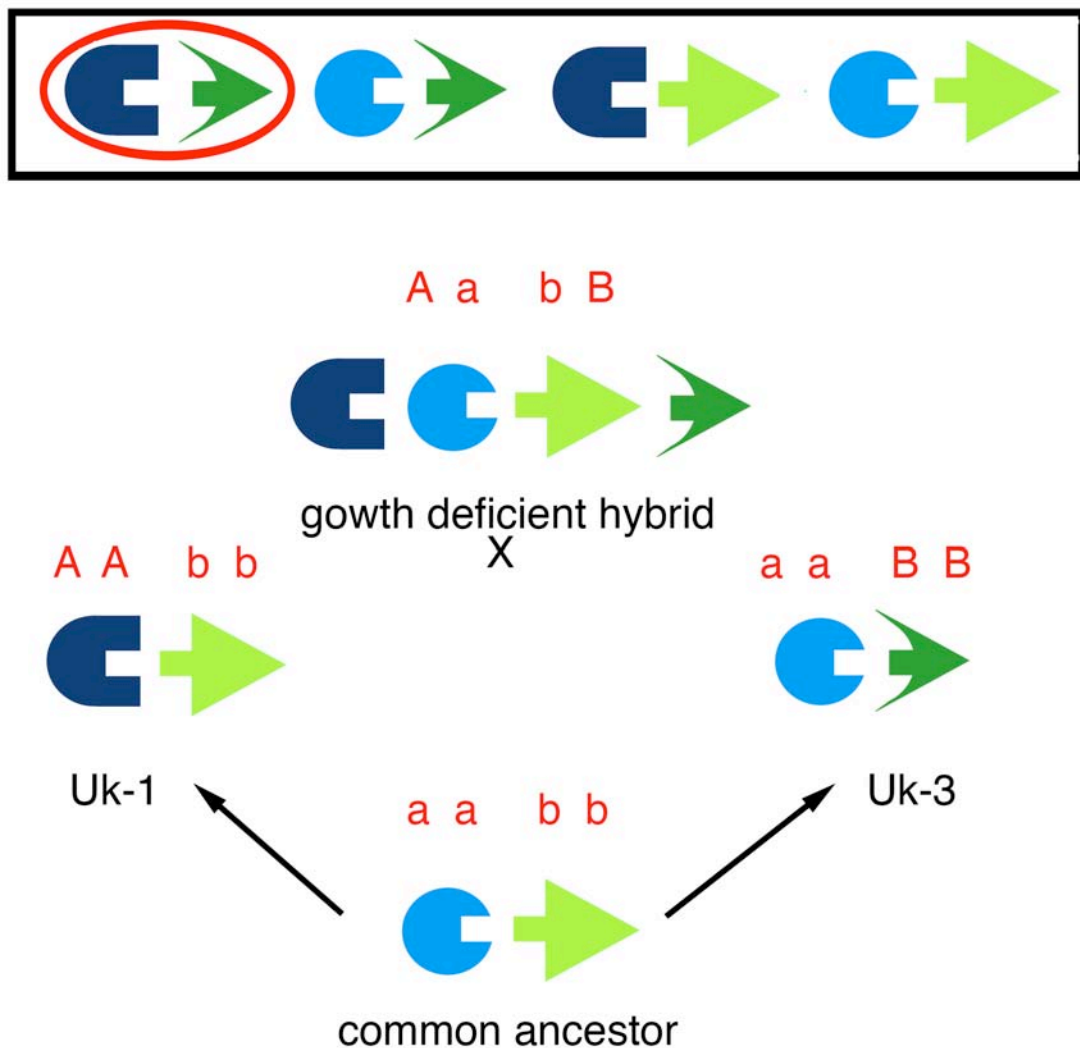
Figure 29: Mechanistic model of interacting Uk-1 and Uk-3 alleles that lead to an incompatibility reaction in F1 hybrids, following the BDM-model mentioned above. Color patterns are same as in Figure 28. The chromosome 3 locus is represented by blue shapes and the R-gene on chromosome 5 by green symbols. Darker colors illustrate evolved alleles. The dark blue symbol represents the problematic Uk-1 allele of chromosome 3 and the dark green symbol the problematic Uk-3 allele of chromosome 5. In the growth deficient F1 hybrids, four different allele combinations are possible (presented within a black box), one of them is detrimental (red ellipse).

The BDM-model suggests an epistatic interaction of two incompatible alleles of two genes involved in the incompatibility. In F1 hybrids, three out of four possible allelic combinations (Ab, aB, ab) result in a positive epistatic interaction and one (AB) is detrimental (Figure 29). The detrimental interaction involves both derived alleles. This model can also explain the lethal phenotype observed in F2 and later generations. Whereas one detrimental

allele combination is present in plants with F1 incompatibility phenotype, at least two detrimental allelic combinations per plant cause lethality (Figure 30).
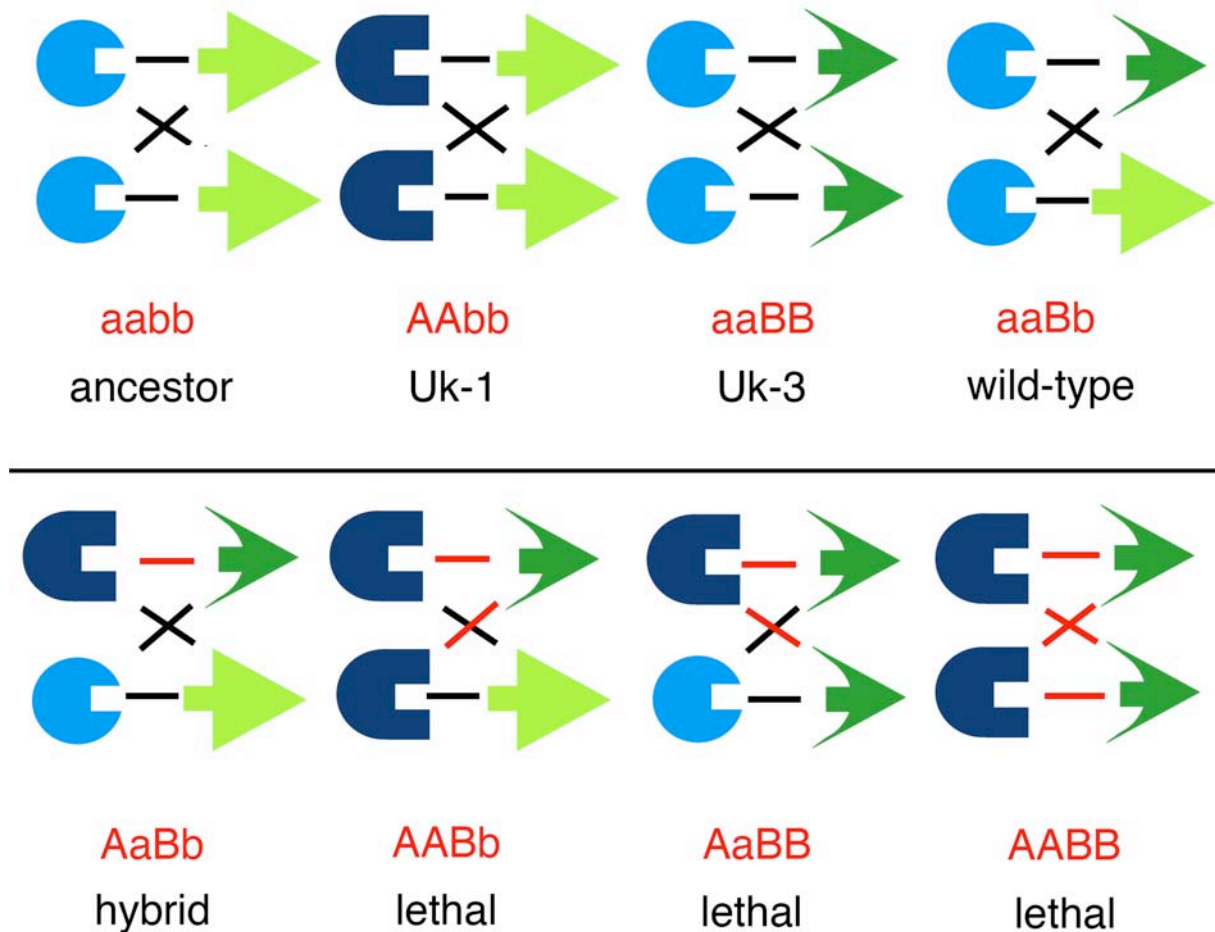
Figure 30: Allele combinations in the F2 generation with genotypes (red letters) and corresponding symbols. Symbols and colors are the same as in previous figures. Black and red lines indicate positive and negative allele interactions respectively. The first row shows parents (AAbb, aaBB), the fictive common ancestor (aabb) and an F2 individual with the genotype aaBb that are wild type phenotype. In the second row are allelic combinations that have at least one detrimental allele interaction: the hybrid (AaBb) and all lethal allele combinations found in F2 generations (AABb, AaBB, AABB).

In R-gene mediated pathogen resistance, the matching of a specific R-gene product with the corresponding pathogen-derived avr protein is crucial. Direct interaction of matching avr and R-gene proteins could only be observed in rare cases. Therefore, a third interacting protein has been proposed to be involved in recognition of the avr protein (Van der Biezen and Jones, 1998; Dangl and Jones, 2001). In this model, known as "guard hypothesis", an avr protein targets a third component, a "guardee", which is monitored by the R-gene (guard). Many avr proteins possess enzymatic activity and are believed to alter the "guardee"

enzymatically. The R-protein would monitor this perturbation (Chisholm et al., 2006). Strong support for the guard hypothesis comes from the case of RIN4 (guardee), which interacts with the R-protein RPM1 (guard) as well as with AvrRmp1 or AvrB (Mackey et al., 2003).

In the observed incompatibility, one could hypothesize the incompatibility loci to be two components of the guarding machinery. The R-gene on chromosome 5 could guard the gene product of the chromosome 3 locus (guardee). In case of a pathogen attack, the avr protein would enzymatically alter the chromosome 3 gene product, which then would result in a defense response. In F1 hybrids, the Uk-1 allele of the chromosome 3 locus might be recognized by the Uk-3 allele of the R-gene as altered by an invasive avr protein, even in absence of such. This would lead to the observed constitutive activation of defense response and to the growth deficiency.

2.4   Speciation aspects

Speciation genes as fast evolving genes

Although the so-called speciation genes identified so far are of diverse functions, a common feature is their fast evolving nature (Coyne and Orr, 2004; Orr, 2005; Mallet, 2006). In several cases, signatures of positive Darwinian selection were found in speciation genes (Ting et al., 1998; Barbash et al., 2003; Presgraves et al., 2003; Presgraves and Stephan, 2007). Therefore, incompatibilities - and in the long run species – arise as a by-product of natural selection.

In line with this trend, the incompatibility locus on chromosome 5 is an R-gene, which are known to be highly variable. In particular the leucine-rich-repeat- regions, which are involved in pathogen recognition are variable in nucleotide and protein sequence (Bergelson et al., 2001; Mauricio et al., 2003; Bakker et al., 2006a).

In contrast to the polymorphic R-genes, only few nucleotide polymorphisms have been identified for *HIRA* on chromosome 3 in the set of 78 accessions that were partially sequenced in this study. Although final proof is required to identify *HIRA* as the second speciation gene in the Uk-1/Uk-3 incompatibility, it is an interesting notion to think about a slowly evolving speciation gene. Even though only fast evolving speciation genes are known yet, the number is rather small and there is only one example for a pair of speciation genes known so far (Brideau et al., 2006).

80

Sympatric versus allopatric speciation

Although Uk-1 and Uk-3 were both collected in Umkirch in Germany by Röbbelen in 1954 (Röbbelen, 1965), it is difficult to find out whether these accessions also originate from the same location and population. The frequency of genome wide nucleotide differences observed between Uk-1 and Uk-3 is not unusual compared to nucleotide differences in a set of randomly chosen accessions. Therefore, a particularly close or distant relationship between Uk-1 and Uk-3 cannot be assumed. In *A. thaliana,* similar densities of nucleotide polymorphisms have been observed in other local regional samples compared to samples throughout the species range. This makes general assumptions of relationships between accessions difficult (Le Corre et al., 2002; Le Corre, 2005; Nordborg et al., 2005; Bakker et al., 2006b).

It is both possible that Uk-1 and Uk-3 have originated and evolved at the same location in Umkirch, or that one or both accessions only recently migrated to Umkirch. Therefore, both sympatric and allopatric population divergence could have occurred in the Umkirch *A. thaliana* population.

It is a long running controversy, whether species can arise in sympatry, or wether geographic barriers (allopatry) are required for speciation. Until the later half of the 20th century, leading evolutionists believed only in allopatric speciation (Dobzhansky, 1937; Mayr, 1963; Coyne and Orr, 2004). However, this view started to change recently, with emerging evidence for sympatric speciation in cichlid fish (Joyce et al., 2005; Barluenga et al., 2006) and palm trees (Savolainen et al., 2006), commented by (Ortiz-Barrientos and Rieseberg, 2006).

Relevance of the observed incompatibility

Although incompatible F1 hybrids were produced and monitored under laboratory conditions, the observed incompatibility is likely to be relevant for the natural population in Umkirch. F1 hybrids are severely compromised at 16°C with obviously reduced fitness. The growth conditions at 16°C resemble much better ecologically relevant conditions compared to higher growth temperatures usually used in the laboratory (Hoffmann, 2002). Although, *A. thaliana* is a selfing plant, hybridization between different accessions has been observed in wild and is likely to matter for population evolution (Hoffmann et al., 2003; Le Corre, 2005).

81

Fortunately, we could still find one of the incompatibility alleles, of Uk-3, in an accession we collected in Umkirch in 2005, about 50 years after Röbbelen collected Uk-1 and Uk-3 (Röbbelen, 1965). This finding proves the natural occurrence of at least one examined incompatibility allele and excludes the possibility that it arose as a spontaneous mutation in the laboratory.

Incipient speciation

It has been difficult for evolutionary biologists to discriminate between (1) genes that are involved in the first step of speciation and (2) gene interactions that lead to incompatibility in present-day species but only have accumulated secondarily after the two lines already had separated (Coyne and Orr, 2004). These questions are difficult to address in crosses between individuals of already separated species. Most speciation research has been performed between *Drosophila* species and revealed many loci that cause incompatibility between two species. The maximum estimate involves more than 190 loci in a single incompatibility (Presgraves, 2003; Mallet, 2006). Very likely, many of these incompatibility loci have only accumulated after the speciation event.

The described Uk-1/Uk-3 incompatibility appears between two accessions of the same species. This incompatibility therefore represents a case of incipient speciation, before complete separation. In this study system, one can hypothesize that the observed minimal number of two loci, of which one is an R-gene, may be sufficient for a first step in speciation. Since evolution is slow, we do not know whether a new distinct species will eventually evolve from the *A. thaliana* population in Umkirch. But combining studies of incipient and completed speciation will certainly help to understand the process of speciation.

V. SUMMARY

The fascinating process of evolution leads to both well-adapted organisms with continuous natural variation and to distinct species. Both processes, adaptation and speciation, are key interests of biological research. This work explores flowering time variation in *A. thaliana* accessions as an adaptive trait and studies the hybrid incompatibility which was observed between the two accessions Uk-1 and Uk-3 and might represent an example of incipient speciation.

Flowering time

For the success of plant populations, the tight control of flowering time is crucial. In this study, the flowering time of 155 natural accessions together with 32 flowering time mutants was measured in four different environments representing the major environmental cues affecting flowering time: daylength, ambient temperature and vernalization, a prolonged winter-like time of cold. The flowering time distribution under long day conditions showed two peaks and confirmed the important role of *FRI* and *FLC* in flowering time regulation. In late-flowering plants, we found no obvious deletions in either gene, but the function of either gene was impaired in all early flowering accessions. Most frequent were the well-known Col-0- or L*er*-type deletion in *FRI*. Additionally, we found 20 accessions with so far unknown, rare alleles of *FRI* or *FLC*. *FLC* expression levels showed a high linear correlation with flowering time. Microarray experiments revealed that *FLC* is indeed the gene most highly correlated with flowering time, among all genes represented on the ATH1 microarray.

Despite the fact that *FRI* and *FLC* functionality determines early or late flowering behavior, we found substantial variation within the two classes, reflecting the action of additional signaling pathways. Also, the effect of *FRI* and *FLC* is much smaller under short day conditions, accounting for only 23% of the observed variation - in contrast to 63% under long day conditions. Through environmental sensitivity analysis we have identified conspicuous accessions that are promising candidates to identify new natural alleles of genes controlling flowering time in future studies. A new natural allele of *PHYD* could indeed be identified successfully from a conspicuous accession identified in this flowering time screen.

Speciation

The most widely accepted species definition by Ernst Mayr defines species as reproductively isolated groups and turned reproductive isolation or hybrid incompatibility into the main target of biological speciation research.

We discovered a dominant hybrid incompatibility between the two *A. thaliana* accessions Uk-1 and Uk-3. The temperature-sensitive, dwarf-like phenotype with leaf malformations and infertility under ecologically relevant conditions (16°C) suggests a severe loss of fitness. The genetic basis of the observed incompatibility fits with a two gene Bateson-Dobzhansky-Muller model: Two derived alleles of distinct genes function well in their own genetic context. however, negative epistatic interaction of these alleles has detrimental effects when they come together in a hybrid plant,. We mapped the two responsible loci to small chromosomal regions on chromosome 3 and 5. The region on chromosome 3 was fine mapped and all candidate genes present within the mapping interval were tested. The gene *HIRA* is likely to be the incompatibility gene, however final proof is still required.

Microarray experiments suggest that the ectopic expression of pathogen response genes underlies the observed incompatibility. This is consistent with other cases of hybrid incompatibilities described in different plant taxa that also involve pathogen response and show symptoms very similar as described here. Taken together, these findings suggest that ectopic activation of pathogen response genes might be a common mechanism involved in the first steps of species separation.

To find such an incompatibility between individuals of the same species, gives the opportunity to study speciation at its beginning. Although we will not find out if eventually a new species will form, we can ask questions about speciation that cannot be addressed by the incompatibilities between already existing species that mainly have been used in speciation research so far.

VI. REFERENCES

Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., Ichinoki, H., Notaguchi, M., Goto, K., and Araki, T. (2005). FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. Science **309,** 1052-1056.

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

Alonso-Blanco, C., Mendez-Vigo, B., and Koornneef, M. (2005a). From phenotypic to molecular polymorphisms involved in naturally occurring variation of plant development. Int J Dev Biol **49,** 717-732.

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

Alonso-Blanco, C., Bentsink, L., Hanhart, C.J., Blankestijn-de Vries, H., and 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., and Martinez-Zapater, J.M. (2005b). 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.

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

Arnold, M.L. (1997). Natural hybridization and evolution. (New York: Oxford University Press).

Arnold, M.L., and Meyer, A. (2006). Natural hybridization in primates: one evolutionary mechanism. Zoology (Jena) **109,** 261-276.

Aukerman, M.J., Hirschfeld, M., Wester, L., Weaver, M., Clack, T., Amasino, R.M., and Sharrock, R.A. (1997). A deletion in the PHYD gene of the Arabidopsis Wassilewskija ecotype defines a role for phytochrome D in red/far-red light sensing. Plant Cell **9,** 1317-1326.

Ausin, I., Alonso-Blanco, C., Jarillo, J.A., Ruiz-Garcia, L., and Martinez-Zapater, J.M. (2004). Regulation of flowering time by FVE, a retinoblastoma-associated protein. Nat Genet **36,** 162-166.

Bakker, E.G., Toomajian, C., Kreitman, M., and Bergelson, J. (2006a). A genome-wide survey of R gene polymorphisms in Arabidopsis. Plant Cell **18,** 1803-1818.

Bakker, E.G., Stahl, E.A., Toomajian, C., Nordborg, M., Kreitman, M., and Bergelson, J. (2006b). Distribution of genetic variation within and among local populations of Arabidopsis thaliana over its species range. Mol Ecol **15,** 1405-1418.

Balasubramanian, S., Sureshkumar, S., Lempe, J., and Weigel, D. (2006a). Potent induction of Arabidopsis thaliana flowering by elevated growth temperature. PLoS Genet **2,** e106.

Balasubramanian, S., Sureshkumar, S., Agrawal, M., Michael, T.P., Wessinger, C., Maloof, J.N., Clark, R., Warthmann, N., Chory, J., and Weigel, D. (2006b). The PHYTOCHROME C photoreceptor gene mediates natural variation in flowering and growth responses of Arabidopsis thaliana. Nat Genet **38,** 711-715.

Barbash, D.A., Roote, J., and Ashburner, M. (2000). The Drosophila melanogaster hybrid male rescue gene causes inviability in male and female species hybrids. Genetics **154,** 1747-1771.

Barbash, D.A., Siino, D.F., Tarone, A.M., and Roote, J. (2003). A rapidly evolving MYB-related protein causes species isolation in Drosophila. Proc Natl Acad Sci U S A **100,** 5302-5307.

Barluenga, M., Stolting, K.N., Salzburger, W., Muschick, M., and Meyer, A. (2006). Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature **439,** 719-723.

Bastow, R., Mylne, J.S., Lister, C., Lippman, Z., Martienssen, R.A., and Dean, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. Nature **427,** 164-167.

Bateson, W. (1909). Heredity and variation in modern lights. In Darwin and modern science, A.C. Steward, ed (Cambridge, U.K.: Cambridge University Press), pp. 85-101.

Baurle, I., and Dean, C. (2006). The timing of developmental transitions in plants. Cell **125,** 655-664.

Bentsink, L., Yuan, K., Koornneef, M., and Vreugdenhil, D. (2003). The genetics of phytate and phosphate accumulation in seeds and leaves of Arabidopsis thaliana, using natural variation. Theor Appl Genet **106,** 1234-1243.

Bentsink, L., Jowett, J., Hanhart, C.J., and Koornneef, M. (2006). Cloning of DOG1, a quantitative trait locus controlling seed dormancy in Arabidopsis. Proc Natl Acad Sci U S A **103,** 17042-17047.

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

Bergelson, J., Kreitman, M., Stahl, E.A., and Tian, D. (2001). Evolutionary dynamics of plant R-genes. Science **292,** 2281-2285.

Blazquez, M.A., Ahn, J.H., and Weigel, D. (2003). A thermosensory pathway controlling flowering time in Arabidopsis thaliana. Nat Genet **33,** 168-171.

Blazquez, M.A., Green, R., Nilsson, O., Sussman, M.R., and Weigel, D. (1998). Gibberellins promote flowering of arabidopsis by activating the LEAFY promoter. Plant Cell **10,** 791-800.

Bomblies, K., Lempe, J., Epple, P., Warthmann, N., Lanz, C., Dangl, J.L., and Weigel, D. (2007). A mechanism for Hybrid Necrosis, a Geneetic Incompatibility Syndrome in plants. PLoS Biol **in press**.

Borevitz, J.O., Maloof, J.N., Lutes, J., Dabi, T., Redfern, J.L., Trainer, G.T., Werner, J.D., Asami, T., Berry, C.C., Weigel, D., and Chory, J. (2002). Quantitative trait loci controlling light and hormone response in two accessions of Arabidopsis thaliana. Genetics **160,** 683-696.

Bowling, S.A., Guo, A., Cao, H., Gordon, A.S., Klessig, D.F., and Dong, X. (1994). A mutation in Arabidopsis that leads to constitutive expression of systemic acquired resistance. Plant Cell **6,** 1845-1857.

Breitling, R., Armengaud, P., Amtmann, A., and Herzyk, P. (2004). Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett **573,** 83-92.

Brideau, N.J., Flores, H.A., Wang, J., Maheshwari, S., Wang, X., and Barbash, D.A. (2006). Two Dobzhansky-Muller genes interact to cause hybrid lethality in Drosophila. Science **314,** 1292-1295.

Broman, K.W., Wu, H., Sen, S., and Churchill, G.A. (2003). R/qtl: QTL mapping in experimental crosses. Bioinformatics **19,** 889-890.

Caicedo, A.L., Stinchcombe, J.R., Olsen, K.M., Schmitt, J., and Purugganan, M.D. (2004). Epistatic interaction between Arabidopsis FRI and FLC flowering time genes generates a latitudinal cline in a life history trait. Proc Natl Acad Sci U S A **101,** 15670-15675.

Caldwell, R.M., and Compton, L.E. (1943). Complementary lethal genes in wheat. Journal of Heredity **34**.

Candela, M., Martinez-Laborda, A., and Micol, J.L. (1999). Venetion pattern formation in *Arabidopsis thaliana* vegetative leaves. Developmental Biology **205,** 205-216.

Canvin, D.T., and McVetty, P.B.E. (1976). Hybrid grass-clump dwarfness in wheat: physiology and genetics. Euphytica **25,** 471-483.

Castillo-Davis, C.I., and Hartl, D.L. (2003). GeneMerge--post-genomic analysis, data mining, and hypothesis testing. Bioinformatics **19,** 891-892.

Cervera, M.T., Ruiz-Garcia, L., and Martinez-Zapater, J.M. (2002). Analysis of DNA methylation in Arabidopsis thaliana based on methylation-sensitive AFLP markers. Mol Genet Genomics **268,** 543-552.

Chase, C.D. (2007). Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. Trends Genet **23,** 81-90.

Chisholm, S.T., Coaker, G., Day, B., and Staskawicz, B.J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. Cell **124,** 803-814.

Christie, P., and Macnair, M.R. (1984). Complementary lethal factors in two north american populations of the yellow monkey flower. Journal of Heredity **75,** 510-511.

Chu, Y., and Oka, H. (1972). The distribution and effects of genes causing F1 weakness in *Oryza breviligulata* and *O. glaberrima*. genetics **70,** 163-173.

Clark, R.M., Schweikert, G., Toomajian, C., Ossowski, S., Zeller, G., Shinn, P., Warthmann, N., Hu, T.T., Fu, G., Hinds, D.A., Chen, H., Frazer, K.A., Huson, D.H., Schöllkopf, B., Nordborg, M., Rätsch, G., Ecker, J.R., and Weigel, D. (2007). Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. Science **accepted**.

Clarke, J.H., and Dean, C. (1994). Mapping FRI, a locus controlling flowering time and vernalization response in Arabidopsis thaliana. Mol Gen Genet **242,** 81-89.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J **16,** 735-743.

Cooley, N.M., Higgins, J.T., Holmes, M.G., and Attridge, T.H. (2001). Ecotypic differences in responses of Arabidopsis thaliana L. to elevated polychromatic UV-A and UV-B+A radiation in the natural environment: a positive correlation between UV-B+A inhibition and growth rate. J Photochem Photobiol B **60,** 143-150.

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., and Coupland, G. (2007). FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of Arabidopsis. Science.

Coyne, J.A., and Orr, H.A. (2004). Speciation. (Sunderland, MA, USA: Sinauer Associates, Inc.).

Cracraft, J. (1989). Speciation and its ontology: The empirical consequences of alternative species concepts for understanding patterns and processes of differentiation In Speciation and its consequences, D. Otte and J.A. Endler, eds (Sunderland, MA: Sinauer Association.

Dangl, J.L., and Jones, J.D. (2001). Plant pathogens and integrated defence responses to infection. Nature **411,** 826-833.

Dobzhansky, T.H. (1937). Genetics and the origin of species. (New York: Columbia University Press).

Doyle, J.J., and Doyle, J.L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull **19,** 11-15.

Doyle, M.R., Bizzell, C.M., Keller, M.R., Michaels, S.D., Song, J., Noh, Y.S., and Amasino, R.M. (2005). HUA2 is required for the expression of floral repressors in Arabidopsis thaliana. Plant J **41,** 376-385.

Edgar, R.C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5,** 113.

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

Ellison, C.K., and Burton, R.S. (2006). Disruption of mitochondrial function in interpopulation hybrids of Tigriopus californicus. Evolution Int J Org Evolution **60,** 1382-1391.

Endler, J.A. (1977). Geographic Variation, Speciation, and the Clines. (Princeton, NJ: Princton University Press).

Eulgem, T. (2005). Regulation of the Arabidopsis defense transcriptome. Trends Plant Sci **10,** 71-78.

Eulgem, T., Rushton, P.J., Robatzek, S., and Somssich, I.E. (2000). The WRKY superfamily of plant transcription factors. Trends Plant Sci **5,** 199-206.

Falconer, D.S., and Mackay, T.F.C. (1996). Introduction to Quantitative Genetics. (Harlow, Essex: Addison Wesley Longman).

Falk, A., Feys, B.J., Frost, L.N., Jones, J.D., Daniels, M.J., and Parker, J.E. (1999). EDS1, an essential component of R gene-mediated disease resistance in Arabidopsis has homology to eukaryotic lipases. Proc Natl Acad Sci U S A **96,** 3292-3297.

Feys, B.J., Moisan, L.J., Newman, M.A., and Parker, J.E. (2001). Direct interaction between the Arabidopsis disease resistance signaling proteins, EDS1 and PAD4. Embo J **20,** 5400-5411.

Fishman, L., and Willis, J.H. (2006). A cytonuclear incompatibility causes anther sterility in Mimulus hybrids. Evolution Int J Org Evolution **60,** 1372-1381.

Flor, H.H. (1971). Current status of the gene-for-gene concept. Annu Rev Phytopathol **9,** 275-296.

Gazzani, S., Gendall, A.R., Lister, C., and Dean, C. (2003). Analysis of the molecular basis of flowering time variation in Arabidopsis accessions. Plant Physiol **132,** 1107-1114.

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

Gerstel, D.U. (1954). A new lethal combination in interspecific cotton hybrids. Genetics **39,** 628-639.

Glazebrook, J. (2001). Genes controlling expression of defense responses in Arabidopsis-- 2001 status. Curr Opin Plant Biol **4,** 301-308.

Gompert, Z., Fordyce, J.A., Forister, M.L., Shapiro, A.M., and Nice, C.C. (2006). Homoploid hybrid speciation in an extreme habitat. Science **314,** 1923-1925.

Grant, V. (1971). Plant Speciation. (New York: Columbia University Press).

Hagenblad, J., Tang, C., Molitor, J., Werner, J., Zhao, K., Zheng, H., Marjoram, P., Weigel, D., and Nordborg, M. (2004). Haplotype structure and phenotypic associations in the chromosomal regions surrounding two Arabidopsis thaliana flowering time loci. Genetics **168,** 1627-1638.

Hannah, M.A., Wiese, D., Freund, S., Fiehn, O., Heyer, A.G., and Hincha, D.K. (2006). Natural genetic variation of freezing tolerance in Arabidopsis. Plant Physiol **142,** 98-112.

Harrison, J.S., and Burton, R.S. (2006). Tracing hybrid incompatibilities to single amino acid substitutions. Mol Biol Evol **23,** 559-564.

Hayama, R., and Coupland, G. (2004). The molecular basis of diversity in the photoperiodic flowering responses of Arabidopsis and rice. Plant Physiol **135,** 677-684.

He, Y., Michaels, S.D., and Amasino, R.M. (2003). Regulation of flowering time by histone acetylation in Arabidopsis. Science **302,** 1751-1754.

Hermsen, J.G.T. (1963). The genetic basis of hybrid necrosis in wheat. Genetica **33,** 245-287.

Hermsen, J.G.T. (1967). Hybrid dwarfness in wheat. Euphytica **16,** 134-162.

Hoekenga, O.A., Vision, T.J., Shaff, J.E., Monforte, A.J., Lee, G.P., Howell, S.H., and Kochian, L.V. (2003). Identification and characterization of aluminum tolerance loci in Arabidopsis (Landsberg erecta x Columbia) by quantitative trait locus mapping. A physiologically simple but genetically complex trait. Plant Physiol **132,** 936-948.

Hoffmann, M.H. (2002). Biogeography of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae). Journal of Biogeography **29,** 125-134.

Hoffmann, M.H., Bremer, M., Schneider, K., Burger, F., Stolle, E., and Moritz, G. (2003). Flower visitors in natural population of *Arabidopsis thaliana* Plant Biology **5,** 491-494.

Hollingshead, L.A. (1930). A lethal factor in *Crepis* effective only in an interspecific hybrid. Genetics **15,** 114-140.

Hua, J., Grisafi, P., Cheng, S.H., and Fink, G.R. (2001). Plant growth homeostasis is controlled by the Arabidopsis BON1 and BAP1 genes. Genes Dev **15,** 2263-2272.

Hutter, P., and Ashburner, M. (1987). Genetic rescue of inviable hybrids between Drosophila melanogaster and its sibling species. Nature **327,** 331-333.

Hutter, P., Roote, J., and Ashburner, M. (1990). A genetic basis for the inviability of hybrids between sibling species of Drosophila. Genetics **124,** 909-920.

Imaizumi, T., and Kay, S.A. (2006). Photoperiodic control of flowering: not only by coincidence. Trends Plant Sci **11,** 550-558.

Jack, T., Brockman, L.L., and Meyerowitz, E.M. (1992). The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. Cell **68,** 683-697.

Jacobsen, S.E., and Olszewski, N.E. (1993). Mutations at the SPINDLY locus of Arabidopsis alter gibberellin signal transduction. Plant Cell **5,** 887-896.

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

Joyce, D.A., Lunt, D.H., Bills, R., Turner, G.F., Katongo, C., Duftner, N., Sturmbauer, C., and Seehausen, O. (2005). An extant cichlid fish radiation emerged in an extinct Pleistocene lake. Nature **435,** 90-95.

Juenger, T., Purugganan, M., and Mackay, T.F. (2000). Quantitative trait loci for floral morphology in Arabidopsis thaliana. Genetics **156,** 1379-1392.

Kaufman, P.D., Cohen, J.L., and Osley, M.A. (1998). Hir proteins are required for position-dependent gene silencing in Saccharomyces cerevisiae in the absence of chromatin assembly factor I. Mol Cell Biol **18,** 4793-4806.

King, J.J., and Stimart, D.P. (1998). Genetic analysis of variation for auxin-induced adventitious root formation among eighteen ecotypes of Arabidopsis thaliana L. Heynh. J Hered **89,** 481-487.

Kliebenstein, D.J., Gershenzon, J., and Mitchell-Olds, T. (2001a). Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in Arabidopsis thaliana leaves and seeds. Genetics **159,** 359-370.

Kliebenstein, D.J., Lambrix, V.M., Reichelt, M., Gershenzon, J., and Mitchell-Olds, T. (2001b). Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in Arabidopsis. Plant Cell **13,** 681-693.

Kliebenstein, D.J., Kroymann, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J., and Mitchell-Olds, T. (2001c). Genetic control of natural variation in Arabidopsis glucosinolate accumulation. Plant Physiol **126,** 811-825.

Kobayashi, Y., and Koyama, H. (2002). QTL analysis of Al tolerance in recombinant inbred lines of Arabidopsis thaliana. Plant Cell Physiol **43,** 1526-1533.

Koch, M.A., Dobes, C., and Mitchell-Olds, T. (2003). Multiple hybrid formation in natural populations: concerted evolution of the internal transcribed spacer of nuclear ribosomal DNA (ITS) in North American Arabis divaricarpa (Brassicaceae). Mol Biol Evol **20,** 338-350.

Koornneef, M., Hanhart, C.J., and van der Veen, J.H. (1991). A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana. Mol Gen Genet **229,** 57-66.

Koornneef, M., Alonso-Blanco, C., and Vreugdenhil, D. (2004). Naturally occurring genetic variation in Arabidopsis thaliana. Annu Rev Plant Biol **55,** 141-172.

Koornneef, M., Blankenstijn-de Vries, H., Hanhart, C., Soppe, W., and Peters, T. (1994). The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg *erecta* wild type Plant Journal **6,** 911-919.

Kroymann, J., Donnerhacke, S., Schnabelrauch, D., and Mitchell-Olds, T. (2003). Evolutionary dynamics of an Arabidopsis insect resistance quantitative trait locus. Proc Natl Acad Sci U S A **100 Suppl 2,** 14587-14592.

Kruger, J., Thomas, C.M., Golstein, C., Dixon, M.S., Smoker, M., Tang, S., Mulder, L., and Jones, J.D. (2002). A tomato cysteine protease required for Cf-2-dependent disease resistance and suppression of autonecrosis. Science **296,** 744-747.

Kubo, T., and Yoshimura, A. (2002). Genetic basis of hybrid breakdown in a Japonica/Indica cross of rice, Oryza sativa L. Theor Appl Genet **105,** 906-911.

Kunkel, B.N., and Brooks, D.M. (2002). Cross talk between signaling pathways in pathogen defense. Curr Opin Plant Biol **5,** 325-331.

Laibach, F. (1943). *Arabidospis thaliana* (L.) Heynh. als Objekt für genetische und entwicklungs-physiologische Untersuchungen. Bot Arch **44,** 439-455.

Lambrix, V., Reichelt, M., Mitchell-Olds, T., Kliebenstein, D.J., and Gershenzon, J. (2001). The Arabidopsis epithiospecifier protein promotes the hydrolysis of glucosinolates to nitriles and influences Trichoplusia ni herbivory. Plant Cell **13,** 2793-2807.

Le Corre, V. (2005). Variation at two flowering time genes within and among populations of Arabidopsis thaliana: comparison with markers and traits. Mol Ecol **14,** 4181-4192.

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

Lemon, W.J., Liyanarachchi, S., and You, M. (2003). A high performance test of differential gene expression for oligonucleotide arrays. Genome Biol **4,** R67.

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

Levin, D.A. (2000). The origin, expansion and demise of plant species. (New York: Oxford University Press).

Li, Y., Roycewicz, P., Smith, E., and Borevitz, J.O. (2006). Genetics of Local Adaptation in the Laboratory: Flowering Time Quantitative Trait Loci under Geographic and Seasonal Conditions in Arabidopsis. PLoS ONE **1,** e105.

Loudet, O., Gaudon, V., Trubuil, A., and 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.

Loudet, O., Chaillou, S., Merigout, P., Talbotec, J., and Daniel-Vedele, F. (2003). Quantitative trait loci analysis of nitrogen use efficiency in Arabidopsis. Plant Physiol **131,** 345-358.

Lynch, M., and Force, A.G. (2000). The origin of interspecific genomic incompatibility via gene duplication. Amerian Naturalist **156,** 590-605.

Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R., and Dangl, J.L. (2003). Arabidopsis RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. Cell **112,** 379-389.

Macknight, R., Duroux, M., Laurie, R., Dijkwel, P., Simpson, G., and Dean, C. (2002). Functional significance of the alternative transcript processing of the Arabidopsis floral promoter FCA. Plant Cell **14,** 877-888.

Macnair, M.R., and Christie, P. (1983). Reproductive isolation as a pleiotropic effect of copper tolerance in *Mimulus guttatus*. Heredity **50,** 295-302.

Magnaghi, P., Roberts, C., Lorain, S., Lipinski, M., and Scambler, P.J. (1998). HIRA, a mammalian homologue of Saccharomyces cerevisiae transcriptional co-repressors, interacts with Pax3. Nat Genet **20,** 74-77.

Malamy, J., Hennig, J., and Klessig, D.F. (1992). Temperature-Dependent Induction of Salicylic Acid and Its Conjugates during the Resistance Response to Tobacco Mosaic Virus Infection. Plant Cell **4,** 359-366.

Malitschek, B., Fornzler, D., and Schartl, M. (1995). Melanoma formation in Xiphophorus: a model system for the role of receptor tyrosine kinases in tumorigenesis. Bioessays **17,** 1017-1023.

Mallet, J. (1995). A species definition for the Mordern Synthesis. Trends Ecol Evol **10,** 294-299.

Mallet, J. (2005). Hybridization as an invasion of the genome. Trends Ecol Evol **20,** 229-237.

Mallet, J. (2006). What does Drosophila genetics tell us about speciation? Trends Ecol Evol **21,** 386-393.

Maloof, J.N., Borevitz, J.O., Dabi, T., Lutes, J., Nehring, R.B., Redfern, J.L., Trainer, G.T., Wilson, J.M., Asami, T., Berry, C.C., Weigel, D., and Chory, J. (2001). Natural variation in light sensitivity of Arabidopsis. Nat Genet **29,** 441-446.

Mandel, M.A., Gustafson-Brown, C., Savidge, B., and Yanofsky, M.F. (1992). Molecular characterization of the Arabidopsis floral homeotic gene APETALA1. Nature **360,** 273-277.

Masly, J.P., Jones, C.D., Noor, M.A., Locke, J., and Orr, H.A. (2006). Gene transposition as a cause of hybrid sterility in Drosophila. Science **313,** 1448-1450.

Mauricio, R., Stahl, E.A., Korves, T., Tian, D., Kreitman, M., and Bergelson, J. (2003). Natural selection for polymorphism in the disease resistance gene Rps2 of Arabidopsis thaliana. Genetics **163,** 735-746.

Mayr, E. (1942). Systematics and the Origin of Species. (New York: Columbia University Press).

Mayr, E. (1963). Animal Species and Evolution (Cambridge, MA: Belknap Press).

McClung, C.R. (2006). Plant circadian rhythms. Plant Cell **18,** 792-803.

McKay, J.K., Richards, J.H., and Mitchell-Olds, T. (2003). Genetics of drought adaptation in Arabidopsis thaliana: I. Pleiotropy contributes to genetic correlations among ecological traits. Mol Ecol **12,** 1137-1151.

Michael, T.P., Salome, P.A., Yu, H.J., Spencer, T.R., Sharp, E.L., McPeek, M.A., Alonso, J.M., Ecker, J.R., and McClung, C.R. (2003). Enhanced fitness conferred by naturally occurring variation in the circadian clock. Science **302,** 1049-1053.

Michaels, S.D., and Amasino, R.M. (1999). FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell **11,** 949-956.

Michaels, S.D., and Amasino, R.M. (2001). Loss of FLOWERING LOCUS C activity eliminates the late-flowering phenotype of FRIGIDA and autonomous pathway mutations but not responsiveness to vernalization. Plant Cell **13,** 935-941.

Michaels, S.D., He, Y., Scortecci, K.C., and Amasino, R.M. (2003). Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in Arabidopsis. Proc Natl Acad Sci U S A **100,** 10102-10107.

Millar, A.A., and Kunst, L. (1999). The natural genetic variation of the fatty-acyl composition of seed oils in different ecotypes of Arabidopsis thaliana. Phytochemistry **52,** 1029-1033.

Mindrinos, M., Katagiri, F., Yu, G.L., and Ausubel, F.M. (1994). The A. thaliana disease resistance gene RPS2 encodes a protein containing a nucleotide-binding site and leucine-rich repeats. Cell **78,** 1089-1099.

Mino, M., Maekawa, K., Ogawa, K., Yamagishi, H., and Inoue, M. (2002). Cell death processes during expression of hybrid lethality in interspecific F1 hybrid between Nicotiana gossei Domin and Nicotiana tabacum. Plant Physiol **130,** 1776-1787.

Mitchell-Olds, T., and Pedersen, D. (1998). The molecular basis of quantitative genetic variation in central and secondary metabolism in Arabidopsis. Genetics **149,** 739-747.

Mithen, R., Clarke, J.H., Lister, C., and Dean, C. (1995). Genetics of aliphatic glucosinolates III. Side chain structure of aliphatic glucosinolates in *Arabidopsis thaliana*. Heredity **74,** 210-215.

Moon, J., Suh, S.S., Lee, H., Choi, K.R., Hong, C.B., Paek, N.C., Kim, S.G., and Lee, I. (2003). The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. Plant J **35,** 613-623.

Mouchel, C.F., Briggs, G.C., and Hardtke, C.S. (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.

Mouradov, A., Cremer, F., and Coupland, G. (2002). Control of flowering time: interacting pathways as a basis for diversity. Plant Cell **14 Suppl,** S111-130.

Moyle, L.C., and Graham, E.B. (2005). Genetics of hybrid incompatibility between Lycopersicon esculentum and L. hirsutum. Genetics **169,** 355-373.

Muller, H.J. (1942). Isolating mechanisms, evolution, and temperature. Biol. Symp. **6,** 71-125.

Mylne, J.S., Barrett, L., Tessadori, F., Mesnage, S., Johnson, L., Bernatavichute, Y.V., Jacobsen, S.E., Fransz, P., and Dean, C. (2006). LHP1, the Arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. Proc Natl Acad Sci U S A **103,** 5012-5017.

Nordborg, M., Borevitz, J.O., Bergelson, J., Berry, C.C., Chory, J., Hagenblad, J., Kreitman, M., Maloof, J.N., Noyes, T., Oefner, P.J., Stahl, E.A., and Weigel, D. (2002). The extent of linkage disequilibrium in Arabidopsis thaliana. Nat Genet **30,** 190-193.

Nordborg, M., Hu, T.T., Ishino, Y., Jhaveri, J., Toomajian, C., Zheng, H., Bakker, E., Calabrese, P., Gladstone, J., Goyal, R., Jakobsson, M., Kim, S., Morozov, Y., Padhukasahasram, B., Plagnol, V., Rosenberg, N.A., Shah, C., Wall, J.D., Wang, J., Zhao, K., Kalbfleisch, T., Schulz, V., Kreitman, M., and Bergelson, J. (2005). The pattern of polymorphism in Arabidopsis thaliana. PLoS Biol **3,** e196.

Orr, H.A. (2005). The genetic basis of reproductive isolation: insights from Drosophila. Proc Natl Acad Sci U S A **102 Suppl 1,** 6522-6526.

Ortiz-Barrientos, D., and Rieseberg, L.H. (2006). Speciation: Splitting when together. Heredity **97,** 2-3.

Paterson, H.E.H. (1985). The recognition concept of species. In Species and Speciation, E.S. Vrba, ed (Pretoria: Transvaal Museum Monograph.

Perchepied, L., Kroj, T., Tronchet, M., Loudet, O., and Roby, D. (2006). Natural Variation in Partial Resistance to Pseudomonas syringae Is Controlled by Two Major QTLs in Arabidopsis thaliana. PLoS ONE **1**, e123.

Perez-Perez, J.M., Serrano-Cartagena, J., and Micol, J.L. (2002). Genetic analysis of natural variations in the architecture of Arabidopsis thaliana vegetative leaves. Genetics **162**, 893-915.

Petersen, M., Brodersen, P., Naested, H., Andreasson, E., Lindhart, U., Johansen, B., Nielsen, H.B., Lacy, M., Austin, M.J., Parker, J.E., Sharma, S.B., Klessig, D.F., Martienssen, R., Mattsson, O., Jensen, A.B., and Mundy, J. (2000). Arabidopsis map kinase 4 negatively regulates systemic acquired resistance. Cell **103**, 1111-1120.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res **29**, e45.

Phelps-Durr, T.L., Thomas, J., Vahab, P., and Timmermans, M.C. (2005). Maize rough sheath2 and its Arabidopsis orthologue ASYMMETRIC LEAVES1 interact with HIRA, a predicted histone chaperone, to maintain knox gene silencing and determinacy during organogenesis. Plant Cell **17**, 2886-2898.

Phillips, L.L. (1977). Interspecific incompatibility in *Gossypium*. IV. Temperature-conditional lethality in hybrids of *G. klotzschianum*. American Journal of Botany **64**.

Poduska, B., Humphrey, T., Redweik, A., and Grbic, V. (2003). The synergistic activation of FLOWERING LOCUS C by FRIGIDA and a new flowering gene AERIAL ROSETTE 1 underlies a novel morphology in Arabidopsis. Genetics **163**, 1457-1465.

Presgraves, D.C. (2003). A fine-scale genetic analysis of hybrid incompatibilities in Drosophila. Genetics **163**, 955-972.

Presgraves, D.C., and Stephan, W. (2007). Pervasive Adaptive Evolution among Interactors of the Drosophila Hybrid Inviability Gene, Nup96. Mol Biol Evol **24**, 306-314.

Presgraves, D.C., Balagopalan, L., Abmayr, S.M., and Orr, H.A. (2003). Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. Nature **423**, 715-719.

Putterill, J., Laurie, R., and Macknight, R. (2004). It's time to flower: the genetic control of flowering time. Bioessays **26**, 363-373.

Quesada, V., Macknight, R., Dean, C., and Simpson, G.G. (2003). Autoregulation of FCA pre-mRNA processing controls Arabidopsis flowering time. Embo J **22**, 3142-3152.

Quesada, V., Garcia-Martinez, S., Piqueras, P., Ponce, M.R., and Micol, J.L. (2002). Genetic architecture of NaCl tolerance in Arabidopsis. Plant Physiol **130**, 951-963.

Rashotte, A.M., Jenks, M.A., Nguyen, T.D., and Feldmann, K.A. (1997). Epicuticular wax variation in ecotypes of Arabidopsis thaliana. Phytochemistry **45**, 251-255.

Rauh, L., Basten, C., and Buckler, S.t. (2002). Quantitative trait loci analysis of growth response to varying nitrogen sources in Arabidopsis thaliana. Theor Appl Genet **104**, 743-750.

Riddle, N.C., and Richards, E.J. (2005). Genetic variation in epigenetic inheritance of ribosomal RNA gene methylation in Arabidopsis. Plant J **41**, 524-532.

Rieseberg, L.H. (1997). Hybrid origin of plant species. Ann Rev Ecol Syst **28**, 359-389.

Rieseberg, L.H., and Carney, S.E. (1998). Plant Hybridization. New Phytologist **140**, 599-624.

Röbbelen, G. (1965). The LAIBACH Standard Collection of Natural Races. Arabidopsis Information Service **2**, 36-47.

Roberts, C., Sutherland, H.F., Farmer, H., Kimber, W., Halford, S., Carey, A., Brickman, J.M., Wynshaw-Boris, A., and Scambler, P.J. (2002). Targeted mutagenesis of the Hira gene results in gastrulation defects and patterning abnormalities of mesoendodermal derivatives prior to early embryonic lethality. Mol Cell Biol **22,** 2318-2328.

Rogers, E.E., and Ausubel, F.M. (1997). Arabidopsis enhanced disease susceptibility mutants exhibit enhanced susceptibility to several bacterial pathogens and alterations in PR-1 gene expression. Plant Cell **9,** 305-316.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning. (Cold Spring Harbor N.Y.: Cold Spring Harbor Laboratory Press).

Sanchez-Moran, E., Armstrong, S.J., Santos, J.L., Franklin, F.C., and Jones, G.H. (2002). Variation in chiasma frequency among eight accessions of Arabidopsis thaliana. Genetics **162,** 1415-1422.

Savant, A.C. (1956). Semilethal Complementary Factors in a Tomato Species Hybrid. Evolution Int J Org Evolution **10,** 93-96.

Savolainen, V., Anstett, M.C., Lexer, C., Hutton, I., Clarkson, J.J., Norup, M.V., Powell, M.P., Springate, D., Salamin, N., and Baker, W.J. (2006). Sympatric speciation in palms on an oceanic island. Nature **441,** 210-213.

Schartl, M., Hornung, U., Gutbrod, H., Volff, J.N., and Wittbrodt, J. (1999). Melanoma loss-of-function mutants in Xiphophorus caused by Xmrk-oncogene deletion and gene disruption by a transposable element. Genetics **153,** 1385-1394.

Schmid, K.J., Torjek, O., Meyer, R., Schmuths, H., Hoffmann, M.H., and Altmann, T. (2006). Evidence for a large-scale population structure of Arabidopsis thaliana from genome-wide single nucleotide polymorphism markers. Theor Appl Genet **112,** 1104-1114.

Schmid, M., Uhlenhaut, N.H., Godard, F., Demar, M., Bressan, R., Weigel, D., and Lohmann, J.U. (2003). Dissection of floral induction pathways using global expression analysis. Development **130,** 6001-6012.

Schmuths, H., Bachmann, K., Weber, W.E., Horres, R., and Hoffmann, M.H. (2006). Effects of preconditioning and temperature during germination of 73 natural accessions of Arabidopsis thaliana. Ann Bot (Lond) **97,** 623-634.

Schwab, R., Ossowski, S., Riester, M., Warthmann, N., and Weigel, D. (2006). Highly specific gene silencing by artificial microRNAs in Arabidopsis. Plant Cell **18,** 1121-1133.

Sergeeva, L.I., Vonk, J., Keurentjes, J.J., van der Plas, L.H., Koornneef, M., and Vreugdenhil, D. (2004). Histochemical analysis reveals organ-specific quantitative trait loci for enzyme activities in Arabidopsis. Plant Physiol **134,** 237-245.

Shah, J., Kachroo, P., and Klessig, D.F. (1999). The Arabidopsis ssi1 mutation restores pathogenesis-related gene expression in npr1 plants and renders defensin gene expression salicylic acid dependent. Plant Cell **11,** 191-206.

Sharbel, T.F., Haubold, B., and Mitchell-Olds, T. (2000). Genetic isolation by distance in Arabidopsis thaliana: biogeography and postglacial colonization of Europe. Mol Ecol **9,** 2109-2118.

Shaw, K.L. (1998). Species and the diversity of natural groups. In Endless forms: Species and Speciation, D.J. Howard and S.J. Berlocher, eds (Oxford: Oxford University Press), pp. 44-56.

Sheldon, C.C., Rouse, D.T., Finnegan, E.J., Peacock, W.J., and Dennis, E.S. (2000). The molecular basis of vernalization: the central role of FLOWERING LOCUS C (FLC). Proc Natl Acad Sci U S A **97,** 3753-3758.

Sheldon, C.C., Burn, J.E., Perez, P.P., Metzger, J., Edwards, J.A., Peacock, W.J., and Dennis, E.S. (1999). The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. Plant Cell **11,** 445-458.

Shindo, C., Lister, C., Crevillen, P., Nordborg, M., and Dean, C. (2006). Variation in the epigenetic silencing of FLC contributes to natural variation in Arabidopsis vernalization response. Genes Dev **20,** 3079-3083.

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

Simpson, G.G., and Dean, C. (2002). Arabidopsis, the Rosetta stone of flowering time? Science **296,** 285-289.

Simpson, G.G., Dijkwel, P.P., Quesada, V., Henderson, I., and Dean, C. (2003). FY is an RNA 3' end-processing factor that interacts with FCA to control the Arabidopsis floral transition. Cell **113,** 777-787.

Singh, A.P., and Gutiérrez, J.A. (1984). Geographical distribution of the DL1 and DL2 genes causing hybrid dwarfism in *Phaseolus vulgaris* L., their association with seed size, and their significance to breeding. Euphytica **33,** 337-345.

Sokal, R.R., and Rohlf, F.J. (1981). Biometry. (New York: Freeman New York).

Spector, M.S., Raff, A., DeSilva, H., Lee, K., and Osley, M.A. (1997). Hir1p and Hir2p function as transcriptional corepressors to regulate histone gene transcription in the Saccharomyces cerevisiae cell cycle. Mol Cell Biol **17,** 545-552.

Staden, R., Beal, K.F., and Bonfield, J.K. (2000). The Staden package, 1998. Methods Mol Biol **132,** 115-130.

Stahl, E.A., Dwyer, G., Mauricio, R., Kreitman, M., and Bergelson, J. (1999). Dynamics of disease resistance polymorphism at the Rpm1 locus of Arabidopsis. Nature **400,** 667-671.

Stebbins, G.L. (1950). Variation and Evolution in Plants. (New York: Columbia Univeristy Press).

Stenoien, H.K., Fenster, C.B., Kuitten, H., and Savolainen, O. (2002). Quantifying latitudinal clines to light responses in natural populations of *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. American Journal of Botany **89,** 1604-1608.

Stinchcombe, J.R., Weinig, C., Ungerer, M., Olsen, K.M., Mays, C., Halldorsdottir, S.S., Purugganan, M.D., and Schmitt, J. (2004). A latitudinal cline in flowering time in Arabidopsis thaliana modulated by the flowering time gene FRIGIDA. Proc Natl Acad Sci U S A **101,** 4712-4717.

Stokes, T.L., Kunkel, B.N., and Richards, E.J. (2002). Epigenetic variation in Arabidopsis disease resistance. Genes Dev **16,** 171-182.

Sun, S., Ting, C.T., and Wu, C.I. (2004). The normal function of a speciation gene, Odysseus, and its hybrid sterility effect. Science **305,** 81-83.

Sung, S., and Amasino, R.M. (2004a). Vernalization and epigenetics: how plants remember winter. Curr Opin Plant Biol **7,** 4-10.

Sung, S., and Amasino, R.M. (2004b). Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. Nature **427,** 159-164.

Sung, S., and Amasino, R.M. (2005). Remembering winter: toward a molecular understanding of vernalization. Annu Rev Plant Biol **56,** 491-508.

Swarup, K., Alonso-Blanco, C., Lynn, J.R., Michaels, S.D., Amasino, R.M., Koornneef, M., and Millar, A.J. (1999). Natural allelic variation identifies new genes in the Arabidopsis circadian system. Plant J **20,** 67-77.

Sweigart, A.L., Fishman, L., and Willis, J.H. (2006). A simple genetic incompatibility causes hybrid male sterility in mimulus. Genetics **172,** 2465-2479.

Tempelton, A.R. (1989). The meaning of species and speciation: A genetic perspective. In Speciation and its consequences, D. Otte and J.A. Endler, eds (Sunderland, MA: Sinauer Associates), pp. 3-27.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res **25,** 4876-4882.

Ting, C.T., Tsaur, S.C., Wu, M.L., and Wu, C.I. (1998). A rapidly evolving homeobox at the site of a hybrid sterility gene. Science **282,** 1501-1504.

Toomajian, C., Hu, T.T., Aranzana, M.J., Lister, C., Tang, C., Zheng, H., Zhao, K., Calabrese, P., Dean, C., and Nordborg, M. (2006). A nonparametric test reveals selection for rapid flowering in the Arabidopsis genome. PLoS Biol **4,** e137.

Valkonen, J.P.T., and Watanabe, K.N. (1999). Autonomous cell death, temperature sensitivity and the genetic control associated with resistance to cucumber mosaic virus (CMV) in diploid potatoes (*Solanum* spp.). Theor Appl Genet **99,** 996-1005.

Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., and Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. Science **303,** 1003-1006.

Van der Biezen, E.A., and Jones, J.D. (1998). Plant disease-resistance proteins and the gene-for-gene concept. Trends Biochem Sci **23,** 454-456.

van Der Schaar, W., Alonso-Blanco, C., Leon-Kloosterziel, K.M., Jansen, R.C., van Ooijen, J.W., and Koornneef, M. (1997). QTL analysis of seed dormancy in Arabidopsis using recombinant inbred lines and MQM mapping. Heredity **79 (Pt 2),** 190-200.

Van Valen, L. (1976). Ecological species, multispecies, and oaks. Taxon **25,** 233-239.

Wang, Z.X., Yamanouchi, U., Katayose, Y., Sasaki, T., and Yano, M. (2001). Expression of the Pib rice-blast-resistance gene family is up-regulated by environmental conditions favouring infection and by chemical signals that trigger secondary plant defences. Plant Mol Biol **47,** 653-661.

Weigel, D., and Glazebrook, J. (2002). Arabidopsis: A Laboratory Manual. (Cold Spring Harbor N.Y.: Cold Spring Harbor Laboratory Press).

Wen, S.Y., Shimada, K., Kawai, K., and Toda, M.J. (2006). Strong purifying selection on the Odysseus gene in two clades of sibling species of the Drosophila montium species subgroup. J Mol Evol **62,** 659-662.

Werner, J.D., Borevitz, J.O., Uhlenhaut, N.H., Ecker, J.R., Chory, J., and Weigel, D. (2005a). FRIGIDA-independent variation in flowering time of natural Arabidopsis thaliana accessions. Genetics **170,** 1197-1207.

Werner, J.D., Borevitz, J.O., Warthmann, N., Trainer, G.T., Ecker, J.R., Chory, J., and Weigel, D. (2005b). Quantitative trait locus mapping and DNA array hybridization identify an FLM deletion as a cause for natural flowering-time variation. Proc Natl Acad Sci U S A **102,** 2460-2465.

West, M.A., Kim, K., Kliebenstein, D.J., van Leeuwen, H., Michelmore, R.W., Doerge, R.W., and St Clair, D.A. (2007). Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. Genetics **175,** 1441-1450.

Wiebe, G.A. (1934). Complementary factors in barley giving a lethal progeny. Journal of Heredity **25,** 272-274.

Wigge, P.A., Kim, M.C., Jaeger, K.E., Busch, W., Schmid, M., Lohmann, J.U., and Weigel, D. (2005). Integration of spatial and temporal information during floral induction in Arabidopsis. Science **309,** 1056-1059.

Wiley, E.O. (1978). The evolutionary species concept reconsidered. Syst. Zool. **27,** 17-26.

Willett, C.S., and Burton, R.S. (2004). Evolution of interacting proteins in the mitochondrial electron transport system in a marine copepod. Mol Biol Evol **21,** 443-453.

Wilson, R.N., Heckman, J.W., and Somerville, C.R. (1992). Gibberellin Is Required for Flowering in Arabidopsis thaliana under Short Days. Plant Physiol **100,** 403-408.

Wittbrodt, J., Adam, D., Malitschek, B., Maueler, W., Raulf, F., Telling, A., Robertson, S.M., and Schartl, M. (1989). Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu locus in Xiphophorus. Nature **341,** 415-421.

Wu, C.I., and Ting, C.T. (2004). Genes and speciation. Nat Rev Genet **5,** 114-122.

Yamada, T., and Marubashi, W. (2003). Overproduced ethylene causes programmed cell death leading to temperature-sensitive lethality in hybrid seedlings from the cross Nicotiana suaveolens x N. tabacum. Planta **217,** 690-698.

Yamada, T., Marubashi, W., Nakamura, T., and Niwa, M. (2001). Possible involvement of auxin-induced ethylene in an apoptotic cell death during temperature-sensitive lethality expressed by hybrid between Nicotiana glutinosa and N. repanda. Plant Cell Physiol **42,** 923-930.

Yang, S., and Hua, J. (2004). A haplotype-specific Resistance gene regulated by BONZAI1 mediates temperature-dependent growth control in Arabidopsis. Plant Cell **16,** 1060-1071.

Yanovsky, M.J., and Kay, S.A. (2002). Molecular basis of seasonal time measurement in Arabidopsis. Nature **419,** 308-312.

Yi, H., Riddle, N.C., Stokes, T.L., Woo, H.R., and Richards, E.J. (2004). Induced and natural epigenetic variation. Cold Spring Harb Symp Quant Biol **69,** 155-159.

Zhang, Y., Goritschnig, S., Dong, X., and Li, X. (2003). A gain-of-function mutation in a plant disease resistance gene leads to constitutive activation of downstream signal transduction pathways in suppressor of npr1-1, constitutive 1. Plant Cell **15,** 2636-2646.

VII. PUBLICATIONS AND CONTRIBUTIONS

During the course of this work, the following articles have been published or were submitted for publication:

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

Bomblies, K.*, Lempe, J.*, Epple, P., Warthmann, N., Lanz, C., Dangl, J.L., and Weigel, D. (2007). A mechanism for Hybrid Necrosis, a Geneetic Incompatibility Syndrome in plants. PLoS Biol, in press.

Balasubramanian, S., Sureshkumar, S., Lempe, J., and Weigel, D. (2006). Potent induction of Arabidopsis thaliana flowering by elevated growth temperature. PLoS Genet **2,** e106.

* authors contributed equally

Contributions to experiments

Without contributions of several colleagues, the described work would not have been possible:

Flowering time of all accessions and mutants was measured by S. Balasubramanian, S. Sureshkumar and J. Lempe

Summary statistics and genetic correlations across environments were calculated by S. Balasubramanian

Expression levels of *FLC* of 155 accessions were measured by S. Balasubramanian

Microarray experiments of 38 accessions were performed by S. Balasubramanian, J. Lempe, J. Lohmann and M. Schmid

Pearson correlations of gene expression with flowering time were calculated by J. Lempe and D. Weigel selected correlation coefficients of flowering time genes

Jürgen Berger prepared samples for Scanning Electron Microscopy

Single nucleotide polymorphism markers were designed by N. Warthmann

Genotyping was performed by Genaissance Pharmacyticals

Christa Lanz constructed a fosmid library and shotgun sequenced two fosmid clones

VIII. ABBREVIATIONS

Chemicals

CTAB	Cetyl-Trimethylammonium Bromide
ddH2O	double destilled water
DIG	Digoxygenin
EDTA	Ethylendiamintetraaceticacid
EtOH	Ethanol
NaCl/ Citrate	Sodium Cloride /Citrate
NaOH	Sodium Hydroxide
SDS	Sodium Dodecyl Sulfate
SSC	Sodium Sodium Citrate
Tris	Tris-[hydroxymethyl]-aminomethan

Flowering time related abbreviations

SD	short days
LD	long days
16LD	16°C long days
16LDV	16°C long days with vernalization treatment
23LD	23°C long days
23SD	23°C short days
DTF	days to flowering
JLN	juvenile leaf number
ALN	adult leaf number
RLN	rosette leaf number
TLN	total leaf number
FRI	FRIGIDA
FLC	Flowering Locus C

Other abbreviations

amiRNA	artificial micro RNA
ANOVA	analysis of variance
bp	base pair
CAPS	cleaved amplified polymorphic sequence marker
cDNA	coding Desoxyribonucleic Acid
DNA	Desoxyribonucleic Acid
gcRMA	guanidine cytosine robust multi-array
H2	heritability
kb	kilo base pair
LOD score	log 10 likelihood ratio comparing a single-QTL model to a zero-QTL model
Mb	mega base pair
qRT-PCR	quantitative real-time PCR
QTL	quantitative trait locus
RNA	Ribonucleic Acid
RNAse	Ribonuclease
rpm	rounds per minute

SEM	scanning electron micrograph
SNP	single nucleotide polymorphism
SSLP	simple sequence length polymorphism marker
v/v	volume per volume
w/v	weight per volume

IX. ACKNOWLEDGEMENTS

First of all, I would like to thank Detlef Weigel for giving me the opportunity to work in this department and for leaving most exiting research projects to me. I am thankful to him for teaching me diverse aspects of "non-pedestrian" science and for sharing his clever ideas. I greatly enjoyed his very supportive and positive attitude.

I want to express my gratitude to Sureshkumar Balasubramanian, who introduced me into natural variation science and guided the first part of my phD. I am also thankful to him for many discussions and for critical reading of this manuscript.

Many thanks to Gerd Jürgens for his interest in my work as second assessor and for being part of my phD advisory committee.

New thoughts and learning diverse techniques requires an active and friendly exchange among colleagues. My special appreciation goes to Carla Schommer for her friendship and for showing me how to do successful cloning. Thanks to Norman Warthmann for introducing me into "R" and for help with diverse computational problems and for. For discussions and friendship I gratefully acknowledge Sandra Stehling, Alexis Maizel, Javier Palatnik, Richard Clark, Frank Küttner, Sridevi Sureshkumar, Kirsten Bomblies, Vava Grbic, Markus Schmid and Jan Lohmann.

I am much obliged to my parents, especially to my mother for taking care of Malte while I was writing this thesis.

I also want to thank Björn Falkenburger for support in various ways and for critical reading of this manuscript.

X. APPENDIX

1 Primer list

| Number | Name | Sequence 5´-> 3´ |

FRIGIDA

	Number	Name	Sequence 5´-> 3´
1	N-0109	CLFRI16	GTG AGT GTA TCT AGT GTT CA
2	N-0107	CLFRI3	GTG GAA ATT AGG GCT TCT G
3	G-0360	FRIP F	AGT ACT CAC AAG TCA CAA C
4	G-0362	FRIP-R	GAA GAT CAT CGA ATT GGC
5	N-0574	JW100	ATG TTA TTG CTT CAG AGG TT
6	N-0831	JW159	AGG GCG TAG AGC ATT TAC
7	N-0832	JW160	TAA TCC AAC TCT CAA TCT TCA
8	N-0105	JW21	CGT ATG ACT TAG CTG TTG GAT
9	N-0106	JW22	CTT GTG AGT CTC CAT ACA CTG
10	N-0572	JW98	AAC GAC TTT TTC CTT TGA G
11	N-1078	N-1078	GGA TAA CTA TTT GGG GTC TAA TGA TGA GTT ACT GCG
12	G-0363	UJ26	AGA TTT GCT GGA TTT GAT AAG G
13	G-0365	UJ34	ATA TTT GAT GTG CTC TCC

1-13	used for sequencing FRI
12,13	Col-type deletion
3,4	Ler-type deletion
1,3	3.3 kb fragment
6,7	for qRT-PCR

Flowering Locus C

	Number	Name	Sequence 5´-> 3´
14	N-1106	CT05	GTT ACT CAG TTA CTC TTT TTG
15	N-1102	CT12	GGA TTT CAT TAT TTC CTT GG
16	N-1103	CT13	CTT TGA ATC ACA ATC GTC GT
17	N-1104	CT14	TG TAT CTT GTG TCT TTT GTC
18	N-1107	CT15	ATT TTG ACA CGA GAT TAC TAA
19	N-1109	CT17	ATG TAA AAG GTA AGG TGT TC
20	G-2513	G-2513	GGA AGA ACA ATG TCG TGA AGA A
21	G-2526	G-2526	GCT AGT ATT GAT GAC CCA TAA GAT
22	G-2527	G-2527	TCC CTC AAA GAA AAG TCA TAC A
23	G-1792	JA01	CAA ATC GTG AAT GAC ATG C
24	G-1793	JA02	AAT TAG ACC AGT TTA TGT ACA GCA
25	G-1794	JA03	GGC ACC AAA GAA ACA AGG CT
26	G-1795	JA04	TAT CAG TCC TAT TGT GAA GTT AAG
27	G-1796	JA05	ATA ATG ATG ATG TGG CGG TAA
28	G-1797	JA06	TGT GAA TCT ATG TTG AAA TAA TTG AT
29	G-1798	JA07	AGG ATC AAA ACT ACT AGC TAA CCC
30	G-1799	JA08	CCT CCA GTT GAA CAA GAG CAT C
31	G-1805	JA11	AGA TTG GGG CTG CGT TTA CAT TTT AT
32	G-2622	JA15	ATT GAT TCA TAT TTT TCA TAC ACA G
33	N-0912	JW177	TGT ATT TGG AGT TTG GCT TC
34	G-1426	YK13	CGA GAA AAG GAA AAA AAA AAA TAG AAA GAG
35	G-1427	YK14	CTC ACA CGA ATA AGG TAC AAA GTT CAT C

36	N-0913	JW178	ACA TTG AGA GAA CAC CTT AC
37	N-0914	JW179	TGG TCA CTT TTT TGG TTG CT
38	N-1124	JW183	ATT GGG GCT GCG TTT ACA
39	G-1586	G-1586	TAA AAA CTT TCT CTC AAT TCT CTC T
40	G-1587	G-1587	TTG TCG ATG GTG TCG GAG CTT
41	G-1981	G-1981	AGA CAA GAA GAC CGA ACT CA
42	G-1982	G-1982	TTT GTC CAG CAG GTG ACA TC
43	G-0868	G-0868	AAA ATA TCT GGC CCG ACG AAG
44	G-0869	G-0869	CGA CGA GAA GAG CGA CGG ATG
45	G-3024	JA23	GGA TCC GAG AAA AGG AAA AAA AAA AAT AGA AAG AG
46	G-3025	JA24	GTC GAC CTC ACA CGA ATA AGG TAC AAA GTT CAT C

14-35	sequencing FLC
31,30	FLC genomic Fragment
34,35	for FLC cDNA
36,37,38	Ler-type insertion
39, 40	UBQ
41, 42	qRT-PCR
43, 44	flc-3 genotyping
45, 46	cloning cFLC

Fine mapping of chromosome 3 incompatibility region

47	G-5111		TAT TTC CCT TGG CTC AGT GG
48	G-5112		CAG GAC CAA ACC ATC GAG TT
49	G-5369		CAA TTT GGG AGT GAC AAC AA
50	G-5370		TGG AGC GTG TTG AAT GTA TC
51	G-5504		GCG AAA TGG TCT GTT TGG AT
52	G-5505		TGG CTC AGT GAT CAC AGC TT
53	G-5550		ATG GAG GAA CAT CTT CAA GGA A
54	G-5551		TTT GAT TTC TTT ACC TTT GAG G
55	G-5552		CGA ATT GAG AGA AGA GAG ATG AT
56	G-5553		TGT CCA AAC TCC CAC TAG CC
57	G-5554		GGA TAG ATA TAA GTT TCT TCA ATG
58	G-5555		TTT TCA ACC GCA GTT TCT CAG
59	G-6065		CAT TGT GGT CCA CTG AGA ATG T
60	G-6066		TCG GTC CAA GCA AAT TTT ATG
61	G-6036		TTG TCC CCT GTC GAT TTT TG
62	G-6037		CGG AGA AGA CCT CGA ACA AG
63	G-5371		GAC TTT CTT GCA ACG TCC TT
64	G-5372		ACA CTT GAG CAC TCC CTT TC
65	G-5564		ATG AAG GTT GCC GCT GCT
66	G-5565		AAA AAC CTT ACT CGA AGA GAC TG
67	G-5288		AAA CAA TCT CAG GAG CCA AG
68	G-5289		AGA GAA CAA TGG CTC TAC GG
69	G-5115		GAG GAA AGC AAC GAG TCA GG
70	G-5368		ATG GCC TTT ATG TGA ATG GA
71	G-5290		AAT TTT CTC CGG GTC ATG TA
72	G-5291		TTA CCG GTT GAG TGT CAA TG
73	G-5292		AGA TGA AGG AAA CAG CCA AG
74	G-5293		GTG CAG GCT GGA ATA GTT CT
75	G-5119		GAC CGC AAA ATC CAA AAA TG
76	G-5120		GGA AAA CCA GAA ATG CGT GT
77	G-5123		TGG TTG CGT ATA TGC GTG TT
78	G-5124		TTC GAA AGG GAA TTC AAT CG
79	G-5481		ATG CAA CTC CGG CAA AGT TA
80	G-5482		ACG CAC AAA CCC TGA AAA TC

47, 48		CAPs marker left side chr3	
69, 70		SSLP marker right side chr3	
79, 80		genotyping chr 5	

Probes for screening a fosmid library

81	G-5550		ATG GAG GAA CAT CTT CAA GGA A
82	G-5551		TTT GAT TTC TTT ACC TTT GAG G
83	G-5556		CTT CCG AAT TAG ACC AAA GTC C
84	G-7890		GTG GCA GCT TCA TTG TCA GA

Artificial micro RNAs

85	G-7134	mir1-1	gaT AGG AAT GCA TGA CGT GTG GGt ctc tct ttt gta ttc c
86	G-7135	mir1-2	gaC CCA CAC GTC ATG CAT TCC TAt caa aga gaa tca atg a
87	G-7136	mir1-3	gaC CAA CAC GTC ATG GAT TCC TTt cac agg tcg tga tat g
88	G-7137	mir1-4	gaA AGG AAT CCA TGA CGT GTT GGt cta cat ata tat tcc t
89	G-7138	mir2-1	gaT ATT GCC GCA GCC GTG AGC GTt ctc tct ttt gta ttc c
90	G-7139	mir2-2	gaA CGC TCA CGG CTG CGG CAA TAt caa aga gaa tca atg a
91	G-7140	mir2-3	gaA CAC TCA CGG CTG GGG CAA TTt cac agg tcg tga tat g
92	G-7141	mir2-4	gaA ATT GCC CCA GCC GTG AGT GTt cta cat ata tat tcc t

Candidate genes at3g44530 & at3g44540 related primers

93	G-5552		CGA ATT GAG AGA AGA GAG ATG AT
94	G-5989		TGT CCA AAC TCC CAC TAG CC
95	G-6061		TGT TTG TGG AGA AAA TAT TGA GGA
96	G-9161		GTT GTG GCG GCT ACA TTG AC
97	G-1586		TAA AAA CTT TCT CTC AAT TCT CTC T
98	G-1587		TTG TCG ATG GTG TCG GAG CTT
99	G-9582		GCA AAA CAA ATC AAT AGT TTA GGA GA
100	G-9583		GGT GGG TCG AAG CTC AAA T
101	G-6056		ACT TTT CGC ATG TTC CTT GG
102	G-7883		TAA GCA AAC AGG TGG TGC AG
103	G-6051		CAG GGG GTG AAG AAG CAT AA
104	G-6052		CAT CCC TCT GTT ACA TCT CAT CC
105	G-6061		TGT TTG TGG AGA AAA TAT TGA GGA
106	G-7885		CAT TTG ATC CAT CAG TTC ATT TG

93,94	qRT-PCR of at3g44530
95,96	qRT-PCR of at3g44540
97,98	qRT-PCR of ubiquitin
99, 100	20 bp deletion in at3g44530
101,102	at3g44530 exon 5
103, 104	at3g44530 exon 11
105, 106	at3g44540 exon 4

2 Flowering time variation in *Arabidopsis thaliana*

2.1 Summary statistics based on individual plants

Condition	Parameters	DTF	TLN
16LD	Mean*	38.5±0.6	34.2±1.1
	Standard deviation	12.18	20.78
	Median	36	26
	CV_g	28.12	56.94
	Range	20-86	8-114
	Sample size	1489	1377
	Heritability	0.81	0.9
23LD	Mean	28.0±0.6	29.6±1.0
	Standard deviation	13.0	20.2
	Median	23	20
	CV_g	43.11	64.45
	Range	14-82	7-136
	Sample size	1653	1643
	Heritability	0.85	0.89
16LDV	Mean	23.8±0.2	20.8±0.3
	Standard deviation	4.45	7.05
	Median	23	20
	CV_g	15.36	30.43
	Range	13-55	8-70
	Sample size	1679	1648
	Heritability	0.70	0.84
23SD	Mean	54.3±1.0	57.2±1.1
	Standard deviation	15.03	17.85
	Median	52	53
	CV_g	23.01	35.17
	Range	22-117	12-148
	Sample size	1187	1410
	Heritability	0.71	0.87

* ± 95% confidence intervals (2 x standard error of the mean)

2.2 Genetic correlation across environments

	16LD	23LD	16LDV
23LD	0.93 (0.90-0.94)		
16LDV	0.47 (0.33-0.68)	0.31 (0.15-0.69)	
23SD	0.78 (0.71-0.84)	0.72 (0.63-0.80)	0.38 (0.23-0.67)

95% confidence intervals (2 x standard error of the mean) in parentheses

2.3 Means of flowering time traits in accessions and mutant strains

The table below lists accession identification data (left columns) and mean flowering time trait values measured under four conditions: 16°C Long Day (16LD), 23°C Long Day (23LD), 16°C Long Day + Vernalization (16LDV), and 23°C Short Day. Traits reported include DTF, JLN, ALN, RLN, CLN, TLN.

No.	Ecotype	Stock No	Country	FRI Status	FLC status
1	Ak-1	N939	Germany	FRI-Ler	
2	Ang-1	N951	Belgium	new FRI allele	
3	Bay-0	N955	Germany	FRI-Ler	
4	Bch-3	N959	Germany	FRI-Col	
5	Bd-0	N963	Germany	FRI functional	new FLC allele
6	BG1	N22341	USA	new FRI allele	
7	BG4	N22344	USA	new FRI allele	
8	BG6	N22346	USA	new FRI allele	
9	BG9	N22349	USA	new FRI allele	
10	Bia-1	N971	Spain	no FRI deletion	
11	Bia-2	N973	Spain	no FRI deletion	
12	Bla-1	N977	Spain	FRI-Ler	
13	Bla-5	N6920	Spain	no FRI deletion	
14	Bla-6	N6921	Spain	no FRI deletion	
15	Bla-11	N985	Spain	no FRI deletion	
16	Bla-12	N987	Spain	FRI-Col	
17	Bth-2	N6657	Czech Republic	no FRI deletion	
18	Bt-0	N995	Czech Republic	FRI-Ler	
19	Bu-0	N997	Switzerland	FRI-Ler	
20	Bu-5	N1001	Switzerland	FRI-Col	
21	Bsch-2	N1025	Germany	FRI-Col	
22	Bu-11	N1029	Germany	no FRI deletion	
23	Ca-0	N1063	Ireland	no FRI deletion	new FLC allele
24	Cal-0	N1065	UK	no FRI deletion	new FLC allele
25	Can-0	N1067	Spain	no FRI deletion	new FLC allele
26	Cen-0	N1069	France	FRI functional	
27	Chi-0	N1073	Switzerland	no FRI deletion	
28	Chi-1	N22220	Russia(?)	no FRI deletion	
29	Chi-2	N22229	Russia(?)	no FRI deletion	
30	CIBC1	N22222	UK	no FRI deletion	
31	CIBC10	N22229	UK	no FRI deletion	
32	Co-1	N1085	Portugal	no FRI deletion	
33	CSHL1	N22419	USA	FRI non-functional	FLC functional
34	Cvi-0	N1097	Cape Verde	FRI non-functional	
35	Db-0	N1101	France	FRI-Col	
36	Di-1	N1109	France	FRI-Ler	
37	Di-0	N1115	Czech Republic	FRI-Ler	
38	Dra-1	N1119	Czech Republic	FRI-Col	
39	Ei-4	N1127	Germany	FRI-Col	
40	El-0	N1135	Germany	FRI-Col	
41	En-0	N1137	Germany	FRI-Col	
42	En-1		Germany	FRI-Ler	
43	Erk-T	N1143	Germany	no FRI deletion	
44	Est-1	N1151	Estonia	no FRI deletion	
45	Est-0	N1153	France	no FRI deletion	
46	Et-0		France	FRI-Ler	
47	Fe-1	N1155	Germany	FRI-Col	
48	Fi-0	N1157	Germany	FRI-Col	
49	FM10	N22391	USA	FRI-Ler	
50	FM11	N22392	USA	FRI-Col	
51	Fr-2	N1169	Germany	FRI-Col	
52	Fr-4	N1173	Germany	FRI-Ler	FLC-Da1
53	Fr-6	N1177	Germany	FRI-Ler	
54	Ga-0	N1185	Germany	FRI-Col	
55	Gd-1	N1189	Switzerland	no FRI deletion	
56	Ge-1	N1191	Switzerland	no FRI deletion	
57	Ge-0	N1197	Germany	no FRI deletion	
58	Go-2	N22277	Germany	no FRI deletion	
59	GOT10		Germany	no FRI deletion	
60	GOT11	N22286	Germany	new FRI allele	FLC-Da1
61	Gr-3	N1199	Austria	FRI-Ler	
62	Gr-6	N1203	Austria	FRI-Col	
63	Gre-0	N6728	Austria	FRI-Ler	
64	Gre-0	N1211	USA	FRI-Ler	
65	Gy-0	N1213	Germany	FRI-Col	
66	H-0	N1217	Denmark	no FRI deletion	
67	Hau-0	N1225	Germany	no FRI deletion	
68	Hh-0	N1229	Germany	no FRI deletion	
69	Hi-0	N1233	Germany	FRI-Col	
70	Hl-3	N6178	Tajikistan	FRI-Ler	
71	Kbodja-Col		Tajikistan	FRI-Ler	
72	HR15	N22214	UK	new FRI allele	
73	HR5	N22205	UK	no FRI deletion	
74	HS1	N1237	Germany	no FRI deletion	
75	Hs-0	N22351	USA	FRI-Col	
76	In-0	N1241	Austria	no FRI deletion	
77	Is-0			no FRI deletion	
78	J-1	N1249	Czech Republic	new FRI allele	
79	Jm-1	N1261	Czech Republic	FRI-Ler	
80	Ki-0	N1271	UK	no FRI deletion	new FLC allele
81	Km-0	N1273	USA	FRI non-functional	FLC non-functional

Mutants with Controls

No.	Genotype	Stock No	Background	16°C Long Day (16LD)						23°C Long Day (23LD)						16°C Long Day + Vernalization (16LDV)						23°C Short Day (2:			
				DTF	JLN	ALN	RLN	CLN	TLN	DTF	JLN	ALN	RLN	CLN	TLN	DTF	JLN	ALN	RLN	CLN	TLN	DTF	JLN	ALN	RLN
1	Col-0	N1093	FRI-Col	30.40	8.73	8.36	17.09	3.45	20.55	18.43	4.25	7.88	12.13	3.00	15.13	22.60	6.40	4.80	11.20	3.40	14.60	57.33	8.67	40.89	49.56
2	Ler	N6928	FRI-Ler	30.33	9.67	2.83	12.50	4.17	16.67	17.50	4.20	6.80	11.00	3.20	14.20	22.00	7.14	2.86	10.00	3.71	13.71	58.70	7.75	21.00	28.30
3	ap1-15	Lab Stock	FRI-Col	28.33	6.86	6.29	12.50	2.83	15.33	17.27	4.27	6.27	10.42	2.58	13.00	23.80	8.60	6.80	15.40	3.60	19.00	52.80	9.50	26.50	34.67
4	co-1	N3122	FRI-Ler	42.17	9.42	26.83	36.25	10.50	46.75	23.92	4.25	15.08	18.33	5.50	23.44	31.75	6.57	14.57	24.40	8.88	29.75	47.58	7.67	20.67	28.33
5	co-2	Lab Stock	FRI-Ler	45.30	9.00	22.00	27.60	9.60	37.20	26.33	4.25	14.00	18.33	5.11	23.44	32.50	7.67	16.33	24.00	10.00	34.40	47.43	8.60	18.40	26.30
6	cry2-1	N3732	FRI-Col	53.45	12.33	28.00	42.17	7.67	49.83	25.08	5.27	13.64	18.91	3.91	22.82	30.91	9.36	8.27	17.50	4.58	22.08	63.82	11.00	46.00	57.50
7	fca-1	N52	FRI-Ler	37.42	11.17	13.17	24.33	7.50	31.83	23.70	3.89	13.33	17.22	4.00	21.22	22.55	5.67	4.33	10.00	4.22	14.22	66.00	6.00	28.00	42.25
8	fca-2	N168	FRI-Ler	36.67	10.17	9.75	19.92	6.58	26.50	19.08	4.67	8.92	13.58	3.58	17.17	22.00	6.20	4.40	10.60	4.30	14.80	57.58	8.33	20.67	28.88
9	fca-3	N169	FRI-Ler	34.50	8.00	9.83	17.83	5.50	23.33	21.58	4.33	10.08	14.42	3.83	18.25	22.60	6.00	3.20	9.20	3.20	12.40	63.45	7.38	30.38	42.55
10	fca-4	N170	FRI-Ler	38.75	10.25	11.92	22.17	6.75	28.92	23.50	5.10	11.42	15.50	5.20	18.92	22.00	6.33	3.33	9.67	3.42	13.08	59.18	9.20	26.60	38.70
11	fca-5	N171	FRI-Ler	45.82	10.33	17.50	27.83	7.50	35.33	32.91	4.75	15.30	20.00	5.33	25.20	24.17	6.25	3.25	9.42	3.58	13.00	68.50	7.00	36.00	51.00
12	fca-6	N172	FRI-Ler	41.92	11.25	20.08	31.33	8.58	39.92	29.33	4.83	19.42	24.17	5.33	29.50	22.50	7.33	7.00	14.88	4.42	15.00	52.27	10.00	16.86	25.91
13	fd-1	Lab Stock	FRI-Ler	38.27	10.00	9.22	19.20	5.90	25.10	20.92	3.00	8.09	12.82	3.64	16.45	27.78	7.83	6.50	14.18	5.13	20.00	49.67	9.57	18.27	59.08
14	fe-1	N54	FRI-Ler	40.27	9.00	13.30	22.36	7.18	29.64	22.67	4.00	11.08	15.92	4.25	20.17	26.00	7.60	6.08	13.67	5.36	19.55	60.36	8.45	27.50	34.20
15	fha-1	N108	FRI-Ler	44.58	10.36	13.36	23.33	6.75	30.08	19.50	3.27	7.50	11.50	2.83	14.33	26.50	7.58	6.25	12.17	4.42	18.08	54.33	8.00	21.00	28.08
16	fha-2	N186	FRI-Ler	37.00	8.58	8.42	17.00	5.33	22.33	18.67	2.91	6.50	9.50	3.58	13.08	23.50	5.92	5.10	12.50	5.25	17.42	51.08	7.08	17.80	25.58
17	fha-3 A	N187	FRI-Ler	33.18	9.45	8.45	17.91	5.82	23.73	17.92	4.08	6.09	9.36	3.27	12.64	23.33	7.40	5.70	13.09	5.30	17.80	50.89	7.30	21.00	29.00
18	fha-3 B	Lab Stock	FRI-Ler	31.18	7.64	6.36	14.00	2.91	16.91	19.64	5.64	6.73	9.64	2.75	12.91	26.55	6.45	6.64	13.00	4.91	16.80	56.50	6.83	40.09	48.09
19	flc-3	N109	FRI-Col	35.00	10.55	9.82	20.36	4.27	24.64	24.08	5.67	6.75	10.83	3.67	13.58	22.00	7.30	5.29	12.71	3.80	17.00	56.33	8.00	22.57	32.14
20	fpa-1	Lab Stock	FRI-Ler	NA	8.89	12.00	20.90	8.10	29.00	26.00	4.73	10.55	16.58	5.83	19.67	22.00	7.43	9.60	18.40	4.29	25.60	NA	9.57	NA	29.57
21	ft-1	Lab Stock	FRI-Ler	39.92	10.33	16.42	26.75	7.25	34.00	23.36	5.09	10.92	16.58	3.82	22.42	22.25	7.18	3.82	11.00	7.20	15.18	63.42	9.88	37.38	45.82
22	fve-1	N107	FRI-Ler	41.67	8.70	12.80	21.50	8.92	30.42	23.73	4.00	12.91	17.64	5.27	21.73	28.86	7.17	8.33	14.63	4.18	22.50	58.75	7.80	21.40	29.58
23	fwa-1	N106	FRI-Ler	30.33	7.92	10.33	18.25	8.00	26.25	24.17	5.17	11.36	16.45	6.75	20.42	27.82	6.73	9.18	15.91	7.88	23.73	NA	NA	NA	28.00
24	fwa-2	Lab Stock	FRI-Ler	39.75	9.00	11.83	21.50	5.25	26.75	24.73	4.78	9.67	14.91	3.73	18.64	24.25	6.00	4.45	10.58	7.82	13.92	63.17	9.25	31.00	37.70
25	fy-1 A	N57	FRI-Ler	35.08	8.00	13.50	21.50	6.17	27.67	26.08	4.42	10.82	16.75	3.92	18.17	24.42	6.67	5.50	12.17	4.25	16.42	NA	8.00	NA	39.80
26	fy-1 B	Lab Stock	FRI-Ler	39.45	9.00	10.71	19.22	5.00	24.22	26.90	4.00	9.08	15.55	3.82	19.36	25.75	7.08	4.75	11.83	3.75	15.58	64.25	NA	NA	35.09
27	fy-1 C	Lab Stock	FRI-Ler	46.09	10.70	19.10	29.55	8.91	38.45	28.45	NA	10.56	17.83	4.92	22.75	31.70	8.00	14.00	18.75	6.75	25.50	55.33	NA	NA	28.33
28	gi-6	N183	FRI-Ler	31.50	9.75	7.83	17.58	5.92	23.50	26.90	3.33	13.42	11.83	3.92	11.83	24.83	7.09	5.45	12.55	5.00	17.55	60.58	9.25	19.13	52.00
29	hy4-1 (cry1)	N70	FRI-Ler	58.38	NA	NA	54.00	12.25	66.25	47.00	4.90	NA	48.67	9.78	58.44	27.60	9.10	9.50	18.60	4.40	23.00	NA	NA	41.22	NA
30	ld-1	Lab Stock	FRI-Col	28.88	5.00	4.75	18.40	2.63	21.03	18.90	3.56	5.33	8.67	2.56	11.22	23.22	5.30	4.90	10.20	3.00	11.22	50.63	7.40	27.20	37.14
31	spy-1	Lab Stock	FRI-Ler	32.00	10.00	9.38	18.40	4.10	22.50	19.27	2.58	7.30	12.20	3.30	15.50	24.00	8.58	5.75	14.33	3.50	17.83	54.00	9.17	30.83	45.13
32	spy-2	Lab Stock	FRI-Ler	29.86	4.40	5.80	10.83	2.83	13.67	18.38	4.00	5.22	8.78	2.78	11.56	22.83	5.09	4.82	9.91	2.73	12.64	52.50	7.83	29.17	42.50
33	spy-3	Lab Stock	FRI-Ler	25.75	3.66	2.29	7.38	3.38	10.75	16.25	3.17	3.75	6.33	3.17	9.50	18.42	3.92	2.33	6.25	3.25	9.50	33.30	4.11	10.56	14.67
34	spy-5	Lab Stock	FRI-Ler	22.92	NA	NA	9.10	0.80	9.90	15.80	NA	NA	9.11	1.22	10.33	16.60	NA	NA	10.90	1.33	12.10	49.00	5.73	3.09	44.43
35	tfl1-1	Lab Stock	FRI-Col	25.25	7.25	1.75	8.38	2.25	10.63	16.63	4.00	4.00	8.00	2.11	10.11	18.13	5.22	2.89	8.11	1.22	10.33	24.73	8.16	26.39	8.82
36	toc1-1	Lab Stock	FRI-Ler	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
			AVG	36.84	8.94	11.49	20.94	5.98	26.93	22.68	4.30	9.64	14.82	3.97	18.72	24.62	6.94	6.27	13.03	4.62	17.65	55.36	8.16	26.39	35.86
			SD	7.73	1.87	6.27	9.46	2.54	11.71	5.90	0.74	3.60	7.06	1.50	8.37	3.74	1.19	3.31	3.80	1.84	5.39	9.16	1.44	9.86	11.25

Multiple stocks and repeated assays are denoted by A, B, C, etc. Latitude information was obtained from Geographic Names Information System (US Geological Survey).

Traits: DTF, days to flowering; JLN, juvenile rosette leaf number; ALN, adult rosette leaf number; CLN, cauline leaf number; TLN, total leaf number.

2.4 Key to ATH1 array data

Accession	ATGE array IDs	Replicates
Bay-0	ATGE_111_A,B,C	3
C24	ATGE_112_A,C,D	3
Col-0	ATGE_113_A,C	2
Cvi	ATGE_114_A,B,C	3
Est-1	ATGE_115_A,B	2
Kin-0	ATGE_116_A,B,C	3
Ler	ATGE_117_B,C	2
Nd-1	ATGE_118_A,B,C	3
Shahdara	ATGE_119_A,C,D	3
Van-0	ATGE_120_A,B,C	3
Ak-1	ATGE_121_A	1
Bch-3	ATGE_123_A	1
Bla-5	ATGE_124_A	1
Can-0	ATGE_125_A	1
Cen-0	ATGE_126_A	1
CIBC10	ATGE_127_A	1
Dra-1	ATGE_128_A	1
Enk-T	ATGE_129_A	1
Er-0	ATGE_130_A	1
Fr-2	ATGE_131_A	1
GOT1	ATGE_132_A	1
HR-5	ATGE_133_A	1
Is-0	ATGE_134_A	1
Kl-0	ATGE_135_A	1
Li2:1	ATGE_136_A	1
Nc-1	ATGE_137_A	1
NFE1	ATGE_138_A	1
Nok-1	ATGE_139_A	1
Nw-1	ATGE_140_A	1
M7323S	ATGE_141_A	1
Ms-0	ATGE_142_A	1
Ob-1	ATGE_143_A	1
Old-1	ATGE_144_A	1
Ove-1	ATGE_145_A	1
Se-0	ATGE_146_A	1
Sf-2	ATGE_147_A	1
Ta-0	ATGE_148_A	1
Uk-3	ATGE_149_A	1

2.5 Pearson correlation of known flowering regulators

Gene	ID	DTF	TLN
FLC	At5g10140	0.7833	0.6397
ELF8	At2g06210	0.3964	0.3100
FRI	At4g00650	0.3826	0.4192
SPL9	At2g42200	0.3491	0.2995
TOE2	At5g60120	0.3349	0.1758
SPL3	At2g33810	0.2635	0.2381
SNZ	At2g39250	0.2464	0.1375
SMZ	At3g54990	0.2369	0.2099
EMF1	At5g11530	0.2218	0.1982
TOC1	At4g18020	0.2107	0.1278
SPL2	At5g43270	0.2076	0.1465
SPL6	At1g69170	0.1984	0.1453
EBS	At4g22140	0.1980	0.1609
GI	At1g22770	0.1795	0.1033
LHY	At1g01060	0.1357	0.2333
FUL	At5g60910	0.1259	0.1626
ELF6	At5g04240	0.1247	-0.0091
VIP2	At5g59710	0.1219	0.0798
FVE	At2g19520	0.1120	-0.0160
ELF4	At2g40080	0.1038	0.1106
FPF1	At5g24860	0.1020	0.1990
VIP3	At4g29830	0.0992	0.2011
CCA1	At2g46830	0.0947	0.2164
FCA	At4g16280	0.0779	0.0856
TFL1	At5g03840	0.0643	0.1812
FD	At4g35900	0.0530	0.1000
TOE1	At2g28550	0.0427	0.1105
FRS6	At1g52520	0.0375	0.0827
SPL4	At1g53160	0.0313	0.0681
ELF5	At5g62640	0.0220	-0.0227
PHYB	At2g18790	0.0209	0.0641
FRL1	At5g16320	0.0183	-0.0094
PHYA	At1g09570	0.0071	-0.0330
PFT1	AT1G25540	0.0023	0.0517
FRS9	At4g38170	0.0006	-0.0385
MAF5	At5g65080	-0.0013	-0.0570
SPL5	At3g15270	-0.0182	0.0649
FKF1	At1g68050	-0.0217	-0.0746
TFL2	At5g17690	-0.0265	-0.1365
FY	At5g13480	-0.0270	0.0008
FLM	At1g77070	-0.0353	0.0120
VRN2	At4g16845	-0.0431	-0.1480
EMF2	At5g51230	-0.0574	-0.0388
FLD	At3g10390	-0.0714	-0.1506
MAF2	At5g65050	-0.0747	-0.0570
SPL10	At1g27370	-0.0752	-0.1330
FT	At1g65480	-0.0793	-0.0118
CRY1	At4g08920	-0.0914	-0.1559
VIP4	At5g61150	-0.1064	-0.0367

FLM	At1g77080	-0.1220	-0.0537
VRN1	At3g18990	-0.1238	-0.1935
SPL11	At1g27360	-0.1290	-0.1075
FLK	At3g04610	-0.1505	-0.0876
LFY	At5g61850	-0.1539	-0.0836
HEN1	At4g20910	-0.1596	-0.3115
ELF7	At1g79730	-0.1861	-0.2738
CO	At5g15840	-0.1901	-0.1453
AGL24	At4g24540	-0.1921	-0.0696
ZTL	At5g57360	-0.1984	-0.1156
HUA2	At5g23150	-0.2085	-0.1597
MAF4	At5g65070	-0.2233	-0.1262
SPL13	At5g50570	-0.2465	-0.1668
FRL2	At1g31814	-0.2681	-0.1691
CRY2	At1g04400	-0.2843	-0.3587
SVP	At2g22540	-0.3355	-0.2931
ELF3	At2g25930	-0.3644	-0.3524
SOC1	At2g45660	-0.4101	-0.3901
SPL15	At3g57920	-0.4887	-0.4463

3 Uk-1/Uk-3 hybrid incompatibility

3.1 Marker information of 64 SNP marker tested on 181 F2 plants

3.2 Marker information of 26 SNP marker tested on 93 Col-0 backcross lines

3.3 Key to ATH1 array data

array number	genotype	treatment
313	Uk-1	0 days at 16°C SD
314	Uk-3	0 days at 16°C SD
315	F1 (Uk1xUk3)	0 days at 16°C SD
316	Uk-1	5 days at 16°C SD
317	Uk-3	5 days at 16°C SD
318	F1 (Uk1xUk3)	5 days at 16°C SD

3.4 Differentially expressed genes in F1 hybrids versus parents at 23°C

AT2G30520	AT2G43570	AT4G35090	AT5G51890
AT2G41100	AT1G70580	AT4G38840	AT5G52250
AT2G39250	AT1G74670	AT4G39030	AT5G52640
AT2G23910	AT1G67865	AT3G43190	AT5G54250
AT2G29290	AT1G66230	AT3G48640	AT5G55450
AT2G18660	AT1G29500	AT3G51240	AT5G57560
AT2G25510	AT1G71030	AT3G56400	AT5G59540
AT2G20750	AT1G43910	AT3G57260	AT5G61290
AT1G03870	AT3G04510	AT5G03350	AT5G27350
AT1G61800	AT3G22840	AT5G10760	AT5G14740
AT2G37040	AT2G20875	AT5G10380	AT4G14365
AT2G46400	AT2G40750	AT5G12110	AT2G26400
AT2G31880	AT4G21380	AT5G24530	
AT1G65060	AT4G21830	AT5G39080	
AT1G18710	AT4G31870	AT5G44620	

3.5 Differentially expressed genes in F1 hybrids at 16°C versus 23°C

AT2G32680	AT1G13210	AT3G17690	AT4G26070	AT5G23870
AT2G19130	AT1G11330	AT3G17700	AT4G26270	AT5G24210
AT2G34940	AT1G34750	AT3G14620	AT4G27280	AT5G24530
AT2G26240	AT1G72940	AT3G25780	AT4G28490	AT5G24540
AT2G26190	AT1G74710	AT3G26210	AT4G31800	AT5G35735
AT2G30140	AT1G56150	AT3G26600	AT4G33300	AT5G37600
AT2G23810	AT1G08450	AT2G45760	AT4G34150	AT5G39670
AT2G30250	AT1G27770	AT3G13100	AT4G35180	AT5G40330
AT2G44180	AT1G27730	AT3G28580	AT4G35600	AT5G42830
AT2G37585	AT1G49750	AT3G22910	AT4G39030	AT5G44820
AT2G37710	AT1G14370	AT3G22231	AT4G39830	AT5G44570
AT2G38470	AT1G21130	AT3G25600	AT3G43800	AT5G45180
AT2G39210	AT1G28480	AT3G13610	AT3G45640	AT5G45110
AT2G44790	AT1G18810	AT3G22240	AT3G45620	AT5G45510
AT2G25810	AT1G35710	AT3G14840	AT3G45970	AT5G45500
AT2G47730	AT1G05880	AT3G12220	AT3G45860	AT5G46230
AT2G41410	AT1G34420	AT1G13750	AT3G46880	AT5G48380
AT2G46680	AT1G21520	AT1G07000	AT3G47480	AT5G49760
AT2G28780	AT1G02450	AT1G58300	AT3G48080	AT5G49680
AT2G38870	AT1G19380	AT1G22280	AT3G48090	AT5G52750
AT2G18660	AT1G08050	AT1G66970	AT3G50770	AT5G52760
AT2G18690	AT2G43570	AT2G33580	AT3G50930	AT5G52740
AT2G24160	AT1G69840	AT2G40600	AT3G50950	AT5G54610
AT2G14560	AT1G70530	AT4G00330	AT3G51300	AT5G54490
AT2G48030	AT1G70520	AT4G01700	AT3G52430	AT5G54860
AT2G46600	AT1G33960	AT4G01920	AT3G52400	AT5G55170
AT1G23840	AT1G73805	AT4G01870	AT3G55980	AT5G60950
AT1G30900	AT1G73800	AT4G02420	AT3G57240	AT5G60900
AT1G23830	AT1G67970	AT4G02380	AT3G57460	AT5G61010
AT1G76970	AT1G76420	AT4G04500	AT3G57950	AT5G61520
AT1G65490	AT1G72280	AT4G04220	AT3G59270	AT5G61900
AT1G08930	AT1G64280	AT4G08850	AT3G60470	AT5G62630
AT1G10340	AT1G52290	AT4G10500	AT3G60420	AT5G64000
AT1G03290	AT1G35230	AT4G11900	AT3G62280	AT5G67110
AT1G03370	AT1G43910	AT4G11890	AT5G04720	AT5G25440
AT1G22650	AT1G15790	AT4G12720	AT5G05460	AT5G26920
AT2G04430	AT1G01560	AT4G18880	AT5G10380	AT5G14930
AT2G40270	AT1G13470	AT5G44810	AT5G11250	AT1G16670
AT2G24600	AT3G01290	AT4G23150	AT5G13320	AT3G51890
AT2G46400	AT3G07520	AT4G23480	AT5G17330	AT4G36500
AT2G17040	AT3G09410	AT4G23220	AT5G18470	AT5G20230
AT2G17120	AT3G09940	AT4G23310	AT5G18780	AT5G09440
AT2G24850	AT3G05660	AT4G23270	AT5G22270	AT4G17660
AT2G40140	AT3G11820	AT4G23210	AT5G22380	AT2G23200
AT1G23850	AT3G08720	AT4G25000	AT5G22630	
AT1G59710	AT3G09830	AT4G25940	AT5G22530	
AT1G59870	AT3G04210	AT4G26090	AT5G22570	

3.6 Accessions of which fragments of at3g44530 and at3g44540 were sequenced

Ak-1	Jl-1	Per-2	2-2-a	Db-2	Mv-0
Aa-0	Jl-2	Pf-0	2-3-a	Dr-0	NFC1
Bd-0	Kn-0	Rsch-0	3-1-a	Edi-0	Old-1
BG1	Lip-0	Rd-0	3-2-a	Ei-2	Pi-0
Bla-1	Lö-1	Sg-1	3-3-a	Fe-1	Pla-2
Uk-3	Lö-2	Sg-2	4-1-a	Gr-1	Pr-0
Bla-3	M3385S	Te-0	5-1-a	Gr-3	PUZ16
Bs-1	Mt-0	Ts-1	5-2-a	Gre-0	Rou-0
Bs-2	Na-1	Ts-6	5-3-a	Hh-0	RP1
Ge-1	Nc-1	Uk-2	6-1-a	HR5	Sav-0
Chi-0	Nd-0	Uk-4	7-1-a	Kro-0	Shahdara
Ang-1	Nok-0	Wt-4	8-1-a	Le-0	St-0
Gy-0	Np-0	1-1-a	ANH1	Li-2	Ste-0
Er-0	Nw-1	1-2-a	Bay-0	Ll-2	Stw-0
Fr-6	Nw-2	1-3-a	CIBC1	Me-0	Wil-1
Fr-7	Oy-1	2-1-a	Db-1	Mir-0	Wü-0

3.8 Collection sites of *Arabidopsis thaliana* in Umkirch

Figure A2: A. thaliana was found and collected on sandy soil next to a soccer field (A) and (C), as weed in a flower bed (B) and on a parking lot (D) and (E).

3.9 Phenotypic variation of accessions collected in Umkirch

Figure A3: Phenotypic variation of lines collected in Umkirch becomes obvious after growing lines under controlled conditions, at 23°C LD. All plants are of same age.
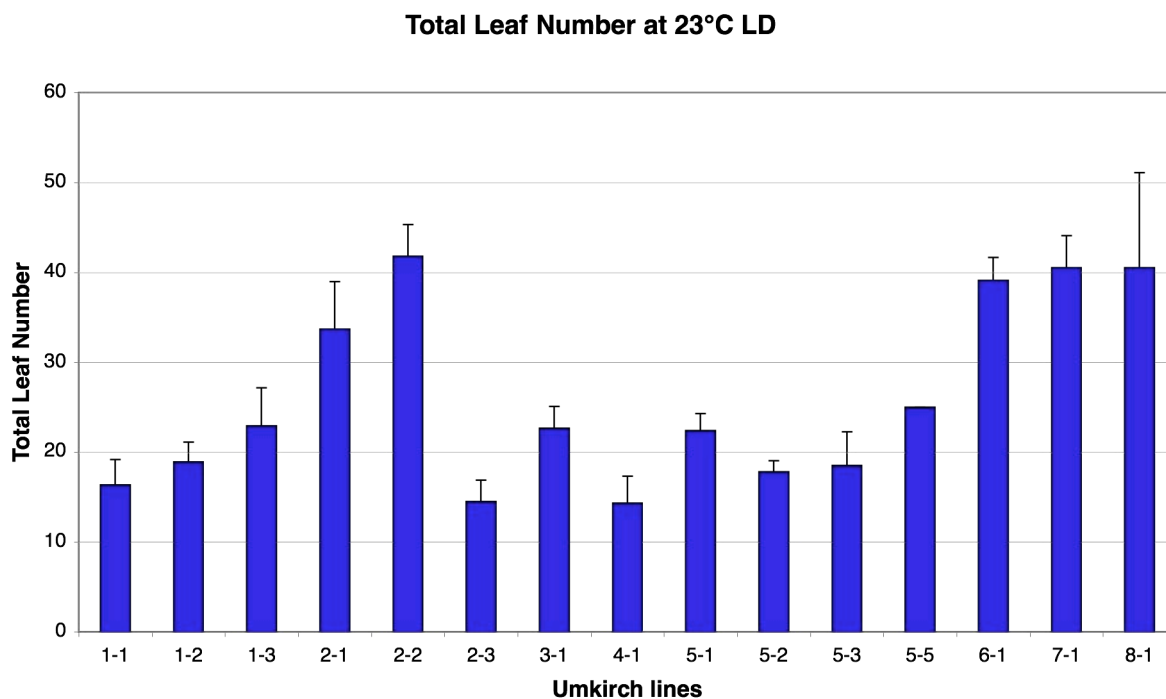
Total Leaf Number at 23°C LD

Figure A4: Flowering time variation observed in Umkirch lines grown at 23LD.

3.10   Crosses of Uk-1 and Uk-3 to different accessions
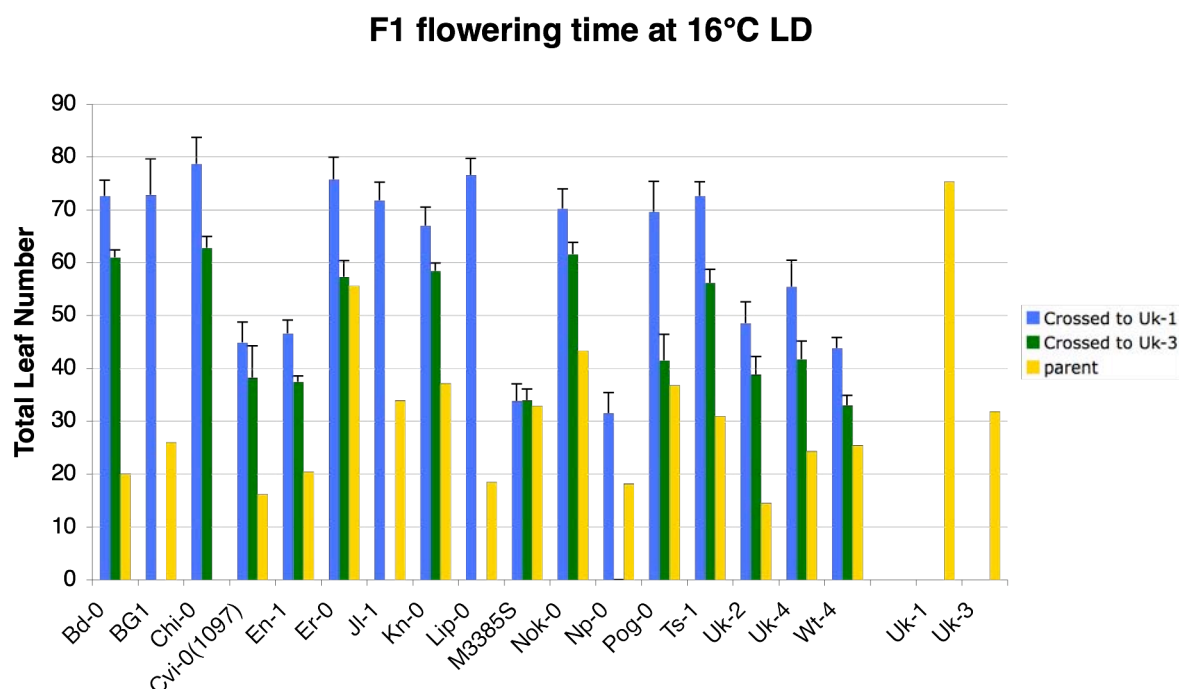
F1 flowering time at 16°C LD

Figure A5: Crosses of different accessions with Uk-1 and Uk-3 reveal rather later flowering F1 generations with the exception of crosses to M3385.

XI. LEBENSLAUF

18.02.1976	geboren in Heilbronn-Neckargartach
1995	Abitur am Gymnasium Eppingen in Eppingen
1996	Freie Kunsthochschule Metzingen
1996-2002	Studium der Biologie (Diplom) an der Eberhard-Karls-Universität Tübingen
1999-2000	Auslandssemester an der „University of Guelph" (Guelph, Kanada)
2000	Forschungsaufenthalt an der „University of Guelph" im Labor von Dr. A. Nassuth
2002	Diplom der Biologie an der Eberhard-Karls-Universität Tübingen mit den zwei Hauptfächern Pflanzenphysiologie und Botanik und dem Nebenfach Biochemie
2002-2006	Doktorarbeit in der Abteilung für Molekularbiologie am Max-Planck-Institut für Entwicklungsbiologie in Tübingen unter Anleitung von Prof. D. Weigel
2006-2007	Elternzeit